Human Recombineering

Melvys Valledor

University of Miami, melvys_valledor@yahoo.com

Follow this and additional works at: https://scholarlyrepository.miami.edu/oa_dissertations

Recommended Citation
https://scholarlyrepository.miami.edu/oa_dissertations/1111

This Embargoed is brought to you for free and open access by the Electronic Theses and Dissertations at Scholarly Repository. It has been accepted for inclusion in Open Access Dissertations by an authorized administrator of Scholarly Repository. For more information, please contact repository.library@miami.edu.
UNIVERSITY OF MIAMI

HUMAN RECOMBINEERING

By

Melvys Valledor

A DISSERTATION

Submitted to the Faculty
of the University of Miami
in partial fulfillment of the requirements for
the degree of Doctor of Philosophy

Coral Gables, Florida

December 2013
UNIVERSITY OF MIAMI

A dissertation submitted in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

HUMAN RECOMBINEERING

Melvys Valledor

Approved:

Paul C. Schiller, Ph.D.
Professor of Orthopedics and
Biochemistry and Molecular Biology
Tissue Bank; Geriatrics, Vascular Biology,
and Interdisciplinary Stem Cell Institutes

Pedro J.I. Salas, M.D., Ph.D.
Professor of Cell Biology
and Anatomy

Richard S. Myers, Ph.D.
Lecturer of Biochemistry and
Molecular Biology

Ramiro E. Verdun, Ph.D.
Assistant Professor of Medicine

Stephan Züchner, M.D., Ph.D.
Interim Chair and Professor,
Dr. John T. Macdonald Foundation
Department of Human Genetics and
Director, Center for Human Molecular Genomics
John P. Hussman Institute for Human Genomics

M. Brian Blake, Ph.D.
Dean of the Graduate School
Abstract of a dissertation at the University of Miami.

Dissertation supervised by Professors Paul C. Schiller and Richard S. Myers.
No. of pages in text. (776)

Genome editing by homologous recombination (HR) is centrally important for functional genomics, for making disease models, for bioengineering and for gene therapy. HR is a consequence of DNA damage repair and induced DNA damage stimulates HR but is toxic and can lead to unintended genetic consequences. While cellular HR is dependent on many enzymes, viral HR relies on a simple SynExo nanomachine comprised of an exonuclease and a synaptase. Using ssDNA substrates, the enterobacteria phage lambda synaptase (Beta protein) has been shown to efficiently and safely edit DNA sequences by Recombineering in its host, *E. coli*, with HR rates on the order of 10%. Unfortunately, Beta protein is 10,000x less efficient when used to modify genomes in mammalian cells. SynExos are common to dsDNA viruses that infect a variety of hosts - including humans - and their mechanisms are conserved, suggesting that direct genome Recombineering can be extended beyond bacteria to humans.

Our hypothesis is that Recombineering is host-specific because viral SynExo recombinases co-evolve with host proteins. We predict that a human viral synaptase can catalyze HR in human cells with efficiencies approaching that of the phage lambda synaptase in *E. coli*.

To test our hypothesis we built a sensitive and quantitative assay for gene targeting by changing the spectral properties of fluorescent proteins with small nucleotide changes.
directed to the chromophore region. Fluorescent protein engineering with oligos in bacteria corroborated a DNA strand bias in Recombineering, that MMR was a potent inhibitor of Recombineering unless oligo sequences were selected to avoid detection, and that the incorporated oligo segregated with cell division to produce a mixed clone of recombinant and non-recombinant cells. The efficiency of optimized Recombineering in bacteria could reach ~40% of transformed cells.

Oligo-mediated gene targeting in human cells was optimized using endogenous cellular functions. Gene targeting was further stimulated by Recombineering using human Herpes virus ICP8. Recombineering-mediated changes in phenotype corresponded to the predicted change in genotype. However, ICP8 expression was toxic to cells and caused loss of target gene expression. Human-optimized phage lambda Beta protein did not stimulate gene targeting, indicating that Recombineering is host specific.

Additional evidence for host-specific protein interactions in Recombineering was obtained using proteomics and in vitro functional assays. Both the λ SynExo Beta synaptase and the E. coli RecA protein were found to protect nascent ssDNA recombination intermediates from a host exonuclease. RecA protein competed with Beta for binding to λ Exo, suggesting that RecA displaced Beta from nascent ssDNA to block strand annealing and promote strand invasion. A set of interacting host proteins identified from co-purification studies suggests that λ SynExo activity is regulated by proteases and chaperones, that Recombineering is coupled to replication via primosome assembly on recombination intermediates, and that the resolution of Recombineering intermediates involves Holliday Junction resolvases.
DEDICATION

To my parents

for all their love and support

To Christina Wagner

for allowing this dream to happen

To the BMB professors

Dr. Rudd, Dr. Myers, Dr. Scott, Dr. Deutscher and Dr. Werner

for motivation and helping me to start in the Ph.D. program
ACKNOWLEDGMENT

I would like to thank my mentors Drs. Richard S. Myers and Paul C. Schiller for taking me as their student, for giving me a long leash during my training, and for moral and financial support. I would like to thank them for helping me carry out this ambitious project and trusting in my experimental design and results. I would like to thank them for assisting in the preparation of my grant applications, helping me attend research conferences, and for writing so many recommendation letters. I especially would like to thank Dr. Myers for his assistance in reviewing all my documents and presentations to help me with my English and logic.

I would like to thank my Ph.D. committee members, the late Dr. Walter Scott, Dr. Pedro Salas and Dr. Ramiro Verdun for their guidance through all these years. The members of my committee were thoughtful and engaging at every meeting and their sage advice helped me move forward and leave for another day the shiny objects at the edge of my path. They opened their doors to my questions and were my mentors on many occasions. They offered me not only recommendations but also experimental support in their labs and with their facilities. I would like to thank Dr. Stephan Züchner for agreeing to serve as the external examiner of this dissertation and to help escort me into my next training in human genetics.

I would like to thank Dr. Toni Barrientos and his wonderful lab members Drs. Flavia Fontanesi, Iliana Soto, Jingjing Liu, Alejandro Ocampo and Darryl Horn for fantastic training in my first rotation that led to the paper “Synthesis of cytochrome c oxidase subunit 1 is translationally downregulated in the absence of functional F1F0-ATP synthase”. I would like to thank Dr. Carlos Moraes and his wonderful lab members Drs.
Francisca Diaz, Adriana Rebele, Sion Williams and Sandra Bacman, for training me in human cell biology during my third rotation in tissue culture and mitochondrial molecular biology techniques. I hope to revisit mitochondria some day to help develop tools for studying mitochondriopathies.

I would like to thank the many scientists and staff of the Geriatric Research Education, and Clinical Center (GRECC) at the Miami Veterans Affairs Medical Center for offering me a home for my thesis, for discussing my experimental designs and progress, for teaching me how to do many experiments and for supporting my work. I especially would like to thank to Drs. Guy Howard, Carlos Perez-Stable, Kevin Curtis, Carmen Rios, Adriana Gomez, Bruce Troen, Priya Rai, Gianluca D'Ippolito, Gaëtan Delcroix, Teresita Reyes, Kristina K. Aenlle, Kenneth L. Seldeen, Manhui Pang, Rose Maria Van Zuilen, Bernard Roos and my colleagues Laurin Pacheco and Alicia De Las Pozas. I would like to especially thank Maria A. Rodriguez, Mireya Hernandez, Nubia Rodriguez and David Vazquez for their love and support, and for helping me when I was most busy and in need. I would also like to thank Yvette LeMasters in the VA Media Department for printing many of posters for my research presentations and for taking the GRECC group picture. Dr. Guy Howard deserves an extra “hurrah” for his unwavering support of my experiments and sage advice.

I would like to thank the faculty, students and staff of the Department of Biochemistry and Molecular Biology at the University of Miami, Miller School of Medicine for accepting me into their Ph.D. program and for their advice, training and friendship throughout these years. I would like to offer special thanks to Dr. Kenn Rudd for offering me a job at UM when I was still new to the USA and fumbling with English,
and for providing me great training in bacterial genetics. I would like to thank him, too, for motivating me and helping me to get into graduate school, for sharing many bacterial strains, for sharing lab resources and allowing me to use his lab equipment. I would like to thank Dr. Bill Whelan for his advice, unwavering support of my career development, and for inviting me to every Thanksgiving dinner with his family and friends. I would like to thank Dr. Murray Deutscher for his wonderful and inspiring class in the History of Molecular Biology, for supporting me in my attempt to join the Ph.D. program, for writing recommendation letters that helped fund my grant and for letting me use his tissue culture facility for some of my studies. I would also like to thank his lab members Drs. Georgeta Basturea, Mike Zundel and Helen A. Vincent for helping me trouble shoot many experiments and an especially loud thank you to my brother from Havana, Ernesto Quesada, for helping me on so many occasions and for being one of the best friends I have had in the last seven years.

I would like to thank Dr. Ralf Landgraf, and his lab members for the many discussions about my project and for excellent advice about how to make it work. Many thanks are offered as well to the other 3rd floor DNA guys (Drs. Yanbin Zhang and Feng Gong), to Drs. Arun Malhotra, T.K. Harris, Walter Scott, Sylvia Daunert and their lab members, Drs. Hassan Al-Ali, Pavan P. Vaidyanathan, T.J. Ragan, Fenghua Yuan, Brad Schmier, Renaud Sicard, Ling Zhang and soon-to-be Drs. Nirav Patel, Liangyue Qian and Zhang Qian for their comraderie. Many thanks are also extended to Dr. Emre Dikici and Sue Matsuura for allowing me to use their equipment and resources and to Sue for being one of the sweetest, most thoughtful human beings. Ever.
I would like to thank the BMB professors for their fantastic lectures with special recognition of Drs. Rik Myers, Walter Scott, T.K. Harris, Murray Deutscher, Tracy Fletcher, Arun Malhotra, Ralf Landgraf and Toni Barrientos for their wonderful lectures and for creating the great curriculum the BMB program used to have. I would like to express my highest recognition to Dr. Myers who made learning fun and interesting, and for expecting so much from the students while maintaining good humor and respect. I would like to thank Dr. Katherin Tosney for her outstanding class on Grant Writing that provided me with a stronger and clearer grant application that got funded, and for her recommendation letters for my grant applications.

I would like to thank the amazing BMB staff: Ida Collazo, Pura Pino, Sylka M. Pérez, Carolina Urroz, Grisel Lopez, Erica Alecio, Kyle Bryan and Maria I. Sacerio for all their help and support during all these years. I would like to thank Drs. Sylvia Daunert and Sapna Deo for their support during the last years.

I would like to thank the help of the Graduate School and give special recognition to the efforts of Dierdre Lacativa and Maria Zayas to always find a solution when everybody else could not and for directing me into the IMSD Program. Dr. Michael Gaines is thanked for accepting me into the NIH IMSD program and for providing numerous opportunities to develop as a beginning graduate student. I credit this support with strengthening my career development and helping me get subsequent opportunities leading to independent support by NIH and travel to meetings. I hope to pay it forward to future generations.

I would like to thank Dr. George McNamara of the UM Analytical Imaging Core Facility for the many microscopy sessions and beautiful pictures he arranged. Dr.
McNamara took a special interest in my project and provided me opportunities to use cutting edge approaches when instrument trials were being conducted at UM and elsewhere. I would like to thank Dr. James Baker in the UM Biology Department for taking the spectacular picture of MIAMI cells (Marrow-isolated adult multi-lineage inducible adult human stem cells) transfected with fluorescent labeled oligos. I would like to thank Dr. Pedro Salas for taking the beautiful picture of fluorescent MIAMI cells that was needed to publish the paper “Fluorescent protein Engineering by in vivo site directed mutagenesis”. I would like to thank Gabriel Gaidosh for helping me process confocal images, Dr. Adriana Gomez for her help with the microscope at the VA, and the Delta Vision group for taking gorgeous pictures of *E. coli* and MIAMI cells when they were visiting UM.

I would like to thank Dr. Sanjoy Bhattacharya and his lab members Katya Aribindi, Dr. Di Ding and Mabel Algeciras for the Mass Spec data collection in the SynExo interacting partners study.

I would like to thank Jim Phillips, Dr. Shannon Opiela and Jay Enten for their help at the Flow Cytometry core facility and for buying the narrow bandpass filters needed to separate Mostaza (yellow) from green fluorescent cells necessary to isolate recombinant cells.

I would like to thank Dr. Chun Chau Sze at the School of Biological Sciences, Nanyang Technological University, Singapore for the strain SCC1; Dr. David Knipe at Harvard for the plasmids pCMV-ICP8 and pCMV-ICP8-GFP; Dr. Jakob Reiser at NIH for the plasmids pNL-eGFP/CEF and pENTR2B/TREPitt and for advice about lentivirus engineering; Dr. Tom Misteli for the HeLa progerin reporter cell line; Dr. Priya Rai at
UM for the 293T cells and lentiviral plasmids pHR’ CMV 8.2ΔR and pCMV-VSV-G; Dr. Rong Wen at UM for the HEK293 cell line and Dr. Ken E. Rudd for many bacterial strains.

I would like to thank the interns, students and collaborators that have worked with me over the years, including Zoe Cosner, Megan Allen, Jahmour Givans, Ciara Major, Valine Desronvil, Mariana Braga, Senen Mendoza, and Stan Ivanov on the SynExo Interactome, Steven Basart for his work on PTD-Recombinases, and Tori Koslow for her work on Human Recombineering.

I would like to thank the USA tax payers, the NIH (IMSD and F31-GM089125-01 fellowships), the UM Interdisciplinary Stem Cell Institute through Qinghua Hu, the VA GRECC, the Biochemistry and Molecular Biology Department, and my mentors for the financial support. I could not have done it without you!
# TABLE OF CONTENTS

LIST OF FIGURES ........................................................................................................... xiv

LIST OF TABLES ............................................................................................................... xvii

LIST OF EQUATIONS ...................................................................................................... xviii

LIST OF ABBREVIATIONS .............................................................................................. xix

## CHAPTER 1. GENOME ENGINEERING INTRODUCTION ........................................ 1

- Significance of efficient human nuclear genome engineering ................................ 1
- Human genome engineering ..................................................................................... 1
- Human genome engineering by gene targeting ..................................................... 7
  - ssDNA oligos as substrates for gene targeting ..................................................... 8
  - Cellular functions that can influence gene targeting ......................................... 12
  - Inhibitors of gene targeting ............................................................................... 17
  - Cellular responses to gene targeting ................................................................ 19
- Bacterial genome engineering by Recombineering .............................................. 21
  - SynExos .............................................................................................................. 29
- Hypothesis ............................................................................................................... 36
- Specific Aims .......................................................................................................... 36

## CHAPTER 2. MATERIALS AND METHODS .................................................... 37

- Reagents ............................................................................................................... 37
- Recipes .................................................................................................................. 40
- Human cell line culture conditions ...................................................................... 43
- Human cell line list ............................................................................................... 46
  - E. coli culture conditions .................................................................................. 47
  - E. coli strain lists ............................................................................................... 48
- Oligo list ............................................................................................................... 54
- Plasmid lists ........................................................................................................... 59
- Plasmid construction method ............................................................................... 63
- Recombineering ................................................................................................... 84
  - E. coli Recombineering and fluorescent protein engineering .......................... 84
  - Human cells Recombineering and fluorescent protein engineering ............... 91
- DNA transformation/transfection ......................................................................... 92
  - E. coli electroporation ....................................................................................... 92
  - E. coli chemical transformation ......................................................................... 93
  - Colony transformation (no need to purify plasmid DNA) ............................... 93
  - Human cell transfections with oligos ............................................................... 93
  - Human cell transfections with plasmids ......................................................... 95
- P1 phage transductions ......................................................................................... 96
  - Producing P1 lysates from donor strain ........................................................... 96
  - P1 transduction ................................................................................................. 96
  - P1 titering ......................................................................................................... 97
CHAPTER 5. HUMAN RECOMBINEERING ........................................ 176

Developing the tool kit for human Recombineering.................................. 176
Assay: Fluorescent protein engineering.................................................. 176
Reporter cell line.................................................................................... 183
Oligo substrates...................................................................................... 191
Viral synaptases: ICP8 and HumBeta....................................................... 196
Human cell Recombineering................................................................. 200
293T cell gene targeting........................................................................... 211
Human Recombineering using transient transfection............................... 212
Recombineering using stable cell lines.................................................... 224
Discussion.............................................................................................. 239
Conclusions............................................................................................ 256

CHAPTER 6. REPAIRING PROGERIA BY RECOMBINEERING.............. 258

Progeria Recombineering reporter strain.................................................. 259
Developing recombination quantification assay........................................ 262
Recombination in Progerin HeLa cells...................................................... 265
Validating recombinants........................................................................... 274
Conclusions.............................................................................................. 282

CHAPTER 7. SYNEXO ALLELE SPECIFIC INTERACTIONS ................. 283

Functional interactions between viral recombination proteins and host proteins. 283
Physical interaction studies using pull downs.......................................... 284
Physical interaction studies using pull downs and mass spec..................... 285
Beta interaction with λ Exo was confirmed............................................. 287
λ Exo Ct α helix is not required for Beta interactions............................... 287
Beta Ct domain seems to be required for interactions with λ Exo.............. 288
λ Exo and Beta host proteins partners..................................................... 289
Coupling of recombination to replication via the primosome.................... 291
Processing of recombination intermediates via resolvases....................... 293
Regulation/disassembly via proteases...................................................... 295
Reverse pull-down experiments to validate MS data................................. 298
Genetic approaches to validate MS data.................................................. 298
Functional interactions between host proteins and viral recombination proteins. 300
λ Exo activity is inhibited by nascent ssDNA ........................................ 301
ssDNA annealing proteins overcome the inhibition of λ Exo by ssDNA...... 302
ssDNA increases λ Exo processivity but blocks reinitiation of DNA digestion by λ Exo upon dissociation from partially degraded DNA.............................. 303
Host RecA blocks ExoI from resecting nascent ssDNA produced by λ Exo and increases λ Exo processivity by increasing dwell time...................... 306
λ SynExo Beta blocks ExoI from resecting nascent ssDNA and increases processivity by increasing dwell time............................................... 307
Lags in progress curves......................................................................... 310
Conclusions ................................................................................................................................. 312

CHAPTER 8. CONCLUSIONS AND FUTURE STUDIES ................................................................. 315

Conclusions ................................................................................................................................. 315
Future directions .......................................................................................................................... 321
Human Recombineering ............................................................................................................. 321
Allele specific interactions ......................................................................................................... 325

APPENDIX A. PLASMID SEQUENCES AND MAPS ................................................................. 326

Acquired plasmids for this study .................................................................................................. 326
Plasmids created in this study or by the lab ................................................................................. 450

APPENDIX B. FLUORESCENT PROTEIN SEQUENCES ......................................................... 720

FP gene sequences ..................................................................................................................... 720
Fluorescent protein sequences ................................................................................................... 730

APPENDIX C. VISUAL ABSTRACT ............................................................................................. 734

Recombineering reporter in bacteria .......................................................................................... 734
What is Recombineering? ............................................................................................................. 735
Recombineering reporter in human stem cells .......................................................................... 736
What it takes to change fluorescent protein gene Mostaza to Green, or vice versa? ............... 736
Recombineering reporter in human cancer cells ......................................................................... 737
What is the genetic condition in Progeria? .................................................................................. 737
How efficient is ssDNA delivery to cells? ..................................................................................... 738
Herpes virus recombinase ICP8 expression in human stem cells ............................................. 739
Herpes virus recombinase ICP8 expression in 293T cells .......................................................... 742
Herpes virus recombinase ICP8 expression in HeLa cells .......................................................... 744
Herpes virus exonuclease UL12.5 expression in human HeLa cells ........................................... 746
Recombineering in 293T cells ...................................................................................................... 747
Recombineering in HeLa cells ..................................................................................................... 748
Recombineering enzyme interactions with viral host proteins .................................................. 749

REFERENCES ........................................................................................................................... 750
LIST OF FIGURES

Figure 1. 1 Enterobacteria phage lambda Gam protein structure .................................. 23
Figure 1. 2 Enterobacteria phage lambda exonuclease protein structure ..................... 25
Figure 1. 3 SynExo genes in most dsDNA viruses that infect Bacteria and Eukaryota .... 30
Figure 1. 4 Model of HSV1 UL12 structure .................................................................. 33
Figure 1. 5 Crystallographic structure of two ICP8 monomers ..................................... 34
Figure 1. 6 ICP8 ring ultrastructure .............................................................................. 35

Figure 2. 1 Cell cycle analysis with propidium iodide using BD Accuri gates .............. 132

Figure 3. 1 Chromophore region determines the spectral properties of fluorescent protein genes ........................................................................................................................................ 137
Figure 3. 2 E. coli Recombineering reporter strain ....................................................... 140
Figure 3. 3 Recombineering strain validated as a λ prophage lysogen ............................ 141
Figure 3. 4 E. coli GFPmut3* Recombineering variants are distinguishable ................. 144
Figure 3. 5 Aqua, Amarillo and GFPmut3* E. coli created by Recombineering, distinguished by confocal microscopy ................................................................. 145
Figure 3. 6 Spectra of fluorescent proteins Aqua, Violeta, Amarillo and Bronze created by Recombineering from gfpmut3* ................................................................. 146

Figure 4. 1 Replication fork model across gfpmut3* and oligo targeting the lagging strand template ................................................................. 152
Figure 4. 2 gfpmut3* gene is in opposite orientation to the replication fork ................. 153
Figure 4. 3 Oligos used for fluorescent protein Recombineering .................................. 155
Figure 4. 4 E. coli Recombineering reporter that avoids MMR ..................................... 156
Figure 4. 5 ssDNA Recombineering knowhow .............................................................. 158
Figure 4. 6 Oligo sequence present in one out of eight DNA strands in exponentially growing E. coli ................................................................. 160
Figure 4. 7 ICP8 expression in E. coli ............................................................................ 174

Figure 5. 1 Lentiviral vector to deliver an eGFP variant Recombineering target transgene. ........................................................................................................................................ 177
Figure 5. 2 Chromophore region determines the spectral properties of fluorescence ... 179
Figure 5. 3 Fluorescent proteins created from egfp by Recombineering and SDM have different spectra ................................................................. 181
Figure 5. 4 eGFP variants at the flow cytometer ........................................................... 182
Figure 5. 5 MIAMI human stem cells expressing eGFP and Mostaza ......................... 187
Figure 5. 6 MIAMI-eGFP and MIAMI-Mostaza cells are distinguishable by flow cytometry ................................................................. 189
Figure 5. 7 Lipofectamine 2000 delivers oligos to the nucleus of human cells ............ 192
Figure 5. 8 Lipofectamine 2000-oligo complexes transfect 90 % of 293T cells .......... 194
Figure 5. 9 Oligos for fluorescent protein engineering in human cells ........................ 196
Figure 5. 10 Mismatch produced during annealing of oligos to Mostaza target .......... 197
Figure 5. 11 ICP8-GFP localizes to the nucleus in MIAMI cells ................................. 200
Figure 5. 12 ICP8 expression in 293T cells detected by Western blot ......................... 201
Figure 5. 13 ICP8 expressed from inducible RheoSwitch plasmids ............................ 202
Figure 5. 14 ICP8-GFP expression correlates with ligand concentration .............................................. 203
Figure 5. 15 Synaptases expressed from Tet inducible lentiviral system ........................................... 206
Figure 5. 16 ICP8 and HumBeta expressed from pSLIK transgenes in 293T cells .................. 207
Figure 5. 17 ICP8 and HumBeta synaptases localize to the nucleus .............................................. 208
Figure 5. 18 293T-Mostaza-pSLIK and 293-Mostaza-pSLIK cells transduction efficiency and inducibility ................................................................. 210
Figure 5. 19 293T cells gene targeting .............................................................................................. 212
Figure 5. 20 ICP8 stimulates recombination .................................................................................... 214
Figure 5. 21 pCMV-ICP8-GFP transfection is lower than for other plasmids ................................. 218
Figure 5. 22 Recombinant phenotype decreases over time .............................................................. 219
Figure 5. 23 ICP8 expression produces green cells .......................................................................... 221
Figure 5. 24 pCMV-ICP8 transfection efficiency correlates with loss of Mostaza gene expression ........................................................................................................ 221
Figure 5. 25 pCMV-ICP8-GFP transfection induces cell death ....................................................... 223
Figure 5. 26 pSLIK transgene expression after zeocin selection ..................................................... 224
Figure 5. 27 Gene targeting promoted by ICP8 and antisense oligos, but not by HumBeta. .......... 226
Figure 5. 28 Enrichment for recombinants by cell sorting ............................................................. 237
Figure 5. 29 Recombinant green genotype detected by allele specific PCR and sequencing ........ 238
Figure 5. 30 Recombinant green cells population decreases over time ........................................... 240
Figure 5. 31 Recombinant green cells decreases at least 35% ....................................................... 241
Figure 5. 32 ICP8 promotes green fluorescent cells ....................................................................... 244
Figure 5. 33 Cell cycle analysis of cells expressing ICP8 ............................................................... 247
Figure 5. 34 ICP8 and ICP8 + oligo reduced the number of Mostaza fluorescence cells. ............. 250
Figure 5. 35 Mostaza fluorescence is impaired upon doxycycline induction in pSLIK4 ............... 255

Figure 6. 1 Mutation in LMNA activates a cryptic splice site, leading to aberrant splicing of pre-lamin A message ........................................................................ 258
Figure 6. 2 Progerin reporter HeLa cell line used to quantify Recombineering ............................. 260
Figure 6. 3 The Progerin-GFP excitation spectrum partially overlaps the dsRed excitation spectrum ................................................................. 261
Figure 6. 4 Progerin reporter HeLa cell line fluoresce green and red ............................................. 262
Figure 6. 5 Recombination assay: loss of green fluorescence ......................................................... 263
Figure 6. 6 Quantification of mock recombinants with Celigo ....................................................... 264
Figure 6. 7 Oligos to target the Progerin minigene ......................................................................... 266
Figure 6. 8 Western Blot analysis of ICP8 expression ................................................................. 267
Figure 6. 9 ICP8-GFP localizes to the nucleus in the progerin reporter cell line ...................... 268
Figure 6. 10 Quantification of putative recombinants by flow cytometry, exp 1 .............. 271
Figure 6. 11 Quantification of putative recombinants by flow cytometry, exp 2 .............. 272
Figure 6. 12 Isolating Green‘Red’ clones with Celigo ................................................................. 274
Figure 6. 13 Scheme of Progerin Reporter .................................................................................. 276
Figure 6. 14 PCR amplification from genomic DNA of isolated clones ......................................... 277
Figure 6. 15 Co-oriented repeated sequence flanking eGFP-5’ and dsRed2-3’ in Progerin reporter ................................. 282

Figure 7. 1 Beta forms a complex with λExo that disappears in the presence of RecA. 284
Figure 7. 2 λ SynExo genes were clone fused to a Halo tag for MS analysis. .................. 285
Figure 7. 3 Host proteins co-purifying with λ SynExo proteins to be identified by MS. .......... 286
Figure 7. 4 λ SynExo and E. coli interacting partners were analyzed by MS. ................. 287
Figure 7. 5 Full and truncated λ Exo protein structures for MS studies. ......................... 288
Figure 7. 6 λ SynExo interactome with host proteins .................................................. 291
Figure 7. 7 λ SynExo indirect interactions with DnaT (PriABC, DnaBC) indicates
initiation of recombination-dependent DNA replication by primosome assembly. 293
Figure 7. 8 λ SynExo direct interactions with UvrD helicase (MutU) indicates initiation
of recombination by primosome assembly and resolution by nucleotide excision repair and DNA unwinding................................................................. 294
Figure 7. 9 λ SynExo indirect interactions with RuvABC and LigA indicate resolution of
recombination by Holliday junction cleavage and sealing of residual ssDNA nicks
following resolution.................................................................................................. 295
Figure 7. 10 λ SynExo indirect interactions with UvrC and RuvC indicate resolution of
recombination by nucleotide excision repair and Holliday Junction cleavage,
initiation of recombination by strand invasion via RecA, and DNA replication via
the DnaN sliding clamp component of the replisome........................................... 296
Figure 7. 11 λ SynExo directly interacts with Lon protease while indirect interaction with
other proteases and chaperones indicate λ SynExo regulation or disassembly from
DNA via proteases. .............................................................................................. 297
Figure 7. 12 Recombineering reporter to evaluate hypomorphic and hypermorphic alleles
of λ SynExo Interactome. .................................................................................... 300
Figure 7. 13 ReDa assay: λ Exo resects dsDNA 5’ → 3’ producing 3’ ssDNA
overhangs and free dNMPs.................................................................................. 301
Figure 7. 14 Nascent ssDNA inhibit λ Exo activity but increases λ Exo dwell time. .... 302
Figure 7. 15 ssDNA annealing proteins overcome the inhibition of λ Exo by ssDNA. 303
Figure 7. 16 ssDNA increases λ Exo processivity but blocks reinitiation of DNA
digestion by λ Exo upon dissociation from partially degraded DNA ...................... 304
Figure 7. 17 RecA increases processivity and dwell time and protects nascent ssDNA
recombination substrate from Exol resection. ................................................. 307
Figure 7. 18 Beta blocks Exol from resecting nascent ssDNA recombination substrate.
......................................................................................................................... 308
Figure 7. 19 Beta improves λ Exo processivity by increasing dwell time .................. 308

xvi
LIST OF TABLES

Table 1. 1 Examples of SynExo Functional Modules.................................................. 31

Table 2. 1 Human cell lines ......................................................................................... 46
Table 2. 2 E. coli strains for bacterial Recombineering................................................... 48
Table 2. 3 E. coli strains for human Recombineering..................................................... 49
Table 2. 4 E. coli strains for SynExo interactome studies .............................................. 52
Table 2. 5 Oligo list .................................................................................................... 54
Table 2. 6 Plasmids for studies in human cells ............................................................. 59
Table 2. 7 Plasmids for studies in bacteria ................................................................. 61
Table 2. 8 Allele specific PCR ..................................................................................... 104

Table 4. 1 MMR differently affects different mismatches. .................................... 169
Table 4. 2 Optimized fluorescent protein engineering frequencies. ........................... 172
Table 4. 3 Rules for high efficiency Recombineering ............................................. 172

Table 5. 1 Oligo transfection frequencies in 293T cells using Lipofectamine 2000 ...... 193
Table 5. 2 Six lentiviral constructions were created .................................................... 206
Table 5. 3 Human Recombineering protocol by synaptase transient transfection..... 213
Table 5. 4 Gene targeting frequencies and statistics for pSLIK1 ............................... 227
Table 5. 5 Gene targeting frequencies and statistics for pSLIK4 ............................... 230
Table 5. 6 Gene targeting frequencies and statistics for pSLIK5 ............................... 233
Table 5. 7 Gene targeting efficiencies after subtracting oligo-independent events ...... 251

Table 6. 1 pCMV-ICP8-GFP transfection efficiency ............................................. 268
Table 6. 2 Primers used to amplify the Progerin reporter .......................................... 276

Table 7. 1 λ SynExo interacting partners identified by MS or by functional assays .... 289
Table 7. 2 Beta increases processivity by increasing dwell time ............................... 310
Table 7. 3 Mg^{2+} requirements by ReDa reactions .................................................... 311
Table 7. 4 SynExo from λ phage and HHV1 show significant protein interaction parallels. .......................... 314
LIST OF EQUATIONS

Equation 2.1 Nucleotides released per complex ......................................................... 118
Equation 2.2 Change in fluorescence ........................................................................... 118
Equation 2.3 Total number of nucleotides released per complex .............................. 119
Equation 2.4 Plateau ...................................................................................................... 119
Equation 2.5 Dwell time ............................................................................................... 119

Equation 7.1 Processivity ............................................................................................. 309
Equation 7.2 Dwell Time ............................................................................................. 309
LIST OF ABBREVIATIONS

293: HEK 293 cell line
293T: HEK 293T cell line
AKA: Also known as
Amp: Ampicillin
Carb: Carbenicillin
CCW: Counter-clockwise
CEF: CMV-EF1a hybrid promoter
CFP: Cyan fluorescent proteins
CH$_3$CN: Acetonitrile
Cml: Chloramphenicol
CMV: Cauliflower mosaic virus promoter
CW: Clockwise
DAPI: 4’,6-diamidino-2-phenylindole
DMEM: Dulbecco’s modified eagle medium
dNMPs: Deoxyribonucleotide monophosphates
Dox: Doxycycline
DSB: dsDNA break
dsDNA: Double-stranded DNA
DTT: Dithiothreitol
E. coli: Escherichia coli
eBFP: Enhanced blue fluorescent proteins
eGFP: Enhanced green fluorescent proteins
FBS: Fetal bovine serum
FL-oligos: Fluorescein-conjugated oligos
HA: Influenza virus hemagglutinin epitope YPYDVPDYA
HEK: Human embryonic kidney
HHV1: Human Herpes Virus Type 1 (AKA HSV1)
HR: Homologous genetic recombination
HSV1: Herpes Simplex Virus Type 1 (AKA HHV1)
HumBeta: Human-codon optimized NLS/HA-Beta for expression in human cells
Hyg: Hygromycin
ICC: Immunocytochemistry
ICP8: HHV1 gene UL29 and its protein product ICP8 (infected cell protein 8)
Kan: Kanamycin
LB: Luria Broth
LNA: Locked nucleic acid-protected ssODNs with a 2′-oxygen 4′-carbon methylene bridge in ribose.
MIAMI: Marrow-isolated adult multi-lineage inducible adult human stem cells
MMR: DNA mismatch repair
MS: Mass spectrometry
MSC: Mesenchymal Stem Cells
NEM: N-Ethyl morpholine acetate
NH₄HCO₃: Ammonium bicarbonate
NHEJ: Non-homologous end-joining pathway of dsDNA break repair
NLS: Nuclear Localization Signal (for SV40 virus T antigen NLS sequence is PKKKRKVED)
nts: Nucleotides
Oligos: Single-stranded DNA oligonucleotides, typically under 100 nucleotides
PAGE: Polyacrylamide gel electrophoresis
PI: Propidium Iodide
ReDA: Real-time DNase assay
RSL: RheoSwitch ligand
SDM: Site directed mutagenesis
ssDNA: Single-stranded DNA
TB: Terrific Broth
TFA: Trifluoroacetic acid
YFP: Yellow fluorescent proteins
Zeo: Zeocin
CHAPTER 1. GENOME ENGINEERING INTRODUCTION

Significance of efficient human nuclear genome engineering

The revolutions in stem cell therapy and DNA sequencing technologies are providing the tools for personalized medicine and a broader understanding of human cell development. Autologous stem cell therapy is really promising for medicine, but even more promising is the ability to repair mutations or DNA damage in our own stem cells for autologous stem cell gene correction. Advances in human genome sequencing technology are providing readily available information for geneticists to be able to identify important contributions of genotype to phenotype, but for a truly deep and wide view of functional genomics it is necessary to manipulate the existing DNA sequences. Therefore, the ability to edit genome sequences is not only promising for personalized genetic medicine, but it is also an essential tool for understanding the foundations of human biology. Much remains to be learned about our genome, and development of efficient and specific tools for genome editing is essential to dissect how different protein coding alleles and allelic combinations affect human health and disease, how different splice variants affect development and cell regulation, how copy number variation modulates phenotype and phenotypic dynamics, and many other important topics.

Human genome engineering

Emerging genome editing technologies have dramatically expanded functional genomic analysis of model organisms and these advances encourage the development of methods to correct the genetic causes behind many human diseases. As high quality genome sequence data has become abundant, site-specific modification of genomes now
provides a powerful approach to functionally characterize alleles to distinguish between phenotypically inconsequential polymorphisms and informative mutations. However, in order to achieve the full potential of this technology, many important questions and challenges must be addressed. Unfortunately, current gene targeting techniques are not efficient enough (in humans) for routine bioengineering and disease model creation, and the most commonly used technologies are too mutagenic (nonspecific) to be suitable for therapeutic use.

Current technologies for human genetic engineering include using viral infections to provide a protein of interest (usually used for gain-of-function), introduction of targeted double-stranded DNA (dsDNA) breaks via Zinc Finger, TALEN and the CRISPR/Cas-based nucleases to disrupt genes via mutagenic repair of dsDNA breaks or by introducing a complementary donor DNA for repair of dsDNA breaks by homologous recombination, and gene targeting using single-stranded DNA (ssDNA) oligonucleotides (oligos) to introduce small nucleotide changes, deletions and small insertions. Wide use of all of these emerging technologies has been limited by their low efficiency and tendencies toward cell toxicity.

The use of viruses to make genetic additions to mammalian cells has been very useful because they can efficiently infect cells and either be maintained as extragenomic episomal DNA elements or get integrated into the genome as permanent additions. In this work we have used several lentiviral constructs to deliver genes to human cells because they efficiently integrate transgenes into the genome. One of the disadvantages of this approach is that integration is essentially random and DNA sequence analysis is necessary to determine the site of transgene integration in the genome. Nevertheless,
some of these viruses (e.g. lentiviruses, (McGinley et al. 2013; Akhtar et al. 2013; Chen et al. 2013) have been developed to treat diseases and have been studied in clinical trials (Biffi et al. 2013; Aiuti et al. 2013). However one unwelcome side effect of integration into the genome is the occasional activation of oncogenes (Themis 2012) that have been associated with tumor formation (Nowrouzi et al. 2013). Another viral system being currently developed for genetic medicine uses Adeno-associated virus (AAV, Willett & Bennett 2013), which is not pathogenic (Zhang et al. 2011b), though it does induce immune responses in humans (Hareendran et al. 2013). Efforts have been put into engineering these viruses to use as episomal replicons as the advantage would be that they would not integrate into the genome and cause insertion mutations or directly modify gene expression in unintended ways. The disadvantage is that episome maintenance is impermanent (Pacak et al. 2008) so their use is transient, which can be seen as a treatment but not as a cure. Clearly there is a need to develop strategies for permanent modification of the genome at precise locations to maintain stably inherited engineered genotypes without causing unintended off-target consequences.

Over the last thirty years, great strides have been made in dissecting the molecular underpinnings of homologous recombination since the seminal publication of the idea that genetic recombination is associated with repair of dsDNA breaks (Szostak et al. 1983a). One of the most used modern technologies for site-specific genome modification is based in inducing dsDNA breaks (DSBs) at a target site using either Zinc Finger or TALEN (Transcription Activator-Like Effector Nuclease) nucleases (Perez-Pinera et al. 2012). Zinc Finger or TALEN nucleases have been developed for a wide range of target regions. The use of these nucleases has been tested to destroy disease-associated alleles in
mitochondrial DNA (Smith & Lightowlers 2011) or to destroy the HIV co-receptor CCR5 gene in bone marrow cells (Lai 2012). These nucleases introduce dsDNA breaks that are subsequently repaired by cellular DNA damage repair mechanisms, predominantly via non-homologous end joining (NHEJ) and less frequently by homologous recombination (HR) (Coluccio et al. 2013; Maier et al. 2013; Gantz et al. 2012; Gaj et al. 2012). NHEJ is useful to create loss of function alleles as it is often associated with deletion mutations introduced by DNA resection at the break site that precedes ligation to repair the break. HR is useful for targeted transgene insertion, precise gene deletion and for sequence editing via gene conversion.

While HR sounds ideal, the frequencies of induced targeted HR are typically on the order of $10^{-7}$ to $10^{-5}$ making it difficult to isolate recombinant clones without an associated selection against non-recombinant clones. Some groups have combined introduction of targeted dsDNA breaks with transfection of a template DNA that has homologous targeting sequences from which to repair the damage, thereby introducing information from the template into the genome. This exogenous template DNA brings the possibility to insert additional sequences (e.g. a drug resistance selection “cassette”) to facilitate isolation of recombinant cell lines. When selection is applied, it is possible to isolate clones with the desired sequence at a frequency of about $10^{-1}$ of cells surviving selection (i.e. 10% of ~$10^6$). Unfortunately, this selection process is not feasible for all genome editing and it is not desirable to introduce selection cassettes as permanent genome features. While it is possible to remove selection cassettes that are flanked by loxP sites via subsequent Cre-mediated site-specific recombination, this too leaves additional foreign DNA at the residual “scar” of the site-specific recombination reaction.
(Davis et al. 2008). For genome modification to be generally acceptable as a therapeutic approach, it would be best to avoid leaving foreign DNA in the recombinant human genome. This motivates the search for much more efficient genome modifying technologies that permit isolation of recombinants without needing a selection.

Recently, RNA-directed degradation of complementary sequences of invading viral or plasmid DNA was developed to engineer human genomes using natural defense mechanisms discovered in Bacteria and Archaea. CRISPR/Cas-based nucleases are so named for the Clustered Regularly Interspaced Short Palindromic Repeat system which consists of targeting RNA in complex with the Cas nuclease (Jinek et al. 2012). Mali et al. (2013b) used a modified CRISPR system for human genome engineering and obtained targeted deletion rates of 10 to 25% in 293T cells, 8 to 13 % in K562 cells, and 2 to 4% in induced pluripotent stem cells. When a homologous donor DNA was introduced to direct HR to repair the CRISPR-induced dsDNA break, investigators were able to repair targeted loci using the supplied donor sequence (Mali et al. 2013b) or multiple donor sequences (Cong et al. 2013), albeit at lower frequencies than obtained for gene disruption (Mali et al. 2013b). Unfortunately, recent reports indicate that, like Zinc-finger and TALEN nucleases (Gabriel et al. 2011), the CRISPR system is plagued with off-target genome modification effects (e.g. Fu et al. 2013; Cradick et al. 2013), though opportunities for improvement have been proposed (Ran et al. 2013; Mali et al. 2013a).

In parallel, gene targeting using ssDNA oligos has been optimized for many organisms, including humans (Ellis et al. 2001; Costantino & Court 2003; Li et al. 2003; Radecke et al. 2004; Kotzamanis et al. 2005; Aarts et al. 2006; Dekker et al. 2006; Maguire & Kmiec 2007; Flagler et al. 2008; Larrea et al. 2008; Sangiuolo et al. 2008;...
Sawitzke et al. 2011; Valledor et al. 2012). Gene targeting using oligos has been employed in many cell lines, including mouse ES, HeLa, CHO, HEK293 and 293T at efficiencies ranging from $10^{-5}$ to $10^{-1}$. Gene targeting in mammalian cells resembles the central ssDNA intermediate in the mechanism proposed for Recombineering in bacteria (Vellani & Myers 2003), where small oligos are directed by ssDNA binding and annealing proteins (synaptases) to complementary regions in the genome exposed in the context of the replication fork. This technology is promising since the ssDNA can be targeted to any region of the genome without the need to produce dsDNA breaks at the target region, which can be very toxic to primary cells. Unfortunately the low frequency of recombinants in mammalian cells, oligo treatment-dependent cell arrest and low survival of recombinants limits the use of this technology for easily generating new cell lines or for gene correction therapies when targeting endogenous genes that do not allow for selection.

Meanwhile, microbial genome manipulation has reached its apex with gene targeting rates so high that selection for recombinants is no longer necessary. The high efficiency and specificity of recombination in bacteria is obtained by using bacteriophage (phage) recombinases to place ssDNA recombination substrates in replicating DNA targets, "tricking" cells into "seeing" recombination intermediates as normal Okazaki fragments. Bacterial Recombineering catalyzes highly efficient in vivo homologous recombination by using a complex of two phage proteins: one protein is an Exonuclease that processes linear substrate dsDNA to create ssDNA ends; the second protein is a Synaptase that is loaded onto the ssDNA product of the nuclease to form a protein-DNA filament that catalyzes homologous DNA pairing and strand exchange reactions. Using those enzymes
it is now possible to engineer replicons in all possible ways in *E. coli* and several other bacteria at rates of $\sim 10^{-3}$. By using oligos in place of dsDNA, Recombineering becomes extremely efficient (Sawitzke *et al.* 2011; Valledor *et al.* 2012) and reproducible site-specific modification of bacterial genomes (or genome segments of other organisms carries as bacterial artificial chromosomes) is possible without requiring selection for recombinants. Because oligos resemble replication intermediates and require no processing by viral recombinases, recombination efficiency is orders of magnitude higher with oligo substrates than with dsDNA substrates, though the overall efficiency is very dependent on the sequence and structure of the encoded substitution and the DNA mismatch created in the recombination intermediate (Sawitzke *et al.* 2011; Valledor *et al.* 2012; Costantino & Court 2003a; Kolodner 1996).

This study builds on the lessons from phage-mediated Recombineering in bacteria to explore gene targeting in human cells. The development of host-specific viral recombinases will bring Recombineering to biomedical research. Efficient *in vivo* engineering of human DNA will improve human functional genomics, creation of disease models, and even gene and cellular therapy. Human stem cell Recombineering will transform biotechnology and medicine.

**Human genome engineering by gene targeting**

Oligo-mediated gene conversion using endogenous cellular recombinases is much more efficient in mammalian cells than in *E. coli* (Ellis *et al.* 2001; Costantino & Court 2003; Li *et al.* 2003; Radecke *et al.* 2004; Kotzamanis *et al.* 2005; Aarts *et al.* 2006; Dekker *et al.* 2006; Maguire & Kmiec 2007; Flagler *et al.* 2008; Larrea *et al.* 2008; Sangiuolo *et al.* 2008). The frequencies of gene targeting vary between different labs and
cell lines from $10^{-5}$ to $10^{-1}$. Unfortunately, there are some issues with recombinant cell survival or recovery that make the use of this technology to do functional genomics or for gene therapy difficult. Interestingly, the mechanism of mammalian cell gene targeting resembles bacterial Recombineering in that they both can use ssDNA oligos as the recombination substrate, there is a strand bias in the efficiency of the oligo substrate (Aarts & te Riele 2010b), and gene targeting is stimulated in replicating cells (Olsen et al. 2005a).

**ssDNA oligos as substrates for gene targeting**

The appearance of ssDNA as a central intermediate is universal to recombination. For gene targeting in mammalian cells, cells are transfected with ssDNA oligos that target a specific region of the genome or a replicon by Watson-Crick base-pairing to a complementary single-stranded region of the genome. The most natural substrate for gene targeting is the same material as is most of our genetic information, unmodified DNA. However, modified oligos and alternative base analogs produced by synthetic chemistry have been shown to increase the efficiency of gene targeting, albeit with undesirable side-effects. The ssDNA oligo substrate is typically designed to be complementary to a target region in the genome, differing in sequence only by the desired change. This change is usually positioned by investigators in the middle of the ssDNA oligo sequence to improve efficiency of incorporation into the genome. The complementary region is necessary for the pairing of the oligo substrate to the target gene, presumably catalyzed by endogenous cellular synaptases. The introduced change could be a small substitution, deletion or small insertion of ssDNA. In bacteria, ssDNA oligos have been used to delete up to 50,000 bp (RS Myers unpublished, DL Court, personal communication).
The effect of substrate ssDNA oligo length has been mostly evaluated over the range of 20 to 80 nucleotides (Nickerson & Colledge 2003). Shorter oligos (20-25 nt) require modification to be effective as they are more susceptible to nucleases (Rios et al. 2012; Andrieu-Soler et al. 2005), but could be more effective in finding exposed complementary regions of ssDNA with which to anneal between Okazaki fragments. Longer oligos might promote the homology search by providing both a larger sequence for target identification and by increasing binding free energy when paired with the target (assuming the target region is single-stranded along the entire length). Longer oligos might benefit from catalysis by endogenous synaptases to locate and pair the oligo substrate to the target sequence. In vitro studies have shown the annealing reaction of ssDNA is 3500-fold faster when the cellular synaptase protein Rad52 is included (Mortensen 1996). Synaptases are also predicted to form more stable oligo/synaptase filaments with longer oligos as, for example, each Rad 52 monomer binds about 10 nts (Shinohara et al. 1998). Stable nucleoprotein complexes of oligos and Rad52 have been reported with ssDNA from 41 nt (Mortensen 1996) to 3,000 nt long (Shinohara et al. 1998).

When designing the oligos and the substitution to be incorporated into the target genome, it is very important to consider the effects of DNA mismatch repair (MMR). When the oligo is annealed to the target region this change will produce a mismatch that can be recognized and reverted by the MMR function in the cell. Most of the studies in mammalian cells have been done where the mismatch formed in the biparental heteroduplex recombination intermediate is from 1 to 4 nucleotides. For example, during the repair of a start codon for gfp in HeLa cells, an oligo including a T.T mismatch
produced more recombinants than an oligo including an A.A mismatch and both were better than an oligo including a G.T mismatch (Rios et al. 2012).

Oligos are usually delivered using chemical transfection reagents like Lipofectamine 2000 (Rios et al. 2012; Flagler et al. 2008; Andrieu-Soler et al. 2005) or Transpass (Aarts et al. 2006). Some other groups also use electroporation, the preferred method for oligo delivery to bacteria (Borjigin et al. 2012).

Oligo substrate protection by chemical modification

To avoid the nucleolytic degradation of the oligo substrate, different modifications in the oligo have been described like 2'-O-methyl RNA residues (Igoucheva et al. 2001; Nickerson & Colledge 2003), phosphorothioate (PTO) linkages (Liu et al. 2002a; Olsen et al. 2005b) or locked nucleic acid (LNA) bases (Andrieu-Soler et al. 2005). While modified oligos usually produce more gene targeting in transformed cell lines, such as in human colon cancer cells (Bonner et al. 2012) and in HeLa cells (Rios et al. 2012), Aarts and te Riele (2010a) found no benefits from modifying oligos when targeting mouse ES cells. What is common is that for shorter (25 nt) oligos, most LNA modifications increase gene targeting (Aarts & te Riele 2010a; Andrieu-Soler et al. 2005).

The influence of modification location has been tested throughout oligo sequences. Radecke et al. (2006) found that modifications at the 3’ end inhibit oligo-mediated gene targeting while modifications in the middle or toward the 5’ end increase frequencies of oligo targeting.

Modifications in the oligos have been shown to have toxic effects in the cells. Several reports note that cells targeted by chemically modified oligos underwent a G2 arrest that precluded their clonal expansion (Olsen et al. 2005a; Andrieu-Soler et al. 2005). Rios et
al. (2012) showed that 1) the higher the content of modified bases, the greater the toxicity or lesser the recovery of recombinants; 2) modified oligos affect all transfected cells independent of the success of gene targeting; 3) there is a direct correlation between the oligo concentration that cells are exposed to and cell arrest; and 4) a higher number of PTO bonds correlates to higher c-H2AX phosphorylation – indicative of introduced dsDNA breaks – and higher cell mortality. Rios et al. (2012) also showed how oligo treatment changes global gene expression with upregulation of genes associated with cellular immunity and the DNA damage response. At this point it is not known if modifications exacerbate innate immunity or the DNA damage response when compared to non-modified oligos. Perhaps the stability of modified oligos provides a prolonged signal indicating foreign DNA invasion and/or the presence of DNA damage (persistent ssDNA is such a signal).

The fact that some groups see no additional effect when using modified oligos might be due to the toxic effect of modified oligos leading to lower recovery of genomes that incorporate the oligo. Consistent with this idea, if gene targeting is evaluated by the ability of the introduced oligo to change gene expression, then modified oligos show improved gene targeting rates (Rios et al. 2012). On the other hand, if gene targeting is evaluated by the ability of recombinant cells to proliferate and form colonies, then the negative effect of modified oligos on cell proliferation results in under-estimates gene targeting rates (Aarts & te Riele 2010a).

As mentioned previously, another potent inhibitor of gene targeting is MMR. Rios et al. (2012) reported that chemically modified base analogs capable of avoiding MMR in E. coli outperform gene targeting with unmodified oligos in HeLa cells. These
investigators reported, that 2-fluorouracil and 2-aminopurine outperform thymine and adenine, respectively, in creating T-T or A-A mismatches leading to A/T transversions when present in heteroduplex intermediates that give rise to recombinant cells.

**Cellular functions that can influence gene targeting**

**Cell cycle**

Gene targeting is more efficient when oligos are delivered during late G1 into early S phase (Olsen et al. 2005a). During this part of the cell cycle, cells prepare for genomic DNA duplication. Chromatin is opened, replisomes are recruited and activated, and ssDNA becomes transiently exposed between Okazaki fragments within traveling replication forks unwound by the replicative helicase. Homologous recombination proteins are more abundant in S phase and ready to assist stalled replication forks and to repair damage to the DNA template while NHEJ proteins are less present (Mao et al. 2008; Huertas et al. 2008; Aylon & Kupiec 2005; Aylon et al. 2004). As a consequence, HR outcompetes NHEJ for repair of DNA breaks in S phase, an observation that can be leveraged to improve gene targeting (Majumdar et al. 2003; Hu et al. 2005; Brachman & Kmiec 2005; Engstrom & Kmiec 2007; Engstrom & Kmiec 2008).

The mechanism of oligo gene targeting in mammalian cells is very similar to that in *E. coli* Recombineering and as has been seen in yeast. It is becoming universally appreciated that introduced oligos are physically incorporated in the context of active replication forks in between Okazaki fragments. It is clear that MMR is a potent inhibitor of oligo-directed gene targeting in all creatures (Dekker et al. 2006; Dekker et al. 2011; Aarts et al. 2006). However what is not known is how oligos cross the cytoplasm, enter the nucleus, find their targets and anneal to the genome. Finally, the consequences of having relatively large amounts of oligos inside cells, how cells eliminate excess oligos,
and the cellular responses to oligo invasion during genetic manipulation remain questions that need to be addressed.

Homologous recombination

Genetic recombination occurs as a consequence of DSB repair. Naturally induced DSBs promote recombination in meiosis to generate viable gametes for sexual reproduction (Boulton et al. 1997). In lymphoid cells, DSBs contribute to immunity via V(D)J recombination and class switch recombination (CSR) at the lymphoid antigen receptor loci (Schumacher et al. 2012; Feeney 2011; Tsai & Lieber 2010). Pathogens also use HR repair of DSBs for surface antigen variation to evade the host immune response (Glover et al. 2013). The realization that introducing a DSB in a cloned gene can target homologous integration of the plasmid into the genome of yeast was fundamental to the canonical view that HR is primarily a product of DSB repair in most organisms and DNA viruses (Szostak et al. 1983b). Subsequently, it has been shown that introducing a DSB into a genomic target and providing an exogenous template off of which to repair can shuffle alleles to the genome via HR. However, most natural DSBs arise during errors of DNA synthesis and most HR occurs in the context of DNA replication.

Formation of DSBs is the rate-limiting step in HR (Orr-Weaver et al. 1981; Puchta et al. 1993) and when applied to a genomic target, the most dangerous. The next rate-limiting step, and second most dangerous, is DNA resection by 5’ → 3’ dsDNA exonuclease activity to form ssDNA (Roman et al. 1991) with coincident “loading” of DNA pairing proteins to form a pre-synaptic filament (San Filippo et al. 2008; Jensen et al. 2013). Perhaps surprisingly, the search for homology and subsequent repair occurs at rates far exceeding that of spontaneous DSB formation and subsequent resection and is
not rate-limiting overall, despite the enormity of the task of sifting through millions or billions of DNA bp to find a homologous target (Yancey-Wrona & Camerini-Otero 1995; Waldman 1994). As a consequence, it is possible to greatly stimulate genome editing by providing linear (i.e. “broken”) ssDNA (i.e. “processed”) as a substrate, as was illuminated in the seminal DSB paper of Szostak et al. (1983a).

How do oligos find their targets?

The rates and frequencies for an oligo to find and anneal to a ssDNA target are strongly stimulated by the presence of a synaptase. The human Rad52 protein catalyzes ssDNA annealing 3500 fold faster than oligos alone in vitro (Mortensen 1996). In the complexity of the human cell nucleus, where DNA is compacted in chromatin structures, the chance for a ssDNA oligo to find the target locus is even more difficult. Only during S phase is chromatin relaxed around most targets. Furthermore, ssDNA is most abundant in the context of replication forks in between Okazaki fragments during S phase. The tight association between genome editing and replication will be illuminated throughout this document.

The ssDNA in replication centers is covered with the RPA protein which melts secondary structures and is a binding partner of Rad52 (Gasior et al. 1998; Haaf et al. 1995; Tashiro et al. 1996; Maser et al. 1997; Scully et al. 1997; Bishop et al. 1998). In the presence of DNA breaks, Rad52 and RPA form punctate foci in subnuclear domains and collaborate in DNA repair in conjunction with many other proteins. One can imagine a scenario where a ssDNA oligo substrate forms a nucleoprotein filament with Rad52 in the nucleus that searches for complementary RPA-coated ssDNA in replication centers.
How does ssDNA annealing occur?

Rad52 is a synaptase that anneals ssDNA to ssDNA. ssDNA in the genome is mostly found during replication between Okazaki fragments and a bit during transcription. Since there is substantial evidence that gene targeting is more effective during S phase, it is likely that annealing occurs in between Okazaki fragments. Furthermore, the Rad52 interacting partner RPA covers the ssDNA at the replication fork, providing strong suggestions that Rad52-oligo filaments could scan the genome DNA at the replication forks. Rad52 also interacts with Rad51, a near-universally conserved RecA family strand invasion protein, which could promote strand invasion into unwound dsDNA targets. Support for such a model was recently obtained in murine myoblasts (Bertoni et al. 2009). In this scenario there should not be a strand bias associated with gene targeting. However, the gene targeting strand bias observed in many groups argues against a general strand invasion mechanism and favors the mechanism for annealing at the lagging strand template in between Okazaki fragments as will be illuminated at length in this document.

Okazaki fragments in mammalian cells are much smaller than in *E. coli* (150-200 bp in mammalian vs. ~1000 bp in *E. coli*). The “ruler” for establishing Okazaki fragment length in mammalian cells is thought to be the length of DNA bound by nucleosomes (Balakrishnan & Bambara 2013). It is interesting that the most used oligos for mammalian gene targeting are smaller than the ones used in Bacteria (~35 vs. ~70 nt). Maybe the optimal size for efficient incorporation of oligos relates to the average ssDNA length in between Okazaki fragments.
Some investigators have reported that the antisense oligo is more efficient in their system than the sense oligo (Nickerson & Colledge 2003; Igoucheva et al. 2001; Liu 2002b; Olsen et al. 2005b). These observations have raised the hypothesis that oligo annealing occurs during transcription. In such a case gene targeting should depend more on the transcription level of a gene than on the cell cycle phase and that gene induction should stimulate targeting. Aarts and te Riele (2010a) did a gene targeting experiment with an inducible promoter and evaluated the rates of gene targeting with and without transcription. They found that transcription did not stimulate gene targeting. In addition, there is evidence that gene targeting is more effective at the G1/S transition (Olsen et al. 2005a). Therefore, we discard the hypothesis that the oligo is annealed to ssDNA exposed during transcription and favor the lagging strand template between Okazaki fragments as the primary target of strand annealing and genome editing by oligos.

Physical integration and strand segregation

The evidence thus far suggests that targeting oligo DNA is annealed and ligated to the nascent DNA strand during DNA synthesis. Evidence includes: stable gene expression and phenotypic output as early as 24 hours after oligo exposure; the ability to develop recombinant mouse lines; and corroboration of genotypes by sequencing and allele specific PCR (Aarts & te Riele 2010b).

If the ssDNA gets incorporated into nascently replicated DNA and does not get removed by MMR, a damage repair nuclease, or a helicase, the oligo will form a biparental heteroduplex with the lagging strand DNA template. If all the strands in a replicating chromosome are counted, the oligo allele will account for 1/4 of the DNA strands in the duplicating DNA at the target locus (assuming a single target locus in the
genome; a diplotype locus will reduce this to 1/8). Evidence for physical incorporation includes the observation that the frequency of cells showing the recombinant phenotype decreases during the first 4 replication cycles. Aarts et al (2010) evaluated the frequency of recombinants during a one week period of strand segregation and their data supports the strand segregation model.

**Inhibitors of gene targeting**

Exonucleases

In bacteria, exonucleases are powerful inhibitors of oligo mediated gene targeting (Mosberg et al. 2012; Sawitzke et al. 2011). In mammalian cells the use of modified protected oligos has been shown to increase the frequency of gene targeting (Rios et al. 2012). The major 3’ DNA exonuclease in mammalian cells is Trex1 (de Silva et al. 2009). Trex1 associates with the endoplasmic reticulum and degrades ssDNA. During S phase, Trex1 translocates to the nucleus upon DNA damage (Christmann et al. 2010), indicating that Trex1 could clear the oligo substrate in the cytoplasm and in the nucleus.

Trex1 deficiency results in constitutive activation of the ATM-dependent DNA-damage checkpoint, as indicated by destabilization of Chk2, a key transducer kinase in the pathway, and increased levels of the tumor suppressor p53 and the downstream CDK inhibitor p21. Trex1 deficient fibroblasts are impaired in the G1/S phase transition (Coscoy & Raulet 2007). The temporary to permanent arrest observed in cells transformed by ssDNA might resemble the G1/S phase transition arrest in Trex1-deficient fibroblasts described in (Coscoy & Raulet 2007). The arrest could be especially true if Trex1 cannot get rid of modified oligos. We imagine that if ssDNA is not quickly removed by Trex1, innate immune response activation promotes cell cycle arrest.
DNA polymerase and Fen1

On the human lagging strand template, Pol δ elongates the primers and closes the gap between two Okazaki fragments. Pol δ has strand displacement activity, therefore on encountering the down-stream Okazaki fragment, Pol δ continues synthesis, unwinding the 5′ end of the preceding Okazaki fragment into a single-stranded 5′ flap structure. FEN1 or Dna2 recognizes this structure, binds to the base of the flap, and cleaves it. In normal DNA synthesis the flap consist of the RNA covalently bonded to nascent DNA (Balakrishnan & Bambara 2013). In gene targeting, Pol δ could displace the 5′ end of the oligo substrate to be incorporated into the flap cleaved by FEN1 or Dna2. Aarts and te Riele (2010b) showed that modifications at the 5′ end promote gene targeting. Maybe strand displacement synthesis or the 5′ to 3′ nuclease cleavage activities are impaired by 5′ modification of oligos.

MMR

MMR is a potent inhibitor of ssDNA oligo-mediated gene targeting. The nature of the hybrid DNA creates a kinked structure that is recognized by the MMR system, specifically by the heterodimers MSH2/MHS6 and MSH2/MHS3. The MSH2/MHS6 complex recognizes small mismatches and loops of 1 and 2 nucleotides while the MSH2/MHS3 heterodimer recognizes larger mismatches and loops of 2 to 5 nucleotides (Jiricny 2006). Upon mismatch recognition, the newly synthesized DNA strand (or the strand containing the annealed oligo in our case) is cut by MLH1 and subsequently removed from the DNA by ExoI. The resulting gap is then filled in with polymerase using the original template sequence and sealed with ligase. In this way, MMR recovers
the original genotype and decreases the final mutation frequency and/or gene targeting frequency.

Knocking down MMR increases gene targeting frequencies 150 to 200 fold but also increases the chance of random mutations in the cell (Dekker et al. 2011; Wei et al. 2002). Deficiency of the MSH6 gene increases probabilities of substitutions of 1-4 nt (Aarts et al. 2006). Deficiency of MSH3 increases gene targeting for insertions (Dekker et al. 2006). Deficiency of the common mismatch detecting subunit MSH2 increases gene targeting efficiency for both substitutions and insertions (Dekker et al. 2006). Deficiency of the master coordinator of MMR, MLH1, increases efficiencies for gene targeting substitutions by ssDNA oligos but also induces frameshift mutagenesis (Dekker et al. 2011). The authors also reported that inserting four nucleotides was more tolerable by MMR than insertion or deletion of one to three nucleotides. Some cell lines have impaired MMR and the degree of MMR proficiency could account for the degree of differences in gene targeting across cell lines. To this end, HEK 293T derived cells are deficient for hMLH1, hPMS2, hMLH3 (Cannavo et al. 2005), consistent with the higher gene targeting rates seen in this cell line.

**Cellular responses to gene targeting**

Almost all the studies of gene targeting report a gradual loss of the cells expressing recombinant phenotypes (Olsen et al. 2005a; Olsen et al. 2009; Ferrara & Kmiec 2006). It seems that most cells successfully transfected by oligos suffer and that it is not specific to the cells doing the gene conversion. However, oligo uptake positively correlates with recombination frequency (Aarts & te Riele 2010b), so finding ways to improve oligo uptake while minimizing induction of cellular responses to foreign DNA is well worth pursuing.
DNA sensors awake the anti-viral response

ssDNA oligo substrates in the cell cytoplasm are recognized by several DNA sensors (Fernandes-Alnemri et al. 2009; Zhang et al. 2011c). This recognition activates a potent inflammatory response with the induction of more than 300 IFN-stimulated genes (Atianand & Fitzgerald 2013; Sadler & Williams 2008) that could lead to apoptosis (Fernandes-Alnemri et al. 2009). The substrate for gene targeting is ssDNA oligos, which expose cells to this response. Indeed, when Rios et al. (2012) compared the transcription profiles of EGFP-corrected and -uncorrected cells via RNA-seq they found upregulation of many genes associated with cellular immunity. Perhaps this immune response is the cause of arrest and death observed in cells transformed by ssDNA and in the resulting recombinants.

DNA damage response

Oligo substrates have also been shown to trigger a DNA damage response, such as an increase in H2AX phosphorylation and activation of the ATM/ATR cell cycle checkpoint (Olsen et al. 2009; Bonner & Kmiec 2009; Ferrara & Kmiec 2006). Aarts et al. (2010) compared mock transfection to transfection by modified and unmodified oligos. The mock transfection produced an increase in the dsDNA break response (~25 fold), while the PTO-modified oligos produced a bigger increase (~125 fold) in H2AX phosphorylation (Aarts & te Riele 2010b). These results suggest that the transfection protocols and modified oligos reduced viability due to high levels of perceived DNA damage, though whether or not transfection reagents and oligos induced actual damage to the genome is not yet known.
Cell cycle arrest

Another consequence of gene targeting is cell cycle arrest (Olsen et al. 2005a; Papaioannou et al. 2009; Olsen et al. 2009). Aarts and te Riele (2010b) reported that cells corrected by ssDNA oligos showed a decrease in their ability to progress into mitosis (from 48% to 23% for corrected cells while from 48% to 44% for uncorrected cells), indicating the cell cycle was arrested for some time in the corrected cells when using non-modified oligos. When using PTO-modified oligos, only 10% of the recombinant cells were able to progress into mitosis, while 36% of the nonrecombinant cells progressed into mitosis (Aarts & te Riele 2010b). These results indicate that oligo transfection produces a moderate cell arrest, which is stronger for PTO-modified oligos. Recombinants showed the strongest arrest, especially for the PTO-modified oligo recombinants, but it is still unclear if the arrest is due to levels of oligo uptake and the ssDNA innate cellular sensing response or a consequence of pairing the ssDNA oligo to the chromosomal DNA and inducing a DNA damage response, or both.

**Bacterial genome engineering by Recombineering**

Pioneering in vitro studies (Cassuto & Radding, 1971; Cassuto et al. 1971), later validated in vivo (Stahl et al. 1997), revealed a simple mechanism for homologous recombination by phage \( \lambda \) operating in *E. coli* cells. Stahl and colleagues (Stahl et al. 1974; Stahl & Stahl 1976; Stahl et al. 1985; Stahl 1986; Stahl & Stahl 1986; Stahl et al. 1990; Stahl et al. 1997; Stahl 1998; Thaler et al. 1987a; Thaler et al. 1987b; Thaler et al. 1987c; Siddiqi et al. 1991; Hill et al. 1997) found that \( \lambda \) exonuclease (\( \lambda \) Exo) and Beta protein cooperate to create ssDNA recombination intermediates from linear dsDNA substrates. These proteins catalyze homologous recombination by pairing the
intermediates with homologous DNA target sequences, promoting the covalent joining of the recombinant DNA products with assistance from host proteins.

The first use of phage-recombination functions for generating recombinant bacterial genomes using transformed dsDNA employed the lambda Red system (Murphy 1998) followed by the Rec E/T system from the Rac prophage (Zhang et al. 1998). The use of viral recombination proteins to catalyze homologous recombination has been termed “Recombineering”. Bacterial genomes may be engineered directly by introducing linear dsDNA PCR products or ssDNA oligos into cells expressing a viral recombinase, like the λ Red system or the RacE/T. Viral genomes, such as that of HIV1 (Geffin and Myers, in preparation 2013) and HCMV (Britt et al. 2004), and mammalian genomes, such as that of the mouse (Copeland et al. 2001), have been manipulated as clones by Recombineering in bacteria. Any kind of modifications can be introduced in replicons or complementary ssDNA using this technology including insertions, deletions and substitutions. A strain collection was built using Recombineering in which each non-essential gene was deleted from E. coli (Baba et al. 2006). This collection has been important for functional studies in E. coli. Recently Recombineering has been automated to facilitate the incorporation of whole pathways into the E. coli genome (Wang et al. 2009). Recombineering has transformed genome engineering (Thomason et al. 2005).

The λ Red recombination functions are encoded by three adjacent genes, gam, bet and exo, expressed within the λ pL operon (Thomason et al. 2005). The gam gene encodes the protein Gam that stimulates recombination by preserving DNA substrates from exonucleolytic degradation by RecBCD and SbcCD (Kulkarni & Stahl 1989; Murphy 1991). The exo gene encodes the protein λ Exo that forms a complex with Beta and
resects dsDNA 5′ → 3′ to produce ssDNA substrates for Beta-mediated homologous DNA pairing. The bet gene encodes the protein Beta, a ssDNA binding protein that both protects ssDNA recombination intermediates from nucleases and catalyzes annealing of complementary ssDNA to promote genetic exchange.

Gam structure and function

Gam is a small protein expressed along with the SynExo of phage λ in the P₇ operon. Gam forms a homodimeric complex that binds to the E. coli RecBCD enzyme and inhibits its nuclease and helicase activities by acting as a DNA mimetic (Court et al. 2007). In the absence of Gam, RecBCD activity inhibits λ development and SynExo-mediated recombination by destroying linear dsDNA. While Gam is not a SynExo recombinase, maximum Recombineering activity in E. coli requires either Gam expression or host mutation that inactivates RecBCD to preserve linear dsDNA substrates such as PCR products. The x-ray crystallographic structure of Gam (Court et al. 2007) is shown in Figure 1.1.

Figure 1.1 Enterobacteria phage lambda Gam protein structure
X-ray model using coordinates from PDB 2UV1 (Court et al. 2007). The helical “wings” insert into the ssDNA binding channels of RecBCD to inhibit exonuclease V activity, preserving dsDNA Recombineering substrates.


\( \lambda \) Exo structure and function

\( \lambda \) Exo forms a toroid composed of 3 identical subunits (Figure 1.2). \( \lambda \) Exo belongs to the SPD(x\(_a\))ExK superfamily of nucleases (Myers & Rudd 1998; Kovall & Matthews 1998; Aravind et al. 2000) that includes restriction enzymes, HJ resolvases and RecB family nucleases. However, \( \lambda \) Exo is most closely related to viral exonucleases like UL12 from human Herpes Virus.

The \( \lambda \) Exo toroid binds to dsDNA ends by recognizing the 5’ phosphate at sites within the central channel (Subramanian et al. 2003). Each monomeric subunit has an active site motif that is competent to catalyze nucleophilic attack on the scissile phosphate of the proximal phosphodiester bond between the terminal nucleotide and its nearest neighbor. Nucleophilic attack is catalyzed by a hydroxide anion abstracted from water bound to Mg\(^{2+}\) in the active site, forming a negatively charged hydrolysis intermediate that is stabilized by Mg\(^{2+}\) acting as a general Lewis acid and by positively charged amino acid side-chains in residues next to the Mg\(^{2+}\) binding site (Subramanian et al. 2003). The reaction releases deoxyribonucleotide monophosphates (dNMPs) one nucleotide at a time by a processive 5’ \( \rightarrow \) 3’ mechanism, creating 3’ ssDNA overhangs that can prime DNA synthesis if paired with a template strand. \( \lambda \) Exo forms a 1 to 1 complex with Beta and loads Beta on the single strand DNA as it resects dsDNA, creating a pre-synapatic filament that protects ssDNA from cellular exonucleases and facilitates homologous DNA pairing. Given the quaternary structures of \( \lambda \) Exo and Beta, the 1:1 stoichiometry likely represents a 1:1 association of subunits on a much larger scale (e.g. 4 \( \lambda \) Exo homotrimers bound to one 12 subunit Beta toroid), as originally proposed by Tolun (2008).
Figure 1. 2 Enterobacteria phage lambda exonuclease protein structure
X-ray model using coordinates from PDB 3SLP (Zhang et al. 2011a). The homotrimeric λexo toroid is shown in a complex with its substrate, linear dsDNA, which is bound to the central channel of the toroid.

Beta structure and function

Beta is a globular protein of the Synaptase or SSAP (single strand annealing protein) superfamily (Iyer et al. 2002). Members of this family form multimeric rings and filaments with DNA and facilitate ssDNA annealing. Beta forms complexes of 12 subunits and forms filaments when bound to ssDNA (Passy et al. 1999). Beta protects ssDNA from cellular exonucleases and promotes annealing of complementary ssDNA sequences. When the substrate is dsDNA, both λ Exo and Beta are required to promote homologous recombination. When the substrate is ssDNA, only Beta is required to promote homologous recombination. Beta binds 4-5 nucleotides per monomer and forms stable filaments with oligos equal or bigger than 36 nucleotides long (Muniyappa & Radding 1986; Mythili et al. 1996), suggesting that cooperative protein-protein...
interactions stabilize the filament (also evident in our lab, Basart and Myers, unpublished).

Recombineering mechanism

The λ Red recombination functions are expressed in the λ pL operon (Thomason et al. 2005) under control of the cl repressor protein. For Recombineering it is common to use a thermolabile cl derivative called cl857, which is composed of the ind1 allele (which makes cl noniducible by DNA damage) and the 857 allele (which makes cl thermolabile). In the preferred configuration, cl inactivation by heat shock leads to induced transcription of gam, bet and exo.

After induction of these proteins the DNA substrate is transformed into cells, usually by electroporation. If the substrate is dsDNA, Exo protein resects it creating 3’ ssDNA tipped recombination intermediates and loads Beta onto the nascent ssDNA (Stahl et al. 1997). If the substrate is ssDNA, then exo and gam expression is not required for Beta loading and recombination. Beta protects the substrate from exonuclease degradation and promotes the annealing of the ssDNA to a complementary ssDNA sequence. Usually, this annealing happens in the context of the replication fork in between Okazaki fragments, where ssDNA is exposed in the lagging strand template. The exposure of ssDNA in between Okazaki fragments facilitates annealing, therefore there is a strand bias in Recombineering of replicating DNA (Lim et al. 2008; Poteete 2008). Recombineering is RecA independent, but Recombineering is stimulated in the presence of RecA (Ellis et al. 2001). In this document, we show that Exo and RecA interact physically and functionally. It is possible that RecA could also facilitate strand invasion, which could
account for the low but evident frequencies observed for leading strand template targeting which would be difficult to do using Beta alone.

The annealed ssDNA intermediate can prime DNA synthesis as if it were an Okazaki fragment and later be ligated to the newly synthesized DNA strand. At this point if the substrate DNA has regions of heterology with the annealed target, then it can be recognized by MMR and physically removed. In fact MMR is a potent inhibitor of Recombineering. If the mismatch is not removed, then the sequence will be copied during the next DNA replication round producing a stable recombinant molecule. The recombinant genomes will segregate from the non-recombinant genomes as the cell divides. Therefore the end result of Recombineering is a mixed population of recombinants and cells bearing the initial parental genotype (Boyle et al. 2013; Sawitzke et al. 2011; Valledor et al. 2012).

ssDNA as a substrate for Recombineering

While dsDNA Recombineering requires the three Red genes (gam, bet and exo), ssDNA requires only Beta and this simpler reaction is much more efficient than when using dsDNA substrates (Ellis et al. 2001; Zhang et al. 2003). ssDNA substrates are designed to be complementary to the target region with the change in the middle of the oligo with flanking homology arms of 35 nucleotides at each side (Sawitzke et al. 2011). The best performance is obtained using oligos with a total length between 60 and 100 nt (Li et al. 2003). Genome editing is more effective when altered base sequences are encoded in the middle since the oligo suffers processing at both ends, probably by exonucleases, before it is ligated to the nascent DNA (Wang et al. 2011).
Oligo-dependent Recombineering in replicating cells has been shown to have a strong strand bias that correlates with annealing the oligo to the lagging strand template in the context of the replication fork (Li et al. 2003). Therefore the most efficient oligos are those designed to target the lagging strand template.

When designing the oligos, it is best to use the less well-recognized substitutions by MMR. C.C mismatches are not recognized by MMR, while other mismatches are recognized to different degrees (Sawitzke et al. 2011, Valledor et al. 2012). MMR recognition of heteroduplexes declines as the heteroduplex mismatch is enlarged, without apparent effect for bubbles of ~8 nt in E. coli (Wang et al. 2011), but the frequency of recombinants also decreases as the heteroduplex size increases, perhaps because Beta catalysis is less efficient when homology arms are more separated. Another strategy used to avoid MMR and that increases Recombineering efficiency up to 20 fold is the use of modified bases that are not recognizable by MMR like 2'-fluorouridine, 5-methyl-deoxycytidine, 2,6-diaminopurine or iso-deoxyguanosine instead of T, C, A or G (Wang et al. 2011).

Phosphorothioate-modified oligos have also been shown to increase Recombineering rates, probably by limiting oligo degradation by cellular exonucleases (Wang et al. 2009; Mosberg et al. 2010). Interestingly, phosphorothioated sites do not change the rates at which the ends of the oligos get incorporated. To explain this observation, (Wang et al. 2011) proposed that cells possess two different sets of nucleases that can act on the ssDNA substrate, where phosphorothioated bases protect oligos from the processive exonucleases but not from distributive exonucleases and endonucleases.

MMR

The MMR system is a potent inhibitor of Recombineering. The mismatch recognition component of MMR in E. coli is MutS. MutS recognizes distorted DNA produced by mismatched bases within heteroduplex intermediates and interacts with MutL, which
directs removal of the newly introduced oligo. Following removal of the newly incorporated oligo, DNA synthesis fills the gap. Deletion of the mutS gene increases Recombineering frequencies for most heteroduplexes (Costantino & Court 2003a), but also increases the general mutation rate, so it is preferred to avoid MMR using careful oligo design.

**SynExos**

Two sets of Recombineering enzymes are commonly used in *E. coli* to engineer genomes: the RecE/T SynExo from the Rac prophage and the λ Red Exo/Beta SynExo system from bacteriophage λ. While RecE/T and λ Exo/Beta are isofunctional complexes composed of exonuclease and synaptase subunits and while they both share the same host (*E. coli*), dsDNA Recombineering requires homospecific pairs of exonuclease and synaptase (*i.e.* RecE + Beta and λ Exo + RecT do not recombine dsDNA and do not form SynExo complexes (Muyrers *et al.* 2000)), indicating that these pairs have distinct allele-specific interactions necessary for SynExo function. This concept will be extended to allele-specific interactions between SynExos and host proteins, below.

Similar SynExo recombinases (Vellani & Myers 2003) have been identified in the genomes of many other DNA viruses (Aravind *et al.* 2000; Myers & Rudd 1998). Putative SynExo nucleases were identified by searching for sequences containing the catalytic motif (SPD(xₙ)ExK) and 5’ end phosphate recognition motifs (“WRS” and “YQQ”). Synaptase subunits are more difficult to identify by primary sequence comparisons via BLASTP and PSIBLAST as there are no distinctive amino acid sequence motifs despite substantial conservation in quaternary structure and function. Instead, Synaptases are primarily identified first by functional assays and data mining and subsequently by primary sequence comparisons using each Synaptase to extend the
protein families. Therefore we and others (Iyer et al. 2002) identified Synapatase Families that comprise an overall Synapatase Superfamily.

While the SynExo recombinases were discovered in lambdoid phage, they are common to many diverse dsDNA viruses (Figure 1.3 and Table 1.1). SynExo orthologs have been found in viruses that infect most types of bacteria, in viruses of eukaryotes including humans, mice, flies, frogs, fish, plants (including rice, corn, soy, sorghum and algae) and in protists. However no SynExo recombinases have yet been detected in Archaea.

Figure 1.3 SynExo genes in most dsDNA viruses that infect Bacteria and Eukaryota.
Orthologous protein sequences were detected in sequenced genomes using reiterative use of PSIBLAST with a cutoff of 0.005 until convergence was reached. Numbers reflect the number of organisms in which viral SynExo subunits were found and in viruses for these organisms. The numbers are color-coded to show the distribution of different SynExo protein families.
Table 1.1 Examples of SynExo Functional Modules

<table>
<thead>
<tr>
<th>Host</th>
<th>Virus</th>
<th>Exonuclease</th>
<th>Synaptase</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bacterium</td>
<td>λ</td>
<td>λ Exo</td>
<td>Redβ</td>
</tr>
<tr>
<td>Bacterium</td>
<td>Rac</td>
<td>RecE</td>
<td>RecT</td>
</tr>
<tr>
<td>Bacterium</td>
<td>SPP1</td>
<td>Chu</td>
<td>35</td>
</tr>
<tr>
<td>Bacterium</td>
<td>P22</td>
<td>ABC2-RecBCD</td>
<td>ERF</td>
</tr>
<tr>
<td>Human</td>
<td>HSV-1</td>
<td>UL12</td>
<td>UL29</td>
</tr>
<tr>
<td>Human</td>
<td>Cellular</td>
<td></td>
<td>Rad52</td>
</tr>
<tr>
<td>Insect</td>
<td>AcNPV</td>
<td>Alk exo</td>
<td>LEF-3</td>
</tr>
<tr>
<td>Plant</td>
<td>PBCV-1</td>
<td>A166R</td>
<td>?</td>
</tr>
</tbody>
</table>

While cellular homologous recombination is performed by many proteins in complex pathways, SynExo recombinases are the most compact and efficient recombination machinery yet described. Some of these SynExos have been studied in detail (Vellani & Myers 2003; Mikhailov et al. 2003; Reuven et al. 2003) and some have been used to edit genomes such as in *Lactobacillus* (van Pijkeren & Britton 2012), *Mycobacterium tuberculosis* (van Kessel & Hatfull 2007) and recently in human cells (Schumacher et al. 2012).

Host specificity

Recombineering was evaluated in mouse ES cells by expressing the bacteriophage synaptase Beta and providing ssDNA oligos that repaired a defective neomycin gene (Zhang et al. 2003). The authors showed that cells expressing Beta produced recombinants that survived neomycin selection (6 of 12 clones yielded recombinants), while no recombination was observed for the endogenous functions alone. This experiment demonstrated the feasibility of extrapolating the bacteriophage recombinases to mammalian cells, although the efficiency of this experiment was lower than the other emerging genome technologies for human cells and much lower than Recombineering rates in *E. coli*.
ssDNA Recombineering using different synaptases in *Mycobacteria* shows evidence of host specificity where its natural synaptase outperformed *E. coli* synaptases RecT and Beta (van Kessel *et al.* 2008). Datta *et al.* (2008) also showed SynExo host specificity when comparing Recombineering catalyzed by Gram positive and Gram negative bacteriophage SynExos expressed in *E. coli*. They saw stringent host specificity for dsDNA Recombineering and less restriction for ssDNA Recombineering.

This evidence suggests that viral Recombineering functions are host-specific. It is possible that as cellular proteins evolved, viral proteins must co-evolve to maintain host/viral protein interactions including those required for efficient viral recombination. Our hypothesis is that Recombineering is host-specific and that viral recombinases co-evolved with host proteins. For human cell Recombineering, a good candidate recombinase is one from a virus infecting human cells. We proposed that the Human Herpes Virus 1 SynExo (UL12 exonuclease + ICP8 DNA binding protein) is a SynExo recombinase and it was shown in the lab of our collaborator Sandy Weller to catalyze strand exchange in an *in vitro* model of recombination (Reuven *et al.* 2003) and to stimulate dsDNA break repair in human cells (Schumacher *et al.* 2012).

**HHV1 SynExo: UL12 and ICP8**

HHV1 proteins UL12 and ICP8 form a complex that colocalizes to the same intranuclear compartments (Thomas *et al.* 1992) and that coordinates their activities in recombination (Reuven *et al.* 2003). UL12 is an exonuclease (a λ exo ortholog) and ICP8 is a synaptase (a functional analog of Beta) and their complex is structurally and mechanistically similar to the λ SynExo.
UL12 structure and function

The HHV1 Ul12 gene expresses two transcripts from two different start sites. Initiation of translation from one start site produces a protein, UL12, that includes a nuclear localization signal, while expression from the second start site produces an isoform, UL12.5 (Reuven et al. 2004), lacking the nuclear localization signal that un masks a mitochondrial localization signal. Upon HHV1 infection, UL12.5 travels to mitochondria and destroys the mitochondrial genome (Saffran et al. 2007), providing in vivo evidence of the hydrolytic processivity characteristic of these exonucleases. UL12 interacts with host proteins (as does λ Exo, this thesis) including the DNA repair complex MRN and with host MMR proteins (Mohni et al. 2011; Balasubramanian et al. 2010) which could indicate cooperation with host proteins to promote homologous recombination. A structural homology model of part of UL12 is shown in Figure 1.4.

Figure 1.4 Model of HSV1 UL12 structure
A crystal structure for Herpes Simplex Virus Type 1 (HSV1) UL12 is not yet available, but the primary sequence conservation allows some of the UL12 sequence to be used for structural homology modeling using the λ exo structure (PDB code 1avq) as the reference protein (Hsiang & Ho 2008).
ICP8 structure and function

ICP8 is a 128 KDa zinc metalloprotein encoded by the UL29 gene (Taylor et al. 2003). ICP8 is the major DNA binding protein for HHV1 and is essential for viral replication (Conley et al. 1981; Boehmer & Lehman 1993). Consequently, ICP8 is commonly used to identify HHV1 replication compartments in the nucleus (Taylor et al. 2003). ICP8 has been successfully expressed in Vero cells from the virus, from transiently transfected plasmids and from stable cell lines (Taylor et al. 2003).

ICP8 forms filaments with DNA as do other ssDNA annealing proteins such as Beta and Rad52 (Passy et al. 1999; Shinohara et al. 1998). The ultrastructure of ICP8 suggests that ICP8 is organized into two rings of nine subunits stacked on top of each other (Figures 1. 5 and 1. 6). Tolun calculated that each ring could bind 140 to 160 nucleotides ssDNA fragments, but was able to obtain ICP8-ssDNA complexes with 38 nucleotides.
oligos that formed a nucleoprotein complex of 2.3 MDa (Tolun et al. 2013). When ICP8 is bound to ssDNA, the rings stack and the C-terminal domain of one ICP8 monomer interacts with the N-terminal domain of the monomer in the other ring suggesting both ends are required for synapsis of one DNA-ICP8 filament (ring) with the other.

Figure 1. 6 ICP8 ring ultrastructure
Electron microscopic images of ICP8 toroids bound to ssDNA. (a) EM field showing representative ICP8 toroids. (b) Averages of images randomly selected from EM fields were used to enhance structural resolution. Seen are twin rings of ICP8 with 9 subunits in each ring (Tolun et al. 2013).
ICP8 is able to catalyze homologous recombination and strand exchange \textit{in vitro} (Reuven \textit{et al.} 2003) and was recently shown to work with UL12 to promote dsDNA break repair in HEK293 \textit{rad52} deficient cells, although the efficiency was not greater than when Rad52 was present (Schumacher \textit{et al.} 2012). Given that the HHV1 SynExo shares many properties with the \textit{\lambda} SynExo and given that Recombineering in bacteria has been so efficient in genome modification, we were drawn to determine if Recombineering could be reconstituted in human cells.

\textbf{Hypothesis}

The hypothesis of this thesis is that Recombineering is host-specific and that viral recombinases co-evolved with host proteins. To test this hypothesis I studied the interaction of phage \textit{\lambda} recombinases with bacterial host proteins and reconstituted viral recombinases in human cells and bacterial cells to evaluate the efficiency of recombination. The prediction is that the HHV1 recombinase, having co-evolved with human cells, will be specific to human cells and will not function well in \textit{E. coli}. In contrast, the \textit{\lambda} system will be specific to \textit{E. coli} and not function well in human cells.

\textbf{Specific Aims}

Aim 1: Identify the network of protein-protein and functional interactions between viral recombination proteins and host proteins.

Aim 2: Evaluate the efficiency and host-specificity of Recombineering in human cells and in bacteria.
CHAPTER 2. MATERIALS AND METHODS

Reagents

All reagents were from Sigma-Aldrich unless indicated.

All oligos were from from Sigma Genosys.

All enzymes were from NEB unless indicated.

QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent Technologies, 210518-5)

Gateway LR Clonase II Enzyme Mix (Invitrogen, 11791-020)

KOD master mix (EMD Millipore, 71842)

FailSafe PCR Enzyme (Epicentre, FS99060)

RNase-Free DNase (Promega, M6101) 1 U/µl = 10 mg/ml

Flexi system Entry/Transfer (Promega, C8640)

DNA ladder: 1 kb and 100 bp (NEB, N3232 and N3231)

Lipofectamine 2000 (Invitrogen, 11668)

Fugene 6 (Roche, 1814443001)

Nucleofector II kit: (Lonza VPE-1001 for MIAMI cells and VCA-1001 for HeLa)

Neon® Transfection System kit (Invitrogen, MPK1096)

0.45 µm low protein binding membrane filter (Pall, 4148)

BugBuster Protein Extract Reagent (Novagen, 70584)

Bugbuster HT (Novagen, 70922)

Benzonase Nuclease (70746, Novagen)

Protease inhibitor cocktail from Sigma (P8340)

cOmplete protease inhibitor tablet (Roche, 05892970001)
HaloTag® Protein Purification System (Promega, G6280)

Ni-NTA Magnetic Agarose Beads (Qiagen, 36111)

Precision Plus Protein™ Kaleidoscope™ Standards (Biorad, 161-0375)

The MagicMark™ XP Western Protein Standard (Life Technologies, LC5602)

Laemmli sample buffer (BioRad, 161-0737)

Glycerol, ultra pure (MP Biomedicals 800687)

Gel code blue stain reagent (Thermo Scientific, 24592)

DTT (Sigma, 43817)

Prolong Antifade (Molecular Probes, P7481 or Invitrogen, P36930)

Vectashield (Vector Laboratories, Inc. H-1000)

Poly-L-lysine hydrobromide, mol wt 30,000-70,000 (Sigma, P9155)

Primary antibodies:

Mouse monoclonal [10A3] to HSV1 ICP8 major DNA binding protein (Abcam, ab20193)

Mouse monoclonal [11E2] to HSV1 ICP8 Mayor DNA binding protein (Abcam, ab20194)

Rabbit polyclonal anti-HA (Abcam, ab9110)

Rabbit polyclonal to full-length GFP (Clontech, 632460)

Mouse anti-human Lamin A+C Monoclonal Antibody (Millipore, MAB3211)

Mouse monoclonal IgM to α Tubulin (Santa Cruz, sc-8035)

Polyclonal to human cleaved PARP (Cell Signaling, 9541)

Mouse monoclonal to GFP (cell signaling, 2955)
Monoclonal to Flag M2 (Sigma, F 1804)

Secondary antibody for ICC:
Alexa Fluor 594 Goat anti mouse IgG (H+L) (Life Technologies, A-11032)

Secondary antibodies for HRP:
Goat anti mouse IgG1-HRP (Santa Cruz, sc-2060)
Goat anti mouse IgM-HRP (Santa Cruz, sc-2064)
Goat anti rabbit IgG-HRP (Santa Cruz, sc-2004)

Media:
DMEM-high glucose, pyruvate (Gibco, 11995)
OptiMEM (Invitrogen, 31985)
RPMI 1640 (Gibco, 11875)
DMEM\textsuperscript{GFP} (Evrogen, MC101)
FBS (Hyclone, sh30396.03)
Ascorbic acid (Sigma, 49752)
Arachidonic acid, (Sigma, A9673)
Cholesterol (Sigma, C3045)
DL-alpha tocopherol-acetate (Sigma, T3376)
Myristic acid (Sigma, M3128)
Oleic acid (Sigma, 01383)
Palmitic acid (Sigma, P5585)
Palmitoleic acid (Sigma, P9417)
Stearic acid (Sigma, S4751)
Pen Strep Glutamine (Gibco, 15140)
TB (Difco, 243820)
SOB (BD, 24310)
Fibronectin (Sigma, F2518)
ECL Plus Western Blotting Detection System (from GE Healthcare, RPN2132)
RheoSwitch® Mammalian Inducible Expression System (NEB #E3000), including the
Gaussia Luciferase Assay Kit (NEB #E3300), the RheoSwitch® Ligand RSL1 (NEB
E3301), the pNEBR-R1 Regulator Plasmid, the pNEBR-X1GLuc Control Plasmid and
the pNEBR-X1Hygro Vector.

Recipes

**MIAMI cells expansion medium:** Marrow-isolated adult multi-lineage inducible adult
human stem were expanded in 3 % FBS, 20 mM ascorbic acid, 1:2000 essential fatty acid
mixture (12.9 nM arachidoncic acid, 1.12 µM cholesterol, 290 nM DL-alpha tocopherol-
acetate, 85.9 nM myristic acid, 69.4 nM oleic acid, 76.5 nM palmitic acid, 77.1 nM
palmitoleic acid and 68.9nM stearic acid), 100 U/mL penicillin and 1 µg/mL
streptomycin in DMEM-low glucose, pyruvate medium. Filter sterilized.

**MIAMI cells freezing medium:** 80 % FBS, 10 % MIAMI expansion medium and 10 %
DMSO. Filter sterilized.

**293T, 293, HeLa cells expansion medium:** 5 to 10 % FBS, 1 % Pen Strep Glutamine,
DMEM-high glucose, pyruvate medium. Filter sterilized.
293T, 293, HeLa cells freezing medium: 50 % FBS, 40% DMEM-high glucose, pyruvate medium and 10 % DMSO. Filter sterilized.

**LB:** Mix 10 g Bacto-tryptone, 5 g yeast extract, 5 g NaCl in 1 L H₂O, adjust pH to 7.5 with NaOH. Sterilize by autoclaving.

**LB agar:** Add 12 g Bactoagar to LB broth before autoclaving.

**BBL Trypticase agar:** Mix 10 g BBL trypticase, 5 g NaCl in 1L H₂O, adjust pH to 7.5 with NaOH. Add 10 g Bactoagar for BBL plates or 7 g Bactoagar for top agar (for phage plating and spot tests).

**SOB:** Dissolve 28 g in 1L H₂O. Sterilize by autoclaving.

**SOC:** Make 100 ml SOC with 99 ml SOB media. After autoclaving, add 1 ml of sterile 2 M glucose or 2 ml of 20 % glucose stock.

**OMBG:** 1.05 % K₂HPO₄, 0.45 % KH₂PO₄, 0.005 % MgSO₄·7H₂O, 0.1 % (NH₄)₂SO₄, 0.05 % sodium citrate and 0.2% glucose.

**TE buffer:** 10mM Tris 1 mM EDTA pH 7.55.

**TMC Buffer:** 10mM Tris pH 7, 5 mM MgCl₂, 5 mM CaCl₂.

**IPTG:** Make 1 M solution by dissolving 1 g in 4196 µl of water. Filter sterilize. IPTG Molecular weight: 238.31 g/mol.

**Colony transfection solution:** 5 % PEG8000, 5 % DMSO, 50 mM CaCl₂, pH 6.5.

**Protamine sulfate stock:** Prepare in water to 10 mg/ml and filter sterilize.

**DNA lysis solution:** 3 % SDS, 0.1M Tris HCl pH 8, 120 mM NaOH.

**PCI:** 25 parts Phenol: 24 parts Chloroform: 1 part Isoamyl alcohol. Store at -20 °C in the dark to avoid phenol oxidation (visible as a pink tint).
Super Magic Medium: Prepare 1 L with the following components added and autoclaved: 35 g Tryptone, 20 g Yeast Extract, 3.3 g (NH₄)₂SO₄, 6.8 g KH₂PO₄, 7.1 g Na₂HPO₄ (or 8.7 g K₂HPO₄), 0.5 g glucose, 2.0 g α-lactose, 0.15 g Mg₂SO₄.

100 mM NH₄HCO₃: 0.395 g in 50 ml of ddH₂O, vortex. Keep at 4 °C.

10 mM DTT/100 mM NH₄HCO₃: 0.154 g of DTT in 10 ml of 100 mM NH₄HCO₃. Keep at 4 °C.

55 mM iodoacetamide/100 mM NH₄HCO₃: 0.102 g of iodoacetamide in 10 ml of 100 mM NH₄HCO₃. Prepare fresh just before use. Make sure the iodoacetamide/NH₄HCO₃ solution tube is wrapped with aluminum foil to protect from the light.

30 mM NEM acetate: in HPLC-H₂O (if stock is 7.86 mol/L, add 191 µl to 50 ml HPLC-H₂O).

Working trypsin: Prepare just before trypsin digestion of proteins in SDS PAGE gel slabs, mix 19 µl of 15 mM NEM with 1 µl of trypsin per sample. Scale up mix for number of samples. Do not vortex or shake, only tap or flick with your finger.

Loading dye: Mix Laemmli sample buffer with the same amount of glycerol ultra pure.

SDS-PAGE running buffer (1X Tris-glycine): 25 mM Tris base, 190 mM glycine, 0.1 % SDS in ultra pure water. Check the pH; it should be around pH 8.3 so adjust if necessary.

TBS 10x: 24.23 g Trizma HCl, 80.06 g NaCl. Dissolve in 800 ml ultra pure water. Adjust pH to 7.6 with pure HCl. Top up to 1 L with ultra pure water.

TBST: 100 ml of TBS 10x, 900 ml ultra pure water, 1ml Tween20.

Ni-NTA Magnetic Agarose Beads Binding Buffer: 50 mM Tris Acetate pH 7.5, 50 mM NaCl, 30 mM Imidazole and 0.5 % Triton X-100.
Ni-NTA Magnetic Agarose Beads Elution buffer: 150 mM Tris Acetate pH 7.5, 100 mM NaCl, 300 mM Imidazole and 0.5 % tween 20.

Transfer buffer: 25 mM Tris, 192 mM glycine, 10% methanol, pH 8.3.

Propidium iodide solution: 0.5 mg/ml Propidium Iodide in PBS, pH7.4.

Complementary ssDNA annealing buffer: 10mM Tris, pH 7.5-8.0, 50mM NaCl, 1mM EDTA.

Human cell line culture conditions

MIAMI 3515 and its derivatives: Grown in expansion medium consisting of DMEM-low glucose containing 3 % fetal bovine serum, 20 mM ascorbic acid, 1:2000 essential fatty acid mixture (12.9 nM arachidonic acid, 1.12 µM cholesterol, 290 nM DL-alpha tocopherol-acetate, 85.9 nM myristic acid, 69.4 nM oleic acid, 76.5 nM palmitic acid, 77.1 nM palmitoleic acid and 68.9 nM stearic acid) and antibiotics (100 U/mL penicillin, 0.1 mg/mL streptomycin) in low oxygen conditions (3% O₂, 5% CO₂ and 92 % N₂). Media were changed every 3-4 days and the cells were detached using trypsin upon reaching ~60 % confluence. Cells were washed out from the trypsin with expansion medium by centrifugation at 400 RCF for 5 minutes at 4 °C. Pelleted cells were resuspended in pre-warmed expansion medium and plated in 10 ng/ml fibronectin-treated flasks at 100 cells/cm². Cells were frozen in filter sterilized freezing medium (80 % FBS, 10 % MIAMI expansion media and 10% DMSO). Cells were resuspended in freezing medium to 1.5x10⁵ to 3x10⁵ cells per ml, aliquotted in freezing vials, frozen in a cryogenic box containing isopropanol at -80 °C overnight and stored for long term in liquid nitrogen at -140 °C. MIAMI cells were thawed by placing the cryovials in a 37 °C water bath until cells were almost completely thawed, then the cells were washed from the
freezing medium by resuspending them in pre-warmed 10 ml expansion medium and centrifugation at 400 RCF for 5 minutes at 4 °C. MIAMI cells were then resuspended in pre-warmed expansion medium, transferred to a fibronectin coated flask and incubated at 37 °C, 3 % O₂, 5 % CO₂.

**HEK 293 and its derivatives:** Human embryonic kidney cells were grown in DMEM high glucose, 10 % FBS and antibiotics-antimycotics (50 units/ml Penicillin G, 0.25 µg/ml Amphotericin B and 10 µg/ml Streptomycin). Cultures were grown with 5 % CO₂ in atmosphere at 37 °C. The pSLIK HEK 293 cell lines were grown as described for HEK 293 except using 10 % Tet free FBS (Clontech, 631106), in the presence of 100 µg/ml zeocin. Cells were frozen in filter sterilized freezing medium (50 % FBS, 40 % DMEM high glucose media and 10 % DMSO). Cells were resuspended in freezing medium to 10⁵ to 10⁷ cells per ml, aliquotted in freezing vials, freezed in a cryogenic box containing isopropanol at -80 °C overnight and storage for long term in liquid nitrogen at -140 °C. Cells were thawed by placing the cryovials in a 37 °C water bath until cells were almost completed defrosted, then the cells were washed from the freezing medium by resuspending them in pre-warmed 10 ml expansion media and centrifugation at 400 RCF for 5 minutes at 4 °C. Cells were then resuspended in pre-warmed expansion medium, transferred to a flask and incubated at 37 °C with 5 % CO₂.

**HEK 293T and its derivatives:** Grown in DMEM high glucose, 5 % FBS and antibiotics-antimycotics (50 units/ml Penicillin G, 0.25 µg/ml Amphotericin B and 10 µg/ml Streptomycin). Cultures were grown with 5 % CO₂ in atmosphere at 37°C. The pSLIK HEK 293T cell lines were grown as described for HEK 293 except using 10% Tet free FBS (Clontech, 631106), in the presence of 100 µg/ml zeocin. Cells were frozen in
filter sterilized freezing medium (50 % FBS, 40 % DMEM high glucose media and 10 % DMSO). Cells were resuspended in freezing medium to $10^5$ to $10^7$ cells per ml, aliquotted in freezing vials, frozen in a cryogenic box containing isopropanol at -80 °C overnight and stored for long term in liquid nitrogen at -140 °C. Cells were thawed by placing the cryovials in a 37 °C water bath until cells were almost completed defrosted, then the cells were washed from the freezing medium by resuspending them in pre-warmed 10 ml expansion medium and centrifugation at 400 RCF for 5 minutes at 4 °C. Cells were then resuspended in pre-warmed expansion media, transferred to a flask and incubated at 37 °C with 5 % CO₂.

**RheoSwith HeLa 11** is part of the RheoSwitch® Mammalian Expression System (NEB C2007). Grown in DMEM, 5 % FBS, antibiotic-antimycotic (50 units/ml Penicillin G, 0.25 µg/ml Amphotericin B and 10 µg/ml Streptomycin) and 500 µg/ml G418. Cultures were grown with 5 % CO₂ in atmosphere at 37°C. Cells were frozen in filter sterilized freezing medium (50 % FBS, 40 % DMEM high glucose media and 10 % DMSO). Cells were resuspended in freezing medium to $10^5$ to $10^7$ cells per ml, aliquotted in freezing vials, frozen in a cryogenic box containing isopropanol at -80 °C overnight and stored for long term in liquid nitrogen at -140 °C. Cells were thawed by placing the cryovials in a 37 °C water bath until cells were almost completed defrosted, then the cells were washed from the freezing medium by resuspending them in pre-warmed 10 ml expansion medium and centrifugation at 400 RCF for 5 minutes at 4 °C. Cells were then resuspended in pre-warmed expansion media, transferred to a flask and incubated at 37 °C with 5 % CO₂.
**pPRO1.1go.** Thawed in DMEM, 10 % FBS, and expanded in 5 % FBS, antibiotic-antimycotic (50 units/ml Penicillin G, 0.25 µg/ml Amphotericin B and 10 µg/ml Streptomycin) and 400 µg/ml zeocin (Invitrogen). Cultures were grown with 5 % CO₂ atmosphere at 37 °C. Cells were frozen in filter sterilized freezing medium (50 % FBS, 40 % DMEM high glucose media and 10 % DMSO). Cells were resuspended in freezing medium to 10⁵ to 10⁷ cells per ml, aliquotted in freezing vials, frozen in a cryogenic box containing isopropanol at -80 °C overnight and stored for long term in liquid nitrogen at -140 °C. Cells were thawed by placing the cryovials in a 37 °C water bath until cells were almost completely defrosted, then the cells were washed from the freezing medium by resuspending them in pre-warmed 10 ml expansion medium and centrifugation at 400 RCF for 5 minutes at 4 °C. Cells were then resuspended in pre-warmed expansion media, transferred to a flask and incubated at 37 °C with 5 % CO₂.

### Human cell line list

**Table 2.1 Human cell lines**

<table>
<thead>
<tr>
<th>Human cell lines</th>
<th>AKA</th>
<th>Source</th>
<th>Phenotype</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIAMI 3515</td>
<td>MIAMI</td>
<td>Dr. Schiller’s lab</td>
<td></td>
<td>(D’Ippolito <em>et al.</em> 2004)</td>
</tr>
<tr>
<td>MIAMI-Mostaza</td>
<td>MIAMI-Yellow</td>
<td>This study</td>
<td>~50 % yellow fluorescent.</td>
<td>MIAMI transduced with pDUAL-eGFP (Y203)</td>
</tr>
<tr>
<td>MIAMI-eGFP</td>
<td>MIAMI-Green</td>
<td>This study</td>
<td>~50 % green fluorescent.</td>
<td>MIAMI transduced with pDUAL-eGFP</td>
</tr>
<tr>
<td>MIAMI-Mostaza+eGFP</td>
<td>MIAMI-Y+G</td>
<td>This study</td>
<td>~50 % yellow and green fluorescent.</td>
<td>MIAMI transduced with pDUAL-eGFP (Y203) and pDUAL-eGFP</td>
</tr>
<tr>
<td>HEK 293</td>
<td>293</td>
<td>Gift of Dr. Rong Wen at UM</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HEK 293-Mostaza</td>
<td>293-Yellow</td>
<td>This study</td>
<td>~50 % yellow fluorescent.</td>
<td>293 transduced with pDUAL-eGFP(Y203)</td>
</tr>
<tr>
<td>HEK 293-eGFP</td>
<td>293-Green</td>
<td>This study</td>
<td>~50 % green fluorescent.</td>
<td>293 transduced with pDUAL-eGFP</td>
</tr>
<tr>
<td>HEK 293-Mostaza+eGFP</td>
<td>293-Y+G</td>
<td>This study</td>
<td>~50 % yellow and green fluorescent.</td>
<td>293 transduced with pDUAL-eGFP(Y203) and pDUAL-eGFP</td>
</tr>
<tr>
<td>HEK 293-Mostaza-pSLIK:Crimson-P2A/ICP8</td>
<td>293-Mostaza-pSLIK1</td>
<td>This study</td>
<td>~90 % yellow fluorescent. Red fluorescent upon Dox induction. Zeocin resistant.</td>
<td>HEK 293-Mostaza transduced with pSLIK::Crimson-P2A/ICP8</td>
</tr>
<tr>
<td>HEK 293-Mostaza-pSLIK:Crimson-P2A/NSL-HA-Hum-Beta</td>
<td>293-Mostaza-pSLIK2</td>
<td>This study</td>
<td>~90 % yellow fluorescent. Red fluorescent upon Dox induction. Zeocin resistant.</td>
<td>HEK 293-Mostaza transduced with pSLIK::Crimson-P2A/NSL-HA-Hum-Beta</td>
</tr>
<tr>
<td>HEK 293-Mostaza-pSLIK:Crimson-P2A</td>
<td>293-Mostaza-pSLIK3</td>
<td>This study</td>
<td>~90 % yellow fluorescent. Red fluorescent upon Dox induction.</td>
<td>HEK 293-Mostaza transduced with pSLIK::Crimson-P2A</td>
</tr>
</tbody>
</table>
### Human cell lines

<table>
<thead>
<tr>
<th>Human cell lines</th>
<th>AKA</th>
<th>Source</th>
<th>Phenotype</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>HEK 293-Mostaza-pSLIK: ICP8-P2A/Crimson</td>
<td>HEK 293T</td>
<td>This study</td>
<td>~90% yellow fluorescent. Red fluorescent upon Dox induction. Neomycin resistant.</td>
<td>HEK 293T-Mostaza transduced with pSLIK::Crimson-ICP8-P2A/Crimson</td>
</tr>
<tr>
<td>HEK 293-Mostaza-pSLIK: NLS-HA-Hum-Beta - P2A/Crimson</td>
<td>HEK 293T</td>
<td>This study</td>
<td>~90% yellow fluorescent. Red fluorescent upon Dox induction. Neomycin resistant.</td>
<td>HEK 293T-Mostaza transduced with pSLIK::NLS-HA-Hum-Beta -P2A/Crimson</td>
</tr>
<tr>
<td>HEK 293-Mostaza-pSLIK: P2A/Crimson</td>
<td>HEK 293T</td>
<td>This study</td>
<td>~90% yellow fluorescent. Red fluorescent upon Dox induction. Neomycin resistant.</td>
<td>HEK 293T-Mostaza transduced with pSLIK::P2A/Crimson</td>
</tr>
<tr>
<td>HEK 293T</td>
<td>HEK 293T</td>
<td>Gift of Dr. Priya Rai at UM</td>
<td>Neomycin resistant.</td>
<td></td>
</tr>
<tr>
<td>HEK 293T-eGFP</td>
<td>HEK 293T</td>
<td>This study</td>
<td>~90% yellow fluorescent. Neomycin resistant.</td>
<td>293T transduced with pDUAL-eGFP(Y203)</td>
</tr>
<tr>
<td>HEK 293T-Y+G</td>
<td>HEK 293T</td>
<td>This study</td>
<td>~90% yellow and green fluorescent. Neomycin resistant.</td>
<td>293T transduced with pDUAL-eGFP(Y203) and pDUAL-eGFP</td>
</tr>
<tr>
<td>HEK 293T-Celete</td>
<td>HEK 293T</td>
<td>This study</td>
<td>~50% cyan fluorescent. Neomycin resistant.</td>
<td>293T transduced with pDUAL-eGFP(W66)</td>
</tr>
<tr>
<td>HEK 293T-Azure</td>
<td>HEK 293T</td>
<td>This study</td>
<td>~50% blue fluorescent. Neomycin resistant.</td>
<td>293T transduced with pDUAL-eGFP(H66)</td>
</tr>
<tr>
<td>HEK 293T-Dark</td>
<td>HEK 293T</td>
<td>This study</td>
<td>Normal. Neomycin resistant.</td>
<td>293T transduced with pDUAL-eGFP(Stop66)</td>
</tr>
<tr>
<td>HEK 293T-Mostaza-pSLIK:Crimson-P2A/ICP8</td>
<td>HEK 293T</td>
<td>This study</td>
<td>~90% yellow fluorescent. Red fluorescent upon Dox induction. Neomycin and zeocin resistant.</td>
<td>HEK 293T-Mostaza transduced with pSLIK::Crimson-P2A/ICP8</td>
</tr>
<tr>
<td>HEK 293T-Mostaza-pSLIK:Crimson-P2A/NLS-HA-Hum-Beta</td>
<td>HEK 293T</td>
<td>This study</td>
<td>~90% yellow fluorescent. Red fluorescent upon Dox induction. Neomycin and zeocin resistant.</td>
<td>HEK 293T-Mostaza transduced with pSLIK::Crimson-P2A/NLS-HA-Hum-Beta</td>
</tr>
<tr>
<td>HEK 293T-Mostaza-pSLIK:Crimson-P2A</td>
<td>HEK 293T</td>
<td>This study</td>
<td>~90% yellow fluorescent. Red fluorescent upon Dox induction. Neomycin and zeocin resistant.</td>
<td>HEK 293T-Mostaza transduced with pSLIK::Crimson-P2A</td>
</tr>
<tr>
<td>HEK 293T-Mostaza-pSLIK: ICP8-P2A/Crimson</td>
<td>HEK 293T</td>
<td>This study</td>
<td>~90% yellow fluorescent. Red fluorescent upon Dox induction. Neomycin and zeocin resistant.</td>
<td>HEK 293T-Mostaza transduced with pSLIK::Crimson-ICP8-P2A/Crimson</td>
</tr>
<tr>
<td>HEK 293T-Mostaza-pSLIK: NLS-HA-Hum-Beta - P2A/Crimson</td>
<td>HEK 293T</td>
<td>This study</td>
<td>~90% yellow fluorescent. Red fluorescent upon Dox induction. Neomycin and zeocin resistant.</td>
<td>HEK 293T-Mostaza transduced with pSLIK::NLS-HA-Hum-Beta -P2A/Crimson</td>
</tr>
<tr>
<td>HEK 293T-Mostaza-pSLIK: P2A/Crimson</td>
<td>HEK 293T</td>
<td>This study</td>
<td>~90% yellow fluorescent. Red fluorescent upon Dox induction. Neomycin and zeocin resistant.</td>
<td>HEK 293T-Mostaza transduced with pSLIK::P2A/Crimson</td>
</tr>
<tr>
<td>RheoSwith HeLa 11</td>
<td>HeLa</td>
<td>Gift of Dr. Thomas Misteli at NIH/NCI</td>
<td>Neomycin resistant.</td>
<td>HeLa stably transfected with pNEBR-X1</td>
</tr>
<tr>
<td>pPRO1.1go</td>
<td>HeLa-</td>
<td>Gift of Dr. Thomas Misteli at NIH/NCI</td>
<td>Green and Red fluorescent. Zecarin resistant.</td>
<td></td>
</tr>
</tbody>
</table>

### E. coli culture conditions

*E. coli* were routinely cultured in LB broth with shaking or on LB agar plates at 37 °C unless otherwise indicated with appropriate drug selection. Cultures were started usually
from single colonies or by diluting overnight (saturated) cultures 1/100 into fresh medium. *E. coli* were cultured in TB at 37 °C with shaking in baffle bottom flasks at least 5x the culture volume for plasmid isolation. Cells were frozen by mixing 1.2 ml of an overnight culture and 0.2 ml sterile DMSO and stored at -80 °C. Cells were revived from frozen stocks by streaking some of the frozen culture on a LB plate with the appropriate drug. Antibiotics used for selection were at the following concentrations: Amp (100 ng/ml), Carb (50 ng/ml), Kan (25-50 ng/ml), Cml (20 ng/ml).

### E. coli strain lists

#### Table 2.2 *E. coli* strains for bacterial Recombineering

<table>
<thead>
<tr>
<th>Strain #</th>
<th>AKA</th>
<th>Source</th>
<th>Relevant Genotype</th>
<th>Construction method</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>RIK385</td>
<td>SW101</td>
<td>Gift of Don Court</td>
<td>As DH10B [λ cl 377 Δ(cro-bioA) &lt;(\rightarrow) [tetRA] recA1]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RIK401</td>
<td>SCC1</td>
<td>Gift of CC Sze</td>
<td>as MG1655 but PA1/04/03-gfpmut3* integrated at nts 312754-31277</td>
<td>(Andersen et al. 1998)</td>
<td>Green fluorescence</td>
</tr>
<tr>
<td>RIK416</td>
<td>KRE12765</td>
<td>Gift of Kenn Rudd at UM</td>
<td>as MG1655 seq* rph* but ΔmutS::Kan</td>
<td>Kan&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>RIK430</td>
<td>Mel120</td>
<td>Gift of Kenn Rudd at UM</td>
<td>as MG1655 seq* rph* but ΔmutS::Kan</td>
<td>KEIO collection (Baba et al. 2006)</td>
<td>Kan&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>RIK428</td>
<td>BW25113, Mel119</td>
<td>Gift of Kenn Rudd at UM</td>
<td>as MG1655 seq* rph* but ΔmutS::Kan</td>
<td>KEIO collection</td>
<td>Kan&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>RIK410</td>
<td>Mel1</td>
<td>This study</td>
<td>as SCC1 but [λ cl 377 Δ(cro-bioA) &lt;(\rightarrow) [tetRA]]</td>
<td>P1vir (SW101) x SCC1&lt;sup&gt;à&lt;/sup&gt; Tet&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Tet&lt;sup&gt;a&lt;/sup&gt;, Green fluorescence</td>
</tr>
<tr>
<td>RIK417</td>
<td>Mel36</td>
<td>This study</td>
<td>as RIK410 but ΔmutS::Kan</td>
<td>P1vir (KRE12765) x RIK410&lt;sup&gt;à&lt;/sup&gt; Kan&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Tet&lt;sup&gt;a&lt;/sup&gt;, Kan&lt;sup&gt;a&lt;/sup&gt;, Aqua fluorescence</td>
</tr>
<tr>
<td>RIK411</td>
<td>Mel14</td>
<td>This study</td>
<td>as RIK410 but gfpmut3* L&lt;sub&gt;65&lt;/sub&gt;T&lt;sub&gt;66&lt;/sub&gt;W&lt;sub&gt;66&lt;/sub&gt; (Ctgaccgg)</td>
<td>Recombineering: oligo 1 x RIK410</td>
<td>Tet&lt;sup&gt;a&lt;/sup&gt;, Aqua fluorescence</td>
</tr>
<tr>
<td>RIK418</td>
<td>Mel54</td>
<td>This study</td>
<td>as RIK411 but ΔmutS::Kan</td>
<td>Recombineering: oligo 1 x RIK417</td>
<td>Tet&lt;sup&gt;a&lt;/sup&gt;, Kan&lt;sup&gt;a&lt;/sup&gt;, Aqua fluorescence</td>
</tr>
<tr>
<td>RIK413</td>
<td>Mel24</td>
<td>This study</td>
<td>as RIK410 but gfpmut3* T&lt;sub&gt;65&lt;/sub&gt;W&lt;sub&gt;66&lt;/sub&gt; (ACGcGgG)</td>
<td>Recombineering: oligo 3 x RIK410</td>
<td>Tet&lt;sup&gt;a&lt;/sup&gt;, Mar fluorescence</td>
</tr>
<tr>
<td>RIK420</td>
<td>Mel75</td>
<td>This study</td>
<td>as RIK413 but ΔmutS::Kan</td>
<td>Recombineering: oligo 3 x RIK417</td>
<td>Tet&lt;sup&gt;a&lt;/sup&gt;, Mar, Kan&lt;sup&gt;a&lt;/sup&gt;, fluorescence</td>
</tr>
<tr>
<td>RIK414</td>
<td>Mel29</td>
<td>This study</td>
<td>as RIK410 but gfpmut3* L&lt;sub&gt;65&lt;/sub&gt;T&lt;sub&gt;66&lt;/sub&gt;W&lt;sub&gt;66&lt;/sub&gt; (GAGCtGgG)</td>
<td>Recombineering: oligo 4 x RIK410</td>
<td>Tet&lt;sup&gt;a&lt;/sup&gt;, Aqua fluorescence</td>
</tr>
<tr>
<td>RIK421</td>
<td>Mel86</td>
<td>This study</td>
<td>as RIK414 but ΔmutS::Kan</td>
<td>Recombineering: oligo 4 x RIK410</td>
<td>Tet&lt;sup&gt;a&lt;/sup&gt;, Aqua, Kan&lt;sup&gt;a&lt;/sup&gt;, fluorescence</td>
</tr>
<tr>
<td>Mel145</td>
<td>Mel145</td>
<td>This study</td>
<td>as RIK410 but gfpmut3* W&lt;sub&gt;66&lt;/sub&gt;(CC)</td>
<td>Recombineering: oligo 56 x RIK410</td>
<td>Tet&lt;sup&gt;a&lt;/sup&gt;, Violeta fluorescence</td>
</tr>
<tr>
<td>RIK419</td>
<td>Mel63</td>
<td>This study</td>
<td>as Mel145 but ΔmutS::Kan</td>
<td>Recombineering: oligo 2 x RIK417</td>
<td>Tet&lt;sup&gt;a&lt;/sup&gt;, Kan&lt;sup&gt;a&lt;/sup&gt;, Violeta fluorescence</td>
</tr>
<tr>
<td>RIK422</td>
<td>Mel101</td>
<td>This study</td>
<td>as RIK410 but gfpmut3* R&lt;sub&gt;23&lt;/sub&gt;Stop&lt;sub&gt;99&lt;/sub&gt; (GTTG)</td>
<td>Recombineering: oligo 8 x RIK410</td>
<td>Tet&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>RIK426</td>
<td>Mel115</td>
<td>This study</td>
<td>as RIK422 but ΔmutS::Kan</td>
<td>P1vir (RIK416) x RIK422&lt;sup&gt;à&lt;/sup&gt; Kan&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Tet&lt;sup&gt;a&lt;/sup&gt;, Kan&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
</tbody>
</table>
### Table 2.3 *E. coli* strains for human Recombineering

<table>
<thead>
<tr>
<th>Strain #</th>
<th>AKA</th>
<th>Source</th>
<th>Relevant Genotype</th>
<th>Construction method</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>RIK473</td>
<td>Mel172</td>
<td>This study</td>
<td>as SW101, but ΔmutS::Kan</td>
<td>Recombineering: Plvir (KRE12765) x SW101</td>
<td>Tet*, Kan*</td>
</tr>
<tr>
<td>RIK450</td>
<td>Mel136</td>
<td>This study</td>
<td>DH5a/pNL-eGFP/CEF</td>
<td>Transformation</td>
<td>Amp*</td>
</tr>
<tr>
<td>RIK498</td>
<td>Mel188</td>
<td>This study</td>
<td>DH5a/pDual-egfp</td>
<td>Transformation</td>
<td>Amp*</td>
</tr>
<tr>
<td>RIK494</td>
<td>Mel231</td>
<td>This study</td>
<td>DH5a/pDual-egfp(W66)</td>
<td>Transformation</td>
<td>Amp*</td>
</tr>
<tr>
<td>RIK495</td>
<td>Mel232</td>
<td>This study</td>
<td>DH5a/pDual-egfp(Stop66)</td>
<td>Transformation</td>
<td>Amp*</td>
</tr>
<tr>
<td>RIK496</td>
<td>Mel235</td>
<td>This study</td>
<td>DH5a/pDual-egfp(H66)</td>
<td>Transformation</td>
<td>Amp*</td>
</tr>
<tr>
<td>RIK497</td>
<td>Mel236</td>
<td>This study</td>
<td>DH5a/pDual-egfp(Y203)</td>
<td>Transformation</td>
<td>Amp*</td>
</tr>
<tr>
<td>RIK276</td>
<td>Rosetta- gami™(DE3)</td>
<td>Novagen</td>
<td>BL21(DE3) [pRARE2]</td>
<td>Cml*</td>
<td></td>
</tr>
<tr>
<td>RIK399</td>
<td>This study</td>
<td>SW102ATet/pRARE</td>
<td></td>
<td>Cml*</td>
<td></td>
</tr>
<tr>
<td>RIK403</td>
<td>This study</td>
<td>SW102ATet/pRAREACml::Kan</td>
<td>Recombineering: Kan PCR x RIK399</td>
<td>Kan*</td>
<td></td>
</tr>
<tr>
<td>RIK404</td>
<td>This study</td>
<td>DH5a/ pRAREACml::Kan</td>
<td>pRAREACml::Kan transformation</td>
<td>Kan*</td>
<td></td>
</tr>
<tr>
<td>RIK443</td>
<td>Mel132</td>
<td>This study</td>
<td>DH5a/maxGFP</td>
<td>Transformation</td>
<td>Kan*</td>
</tr>
<tr>
<td>RIK448</td>
<td>Mel132</td>
<td>This study</td>
<td>DH5a/pCMV-ICP8</td>
<td>Transformation</td>
<td>Amp*</td>
</tr>
<tr>
<td>RIK449</td>
<td>Mel134</td>
<td>This study</td>
<td>DH5a/pCMV-ICP9GFP</td>
<td>Transformation</td>
<td>Amp*</td>
</tr>
<tr>
<td>Mel317</td>
<td>RSM274</td>
<td>This study</td>
<td>NEB 5a/pCMV</td>
<td>pCMV-ICP8 Smal. Agarose gel extraction of 4 Kb band. Ligation and transformation</td>
<td>Amp*</td>
</tr>
</tbody>
</table>

### Table 2. E. coli strains for Bacterial Recombineering

<table>
<thead>
<tr>
<th>Strain #</th>
<th>AKA</th>
<th>Source</th>
<th>Relevant Genotype</th>
<th>Construction method</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>RIK423</td>
<td>Mel102</td>
<td>This study</td>
<td>as RIK410 but gfpmut3* Stop65 (G)</td>
<td>Recombineering: oligo 10 x RIK410</td>
<td>Tet*</td>
</tr>
<tr>
<td>RIK427</td>
<td>Mel117</td>
<td>This study</td>
<td>as RIK423 but AmutS::Kan</td>
<td>Plvir (RIK416) x RIK423 → Kan*</td>
<td>Tet*, Kan*</td>
</tr>
<tr>
<td>RIK424</td>
<td>Mel111</td>
<td>This study</td>
<td>as RIK410 but gfpmut3* Yso(taC)</td>
<td>Recombineering: oligo 13 x RIK411</td>
<td>Tet*, Green fluorescence</td>
</tr>
<tr>
<td>RIK425</td>
<td>Mel113</td>
<td>This study</td>
<td>as RIK410 but gfpmut3* LasT(3) Yso( (CtgACtCtaC)</td>
<td>Recombineering: oligo 16 x RIK411</td>
<td>Tet*, Green fluorescence</td>
</tr>
<tr>
<td>RIK445</td>
<td>Mel127</td>
<td>This study</td>
<td>as RIK410 but gfpmut3* T64W66, Y203 (CCaCGT; TTAC)</td>
<td>Recombineering: oligo 25 x RIK413</td>
<td>Tet*, Bronze fluorescence</td>
</tr>
<tr>
<td>Mel166</td>
<td>This study</td>
<td>as RIK445 but ΔmutS::Kan</td>
<td>Recombineering: oligo 57 x RIK470</td>
<td>Tet*, Cml*, ? fluorescence</td>
<td></td>
</tr>
<tr>
<td>RIK446</td>
<td>Mel129</td>
<td>This study</td>
<td>as RIK410 but gfpmut3* Y203 (TTAC)</td>
<td>Recombineering: oligo 13 x RIK445</td>
<td>Tet*, Amarillo fluorescence</td>
</tr>
<tr>
<td>RIK470</td>
<td>Mel143</td>
<td>This study</td>
<td>as RIK446 but ΔmutS::Kan</td>
<td>Plvir (RIK416) x RIK446 → Kan*</td>
<td>Tet*, Kan*, Amarillo fluorescence</td>
</tr>
<tr>
<td>Mel153</td>
<td>This study</td>
<td>as RIK410 but gfpmut3* W66, Y203 (CC, TTAC)</td>
<td>Recombineering: oligo 56 x RIK446</td>
<td>Tet*, fluorescence undetermined</td>
<td></td>
</tr>
<tr>
<td>Mel149</td>
<td>This study</td>
<td>as Mel153 but ΔmutS::Kan</td>
<td>Recombineering: oligo 56 x RIK470</td>
<td>Tet*, Kan*, fluorescence undetermined</td>
<td></td>
</tr>
<tr>
<td>Mel157</td>
<td>This study</td>
<td>as RIK410 but gfpmut3* LasT(3) W66, Y203 (CtgACtCtaC; TTAC)</td>
<td>Recombineering: oligo 1 x RIK446</td>
<td>Tet*, fluorescence undetermined</td>
<td></td>
</tr>
<tr>
<td>Mel161</td>
<td>This study</td>
<td>as Mel157 but ΔmutS::Kan</td>
<td>Recombineering: oligo 1 x RIK470</td>
<td>Tet*, Kan*, fluorescence undetermined</td>
<td></td>
</tr>
<tr>
<td>RIK431</td>
<td>Mel121</td>
<td>This study</td>
<td>as RIK423 but ΔsulA::Kan</td>
<td>P1vir (RIK430) x RIK423 → Kan*</td>
<td>Tet*, Kan*</td>
</tr>
<tr>
<td>RIK432</td>
<td>Mel123</td>
<td>This study</td>
<td>as RIK423 but ΔrecA::Kan</td>
<td>Plvir (RIK428) x RIK423 → Kan*</td>
<td>Tet*, Kan*</td>
</tr>
<tr>
<td>Strain #</td>
<td>AKA</td>
<td>Source</td>
<td>Relevant Genotype</td>
<td>Construction method</td>
<td>Comments</td>
</tr>
<tr>
<td>---------</td>
<td>-----</td>
<td>--------</td>
<td>------------------</td>
<td>---------------------</td>
<td>----------</td>
</tr>
<tr>
<td>Mel233</td>
<td>This study</td>
<td>DH5α/pCMV-VSVG</td>
<td>Transformation</td>
<td>Amp⁺</td>
<td></td>
</tr>
<tr>
<td>Mel234</td>
<td>This study</td>
<td>DH5α/pCMV-ΔR8.2</td>
<td>Transformation</td>
<td>Amp⁺</td>
<td></td>
</tr>
<tr>
<td>RIK544</td>
<td>Mel314</td>
<td>This study</td>
<td>DH5α/pTEC19</td>
<td>Transformation</td>
<td></td>
</tr>
<tr>
<td>NEB 5α</td>
<td>NEB (C2987I)</td>
<td>fluA2Δ(argF- lacZ)U169 phoA glnV44 Δ80 Δ(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17</td>
<td>Transformation</td>
<td>Chemo-competent cells for transformation</td>
<td></td>
</tr>
<tr>
<td>Mel316</td>
<td>NEB 10-beta (C3019I, C3019H)</td>
<td>araD139 Δ(ara-leu)7697 fluA lacX74 galK (Δ80 Δ(lacZ)M15) mcrA galU recA1 endA1 nupG rpsL (StrR) Δ(mrr-hsdRMS-mcrBC)</td>
<td>Transformation of ligation of ICP8 PCR product with primers ICP8-Flexi7 and ICP8-Flexi8 from pCMV-ICP8 with AsISI and Pmel cut pFN22K</td>
<td>Cloning large plasmids</td>
<td></td>
</tr>
<tr>
<td>RIK499</td>
<td>RSM13</td>
<td>This study</td>
<td>NEB 5α/pFN22K::ICP8</td>
<td>Transformation of ligation of ICP8 PCR product with primers ICP8-Flexi7 and ICP8-Flexi8 from pCMV-ICP8 with AsISI and Pmel cut pFN22K</td>
<td>Kan⁺</td>
</tr>
<tr>
<td>RIK500</td>
<td>RSM16</td>
<td>This study</td>
<td>NEB 5α/pFN22K::ICP8GFP</td>
<td>Transformation of ligation of ICP8GFP PCR product with primers ICP8-Flexi7 and ICP8-Flexi8 from pCMV-ICP8GFP with AsISI and Pmel cut pFN22K</td>
<td>Kan⁺</td>
</tr>
<tr>
<td>RIK580</td>
<td>RSM195</td>
<td>This study</td>
<td>NEB 5α/pUC57-Kan::NLS/HA-HumBeta</td>
<td>Sequence design, this study. Synthesis by Genewiz</td>
<td>Kan⁺</td>
</tr>
<tr>
<td>RIK581</td>
<td>RSM196</td>
<td>This study</td>
<td>NEB 5α/pUC57-Kan::NLS/Crimson-HumBeta</td>
<td>Sequence design, this study. Synthesis by Genewiz</td>
<td>Kan⁺</td>
</tr>
<tr>
<td>RIK588</td>
<td>RSM255</td>
<td>This study</td>
<td>NEB 5α/pFN24K::Crimson/P2A</td>
<td>Transformation of NEB 5α with ligation product of Product N PCR and pFN22K/BsaAI/Phosphatase</td>
<td>Kan⁺</td>
</tr>
<tr>
<td>RIK608</td>
<td>RSM307</td>
<td>This study</td>
<td>NEB 5α/pFN22K::Crimson/P2A-ICP8</td>
<td>Transformation of NEB 5α with ligation product of Crimson/P2A from RSM255/AsISI and pFN22K::ICP8/AsISI/Phosphatase</td>
<td>Kan⁺</td>
</tr>
<tr>
<td>RIK607</td>
<td>RSM283</td>
<td>This study</td>
<td>NEB 5α/pUC19::Crimson/P2A-NLS/HA-HumBeta</td>
<td>Transformation of NEB 5α with ligation product of Production N released by PciI digestion of RSM255 and pUC19::NLS/HA-HumBeta/NcoI/Phosphatase</td>
<td>Kan⁺</td>
</tr>
<tr>
<td>RIK594</td>
<td>RSM262</td>
<td>This study</td>
<td>NEB 5α/pFN24K::Crimson/P2A-ICP8</td>
<td>Transformation of NEB 5α with ligation product of Production N PCR and pFN22K/BsaAI/Phosphatase</td>
<td>Kan⁺</td>
</tr>
<tr>
<td>RIK623</td>
<td>RSM326</td>
<td>This study</td>
<td>NEB 5α/pFN22K::ICP8-P2A/Crimson</td>
<td>Transformation of NEB 5α with ligation product of Production C PCR and pFN22K::ICP8/Pmel/Phosphatase</td>
<td>Kan⁺</td>
</tr>
<tr>
<td>RIK601</td>
<td>RSM266</td>
<td>This study</td>
<td>NEB 5α/pUC19::NLS/HA-HumBeta::Product-C</td>
<td>Transformation of NEB 5α with ligation product of Production C PCR and pUC19::NLS/HA-HumBeta/Pmel/Phosphatase</td>
<td>Kan⁺</td>
</tr>
<tr>
<td>RIK622</td>
<td>RSM363</td>
<td>This study</td>
<td>NEB 5α/pFN24K::Product-C</td>
<td>Transformation of NEB 5α with ligation product of Production C-Control PCR and pFN24K/BsaAI/Phosphatase</td>
<td>Kan⁺, red fluorescent.</td>
</tr>
<tr>
<td>RIK 614</td>
<td>RSM348</td>
<td>This study</td>
<td>NEB 5α/pENTR2B/TREPit</td>
<td>Transformation</td>
<td>Kan⁺</td>
</tr>
<tr>
<td>RIK 631</td>
<td>RSM376</td>
<td>This study</td>
<td>NEB 5α/pENTR2B/TREPit::Crimson/P2A-ICP8</td>
<td>Transformation of NEB 5α with ligation product of pENTR2B/TREPit/BamHI and EcoRI/Phosphatase with Crimson/P2A-ICP8</td>
<td>Kan⁺</td>
</tr>
<tr>
<td>Strain #</td>
<td>AKA</td>
<td>Source</td>
<td>Relevant Genotype</td>
<td>Construction method</td>
<td>Comments</td>
</tr>
<tr>
<td>---------</td>
<td>---------</td>
<td>--------</td>
<td>------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------</td>
<td>----------</td>
</tr>
<tr>
<td>RIK 647</td>
<td>RSM392</td>
<td>This study</td>
<td>NEB 5α/ pENTR2B/TREPitt::Crimson/P2A-NLS/HA-HumBeta</td>
<td>Transformation of NEB 5α with ligation product of pENTR2B/TREPitt/BamHI and EcoRI/Phosphatase with Crimson/P2A-P2A-NLS/HA-HumBeta</td>
<td>Kan^R</td>
</tr>
<tr>
<td>RIK 655</td>
<td>RSM400</td>
<td>This study</td>
<td>NEB 5α/ pENTR2B/TREPitt::Crimson/P2A</td>
<td>Transformation of NEB 5α with ligation product of pENTR2B/TREPitt/BamHI and EcoRI/Phosphatase with Crimson/P2A-P2A</td>
<td>Kan^R</td>
</tr>
<tr>
<td>RIK 663</td>
<td>RSM408</td>
<td>This study</td>
<td>NEB 5α/ pENTR2B/TREPitt::ICP8/P2A-Crimson</td>
<td>Transformation of NEB 5α with ligation product of pENTR2B/TREPitt/BamHI and EcoRI/Phosphatase with ICP8/P2A-Crimson</td>
<td>Kan^R</td>
</tr>
<tr>
<td>RIK 664</td>
<td>RSM409</td>
<td>This study</td>
<td>NEB 5α/ pENTR2B/TREPitt::ICP8/P2A-Crimson</td>
<td>Transformation of NEB 5α with ligation product of pENTR2B/TREPitt/BamHI and EcoRI/Phosphatase with ICP8/P2A-Crimson</td>
<td>Kan^R</td>
</tr>
<tr>
<td>RIK 671</td>
<td>RSM416</td>
<td>This study</td>
<td>NEB 5α/ pENTR2B/TREPitt::ICP8/P2A-Crimson</td>
<td>Transformation of NEB 5α with ligation product of pENTR2B/TREPitt/BamHI and EcoRI/Phosphatase with P2A-Crimson</td>
<td>Kan^R</td>
</tr>
<tr>
<td>RIK 710</td>
<td>Stbl3</td>
<td>Invitrogen (C737303)</td>
<td>F- merB mrr hsdS20(rB-, mB-) recA13 supE44 ara-14 galK2 lacY1 proA2 rpsL20(StrR) xyl-5 λ-leu mt1</td>
<td>Chemically competent cells, good for cloning lentiviruses</td>
<td></td>
</tr>
<tr>
<td>RIK 579</td>
<td>RSM222</td>
<td>Addgene (108480)</td>
<td>pSLIK-Zeocin</td>
<td>Unknown strain containing the plasmid.</td>
<td>Amp^R, Zeo^R</td>
</tr>
<tr>
<td>RIK 711</td>
<td>RSM479</td>
<td>This study</td>
<td>Stbl3/pSLIK/TREPitt::Crimson/P2A-ICP8</td>
<td>Transformation of Stbl3 with recombination product of pENTR2B/TREPitt::Crimson/P2A-ICP8 clone 1 (pENTER) with pSLIK-Zeocin (pDEST)</td>
<td>Amp^R, Zeo^R</td>
</tr>
<tr>
<td>RIK 715</td>
<td>RSM480</td>
<td>This study</td>
<td>Stbl3/pSLIK/TREPitt::Crimson/P2A-NLS/HA-HumBeta</td>
<td>Transformation of Stbl3 with recombination product of pENTR2B/TREPitt:Crimson/P2A-NLS/HA-HumBeta clone 17 (pENTER) with pSLIK-Zeocin (pDEST)</td>
<td>Amp^R, Zeo^R</td>
</tr>
<tr>
<td>RIK 719</td>
<td>RSM481</td>
<td>This study</td>
<td>Stbl3/pSLIK/TREPitt::Crimson/P2A</td>
<td>Transformation of Stbl3 with recombination product of pENTR2B/TREPitt::Crimson/P2A clone 25 (pENTER) with pSLIK-Zeocin (pDEST)</td>
<td>Amp^R, Zeo^R</td>
</tr>
<tr>
<td>RIK 723</td>
<td>RSM482</td>
<td>This study</td>
<td>Stbl3/pSLIK/TREPitt::ICP8/P2A-Crimson</td>
<td>Transformation of Stbl3 with recombination product of pENTR2B/TREPitt::ICP8/P2A-Crimson clone 43 (pENTER) with pSLIK-Zeocin (pDEST)</td>
<td>Amp^R, Zeo^R</td>
</tr>
<tr>
<td>RIK 728</td>
<td>RSM483</td>
<td>This study</td>
<td>Stbl3/pSLIK/TREPitt::NLS/HA-HumBeta-P2A-Crimson</td>
<td>Transformation of Stbl3 with recombination product of pENTR2B/TREPitt::NLS/HA-HumBeta/P2A-Crimson clone 50 (pENTER) with pSLIK-Zeocin (pDEST)</td>
<td>Amp^R, Zeo^R</td>
</tr>
</tbody>
</table>
Table 2.4 *E. coli* strains for SynExo interactome studies

<table>
<thead>
<tr>
<th>Strain #</th>
<th>AKA</th>
<th>Source</th>
<th>Relevant Genotype</th>
<th>Construction method</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>RIK731</td>
<td>RSM484</td>
<td>This study</td>
<td>Stbl3:pSLIK/TREPitt::P2A/Crimson-Control</td>
<td>Transformation of Stbl3 with recombination product of pENTR2B/TREPitt::P2A/Crimson clone 33 (pENTER) with pSLIK-Zeocin (pDEST)</td>
<td>Amp(^a), Zeo(^b)</td>
</tr>
<tr>
<td>RIK124</td>
<td></td>
<td></td>
<td>Cosmid with HSV1 strain 17 DNA nucleotides: 2995-45035</td>
<td></td>
<td>Amp(^b)</td>
</tr>
<tr>
<td>RIK452</td>
<td>RSM253</td>
<td>NEB, # E3000</td>
<td>DH5α/pNEBR-X1Hygro</td>
<td>RheoSwitch® Mammalian Inducible Expression System</td>
<td>Amp(^b), Hygro(^b)</td>
</tr>
<tr>
<td>RIK373</td>
<td>RSM150</td>
<td>NEB, # E3000</td>
<td>XL10-Gold/pNEBR-R1</td>
<td>RheoSwitch® Mammalian Inducible Expression System</td>
<td>Amp(^b), Neo(^b)</td>
</tr>
<tr>
<td>RIK374</td>
<td>RSM152</td>
<td>NEB, # E3000</td>
<td>DH5α/pNEBR-X1Glc</td>
<td>RheoSwitch® Mammalian Inducible Expression System</td>
<td>Amp(^b)</td>
</tr>
<tr>
<td>RIK372</td>
<td>RSM118</td>
<td></td>
<td>XL10-Gold/pNEBRX1-UL12.5</td>
<td>UL12.5 PCR amplified with oligos 164/165 and ligated EcoRV and NotI</td>
<td>Amp(^b), Hygro(^b)</td>
</tr>
<tr>
<td>RIK368</td>
<td>RSM107</td>
<td></td>
<td>XL10-Gold/pNEBRX1-UL12.5-HA</td>
<td>UL12.5-HA PCR amplified with oligos 164/166 and ligated EcoRV and NotI</td>
<td>Amp(^b), Hygro(^b)</td>
</tr>
<tr>
<td>RIK370</td>
<td>RSM113</td>
<td></td>
<td>XL10-Gold/pNEBRX1-MLS</td>
<td>oligos 173 and 174 hybridized and cloned HindIII/Apal</td>
<td>Amp(^b), Hygro(^b)</td>
</tr>
<tr>
<td>Mel302</td>
<td>RSM158</td>
<td></td>
<td>DH5α/pNEBR-X1Hygro::ICP8-GFP</td>
<td>Ligation of pNEBR-X1/HindIII/EcoRVand pCMV-ICP8-GFP/Hpal/HindIII</td>
<td>Amp(^b), Hygro(^b)</td>
</tr>
<tr>
<td>Mel301</td>
<td>RSM185</td>
<td></td>
<td>DH5α/pNEBR-X1Hygro::ICP8</td>
<td>Ligation of pNEBR-X1/HindIII/EcoRVand pCMV-ICP8/Hpal/HindIII</td>
<td>Amp(^b), Hygro(^b)</td>
</tr>
</tbody>
</table>

Table 2.4 *E. coli* strains for SynExo interactome studies

<table>
<thead>
<tr>
<th>Strain #</th>
<th>AKA</th>
<th>Source</th>
<th>Relevant Genotype</th>
<th>Construction method</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mel241</td>
<td>Gift of Kenn Rudd at UM (HIM2694)</td>
<td>MG1655* ΔdnaT::Kan</td>
<td>KEIO collection</td>
<td>Kan(^a)</td>
<td></td>
</tr>
<tr>
<td>Mel242</td>
<td>Gift of Kenn Rudd (KRE13223)</td>
<td>MG1655* ΔkdsD::Kan</td>
<td>KEIO collection</td>
<td>Kan(^a)</td>
<td></td>
</tr>
<tr>
<td>Mel243</td>
<td>Gift of Kenn Rudd (KRE12212)</td>
<td>MG1655* ΔyejH::Kan</td>
<td>KEIO collection</td>
<td>Kan(^a)</td>
<td></td>
</tr>
<tr>
<td>Mel244</td>
<td>Gift of Kenn Rudd (KRE10934)</td>
<td>MG1655* ΔompF::Kan</td>
<td>KEIO collection</td>
<td>Kan(^a)</td>
<td></td>
</tr>
<tr>
<td>Mel245</td>
<td>Gift of Kenn Rudd (KRE10435)</td>
<td>MG1655* Δlon::Kan</td>
<td>KEIO collection</td>
<td>Kan(^a)</td>
<td></td>
</tr>
<tr>
<td>Mel246</td>
<td>Gift of Kenn Rudd (KRE10433)</td>
<td>MG1655* ΔclpP::Kan</td>
<td>KEIO collection</td>
<td>Kan(^a)</td>
<td></td>
</tr>
<tr>
<td>Mel247</td>
<td>Gift of Kenn Rudd (KRE10434)</td>
<td>MG1655* ΔclpP::Kan</td>
<td>KEIO collection</td>
<td>Kan(^a)</td>
<td></td>
</tr>
<tr>
<td>Mel248</td>
<td>Gift of Kenn Rudd (KRE10963)</td>
<td>MG1655* ΔulA::Kan</td>
<td>KEIO collection</td>
<td>Kan(^a)</td>
<td></td>
</tr>
<tr>
<td>Mel249</td>
<td>Gift of Kenn Rudd (KRE12116)</td>
<td>MG1655* ΔyegS::Kan</td>
<td>KEIO collection</td>
<td>Kan(^a)</td>
<td></td>
</tr>
<tr>
<td>Mel250</td>
<td>Gift of Kenn Rudd (KRE10197)</td>
<td>MG1655* ΔgmhB::Kan</td>
<td>KEIO collection</td>
<td>Kan(^a)</td>
<td></td>
</tr>
<tr>
<td>Mel251</td>
<td>Gift of Kenn Rudd (BW25113)</td>
<td>MG1655* ΔruvC::Kan</td>
<td>KEIO collection</td>
<td>Kan(^a)</td>
<td></td>
</tr>
<tr>
<td>RIK1</td>
<td>C600</td>
<td>Lab</td>
<td>sull, thi, thr, leu</td>
<td></td>
<td>To grow amber mutant λ</td>
</tr>
<tr>
<td>VDL1</td>
<td>This study</td>
<td>As RIK410 but ΔrusB11::Kan</td>
<td>P1 transduction</td>
<td>Kan(^a)</td>
<td></td>
</tr>
<tr>
<td>VDL2</td>
<td>This study</td>
<td>As RIK410 but ΔrusB12::Kan</td>
<td>P1 transduction</td>
<td>Kan(^a)</td>
<td></td>
</tr>
<tr>
<td>VDL3</td>
<td>This study</td>
<td>As RIK423 but ΔrusB21::Kan</td>
<td>P1 transduction</td>
<td>Kan(^a)</td>
<td></td>
</tr>
</tbody>
</table>
### E. coli strains for SynExo Interactome studies

<table>
<thead>
<tr>
<th>Strain #</th>
<th>AKA</th>
<th>Source</th>
<th>Relevant Genotype</th>
<th>Construction method</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>VDL4</td>
<td>This study</td>
<td>As RIK423 but ΔvwB22::Kan</td>
<td>P1 transduction</td>
<td>Kan&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>VDL5</td>
<td>This study</td>
<td>As RIK410 but ΔwvA11::Kan</td>
<td>P1 transduction</td>
<td>Kan&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>VDL6</td>
<td>This study</td>
<td>As RIK410 but ΔvwA12::Kan</td>
<td>P1 transduction</td>
<td>Kan&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>VDL7</td>
<td>This study</td>
<td>As RIK423 but ΔwvA21::Kan</td>
<td>P1 transduction</td>
<td>Kan&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>VDL8</td>
<td>This study</td>
<td>As RIK423 but ΔvwA22::Kan</td>
<td>P1 transduction</td>
<td>Kan&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>VDL9</td>
<td>This study</td>
<td>As RIK410 but ΔuvwC11::Kan</td>
<td>P1 transduction</td>
<td>Kan&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>VDL10</td>
<td>This study</td>
<td>As RIK410 but ΔuvwC12::Kan</td>
<td>P1 transduction</td>
<td>Kan&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>VDL11</td>
<td>This study</td>
<td>As RIK423 but ΔuvwC21::Kan</td>
<td>P1 transduction</td>
<td>Kan&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>VDL12</td>
<td>This study</td>
<td>As RIK423 but ΔuvwC22::Kan</td>
<td>P1 transduction</td>
<td>Kan&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>VDL13</td>
<td>This study</td>
<td>As RIK410 but ΔuvwD11::Kan</td>
<td>P1 transduction</td>
<td>Kan&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>VDL14</td>
<td>This study</td>
<td>As RIK410 but ΔuvwD12::Kan</td>
<td>P1 transduction</td>
<td>Kan&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>VDL15</td>
<td>This study</td>
<td>As RIK423 but ΔuvwD21::Kan</td>
<td>P1 transduction</td>
<td>Kan&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>VDL16</td>
<td>This study</td>
<td>As RIK423 but ΔuvwD22::Kan</td>
<td>P1 transduction</td>
<td>Kan&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>VDL17</td>
<td>This study</td>
<td>As RIK410 but Δpici11::Kan</td>
<td>P1 transduction</td>
<td>Kan&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>VDL18</td>
<td>This study</td>
<td>As RIK410 but Δpici12::Kan</td>
<td>P1 transduction</td>
<td>Kan&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>VDL19</td>
<td>This study</td>
<td>As RIK423 but Δpici21::Kan</td>
<td>P1 transduction</td>
<td>Kan&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>VDL20</td>
<td>This study</td>
<td>As RIK423 but Δpici22::Kan</td>
<td>P1 transduction</td>
<td>Kan&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
</tbody>
</table>

Genetic interactions in Recombineering, Hypermorphic alleles

| Mel253 | JW1850 | Gift of Kenn Rudd | As AG1 recA1 endA1 gryA96 thi-1 hsdR17(rK-mK+.supI) supA4 relA1::pCA24N::ruvA | ASKA collection (Katagawa et al. 2005) | Cm<sup>a</sup> |
| Mel254 | JW1849 | Gift of Kenn Rudd | AG1/pCA24N::ruvB | ASKA collection | Cm<sup>a</sup> |
| Mel255 | JW2403 | Gift of Kenn Rudd | AG1/pCA24N::ils4 (ilo) | ASKA collection | Cm<sup>a</sup> |
| Mel256 | JW3038 | Gift of Kenn Rudd | AG1/pCA24N::dhaA | ASKA collection | Cm<sup>a</sup> |
| Mel257 | JW2669 | Gift of Kenn Rudd | AG1/pCA24N::recA | ASKA collection | Cm<sup>a</sup> |
| Mel258 | JW3627 | Gift of Kenn Rudd | AG1/pCA24N::recG | ASKA collection | Cm<sup>a</sup> |
| Mel259 | JW4020 | Gift of Kenn Rudd | AG1/pCA24N::dshB | ASKA collection | Cm<sup>a</sup> |
| Mel260 | JW1993 | Gift of Kenn Rudd | AG1/pCA24N::iscB | ASKA collection | Cm<sup>a</sup> |
| Mel261 | JW3906 | Gift of Kenn Rudd | AG1/pCA24N::priA | ASKA collection | Cm<sup>a</sup> |
| Mel262 | JW4159 | Gift of Kenn Rudd | AG1/pCA24N::priB | ASKA collection | Cm<sup>a</sup> |
| Mel263 | JW0456 | Gift of Kenn Rudd | AG1/pCA24N::priC | ASKA collection | Cm<sup>a</sup> |
| Mel264 | JW3786 | Gift of Kenn Rudd | AG1/pCA24N::uvrD | ASKA collection | Cm<sup>a</sup> |
| Mel265 | JW1898 | Gift of Kenn Rudd | AG1/pCA24N::yviC | ASKA collection | Cm<sup>a</sup> |
| RIK506 | JW0429 | Gift of Kenn Rudd | AG1/pCA24N::lon | ASKA collection | Cm<sup>a</sup> |
| RIK507 | JW0428 | Gift of Kenn Rudd | AG1/pCA24N::clpX | ASKA collection | Cm<sup>a</sup> |
| RIK508 | JW0427 | Gift of Kenn Rudd | AG1/pCA24N::clpP | ASKA collection | Cm<sup>a</sup> |
| RIK509 | JW1056 | Gift of Kenn Rudd | AG1/pCA24N::mviN | ASKA collection | Cm<sup>a</sup> |
| RIK510 | JW3145 | Gift of Kenn Rudd | AG1/pCA24N::fsH | ASKA collection | Cm<sup>a</sup> |
| RIK511 | JW0941 | Gift of Kenn Rudd | AG1/pCA24N::sulA | ASKA collection | Cm<sup>a</sup> |
| RIK512 | JW2070 | Gift of Kenn Rudd | AG1/pCA24N::segS | ASKA collection | Cm<sup>a</sup> |
| RIK513 | JW4326 | Gift of Kenn Rudd | AG1/pCA24N::dneT | ASKA collection | Cm<sup>a</sup> |
| RIK514 | JW2070 | Gift of Kenn Rudd | AG1/pCA24N::adsD | ASKA collection | Cm<sup>a</sup> |
| RIK515 | JW0912 | Gift of Kenn Rudd | AG1/pCA24N::ompF | ASKA collection | Cm<sup>a</sup> |
| RIK516 | JW0196 | Gift of Kenn Rudd | AG1/pCA24N::gmhB | ASKA collection | Cm<sup>a</sup> |
| RIK517 | JW1852 | Gift of Kenn Rudd | AG1/pCA24N::ruvC | ASKA collection | Cm<sup>a</sup> |

SynExo protein expression and pull downs

| RIK343 | Gift of Don Court (NCI) | JM109/pSIM5 | Cm<sup>a</sup> |
| RIK60  | Mel266 | Lab | BL21(DE3) | Cm<sup>a</sup> |
| RIK311 | This study | BL21(DE3)pLysS | Electroporation | Cm<sup>a</sup> |
| RIK490 | Mel267 | This study | BL21(DE3)pFC15K::β | Electroporation | Kan<sup>a</sup> |
| RIK491 | Mel268 | This study | BL21(DE3)pFC15K::β1-177 | Electroporation | Kan<sup>a</sup> |
| RIK492 | Mel269 | This study | BL21(DE3)pFC15K::Exo | Electroporation | Kan<sup>a</sup> |
| RIK493 | Mel270 | This study | BL21(DE3)pFC15K::Exo(1-189) | Electroporation | Kan<sup>a</sup> |
| RIK532 | Mel272 | This study | BL21(DE3)pFN22K::β | Electroporation | Kan<sup>a</sup> |
| RIK533 | Mel273 | This study | BL21(DE3)pFN22K::Exo | Electroporation | Kan<sup>a</sup> |
| RIK534 | Mel274 | This study | BL21(DE3)pFN22K::Exo(1-189) | Electroporation | Kan<sup>a</sup> |
Table 2. 

<table>
<thead>
<tr>
<th>Strain #</th>
<th>AKA</th>
<th>Source</th>
<th>Relevant Genotype</th>
<th>Construction method</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>RIK518</td>
<td>Mel275, MA1</td>
<td>This study</td>
<td>DH5α/pFN22K::β</td>
<td>Electroporation Kanα</td>
<td></td>
</tr>
<tr>
<td>RIK519</td>
<td>Mel276, ZC1</td>
<td>This study</td>
<td>DH5α/pFN22K::β(1-177)</td>
<td>Electroporation Kanα</td>
<td></td>
</tr>
<tr>
<td>Mel277</td>
<td>JG1</td>
<td>This study</td>
<td>DH5α/pFN22K::Exo</td>
<td>Electroporation Kanα</td>
<td></td>
</tr>
<tr>
<td>Mel278</td>
<td>KK1</td>
<td>This study</td>
<td>DH5α/pFN22K::Exo(1-189)</td>
<td>Electroporation Kanα</td>
<td></td>
</tr>
<tr>
<td>RIK483</td>
<td>Mel279</td>
<td>This study</td>
<td>JM109/pFC15K::β</td>
<td>Electroporation Kanα</td>
<td></td>
</tr>
<tr>
<td>RIK484</td>
<td>Mel280</td>
<td>This study</td>
<td>JM109/pFC15K::β(1-177)</td>
<td>Electroporation Kanα</td>
<td></td>
</tr>
<tr>
<td>RIK485</td>
<td>Mel281</td>
<td>This study</td>
<td>JM109/pFC15K::Exo</td>
<td>Electroporation Kanα</td>
<td></td>
</tr>
<tr>
<td>RIK486</td>
<td>Mel282</td>
<td>This study</td>
<td>JM109/pFC15K::Exo(1-189)</td>
<td>Electroporation Kanα</td>
<td></td>
</tr>
<tr>
<td>RIK187</td>
<td>Mel283</td>
<td>Myers’ Lab</td>
<td>XL-1 Blue/pET28a+::his-redA(R28A)</td>
<td>(Subramanian et al. 2003)</td>
<td>Kanα</td>
</tr>
<tr>
<td>RIK188</td>
<td>Mel284</td>
<td>Myers’ Lab</td>
<td>XL-1 Blue/pET28a+::his-redA(D119A)</td>
<td>(Subramanian et al. 2003)</td>
<td>Kanα</td>
</tr>
<tr>
<td>RIK189</td>
<td>Mel285</td>
<td>Myers’ Lab</td>
<td>XL-1 Blue/pET28a+::his-redβ</td>
<td>(Tolun 2008)</td>
<td>Kanα</td>
</tr>
<tr>
<td>RIK254</td>
<td>Mel286</td>
<td>Myers’ Lab</td>
<td>DH5α/pET28a+::his-redβ</td>
<td>(Subramanian et al. 2003)</td>
<td>Kanα</td>
</tr>
<tr>
<td>Mel287</td>
<td>This study</td>
<td>BL21(DE3)pCA24N::zvuA</td>
<td>Electroporation CmⅨ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mel288</td>
<td>This study</td>
<td>BL21(DE3)pCA24N::zuvA+::pFN22K::β</td>
<td>Electroporation CmⅨ, Kanα</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mel289</td>
<td>This study</td>
<td>BL21(DE3)pCA24N::zuvA+::pFN22K::β</td>
<td>Electroporation CmⅨ, Kanα</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mel290</td>
<td>This study</td>
<td>BL21(DE3)pCA24N::zuvA+::pFN22K::β</td>
<td>Electroporation CmⅨ, Kanα</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mel292</td>
<td>This study</td>
<td>BL21(DE3)pCA24N::lon</td>
<td>Electroporation CmⅨ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mel293</td>
<td>This study</td>
<td>BL21(DE3)pCA24N::clpX</td>
<td>Electroporation CmⅨ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mel294</td>
<td>This study</td>
<td>BL21(DE3)pCA24N::clpP</td>
<td>Electroporation CmⅨ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mel295</td>
<td>This study</td>
<td>BL21(DE3)pCA24N::zmaT</td>
<td>Electroporation CmⅨ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mel296</td>
<td>This study</td>
<td>BL21(DE3)pCA24N::lgaA</td>
<td>Electroporation CmⅨ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mel297</td>
<td>This study</td>
<td>BL21(DE3)pCA24N::recA</td>
<td>Electroporation CmⅨ, Kanα</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mel312</td>
<td>This study</td>
<td>BL21(DE3)pLYSs and pFN22K::Exo</td>
<td>Electroporation CmⅨ, Kanα</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mel313</td>
<td>This study</td>
<td>BL21(DE3)pLYSs and pFN22K::Exo1-189</td>
<td>Electroporation CmⅨ, Kanα</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZC9</td>
<td>This study</td>
<td>BL21(DE3)pCA24N::ClpX-pFN22K::β</td>
<td>Electroporation CmⅨ, Kanα</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZC10</td>
<td>This study</td>
<td>BL21(DE3)pCA24N::ClpX-pFN22K::Exo</td>
<td>Electroporation CmⅨ, Kanα</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZC11</td>
<td>This study</td>
<td>BL21(DE3)pCA24N::β</td>
<td>Electroporation Kanα</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZC12</td>
<td>This study</td>
<td>BL21(DE3)pCA24N::Exo</td>
<td>Electroporation Kanα</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RIK535</td>
<td>This study</td>
<td>As RIK 410 but gfpMut3+::kanα</td>
<td>Recombineering Kanα as fused to GFP; Kanα arises from knocking out GFP translation. Recombineering reporter for hypo- and hypermorphic allele studies.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Oligo list

Table 2.5 Oligo list

<table>
<thead>
<tr>
<th>#</th>
<th>Oligo name</th>
<th>Oligo sequence (5’ to 3’)</th>
<th>Oligo function/description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>GFPmut3-CFP</td>
<td>actactggaaaaacactgcgttcattgcaacatctgaactgggggtttgcatgataggataccagatcatatgaaacagcat</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>99 nt, sense oligo to change GFPmut3+::Fd,Gs,Y66 to Fd,Ms,W66 (Aqu) by Recombineering.</td>
</tr>
<tr>
<td>2</td>
<td>Y66W</td>
<td>ggatgcttgcattatcgctgggtatctgccgaaacagttgcatgataggataccagatcatatgaaacagcatcatatgaaacagcat</td>
<td>90 nt, anti-sense oligo to change GFPmut3+::Y66W (Violeta) by Recombineering.</td>
</tr>
<tr>
<td>3</td>
<td>G651-Y66W</td>
<td>ggatgcttgcattatcgctgggtatctgccgaaacagttgcatgataggataccagatcatatgaaacagcatcatatgaaacagcat</td>
<td>90 nt, anti-sense oligo to change GFPmut3+::G651,Y66 to Fd,Ms,W66 (Mat) by Recombineering.</td>
</tr>
<tr>
<td>Oligo name</td>
<td>Oligo sequence (5’ to 3’)</td>
<td>Oligo function/description</td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>---------------------------</td>
<td>----------------------------</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>f64L/G65T/Y66W</td>
<td>cagctgcctatatgctggatctgcaaacatggaac atccaccaaatggaacagtgtggtgccatggaacaaggtgctggt</td>
<td>90 nt, anti-sense oligo to change GFPmut3* F&lt;sub&gt;64&lt;/sub&gt;G&lt;sub&gt;65&lt;/sub&gt;T&lt;sub&gt;66&lt;/sub&gt;W towards 5' UTR (for Recombineering)</td>
</tr>
<tr>
<td>5</td>
<td>GFPmut3-F</td>
<td>tttctactgaggtgttcccaaa</td>
<td>For PCR amplification of chromophore region of GFPmut3*.</td>
</tr>
<tr>
<td>6</td>
<td>GFPmut3-R</td>
<td>tcattaacaggggtacctacccaa</td>
<td>For PCR amplification of chromophore region of GFPmut3*.</td>
</tr>
<tr>
<td>7</td>
<td>GFPmut3-seq</td>
<td>ggtggtgtaagcttggcaaca</td>
<td>For sequencing the chromophore region of GFPmut3*.</td>
</tr>
<tr>
<td>8</td>
<td>GFPmut3*P58R/T59 stop</td>
<td>gtatcgcgaacagcaacaaaaactgagagtga caagtctacccagaatctgagattgtctctctctggtgctgccaa ataaatatttggg</td>
<td>95 nt, anti-sense oligo to change GFPmut3* P&lt;sub&gt;58&lt;/sub&gt;R/T&lt;sub&gt;59&lt;/sub&gt; to R&lt;sub&gt;59&lt;/sub&gt;T&lt;sub&gt;59&lt;/sub&gt; (Dark) by Recombineering.</td>
</tr>
<tr>
<td>9</td>
<td>GFPmut3*R58P/stop 59T</td>
<td>gtatcgcgaacagcaacaaaaactgagagtga caagtctacccagaatctgagattgtctctctggtgctgccaa ataaatatttggg</td>
<td>95 nt, anti-sense oligo as GFPmut3*, to revert RIK422 and RIK426 by Recombineering.</td>
</tr>
<tr>
<td>10</td>
<td>GFPmut3*Y39stop</td>
<td>ggaacaggtttttcactgatgcaaaatatttaaggtg agttcttcattgtactcactctacacctctagcagcaag</td>
<td>97 nt, anti-sense oligo to change GFPmut3* Y&lt;sub&gt;39&lt;/sub&gt; to Y&lt;sub&gt;39&lt;/sub&gt;Stop (Dark) by Recombineering.</td>
</tr>
<tr>
<td>11</td>
<td>GFPmut3*stop39Y</td>
<td>ggccaatacttctctcagagggggtagatgctgca acacgaaacttcacactaattttttctctctactctgtgactgcagcaag</td>
<td>97 nt, sense oligo as GFPmut3*, to revert RIK423 and RIK427 by Recombineering.</td>
</tr>
<tr>
<td>12</td>
<td>BFPmut3*W66Y_L1</td>
<td>gatgcgtctgcattatgctggatctgcaaacatggaac atccaccaaatggaacagtgtggtgccatggaacaaggtgctggt</td>
<td>98 nt, anti-sense oligo as GFPmut3*, with a silent mutation of Y66Y to revert mutations in chromophore region by Recombineering with a less recognizable mismatch AC/CC vs original AT/CC.</td>
</tr>
<tr>
<td>13</td>
<td>BFPmut3*W66Y_L2</td>
<td>ctggaaccacctctctctcagagggggtagatgctgca acacgaaacttcacactaattttttctctctactctgtgactgcagcaag</td>
<td>98 nt, sense oligo as GFPmut3*, with a silent mutation of Y66Y to revert mutations in chromophore region by Recombineering with a less recognizable mismatch AC/CC vs original AT/CC.</td>
</tr>
<tr>
<td>14</td>
<td>BFPmut3*W66Y_S1</td>
<td>catatgatgcgtctgcattatgctggatctgcaaacatggaac atccaccaaatggaacagtgtggtgccatggaacaaggtgctggt</td>
<td>74 nt, anti-sense oligo as GFPmut3*, with a silent mutation of Y66Y to revert mutations in chromophore region by Recombineering with a less recognizable mismatch AC/CC vs original AT/CC.</td>
</tr>
<tr>
<td>15</td>
<td>BFPmut3*W66Y_S2</td>
<td>ctggtctggccacacctctctcagagggggtagatgctgca acacgaaacttcacactaattttttctctctactctgtgactgcagcaag</td>
<td>74 nt, sense oligo as GFPmut3*, with a silent mutation of Y66Y to revert mutations in chromophore region by Recombineering with a less recognizable mismatch AC/CC vs original AT/CC.</td>
</tr>
<tr>
<td>16</td>
<td>CFPmut3*W66Y_L2</td>
<td>ctggaaccacctctctctcagagggggtagatgctgca acacgaaacttcacactaattttttctctctactctgtgactgcagcaag</td>
<td>98 nt, sense oligo to change Aqua L&lt;sub&gt;64&lt;/sub&gt;T&lt;sub&gt;65&lt;/sub&gt;W&lt;sub&gt;66&lt;/sub&gt; to L&lt;sub&gt;64&lt;/sub&gt;T&lt;sub&gt;65&lt;/sub&gt;W&lt;sub&gt;66&lt;/sub&gt; (Green) by Recombineering with mismatch AC/CC.</td>
</tr>
<tr>
<td>17</td>
<td>GFP_left_flank</td>
<td>tcgtacttcatctctcttctgcatctgtactgtca</td>
<td>To amplify GFP junctions in E. coli MG1655.</td>
</tr>
<tr>
<td>18</td>
<td>GFP_right_flank</td>
<td>cgaacttctcaacaccctgcaaca</td>
<td>To amplify GFP junctions in E. coli MG1655.</td>
</tr>
<tr>
<td>19</td>
<td>GFPmut3*stop39Y_7/nt</td>
<td>aaattttcttctctctcagagggggtagatgctgca acacgaaacttcacactaattttttctctctactctgtgactgcagcaag</td>
<td>87 nt, sense oligo as GFPmut3*, to revert RIK423 and RIK427 by Recombineering. To study effect of oligo length.</td>
</tr>
<tr>
<td>20</td>
<td>GFPmut3*stop39Y_7/nt</td>
<td>ttctctgtgggaaccacctctctcagagggggtagatgctgca acacgaaacttcacactaattttttctctctactctgtgactgcagcaag</td>
<td>77 nt, sense oligo as GFPmut3*, to revert RIK423 and RIK427 by Recombineering. To study effect of oligo length.</td>
</tr>
<tr>
<td>21</td>
<td>GFPmut3*stop39Y_6/nt</td>
<td>tcggtactggaggtgctggaacatctggacaaacctctg tcccaattttttctctctactctgtgactgcagcaag</td>
<td>67 nt, sense oligo as GFPmut3*, to revert RIK423 and RIK427 by Recombineering. To study effect of oligo length.</td>
</tr>
<tr>
<td>22</td>
<td>GFPmut3*stop39Y_6/nt</td>
<td>gcacaaaattttttctctctcagagggggtagatgctgca acacgaaacttcacactaattttttctctctactctgtgactgcagcaag</td>
<td>97 nt, sense oligo as GFPmut3*, to revert RIK423 and RIK427 by Recombineering. To compare C.C mismatch to T.C (oligo11 vs 22).</td>
</tr>
<tr>
<td>23</td>
<td>CFPmut3*W66Y_AT.CC</td>
<td>cgtggaaccacctctctctcagagggggtagatgctgca acacgaaacttcacactaattttttctctctactctgtgactgcagcaag</td>
<td>98 nt, sense oligo to change Aqua L&lt;sub&gt;64&lt;/sub&gt;T&lt;sub&gt;65&lt;/sub&gt;W&lt;sub&gt;66&lt;/sub&gt; to L&lt;sub&gt;64&lt;/sub&gt;T&lt;sub&gt;65&lt;/sub&gt;W&lt;sub&gt;66&lt;/sub&gt; (Green) by Recombineering with mismatch AT/CC. To compare mismatch AC/CC (oligo 16) to AT/CC.</td>
</tr>
<tr>
<td>24</td>
<td>GFPmut3*R58P/stop 59T-lagging</td>
<td>ttaccccaattttttctctctcagagggggtagatgctgca acacgaaacttcacactaattttttctctctactctgtgactgcagcaag</td>
<td>99 nt, sense oligo as GFPmut3*, to revert RIK422 and RIK426 by Recombineering. To study effect of mismatch size (4nt) in Recombineering efficiency</td>
</tr>
<tr>
<td>25</td>
<td>GFPmut3*(T203Y)</td>
<td>ccaaggtgaggggtgccctcgttcttcacgcacaaaccttctccttctctctcagagggggtagatgctgca acacgaaacttcacactaattttttctctctactctgtgactgcagcaag</td>
<td>99 nt, sense oligo to change GFPmut3* T&lt;sub&gt;203&lt;/sub&gt;Y to change GFPmut3* T&lt;sub&gt;203&lt;/sub&gt;Y (yellow) by Recombineering.</td>
</tr>
<tr>
<td>26</td>
<td>YFP-PCR (Y203)</td>
<td>gtcttttcacgcacaaaccttctccttctctctcagagggggtagatgctgca acacgaaacttcacactaattttttctctctactctgtgactgcagcaag</td>
<td>To screen for GFPmut3*(203Y) recombinants.</td>
</tr>
<tr>
<td>27</td>
<td>GFP-PCR (T203)</td>
<td>gtcttttcacgcacaaaccttctccttctctctcagagggggtagatgctgca acacgaaacttcacactaattttttctctctactctgtgactgcagcaag</td>
<td>To screening for GFPmut3*(T203) non-recombinants.</td>
</tr>
<tr>
<td>44</td>
<td>EGF N_out</td>
<td>cgtggtggcagcttcagcagcaag</td>
<td>To sequence from eGFP 5' end towards 5' UTR.</td>
</tr>
<tr>
<td>#</td>
<td>Oligo name</td>
<td>Oligo sequence (5' to 3')</td>
<td>Oligo function/description</td>
</tr>
<tr>
<td>----</td>
<td>----------------------------</td>
<td>------------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>45</td>
<td>EGFP_forward</td>
<td>gacagacggcggagggagatgtc</td>
<td>To sequence eGFP from 5' end and also used for allele-specific PCR.</td>
</tr>
<tr>
<td>46</td>
<td>EGFP_reverse</td>
<td>gtctagctgggagatgtccg</td>
<td>To sequence eGFP from 3' end and also used for allele-specific PCR.</td>
</tr>
<tr>
<td>49</td>
<td>Y67_eGFP_forward</td>
<td>cctgctgacacccgctgaa ctgacccctgaa ctgacccctgaa</td>
<td>Allelic specific PCR Y67.</td>
</tr>
<tr>
<td>50</td>
<td>W67_eGFP_forward</td>
<td>tctgtgacacccgctgaa ctgacccctgaa</td>
<td>Allelic specific PCR W67.</td>
</tr>
<tr>
<td>51</td>
<td>Y67_eGFP_reverse</td>
<td>ctagaagcgactagcacgg</td>
<td>Allelic specific PCR Y67.</td>
</tr>
<tr>
<td>52</td>
<td>W67_eGFP_reverse</td>
<td>ctagaagcgactagcacgg</td>
<td>Allelic specific PCR W67.</td>
</tr>
<tr>
<td>53</td>
<td>Y67/eGFP_reverse</td>
<td>ctagaagcgactagcacgg</td>
<td>Allelic specific PCR Y67.</td>
</tr>
<tr>
<td>54</td>
<td>Stop67/eGFP_forwa rd</td>
<td>cctgctgacacccgctgaa ctgacccctgaa</td>
<td>Allelic specific PCR STOP67.</td>
</tr>
<tr>
<td>55</td>
<td>Stop67/eGFP_revers e</td>
<td>ctagaagcgactagcacgg</td>
<td>Allelic specific PCR STOP67.</td>
</tr>
<tr>
<td>56</td>
<td>Y66W_lagging</td>
<td>ttctagctggacacccgctgaa ctgacccctgaa ctgacccctgaa</td>
<td>72 nt, sense oligo to change GFPmut3* Y66V (Violeta) by Recombineering.</td>
</tr>
<tr>
<td>57</td>
<td>G65T/Y66W_lagging</td>
<td>ttctagctggacacccgctgaa ctgacccctgaa ctgacccctgaa</td>
<td>76 nt, sense oligo to change GFPmut3* G65Y to T67Y (Mar) by Recombineering.</td>
</tr>
<tr>
<td>58</td>
<td>GFPmut3* Y39stop1_lagging</td>
<td>ttctagctggacacccgctgaa ctgacccctgaa ctgacccctgaa</td>
<td>71 nt, sense oligo to change GFPmut3* Y39Stop (Dark) by Recombineering.</td>
</tr>
<tr>
<td>59</td>
<td>W67_eGFP</td>
<td>gctgctgacacccgctgaa ctgacccctgaa ctgacccctgaa</td>
<td>To change eGFP to CFP.</td>
</tr>
<tr>
<td>60</td>
<td>Stop67_eGFP</td>
<td>gctgctgacacccgctgaa ctgacccctgaa ctgacccctgaa</td>
<td>To create a nonsense eGFP mutant.</td>
</tr>
<tr>
<td>61</td>
<td>T7_Fw</td>
<td>agatctgatccgccggtaaat actgc</td>
<td>To amplify from the pET28a MCS (BglII, T7 promoter, Lac operator, Shine-Dalgrano, His6, T7 tag).</td>
</tr>
<tr>
<td>62</td>
<td>T7_Rv</td>
<td>ctcgagcttcgctgacccgctgaa ctgacccctgaa ctgacccctgaa</td>
<td>To amplify from the pET28a MCS (BglII, T7 promoter, Lac operator, Shine-Dalgrano, His6, T7 tag).</td>
</tr>
<tr>
<td>63</td>
<td>W67_eGFP</td>
<td>gctgctgacacccgctgaa ctgacccctgaa ctgacccctgaa</td>
<td>To change eGFP to CFP by site-directed mutagenesis.</td>
</tr>
<tr>
<td>64</td>
<td>W67_eGFP_R</td>
<td>gctgctgacacccgctgaa ctgacccctgaa ctgacccctgaa</td>
<td>To change eGFP to CFP by site-directed mutagenesis.</td>
</tr>
<tr>
<td>65</td>
<td>Stop67_eGFP</td>
<td>gctgctgacacccgctgaa ctgacccctgaa ctgacccctgaa</td>
<td>To create a nonsense eGFP mutant by site-directed mutagenesis.</td>
</tr>
<tr>
<td>66</td>
<td>Stop67_eGFP_R</td>
<td>gctgctgacacccgctgaa ctgacccctgaa ctgacccctgaa</td>
<td>To create a nonsense eGFP mutant by site-directed mutagenesis.</td>
</tr>
<tr>
<td>67</td>
<td>Y67_eGFP</td>
<td>gctgctgacacccgctgaa ctgacccctgaa ctgacccctgaa</td>
<td>To change to eGFP by Recombineering.</td>
</tr>
<tr>
<td>68</td>
<td>H67_eGFP</td>
<td>gctgctgacacccgctgaa ctgacccctgaa ctgacccctgaa</td>
<td>To change eGFP to BFP by site-directed mutagenesis.</td>
</tr>
<tr>
<td>69</td>
<td>H67_eGFP_R</td>
<td>gctgctgacacccgctgaa ctgacccctgaa ctgacccctgaa</td>
<td>To change eGFP to BFP by site-directed mutagenesis.</td>
</tr>
<tr>
<td>70</td>
<td>Y204_eGFP</td>
<td>gctgctgacacccgctgaa ctgacccctgaa ctgacccctgaa</td>
<td>To change eGFP to YFP by site-directed mutagenesis.</td>
</tr>
<tr>
<td>71</td>
<td>Y204_eGFP</td>
<td>gctgctgacacccgctgaa ctgacccctgaa ctgacccctgaa</td>
<td>To change eGFP to YFP by site-directed mutagenesis.</td>
</tr>
<tr>
<td>72</td>
<td>3Y67/eGFP_F</td>
<td>cctgctgacacccgctgaa</td>
<td>Allele-specific PCR to detect &quot;Green&quot;.</td>
</tr>
<tr>
<td>73</td>
<td>3H67/eGFP_F</td>
<td>cctgctgacacccgctgaa</td>
<td>Allele-specific PCR to detect &quot;Blue&quot;.</td>
</tr>
<tr>
<td>74</td>
<td>3T204/eGFP_R</td>
<td>ttgtctgacccgctgaa</td>
<td>Allele-specific PCR to detect &quot;Green&quot;.</td>
</tr>
<tr>
<td>75</td>
<td>3Y204/eGFP_R</td>
<td>ttgtctgacccgctgaa</td>
<td>Allele-specific PCR to detect &quot;Mostaza&quot;.</td>
</tr>
<tr>
<td>76</td>
<td>dsRed_Fw</td>
<td>gactgctgacacccgctgaa ctgacccctgaa ctgacccctgaa</td>
<td>To sequence dsRed.</td>
</tr>
<tr>
<td>77</td>
<td>dsRed_Rv</td>
<td>gactgctgacacccgctgaa ctgacccctgaa ctgacccctgaa</td>
<td>To sequence dsRed.</td>
</tr>
<tr>
<td>78</td>
<td>Y203T-4s/35</td>
<td>gacagacggcggagggagatgtc</td>
<td>For Recombineering 203T (Mostaza to green).</td>
</tr>
<tr>
<td>79</td>
<td>Y203T-4s/45</td>
<td>gacagacggcggagggagatgtc</td>
<td>For Recombineering 203T (Mostaza to green).</td>
</tr>
<tr>
<td>80</td>
<td>Y203T-4s/45</td>
<td>gacagacggcggagggagatgtc</td>
<td>For Recombineering 203T (Mostaza to green).</td>
</tr>
<tr>
<td>81</td>
<td>Y203T-4s/45</td>
<td>gacagacggcggagggagatgtc</td>
<td>For Recombineering 203T (Mostaza to green).</td>
</tr>
<tr>
<td>82</td>
<td>Y203T-4s/55</td>
<td>gacagacggcggagggagatgtc</td>
<td>For Recombineering 203T (Mostaza to green).</td>
</tr>
<tr>
<td>83</td>
<td>Y203T-4s/65</td>
<td>gacagacggcggagggagatgtc</td>
<td>For Recombineering 203T (Mostaza to green).</td>
</tr>
<tr>
<td>84</td>
<td>Y203T-4s/65</td>
<td>gacagacggcggagggagatgtc</td>
<td>For Recombineering 203T (Mostaza to green).</td>
</tr>
<tr>
<td>85</td>
<td>Y203T-4s/65</td>
<td>gacagacggcggagggagatgtc</td>
<td>For Recombineering 203T (Mostaza to green).</td>
</tr>
<tr>
<td>86</td>
<td>TR-Y203T-4s/45</td>
<td>gacagacggcggagggagatgtc</td>
<td>Sense strand targeting oligo for YFP-GFP. For looking at colocalization with ICP8 and off-target effects.</td>
</tr>
<tr>
<td>87</td>
<td>UL29_upst</td>
<td>gtaacgctgacacccgctgaa ctgacccctgaa ctgacccctgaa</td>
<td>To sequence upstream region (5' of Knipe plasmids (pCMV-ICP8 and pCMV-ICP8-GFP).</td>
</tr>
<tr>
<td>88</td>
<td>UL29_downst</td>
<td>gtaacgctgacacccgctgaa ctgacccctgaa ctgacccctgaa</td>
<td>To sequence downstream region (3' of Knipe plasmid (pCMV-ICP8).</td>
</tr>
</tbody>
</table>
To amplify Crimson from pTEC19 with primer 150

To amplify Crimson from product of primers (145 and 146) with primer 148 to produce Product N

To amplify Crimson from product of primers (145 and 146) with primer 147 to produce Product N

To amplify Crimson from pTEC19 with primer 146

5' end phosphorylated.

To amplify Crimson from pTEC19 with primer 146. 5' end phosphorylated.

To amplify Crimson from product of primers (145 and 146) with primer 148 to produce Product N. 5' end phosphorylated.

To amplify Crimson from product of primers (145 and 146) with primer 147 to produce Product N. 5' end phosphorylated.

To amplify Crimson from product N with primer 145 to produce Product N-Control. 5' end phosphorylated.

To amplify Crimson from pTEC19 with primer 151.

To amplify Crimson from pTEC19 with primer 150. 5' end phosphorylated.
58
#
152

Oligo name
P2A/Ccrim-F2

Oligo sequence (5’ to 3’)
ccggatccggagccacgaacttctctctgttaaagcaagca
ggagacg

153

P2A/Ccrim-F3

154

P2A/Ccrim-R2

cagatctaggaggtccgccaccatggccggatccggagc
cacg
tgcggccgcgaattcctactggaacaggtggtg

163

EcoRV-UL12
forward
EcoRV-UL12.5
forward
NotI-UL12 reverse

164
165

cattagatatcgccatggagtccacggtaggcccagca
cattagatatcgccgccaccatgtggtcggcgtcggtga

174

Mito-Oligo-Crick
apaI HindIII

109
110
183

R-X1
T7-min
CAT->KAN 5'

184

CAT->KAN 3'

190

pSLIK-att-f

caattaatgcggccgctcagcgagacgacctccccgtcgt
cggt
caattaatgcggccgctcaagcgtaatcagggacgtcata
ggggtagcgagacgacctccccgt
cgccaccatgtccgtcctgacgccgctgctgctgcggggc
ttgacaggctcggcccggcggctcccagtgccgcgcgcc
aagatccattcgttga
agcttcaacgaatggatcttggcgcgcggcactgggagcc
gccgggccgagcctgtcaagccccgcagcagcagcggc
gtcaggacggacatggtggcgggcc
ggtatataatgggggccgtca
cgtaatacgactcactataggcctt
atggagaaaaaaatcactggatataccaccgttgatatatcc
caagtgtaggctggagctgcttc
ttacgccccgccctgccactcatcgcagtactgttgtaattc
attcatatgaatatcctccttag
cagggacagcagagatccag

191

pSLIK-att-r

gccagatcttgggtgggttaat

192

Crimson5-r1

gtgccctcgtagggcttg

193

Crimson3-r2

tggaacaggtggtggcgg

194

Crimson5-f1

cactgagaacgtcatcaagcc

195
196
197
198
199

HumBeta-r1
HumBeta3-r2
HumBeta5-f1
HumBeta3-f2
3ST204newgreen_R

ggttcctgtccttcctgtagat
aaacggcggcgaccttc
atggtgcctcccaagaagaa
gctgcacatgcaggatctac
cagggcggactgggtgga

200

3SY204yellow_R

cagggcggactggtagct

201

T203_as/65

203

7-ICP8-Flexi7

tcgttggggtctttgctcagggcggactgggtgctcaggta
gtggttgtcgggcagcagcacggg
aggagcgatcgccatggagacaaagcccaagacg

204

8-ICP8-Flexi8

gtcggtttaaaccagcatatccaacgtcaggtctc

205

9-ICP8-GFP-Flexi9

gtcggtttaaaccttgtacagctcgtccatgcc

206

beta-flex1

aggagcgatcgccatgagtactgcactcgcaacg

207

beta-flex2

gtcggtttaaactgctgccaccttctgctct

208

beta(1-177)-flexi3

gtcggtttaaacgcgctcggcttcatccttgt

209

exo-flexi4

aggagcgatcgccatgacaccggacattatcctg

210

exoflexi-5

gtcggtttaaactcgccattgctccccaaat

211

exo(1-189)-flexi5

gtcggtttaaacaatcacgacataatgcaggc

212

FITC-CHI

gctggtggcgtggtgggctggtgggctggtgg

166
173

NotI-HA-UL12
reverse
Mito-Oligo-Watson

Oligo function/description
To amplify Crimson from product of primers (150 and
151) with primer 151 to produce Product C. 5’ end
phosphorylated.
To amplify Crimson from product C with primer 154
to produce Product C-Control. 5’ end phosphorylated.
To amplify Crimson from product C with primer 153
to produce Product C-Control. 5’ end phosphorylated.
To amplify UL12 by PCR and clone in pNEBRX1Hygro (EcoRV and NotI)
To amplify UL12.5 by PCR and clone in pNEBRX1Hygro (EcoRV and NotI)
To amplify UL12 and UL12.5 by PCR and clone in
pNEBRX1-Hygro (EcoRV and NotI)
To amplify UL12-HA and UL12.5-HA by PCR and
clone in pNEBRX1-Hygro (EcoRV and NotI)
Mitochondrial Localization Signal, Watson, to add to
genes to import into mitochondria
Mitochondrial Localization Signal, Crick, to add to
genes to import into mitochondria
To amplify insert in the pNEBRX1-Hygro plasmid
To amplify insert in the pNEBRX1-Hygro plasmid
To amplify Kan cassette by PCR. PCR product used to
delete Cml and replace with Kan
To amplify Kan cassette by PCR. PCR product used to
delete Cml and replace with Kan
Sequencing of pSLIK, upstream of att site, directed
toward inserts.
Sequencing of pSLIK, downstream of att site, directed
toward inserts.
Sequencing of pSLIK-Crimson, 5' end of Crimson,
directed out.
Sequencing of pSLIK-Crimson, 3' end of Crimson,
directed toward its 5' end.
Sequencing of pSLIK-Crimson, 5' end of Crimson,
directed toward its 3' end.
Sequencing of HumBeta, directed out to 5'.
Sequencing of HumBeta, 3' end, directed toward its 5’.
Sequencing of HumBeta, 5' end, toward its 3'.
Sequencing of HumBeta, 3' end directed out.
Allele-specific PCR to amplify green recombinants
using oligo 85. PCR Rx with oligo # 45.
Allele-specific PCR to amplify Mostaza non
recombinants. PCR Rx with oligo # 45.
For Recombineering 203T Mostaza to eGFP (2 nt
bubble with a C.C). PAGE purified.
Forward primer to amplify ICP8 and ICP8-GFP from
pCMV-ICP8 and pCMV-ICP8-GFP and ligate to
pFN22K
Reverse primer to amplify ICP8 from pCMV-ICP8 and
ligate to pFN22K
Reverse primer to amplify ICP8-GFP from pCMVICP8-GFP and ligate to pFN22K
To PCR amplify the Beta gene and clone it in the
Promega Flexi System
To PCR amplify the Beta gene and clone it in the
Promega Flexi System
To PCR amplify the Beta gene and clone it in the
Promega Flexi System
To PCR amplify the Exo gene and clone it in the
Promega Flexi System
To PCR amplify the Exo gene and clone it in the
Promega Flexi System
To PCR amplify the Exo gene and clone it in the
Promega Flexi System
FITC labeled to evaluate oligo transfection


### Table 2.6 Plasmids for studies in human cells

<table>
<thead>
<tr>
<th>Plasmid #</th>
<th>Name</th>
<th>Reference/source</th>
<th>Strain</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>RSM57</td>
<td>pNL-eGFP/CEF</td>
<td>Gift of Jakob Reiser (USDA)</td>
<td>RIK450</td>
<td>Amp&lt;sup&gt;®&lt;/sup&gt;. (Reiser et al. 2000) Used to transfect and transduce mammalian cells and to create the Recombineering reporters</td>
</tr>
<tr>
<td>RSM53</td>
<td>pDual-eGFP</td>
<td>This study</td>
<td>RIK498</td>
<td>Amp&lt;sup&gt;®&lt;/sup&gt;. Mammalian Recombineering reporter</td>
</tr>
<tr>
<td>RSM54</td>
<td>pDual-eGFP(T204Y)</td>
<td>This study</td>
<td>RIK497</td>
<td>Amp&lt;sup&gt;®&lt;/sup&gt;. Mammalian Recombineering reporter</td>
</tr>
<tr>
<td>RSM55</td>
<td>pCMV-VSV-G</td>
<td>Gift of Priya Rai (UM)</td>
<td>RIK494</td>
<td>Amp&lt;sup&gt;®&lt;/sup&gt;. Mammalian Recombineering reporter</td>
</tr>
<tr>
<td>RSM103</td>
<td>pCMV-ICP8 (pCMV8.3)</td>
<td>Gift of Dr. David Knipe (Harvard)</td>
<td>RIK496</td>
<td>Amp&lt;sup&gt;®&lt;/sup&gt;. Mammalian Recombineering reporter</td>
</tr>
<tr>
<td>RSM104</td>
<td>pCMV-ICP8-GFP (pCMV-ICP8GFP)</td>
<td>Gift of Dr. David Knipe (Harvard)</td>
<td>RIK495</td>
<td>Amp&lt;sup&gt;®&lt;/sup&gt;. Mammalian Recombineering reporter</td>
</tr>
<tr>
<td>RSM274</td>
<td>pCMV</td>
<td>This study</td>
<td>RIK495</td>
<td>Amp&lt;sup&gt;®&lt;/sup&gt;. Mammalian Recombineering reporter</td>
</tr>
<tr>
<td>RSM49</td>
<td>pmaxGFP</td>
<td>Lonza, GI:33243028</td>
<td>RIK443</td>
<td>Amp&lt;sup&gt;®&lt;/sup&gt;. Packaging plasmid for Lentiviral constructs (An et al. 1999).</td>
</tr>
<tr>
<td>RSM177</td>
<td>pTEC19</td>
<td>Addgene</td>
<td>RIK443</td>
<td>Amp&lt;sup&gt;®&lt;/sup&gt;. Packaging plasmid for Lentiviral constructs (An et al. 1999).</td>
</tr>
<tr>
<td>RSM196</td>
<td>pUC57-Kan::NLS/Crimson-HumBeta</td>
<td>Genewiz</td>
<td>RIK581</td>
<td>Amp&lt;sup&gt;®&lt;/sup&gt;. Sequence design, this study. Synthesis by Genewiz</td>
</tr>
<tr>
<td>RSM195</td>
<td>pUC57-Kan::NLS/HA-HumBeta</td>
<td>Genewiz</td>
<td>RIK580</td>
<td>Amp&lt;sup&gt;®&lt;/sup&gt;. Sequence design, this study. Synthesis by Genewiz</td>
</tr>
<tr>
<td>RSM224</td>
<td>pUC19::NLS/HA-HumBeta</td>
<td>This study</td>
<td>RIK582</td>
<td>Amp&lt;sup&gt;®&lt;/sup&gt;.</td>
</tr>
<tr>
<td>RSM13</td>
<td>pFN22K::ICP8</td>
<td>This study</td>
<td>RIK499</td>
<td>Kan&lt;sup&gt;®&lt;/sup&gt;. Expression of ICP8 in E. coli. ICP8 purification for PTD transfection and chromatin remodeling experiments. HHV1 Recombineering in E. coli.</td>
</tr>
<tr>
<td>RSM16</td>
<td>pFN22K::ICP8-GFP</td>
<td>This study</td>
<td>RIK582</td>
<td>Kan&lt;sup&gt;®&lt;/sup&gt;. Expression of ICP8-GFP in E. coli. ICP8-GFP purification for PTD transfection and chromatin remodeling experiments. HHV1 Recombineering in E. coli.</td>
</tr>
<tr>
<td>Plasmid #</td>
<td>Name</td>
<td>Reference/ source</td>
<td>Strain</td>
<td>Comments</td>
</tr>
<tr>
<td>----------</td>
<td>-----------------------------</td>
<td>-------------------</td>
<td>--------</td>
<td>--------------------------------------------------------------------------</td>
</tr>
<tr>
<td>RSM255</td>
<td>pFN24K::Crimson/P2A</td>
<td>This study</td>
<td>RIK588</td>
<td>Kan&lt;sup&gt;a&lt;/sup&gt;. Expresses Crimson from T7 promoter verified by flow cytometry and Fluorescence microscopy</td>
</tr>
<tr>
<td>RSM307</td>
<td>pFN22K::Crimson/P2A-ICP8</td>
<td>This study</td>
<td>RIK608</td>
<td>Kan&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>RSM348</td>
<td>pENTR2B/TREPitt</td>
<td>Gift from Jakob Reiser at NIH</td>
<td>RIK614</td>
<td>Kan&lt;sup&gt;a&lt;/sup&gt;. Tet responsive promoter driving expression of transgene, positive selection for insert and Gateway donor (Ou et al. 2012). Plasmid needs to be grown in E. coli DB3.1 strain</td>
</tr>
<tr>
<td>RSM376</td>
<td>pENTR2B/TREPitt::Crimson/P2A</td>
<td>This study</td>
<td>RIK631</td>
<td>Kan&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>RSM222</td>
<td>pSLIK-zoeclin</td>
<td>Addgene (108480)</td>
<td>RIK579</td>
<td>Amp&lt;sup&gt;b&lt;/sup&gt;, Zeo&lt;sup&gt;b&lt;/sup&gt;. Gateway acceptor lentiviral vector with Tet-ON transactivators. Use to capture TREPitt driven constructs for Tet inducible expression in mammalian cells (Shin et al. 2006).</td>
</tr>
<tr>
<td>RSM479</td>
<td>pSLIK2B/TREPitt::Crimson/P2A</td>
<td>This study</td>
<td>RIK711</td>
<td>Lentiviral vector plasmid used to create lentiviral particles expressing Crimson and ICP8 in mammalian cells.</td>
</tr>
<tr>
<td>RSM283</td>
<td>pUC19::Crimson/P2A-NLS/HA-HumBeta</td>
<td>This study</td>
<td>RIK607</td>
<td>Kan&lt;sup&gt;a&lt;/sup&gt;. Expresses Crimson from Lac UV5 promoter-mediated transcription in pUC19. Verified by flow cytometry and Fluorescence microscopy. Cells are blue under lab room lighting.</td>
</tr>
<tr>
<td>RSM392</td>
<td>pENTR2B/TREPitt::Crimson/P2A</td>
<td>This study</td>
<td>RIK647</td>
<td>Kan&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>RSM480</td>
<td>pSLIK/TREPitt::Crimson/P2A-NLS/HA-HumBeta</td>
<td>This study</td>
<td>RIK715</td>
<td>Amp&lt;sup&gt;b&lt;/sup&gt;, Zeo&lt;sup&gt;b&lt;/sup&gt;. Lentiviral vector plasmid used to create lentiviral particles expressing Crimson and NLS/HA-HumBeta in mammalian cells.</td>
</tr>
<tr>
<td>RSM262</td>
<td>pFN24K::Crimson/P2A-Control</td>
<td>This study</td>
<td>RIK594</td>
<td>Kan&lt;sup&gt;a&lt;/sup&gt;. Expresses Crimson from T7 promoter verified by flow cytometry and Fluorescence microscopy</td>
</tr>
<tr>
<td>RSM400</td>
<td>pENTR2B/TREPitt::Crimson/P2A-Control</td>
<td>This study</td>
<td>RIK655</td>
<td>Kan&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>RSM481</td>
<td>pSLIK/TREPitt::Crimson/P2A</td>
<td>This study</td>
<td>RIK719</td>
<td>Amp&lt;sup&gt;b&lt;/sup&gt;, Zeo&lt;sup&gt;b&lt;/sup&gt;. Lentiviral vector plasmid used to create lentiviral particles expressing Crimson/P2A as the non-recombinase control in mammalian cells.</td>
</tr>
<tr>
<td>RSM326</td>
<td>pFN22K::ICP8-P2A/Crimson</td>
<td>This study</td>
<td>RIK623</td>
<td>Kan&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>RSM408</td>
<td>pENTR2B/TREPitt::ICP8-P2A/Crimson</td>
<td>This study</td>
<td>RIK663</td>
<td>Kan&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>RSM482</td>
<td>pSLIK/TREPitt::ICP8-P2A/Crimson</td>
<td>This study</td>
<td>RIK723</td>
<td>Amp&lt;sup&gt;b&lt;/sup&gt;, Zeo&lt;sup&gt;b&lt;/sup&gt;. Lentiviral vector plasmid used to create lentiviral particles expressing ICP8 and Crimson in mammalian cells.</td>
</tr>
<tr>
<td>RSM266</td>
<td>pUC19-Kan::NLS/HA-HumBeta-P2A/Crimson</td>
<td>This study</td>
<td>RIK601</td>
<td>Amp&lt;sup&gt;b&lt;/sup&gt;. Expresses Crimson from Lac UV5 promoter-mediated transcription in pUC19. Verified by flow cytometry and Fluorescence microscopy</td>
</tr>
<tr>
<td>RSM409</td>
<td>pENTR2B/TREPitt::NLS/HA-HumBeta-P2A/Crimson</td>
<td>This study</td>
<td>RIK664</td>
<td>Kan&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>RSM483</td>
<td>pSLIK/TREPitt::NLS/HA-HumBeta-P2A/Crimson</td>
<td>This study</td>
<td>RIK728</td>
<td>Amp&lt;sup&gt;b&lt;/sup&gt;, Zeo&lt;sup&gt;b&lt;/sup&gt;. Lentiviral vector plasmid used to create lentiviral particles expressing NLS/HA-HumBeta and Crimson in mammalian cells.</td>
</tr>
<tr>
<td>RSM363</td>
<td>pFN24K::P2A/Crimson-control</td>
<td>This study</td>
<td>RIK622</td>
<td>Kan&lt;sup&gt;a&lt;/sup&gt;. Expresses Crimson from T7 promoter verified by Fluorescence microscopy</td>
</tr>
<tr>
<td>RSM416</td>
<td>pENTR2B/TREPitt::P2A/Crimson-control</td>
<td>This study</td>
<td>RIK671</td>
<td>Kan&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
</tbody>
</table>
### Plasmids for studies in human cells

<table>
<thead>
<tr>
<th>Plasmid #</th>
<th>Name</th>
<th>Reference/ source</th>
<th>Strain</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>RSM484</td>
<td>pSLIK/TREPitt::P2A/Crimson</td>
<td>This study</td>
<td>RIK731</td>
<td>Amp&lt;sup&gt;+&lt;/sup&gt;, Zeo&lt;sup&gt;+&lt;/sup&gt;. Lentiviral vector plasmid used to create lentiviral particles expressing P2A/Crimson as the non-recombinase control in mammalian cells.</td>
</tr>
<tr>
<td>RSM253</td>
<td>pNEBR-X1Hygro</td>
<td>NEB, E3000, N8083</td>
<td>RIK452</td>
<td>Amp&lt;sup&gt;+&lt;/sup&gt;, Hygro&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>RSM150</td>
<td>pNEBR-R1</td>
<td>NEB, E3000</td>
<td>RIK373</td>
<td>Amp&lt;sup&gt;+&lt;/sup&gt;, Neo&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>RSM152</td>
<td>pNEBR-X1Glu</td>
<td>NEB, E3000</td>
<td>RIK374</td>
<td>Amp&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>RSM118</td>
<td>pNEBRX1-UL12.5</td>
<td>RIK372</td>
<td>Amp&lt;sup&gt;+&lt;/sup&gt;, Hygro&lt;sup&gt;+&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>RSM107</td>
<td>pNEBRX1-UL12.5-HA</td>
<td>RIK368</td>
<td>Amp&lt;sup&gt;+&lt;/sup&gt;, Hygro&lt;sup&gt;+&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>RSM113</td>
<td>pNEBRX1-MLS</td>
<td>RIK370</td>
<td>Amp&lt;sup&gt;+&lt;/sup&gt;, Hygro&lt;sup&gt;+&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>RSM158</td>
<td>pNEBR-X1Hygro::ICP8-GFP</td>
<td>Mel302</td>
<td>Amp&lt;sup&gt;+&lt;/sup&gt;, Hygro&lt;sup&gt;+&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>RSM185</td>
<td>pNEBR-X1Hygro::ICP8</td>
<td>Mel301</td>
<td>Amp&lt;sup&gt;+&lt;/sup&gt;, Hygro&lt;sup&gt;+&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>RSM59</td>
<td>pCA24N::ruvA</td>
<td>ASKA(Kitagawa et al. 2005)</td>
<td>Mel253</td>
<td>Cml&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

### Table 2.7 Plasmids for studies in bacteria

Plasmids for studies in bacteria

<table>
<thead>
<tr>
<th>Plasmid #</th>
<th>Name</th>
<th>Reference/ source</th>
<th>Strain</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>RSM134</td>
<td>pUC19AKpnI</td>
<td>Lab</td>
<td>RIK184</td>
<td>Amp&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>RSM88</td>
<td>pFC15K::β</td>
<td>This study</td>
<td>RIK483</td>
<td>Kan&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>RSM89</td>
<td>pFC15K::β(1-177)</td>
<td>This study</td>
<td>RIK484</td>
<td>Kan&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>RSM90</td>
<td>pFC15K::Exo</td>
<td>This study</td>
<td>RIK485</td>
<td>Kan&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>RSM91</td>
<td>pFC15K::Exo(1-189)</td>
<td>This study</td>
<td>RIK486</td>
<td>Kan&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>RSM84</td>
<td>pFN22K::β</td>
<td>This study</td>
<td>RIK518</td>
<td>Kan&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>RSM85</td>
<td>pFN22K::β(1-177)</td>
<td>This study</td>
<td>RIK519</td>
<td>Kan&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>RSM86</td>
<td>pFN22K::Exo</td>
<td>This study</td>
<td>Mel277</td>
<td>Kan&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>RSM87</td>
<td>pFN22K::Exo(1-189)</td>
<td>This study</td>
<td>Mel278</td>
<td>Kan&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>RSM92</td>
<td>pET28a::his-reda(R28A)</td>
<td>Lab</td>
<td>RIK187</td>
<td>Kan&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>RSM93</td>
<td>pET28a::his-reda(D119A)</td>
<td>Lab</td>
<td>RIK188</td>
<td>Kan&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>RSM94</td>
<td>pET28a::his-redji</td>
<td>Lab</td>
<td>RIK189</td>
<td>Kan&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>RSM95</td>
<td>pET28a::his-reda</td>
<td>Lab</td>
<td>RIK254</td>
<td>Kan&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>RSM175</td>
<td>pLYSs</td>
<td>Lab</td>
<td>Mel311</td>
<td>Cml&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

Plasmids for hypermorphic allele in Recombineering and for pull down studies
<table>
<thead>
<tr>
<th>Plasmid #</th>
<th>Name</th>
<th>Reference/ source</th>
<th>Strain</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>RSM60</td>
<td>pCA24N::ruvB</td>
<td>ASKA</td>
<td>Mel254</td>
<td>Cml*</td>
</tr>
<tr>
<td>RSM61</td>
<td>pCA24N::ligA (lig)</td>
<td>ASKA</td>
<td>Mel255</td>
<td>Cml*</td>
</tr>
<tr>
<td>RSM62</td>
<td>pCA24N::dnaG</td>
<td>ASKA</td>
<td>Mel256</td>
<td>Cml*</td>
</tr>
<tr>
<td>RSM63</td>
<td>pCA24N::recA</td>
<td>ASKA</td>
<td>Mel257</td>
<td>Cml*</td>
</tr>
<tr>
<td>RSM64</td>
<td>pCA24N::recG</td>
<td>ASKA</td>
<td>Mel258</td>
<td>Cml*</td>
</tr>
<tr>
<td>RSM65</td>
<td>pCA24N::ssb</td>
<td>ASKA</td>
<td>Mel259</td>
<td>Cml*</td>
</tr>
<tr>
<td>RSM66</td>
<td>pCA24N::sbcB</td>
<td>ASKA</td>
<td>Mel260</td>
<td>Cml*</td>
</tr>
<tr>
<td>RSM67</td>
<td>pCA24N::priA</td>
<td>ASKA</td>
<td>Mel261</td>
<td>Cml*</td>
</tr>
<tr>
<td>RSM68</td>
<td>pCA24N::priB</td>
<td>ASKA</td>
<td>Mel262</td>
<td>Cml*</td>
</tr>
<tr>
<td>RSM69</td>
<td>pCA24N::priC</td>
<td>ASKA</td>
<td>Mel263</td>
<td>Cml*</td>
</tr>
<tr>
<td>RSM70</td>
<td>pCA24N::avrD</td>
<td>ASKA</td>
<td>Mel264</td>
<td>Cml*</td>
</tr>
<tr>
<td>RSM71</td>
<td>pCA24N::avrC</td>
<td>ASKA</td>
<td>Mel265</td>
<td>Cml*</td>
</tr>
<tr>
<td>RSM23</td>
<td>pCA24N::lon</td>
<td>ASKA</td>
<td>RIK506</td>
<td>Cml*</td>
</tr>
<tr>
<td>RSM24</td>
<td>pCA24N::clpX</td>
<td>ASKA</td>
<td>RIK507</td>
<td>Cml*</td>
</tr>
<tr>
<td>RSM25</td>
<td>pCA24N::clpP</td>
<td>ASKA</td>
<td>RIK508</td>
<td>Cml*</td>
</tr>
<tr>
<td>RSM26</td>
<td>pCA24N::mviN</td>
<td>ASKA</td>
<td>RIK509</td>
<td>Cml*</td>
</tr>
<tr>
<td>RSM27</td>
<td>pCA24N::ftsH</td>
<td>ASKA</td>
<td>RIK510</td>
<td>Cml*</td>
</tr>
<tr>
<td>RSM28</td>
<td>pCA24N::ftsI</td>
<td>ASKA</td>
<td>RIK511</td>
<td>Cml*</td>
</tr>
<tr>
<td>RSM29</td>
<td>pCA24N::yegS</td>
<td>ASKA</td>
<td>RIK512</td>
<td>Cml*</td>
</tr>
<tr>
<td>RSM30</td>
<td>pCA24N::dnaT</td>
<td>ASKA</td>
<td>RIK513</td>
<td>Cml*</td>
</tr>
<tr>
<td>RSM31</td>
<td>pCA24N::adsD</td>
<td>ASKA</td>
<td>RIK514</td>
<td>Cml*</td>
</tr>
<tr>
<td>RSM32</td>
<td>pCA24N::ompF</td>
<td>ASKA</td>
<td>RIK515</td>
<td>Cml*</td>
</tr>
<tr>
<td>RSM33</td>
<td>pCA24N::gmbB</td>
<td>ASKA</td>
<td>RIK516</td>
<td>Cml*</td>
</tr>
<tr>
<td>RSM34</td>
<td>pCA24N::rowC</td>
<td>ASKA</td>
<td>RIK517</td>
<td>Cml*</td>
</tr>
<tr>
<td>pKD4</td>
<td>Gift of Barry Wanner (Purdue)</td>
<td>RIK212</td>
<td></td>
<td>Amp*, Kan*</td>
</tr>
<tr>
<td>pSIM5</td>
<td>Gift of Don Court (NCI)</td>
<td>RIK343</td>
<td></td>
<td>Cml*</td>
</tr>
<tr>
<td>pSIM6</td>
<td>Gift of Don Court (NCI)</td>
<td>RIK344</td>
<td></td>
<td>Amp*</td>
</tr>
</tbody>
</table>
Plasmid construction method

Plasmid sequences and maps are presented in the APPENDIX A. PLASMID SEQUENCES AND MAPS.

Plasmids for lentiviral Recombineering targets (pDual-eGFP variants)

pDual-eGFP

A PCR product amplified from pET28a was ligated to the SmaI digested pNL-eGFP/CEF (Zhang et al. 2002; Zhang et al. 2004; Reiser et al. 2000) and screened for gain of green fluorescence using a Dark Reader light box. The PCR product starts with a BglII sequence and then amplifies the T7 promoter, the LacI binding site (Lac operator), a Shine-Dalgarno sequence, His6, T7 tag, and ends at the beginning of the pET28a MCS. The PCR product was produced with 5’ phosphorylated primers (61 and 62). eGFP was strongly expressed from the Rosetta-gami™2(DE3) strain, which bears the T7 RNA polymerase gene. If desired, the insert may be excised again by cleavage with BglII and BamHI and religating. NcoI/XhoI excises eGFP and produces a product that can by ligated to pET28a (NcoI/XhoI) to create a T7 expressed eGFP with a C-terminal His tag for purification. The portion of the pNL-eGFP/CEF construct germane to these studies was verified by sequencing using oligo 44 as a primer.

pDual-eGFP(W66)

pDual-eGFP was modified by Recombineering with oligo 59 in strain RIK473 and screened for loss of green fluorescence using a Dark Reader light box (the eGFP W66 variant has a fluorescence emission spectrum that is below the wavelength cut-on the Dark Reader filter). Mutagenesis was confirmed by allele-specific PCR (Y66 with oligos 49 and 46 and W66 with oligos 50 and 46) and sequencing using oligos 45 and 46.
pDual-eGFP(Stop<sub>66</sub>)

pDual-eGFP was modified by Recombineering with oligo 60 in strain RIK473 and screened for loss of green fluorescence using a Dark Reader light box. Mutagenesis was confirmed by allele-specific PCR (Y<sub>66</sub> with oligos 49 and 46 and Stop<sub>66</sub> with oligos 55 and 46) and sequencing using oligos 45 and 46.

pDual-eGFP(H<sub>66</sub>)

pDual-eGFP was modified by site-directed mutagenesis with oligos 68 and 69 and the QuikChange Lightning Kit (Agilent Technologies) and screened for loss of green fluorescence using a Dark Reader light box (the eGFP H<sub>66</sub> variant has a fluorescence emission spectrum that is below the wavelength cut-on the Dark Reader filter). Mutagenesis was confirmed by allele-specific PCR (Y<sub>66</sub> with oligos 49 and 46 and H<sub>66</sub> with oligos 73 and 46) and sequencing using oligos 45 and 46.

pDual-eGFP(Y<sub>203</sub>)

pDual-eGFP was modified by site-directed mutagenesis with oligos 70 and 71 and the QuikChange Lightning Kit (Agilent Technologies). Mutagenesis was confirmed by allele-specific PCR (T<sub>203</sub> with oligos 45 and 74 and Y<sub>203</sub> with oligos 45 and 75) and sequencing using oligos 45 and 46.

Plasmids for expressing viral recombinases in mammalian cells

Plasmids for expressing viral recombinases in mammalian cells by transient transfection

*PCMV empty vector control for pCMV-ICP8*

pCMV-ICP8 (Taylor & Knipe 2003) was digested with SmaI, which cut the ICP8 insert at several sites. The 4 Kb fragment of vector backbone lacking most of the ICP8
gene was purified by agarose gel extraction and circularized by ligation. The empty vector was validated by restriction digestion and sequenced.

Plasmids for expressing viral recombinases in stable cell lines

\textit{pFN22K::ICP8}

ICP8 was amplified from pCMV-ICP8 (Taylor & Knipe 2003) using a primer that introduced AsiSI at the 5' end of ICP8 to put it in frame with the Halo tag and with a second primer containing PmeI engineered at the 3' end of ICP8 which adds one more codon before the TAA stop (primers: ICP8-Flexi7 and ICP8-Flexi8). The product was digested with a mix of AsiSI and PmeI and ligated to pFN22K also cut with AsiSI and PmeI. The ligation reaction removes the toxic barnase gene. Kan resistant clones were selected and screened for inserts.

\textit{pFN22K::ICP8-GFP}

ICP8-GFP was amplified from pCMV-ICP8-GFP (Taylor & Knipe 2003) using a primer that introduced AsiSI at the 5' end of ICP8 to put it in frame with the Halo tag and with a second primer containing PmeI engineered at the 3' end of ICP8-GFP which adds one more codon before the TAA stop (primers: ICP8-Flexi7 and ICP8-GFP-Flexi9). The product was digested with a mix of AsiSI and PmeI and 100 ng were ligated to 50 ng pFN22K also cut with AsiSI and PmeI. The ligation removes the toxic barnase gene. Kan resistant clones were selected and screened for inserts.
**Humanized Beta genes**

**pUC57-Kan::NLS/HA-HumBeta**

The sequence of the enterobacteria phage $\lambda$ Beta gene ($bet$) was obtained from NCBI gi: 9626243. To that sequence was added the Nuclear Localization Signal (NLS) sequence from the SV40 large T antigen [here](http://www.uniprot.org/uniprot/P03070), an Influenza virus hemagglutinin epitope (HA) tag [here](http://en.wikipedia.org/wiki/HA-tag) and then a linker (GGGGGSGGGGSGGGS) to reduce steric hinderence with Beta protein. The resulting fusion protein sequence is:

```
MVPPKKKRKVDPKYPYDVPYAGGGGSGGGGSGGGSMSSTALATLAGK
LAERVGMDSVPQELITTLRQTAFAKGDASDAQFIALIVANQYGLNPWTKEIYAF
PDKQNGIVPVVGVDGWRIINENQQFDGMDFEQDNESCTCRIYRKDRNHPCVTE
WMDECRRPEFKTREGREITGPWQSHPKRMLRHKAMIQCARLAFGAGIYDKE
AERVENTAYTAERQPERDITPVNDETMQEINTLLIALDKTWDDDLPLCSQIFRR
DIRASSELITQAEAVKALGFLKQKAAEQKVAAV
```

Restriction sites were added to the end of the sequence to facilitate cloning. This sequence was submitted to Genewiz where they used a software to optimize the gene sequence to match the codon usage pattern of human cells. Genewiz synthesized the DNA and ligated it into plasmid pUC57-Kan.

**pUC57-Kan::NLS/Crimson-HumBeta**

The sequence of the enterobacteria phage $\lambda$ Beta gene ($bet$) was obtained from NCBI gi: 9626243. To that sequence was added the NLS sequence from the SV40 large T antigen [here](http://www.uniprot.org/uniprot/P03070), the E2/Crimson far red fluorescence protein sequence [here](http://www.addgene.org/browse/sequence/13866/) and then a linker...
(GGGGSGGGGGSGGGS) to reduce steric hinderence with Beta protein. The resulting
fusion protein sequence is:

MVPPKKKRKVQEDPKMDSTENVIKPFMRKHKMEGSVNGHEFEIEGVGEGKP
YEGTQTKLQVTGGLPFAWDSLQFYYGSKAYIKHPADIPDYLKQSFPEGFK
WERVMNFEDGGVVTVTQDSSLQDGTLYHVKFIGNFSPGPMQKKTGLGVEP
STERNYPRDGVLKGENHMALKLKGGGHLCEFKSIYMAKPKVLPGYHYVDY
KLDITSHNEDYTVEQYERAEARHHLFQGGGGSGGGGSGGGGSMSTALATLA
GKLAEVRGMDVDQPELITTLRQFTKGDSAQFIALLIVANQYGILNPWTKEIIY
AFPDQKQNGIPVVGVDWSRIINENQQFQGMDFEQDNESCTCRIYRKDRNHPI
VTEWMDECRREPFTREGTIGPWQSHPKRLRHKAMIQCARLAFGIAIYD
KDEAERIVENTAYTAERQPERDITPVNDTEMQEINTLIALDKTWDJDLPLCSQI
FRRDIRASELTQAEAVKALGFLKQAAEQKVAAB

Restriction sites were added to the end of the sequence to facilitate cloning. This
sequence was submitted to Genewiz where they used software to optimize the gene
sequence to match the codon usage pattern of human cells. Genewiz synthesized the
DNA and ligated it into plasmid pUC57-Kan.

\textit{pUC19::NLS/HAs-HumBeta}

NLS/HAs-HumBeta was released from pUC57-Kan::NLS/HAs-HumBeta by cleavage
with HindIII and EcoRI and ligated to pUC19(delta-KpnI) that had been digested with
HindIII and EcoRI and treated with alkaline phosphatase to eliminate the 5’ end
phosphates to prevent pUC19(delta-KpnI) self-ligation. 100 ng of insert and 50 ng of
vector were ligated, ligase was heat inactivated at 70 °C for 30 min, and 2 µl of the
ligation reaction was used to transform chemically competent NEB5α cells to AmpR and
selected on LB+Carb(100) plates. The correct plasmid **pUC19::NLS/HA-HumBeta** was validated with EcoRI digestion that produced a 3560 linear plasmid as expected. This plasmid places expression of the HumBeta construction under control of $P_{lac}/Iac/IPTG$, a property that was used in screens for Crimson fusions to HumBeta.

**pSLIK1**

**pFN24K::Crimson/P2A (AKA pFN24K::Product N, RSM255, RIK588)**

Crimson was amplified from pTEC19 (RSM177) with primers 145 and 146. One ng of the resulting PCR product (742 bp) was used as template for a second PCR using phosphorylated primers 147 and 148 to incorporate the P2A sequence and to phosphorylate the 5’-ends. The final PCR product of 795 bp was purified by gel extraction and was named “Product N”.

pFN24K (Promega) was cut with BsaAI (NEB) and then treated with alkaline phosphatase (NEB) to eliminate the 5’ end phosphates to prevent pFN24K self-ligation. The largest linear pFN24K fragment of 2477 bp was purified by gel extraction, 50 ng was ligated to 100 ng PCR “Product N”, and ligase was heat inactivated at 70 °C for 30 min. 2 µl of the ligation reaction was used to transform chemically competent NEB5α cells and Kan$^R$ clones were selected on LB+Kan25 plates. The correct plasmid **pFN24K::Crimson/P2A** is 3282 bp, expresses Crimson from T7 promoter and can release a 779 bp insert when cleaved with AsiSI. Clones were screened for red fluorescence by fluorescence microscopy and validated by AsiSI restriction digestion analysis.
pFN22K-Crimson/P2A-ICP8 (AKA: RSM307, RIK608)

The insert Crimson/P2A was released from pFN24K-Crimson/P2A with AsiSI and gel purified.

pFN22K::ICP8 was cut with AsiSI and treated with alkaline phosphatase (NEB) to eliminate the 5’ end phosphates to prevent pFN22K::ICP8 self-ligation. Enzymes were removed using a PCR cleanup kit (USB PrepEase).

50 ng of dephosphorylated pFN22K::ICP8/AsiSI and 100 ng of the pFN24K-Crimson/P2A/AsiSI 779 bp insert were ligated, ligase was heat inactivated at 70 °C for 30 min, and 2 µl of the ligation reaction was used to transform chemically competent NEB5α cells and KanR clones were selected on LB+Kan25 plates. The correct plasmid pFN22K-Crimson/P2A-ICP8 is 8191 bp, should expresses Crimson due to leaky T7 promoter-mediated transcription and can release a 4376 bp Crimson/P2A-ICP8 insert when cleaved with BglII and EcoRI. Clones were screened for red fluorescence by fluorescence microscopy but they were not red, possible because ICP8 is not well expressed in E. coli due to differences in codon usage in human and bacterial cells. Positive clones were identified by digestion with BglII and EcoRI.

pENTR2B/TREPPitt::Crimson/P2A-ICP8 (AKA: RSM376, RIK631)

The insert Crimson/P2A-ICP8 was released from pFN22K-Crimson/P2A-ICP8 with BglIII and EcoRI-HF. The largest band of 4376 bp was gel purified.

pENTR2B/TREPPitt was first cut with BamHI and then cut with EcoRI-HF. The largest band of 2685 bp was gel purified.

The pENTR2B/TREPPitt/BamHI 2685 bp fragment and the pFN22K-Crimson/P2A-ICP8/BglII/EcoRI-HF 4376 bp insert were ligated, ligase was heat inactivated at 70 °C
for 30 min, and 2 µl of the ligation reaction was used to transform chemically competent NEB5α cells and Kan\textsuperscript{R} clones were selected on LB+Kan25 plates. The correct plasmid pENTR2B/TREPitt::Crimson/P2A-ICP8 is 7061 bp. Clones were confirmed by restriction analysis with EcoRI and AsiSI.

\textit{pSLIK/TREPitt::Crimson/P2A-ICP8 (AKA: RSM479, RIK711)}

pENTR2B/TREPitt::Crimson/P2A-ICP8 (pENTER) and pSLIK-Zeo (pDEST) were mixed with the Gateway LR clonase II reaction. The products of site-specific recombination were used to transform NEB 5α to Kan\textsuperscript{R}. The correct plasmid pSLIK/TREPitt::Crimson/P2A-ICP8 is 16133 bp. Correct clones were screened by restriction digestion with EcoRI. One of the clones (RSM427) was used to transform Stbl3 chemically competent cells for higher production of the lentivirus vector than in our other \textit{E. coli} strains. The lentivirus vector was isolated using a maxiprep-endotoxin free kit (Qiagen) and submitted for sequencing to Genewiz. The sequence obtained confirmed the entire sequence of the insert between the recombined \textit{att} site in the pSLIK-zeo vector.

\textit{pSLIK2}
\textit{pUC19::Crimson/P2A-NLS/HA-HumBeta (AKA: RSM283, RIK607)}

The 777 bp Crimson/P2A fragment was released from pFN24K::Crimson/P2A with PciI and gel purified.

pUC19ΔKpnI::NLS/HA-HumBeta was linearized with Neol and then treated with alkaline phosphatase (NEB) to eliminate the 5’ end phosphates to prevent self-ligation. The enzymes were removed using the PCR cleanup kit.
Dephosphorylated pUC19ΔKpnI::NLS/HA-HumBeta/NcoI was ligated to the 777 bp pFN24K::Crimson/P2A/ PciI fragment, ligase was heat inactivated at 70 °C for 30 min, and 2 µl of the ligation reaction was used to transform chemically competent NEB5α cells and AmpR clones were selected on LB+Carb50 plates. The correct plasmid pUC19::Crimson/P2A-NLS/HA-HumBeta is 4337 bp, expresses Crimson from the Lac UV5 promoter and can release a 1673 bp insert when cleaved with BglII and EcoRI. Clones were screened for red fluorescence. Interestingly, positive clones were blue under standard fluorescent lab room lighting, reflecting the absorbance spectrum of the strongly expressed Crimson fluorescent protein. The clones were also super bright on the Dark Reader imaging box. Clones were validated by BglII and EcoRI restriction analysis.

pENTR2B/TREPitt::Crimson/P2A-NLS/HA-HumBeta (AKA: RSM392, RIK647)

The 1673 bp Crimson/P2A-NLS/HA-HumBeta fragment was released from pUC19::Crimson/P2A-NLS/HA-HumBeta with BglII and EcoRI-HF.

pENTR2B/TREPitt was first cut with BamHI and then cut with EcoRI-HF. The largest band of 2685 bp was gel purified.

pENTR2B/TREPitt/BamHI/EcoRI 2685 and the pUC19::Crimson/P2A-NLS/HA-HumBeta/BglII/EcoRI-HF 1673 bp insert were ligated, ligase was heat inactivated at 70 °C for 30 min, and 2 µl of the ligation reaction was used to transform chemically competent NEB5α cells and KanR clones were selected on LB+Kan25 plates. The correct plasmid pENTR2B/TREPitt::Crimson/P2A-NLS/HA-HumBeta is 4358 bp. Clones were validated by EcoRI restriction analysis.
pSLIK/TREPitt::Crimson/P2A-NLS/HA-HumBeta (AKA: RSM480, RIK715)

pENTR2B/TREPitt::Crimson/P2A-NLS/HA-HumBeta (pENTER) and pSLIK-Zeocin (pDEST) were mixed with the Gateway LR clonase II reaction. The products of site-specific recombination were used to transform NEB 5α. The correct plasmid pSLIK/TREPitt::Crimson/P2A-NLS/HA-HumBeta is 13617 bp. Correct clones were confirmed by restriction digestion with EcoRI. One of the clones (RSM435) was used to transform Stbl3 chemically competent cells for higher production of the lentivirus vector than in our other E. coli strains. The lentivirus vector was isolated using a maxiprep-endotoxin free kit (Qiagen) and submitted for sequencing to Genewiz. The sequence obtained confirmed the entire sequence of the insert between the recombined att site in the pSLIK-zeocin vector.

pSLIK3

pFN24K::Crimson/P2A-Control (AKA: RSM262, RIK594)

One ng of the 795 bp PCR “Product N” described above was used as template for a third PCR to create “Product N-control”. The primers used for this PCR were phosphorylated primers 145 and 149. The final PCR product of 784 bp was purified by gel extraction.

pFN24K (Promega) was cut with BsaAI (NEB) and then treated with alkaline phosphatase (NEB) to eliminate the 5’ end phosphates to prevent self pFN24K backbone ligation. The largest linear fragment of 2477 bp was purified by gel extraction. Dephosphorylated pFN24K/BsaAI and PCR “Product N-control” were ligated, ligase was heat inactivated at 70 °C for 30 min, 2 µl of the ligation reaction was used to transform chemically competent NEB5α cells and KanR clones were selected on LB+Kan25 plates.
The correct plasmid **pFN24K::Crimson/P2A-Control** is 3261 bp, expresses Crimson from the T7 promoter, and can release a 768 bp insert when cleaved with with BglII and EcoRI. Clones were screened for red fluorescence by fluorescence microscopy and validated by BglII and EcoRI restriction analysis.

**pENTR2B/TREPitt::Crimson/P2A-Control (AKA: RSM400, RIK655)**

The insert Crimson/P2A-control was released from pFN24K::Crimson/P2A-Control with BglII and EcoRI-HF. The insert of 768 bp was gel purified.

pENTR2B/TREPitt was first cut with BamHI and then cut with EcoRI-HF. The largest band of 2685 bp was gel purified.

pENTR2B/TREPitt/BamHI/EcoRI 2685 and the 768 bp pFN24K::Crimson/P2A-Control/BglII/EcoRI-HF insert were ligated, ligase was heat inactivated at 70 °C for 30 min, 2 µl of the ligation reaction was used to transform chemically competent NEB5α cells and KanR clones were selected on LB+Kan25 plates. The correct plasmid **pENTR2B/TREPitt::Crimson/P2A-Control** is 3453 bp. Clones were confirmed by EcoRI and by AsiSI + NcoI restriction analysis.

**pSLIK/TREPitt::Crimson/P2A-Control (AKA: RSM481, RIK719)**

pENTR2B/TREPitt:: Crimson/P2A-Control (pENTER) and pSLIK-Zeocin (pDEST) were mixed with the Gateway LR clonase II reaction. The products of site-specific recombination were used to transform NEB 5α. The correct plasmid **pSLIK/TREPitt::Crimson/P2A-Control** is 12712 bp. Correct clones were confirmed by EcoRI restriction analysis. One of the clones (RSM439) was used to transform Stbl3 chemically competent cells for higher production of the lentivirus vector than in our other *E. coli* strains. The lentivirus vector was isolated using a maxiprep-endotoxin free kit
(Qiagen) and the insert was submitted for sequencing to Genewiz. The sequence obtained confirmed the entire sequence of the insert between the recombined att site in the pSLIK-zeocin vector.

\[ pSLIK4 \]

\[ pFN22K::ICP8-P2A/Crimson \] (AKA: RSM326, RIK623)

P2A/Crimson was amplified from pTEC19 (RSM177) with primers 150 and 151. One ng of the resulting PCR product (717 bp) was used as template for a second PCR to incorporate the P2A sequence and phosphorylate the 5’ ends. The primers used for the second PCR were phosphorylated 152 and 151. The final PCR product of 746 bp was purified by gel extraction and was named “Product C”.

\[ pFN22K::ICP8 \] was cut with PmeI and treated with alkaline phosphatase (NEB) to eliminate the 5’ end phosphates to prevent self-ligation. Enzymes were removed using a PCR cleanup kit.

The pFN22K::ICP8/ PmeI/Phosphatase backbone vector and PCR ‘Product C’ insert were ligated, ligase was heat inactivated at 70 °C for 30 min, 2 µl of the ligation reaction was used to transform chemically competent NEB5α cells and KanR clones were selected on LB+Kan25 plates. The correct plasmid \[ pFN22K::ICP8-P2A/Crimson \] is 8158 bp, expresses Crimson in RIK276 with IPTG, and can be cleaved once with AsiSI. Clones were validated by BamHI restriction analysis.

\[ pENTR2B/TREPitt::ICP8-P2A/Crimson \] (AKA: RSM408, RIK663)

\[ pFN22K::ICP8-P2A/Crimson \] was cut with AsiSI. Then, the ends were blunted by a fill-in reaction with T4 DNA Polymerase (+ 100 μM dNTPs, 0.1 U/µg DNA T4 Pol, 100
μg/ml BSA) for 20 minutes at 12 °C and promptly heat inactivated at 65 °C for 10 minutes. Enzymes were removed using the PCR cleanup kit (USB PrepEase). The modified plasmid was then cut with EcoRI-HF to release the ICP8-P2A/Crimson insert of 4348 bp, which was then gel, purified.

pENTR2B/TREPitt was cut with BamHI. Then, the ends were blunted with T4 DNA Polymerase as described above. Enzymes were removed using the PCR cleanup kit (USB PrepEase). The modified plasmid was then cut with EcoRI-HF, treated with alkaline phosphatase and the largest band of 2685 bp was gel purified.

Dephosphorylated pENTR2B/TREPitt/BamHI/Blunted/EcoRI and the pFN22K::ICP8-P2A/Crimson/AsiSI/Blunted/EcoRI 4348 bp insert were ligated, ligase was heat inactivated at 70 °C for 30 min, 2 µl of the ligation reaction was used to transform chemically competent NEB5α cells and KanR clones were selected on LB+Kan25 plates. The correct plasmid pENTR2B/TREPitt::ICP8-P2A/Crimson is 7037 bp. Clones were confirmed by EcoRI and Posit restriction analysis.

*pSLIK/TREPitt::ICP8-P2A/Crimson (AKA: RSM482, RIK723)*

pENTR2B/TREPitt::ICP8-P2A/Crimson (pENTER) and pSLIK-Zeocin (pDEST) were mixed with the Gateway LR clonase II reaction. The products of site-specific recombination were used to transform NEB 5α. The correct plasmid pSLIK/pSLIK/TREPitt::ICP8-P2A/Crimson is 16297 bp. Correct clones were confirmed by restriction digestion with EcoRI. One of the clones (RSM443) was used to transform was used to transform Stbl3 chemically competent cells for higher production of the lentivirus vector than in our other E. coli strains. The lentivirus vector was isolated using a maxiprep-endotoxin free kit (Qiagen) and the insert was submitted for sequencing to
Genewiz. The sequence obtained confirmed the entire sequence of the insert between the recombined \textit{att} site in the pSLIK-zeocin vector.

\textit{pSLIK5}

\textit{pUC19::NLS/HA-HumBeta- P2A/Crimson (AKA:RSM266, RIK601)}

\textit{pUC19::NLS/HA-HumBeta} was cut with Pmel and treated with alkaline phosphatase (NEB) to eliminate the 5' end phosphates to prevent self-ligation. Enzymes were removed with the PCR cleanup kit (USB PrepEase).

Dephosphorylated \textit{pUC19::NLS/HA-HumBeta/Pmel} and the PCR “Product C” insert were ligated, ligase was heat inactivated at 70 °C for 30 min, 2 µl of the ligation reaction was used to transform chemically competent NEB5α cells and Amp\textsuperscript{R} clones were selected on LB+Carb50 plates. The correct plasmid \textit{pUC19::NLS/HA-HumBeta-P2A/Crimson} is 4306 bp, expresses Crimson from the Lac UV5 promoter in pUC19 in the presence of IPTG, and can release a 1663 bp insert when cleaved with SnaBI and EcoRI. Clones were screened for red fluorescence by fluorescence microscopy and confirmed with SnaBI and EcoRI restriction analysis.

\textit{pENTR2B/TREPitt::NLS/HA-HumBeta-P2A/Crimson (AKA: RSM409, RIK664)}

The 1663 bp NLS/HA-HumBeta-P2A/Crimson insert was released from \textit{pUC19ΔKpnI-Kan::NLS/HA-HumBeta-P2A/Crimson} with SnaBI and EcoRI.

\textit{pENTR2B/TREPitt} was cut with BamHI. Then, the ends were blunted with T4 DNA Polymerase as described above. Enzymes were removed using the PCR cleanup kit (USB PrepEase). The modified plasmid was then cut with EcoRI-HF, treated with alkaline phosphatase and the largest band of 2685 bp was gel purified.
Dephosphorylated pENTR2B/TREPitt/BamHI/Blunted/EcoRI and the 1663 bp pUC19ΔKpnI-Kan::NLS/HA-HumBeta-P2A/Crimson/SnaBI/EcoRI insert were ligated, ligase was heat inactivated at 70 °C for 30 min, and 2 µl of the ligation reaction was used to transform chemically competent NEB5α cells and KanR clones were selected on LB+Kan25 plates. The correct plasmid pENTR2B/TREPitt::NLS/HA-HumBeta-P2A/Crimson is 4358 bp. Clones were confirmed by EcoRI and PstI restriction analysis.

pSLIK/TREPitt::NLS/HA-HumBeta-P2A/Crimson (AKA: RSM483, RIK728)

pENTR2B/TREPitt::NLS/HA-HumBeta- P2A/Crimson (pENTER) and pSLIK-Zeocin (pDEST) were mixed with the Gateway LR clonase II reaction. The products of site-specific recombination were used to transform NEB 5α. The correct plasmid pSLIK/TREPitt::NLS/HA-HumBeta-P2A/Crimson is 13612 bp. Correct clones were confirmed by restriction analysis with EcoRI. One of the clones (RSM447) was used to transform Stbl3 chemically competent cells for higher production of the lentivirus vector than in our other E. coli strains. The lentivirus vector was isolated using a maxiprep-endotoxin free kit (Qiagen) and the insert was submitted for sequencing to Genewiz. The sequence obtained confirmed the entire sequence of the insert between the recombined att site in the pSLIK-zeocin vector.

pSLIK6

pFN24K::P2A/Crimson-Control (AKA:RSM363, RIK622)

P2A/Crimson-Control was amplified from “Product C” with phosphorylated primers 153 and 154. The resulting PCR product of 787 bp was purified by gel extraction and was named “Product C-Control”.
pFN24K was cut with BsaAI and then treated with alkaline phosphatase to eliminate the 5’ end phosphates to prevent self-ligation of pFN24K. The largest linear fragment of 2477 bp was purified by gel extraction.

Dephosphorylated pFN24K/BsaAI and the PCR “Product C-Control” insert were ligated, ligase was heat inactivated at 70 °C for 30 min, 2 µl of the ligation reaction was used to transform chemically competent NEB5α cells and KanR clones were selected on LB+Kan25 plates. The correct plasmid **pFN24K::P2A/Crimson-Control** is 3264 bp, expresses Crimson from the T7 promoter and can release a 771 bp insert when cleaved with BglIII and EcoRI. Clones were confirmed for red fluorescence by fluorescence microscopy and by BglIII and EcoRI restriction analysis.

**pENTR2B/TREPitt::P2A/Crimson-Control (AKA: RSM416, RIK671)**

The 771 bp P2A/Crimson-Control insert was released from pFN24K::P2A/Crimson-Control with BglIII and EcoRI-HF and gel purified.

pENTR2B/TREPitt was cut with BamHI. Then, the ends were blunted with T4 DNA Polymerase as described above. Enzymes were removed using the PCR cleanup kit (USB PrepEase). The modified plasmid was then cut with EcoRI-HF, treated with alkaline phosphatase and the largest band of 2685 bp was gel purified.

Dephosphorylated pENTR2B/TREPitt/BamHI/EcoRI and the 771 bp pFN24K::P2A/Crimson-Control insert were ligated, ligase was heat inactivated at 70 °C for 30 min, 2 µl of the ligation reaction was used to transform chemically competent NEB5α cells and KanR clones were selected on LB+Kan25 plates. The correct plasmid **pENTR2B/TREPitt::P2A/Crimson-Control** is 3456 bp. Clones were confirmed by EcoRI and PstI restriction analysis.
pSLIK/TREPitt::P2A/Crimson-Control (AKA: RSM484, RIK731)

pENTR2B/TREPitt:: P2A/Crimson-Control (pENTER) and pSLIK-Zeocin (pDEST) were mixed with the Gateway LR clonase II reaction. The products of site-specific recombination were used to transform NEB 5α. The correct plasmid pSLIK/TREPitt::P2A/Crimson-Control is 12715 bp. Correct clones were confirmed by restriction digestion with EcoRI. One of the clones (RSM451) was used to transform Stbl3 chemically competent cells for higher production of the lentivirus vector than in our other E. coli strains. The lentivirus vector was isolated using a maxiprep-endotoxin free kit (Qiagen) and the insert was submitted for sequencing to Genewiz. The sequence obtained confirmed the entire sequence of the insert between the recombined att site in the pSLIK-zeocin vector.

Expressing recombinases from stably or transiently transfected cells using the RheoSwitch® Mammalian Inducible Expression System

pNEBR-X1Hygro-UL12.5

UL12.5 was PCR amplified from a cosmid that contained a segment of the Human Herpes Simplex Virus (HSV1) strain 17 from nucleotides 2995-45035 isolated from RIK124 with oligos 164 and 165. PCR was performed using FailSafe PCR Enzyme and PreMix buffers J, K and L (they all worked). The program consisted of 30 cycles denaturing at 95 °C for 1 min, annealing at 65 °C for 1 min and extension at 73 °C for 2 min. The program also included final extension of 10 min. The PCR product was agarose gel purified and subcloned into the pGEM-T vector. The UL12.5 insert and pNEBRX1 were each cut with EcoRV and NotI, gel purified and ligated. pNEBR-X1Hygro-UL12.5 plasmid was validated by colony PCR.
pNEBR-X1Hygro-UL12.5-HA

UL12.5 was PCR amplified from a cosmid that contained a segment of the HSV1 strain 17 from nucleotides 2995-45035 isolated from RIK124 with oligos 164 and 166. PCR was performed using FailSafe PCR Enzyme and PreMix buffers J, K and L (they all worked). The program consisted of 30 cycles denaturing at 95 °C for 1 min, annealing at 65 °C for 1 min and extension at 73 °C for 2 min. The program also included final extension of 10 min. The PCR product was agarose gel purified and subcloned into the pGEM vector. The UL12.5-HA insert and pNEBR-X1Hygro were each cut with EcoRV and NotI, gel purified and ligated. pNEBR-X1Hygro-UL12.5-HA plasmid was validated by colony PCR. UL12.5-HA was validated by Western Blot and Immunocytochemistry using a polyclonal antibody against the HA epitope.

pNEBR-X1Hygro-MLS

Complementary oligos 173 and 174, coding for the Mitochondria Localization Signal (MLS) were hybridized by resuspending them in an equimolar amount in annealing buffer (10 mM Tris, pH 7.5-8.0, 50 mM NaCl, 1 mM EDTA). The oligos were then boiled for 2 minutes, the heater was turned off and the oligos were left in the water until it reached room temperature. pNEBR-X1Hygro was cut with HindIII and ApaI and agarose gel purified. Oligos and vector were ligated and transform into XL10 Gold. Colonies from transformation were examined by colony PCR using primers 109 and 110. Primers 109 and 110 flank the multiple cloning site. pNEBR-X1Hygro-MLS and should produce a fragment of 309 bp while the empty pNEBR-X1Hygro should produce a fragment of 214 bp.
**pNEBR-X1Hygro-ICP8-GFP**

Ligation of pNEBR-X1/HindIII/EcoRV and pCMV-ICP8-GFP/HpaI/HindIII.

The insert ICP8-GFP was cut from pCMV-ICP8-GFP with Hpal and HindIII, while pNEBR-X1Hygro was cut with HindIII and EcoRV. Both insert and plasmid were gel purified and ligated. NEB5α cells were transformed with the ligation mix. pNEBR-X1Hygro::ICP8-GFP was validated by restriction enzyme digestion. Expression of ICP8-GFP was evaluated by fluorescent microscopy, flow and Western Blot using monoclonal antibody to ICP8.

**pNEBR-X1Hygro-ICP8**

The insert ICP8 was cut from pCMV-ICP8 with Hpal and HindIII, while pNEBR-X1Hygro was cut with HindIII and EcoRV. Both insert and plasmid were gel purified and ligated. NEB5α cells were transformed with the ligation mix. pNEBR-X1Hygro::ICP8 was validated by restriction enzyme digestion. Expression of ICP8 was evaluated by immunocytochemistry, flow and Western Blot using monoclonal antibody to ICP8.

Plasmids for SynExo interactome studies

**pFN22K::β**

pFN22K clones were made following the Promega Flexi system Entry/Transfer (C8640) manual protocol. Briefly, the Beta gene was amplified by PCR from pSIM5 using primers Beta-Flex1 and Beta-Flex2 using KOD Hot Start master mix from EMD Millipore (71842). The PCR product was cleaned up using the Promega Wizard SV Gel and PCR Clean-Up System and then digested with the Flexi Enzyme Blend (Sgfl/Pmel). 100 ng of Flexi Enzyme Blend digested PCR product was ligated with 50 ng of Flexi
Enzyme Blend digested pFN22K and transformed into NEB5α, selecting for Kan\(^R\) on LB+Kan25 plates. 100% of the clones were correct.

\textbf{pFN22K::β (1-177)}

pFN22K clones were made following the Promega Flexi system Entry/Transfer manual protocol. Briefly, the Beta gene was amplified by PCR from codon 1 to codon 177 from pSIM5 using primers Beta-Flex1 and Beta (1-177)-Flexi3 using KOD Hot Start master mix from EMD Millipore (71842). The PCR product was cleaned up using the Promega Wizard SV Gel and PCR Clean-Up System and then digested with the Flexi Enzyme Blend (SgfI/PmeI). 100 ng of Flexi Enzyme Blend digested PCR product was ligated with 50 ng of Flexi Enzyme Blend digested pFN22K and transformed into NEB5α, selecting for Kan\(^R\) on LB+Kan25 plates. 100% of the clones were correct.

\textbf{pFN22K::Exo}

pFN22K clones were made following the Promega Flexi system Entry/Transfer manual protocol. Briefly, the Exo gene was amplified by PCR from pSIM5 using primers Exo-Flexi4 and ExoFlexi-5 using KOD Hot Start master mix from EMD Millipore (71842). The PCR product was cleaned up using the Promega Wizard SV Gel and PCR Clean-Up System and then digested with the Flexi Enzyme Blend (SgfI/PmeI). 100 ng of Flexi Enzyme Blend digested PCR product was ligated with 50 ng of Flexi Enzyme Blend digested pFN22K and transformed into NEB5α, selecting for Kan\(^R\) on LB+Kan25 plates. 100% of the clones were correct.

\textbf{pFN22K::Exo (1-189)}

pFN22K clones were made following the Promega Flexi system Entry/Transfer manual protocol. Briefly, the Exo gene was amplified by PCR from codon 1 to codon 189
from pSIM5 using primers Exo-Flexi4 and Exo(1-189)-Flexi5 using KOD Hot Start master mix from EMD Millipore (71842). The PCR product was cleaned up using the Promega Wizard SV Gel and PCR Clean-Up System and then digested with the Flexi Enzyme Blend (SgfI/Pmel). 100 ng of Flexi Enzyme Blend digested PCR product was ligated with 50 ng of Flexi Enzyme Blend digested pFN22K and transformed into NEB5α, selecting for KanR on LB+Kan25 plates. 100% of the clones were correct.

pFC15K::β

pFC15K clones were made following the Promega transfer protocol. Briefly, the β insert was released from pFN22K::β using Carboxy Flexi Enzyme Blend (SgfI/EcoICRI) followed by gel purification. 100 ng of the purified digested insert was ligated with 50 ng of Carboxy Flexi Enzyme Blend digested pFC15K. NEB5α was transformed with the ligation reaction and clones were selected for KanR on LB+Kan25 plates. 100% of the clones were correct.

pFC15K::β (1-177)

pFC15K clones were made following the Promega transfer protocol. Briefly, the β(1-177) insert was released from pFN22K::β (1-177) using Carboxy Flexi Enzyme Blend (SgfI/EcoICRI) followed by gel purification. 100 ng of the purified digested insert was ligated with 50 ng of Carboxy Flexi Enzyme Blend digested pFC15K. NEB5α was transformed with the ligation reaction and clones were selected for KanR on LB+Kan25 plates. 100% of the clones were correct.

pFC15K::Exo

pFC15K clones were made following the Promega transfer protocol. Briefly, the Exo insert was released from pFN22K::Exo using Carboxy Flexi Enzyme Blend
(SgfI/EcoICRI) followed by gel purification. 100 ng of the purified digested insert was ligated with 50 ng of Carboxy Flexi Enzyme Blend digested pFC15K. NEB5α was transformed with the ligation reaction and clones were selected for KanR on LB+Kan25 plates. 100% of the clones were correct.

pFC15K::Exo (1-189)

pFC15K clones were made following the Promega transfer protocol. Briefly, the Exo (1-189) insert was released from pFN22K::Exo(1-189) using Carboxy Flexi Enzyme Blend (SgfI/EcoICRI) followed by gel purification. 100 ng of the purified digested insert was ligated with 50 ng of Carboxy Flexi Enzyme Blend digested pFC15K. NEB5α was transformed with the ligation reaction and clones were selected for KanR on LB+Kan25 plates. 100% of the clones were correct.

**Recombineering**

**E. coli Recombineering and fluorescent protein engineering**

Recombineering targets in this study included the *gfpmut3* gene integrated in the *E. coli* genome and the *egfp* gene and variants from the pDual-egfp plasmid which replicates in *E. coli*. Recombineering in *E. coli* was performed essentially as in (Yu et al. 2000) and with some adjustments for plasmid Recombineering, as described by (L. C. Thomason et al. 2007)

DNA substrate for Recombineering could be linear ssDNA or dsDNA. Fluorescent protein engineering was performed mostly with ssDNA oligos. Gene knockouts were performed mostly with dsDNA PCR cassettes and also with P1 phage transducing lysates.
Recombineering oligos were designed to hybridize to the lagging strand template of the replication fork, unless specified. For both \textit{gfpmut3*} and pDual-\textit{egfp} the oligo that hybridizes to the lagging strand template was the sequence of the sense strand.

The direction of replication for chromosomal \textit{gfpmut3*} was determined by the coordinates of the insertion of \textit{gfpmut3*} published by the creator (Andersen et al. 1998) with respect to the \textit{E. coli} bidirectional origin of replication (\textit{gfpmut3*} integrated at nts 312754-312771). Initial results of Recombineering indicated that the \textit{gfpmut3*} gene was likely in inverted orientation relative to the published report, so the orientation of \textit{gfpmut3*} with respect of the replication fork was determined by PCR. This analysis confirmed our prediction that the gene was integrated in opposite orientation to the published report, so Recombineering studies were performed in light of this knowledge.

The direction of the replication fork passing through \textit{egfp} in pDual-\textit{egfp} was predicted by the considering the orientation of the unidirectional ColE1 origin of replication in pDual-\textit{egfp} (Reiser et al. 2000).

Therefore, oligos were designed with the sense sequence of \textit{gfpmut3*} (mostly) and \textit{egfp}, with desired change in the middle and 35-49 nt homology at each side.

When designing the change towards a particular protein sequence, MMR recognition for different (DuBridge \textit{et al.} 1987; Joshi \\& Rao 2001; Costantino \\& Court 2003a) and \textit{E. coli} codon usage frequencies (http://www.kazusa.or.jp/codon/cgi-bin/showcodon.cgi?species=83333) were taken into account to maintain expression levels in recombinants. Oligos were ordered from Sigma Genosys, PAGE purified to avoid incompletely synthesized oligos, and resuspended in TE buffer (10mM Tris 1 mM EDTA pH 7.55) to 100 ng/\textmu l and stored at -20°C until used.
The dsDNA PCR products were designed to incorporate a drug resistant cassette usually Kan 100% linked to the gene that was been deleted. The Kan cassette was flanked by ~35 nucleotides of homologous sequence to the flanking region that was desired to remove from a replicon. The primers for the PCR were designed for ~20 nucleotides of the 3’ end to hybridize to the PCR template (e.g. the Kan cassette) and ~35 nucleotides of homologous sequence to the Recombineering target towards the 5’ end. The PCR reaction was then treated with DpnI to digest the Dam-methylated PCR template DNA and the unmethylated PCR product was purified by gel extraction before Recombineering.

P1 lysates were grown on donor strains that carried the desired knockouts to permit movement into other genetic backgrounds. Recombineering strains were revived from frozen stocks and propagated to mid log

Recombineering strains encoding the recombinases were thawed from a -80°C stock by streaking a chip of frozen culture for single colony isolation on an LB plate and incubated overnight at 32 °C. When gfpmut3* was the target, Recombineering was performed in RIK410 (λ cl\textsuperscript{857} Δ(cro-bioA) PA1/04/03-gfpmut3*) or a derivative of this strain; pDual-egfp Recombineering was performed in RIK473 ([λ cl\textsuperscript{857} Δ(cro-bioA)<=tetRA] recA1 ΔmutS::Kan). RIK473 was used to engineer pDual-egfp because the recA1 mutation decreases LTR recombination in the lentivirus sequence and plasmid concatemerization and the mutS deletion disables MMR and stabilizes heteroduplex recombination intermediates.

The next day, 2 ml LB broth was inoculated with one colony of the Recombineering strain and grown overnight at 32 °C with shaking. The culture was then diluted 1/100 and
grown with shaking at 32 °C to OD 0.5-0.6 (spectrophotometer) or 23 KU (Klett meter), usually ~2 hr 15 minutes for RIK410. The highest electroporation and Recombineering efficiencies were obtained with cells in early-log phase.

Recombinases were induced with a sharp heat shock

Cultures were then transferred to a 42 °C water-bath with vigorous shaking (200 RPM) for exactly 15 minutes to induce the expression of the recombinases. Immediately after, cultures were transferred to an ice water slurry and swirled for 2 minutes, then cultures were left sitting on the ice water slurry for 8 minutes more to stop induction. Additional incubation at 42 °C reduced cell viability and overall Recombineering efficiency. A water bath was preferred to an air shaker as the heat transfer is more rapid in water, allowing better control of the induction time.

DNA substrate delivery to cells for Recombineering by electroporation of competent cells or by P1 transduction

Cells were made competent for electroporation via washes in cold water in the shortest time possible. Cells were transferred to pre-cooled centrifuge tubes and spun in a refrigerated centrifuge at 5000 rpm for 5 minutes. Supernatants were decanted and the tubes with the cell pellets were transferred to an ice bucket and taken to the cold room to keep low temperatures while preparing the competent cells. In the cold room, cells were resuspended in 1 ml cold water by swirling the tube gently. Cells were then transferred to a 2 ml microfuge tube using a chilled pasteur pipette. Cells were spun in a cold microfuge at 9,000 rpm for 30 seconds. The supernate solution was removed carefully using a pasteur pipette avoiding aspiration of the soft cell pellet. The cells were washed in this
way 2 more times and resuspended 1/200 of the initial culture volume with ice-cold sterile water.

At this point the cells were very fragile and were dying from the osmotic pressure. Therefore, cells were taken immediately for electroporation to introduce the DNA substrate.

DNA substrate delivery to cells for Recombineering by electroporation 1-2 µl of DNA was aliquotted in microfuge tubes on ice. 100-200 ng was used for DNA substrates like oligos or PCR products; for plasmid Recombineering, 10 ng of plasmid was also added keeping the total DNA volume at ~2-3 µl.

40 µl of induced competent cells were added to the first tube containing DNA. The mix was subsequently transferred to a prechilled 1 mm gap electroporation cuvette, avoiding air bubbles. The cuvette was briefly tapped 3 times to settle the mixture between the plates of the cuvette, dried with the lab coat, moved to the electroporation chamber (e.g. Bio-Rad Gene Pulser) and pulsed with 1.75 KV at 200 ohms and 25 microfarads. Immediately after the discharge, 1 ml LB medium was added to the cuvette and cells were then transferred to a culture tube. The time constants during the electroporation were recorded since cells were most likely to survive the current with the highest transformation efficiencies when time constants were between 4 and 4.8 ms.

If there were more samples for electroporation, cells were added to the second DNA tube and electroporation was continued. When all the samples were ready, the culture tubes were incubated at 32 °C with shaking.

When targeting fluorescent protein genes or transforming with an AmpR marker, 30 minutes of outgrowth was enough time to allow the cells to recover from electroporation.
After outgrowth, 100 µl of a 5x10^{-6} dilution of the outgrowth was plated per LB (or LB + Amp) plate to seeking 30-100 colonies per plate. The plates were incubated at 32 °C overnight.

When a Kan^R cassette was introduced by a PCR product, 4 hours outgrowth was necessary to allow recombinants to become resistant to kanamycin. After outgrowth, 100 µl of the culture were plated on LB-Kan and incubated at 32 °C overnight.

When targeting multicopy plasmid replicons like pDual-egfp, the culture was grown overnight, plasmid DNA was isolated and new cells were electroporated at substoichiometric plasmid amounts to cells (usually ~10 ng plasmid was used). pDual-egfp was transformed into Rosetta-gami™2, which expresses eGFP from the T7 promoter. Cells were again let to recover for about 30 minutes and 100 µl of the outgrowth culture were plated on carbenicillin-chloramphenicol-LB agar to maintain pDual-egfp variants and pRARE2. Plates were incubated at 32 °C overnight.

It is also possible to use as a substrate for Recombineering the injected DNA from a P1 transducing particle. In this case, the cells were transduced with P1 shortly after induction of the λ SynExo recombinases. Briefly, after cells were chilled on ice to stop expression of recombinases, 100 µl of cells were incubated with 100 µl of phage lysate in the presence of 5 mM CaCl_2 for 20 minutes at 32 °C. Infection was stopped by growing cells in LB + 10 mM sodium citrate for the time required for outgrowth (e.g. 1 hour for Amp and 4 hours for Kan). Cells were plated on solid media with the appropriate selection plus 10 mM sodium citrate to reduce superinfection by P1.
Identification and scoring of recombinants

Fluorescent recombinant colonies were identified using a plate imager (DR46B Dark Reader transilluminator from Clare Chemical Research). Recombinants generally appeared as sectored colonies comprised of recombinant and nonrecombinant cells. Recombinants were isolated from non-recombinants by two rounds of single colony isolation by streaking on LB medium. Recombinants were validated by allele-specific PCR where the 3’ end of one of the primers overlaps with the introduced sequence change and by sequencing. Recombination frequencies were determined by dividing the recombinant colony forming units by the total colony forming units on the plate.

Transformation efficiency control

A transformation efficiency control was performed in parallel to the Recombineering experiments to determine the percent of cells competent to uptake DNA. For these controls, the same competent cells used for Recombineering were electroporated with 12.7 µM pUC19 (saturating DNA). After 30 minutes outgrowth, 200 µl of a 10⁻⁵ culture dilution was plated on LB + carbenicillin plates and 100 µl of 5x10⁻⁶ dilution on LB plates. The next day, colonies on both plates were counted and the total transformants and viable titters were determined. Transformation efficiency was calculated as the titer of Carb<sup>R</sup> cells divided by the total cell titer.

Recombineering efficiency was frequently normalized to the number of cells that could uptake the DNA substrate. Recombineering efficiency was calculated as recombination efficiency divided by transformation efficiency.
Human cells Recombineering and fluorescent protein engineering

Transient expression of recombinase

The Recombineering protocol was evaluated and different aspects of it were modified in different experiments. As a general guide, 293T-Mostaza cells were seeded in 6 well plates wells and transfected the next day with pCMV-ICP8 using Fugene at a ratio 6:2 in triplicates (day-2). On the next day (day-1), each well was passed to a 24 well plate. On the next day (day 0), media were changed to 0.75 ml of DMEM, 5 % FBS and cells were transfected with 33 pmol/well oligos using Lipofectamine 2000 in triplicates (as done in Andrieu-Soler, 2005). 3 hours later 0.75 ml of DMEM, 15 % FBS was added to each well and cells were left overnight. The next day (day 1), media were changed and the following day cells from each well were transferred to 6 well plates. Recombinants were quantified using flow cytometry (day 2 and the following days).

Inducible recombinase

The Recombineering protocol was evaluated and different aspects of it were modified in different experiments. As a general guide, 24 well plates were coated with poly-L-Lysine 30-70 kDa (Sigma, P 4707), washed and allowed to dry for some hours. 293T-Mostaza-pSLIK cells were seeded 1/5 to 1/4 so they were ~ 40 % confluent or in exponential growth next day. Seeding media was 0.5 ml DMEM 10 % Tet off FBS per well. 1000 ng/ml Dox was added when indicated to each well unless specified other [Dox], (day-1). Media was mixed and cells incubated with 5 % CO₂ in atmosphere at 37 °C. 16-24 hours later, (day 0), oligo-Lipofectamine 2000 complexes were added to wells, mixed with the media by shaking the plate and cells were incubated as once again as indicated. When cells reached about 80% confluency, each experiment was passed to a well in a 6 well plate, usually between 1 or 2 days after the oligo treatment (day 1-2).
When cells were about 80% confluent in the 6 well dish (day 2 and the following days), some of the cells were analyzed using the Accuri and some others were re-seeded to evaluate stability of the green phenotype.

**DNA transformation/transfection**

**E. coli electroporation**

Making *E. coli* competent for electroporation

*E. coli* cells were grown to mid log and then made competent by washing them 3 times with cold water (for higher transformation efficiency) or 10% glycerol (to increase viability and to freeze leftover aliquots at the expense of about 5-10 fold reduction in Recombineering efficiency). Centrifugations were performed at low speed for short periods of time (*e.g.* 5000 rpm for 5 minutes in a 4 °C clinical centrifuge or 9000 rpm for 30 seconds in a 4 °C microcentrifuge). Cells were concentrated 200 fold and kept on ice during the washes.

**E. coli electroporation**

40 µl competent *E. coli* cells were mixed with the DNA and transferred to a chilled 1 mm gap electroporation cuvette. Cells were electroporated using a Bio-Rad Gene Pulser at 1.75 KV, 200 ohms and 25 microfarads. Immediately after the discharge, media was added to cuvette to recover cells. Cells were recovered with a pasteur pipette and transferred to a culture tube for recovery at normal growth conditions. The preferred transfection protocol for *E. coli* is by electroporation because the efficiency was usually 10 times higher than chemical transformation.
*E. coli* chemical transformation

Making *E. coli* competent for chemical transformation

*E. coli* cells were grown to mid log and then made competent by washing them 3 times with cold 5% CaCl₂. Cells were concentrated 200 fold and kept on ice during the washes.

*E. coli* chemical transformation

Chemical transformations were usually performed as indicated by the company selling the competent cells. Briefly, 50 µl of cells were incubated with DNA for 30 minutes on ice. Then transferred to a 42 °C water bath for a 30-40 second heat shock. Cells were immediately moved to ice and cooled down for 2 minutes. Cells were allowed to recover and to express the protein that confers resistance to a drug on SOC for 30 minutes to 1 hour before plating for selection in solid media. Chemical transformations were preferred after ligations because it was not necessary to clean up the reaction.

**Colony transformation (no need to purify plasmid DNA)**

Resuspend 1 colony with the donor DNA in 30 µl of transfection solution (5 % PEG8000, 5 % DMSO, 50 mM CaCl₂, pH 6.5). Heat at 100 °C for 30 minutes. Grow recipient cells to mid log, pellet and resuspend in 5 % of the original culture with 20 % glycerol. Add 30 µl of recipient cells to heated donor cells and incubate on ice for 30 minutes. Heat shock cells at 42 °C for 40 seconds. Incubate on ice for 10 minutes and add LB for outgrowth. Plate 30 minutes to 1 hour later on the appropriate plates. This is a low efficiency protocol but adequate for moving plasmids between genetic backgrounds.

**Human cell transfections with oligos**

293T cell oligo transfections were performed with Lipofectamine 2000 (Invitrogen), as recommended by the manufacturers. Briefly, 293T cells were seeded in 24 well plates.
1 µl of Lipofectamine 2000 was diluted in 25 µl OptiMEM and 15 pmol oligo was
diluted in 25 µl OptiMEM. Diluted oligo was added to diluted Lipofectamine 2000, the
mix was incubated for 20 minutes at room temperature. Oligo Fugene complexes were
added drop wise to cells in 0.5 ml media without antibiotics. Cells were incubated with
transfection 1 hour to overnight.

MIAMI cell transfections with oligos were performed with Lipofectamine 2000, as
recommended by the manufacturer. Briefly, MIAMI cells were seeded in a 6 well plate
well 4 days before transfection. The day of transfection media was changed to no
antibiotic free media. 7.5 µg of FITC-CHI Mut oligos were diluted in 500 µl OptiMEM
and the 10 µl Lipofectamine 2000 were also diluted in 500 µl OptiMEM. Diluted oligo
and Lipofectamine 2000 were mixed and incubated 20 minutes at room temperature.
Oligo-Lipofectamine 2000 mix was added to cells drop wise. The transfection mix was
dispersed in the wells by rocking the plates back and forth. Cells and transfection mix
were incubated for 5 hours.

Progerin-HeLa cell oligo transfections were performed by electroporation. Cells were
seeded at low density (~20 %), and let grow to 70-80 % confluency. Cells were then
trypsinized, resuspended in media, counted and transferred 5x10^5-10^6 cells per
microcentrifuge tubes. Cells were centrifuged at 300 x g for 10 minutes, washed with
PBS and centrifuged again. Cells were mixed with 50 µl of transfection reagent (Staudt et
al., 2007) and resuspended by pipetting up and down. DNA (1 µg oligos, 10 µg plasmid,
or both) was added as indicated for particular experiments, and the mix was transferred to
electroporation cuvette. Cells were electroporated using the Nucleofector II and the
indicated program (e.g Q-013). Cell growth medium (90 % EMEM from ATCC, 30-
2003, 100 U/ml Penicillin, 100 µg/ml Streptomycin, and 10 % FBS) was added to the cuvette, and then the mix was transferred to a 6 well plate containing media. Plates were incubated at 21 % O₂, 5 % CO₂, 37 °C usually overnight.

**Human cell transfections with plasmids**

293T cells were transfected using Fugene 6 Transfection Reagent (Roche 1814443001) as indicated by the manufacturer. Briefly, cells were seeded in 6 well plates the day before transfection. Fugene 6 was vortexed every time before used. 3-6 µl of Fugene 6 was pipetted into a tube with 100 µl OptiMEM, the reagents were mixed by flicking and incubated for 5 minutes at room temperature. 1-2 µg of plasmid was added to the diluted Fugene 6. DNA and Fugene 6 reagents were mixed by flicking and incubated at room temperature for 15 minutes. The transfection mix was added to cells in normal media without antibiotics in a drop wise manner. The best transfection efficiencies for pCMV-ICP8 and pCMV-ICP8-GFP were with ~50 % cell confluency and a Fugene-plasmid ratio of 6:2.

MIAMI cells were transfected using electroporation (Digital Bio/BTX Microporator, Holliston, MA, #MPK-1096). The electroporation program consisted of 1400 V pulse voltage, 10 ms pulse width, 3 pulses. 10⁵–5x10⁵ cells and 0.5 µg of plasmid or 10–100 µM oligos were resuspended in 10 µl of transfection solution (Staudt et al, 2007). Transfection solutions were prepared with fresh ATP, aliquotted, and stored at –80 °C. MIAMI cells were reseeded at 100 to 3,000 cells/cm² and incubated overnight at 3% pO₂.

HeLa 11 cells were transfected using the Lonza Nucleofector following manufacturer specifications for transfecting HeLa cells. Briefly, cells were grown to 70-80 % confluence and harvested by trypsin-EDTA. 5 x 10⁵ to 10⁶ cells were mixed with 2 µg of plasmid DNA and the transfection reagent (Lonza Nucleofector KitR or as described by
Staudt et al. 2007). The mixture was electroporated using the program I-13. Medium was added to cells and the cells were immediately transferred to an incubator at 37 °C, 5 % CO₂ to allow recovery.

**P1 phage transductions**

**Producing P1 lysates from donor strain**

Phage lysates were produced by inoculating LB with a colony of the donor strain and culturing overnight. The next day, 50 µl of the overnight culture was mixed with 50 µl of 10⁶ pfu P1 phage in TMC (usually 1-10 µl stock). Phage were allowed to infect an *E. coli* donor strain for 5 minutes at 37 °C. The cells and phage were then transferred to 2.5 ml LB, 5 mM CaCl₂ and shaken at 37 °C for 3 to 5 hours. When most of the cells were lysed, usually by 5 hours, 6 drops of chloroform were added to the lysate and vortexed vigorously for 10 seconds. Lysate-chloroform emulsions were stored in the cold room overnight or spun down in a microfuge tube for 5 minutes. The upper phase with lysate was transferred to a detergent-free glass screw-capped tube and stored at 4 °C in the dark. To estimate the phage titer, see below.

**P1 transduction**

Transduction was used to move genetic material from one *E. coli* cell (donor) to another *E. coli* cell (recipient) using P1 phage. The first step was to produce a phage lysate with the donor strain as described above. Then, 100 µl of 10⁷ pfu/ml P1 phage in TMC grown on the donor strain was used to infect 100 µl of a freshly saturated culture of the recipient strain (~8 hours culture of 1/50 dilution from an overnight culture). Phage were allow to infect *E. coli* donor strains for 20 minutes at 37 °C. Infection was stopped by addition of sodium citrate to 10 mM with shaking incubation for the time required for
outgrowth. Cells were then plated in solid media with the appropriate selection plus 10 mM sodium citrate, unless the selection is Tetracycline. Tetracycline and sodium citrate do not allow cells to grow unless tetracycline concentrations are reduced to ¼ the usual amount. In order to completely get rid of the phage, a single colony without plaques from the transduction plate was struck on an LB-selection plate containing 10 mM sodium citrate two subsequent times. A final colony without plaques was then collected as a transduced clone.

**P1 titering**

Phage titers (PFU/ml) were determined by carrying out infections with different dilutions of the phage lysate, usually $10^{-2}$, $10^{-4}$ and $10^{-6}$ in TMC. 100 μl of the phage dilutions were added to 100 μl of *E. coli* cells and incubated for 5 minutes at 37 °C. The mixtures were then added to 2.5 ml LB + 5 mM CaCl$_2$ + 0.6 % agar preheated to 53 °C (usually for several hours to equilibrate the temperature), vortexed and poured on LB plates. On the next day, plaques were counted to determine the PFU. PFU/ml = 10 x PFU/dilution (*e.g* $10^{-2}$).

**P1 transduction of recA$^-$ hosts**

Transduction does not usually work in RecA-deficient strains. We developed a protocol where loss of RecA function is supressed by transient induction of the λ Red SynExo genes in the *E. coli* recipient strains. P1 transduction followed induction of Recombineering enzymes. Briefly, cells were grown to the exponential phase, heat-induced for 15 minutes at 42 °C to express the λ SynExo recombinases. Cells were cooled down on ice for 10 minutes to stop expression of recombinases. 100 μl of cells were incubated with 100 μl of $10^9$-$10^{11}$ PFU/ml phage lysate in the presence of 5 mM CaCl$_2$ for 20 minutes at 32 °C. Infection was stopped by growing cells in LB + 10 mM
sodium citrate for the time required for outgrowth (e.g. 1 hour for Amp and 4 hours for Kan). Cells were plated on solid media with the appropriate selection + 10 mM sodium citrate. Transduced recombinant clones were struck on LB + selection + 10 mM sodium citrate two subsequent times to get rid of all the P1 donor phage. Frequencies were relatively low, but high enough to move markers.

**Spot tests to detect λ lysogens**

Cells were grown in LB overnight at 32 °C with shaking. The next morning, cells were diluted 1/50 in PCB (1 % Bacto tryptone, 0.5 % NaCl, 0.2 % maltose, 10 mM MgSO$_2$ or MgCl$_2$) and grown to mid log. Maltose induces expression of LamB (λ receptor) and Mg$^{2+}$ facilitates interaction between the λ receptor and λ tail proteins. When the culture reached the exponential phase, an equal volume of TM buffer (10 mM Tris, 10 mM MgSO$_4$) was added. Cultures were grown for 15-30 more minutes at 32 °C shaking while cells starved. At this point cells were ready for infection, spot tests, single plaque isolation, and can be stored at 4 °C for 1-2 weeks.

The spot tests were performed with 0.3 ml of cell culture mixed with 3 ml of 53 °C melted BBL trypicase top agar to create a lawn. Plates were set until agar solidified. Then, wild type λ phage (MMS5) was used to test susceptibility to λ phage infection. Serial dilutions of λ phage were performed in λ Tryp medium (1 % Bacto tryptone, 0.5 % NaCl) supplemented with 10 mM MgSO$_2$. 50 µl of the serial dilution were dropped onto the solid cellular lawn. On the next day, plates were evaluated for the presence of plaques. It is expected that cells containing the $\lambda$ cl$^{857}$ Δ(cro-bioA) ↔ tetRA lysogen with the SynExo Red recombinases should be resistant to superinfection by phage $\lambda$ since the
regulatory protein cI is expressed. It is expected that cells that do not contain the λ lysogen should enter the lytic cycle and form plaques upon λ infection.

**Lentiviral transductions**

The recombination target genes and the viral recombinases were integrated in human cells using lentiviral transduction, which uses HIV genes to deliver and integrate the gene of interest. The transfer vectors encoding the target for Recombineering were the pDuals and the transfer vectors encoding the recombinases were the pSLIK-Zeos.

The lentiviral particles were created using the 3 plasmid second generation packaging system, which provides high transduction efficiency but lacks the HIV replication genes that make working with this virus dangerous. The 3 plasmids were the transfer plasmid encoding our gene of interest (pDual and pSLIK) flanked by LTRs, a packaging plasmid for producing viral particles, includes Gag, Pol, Rev, and Tat (pHR’CMV 8.2ΔR, (An et al. 1999)) and the Envelope plasmid for producing pseudotyped lentiviral particles (pCMV-VSV-G, (Stewart et al. 2003)).

Prepare cells for transfection

Lentivirus particles were produced in 293T because they are easy to transfect and they express the T antigen which may make the lentiviral vectors replicate so more copies of the viral protein functions might be transcribed in each cell and more lentiviral particles produced. Low passage 293T cells were thawed a week before the transfection and kept growing in stationary phase until Fugene 6 transfection. The day before transfection 293T cells were seeded ¼ on a 100 mm dish.
The recipient cells were also thawed about one week before the transduction to ensure the cell line will be OK. Cell confluency was not a concern since lentiviral particles infect dividing and non-dividing cells.

293T Fugene transfection with viral vectors

Lentiviral vectors were transfected with Fugene 6 following the manufacturer’s instructions. Briefly, 24 µl Fugene 6 was added to a tube containing 400 µl OptiMEM, and incubated for 5 minutes at room temperature. The 3 plasmids were added to the diluted Fugene 6. 4 µg of the lentiviral plasmid with the gene of interest, 4 µg of pHRCMV8.2ΔR and 0.4 µg of pCMV-VSV-G. The solution was mixed and incubated for 15 minutes at room temperature. 293T media was changed to media without antibiotics (no PenStrep). Transfection mixture was added drop wise to the cells. Plates with cells were rocked back and forth to spread the transfection mix over all the cells. Cells were incubated overnight in a BL2+ incubator with a biohazard sign posted at the door.

Safety measures

The following days, experiments were set in a Biological Safety Level 2+ facility, 2 doors away from the hallway. 293T cells producing virus and cells incubated with virus were always kept in a BL2+ incubator. Access to the laboratory was limited while working with the lentivirus and a biohazard sign was posted at the door of the laboratory and on the hood. The investigator was completely clothed and also wearing a disposable mask, lab coat, goggles, hair net and double gloves. The hood was prepared with a disposable pad to pick any possible spill, 10 % bleach in a beaker, sharp glass was avoided and all the other materials were set next to the hood. Two double Biosafety bags
were set next to the hood, one within a box resistant to punctuation for the plastic pipettes and tips. All the materials that were exposed to virus were disinfected with 10% bleach before disposal. The work surface was disinfected with bleach at completion. At the end of each experiment the waste bags were autoclaved.

Change media to 293T cells to replace transfection reagents for recipient cell media

One overnight after the Fugene 6 and plasmid transfection, the medium was changed to the media corresponding to cell that will be infected. The cells to be infected were plated in a T75 flask.

Lentiviral infection

The following day, media containing transducing particles in the 293T culture was collected to a conical tube, where cell debris were allowed to settle for about 1 minute. In the mean time, recipient cell media was added, very carefully, to the 293T cells to recover more lentivirus particles the next day (at this point, 293T cells were very easy to detach). Lentiviral suspensions were pipetted into another tube, avoiding the last 0.5 ml with cell debris in the bottom of the tube. The viral suspension was filtered with a 0.45 µm low protein binding membrane filter (Pall, 4148) to avoid contamination of the transduced cell lines with the 293T infected cells. When 293T were being the transduced cells, the filtered viral suspension was treated with 1 µg/ml (or 1 U/ml) RNase-Free DNase (Promega, M6101) and 1mM MgCl₂ for 20 minutes at 37°C to destroy plasmids and avoid any further lentiviral particle production. Infection was carried out by replacing media with the lentiviral suspension. The target cells were then incubated with the lentiviral particles for 8 hours (for ~50% transduction efficiency) to overnight (for ~90% transduction efficiency).
When MIAMI cells were the transduced cells, the cells and the virus were treated with 10 µg/ml protamine sulfate. Protamine sulfate stock was 10 mg/ml in water and filtered.

Wash lentiviral particles away

After the lentiviral suspension was removed, cells were then washed with fresh medium 2 times to wash out any left over virus or envelope proteins that can bind to cell receptors. Cells were then supplemented with normal medium under normal conditions 8 hours to overnight before a second round or infection.

2nd infection

A second round of infection was done as described for the first infection. 293T cell producing the lentiviral particles were treated with 10 % bleach after the lentivirus was collected and then discarded. Filtered virus left over was stored at -80 °C for future infections.

Gene of interest expression or drug selection

Phenotypic expression (eGFP, Mostaza, Celeste, Azure and Crimson) was evaluated 48 hours after viral removal. Selection drug was applied also 48 hours after viral removal, in this case 100 ng/ml Zeocin. A non-transduced control was kept to evaluate phenotypic expression and drug resistance of the target cell line.

Safety validation of lenti production

As a safety procedure, once the transduced cell lines were established, the cells were tested for undesired acquired lentiviral production. The supernatant of the cell lines was used to transduce 293T. The 293T transduced cells were evaluated for expression of the gene of interest.
PCR

PCR design
Oligos were designed and validated with OligoAnalyzer 3.1 (http://www.idtdna.com/analyzer/applications/OligoAnalyzer/). PCR was run in silico to validate primers and products using jPCR software (http://primerdigital.com/tools/pcr.html). In general PCR was performed as indicated by the polymerase manufacturer. In this study we used Taq (NEB) MM for screenings like colony PCR and allele specific PCR. We used KOD hot start polymerase for subcloning and sequencing. For colony PCR did an initial 5 minutes denaturing at 95 °C, while for isolated DNA usually 2 minutes. PCR troubleshooting was guided by the KOD troubleshooting section of the general protocol.

ICP8 PCR
The HHV1 UL29 (ICP8) gene was cloned into different vectors during this study. ICP8 had a high GC content and it was difficult to PCR amplify. In general, PCR was performed with the high fidelity hot start polymerase KOD. The program consisted of an initial denaturation of the plasmid DNA template (pCMV-ICP8 or pCMV-ICP8-GFP) at 98 °C for 5 seconds. Then, there were 10 cycles of denaturing at 98 °C for 5 seconds, annealing at 63 °C for 30 seconds and extension at 72 °C for 2 minutes. Finally, there were 15 cycles of denaturing at 98 °C for 5 seconds and extension at 72 °C for 2 minutes.

Site directed mutagenesis
Site-directed mutagenesis was performed using the QuikChange Lightning Site-Directed Mutagenesis Kit (Stratagene 210518-5). Oligos were designed as indicated by the manufacturer. Reactions and protocol followed the manufacturer’s indication. PCR was run with initial denaturation at 95 °C for 2 minutes, 18 cycles of 95 °C for 20
seconds, 60 °C for 10 seconds and 68 °C for 5 minutes 30 seconds; and to finish a final extension at 68 °C for 5 minutes. After the PCR mix was treated with DpnI, 2 µl samples were used to transform RIK276. Primers used were 63/64 for W<sub>66</sub>, 65/66 for Stop<sub>66</sub>, 68/69 for H<sub>66</sub> and 70/71 for Y<sub>203</sub>.

**Allele specific PCR**

Colony PCR was performed using Taq 2X Master Mix (NEB M0270S). For each allele-specific PCR reaction, one of the primers in the pair is complementary at the 3’ end to the template sequence of one allele, to specifically amplify that allele sequence. As a control, the same primer pair was run with the other allele’s template. The annealing temperature was optimized to where amplification was selective for each sequence. Reactions were performed as recommended by the enzyme manufacturer.

<table>
<thead>
<tr>
<th>Target</th>
<th>Allele</th>
<th>Primers</th>
<th>Control</th>
<th>Primers</th>
<th>Product (bp)</th>
<th>TM (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pDual-eGFP</td>
<td>W&lt;sub&gt;66&lt;/sub&gt;</td>
<td>50/46</td>
<td>Y&lt;sub&gt;66&lt;/sub&gt;</td>
<td>49/46</td>
<td>526</td>
<td>70.8</td>
</tr>
<tr>
<td>pDual-eGFP</td>
<td>W&lt;sub&gt;66&lt;/sub&gt;</td>
<td>45/52</td>
<td>Y&lt;sub&gt;66&lt;/sub&gt;</td>
<td>45/51</td>
<td>221</td>
<td></td>
</tr>
<tr>
<td>pDual-eGFP</td>
<td>Stop&lt;sub&gt;66&lt;/sub&gt;</td>
<td>55/46</td>
<td>Y&lt;sub&gt;66&lt;/sub&gt;</td>
<td>49/46</td>
<td>526</td>
<td>72</td>
</tr>
<tr>
<td>pDual-eGFP</td>
<td>Stop&lt;sub&gt;66&lt;/sub&gt;</td>
<td>45/54</td>
<td>Y&lt;sub&gt;66&lt;/sub&gt;</td>
<td>45/53</td>
<td>213</td>
<td></td>
</tr>
<tr>
<td>pDual-eGFP</td>
<td>H&lt;sub&gt;66&lt;/sub&gt;</td>
<td>73/46</td>
<td>Y&lt;sub&gt;66&lt;/sub&gt;</td>
<td>72/46</td>
<td>526</td>
<td>72</td>
</tr>
<tr>
<td>pDual-eGFP</td>
<td>Y&lt;sub&gt;203&lt;/sub&gt;</td>
<td>45/75</td>
<td>T&lt;sub&gt;203&lt;/sub&gt;</td>
<td>45/74</td>
<td>625</td>
<td>70.8</td>
</tr>
<tr>
<td>pDual-eGFP</td>
<td>T&lt;sub&gt;203&lt;/sub&gt;</td>
<td>45/74</td>
<td>Y&lt;sub&gt;203&lt;/sub&gt;</td>
<td>45/75</td>
<td>625</td>
<td>70.8</td>
</tr>
<tr>
<td>pDual-eGFP</td>
<td>ST&lt;sub&gt;203&lt;/sub&gt;</td>
<td>45/199</td>
<td>Y&lt;sub&gt;203&lt;/sub&gt;</td>
<td>45/200</td>
<td>625</td>
<td></td>
</tr>
</tbody>
</table>

**PCR for sequencing**

All sequencing reactions were submitted to Genewiz following their protocols.

Primers 87-91 and 104-112 were used to sequence ICP8. ICP8 was hard to sequence. The best protocol they recommended was program P3. Primers 44-46 were used to sequence eGFP. Primers 5-7 were used to sequence GFPmut3*. Primers 126 and 127 were used to sequence LaminA and the Progeria reporter. Primers 76, 77 and 127 were used to
sequence dsRed. Primers 190-198 were used to sequence the pSLIK clones (including hum-Beta gene).

**Cloning**

For the cloning design NEBcutter V2.0 ([http://tools.neb.com/NEBcutter2/](http://tools.neb.com/NEBcutter2/)) was used to map and choose restriction sites. NEBcutter V2.0 and Snapgene V 1.5.3 ([http://www.snapgene.com/](http://www.snapgene.com/)) were used to generate plasmid maps. Cloning was performed generally using NEB restriction enzymes, phosphatase, T4 DNA Polymerase and T4 ligase under conditions indicated by NEB. Ligation mixes were usually transformed into chemically competent NEB 5α. Some other cloning strategies were used, like the Flexi System from Promega, the Invitrogen Gateway site-specific recombination system, the Gibson Assembly reaction from NEB and the In-fusion Dry-Down PCR system from Clontech as recommended by the manufacturer.

A common strategy when cloning a plasmid to express a gene in mammalian cells in this study was to add a fluorescent reporter gene to validate expression and to use a bacterial promoter to have a phenotypic clue that the gene was expressed. In cases where mammalian genes were expressed in *E. coli*, protein expression was evaluated in BL21(DE3) cells containing pRARE. In cases the gene of interest was toxic to *E. coli*, expression was evaluated in BL21(DE3) pLysS. When bacterial genes were expressed in human cells, genes were optimized for human codon usage and synthesized by Genewiz.

DNA was purified using Qiagen kits or by CsCl gradients. Plasmids to transfect mammalian cells were purified using the endotoxin free kit from Qiagen. DNA analyses were sometimes done using the toothpick screening protocol.

Sequence elements included in cloning designs:
Kozak sequence

A Kozak sequence (consensus (gcc)gccRecAUGG) was added at the 5’ end of genes to be expressed in human cells in order to optimize translation initiation.

Ref: http://en.wikipedia.org/wiki/Kozak_consensus_sequence

SD sequence

A Shine-Dalgarno (SD) sequence was added at the 5’ end of genes to be expressed in bacteria in order to optimize translation initiation if a decent ribosome binding site was not already present.

Linker between fused genes

To reduce the chance that peptide fusions interfered with the protein to which they were permanently fused, a flexible, disordered 15 amino acid peptide linker was occasionally included to spatially separate the fusion domain from the protein of interest. The amino acid sequence of this linker was typically GGGGSGGGGGSGGGS.

Alternatively, the P2A linker was used to produce two proteins from a single transcript. Picornaviruses use 2A peptides to mediate “cleavage” between two proteins between the glycine and the proline residue. We used the porcine teschovirus-1 sequence (P2A) because it was the shortest and most efficient of the 2A variants (Szymczak-Workman et al. 2012). The DNA sequence used to introduce the P2A linker was GGATCCGGAGCCACGAACTTCTCTCTGTTAAAGCAAGCAGGAGACGTGGAAAGAAAACCCCGGTCCT, encoding a 3 amino acid linker (GSG) followed by the P2A peptide itself: ATNFSLLKQAGDVEENPGP.
Poly A sequence at the 3’ end

To promote transcription termination, nuclear export, message stability and translation, a poly A signal sequence was included at the 3’ end of gene to be expressed in human cells.

Mitochondria localization sequence

Targeting of ICP8 to mitochondria was accomplished through incorporation of a Mitochondrial localization sequence (MSVLTPLLRRGLTSARRLPVPRAKIHSL).

Tags for Western blot and immunocytochemistry

HA: YPYDVPDYA

3xFLAG: dykdhdgdykdhdidykdddk

Codon usage tables

The degenerate genetic code is decoded by the tRNA pool. The relative concentration of different tRNA species varies between organisms and this has important implications for translation of “foreign” genes. Expression of E. coli proteins in human cells is improved by recoding bacterial or bacteriophage genes using the codon usage bias of humans. Likewise, human viral protein expression is improved in E. coli by expressing the genes in cells carrying extra copies of the rare tRNAs that are commonly used in translation of human genes.
Codon usage tables for *E. coli* K12

*Escherichia coli K12* [gbct]: 14 CDS's (5122 codons)

fields: [triplet] [frequency: **per thousand**] ([number])

<table>
<thead>
<tr>
<th>Triplet</th>
<th>Frequency per thousand</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>UUU</td>
<td>19.7 ( 101)</td>
<td>86</td>
</tr>
<tr>
<td>UUC</td>
<td>15.0 (  77)</td>
<td>75</td>
</tr>
<tr>
<td>UUA</td>
<td>15.2 (  78)</td>
<td>67</td>
</tr>
<tr>
<td>UUG</td>
<td>11.9 (  61)</td>
<td>54</td>
</tr>
<tr>
<td>CUU</td>
<td>11.9 (  61)</td>
<td>55</td>
</tr>
<tr>
<td>CUC</td>
<td>10.5 (  54)</td>
<td>54</td>
</tr>
<tr>
<td>CUA</td>
<td>5.3 (  27)</td>
<td>55</td>
</tr>
<tr>
<td>CUG</td>
<td>46.9 (  240)</td>
<td>41</td>
</tr>
<tr>
<td>AUU</td>
<td>30.5 (  156)</td>
<td>59</td>
</tr>
<tr>
<td>AUC</td>
<td>18.2 (  93)</td>
<td>59</td>
</tr>
<tr>
<td>AUA</td>
<td>3.7 (  19)</td>
<td>59</td>
</tr>
<tr>
<td>AUG</td>
<td>24.8 ( 127)</td>
<td>59</td>
</tr>
<tr>
<td>GUU</td>
<td>16.8 (  86)</td>
<td>59</td>
</tr>
<tr>
<td>GUC</td>
<td>11.7 (  60)</td>
<td>59</td>
</tr>
<tr>
<td>GUA</td>
<td>11.5 (  59)</td>
<td>59</td>
</tr>
<tr>
<td>GUG</td>
<td>26.4 ( 135)</td>
<td>59</td>
</tr>
</tbody>
</table>

**Note:** The fields provide the frequency of each codon per thousand codons and the number of occurrences.
Codon usage tables for *Homo sapiens*

*Homo sapiens* [gbpri]: 93487 CDS's (40662582 codons)

fields: [triplet] [frequency: per thousand] ([number])

<table>
<thead>
<tr>
<th>Codon</th>
<th>Frequency</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>UUU</td>
<td>17.6</td>
<td>714298</td>
</tr>
<tr>
<td>UCU</td>
<td>15.2</td>
<td>618711</td>
</tr>
<tr>
<td>UAU</td>
<td>12.2</td>
<td>495699</td>
</tr>
<tr>
<td>UGU</td>
<td>10.6</td>
<td>430311</td>
</tr>
<tr>
<td>UUC</td>
<td>20.3</td>
<td>824692</td>
</tr>
<tr>
<td>UCC</td>
<td>17.7</td>
<td>718892</td>
</tr>
<tr>
<td>UAC</td>
<td>15.3</td>
<td>622407</td>
</tr>
<tr>
<td>UGC</td>
<td>12.6</td>
<td>513028</td>
</tr>
<tr>
<td>UUA</td>
<td>7.7</td>
<td>311881</td>
</tr>
<tr>
<td>UCA</td>
<td>12.2</td>
<td>496448</td>
</tr>
<tr>
<td>UAA</td>
<td>1.0</td>
<td>40285</td>
</tr>
<tr>
<td>UGA</td>
<td>1.6</td>
<td>63237</td>
</tr>
<tr>
<td>UUG</td>
<td>12.9</td>
<td>525688</td>
</tr>
<tr>
<td>UCG</td>
<td>4.4</td>
<td>179419</td>
</tr>
<tr>
<td>UAG</td>
<td>0.8</td>
<td>32109</td>
</tr>
<tr>
<td>UGG</td>
<td>13.2</td>
<td>535595</td>
</tr>
<tr>
<td>CUU</td>
<td>13.2</td>
<td>536515</td>
</tr>
<tr>
<td>CCU</td>
<td>17.5</td>
<td>713233</td>
</tr>
<tr>
<td>CAU</td>
<td>10.9</td>
<td>441711</td>
</tr>
<tr>
<td>CGU</td>
<td>4.5</td>
<td>184609</td>
</tr>
<tr>
<td>CUC</td>
<td>19.6</td>
<td>796638</td>
</tr>
<tr>
<td>CCC</td>
<td>19.8</td>
<td>804620</td>
</tr>
<tr>
<td>CAC</td>
<td>15.1</td>
<td>613713</td>
</tr>
<tr>
<td>CGC</td>
<td>10.4</td>
<td>423516</td>
</tr>
<tr>
<td>CUA</td>
<td>7.2</td>
<td>290751</td>
</tr>
<tr>
<td>CCA</td>
<td>16.9</td>
<td>688038</td>
</tr>
<tr>
<td>CAA</td>
<td>12.3</td>
<td>501911</td>
</tr>
<tr>
<td>CGA</td>
<td>6.2</td>
<td>250760</td>
</tr>
<tr>
<td>CUG</td>
<td>39.6</td>
<td>1611801</td>
</tr>
<tr>
<td>CCG</td>
<td>6.9</td>
<td>281570</td>
</tr>
<tr>
<td>CAG</td>
<td>34.2</td>
<td>1391973</td>
</tr>
<tr>
<td>CGG</td>
<td>11.4</td>
<td>464485</td>
</tr>
<tr>
<td>AUU</td>
<td>16.0</td>
<td>650473</td>
</tr>
<tr>
<td>ACU</td>
<td>13.1</td>
<td>533609</td>
</tr>
<tr>
<td>AAU</td>
<td>17.0</td>
<td>689701</td>
</tr>
<tr>
<td>AGU</td>
<td>12.1</td>
<td>493429</td>
</tr>
<tr>
<td>AUC</td>
<td>20.8</td>
<td>846466</td>
</tr>
<tr>
<td>ACC</td>
<td>18.9</td>
<td>768147</td>
</tr>
<tr>
<td>AAC</td>
<td>19.1</td>
<td>776603</td>
</tr>
<tr>
<td>AGC</td>
<td>19.5</td>
<td>791383</td>
</tr>
<tr>
<td>AUA</td>
<td>7.5</td>
<td>304565</td>
</tr>
<tr>
<td>ACA</td>
<td>15.1</td>
<td>614523</td>
</tr>
<tr>
<td>AAA</td>
<td>24.4</td>
<td>993621</td>
</tr>
<tr>
<td>AGA</td>
<td>12.2</td>
<td>494682</td>
</tr>
<tr>
<td>AUG</td>
<td>22.0</td>
<td>896005</td>
</tr>
<tr>
<td>ACG</td>
<td>6.1</td>
<td>246105</td>
</tr>
<tr>
<td>AAG</td>
<td>31.9</td>
<td>1295568</td>
</tr>
<tr>
<td>AGG</td>
<td>12.0</td>
<td>486463</td>
</tr>
<tr>
<td>GUU</td>
<td>11.0</td>
<td>448607</td>
</tr>
<tr>
<td>GCU</td>
<td>18.4</td>
<td>750096</td>
</tr>
<tr>
<td>GAU</td>
<td>21.8</td>
<td>885429</td>
</tr>
<tr>
<td>GGU</td>
<td>10.8</td>
<td>437126</td>
</tr>
<tr>
<td>GUC</td>
<td>14.5</td>
<td>588138</td>
</tr>
<tr>
<td>GCC</td>
<td>27.7</td>
<td>1127679</td>
</tr>
<tr>
<td>GAC</td>
<td>25.1</td>
<td>1020595</td>
</tr>
<tr>
<td>GGC</td>
<td>22.2</td>
<td>903565</td>
</tr>
<tr>
<td>GUA</td>
<td>7.1</td>
<td>287712</td>
</tr>
<tr>
<td>GCA</td>
<td>15.8</td>
<td>643471</td>
</tr>
<tr>
<td>GAA</td>
<td>29.0</td>
<td>1177632</td>
</tr>
<tr>
<td>GGA</td>
<td>16.5</td>
<td>669873</td>
</tr>
<tr>
<td>GUG</td>
<td>28.1</td>
<td>1143534</td>
</tr>
<tr>
<td>GCG</td>
<td>7.4</td>
<td>299495</td>
</tr>
<tr>
<td>GAG</td>
<td>39.6</td>
<td>1609975</td>
</tr>
<tr>
<td>GGG</td>
<td>16.5</td>
<td>669768</td>
</tr>
</tbody>
</table>
Gateway reactions

Site specific-recombination was performed between pENTR2B/TREPitt clones and pSLIK-Zeo using LR Clonase II Enzyme Mix from Invitrogen (11791-020). The reaction was performed as recommended by the provider. Briefly, 100 ng of entry clone was mixed with 150 ng of the destination vector in a final volume of 8 µl in TE, pH 8.0. The LR Clonase was first thawed and mixed for 2 seconds, 2 times, and 2 µl were then added to the DNA. The reaction was mixed twice and incubated at 25 °C for 1 hour. The reaction was terminated with 1 µl of proteinase K solution, vortex and incubated at 37 °C for 10 minutes. 1 µl of the reaction was used to transform competent DH5α cells. After 30 minutes outgrowth cells were plated in the appropriate LB-agar plate (carbenicillin for pSLIK-Zeo). The next day, recombinant plasmids were isolated from the resistant clones.

Blunting ends

Sticky dsDNA ends were blunted with T4 DNA Polymerase (+ 100 µM dNTPs, 0.1 U/µg DNA T4 Pol, 100 µg/ml BSA). Reaction was incubated for 20 minutes at 12°C and promptly heat inactivated at 65 °C for 10 minutes. Enzymes were removed using the PCR cleanup kit (USB PrepEase).

Annealing oligos

Complementary oligos were hybridized by resuspending them in equimolar concentrations in 100 µl of annealing buffer (10 mM Tris, pH 7.5-8.0, 50 mM NaCl, 1 mM EDTA). The oligos were then boiled for 2 minutes, the heater was turned off and the oligos were left in the water until it reached room temperature. The annealed products were then stored at 4 °C.
Recovery of mailed DNA from filter paper

Mailed DNA on filter paper was recovered by submerging the paper disk in TE, which was used to transform NEB5α cells to produce more plasmid.

CsCl plasmid purification

This protocol was adapted from (Maloy 1990). Briefly, 2 ml LB with the appropriate drug was inoculated with a colony containing the plasmid DNA of interest. When the resulting culture was saturated, it was transferred to a flask containing 500 ml TB-drug and grown for ~16 hours. The culture was then moved to the cold room and let it be for 2-24 hours. Cells were then spun down at 5000 rpm for 10 minutes. Pellets were washed in saline solution by resuspension in 100 ml 0.85 % NaCl. Pellets were resuspended in 40 ml of lysis solution (25 mM Tris HCl pH 8, 10 mM EDTA, 50 mM glucose) and incubated for 5 minutes at room temperature. 80 ml of freshly prepared SDS/NaOH (1 % SDS, 0.2 N NaOH) was added to the suspensions and mixed by inverting the tubes. The suspensions were incubated on ice for 5-10 minutes. 60 ml of 7.5 M ammonium acetate was added and mixed by gently inverting the tubes. The suspensions were incubated on ice for 10 minutes. The tubes were centrifuged at 10,000 rpm for 30 minutes at 4 °C to pellet chromosomal DNA and cell debris. The supernatant was filtered through a Kimwipe into clean SS34 tubes. DNA was precipitated by adding 0.6 volumes of room temperature isopropanol, mixing by inverting. The precipitated DNA was pelleted by centrifuging at 12,000 rpm for 30 minutes at room temperature. Pellets were covered with ~10 ml 70 % ethanol and left on ice for 5 minutes. Tubes were centrifuged at 12,000 rpm for 5 minutes, the super was poured off and the pellets were allowed to dry in a vacuum chamber for ~10 minutes. When pellets looked semi-dry and did not smell like ethanol, 3
ml TE was added to resuspend the pellet, and then more TE was added to bring the volume to 32 ml. The solution was transferred to 50 ml falcon tubes and centrifuged at 2500 rpm for 10 minutes at room temperature. 32 g of CsCl was added and gently mixed. The suspension was warmed to 37 °C to facilitate the resuspension and centrifuged at 4000 rpm for 5 minutes at room temperature. 0.6 ml of 10 mg/ml ethidium bromide was added. The mixture was added to a medium size quick seal tube. 1 % CsCl in TE was used to fill the tube. The tubes were sealed, placed in a T70 rotor, balanced with another sample, covered with spacers and spun in the ultracentrifuge at 45,000 rpm for 40 hours at 20-25 °C. The rotor was transported back to the lab very gently. Tubes were put in a support. A UV lamp was used to illuminate the ethidium bromide stained DNA. Lab lights were off. The top of the tube was punched with a needle and the ethidium bromide-supercoiled DNA band was extracted with a syringe, punching right below the band with an angle pointing to the band. Samples were collected to tubes and mixed with equal volumes of Isobutanol saturated in TE to extract the ethidium bromide from the DNA in the dark. The contents were mixed and let it rest until phases separated. The pink upper phase with the isobutanol and ethidium bromide was discarded to a special waste container. The ethidium bromide extraction was repeated until the isobutanol was not pink any more. Samples were then dialyzed in 100 fold TE 3 times for 3-4 hours. DNA was precipitated using 1/2 sample volume of 7.5 M NH4OAc and 2.5 ml 95 % ethanol and incubated on ice for 30 minutes. DNA was pelleted at 10,000 rpm for 30 minutes at 4 °C, washed in 70 % ethanol, left to dry and resuspended in TE.
Toothpick plasmid screens

The toothpick screening protocol was used as a fast and inexpensive way to screen plasmids. This protocol was adapted from (Maloy 1990). Briefly, the cells were picked from a colony or from a pelleted culture. Cells were resuspended in 75 µl of DNA lysis solution (3 % SDS, 0.1 M Tris HCl pH 8, 120 mM NaOH) in a 0.5 ml microfuge tube and were incubated for 5 minutes at room temperature. Then, 75 µl of PCI (phenol: chloroform:isoamyl alcohol) was added in the chemical fume hood. The suspension was mixed by inverting the tube and it was then centrifuged at top speed for 3 minutes. The tubes were transferred to the bench very carefully to not disrupt the aqueous and organic phases. The DNA was in the aqueous phase on the top, from which 10 µl were transferred to another tube with loading dye for gel electrophoresis analysis.

Crude extracts for protein spectra

GFPmut3* protein and variants (Mar, Aqua, Violeta, Amarillo and Bronze) extracts were prepared from overnight E. coli cultures. Cells were lysed using BugBuster Protein Extract Reagent as indicated by the manufacturer (70584, Novagen). The lysate mixture also included 25 U/µl Benzonase Nuclease (70746, Novagen), 20 U/ml lysozyme and 0.1 M DTT. The mixture was incubated for 20 minutes at room temperature in a rocking platform. The mixture was centrifuged for 30 minutes at 4 °C, the supernatant was removed from the tube and centrifuged again. The supernatant was used to measure the spectra. If necessary, samples were diluted in TE containing 0.5 mM DTT. Excitation and emission spectra were determined using a PTI QuantumMaster spectrofluorometer with 2 nm slits. The average integrated one second samples detected ~10^4 -10^6 emitted photons. The data were collected using Felix (PTI), and analyzed using Excel (Microsoft) and
PRISM (GraphPad). Each spectrum was corrected for background, normalized to its peak value and plotted.

eGFP protein and variants (Celeste, Azure and Mostaza) extracts were prepared from overnight E. coli cells transfected with the specific pDual plasmid. Briefly, RIK276 cells were transformed with the specific pDual variant and plated overnight on LB-carbenicillin-chloramphenicol plates. LB-carbenicillin-chloramphenicol broth was inoculated with a fresh colony and let to grow overnight. Cells were lysed using BugBuster Protein Extract Reagent as indicated by the manufacturer (70584, Novagen). The lysate mixture also included 25 U/µl Benzonase Nuclease (70746, Novagen), 20 U/ml lysozyme and 0.1 M DTT. The mixture was incubated for 20 minutes at room temperature in a rocking platform. The mixture was centrifuged for 30 minutes at 4 °C, the supernatant was removed from the tube and centrifuged again. The supernatant was used to measure the spectra. If necessary, samples were diluted in TE containing 0.5 mM DTT. The excitation and emission spectra of eGFP fluorescence and its variants in the extracts were determined with a spectrofluorometer (PTI QuantumMaster) with 2 nm slits at a slew rate of 1 nm/second (the average integrated one second samples collected ~10^4 - 10^6 emitted photons). The data were collected using Felix32 software, and later analyzed using Microsoft Excel and GraphPad Prism. Each spectrum was normalized to its peak value and plotted.

The pPRO1.1go HeLa cell line carrying the Progerin Reporter expressed both dsRed and eGFP. Protein extracts from this cell line were prepared as reported by Paillusson et al. 2005. Briefly, proteins were extracted from HeLa cells in 1% deoxycholic acid, 1% NP40, 25 mM Tris–HCl (pH 8), 50 mM NaCl and protease inhibitor cocktail from Sigma
The cells were incubated at 4 °C for 30 minutes on a shaker. The mixture was scraped from the dish and transferred to a tube. Samples were sonicated (at power 7 for 1 minute on and 1 minute off for 3 pulses) and then centrifuged at 16,000 g for 5 minutes. The excitation and emission spectra of eGFP and dsRed fluorescence in the extracts were determined with our spectrofluorometer (PTI QuantumMaster) with 2 nm slits at a slew rate of 1 nm/second (the average integrated one second samples collected ~10⁴-10⁶ emitted photons). To collect eGFP excitation spectra, the excitation monochromator scanned from 430 to 530 nm and emission was monitored at 540 nm. To collect the GFP emission spectra, samples were excited at 482 nm, and the emission spectra were collected from 490 to 690 nm. To collect dsRed excitation spectra, the excitation monochromator scanned from 430 to 590 nm and the emission was monitored at 650 nm. To collect the dsRed emission spectra, samples were excited at 530 nm, and the emission spectra were collected from 540 to 700. The data were collected using the Felix software, and later analyzed using Microsoft Excel and GraphPad Prism. Each spectrum was normalized to its peak value and plotted.

**Protein-Protein binding assay**

Protein-protein interactions were analyzed by formation of stable complexes between untagged proteins (λExo, SSB, RecA) with His-β or with His-λExo (Spies & Kowalczykowski 2006). 1 µM His-β protein was mixed with 1 µM each of λ Exo (R28A), RecA and SSB in all possible combinations in a binding buffer (50 mM Tris Acetate pH 7.5, 50 mM NaCl, 30 mM Imidazole, 0.5% Tween 20) in a total volume of 100 µl. The mix was incubated for 15 min at 37°C. 20 µl of 5% Ni-NTA magnetic agarose beads (Qiagen) were washed with the binding buffer and then mixed with the
protein solution. Beads and proteins were incubated for 1 hour at room temperature. His-tagged proteins and associated proteins were pulled down by placing the tubes with beads and proteins in a magnetic tube holder. Unbound fractions were carefully removed from the tubes and later analyzed by SDS PAGE. Beads were washed twice with binding buffer and then the bound proteins were eluted with elution buffer (150 mM Tris Acetate pH 7.5, 100 mM NaCl, 300 mM Imidazole, 0.5% tween 20). Bound fractions were carefully removed from the tubes and later analyzed by SDS PAGE in parallel with the unbound fraction. His-λExo protein was analyzed also as its ability to form complex with RecA and SSB in all possible combinations as described above when His-β protein was used as the bait. SDS PAGE was conducted using a Criterion cell and precast 10-20 % Tris-Glycine Gels (Bio-Rad-345-0067), run at room temperature and 0.03 Amps for 4 hours. Two buffers were used for the electrophoresis, the Cathode Buffer (Top) (0.1 M Tris, 0.1 M Tricine and 0.1 % SDS) and Anode Buffer (Bottom) (0.2 M Tris-Cl pH 8.9). Gels were stained for 1 hour in stain solution (10 % acetic acid, methanol (40%), and 0.25 % Coomassie Blue G-250), rinsed and distained 10 % acetic acid several times until bands were visible and background clear. Gels were visualized using a scanner.

**Real - time DNase assay (ReDA)**

The ReDa assay used to study the effect of ssDNA binding proteins on λExo activity was performed as in (Tolun & Myers 2003). Reaction time courses were recorded on the spectrofluorometer (PTI QuantumMaster) using the following settings: the excitation wavelength was 484 nm and emission was monitored at 522 nm with 5-6 nm slit widths and the data sampling frequency was 1 s⁻¹ and the assay period, from 600 - 1500 s, depending on the experiment. All reactions were carried out at the ambient temperature
inside the sample chamber of the spectrofluorometer (~22 – 25 °C). Reaction mixtures (2 ml) consisted of 50 pM PstI linearized pUC19::ΔKpnI molecules (100 pM ends) or 50 pM XhoI linearized pKIS1 molecules (100 pM ends), 1/20 000 diluted PicoGreen®, 1× buffer (50 mM HEPES pH 7.5, 1 mM DTT), 5 mM MgCl₂ and the proteins and enzymes as indicated in each experiment, at the following concentrations: 13 nM λExo monomers, 133 nM β (when β was used at a concentration that saturates ssDNA at the limit digest), 13 nM β (when β was equimolar to λExo), 624 nM Rec A, 200 U (14.6 nM) ExoI. Buffer, enzymes, substrate and PicoGreen® were mixed in a 3 ml (1cm per side) quartz cuvette (Starna) and immediately placed in the fluorometer chamber in the dark with continuous stirring. Data collection was started and after 60 s, MgCl₂ was added through the injection port by means of a Hamilton syringe to initiate λExo activity, and the fluorescence output was continuously measured. To determine processivity, 23.4 mg/ml Heparin (determined sufficient to bind all the λExo under these conditions, data not shown) was added after 100 seconds through the injection port using a Hamilton syringe as well.

As the negative control, a reaction mixture without the enzyme was prepared and read in the fluorometer. The positive control for complete digestion of dsDNA by λExo was prepared as follows: Substrate DNA was boiled for 10–15 min, immediately transferred into ice/water and let to stand for at least 3 min before being added into a reaction mixture to a final concentration equivalent to half the concentration of the starting material to mimic 100% digestion of the substrate by λExo (λExo takes two mol equivalents of ssDNA as one mol of dsDNA and renders it one mol equivalent of ssDNA and the dNMPs equal to one mol equivalent of ssDNA. The dNMPs were previously
shown not to contribute to the signal, data not shown). This solution was then read in the fluorometer. Three samples of each substrate were scanned and the average values for each time point in the scans were determined. The average differential fluorescence of the dsDNA substrate and of half the mass of the dsDNA substrate, denatured, sets the bandwidth of the assay and was used to convert fluorescence into nucleotides released.

The raw fluorescence data were manipulated as follows. First, fluorescence data were corrected for photobleaching by normalizing the signal to the corresponding points in the dsDNA control scan. Second, corrected fluorescence was converted to nucleotides released per complex by dividing the bandwidth of the assay per half the number of nucleotides of the linear substrate for each point:

**Equation 2.1 Nucleotides released per complex**

\[
\text{Nucleotides released per complex} = \frac{(\text{Fluorescence of dsDNA} - \text{Fluorescence of ssDNA})}{(1/2 \times \# \text{ nt in the substrate})}
\]

We determined the Change in Fluorescence for each data point measured at one second intervals by subtracting the sample fluorescence from the average fluorescence of each corresponding point of the negative control dsDNA:

**Equation 2.2 Change in fluorescence**

\[
\text{Change in fluorescence} = \text{Fluorescence of dsDNA} - \text{Fluorescence of sample}
\]

Finally the total number of nucleotides released per complex up to each point in the reaction corresponds to:
Equation 2.3 Total number of nucleotides released per complex

\[ = \Sigma \text{nucleotides released up to each time point} = EQ1/EQ2 \]

Then, the data were truncated to eliminate the non-linear part of the progress curve, including the spikes due to insertion of the needle with MgCl₂ or Heparin. The data points were then normalized so that all the initial values were the same arbitrary value of “0”. To calculate the rates, the initial linear parts of the curves were fit to a linear equation to approximate the first derivative of the progress curve in the steady-state. The slope in the equation gives the rate of the reaction in terms of nucleotides digested per second. Processivity was obtained from the reactions where Heparin was injected as the extent of digestion, either averaged over the last 50-100 seconds of the progress curve or calculated by fitting the progress curve to a single exponential equation and solving for the plateau:

Equation 2.4 Plateau

\[ = Po + (\text{Plateau} - Yo) \times (1-e^{-kt}) \]

The dwell time was then determined by dividing the processivity by the rate of digestion:

Equation 2.5 Dwell time

\[ = \text{Processivity/rate} \]
Halo tag fused proteins purification

Genes for four SynExo recombination proteins (β, β(1-177), Exo and Exo(1-189)) were cloned as fused to a Halo tag in the pFN22K Flexi vector, expressed independently in *E. coli* strain BL21(DE3), and purified as complexes with associated *E. coli* host factors following the Promega HaloTag Purification protocol, except the lysates were prepared using Bugbuster HT (Novagen, 70922; contains Benzonase) and Lysozyme as recommended by the vendor with the addition of one tablet complete protease inhibitor (Roche, 05892970001) to the lysis buffer to minimize proteolysis. ATP was added to 2 mM and MgCl$_2$ (or MgSO$_4$) was added to 10 mM to lysates to dissociate bacterial chaperoning from partially folded proteins. Flowing clarification by centrifugation (10,000 x g for 15-30 minutes at 4 °C), lysates were diluted with an equal volume of ice-cold HaloTag purification buffer (50 mM Hepes, pH 7.5, 150 mM NaCl) and placed on ice.

Mass spectrometry: sample preparation and analysis

Genes for four SynExo recombination proteins (β, β (1-177), Exo and Exo (1-189)) were cloned fused to a Halo tag in pFN24 vectors, expressed independently in *E. coli* strain BL21(DE3), and purified as complexes with *E. coli* host factors following the vendor protocol and as described above. 10 µg of the proteins co-purifying with SynExo recombination proteins were fractionated by SDS PAGE. The gel was then washed 3 times with double distilled water. Gels were stained overnight with rocking using Gel code blue stain reagent. Gels were then washed 3 times with double distilled water. Gels were then scanned and all the bands marked. The bands were cut and the resulting gel pieces transferred to an ELISA plate well with tweezers. As a control, a piece of the gel
without protein and a piece of gel from a known protein were added. 100-150 µl of solution 1 (CH₃CN/H₂O (1:1)) was added to the wells containing gel slices. The plate was centrifuged for 5-10 seconds and shaken overnight at room temperature or repeated in 30 minutes intervals until the blue color of the gel disappeared. The solution was pipetted out. 100-150 µl of solution 3 (CH₃CN/100 mM NH₄HCO₃ (1:1)) was added to the wells to create an alkaline environment. The plate was centrifuged for 5-10 seconds and shaken for 30 minutes at room temperature. The solution was pipetted out. 80 µl of solution 4 (10 mM DTT/100 mM NH₄HCO₃) was added to the samples, the plate was centrifuged for 5-10 seconds and shaken for 45 minutes sharp at room temperature. Solution 4 breaks and reduces disulfide bonds between cysteine residues. Liquid was removed and quickly replaced with 80 µl of freshly prepared solution 5 (55 mM iodoacetamide/100 mM NH₄HCO₃) added in the dark. Solution 5 alkylates the reduced cysteine residues. The plate was incubated for 30 minutes sharp at room temperature in the dark and the solution was removed entirely. 100-150 µl of solution 7 (CH₃CN/30 mM NEM (1:1)) was added to the samples. The plate was centrifuged for 5-10 seconds and shaken for 30 minutes at room temperature. The solution was removed and 20 µl trypsin solution 9 (0.1 mg/band) was added and incubated overnight at 37 °C. Samples were transferred to a new plate where 50 µl of solution 10 (0.1 % TFA/60 % CH₃CN) was added. The plate was centrifuged for 5-10 seconds and shaken at room temperature for 30 minutes. The solution was placed in a new ELISA plate, while a new extraction from the gels was performed. 50 µl of pure CH₃CN was added to ELISA plate with gels. The plate was centrifuged for 5-10 seconds and shaken for 30 minutes at room temperature. The solution was transferred to the new ELISA plate and samples were dried in a vacuum
concentrator (SpeedVac; Thermo Fisher Scientific). 150 µl of solution 11 (CH₃CN:H₂O (1:1)) was added to resuspend the proteins. Proteins were pipetted up and down ~5 times and transferred to the glass capped tubes designated for mass spectrometry. Samples were separated by HPLC and fed into a LCQ Deca XP electrospray ionization/ion trap mass spectrometer (MS) for analysis.

Bioinformatic tools used to identify proteins included a MASCOT query of the BL21(DE3) peptide mass/charge predication database (http://www.matrixscience.com/search_form_select.html) and follow-up using MS tools at EXPASY (http://expasy.org/tools/) to confirm assignments employed to identify putative host-specificity factors for SynExo-mediated Recombineering. Proteins found in complexes with one or more SynExo proteins were evaluated in the STRING database of protein interactions in BL21(DE3) (http://string-db.org/) and by text mining in PUBMED (http://www.ncbi.nlm.nih.gov/pubmed) to extend the interaction network to include proteins known or predicted to interact with the SynExo indirectly through other proteins or determined to functionally interact through genetic and enzymological studies.

**Western blots to evaluate recombinase expression**

Western Blots were performed following the Abcam “Western Blotting - A beginner’s guide” protocol. Briefly, 293T cells were transiently transfected with lentiviral vectors in the presence and absence of Dox (1000 ng/ml). 24 hours later, cells were lysed over ice with lithium dodecyl sulfate loading buffer and 1/100 dilution of Protease inhibitor cocktail from Sigma (P8340). Samples were scraped and collected from plates. Samples were sonicated 5 times at power 7, 1 minute on, 1 minute off, in the Sonicator 3000 Misonix. Extracts were incubated on ice for 10 minutes and centrifuged at 14,000
rpm for 20 minutes. The samples were transferred to other tubes and stored at -20 °C. Protein concentrations were determined using the Pierce 660 Protein assay reagent (22660) with BSA as a control. ICP8 samples were separated on an 8 % SDS PAGE gel and the Beta samples were separated on a 12 % SDS PAGE gel at 100 V, open current. Together with the samples two protein ladders were run: the Precision Plus Protein™ Kaleidoscope™ pre-stained ladder (Biorad, #161-0375) for visualizing protein migration in the gel and the protein transfer from the gel to the PVFD membrane and the Maggic marker XP (Live Technologies, LC5602) where each protein has a IgG binding site which is recognized by the secondary antibody and the ladder is visualized with the chemiluminescent reagent and on the film. Proteins were transferred to a methanol activated PVFD membrane for one hour at 100 V, open current or overnight at 30 V and 90 mA in transfer buffer (25 mM Tris, 192 mM glycine, 10% methanol, pH 8.3). Protein transfer was corroborated by the transfer of the Kaleidoscope pre-stained ladder, as well as by examining the lack of protein after staining the gel with Coomassie brilliant blue dye. The membrane was blocked with 5 % fat-free milk-TBST and incubated with primary Herpes Virus I ICP8 Major DNA Binding Protein antibody [1OA3](ab20193) Mouse monoclonal IgG₁ in 5% milk 1/2000 and rabbit polyclonal anti HA antibody 1/10000 (Abcam ab9110). Both membranes were incubated overnight at 4 °C with shaking. Membranes were washed three times in TBST for 20 minutes and incubated with secondary HRP labeled antibodies, goat anti-mouse IgG₁ (Santa Cruz, sc-2064) and goat anti rabbit IgG (Santa Cruz, sc-2004) in 5 % milk 1/5000 for 2 hours at room temperature shaking. The membrane was washed three times in TBST for 20 minutes and incubated with the ECL Plus Western Blotting Detection System (from GE Healthcare,
RPN2132) for 5 minutes or until the bands started glowing in the dark. A Magic Marker XP (Life Technologies, LC5602) was used as an IgG binding site. Films were exposed to membranes for varying amounts of time and then developed. Films were scanned and analyzed using Photoshop or Image J software.

The Rheoswitch HeLa 11 cell line was transfected with pNEBRX1-Hygro-UL12.5-HA using the Transpass reagent supplied in the Rheoswitch kit. Expression of UL12.5-HA (55 kDa) was induced with 500 nM of the Rheoswitch Ligand (RSL). Cells were collected 24, 48 and 72 hours after induction and sonicated 3 times. Protein extracts were run in 10% SDS PAGE gels at 80 V. Proteins were transferred to PVDF membranes in a semi-dry transfer for 2 hours at 23 V. Membranes were incubated with primary polyclonal HA antibody (Roche) diluted 1/1000 in Odyssey buffer (LI-COR, 927-40100). Membranes were washed with PBS for 2 hours at room temperature with shaking. Membranes were then washed with PBS Twin for 15 minutes and incubated with the secondary anti-rat IRDye (LI-COR) antibody diluted 1/3000 for 1 hour at room temperature in the dark. Membranes were then washed 3 times in PBS and exposed to x-ray film. 48 hours of induction produced the most amount of protein.

**Imaging E. coli**

*E. coli* expressing the gfpmut3* variants were grown to OD<sub>600</sub> of 0.4, pelleted and resuspended in minimal medium OMBG (1.05 % K<sub>2</sub>HPO<sub>4</sub>, 0.45 % KH<sub>2</sub>PO<sub>4</sub>, 0.005 % MgSO<sub>4</sub>.7H<sub>2</sub>O, 0.1 % (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0.05 % sodium citrate and 0.2 % glucose) to a 10 % of the initial volume. 1 µl of the concentrated cells were added to a slide with mounting medium Prolong Antifade (Molecular Probes, P7481). The slide was covered and left to sit overnight before imaging. Fluorescent images of individual *E. coli* expressing
GFPmut3*, Amarillo and Aqua were collected with a Zeiss LSM710 confocal microscope. PerkinElmer/Improvision Velocity 64-bit software was used to create a maximum projection and Adobe Photoshop CS3 was used to merge the colors. Fluorescent images of individual *E. coli* cells expressing GFPmut3* and Aqua was collected with a DeltaVision microscope, using an Olympus 60X objective, a CoolSNAP_HQ / ICX285 camera and 3D deconvolution software. *E. coli* GFPmut3* was imaged using FITC emission and excitation filters. *E. coli* Aqua was imaged using CFP emission and excitation filters.

*E. coli* Recombineering experiments were imaged from agar petri plates.

Fluorescent images were taken with a Nikon Diaphot microscope or using a Dark Reader transilluminator (Clare Chemical, excitation 420-500 nm, emission >520 nm) and a CCD camera.

**Immunocytochemistry and imaging human cells**

Human cells were seeded onto coated coverslips in 6 well plates and media, poly-L-Lysine (Sigma, P9155) for 293T cells or fibronectin for MIAMI cells. When cells were ready for imaging, cell were washed 3 times with PBS and then fixed in 4 % paraformaldehyde in PBS pH 7.4 for 15 min at room temperature. Cells were washed 3 times with PBS and permeabilized with 0.25 % Triton X-100 for 10 min. Cells were washed again and blocked with 1 % BSA, 0.3 M glycine in PBST for 30 min. Cells were incubated with the primary antibody in 1 % BSA in PBST in a humidified chamber overnight at 4 °C. Cells were washed 3 times with PBS and incubated with the secondary antibody (which were labeled by Alexa Fluor) in 1 % BSA for 1 hr at room temperature in the dark. Cells were washed and incubated with 0.5 µg/ml DAPI for 10 min. Cells
were washed, mounted with Prolong Antifade or Vectashield (Vector Laboratories, Inc. H-1000). Cells were viewed with different microscopes including a Nikon Diaphot equipped with a Retiga 1300 camera. A Nikon 20X objective was used. Images were collected and analyzed using different softwares, including the IP-Lab software package.

Nuclear delivery of oligos to human cells was evaluated in MIAMI cells. Briefly, MIAMI cells were transfected with 7.5 µg of FITC-labeled oligos with Lipofectamine 2000 and incubated for 5 hours. Cells were fixed and mounted with DAPI (4',6-diamidino-2-phenylindole) Vectashield mounting media. The images were taken using a Leica SP5 confocal microscope with a 20 X objective.

Distinction between cells expressing the Mostaza Recombineering target and the Green fluorescent protein recombinant was evaluated in MIAMI cells. Briefly, MIAMI cells transduced with pDual-egfp and pDual-Mostaza were grown on Fibronectin-coated cover slips. Before imaging, cells were fixed and stained with DAPI. Slides were mounted using Prolong Antifade. Image acquisition was performed with a Leica SP5 confocal microscope using a 63X oil-immersion 1.3 numerical aperture objective. Confocal acquisition parameters were determined at the beginning of the study using the same parameters (e.g., gains, slit aperture, laser intensity) as for all the images. Confocal optical sections were 0.6 µm thick. Field selection was performed using the DAPI channel to identify MIAMI cells. Adobe Photoshop was used to merge the colors.

Expression and localization of Recombinase ICP8-GFP from pCMV-ICP8-GFP was evaluated in MIAMI cells. Briefly, MIAMI cells were transfected with pCMV-ICP8-GFP using the Nucleofector II electroporator (Lonza). 16 hours later cells were fixed with
4 % PFA, stained with DAPI and mounted with Prolong Gold. Images were taken with a Delta Vision fluorescent microscope.

Expression and localization of Recombinase ICP8-GFP from pCMV-ICP8-GFP and Recombination target Celeste were evaluated in 293T cells. Briefly, 293T cells were grown on coverslips covered with Poly-L-lysine. The next day, cells were transfected with pDual-eGFP variants and pCMV-ICP8-GFP using Fugene 6. 24 hours later, slides were washed with PBS and mounted in ProLong® Gold antifade reagent. 293T/pDual-eGFP(Y67W) images were taken on the DRI Leica SP5 inverted confocal microscope with a 450 nm laser and a 60 X objective. 293T/pCMV-ICP8-GFP images were obtained on an OMX Nanoscope. Expression and localization of Recombinase ICP8 and ICP8-GFP were also evaluated from the inducible Rheoswitch system in 293T. Briefly, 293T were co-transfected with pNEBR-R1 and pNEBR-X1-ICP8. 24 hours later inducer (RSL) was added to cells in the top panel. After 48 hours incubation, cells were processed for immunofluorescence. ICP8 protein was detected using an anti-ICP8 antibody (red) and nucleus was stained with DAPI. Expression and localization of Recombinase ICP8 and HumBeta were also evaluated from the inducible lentivirus plasmid pSILK in 293T. Briefly, 293T cells transiently transfected with the pSLIK vectors and incubated with 1 µg/ml Dox, were seeded on poly-L-Lysine treated coverslips in 6 well plates. Cells were then fixed and permeabilized. Cells were incubated with primary mouse monoclonal anti-HSV1 ICP8 antibody or with rabbit polyclonal anti-HA antibody (1/200 or 1/1000 in 1 % BSA respectively). Cells were incubated with the secondary antibody Alexa Fluor 594. Cells were incubated with DAPI. Cells were mounted with Vectashield. Cells were
imaged with a Nikon Diaphot equipped with a Retiga 1300 camera. A Nikon 20X objective was used. Images were collected with the IP-Lab software package.

The HeLa Progerin reporter cell line obtained from Dr. Misteli at NIH was imaged and analyzed using the Celigo (Cyntellec). Cells were seeded in a Corning 96 Well Flat Clear Bottom Black Polystyrene TC-Treated Microplates. The medium was changed to Hanks’ balanced salt solution from (HBSS) for imaging. Images and data were recorded using the Cyntellec Celigo adherent cell cytometer. Green fluorescence was imaged with filters 483 for excitation and 536 for emission. Red fluorescence was imaged with filters 531 for excitation and 629 for emission. Images were collected with a 4 Mega-pixel CCD camera. Expression and localization of Recombinase ICP8-GFP from pCMV-ICP8-GFP was also evaluated in the HeLa Progerin reporter cell line. Briefly, the HeLa Progerin reporter cell line was sorted to isolate a derivative of the HeLa Progerin reporter that expresses DsRed but not GFP. This strain was then transfected with 1, 2 and 5 µg of pCMV-ICP8GFP using the Amaxa Nucleofector II and and a program recommended by Amaxa (0.005) for high viability during transfection of HeLa cells. Cells were seeded in glass bottom chambers covered with Poly-L-Lysine. Eight hours later, cells were visualized in vivo using a Zeiss confocal microscope LSM 700. The objective used was EC Plan-Neofluar 40X/1.30 Oil DIC M27. Images were collected with the ZenLE software. AxioVisionLE was used to process images.

The Rheoswitch HeLa 11 cell line was transfected with pNEBRX1-Hygro-UL12.5-HA using Fugene 6 transfection reagent. Expression of UL12.5-HA was induced with 500 nM Rheoswitch Ligand (RSL) for 48 hours. Cells were then incubated with 200 nM red Mitotracker for 30 minutes in complete media. Cells were fixed and permeabilized
with cold methanol for 5 to 8 minutes at -20 °C. Cells were incubated with primary mouse polyclonal HA antibody (Roche) diluted 1/200 in 2% BSA overnight at 4 °C. Cells were washed and incubated with the secondary anti-rat Alexa green 488 antibody diluted 1/200.

**Flow cytometric analysis**

About 10^7 cells were harvested, washed with PBS and resuspended in 0.5 ml DMEM^{gfp} or PBS. Just before flow analysis, cells were vortexed and filtered. 3 flow cytometers were used during these studies:

1. BD LSR-Fortessa-HTS
2. BD FACS Aria-II
3. BD Accuri C6

**Bacteria**

*E. coli* cells were harvested and resuspended in 1.05 % K_2HPO_4, 0.45 % KH_2PO_4, 0.005 % MgSO_4•7H_2O, 0.1 % (NH_4)_2SO_4, 0.05 % sodium citrate supplemented with 0.1% thiamine and 0.2% glucose and then analyzed using a BD FACS Aria-II. *E. coli* GFPmut3* cells were evaluated with the FITC filter and *E. coli* Aqua cells were evaluated with the Alexa Fluor 430 nm filter.

**Mammalian cells**

pDual-eGFP variants were first evaluated at the BD LSR-Fortessa-HTS. A UV laser of 350 nm was used to analyze the blue cells (expressing eGFP(Y66H)) and the blue laser of 488 nm was used to analyze the Mostaza and green cells. The data were analyzed using FACSDiva Version 6.1.3 software. Mostaza to Green cell Recombineering studies were quantified using the Accuri flow cytometer. Samples were analyzed with the Accuri
cytometer using a 488 nm blue laser and filters with narrow band passes: 510 (15 nm band pass) for eGFP and 540 (20 nm band pass) for eGFP(T204Y) for the Recombineering experiments. The software was programmed to collect 50,000 cells at the lowest speed for 293T and at the highest speed for MIAMI. Data from the Accuri was analyzed using the C-flow software, Excel spreadsheets and GraphPad PRISM. Green recombinants were sorted with the BD LSR-Fortessa-HTS cytometer. Samples were sorted using a 488 nm blue laser and filters with narrow band passes: HQ510 (20 nm band pass) for eGFP and 550 (30 nm band pass) for eGFP(T204Y). These sets of filters were combined with the dichroic mirror Q495lpxr_11.25 and Q525lpxr_11.25. Gfp expression studies used to estimate transfection efficiency were performed on the Accuri flow cytometer using the 488 nm blue laser and the standard FL1 filter for GFP 533 (30 nm band pass). Crimson fluorescent protein expression was evaluated at the Accuri flow cytometer using the 633 nm red laser and the standard FL4 filter 675 (25 nm band pass). Propidium Iodide staining for viability assays were evaluated with the Accuri flow cytometer using the 488 nm blue laser and the standard FL3 filter 670 nm LP. HeLa Progerin studies were analyzed using a BD FACS Aria-II. eGFP fluorescence was excited with a 488 nm laser and emission was measured with a 525 (50 nm band pass) band pass filter. dsRed fluorescence was detected by excitation with 532 nm laser and emission was measured with a 610 (20 nm band pass) filter. HeLa 11 was used as a control for autofluorescence. A Green’ Red’ derivative of the Progerin reporter isolated in this study was used to set the gates for quantifying Recombineering efficiency. The data were processed using Facs Diva 6.1.3 and FlowJo software (Tree Star, Ashland, OR).
The geometric mean fluorescence values were used to compare the amount of fluorescence in the different samples.

**Cell viability**

Cell viability analysis were evaluated by staining dead cells with Propidium Iodide. Briefly, about $10^6$ 293T cells were harvested, washed with PBS and resuspended 0.5 ml PBS. 20 µl of propidium iodide solution (0.5 mg/ml propidium iodide in PBS, pH 7.4) was added to cell suspension just 10 minutes before flow analysis. During the 10 minutes incubation time, cells were kept on the dark on ice. Just before flow cells were vortexed and filtered. As a control for all dead cells, cells were fixed in 70 % ethanol, prior to propidium iodide staining. The Propidium Iodide viability assays were done in the Accuri flow cytometry. The standard filters were used to detect GFP (530 ± 15 nm) and to detect propidium iodide (670 nm LP).

**Cell cycle analysis**

Cell cycle analysis was performed using a protocol from http://www.meduniwien.ac.at/user/johannes.schmid/PIstaining3.htm. Briefly, $10^5$ to $10^6$ cells were collected and the pellets were suspended in 1 ml PBS. Cells were fixed with ethanol at a final concentration of ~ 70 %, by pipetting the cell suspension into 2.5 ml absolute ethanol while manually vortexing the ethanol tube. The cell suspension was incubated on ice for 15 minutes or overnight at -20 °C. Cells were collected by centrifugation at 1500 rpm for 5 minutes. Cells were then stained by resuspending them in 500 µl Propidium Iodide -solution (50 µg/ml Propidium Iodide, 0.1 mg/ml RNase A, 0.05 % Triton in PBS) and incubating them for 40 minutes at 37°C. Optionally, 3 ml PBS was added, and cells were collected by centrifugation at 1500 rpm for 5 minutes. The
pellets were suspended in 500 µl PBS for flow analysis. Accuri analysis was done as indicated in Figure 2.1. Briefly, first cells were plotted FSC-A vs SSC-A to gate the cell population. Gated cells were from the first plot then plotted on FL3-H vs FL3-A to gate for single cells. Gated cells from the second plot were then plotted on a FL3-A vs count to look at cell cycle patterns or DNA intensity per cell.

Figure 2.1 Cell cycle analysis with propidium iodide using BD Accuri gates

This plot looks at the cell population and is gated to remove debris. This plot is used to remove doublets from the analysis. This plot looks at the cell cycle data, and uses roughly placed markers to estimate percentages. FL3-A works as well.
Rationale for fluorescent protein engineering

We were interested in finding a sensitive and quantitative reporter for studying Recombineering in human cells. Previous studies have used selection for drug resistance ($G418^S \rightarrow G418^R$, Zhang et al. 2003), repair of a $lacZ$ gene or a screen for acquired fluorescence (Pierce et al. 2003). Selection for drug resistance is very sensitive and permits detection of rare recombinants (on the order of $10^{-6}$ for mammalian cells and down to about $10^{-8}$ for bacteria). There is also another advantage to drug selection. Estimates of recombination frequencies could be more accurate, as it is possible to determine the recombination frequency by comparing the number of resistant clones to the total number of viable cells as in (Gao & Knipe 1992; Rios et al. 2012). This is especially an issue if the temporary growth arrest sometimes observed in recombinants (Rios et al. 2012) is more common than appreciated. In assays based on phenotypic changes that do not select for recombinants, temporary growth arrest of recombinants will lead to underestimation of recombination frequencies due to continued growth of the non-recombinant population (the arrested recombinant titer would get diluted by the growing non-recombinant titer). In such assays, recombination frequencies could vary with the elapsed time after transfection before recombinants are scored. While selection for drug resistance allows recovery of rare positive clones and may be a closer indicator of the frequency of recombinants in some circumstances, non-recombinants are killed by this method, which can artifactually decrease the viability of recombinants by sending pro-apoptotic or pro-necrotic signals in the media that lead to death of resistant
recombinants through “collateral damage” (Galluzzi et al. 2012; Lenna & Trojanowska 2012; Tanjore et al. 2013). In addition, the recombination frequency must be determined by separately tittering resistant colonies and total colonies which is not as statistically powerful as enumerating recombinants and non-recombinants together in the same interrogated sample due to sampling errors.

Repair of the lacZ gene has also been used to assess recombination rates (Pluta et al. 2005; Pierce et al. 2003; Igoucheva et al. 2006). The inconvenience of this method is that cells need to be histochemically stained with X-gal to enumerate recombinants. Additional inconveniences associated with the Lac+ selection method include that microscopic quantification is less statistically powerful than flow cytometry as the sample size is typically much lower, histochemical staining is time consuming, and cells need to be killed and fixed to identify recombinants. In addition, use of Lac+ detection in mammalian Recombineering is hampered by endogenous mammalian β-galactosidase so the protocol might produce false positive unless carefully controlled for environmental factors such as pH.

In contrast, screening for fluorescent recombinants among non-recombinant cells is statistically stronger than is selection of any type, does not require unusual growth media, and uses readily available technology like fluorescence plate imagers, fluorescence microscopes, spectrofluorometers and flow cytometry. So long as cell cycle arrest is a rare phenomenon, screening for fluorescent recombinants would appear to be a near ideal method for quantifying recombination.

Repair of a GFP nonsense mutation (Pierce et al. 2003; Gurskaya et al. 2001; Subach et al. 2008; Ai et al. 2008; Shaner et al. 2008; Heim et al. 1994) proved to be
useful for oligo mediated recombination assays in mouse embryonic stem cells, but there is a weakness with this target: there is no direct confirmation that the targeted transgene is transcriptionally active prior to recombination, possibly leading to underestimate of Recombineering efficiency as only expressed recombinant genes are detectable while both expressed and non-expressed genes contribute to the total titer (Disterer et al. 2012). A better assay would be to change the emission spectrum of a fluorescent transgene target so one could be sure that the transgene is expressed in both recombinant and non-recombinant cells. Fluorescent protein genes have been mutated to express many spectral variants of the initial fluorescent protein gene as illustrated by Roger Tsien in his Nobel Laureate award lecture presentation. We thought that the perfect reporter for human Recombineering would be to shift the spectral properties of a fluorescent transgene target. A fluorescent protein gene target would allow us to quantify how many cells express the recombination target and to use flow cytometry to interrogate a large number of cells in a recombination experiment. Therefore we decided to use fluorescent protein genes as recombination targets and to tune protein fluorescence emission spectra by targeted allelic swaps as a means of assessing Recombineering efficiency expressed as fraction of active targets so modified. To better estimate recombination frequencies, time-courses were performed to determine to what extent cell cycle arrest influences quantification.

We idealized this reporter by using the most efficient Recombineering protocol, which uses ssDNA oligos to modify target genes. Oligo Recombineering is convenient because the only SynExo function required to catalyze gene conversion is the Synaptase and the substrate ssDNA oligos are inexpensive and readily obtained. Oligo substrates are
typically designed with ~35 nucleotides homology at each end and the desired change (substitution, insertion or deletion) located in the middle of the oligo, between the two homologous targeting arms. When MMR is not active (e.g. by incorporation of a non-correctable allele), the efficiency of gene targeting by ssDNA oligos is highest when single-nucleotide mismatches are created and lower when there is a larger number of mismatches creating a ssDNA “bubble” between the oligo and the target (Swaminathan et al. 2000; Valledor et al. 2012). Therefore, we were especially interested in protocols that swap phenotypes with the minimum number of changes at the nucleotide level.

**Amino acid residues determining the fluorescent spectra**

We then sought to determine which amino acid residues could be altered by Recombineering to change the spectral properties of fluorescent transgene targets. In this study, multiple alignments of the most closely related AcGFP fluorescent protein spectral variants were performed to identify such amino acid residues (Figure 3.1). A ClustalW alignment of sequences encoding green, cyan, blue and yellow fluorescent protein gene variants showed many changes from one variant to the other. We were aware that many of these proteins were created by random mutagenesis and protein evolution. These techniques were important to find novel variants with different colors in addition to substitutions related to physico-chemical properties such as improved maturation time and photostability that were not directly relevant to spectral wavelengths (Heim et al. 1994; Orm et al. 1996; Gurskaya et al. 2001; Ai et al. 2008; Subach et al. 2008; Shaner et al. 2008).
gi|60115915  GFPmut3  M-RKGEELFTGVPIVLDGDVGHIKFSVSCEGEGD
gi|11321082  ECFP  M-SKGEELFTGVPIVLDGDVGHIKFSVSCEGEGD
gi|14566499  EBFP2  MVSKGEELFTGVPIVLDGDVGHIKFSVREGEVD
gi|156106790  YFP  MVSKGEELFTGVPILVLDGDVGHIKFSVSCEGEGD

6

6

ATYGKLTLEICTTGKLPVPWPTLVTTFGYQCFAR
ATYGKLTLEICTTGKLPVPWPTLVTTFGYQCFAR
ATYGKLTLEICTTGKLPVPWPTLVTTFGYQCFAR
ATYGKLTLEICTTGKLPVPWPTLVTTFGYQCFAR

YPDHMKQHDFDSAMPEGYVQERTIFFKFDGNYKTRA
YPDHMKQHDFDSAMPEGYVQERTIFFKFDGNYKTRA
YPDHMKQHDFDSAMPEGYVQERTIFFKFDGNYKTRA
YPDHMKQHDFDSAMPEGYVQERTIFFKFDGNYKTRA

EVKFEGDTLVDKGIKGDKNLIGHLKELYNNS
EVKFEGDTLVDKGIKGDKNLIGHLKELYNNS
EVKFEGDTLVDKGIKGDKNLIGHLKELYNNS
EVKFEGDTLVDKGIKGDKNLIGHLKELYNNS

HNVYIMADKQKGNIKVNFIRHNEDGSVQLADHYQQ
HNYVIMADKQKGNIKVNFIRHNEDGSVQLADHYQQ
HNYVIMADKQKGNIKVNFIRHNEDGSVQLADHYQQ
HNYVIMADKQKGNIKVNFIRHNEDGSVQLADHYQQ

NTPIGDGPVLLPDNHLYLSTQSALSKDPEKRDMVLL
NTPIGDGPVLLPDNHLYLSTQSALSKDPEKRDMVLL
NTPIGDGPVLLPDNHLYLSTQSALSKDPEKRDMVLL
NTPIGDGPVLLPDNHLYLSTQSALSKDPEKRDMVLL

EFVTAAGITHGMDELYK
EFVTAAGITHGMDELYK
EFVTAAGITHGMDELYK
EFVTAAGITHGMDELYK

Figure 3. 1 Chromophore region determines the spectral properties of fluorescent protein genes
ClustalW alignment of fluorescent protein sequences related to the cycle 3 AcGFP variant GFPmut3* but having
unique spectra (green, cyan, blue and yellow) identify amino acids to target for making novel fluorescent protein genes
by Recombineering.
We noticed that the amino acids located at the chromophore region were different for different color variants, which highlighted the importance of an aromatic amino acid residue at position 66 and its neighboring residues on differences in fluorescence spectra. With the exception of position 203 which influences the environment of the chromophore, most other changes in protein sequence were not universally supported as important for spectral shifts.

Based on previous work (Levine 1997; Miao et al. 2009; Heim et al. 1994; Orm et al. 1996; Ai et al. 2008), we hypothesized that the amino acids at the chromophore region of positions 64 (F|L), 65 (G|T), 66 (Y|W|H) and a residue at position 203 (T|Y) which interacts with the fluorophore, were primarily responsible for the “color” of each protein while other residues are less likely to be important for spectral shifts. We surmised that it might be possible to produce several spectrally distinct fluorescent proteins by oligo-directed mutagenesis in vivo by Recombineering.

**Fluorescent protein engineering plan**

To test the hypothesis that amino acids 64-66 and 203 were the color determinants in the fluorescent protein genes we modified these amino acids using Recombineering and subsequently evaluated the spectral properties of the new proteins. Even though the final goal of this study was to study Recombineering in human cells, we tested this hypothesis where Recombineering works best at this time, in *E. coli*. We therefore sought an *E. coli* cell line expressing a suitable fluorescent protein target.

The *E. coli* strain SCC1, was described as a expressing a bright green fluorescent protein called GFPmut3* (Trojan et al. 2002; Andersen et al. 1998; Miao et al. 2009). SCC1 has the gfpmut3* gene inserted between base pairs 312754 and 312771 of the MG1655 *E. coli* genome. The gfpmut3* gene is expressed from the LacI-repressible
promoter PA1/04/03 (Valledor et al. 2012; Cormack et al. 1996; Andersen et al. 1998; Tombolini et al. 1997). GFPmut3* is derived from the jellyfish *Aequorea victoria* GFP (S2R, S65G, S72A) (Pluta et al. 2005; Cormack et al. 1996; Tombolini et al. 1997), but is 20 times more fluorescent than GFP when excited at 488 nm and has a much faster maturation rate (Shin et al. 2006; Warming et al. 2005; Cormack et al. 1996; Tombolini et al. 1997). We obtained the SCC1 strain from Dr. Sze for our Recombineering studies.

**Building a bacterial Recombineering reporter strain**

We introduced the Red Recombineering enzymes from a defective λ prophage by P1 transduction into SCC1 (Figure 3.2). The strain SW101 (Strack et al. 2009; Thomason et al. 2005; Warming et al. 2005) carries a defective λ prophage [cI<sup>857</sup> Δ(cro-bioA)] expressing the Red system under control of a thermosensitive cI repressor variant (cI<sup>857</sup>). The λ Red recombination functions are encoded by three adjacent genes, *gam*, *bet* and *exo*, in the λ *pL* operon (Poteete 2011; Thomason et al. 2005). Expression of the Red system is tightly regulated from the λ *pL* promoter under the control of the temperature sensitive cI<sup>857</sup> repressor. A Tn10 transposon, encoding the tetracycline resistance gene, is genetically linked to the defective λ prophage in SW101 via *nadA::Tn10*, which allowed us to move the prophage into other strains by P1 transduction with selection for tetracycline resistance.
The bacterial Recombineering reporter strain was constructed by producing P1 transducing lysates from P1 infected SW101. SCC1 was then transduced with the P1(SW101) lysate. Transduced colonies were selected by growth on tetracycline LB media. The resulting Recombineering reporter strain, RIK410, was validated by impaired growth at 42 °C (Figure 3.3) and by immunity to lambda super infection (not shown).
In RIK410, the Red system is induced at 42 °C and impairs growth when expressed for prolonged times. The Kil and N proteins are also expressed from the prophage. Kil blocks progression of the cell cycle and impairs colony formation. N protein is an antitermination factor that increases expression of the \textit{rrn} loci, producing more rRNA and disrupting ribosomal protein stoichiometry. The recipient strain SCC1 showed a normal growth rate at 42 °C with a doubling time of 24 minutes while the donor strain SW101 and RIK410 had growth impairment at 42 °C with doubling times of 220 and 123 minutes respectively (SW101 is also \textit{recA}− which reduces growth rates by about half, so this difference is expected). The growth curves are in Figure 3.3.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure3.png}
\caption{Recombineering strain validated as a \textit{\lambda} prophage lysogen.}
SCC1, SW101 and RIK410 were grown at 42 °C in LB broth in a shaking incubator. Growth rate was determined by measuring the turbidity of the culture using a Klett meter over time. The growth is reported in KU. The doubling time is reported in the right side of the figure for each of the strains evaluated.
\end{figure}

\textit{\lambda} phage are unable to form plaques on \textit{\lambda} lysogens due to superinfection immunity. The \textit{\lambda} prophage is maintained in a silent state by the repression of phage genes by cI repressor protein. When a new \textit{\lambda} phage infects a bacterium carrying a prophage, the
expression of its lytic genes is repressed by cI expressed from the prophage. Therefore it is possible to test if a strain carries a λ prophage by its inability to produce plaques upon λ phage infection. The SW101 strain, RIK410, and SCC1 strains were then infected with λ phage MMS5. Plaques were formed in SCC1 strain indicating successful infection, but no plaques were formed in SW101 or in the RIK410 strains, as predicted.

The resulting bacterial Recombineering reporter strain, RIK410, is an *E. coli* MG1655 derivative that expresses GFPmut3* and carries a defective λ prophage (cIΔ857Δ(cro-bioA)) that encodes the Red SynExo recombinases (Figure 3. 2A). The expression of the Red SynExo system is tightly repressed when cells are grown at 32 °C, but are highly expressed after a 42 °C heat shock induction for 15 minutes (Figure 3. 2B). Expression of the Red system includes Gam (an inhibitor of RecBCD and SbcCD nucleases), Beta (the SynExo synaptase protein) and Exo (a 5’ → 3’ dsDNA exonuclease that makes ssDNA substrates for Beta-mediated homologous DNA pairing). Gam stimulates recombination by preserving DNA substrates from exonucleaseolytic degradation. Beta is a ssDNA binding protein that both protects ssDNA recombination intermediates from nucleases and catalyzes annealing of complementary ssDNA to promote genetic exchange. Exo forms a complex with Beta and resects dsDNA to produce ssDNA substrates for Beta-mediated homologous DNA pairing. The strain was kept frozen to minimize unwanted expression of the Red system. Aliquots were thawed for each individual assay.

**Bacterial Recombineering targeting the chromophore region**

To evaluate if changing the sequence of the chromophore region to the predicted amino acids will change the spectral properties of GFP, Recombineering experiments
were performed by targeting the *gfpmut3* reporter gene in RIK410 with oligos carrying the desired change. Oligos were designed to target the chromophore region with 35 to 44 nucleotide arms of homology to the chromophore region and the desired change in the middle. Oligos 1 to 4 were designed to produce a blue spectral shift of GFPmut3*. Oligo 25 was designed to produce a red spectral shift of GFPmut3*.

Recombineering assays were performed by streaking out a fresh colony of RIK410 obtained from the freezer. One colony was used to inoculate growth medium and incubated overnight. The culture was diluted 1/100 and left to grow until early exponential phase (OD$_{600}$ = 0.5 – 0.6, Klett = 20 KU). Cells were heat-shocked at 42 °C for 15 minutes to induce a burst of recombinase expression and promptly cooled to shut down the expression. Cells were concentrated 200-fold and made competent for DNA uptake by electroporation by washing them 3 times with ice-cold water.

Recombineering oligos were electroporated into cells, and cells were allowed to recover for 30 minutes in liquid medium prior to plating. Cells were diluted and plated to about 30 cells per plate. *E. coli* fluorescence was evaluated using fluorescence microscopy, a plate reader and by flow cytometry (Figure 3. 4). When cells were treated with oligo 1, sectors of Cyan cells within the Green colony were obtained (Figure 3. 4b).

In subsequent experiments *gfpmut3* was successfully changed to express proteins in the blue side of the spectrum using oligos 2 and 56, on the cyan side of the spectrum using oligo 3 and 4 and to the red side of the spectrum using oligo 25. Recombinant cells were imaged using fluorescent microscopy and confocal as well (Figure 3. 5).
Figure 3.4 *E. coli* GFPmut3*+Aqua Recombineering variants are distinguishable

(a) Fluorescent image of individual *E. coli* expressing GFPmut3* and Aqua imaged with a DeltaVision microscope, using an Olympus 60X objective, a CoolSNAP HQ / ICX285 camera and 3D deconvolution software. *E. coli* GFPmut3* was imaged using FITC emission and excitation filters. *E. coli* Aqua was imaged using CFP emission and excitation filters. (b) Fluorescent image of *E. coli* colony sectors obtained after a Recombineering experiment (gfpmut3* to aqua). The top picture is a phase contrast image. The central picture was obtained using a Chroma green/mCherry 51004v2 F/R C71565 filter, which visualizes GFPmut3* fluorescence but excludes Aqua. The bottom picture was obtained by using a Chroma 31044v2 CY GFP C904,12 filter, which visualizes Aqua fluorescence but excludes GFPmut3*. Shown in the picture are two colonies, one non-recombinant (GFPmut3*) and the other a “sectored” recombinant colony (~75% GFPmut3*, ~25% Aqua). The pictures were taken with a Nikon Diaphot microscope, and a Nikon 10X/0.25 Ph1 DL objective. (c) Fluorescent black and white image of a mixture of *E. coli* GFPmut3*+Aqua colonies (bright colonies) and Aqua recombinant colonies (darker colonies). The plate was imaged using a Dark Reader transilluminator (Clare Chemical, excitation 420-500 nm, emission >520nm) and a CCD camera. (d) *E. coli* GFPmut3* and *E. coli* Aqua can be separated and quantified by FACS. About 10^7 cells were harvested and resuspended in 1.05 % K2HPO4, 0.45 % KH2PO4, 0.005 % MgSO4.7H2O, 0.1 % (NH4)2SO4, 0.05 % sodium citrate supplemented with 0.1% thiamine and 0.2% glucose and then analyzed using a BD FACS Aria-II (BD Biosciences). Non-fluorescent cells were analyzed to set the gates (first panel). *E. coli* GFPmut3*+Aqua were separated as positive controls (second and third panels, respectively). A mixture of *E. coli* GFPmut3* and *E. coli* Aqua cells were analyzed as indicated in the fourth panel, showing that it is possible to isolate recombinant from non-recombinant fluorescent *E. coli* cells.

Protein extracts were prepared from isolates of each recombinant class and analyzed by fluorescence spectroscopy as illustrated in (Figure 3.6). We corroborated our hypothesis that changes in the amino acid sequence of the chromophore are sufficient to change the fluorescence properties of fluorescent proteins. The primary determinant was
found to be the identity of the aromatic residue at position 66 with neighboring residues altering the spectral properties in more subtle ways. Recombinant genotypes were validated by allele-specific PCR and DNA sequencing.

Figure 3.5 Aqua, Amarillo and GFPmut3* E. coli created by Recombineering, distinguished by confocal microscopy.
Aqua, Amarillo and GFPmut3* E. coli were grown independently to OD600 of 0.4. Cells were collected from each culture and resuspended in minimal medium (OMBG) to a 10% of the initial volume. Cells were mixed and 1 µl of the mixed concentrated cells were added to a slide with mounting medium (P7481 Prolong Antifade, from Molecular Probes). The slide was covered and left to sit overnight before imaging. Tile scanning was taken with a Zeiss LSM710 confocal microscope. Maximum projection was performed with PerkinElmer/Improvision Volocity 64-bit software. Color merges were done in Adobe Photoshop CS3.
Figure 3. Spectra of fluorescent proteins Aqua, Violeta, Amarillo and Bronze created by Recombineering from \textit{gfpmut3*}.

\textit{E. coli} expressing \textit{GFPmut3*} and the SynExo recombinases (\textit{λ}, \textit{cI}^{857} \textit{Acro-bioA::Tn10}) were transformed with oligos complementary to a ~ 70 nt region of target \textit{gfpmut3*} gene but differing in the nucleotide sequence of selected amino acids, as indicated next to the spectra for each protein. Cultures were diluted and plated on nonselective medium and recombinants were identified by a change in colony fluorescence using a “Dark Reader” light box or by colony PCR (Bronze). Protein extracts were prepared from isolates of each recombinant class by lysing the cells with BugBuster Protein Extract Reagent (70584, Novagen), 25 U/µl Benzonase Nuclease (70746, Novagen), 20 U/ml lysozyme and 0.1 M DTT. The mixture was incubated for 20 minutes at room temperature in a rocking platform. The mixture was then centrifuged for 30 minutes at 4 °C, the supernatant was removed from the tube and centrifuged again. The supernatant was used to measure the protein spectra. If necessary, samples were diluted in TE containing 0.5 mM DTT. The excitation and emission spectra of eGFP fluorescence and its variants in the extracts were determined with a spectrofluorometer (PTI QuantumMaster) with 2 nm slits at a slew rate of 1 nm/second (the average integrated one second samples collected ~ 10^4 - 10^6 emitted photons). The data were collected using Felix32 software, and later analyzed using Microsoft Excel and GraphPad Prism. Each spectrum was normalized to its peak value and plotted.
We corroborated that changes in the chromophore region of GFPmut3* (F<sub>64</sub>G<sub>65</sub>Y<sub>66</sub>…T<sub>203</sub>) produced spectral changes from green fluorescence. GFPmut3* (F<sub>64</sub>G<sub>65</sub>W<sub>66</sub>…T<sub>203</sub>) produced a blue shifted variant (Violeta). GFPmut3* (F<sub>64</sub>T<sub>65</sub>W<sub>66</sub>…T<sub>203</sub> and L<sub>64</sub>T<sub>65</sub>W<sub>66</sub>…T<sub>203</sub>) produced cyan variants (Aqua and Mar, respectively). GFPmut3* (F<sub>64</sub>G<sub>65</sub>Y<sub>66</sub>…Y<sub>203</sub>) produced a yellow variant (Amarillo) while GFPmut3* (L<sub>64</sub>T<sub>65</sub>W<sub>66</sub>…Y<sub>203</sub>) produced a greener cyan variant (Bronze). We determined that these variants could be distinguished by microscopy and flow cytometry. Even the modestly shifted Aqua variant is easily resolved from GFPmut3*, demonstrating the utility of ssDNA oligo Recombineering for producing useful variants for multi-color studies.

**Conclusions**

In conclusion, we were able to change the spectral properties of GFPmut3* by producing small nucleotide changes in the chromophore region of the gfpmut3* gene. The variants can be differentiated using microscopy, flow cytometry and a plate reader. The changes were introduced by Recombineering, thereby demonstrating that targeting the chromophore region is an excellent reporter for Recombineering.

A systematic study of recombination parameter optimization by swapping color was performed and will be discussed in the next chapter. Different aspects of the Recombineering protocol were optimized for fluorescent protein engineering and general aspects of the Recombineering mechanism were studied to develop Recombineering in human cells. A humanized version of GFP was chosen as the reporter for human Recombineering. Mutagenesis was performed in *E. coli* to validate the fluorescent
properties of the human Recombineering model. The details for the human Recombineering reporter will be discussed in Chapter 5.

**Significance**

We have demonstrated that fluorescent protein sequence changes easily introduced *in vivo* via Recombineering produce significant spectral shifts. We showed that this method is useful for generating multi-color reporters and novel fluorescent proteins. The Bronze variant looks especially promising for imaging studies, as the excitation maximum (468 nm) overlaps a common laser wavelength. This approach is not restricted to modifying fluorescent protein genes in the *E. coli* genome, but also works for altering fluorescent protein reporter genes in plasmid and viral sequences propagated in *E. coli*. 
CHAPTER 4. FLUORESCENT PROTEIN BACTERIAL RECOMBINEERING

Rational for fluorescent protein Recombineering optimization in *E. coli*

The previous Chapter examined if it was possible to engineer fluorescent protein spectra via Recombineering. We tested this feasibility in *E. coli* since Recombineering is a technique developed in *E. coli*. We learned that it was possible to engineer fluorescent proteins by Recombineering, and that fluorescent proteins were an excellent assay to quantify Recombineering. Nevertheless, there were features from the data that suggested that the assay needed to be improved. For example, the recombination frequencies were low (<0.1%) compared to others (~1%) (Sawitzke *et al.* 2007). Also, while a strand bias for one of the oligos was observed, it appeared to be opposite to what was observed in previous studies (Ellis *et al.* 2001; Constantino and Court 2003; Li *et al.* 2003). Finally, Recombineering frequencies varied from one experiment to another. The low efficiency and low reproducibility of the preliminary results was a problem since the motivation for expanding this technology to human cells is based on the high efficiency of bacterial Recombineering. In order to recapitulate Recombineering in human cells, we really needed to understand the Recombineering mechanism and how to make the protocol as efficient as possible. Therefore, we evaluated the fluorescent engineering protocol in *E. coli* and optimized it to the best frequencies for Recombineering reported in the field.

For the general Recombineering protocol, the reporter strain expressing a fluorescent protein gene and the Red recombination system was electroporated with an oligo targeted to the fluorescent protein gene (*Figure 3. 2*). Recombinants were identified by a change in fluorescent. The Recombineering protocol was examined at many levels, for example we compared different methods to quantify Recombineering efficiency with sufficient
cells to do statistical analysis. We evaluated the orientation of \textit{gfpmut3}* transcription with respect to the \textit{E. coli} origin of replication in order to properly target the lagging strand template as recommended by Ellis \textit{et al.} (2001), Constantino and Court (2003) and Li \textit{et al.} (2003). Since we noticed that Recombineering frequencies fluctuated from one experiment to another, we did a time course of the outgrowth prior to plating to evaluate the influence of chromosome segregation during outgrowth on recombinant frequencies. We evaluated if oligo transformation was a rate limiting step and how to improve transformation efficiencies. Finally, we evaluated how the oligo concentration, oligo length, oligo strand and the nature of the introduced change by the oligo to produce the color swap affected the recombination efficiency.

\textbf{Scoring Recombineering frequencies in \textit{E. coli}}

We were interested in being able to quantify Recombineering frequencies by using a sufficient number of cells to do statistical analysis to compare experimental results. Microscopy was not suitable because only a small amount of cells could be evaluated at the time, and was subject to investigator error during counting. Flow cytometry at the Core Facility was able to differentiate the variants, but was not optimized for enumerating the small \textit{E. coli} cells. We then used a Dark Reader transilluminator to screen for green fluorescent colonies or sectors on petri plates. This system was convenient since it was possible to interrogate many plates with many colonies in the lab. Looking at colonies on the Dark Reader also offered information about how recombinant cells segregate into recombinants and non-recombinants (\textbf{Figure 3. 4c}).
**Determination $gfpmut3^*$ orientation with respect to replication fork**

In Recombineering, the substrate oligo is incorporated into genomes at the highest rate when it anneals to the lagging strand template during replication of the target DNA sequence (Ellis *et al.* 2001; Constantino and Court 2003; Li *et al.* 2003). A scheme describing insertion of $gfpmut3^*$ into the *E. coli* genome (Sawitzke *et al.* 2011; Miao *et al.* 2009), indicated that the $gfpmut3^*$ coding sequence was inserted in the clockwise (CW) orientation of the *E. coli* genome annotation relative to the genetic map. Miao *et al.* (2009) indicated that the $gfpmut3^*$ gene was inserted by homologous recombination at positions 312754 to 312771 base pairs in the *E. coli* genome; this position corresponds to 6.75 minutes on the genetic map (http://www.ecogene.org/). The origin of replication is located at the 84.57 minutes. The $gfpmut3^*$ gene is positioned to the right of the origin of replication at the 84.57 minutes (Figure 3. 2). Therefore, $gfpmut3^*$ is replicated as the right replication fork translocate in the *E. coli* genome. In this scenario, continuous DNA synthesis uses as a template the antisense sequence of $gfpmut3^*$ and discontinuous DNA synthesis uses as a template the sense sequence of $gfpmut3^*$. The oligo targeting the lagging strand template will be the same sequence as the antisense strand of the $gfpmut3^*$ coding sequence, save the altered nucleotides (Figure 4. 1 top).
Figure 4.1 Replication fork model across *gfpmut3* and oligo targeting the lagging strand template

When *gfpmut3* is in the Clockwise (CW) orientation the coding sense strand is the lagging strand template and oligos targeting the lagging strand template are the reverse complement to the sense strand (antisense). When *gfpmut3* is in the Counter Clockwise (CCW) orientation the coding sense strand is the leading strand template and oligos targeting lagging strand template are the sense strand.

Interestingly, the initial rates for fluorescent protein engineering were more efficient with sense oligos than with the antisense, indicating strand bias, but opposite to our interpretation of the published figure map (Miao et al. 2009). The data was consisted with *gfpmut3* expressed in a counter clockwise (CCW) orientation in the *E. coli* genome. Since the fluorescent protein engineering data consistently contradicted the previously reported strand bias, the orientation of *gfpmut3* in the *E. coli* genome relative to the origin of replication was re-examined. The orientation was examined by PCR amplification within *gfpmut3* towards the chromosomal flanking regions. PCR was performed with either a forward primer (5) or a reverse primer (6) amplifying from within *gfpmut3* and either a left flanking primer (17) hybridizing to base pairs 31244-
312480 of *E. coli*, or a right flanking region primer (18) hybridizing to base pairs 313151-313126 of *E. coli*. The strategy and results are shown in Figure 4.2.

**Figure 4.2** *gfpmut3* gene is in opposite orientation to the replication fork
*gfpmut3* orientation was examined by PCR amplification within the gene and towards the chromosomal flanking regions. PCR was performed with either a forward primer 5 or a reverse primer 6 amplifying from within *gfpmut3* and either a primer 17 hybridizing to base pairs 312444-312480 of *E. coli*, corresponding to the left flanking region of the *gfpmut3* insert or a primer 18 hybridizing to base pairs 313151-313126 of *E. coli*, corresponding to the right flanking region of the *gfpmut3* insert. The *gfpmut3* was amplified by primer pairs 5+17 and 6+18.

In one model *gfpmut3* transcription unit is co-oriented with the direction of the replication fork travel (as previously reported) and should be amplified with primer pair 5+18 and 6+17. In the other model, the *gfpmut3* transcription unit orientation is opposite to the direction of the replication fork and should be amplified with primers 5+17 and 6+18. Only the primer pairs 5+17 and 6+18 produced products indicating that *gfpmut3* is opposite to the direction of the replication fork travel and opposite to the suggested orientation. In this scenario the sense oligos should hybridize to the lagging strand.
template and be the most efficient in Recombineering, which agrees with the fluorescent protein engineering data (Figures 4.1 bottom and 4.5).

**Improving reproducibility of the Recombineering protocol**

Recombineering frequencies are reported often with huge fluctuations depending on the lab and/or the person performing the experiment. For instance, 50% by Dr. Andres Larrea in Rik Myers’ lab (Larrea et al. 2008), 10-30% by Dr. Court’s lab (Sawitzke et al. 2007) and less than 1% from many other reports in the field. At this point fluorescent protein engineering efficiencies were less than 1%. The fluctuations in Recombineering efficiencies are due to the complexity of recombination mechanism, the several steps in the protocol, and several factors that influence the outcome, such as the nature of the mismatch and the sensitivity of the assay used. Our assay was really sensitive, but was affected by MMR in different magnitudes depending of the mismatch introduced by the oligo substrate. The oligos used to modify the fluorescent protein genes are listed in Figure 4.3.

**Avoiding MMR**

Since we were interested in optimizing all the steps in the protocol to reach or improve the maximal reported Recombineering frequencies, we created a non-fluorescent version of GFPmut3* that could be repaired by introducing a C.C mismatch, which is not recognized by MMR. The non-fluorescent version of GFPmut3* was created by introducing a stop codon at position 39 by Recombineering (RIK410 x oligo10 = strain RIK423). RIK423 gfpmut3* could then be repaired by Recombineering with oligo 11 via a C.C mismatch formed in the intermediate to restore GFPmut3* expression (Figure 4.4). The C.C mismatch has been reported to be resistant to MMR (Joshi & Rao 2001).
155
Repairing the gfpmut3* with a C.C mismatch proved to be an improvement in the
Recombineering efficiency to about 1 %. Nevertheless, the frequencies were not yet as
high as reported ~10%.

Figure 4. 3 Oligos used for fluorescent protein Recombineering.
Oligos were designed 67 to 99 nucleotides long either as the sense or antisense strand of the target gene. The
modification or target region was placed right in the middle of the oligo. Four regions of gfpmu3* were targeted: A and
E) oligos to modify the chromophore region or to revert to normal; B) oligo to change Tyrosine 203 to Tryptophan 203;
C, D, F and G) oligos to introduce and stop codon or revert to normal. All oligos were synthesized by Sigma-Aldrich.


RIK423  

**E. coli** Recombineering Reporter

**Figure 4.** 
**E. coli** Recombineering reporter that avoids MMR.

RIK423 encodes a non-fluorescent version of GFPmut3* that can be repaired by introducing a C.C mismatch, which is not recognized by MMR. The non-fluorescent version of GFPmut3* was created by introducing a stop codon at position 39 by Recombineering (RIK410 x oligo 10 = strain RIK423). RIK423 GFPmut3* is repairable by Recombineering with oligo 11.

**Determination of optimal post-Recombineering outgrowth time**

Despite avoiding mismatch repair and using a sensitive assay, the frequency of fluorescent protein engineering was ~1%, not yet the maximal frequencies reported. One difference between our protocol and others was the duration of the outgrowth time after electroporation. In our experiments, we were letting the cells recover and the recombinants to segregate through an overnight incubation period while other labs typically plated cells at different time points starting after 30 min incubation (Sawitzke et al. 2007). In order to evaluate the shortest time cells need to recover after electroporation...
and when to evaluate recombination frequency, we performed outgrowth time course experiments by plating dilutions of cultures in intervals of 30 minutes to one hour. Recombinant colonies were evident 24 hours later by visualizing green fluorescent colonies using a Dark Reader imaging apparatus.

Shortly after electroporation, the Recombineering enzyme Beta is thought to bind an oligo and anneal it a complementary target sequence as the target is replicated (Sawitzke et al. 2011; Maresca et al. 2010), creating a heteroduplex recombination intermediate with the mutation present on one of the two DNA strands. If bacteria are plated before mutant and non-mutant DNA molecules segregate, recombinant colonies are “sectored” with part of the colony comprised of mutant cells and the remainder composed of non-mutant cells (Figure 4.5A and Sergueev et al. 2002; Sawitzke et al. 2011).

We found that plating cells 30-60 minutes after electroporation is an optimal balance of high viability and high recovery of recombinant (mutant) colonies. Plating during the first hour of outgrowth allows one to accurately assess recombination frequencies and also to most efficiently identify mutants. On the other hand, if bacteria are plated after segregation, pure mutant colonies appear at a frequency equal to the initial recombination frequency divided by the number of DNA chains in the original recombinant cells (Sharan et al. 2009; Sergueev et al. 2002), increasing the number of colonies one must screen to recover mutants.

Recombinants were scored as sectored when only part of the colony was fluorescent or solid when the whole colony was fluorescent. The total number of colonies was counted on a 1948 issue colony counter. As shown in Figure 4.5A, there is an initial lag
in the culture where cells are not replicating, then the total cell culture starts to grow but
the recombinants within the culture appear to stall or not to divide for 3 to 4 hours.

Figure 4. 5 ssDNA Recombineering knowhow.
Recombination frequency is reported in each panel as the % of recombinant colonies among the total number of colonies. **A**) Recombinant DNA segregates from non-recombinant DNA as cells divide. After oligo electroporation (oligo 11 x RIK423), cells were allowed to recover on media (outgrowth). Cells were plated at the indicated time on outgrowth. If cells were plated before recombinant DNA segregated from non-recombinant DNA, the resulting recombinant colonies have a sectored phenotype, which is a mixture of recombinant and non-recombinant cells (Sectored Recombinants). If cells were plated after recombinant DNA segregated from non-recombinant DNA, the resulting recombinant colonies have a homogenous phenotype (Solid Recombinants).

**B**) Recombineering is more efficient when targeting the lagging strand template. Recombineering targeting the lagging strand template (oligo 11 x RIK423) was 72 times more efficient than recombination targeting the leading strand template (oligo 10 x RIK410).

**C**) The average Recombineering frequency approaches 40% when normalized to transformation frequency. Transformation frequency was estimated by transformation with pUC19 in a parallel experiment to Recombineering (oligo 11 x RIK423), represented in green. Assuming that the percent of cells that became Amp\(^R\) were the total amount of transformable cells able to uptake oligos, the Recombineering % was normalized to the % of transformable cells represented in red. **D**) MMR can be a potent inhibitor of recombination. Recombineering was performed in isogenic MMR\(^+\) and MMR\(^-\) (\Delta mutS::Kan) strains. A one-tailed unpaired t test with Welch's correction was performed for each of the MutS\(^+\) and MutS\(^-\) pairs. **Table 4.** shows mismatch, assay and P values for each of the MutS\(^+\) and MutS\(^-\) pair.
Sergueev et al. (2002) showed that *E. coli* replicates with 2-4 chromosomes per cell when grown exponentially in LB medium. They showed that upon λ prophage induction, DNA replication continues while there is an arrest to cell division for about one hour. As Recombineering only targets one of the strands in a target chromosome and only happens in the context of DNA synthesis, the number of recombinants per recombination event will never be more than \( \frac{1}{4} \), while if there are two bacterial chromosomes dividing we will expect \( \frac{1}{8} \) recombinants, and so forth. To determine the ratio of initial chromosomes to recombinants, the recombinant frequency as a function of cell divisions (generation) was compared to a simple mathematical model of recombinant segregation as a function of exponential decay that takes into account the dominant phenotype of GFP positive recombinants arising from correction of nonfluorescent mutants. The pattern of GFP positive recombinant segregation as a function of chromosome copy number is shown in Figure 4.6. The recombinant segregation pattern of the data reaches a limit predicted by \( N = 4 \), as if the initial state of *E. coli* were of 4 chromosomes equivalents of the target gene. Our data are most consistent with the model that there are 2-4 chromosome copies present during Recombineering, as suggested from the work of Sergueev et al. (2002).

Assumptions in the strand segregation model:

1) Each chromosome is double-stranded.

2) ssDNA recombination creates a heteroduplex with one recombinant strand and one non-recombinant strand.

3) The number of recombinant strands at \( t = 0 \) could be from 0 to the number of target gene copies (N).
4) There are no segmental duplications, so \( N \) = both the target copy number and the chromosome number.

5) The GFP\(^+\) allele is dominant to the GFP\(^-\) allele in polyploid cells.

6) The ssDNA changes a GFP\(^-\) allele to GFP\(^+\). (*NB: The results would be different for GFP\(^+\) to GFP\(^-\) as the recombinant phenotype will be masked in polyploid cells.*)

7) Based on previous results, under optimum conditions, recombination and transformation efficiencies are equivalent. For Figure 4.6, however, the values were scaled to sample timecourse data where the maximum recombination frequency (R) was 27%.

**Figure 4.6** Oligo sequence present in one out of eight DNA strands in exponentially growing *E. coli*. Chromosome segregation was modeled and compared to recombination data. Data from one experiment is shown but the experiment gave consistent results in multiple trials. Recombinants represent the % of colonies that shows the recombinant phenotype (sectored and solid colonies) when cells were allowed to recover and divide during outgrowth in liquid medium before spreading samples on agar plates.

**Method used in the model:** Let \( N \) vary from 1 to 8 chromosome copies. At generation 0, recombination creates a heteroduplex intermediate at one target copy on one chromosome. In generation 1, cells have divided, halving the DNA content and then replicate the chromosomes once to return to \( N \). At this point, one half of the cells express...
the + phenotype no matter what the value of N is because + is dominant to -. In generation 2, the results diverge:

For N = 1, R = 0.5 in generation 1 and will remain at 0.5 in subsequent generations as all + alleles are in homozygous cells.

For N = 2, R = 0.5 in generation 1 because one half of the cells are homozygous -/- and one half are heterozygous +/- . R = 0.25 in generation 2 because division of one half of the cells at generation 1 can give rise to cells with the + phenotype. R will remain at 0.25 in subsequent generations as all + alleles are in 25% population of +/+ homozygous cells.

For N = 4, R = 0.5 in generation 1 because one half of the cells are homozygous -/-/- and one half are heterozygous +/-/-/. R = 0.25 in generation 2 because division of one half of the cells at generation 1 give rise to cells with the + phenotype (these cells are +/-/-/). In generation 3, the 25% of the population generated in generation 2 can segregate 25% +/- → +/+/++ that will remain invariant as they are now homozygous. Another 50% of the 25% of the population generated in generation 2 can segregate +/- → +/-/-/- heterozygotes, leading to a total + population in generation 3 of R = 0.25 * (0.25+0.5) = 0.1875. This pattern will repeat itself in subsequent generations with the pool of heterozygotes (HET) being steadily depleted and the pool of homozygous (HOM) +/+/+/+ cells rising to an asymptote.

This is described by:

\[ \text{HET}_3 = 0.25 \times 0.5 = 0.125 \]

\[ \text{HET}_4 = 0.25^2 \times 0.5 = 0.03125 \]
HET_N = 0.25^{N-2} \times 0.5

AND

HOM_3 = 0.25^2 = 0.0625

HOM_4 = HOM_3 + 0.25^3 = 0.078125

HOM_N = \text{sum}(HOM_4; HOM_{N-1}) + 0.25^{N-1}

To calculate R at any generation, sum HET and HOM for that generation number: (e.g. \( R_3 = HET_3 + HOM_3 = 0.125 + 0.0625 = 0.1875 \)). As HET approaches 0, HOM fixes at a value approaching R = 0.083333335

For \( N = 8 \), R = 0.5 in generation 1 because one half of the cells are homozygous -/-/-/-/-/-/-/- and one half are heterozygous +/-/-/-/-/-/-/. In generation 2, division of one half of the cells at generation 1 can give rise to cells with the + phenotype (these cells are +/-/-/-/-/-/-/) leading to R = 0.25. In generation 3, the 25% + phenotype population from generation 2 can segregate 1/8 + 1/8 = 25% +/-/-/- → +/-/+/-/-/-/-/- = 6.25% and 1/8 * 1/8 = 1.5625% +/-/-/- → +/-/+/-/-/-/-/-, leading to a total + population in generation 3 of R = 0.25 * (0.25 + 0.015625) = 0.0625 + 0.00390625 = 0.06640625.

NB: These two heterozygote populations need to be treated separately as their segregation probabilities differ based on the different allele frequencies, so in generation 4 we have: 1) 25% of 25% + phenotype population from generation 3 can segregate just
as in generation 3 so 25% of 25% of 25% regenerates +/+-/+-/+-/-/- = 1.5625%; 2) 25% of 25% of 37.5% regenerates +/++/+/-/-/-/- = 2.34375%; 3) 25% of 25% of 18.75% regenerates +/++/+/-/-/-/-/- = 1.171875%; 4) 25% of 25% of 6.25% generates a homozygous population of +/++/+/-/-/-/-/- = 0.003906% that will only increase and 5) 25% of 25% of 6.25% generates a homozygous population of -/-/-/-/-/-/-/- that will only increase. Now that we have populated the entire matrix of allelic states, this pattern will run to fixation of the homozygous + and homozygous – populations as above.

This is described by:

\[ HET_4 = 0.25^3 + 0.25^2 * 0.375 + 0.25^2 * 0.1875 = 0.25^2 * 0.8125 = 0.0546875 \]

\[ HET_5 = 0.25^3 * 0.8125 = 0.012695313 \]

\[ HET_N = 0.25^{N-2} * 0.8125 \]

AND

\[ HOM_4 = 0.25^2 * 0.0625 = 0.00390625 \]

\[ HOM_5 = HOM_4 + (0.25^3 * 0.0625) = 0 \]

\[ HOM_N = \text{sum}(HOM_4: HOM_{N-1}) + 0.25^{N-2} * 0.0625 \]
To calculate $R$ at any generation, sum HET and HOM for that generation number:

\[
\text{e.g. } R_4 = \text{HET}_4 + \text{HOM}_4 = 0.0546875 + 0.00390625 = 0.0546875.
\]

As HET approaches 0, HOM fixes at a value approaching $R = 0.005208336$.

As stated above, the recombinant segregation pattern of the data reaches a limit predicted by $N = 4$. This is the anticipated result if there are two genome equivalents of the target gene replicating during the experiment, as was reported by Sergueev et al. (2002).

The time course experiments showed that to quantitatively account for all the cells that experienced Recombineering, cells should be plated before they start dividing during the initial lag period. All the subsequent experiments were performed with outgrowth time of 30-60 minutes before plating. The following day, recombinant colonies appeared “sectored” with a mixture of recombinant cells and non-recombinant cells. Note that once we were able to record all the cells that experienced Recombineering, the data indicated about 15 to 30% Recombineering efficiency. These frequencies were among the highest Recombineering frequencies ever reported and indicated that we had optimized the Recombineering protocol, although not to 100%.

**Transformation efficiency**

After optimizing the above parameters for Recombineering, there were still fluctuations between experiments. The efficiency of introducing oligos into cells determines the amount of substrate available for Recombineering. Therefore, an important step of the protocol is getting the oligo Recombineering substrates inside the cells. We saw a positive correlation (by a one-tailed nonparametric correlation test; $P = 0.0215$) between the efficiency of plasmid uptake (to estimate transformation) and Recombineering. We found that working with cells in early log phase improved both
transformation and Recombineering efficiencies. The efficiency of transformation was correlated to how competent and viable the cells were before electroporation. We determined that washes in water produced the highest transformation efficiency, but that cell viability was compromised within minutes (viability dropped by 50% in 20 minutes). When cells were washed in 10% glycerol the viability was more stable, but the transformation efficiency was less than with just water. From the results of these studies, the protocol was adjusted to be rigorous in stipulating that cells be grown to early log, made competent with washes in cold water for the highest transformation efficiency and electroporation was to be performed no more than 30 minutes after the first wash to maintain high viability.

Since Recombineering can only happen when cells are transformed with the oligo substrate, the Recombineering experiments were normalized to the transformation efficiency of the cell preparation. The percent of competent cells within a preparation was examined by the ability of cells to produce colonies on Ampicillin plates after transformation with pUC19. The transformation efficiency was calculated by dividing the titer of Ampicillin resistant colonies by the total cell titer of the culture. Recombineering data fluctuations decreased after data were normalized to transformation efficiency. Of the cells that uptake DNA, ~ 40 % did Recombineering (Figure 4.5 C).

A co-transformation experiment demonstrated that selection for transformants increases Recombineering efficiency, but further studies need to be done to titrate the oligo and the dsDNA plasmid used in co-transformation as there may be competition between them.
Oligo substrate optimization

Oligo optimization was performed mostly at the design level. Four regions of the *gfpmut3* were targeted for mutagenesis (Figure 4.3). Two regions were targeted for color changing and two regions were targeted to create or fix non-sense mutations (stop codons). The color changes were created with oligos that targeted the chromophore region to change amino acids 64-66 and amino acid 203, which interacts with the chromophore affecting the spectra of the protein. The nonsense mutations were created at amino acid positions 39 and 59. The oligos were designed with identical homologous targeting “arms” on both sizes of the modified sequence. The target sequence *per se* was not found to affect the efficiency of Recombineering. However, the Recombineering frequency was indeed affected by which DNA strand was targeted and the mismatch introduced to produce the change.

Strand bias in Recombineering

There is a strand bias reported in Recombineering (Ellis *et al.* 2001; Constantino and Court 2003; Li *et al.* 2003) where oligos targeting the lagging strand template during replication produce more recombinants. In this study, we evaluated the strand bias by comparing Recombineering with oligos targeting each strand. One oligo was designed as the sense DNA strand (also known as the coding sequence or Crick strand) and the other oligo was designed as the antisense DNA strand (also know as the reverse complement of the sense DNA strand or the Watson strand, or the template for RNA transcription). We compared the Recombineering frequencies when recombination intermediates contained a C.C mismatch. Cells were plated one hour after electroporation. Recombinants were scored on a plate reader for the lost of green fluorescence in the first experiment and for gain of green fluorescence in the second experiment. As in the previous experiment, the
sense oligo, predicted to hybridize to the lagging strand template, produced more recombinants than the antisense oligo targeted to the leading strand template (Figure 4C). Targeting the lagging strand template was 70 fold more efficient than targeting the leading strand template.

While we obtained a 70-fold recombination increase when targeting the lagging strand template, other labs (eg. Costantino and Court 2003) had reported a 40-fold increase when targeting the lagging strand template. This discrepancy might be because we used different assays to evaluate leading vs lagging strand targeting. In the experiment where the leading strand template was targeted, GFPmut3* codon 39 was changed from tyrosine to a stop codon and the colonies with dark sectors were scored as recombinants. In the experiment where the lagging strand template was targeted, the stop codon was reverted to tyrosine and colonies with bright green sectors were scored as recombinants. It is possible that the bright sectors were easier to observe than the dark sectors, therefore we cannot discard the idea that targeting the leading strand template Recombination frequencies was underestimated. These data guided the design for human Recombineering studies by indicating that it was ideal to use a reporter that screen for “bright” recombinants and substrates targeting the lagging strand template when known.

MMR effect in Recombineering

Costantino and Court (2003) have shown that DNA MMR is a potent inhibitor of heteroduplex recombination intermediate stability. They have also shown that the rate of correction differs by the nature of heteroduplex, including bases on the mismatch and length of the heteroduplex. Our purpose was to evaluate the extent to which the MMR system impairs recombination in the process of changing the coding sequences of the
fluorescent genes. The mutS gene encodes the specificity subunit of the MMR complex (Kunkel and Erie 2005). In order to study the effect of the MMR complexes on Recombineering, the mutS gene was replaced with the Kan gene via transduction using a P1 phage lysate grown on MG1655* ΔmutS::Kan. Recombineering experiments were carried out in MutS⁺ and MutS⁻ strains using the same oligos and experimental conditions. The recombination frequencies are shown in Figure 4. 4D and the P values comparing each isogenic MMR⁺ and MMR⁻ strain are shown in Table 4. 1. In Figure 4. 4D the data for the MutS⁺ strains are in black and the data of the MutS⁻ strains are in red. C.C and C.T mismatches were formed in heteroduplexes that change a stop codon to tyrosine in gfpmut3*, restoring the expression of GFP. CC.AC and CC.AT mismatches were targeted to change a tryptophan to a tyrosine at the same site, changing CFP to GFP. A GaCTGGaCC.TtCGGTtAC mismatch bubble was targeted to change amino acids LTW to FGY, changing CFP to GFP. The data show that C.C is not affected by MMR, but the other single base substitution C.T was significantly affected. CC.AC was not affected but CC.AT was significantly affected. The data also shows that longer mismatches are not very correctable by MMR (the median recombination frequencies are not different by Kruskal-Wallis ANOVA; P = 0.4366).

The observations from these experiments reveal that oligos making larger mismatches have lower efficiency than do poorly corrected small mismatches. Well-corrected small mismatches have the lowest efficiency, corroborating the claim that MMR is a potent inhibitor of Recombineering. C.C mismatches always help to increase the efficiency of Recombineering. If it is desirable to use Recombineering to introduce single nucleotide
substitutions, well-corrected alleles may be introduced at high efficiency in a mutS mutant host.

**Table 4.1 MMR differently affects different mismatches.**
The influence of MMR on recombination was evaluated in isogenic MMR⁺ and MMR⁻ (ΔmutS::Kan) strains (Figure 4. SC). A one-tailed unpaired t test with Welch's correction was performed for crosses in each of the MutS⁺ and MutS⁻ pairs. This table reports the P values for each MMR⁺ and MMR⁻ pair and the degree of significance is noted using (*).

<table>
<thead>
<tr>
<th>Mismatch</th>
<th>Oligo</th>
<th>MMR⁺</th>
<th>MMR⁻</th>
<th>Phenotype</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>C.C</td>
<td>11</td>
<td>RIK423</td>
<td>RIK427</td>
<td>dark to green</td>
<td>0.4972</td>
</tr>
<tr>
<td>T.C</td>
<td>22</td>
<td>RIK423</td>
<td>RIK427</td>
<td>dark to green</td>
<td>0.003 (**)</td>
</tr>
<tr>
<td>G.G</td>
<td>58</td>
<td>RIK410</td>
<td>RIK417</td>
<td>dark to green</td>
<td>0.0417 (*)</td>
</tr>
<tr>
<td>AC.CC</td>
<td>16</td>
<td>RIK411</td>
<td>RIK418</td>
<td>aqua to green</td>
<td>0.4607</td>
</tr>
<tr>
<td>AT.CC</td>
<td>23</td>
<td>RIK411</td>
<td>RIK418</td>
<td>aqua to green</td>
<td>0.3013</td>
</tr>
<tr>
<td>GG.TG</td>
<td>56</td>
<td>RIK410</td>
<td>RIK417</td>
<td>green to violeta</td>
<td>0.0763</td>
</tr>
<tr>
<td>GGA.TA</td>
<td>56</td>
<td>RIK446</td>
<td>RIK470</td>
<td>amarillo to ultramar</td>
<td>0.0196 (*)</td>
</tr>
<tr>
<td>ACGtGG. CCAaTA</td>
<td>57</td>
<td>RIK410</td>
<td>RIK417</td>
<td>green to mar</td>
<td>0.32</td>
</tr>
<tr>
<td>ACGtGG.CCaTG</td>
<td>57</td>
<td>RIK446</td>
<td>RIK470</td>
<td>amarillo to bronze</td>
<td>0.0491 (*)</td>
</tr>
<tr>
<td>TtCGGtAC.GaCTGGaCC</td>
<td>13</td>
<td>RIK411</td>
<td>RIK418</td>
<td>aqua to green</td>
<td>0.427</td>
</tr>
<tr>
<td>CtgACCtGG.AaGCCaTA</td>
<td>1</td>
<td>RIK410</td>
<td>RIK417</td>
<td>green to aqua</td>
<td>0.0525</td>
</tr>
<tr>
<td>CtgACCtGG.AaGCCAaTG</td>
<td>1</td>
<td>RIK446</td>
<td>RIK470</td>
<td>amarillo to aqua</td>
<td>0.0754</td>
</tr>
</tbody>
</table>

**Oligo length**

Oligos were designed as suggested by (Sawitzke et al. 2011) with ~35-48 nt homology arms to the target locus with the desired change in between the arms (Figure 4. 3). We briefly examined the effect of the homology arm size on Recombineering efficiency. We found no statistically significant impact on efficiency over a range of 67-97 nucleotides. Smaller ~70 nt oligos were preferred substrates as they were easier to design and presented fewer issues with secondary structure than longer ~90 nt oligos.

**Oligo concentration**

During the Recombineering experiments ~100 ng of oligo was added to 10¹¹ cells. We did some preliminary titration with oligo concentration ranging from 100 to 400 ng and found no advantage to addition of more oligos.
Optimized protocol for fluorescent protein engineering

Optimal fluorescent protein engineering rates are reached when attention is placed in the design of the Recombineering substrate. Oligos that target the lagging strand template and avoid MMR will produce the highest Recombineering frequencies in agreement with Pear et al. (1993), Costantino and Court (2003b) and Rio et al. (1985).

Fluorescent protein engineering was ~70-fold more efficient when the oligos were targeted to anneal to the lagging strand template. Therefore, when Recombineering, it is important to consider the direction of the replication fork related to the target gene. The gfpmut3* transcription unit was found to be in the opposite orientation to the replication fork travel in the strain SCC1 and RIK410. For targeting gfpmut3* in RIK410, oligos should be designed with the sequence of the sense strand of the gene which could be paired with the lagging strand template.

The efficiency of fluorescent protein engineering depends a great deal on the mismatch required for the color swap. Mutagenesis efficiency could be improved by altering the nature of the mispaired heteroduplex recombination intermediate. When possible it is better to work with less correctable mismatches or by introducing a “bubble” of 2-9 unpaired bases at the center of the oligo. When the desired mutations are predicted to be MMR-correctable, it is recommended to work in an MMR negative strain such as the ΔmutS::Kan strain RIK417.

The optimized protocol for fluorescent protein engineering is described in the Materials and Methods chapter. In this study we corroborated that Recombineering reporter strain growth is reproducible if the Recombineering strain is kept in the freezer and thawed for each experiment. Growing cells until early exponential phase (OD$_{600}$ = 0.5 - 0.6) produces more recombinants than when cells are in the late exponential phase.
We found that transformation efficiency is a rate-limiting step and that transformation efficiency is the highest when cells are washed in water but that viability is maintained by washing cells with 10% glycerol at the expense of some transformation efficiency. By plating cells 30 minutes after electroporation, it is possible to enumerate all the recombinant cells with recombinant colonies developing with sectors of recombinants and non-recombinants. We found that the Dark Reader imaging box was a convenient tool to identify fluorescent *E. coli* colonies.

With the improved protocol the efficiency of *in vivo* site-directed mutagenesis approaches 40% of all transformed cells. The optimized frequencies for fluorescent protein engineering using Recombineering are reported in Table 4.2. The most critical parameters for successful mutagenesis via Recombineering are summarized in Table 4.3.

**Recombineering mechanism**

Fluorescent protein engineering using Recombineering corroborated the strand bias during Recombineering. Oligos that hybridize to the lagging strand template at the replicating target sequence are the best substrate for fluorescent protein engineering by Recombineering.

Changing color of fluorescent proteins involving different heteroduplex intermediates corroborated that MMR is a potent inhibitor of Recombineering. Different mismatches are removed with different frequencies, while C.C mismatches escape MMR. When MMR is removed, single nucleotide changes are the most efficient and longer nucleotide changes are lest efficient.
Finally, if the heteroduplex is not repaired, the strand that incorporated the oligo will segregate with cell division to produce a mixed population of recombinants and non-recombinants. Recombinants will segregate from the non-recombinant DNA to about 1/4 - 1/8 in experiments performed in exponentially growing *E. coli*.

**Table 4.2 Optimized fluorescent protein engineering frequencies.**
P₀ indicates the initial phenotype and Pᵣ indicates the phenotype of Recombinants. All these experiments were performed targeting the lagging strand template.

<table>
<thead>
<tr>
<th>P₀</th>
<th>Pᵣ</th>
<th>% Rec in MMR⁺</th>
<th>% Rec in MMR⁻</th>
<th>Strain: MMR⁺/MMR⁻</th>
<th>Oligo</th>
</tr>
</thead>
<tbody>
<tr>
<td>GFPmut3*</td>
<td>Aqua</td>
<td>40.1 ± 3.97 (3)</td>
<td>32.5 ± 3.20 (3)</td>
<td>RIK410 / RIK417</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Violeta</td>
<td>7.38 ± 2.40 (3)</td>
<td>31.8 ± 16.0 (3)</td>
<td>RIK410 / RIK417</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>Mar</td>
<td>30.1 ± 4.91 (3)</td>
<td>18.6 ± 11.7 (3)</td>
<td>RIK410 / RIK417</td>
<td>57</td>
</tr>
<tr>
<td></td>
<td>Dark</td>
<td>3.39 ± 2.02 (3)</td>
<td>53.6 ± 18.8 (3)</td>
<td>RIK410 / RIK417</td>
<td>58</td>
</tr>
<tr>
<td>Aqua</td>
<td>GFPmut3*</td>
<td>19.5 ± 6.66 (3)</td>
<td>20.0 ± 2.40 (3)</td>
<td>RIK411 / RIK418</td>
<td>16</td>
</tr>
<tr>
<td>Amarillo</td>
<td>Bronze</td>
<td>8.03 ± 3.50 (3)</td>
<td>19.1 ± 4.53 (3)</td>
<td>RIK446 / RIK470</td>
<td>57</td>
</tr>
<tr>
<td></td>
<td>Ultramar</td>
<td>6.60 ± 3.10 (3)</td>
<td>29.4 ± 8.49 (3)</td>
<td>RIK446 / RIK470</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>Aqua</td>
<td>32.4 ± 7.23 (3)</td>
<td>25.5 ± 4.51 (3)</td>
<td>RIK446 / RIK470</td>
<td>1</td>
</tr>
<tr>
<td>Dark</td>
<td>GFPmut3*</td>
<td>40.9 ± 8.92 (4)</td>
<td>40.5 ± 10.8 (4)</td>
<td>RIK423 / RIK427</td>
<td>11</td>
</tr>
</tbody>
</table>

**Table 4.3 Rules for high efficiency Recombineering.**

<table>
<thead>
<tr>
<th>Rule</th>
<th>Advice</th>
</tr>
</thead>
<tbody>
<tr>
<td>One</td>
<td>Use dividing cells to enrich for replication intermediates.</td>
</tr>
<tr>
<td>Two</td>
<td>Use oligos complementary to the lagging strand template.</td>
</tr>
<tr>
<td>Three</td>
<td>Use oligos that incorporate a C•C mismatch in the paired recombination intermediate. Alternatively, use a 2+ nucleotide mismatch or knock out MMR.</td>
</tr>
<tr>
<td>Four</td>
<td>Plate transformants within 30-60 minutes of outgrowth.</td>
</tr>
</tbody>
</table>

**Things to consider for Recombineering in human cells**

During the fluorescent protein engineering studies in bacteria it was evident that Recombineering frequencies are more accurately estimated when the reporter is associated with a gain of function upon recombining. The reporter is even more precise when the oligo produces a change that is brightest with respect to the non-recombinants.
MMR is a potent inhibitor of Recombineering. If it is possible, one should develop a reporter to correct a C.C mismatch. If a C.C mismatch does not produce the desired outcome, one should design oligos that produce a bubble that is less correctable as suggested by (Dekker et al. 2006), knock down MMR, or use a cell line that is deficient in MMR.

Since in Recombineering the oligo substrate is annealed to the nascent DNA strand on the context of the replication fork, it is very important that the oligo is provided just when cells start duplicating the DNA (early S phase). The positive recombinant cell that escapes MMR is expected to segregate into 3 non-recombinants and 1 recombinant cell. Therefore, we should expect that the frequency of Recombineering would look higher right after Recombineering and before 2 duplication cycles. We should expect that after 2 duplication cycles the percent of recombinants should stay constant, assuming one target per cell.

Transformation of ssDNA is a limiting step. Transformation of human cells with the oligo substrate should be optimized and evaluated to provide the highest percent of cells with the Recombineering substrate. It is recommended to initially evaluate human Recombineering in a cell line that is easy to transform, divides quickly and is MMR deficient. Such a cell line is HEK293T, as we will describe later.

**Host-specificity of Recombineering in bacteria**

The “host specific Recombineering” hypothesis was intended to be tested by comparing bacterial Recombineering using the phage λ recombinase vs the HHV1 recombinase. For these purpose, plasmids expressing inducible ICP8 in *E. coli* were created (pFN-ICP8 and pFN-ICP8-GFP) and introduced into bacteria engineered to
improve human gene expression (BL21 (DE3) pRARE6). ICP8 expression was evaluated and faint bands of the expected size were detected (Figure 4.7). ICP8 expression in *E. coli* needs to be optimized, or it might be that with only small expression is enough to test Recombineering in *E. coli*. Another possibility is to test Herpes SynExo activity in *E. coli* by performing phage λ crosses in a λ hybrid with UL12 and UL29 genes (Zissler *et al.* 1971). So far, the functionality of the Herpes SynExo activity in *E. coli* has not been tested, therefore the question about Host Specificity of Recombineering in *E. coli* remains to be tested.

<table>
<thead>
<tr>
<th>pFN22K::ICP8-GFP</th>
<th>-</th>
<th>+</th>
<th>+</th>
<th>+</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPTG (mM)</td>
<td>0</td>
<td>0</td>
<td>0.1</td>
<td>1</td>
</tr>
</tbody>
</table>

*Figure 4.7 ICP8 expression in E. coli*
BL21 (DE3)/pRARE6 was transformed with pFN22K::ICP8-GFP. Colonies were grown o/n. Cultures were diluted 1/100 and grown for 2 hours. 10 ml culture was induced with 0.1 or 1 mM IPTG for 2 hours. Cell pellets were collected, washed and frozen. Protein was extracted with BugBuster, Benzonase, 100 μg/ml Lysozyme and 1 mM DTT.
Conclusions

This is the first description of using fluorescent protein engineering as a quantitative assay for Recombineering efficiency. Recombineering efficiency was optimized and appears to be limited by transformation efficiency and the structure of the heteroduplex recombination intermediate. The efficiency of fluorescent protein engineering could reach ~40 % of transformed cells.
CHAPTER 5. HUMAN RECOMBINEERING

Developing the tool kit for human Recombineering

Human Recombineering was evaluated by targeting ssDNA oligos to fluorescent protein transgenes in a human cell line in the presence of a viral Synaptase. The minimal components of the human Recombineering tool kit were:

1. Assay: Fluorescent protein engineering
2. Reporter cell lines: 293T and MIAMI cells
3. Recombineering substrates: ssDNA oligos
4. Viral Synaptases: ICP8 and HumBeta

Assay: Fluorescent protein engineering

In previous chapters we described the development of fluorescent protein engineering by Recombineering in *E. coli* as an assay for Recombineering in human cells. Fluorescent protein engineering was performed using the gfpmut3* gene variant, well expressed in *E. coli*. In this chapter we extrapolated the lessons from Recombineering in bacteria to Recombineering in human cells. To develop human Recombineering we sought to use a fluorescent protein reporter that is well expressed in human cells by including a human codon optimized fluorescent protein gene sequence and a Kozak sequence that promotes translation initiation.

As a reporter for human cell Recombineering, we chose the enhanced green fluorescent protein gene (*eGFP*) since it is well expressed in human cells. As a vector backbone for the transgene we choose a lentivirus since they are very efficient at integrating the transgene into the genome of human cells and because transduction is less toxic to cells than are transfections. We were also interested in a stable promoter that is
not shut down by heterochromatin. pNL-eGFP/CEF was the lentivirus vector of choice, since it expressed \textit{eGFP} from a CMV-EF1\textsubscript{\alpha} hybrid (CEF) promoter, which was reported to be the best for stable expression of transgenes when compared to the PGK, CMV, Ub, EF1\textsubscript{\alpha} and MHC promoters (Arsenault 2004). A schematic of this lentiviral vector is shown in \textbf{Figure 5. 1}.

![Lentiviral vector schematic](image)

\textbf{Figure 5. 1 Lentiviral vector to deliver an eGFP variant Recombineering target transgene.} pDUAL-eGFP was created from pNL-eGFP/CEF by adding a T7 promoter, a Shine-Dalgarno sequence, and a Histidine peptide at the 5' end of eGFP. The transgene transgenic construct includes a SV40 origin of replication and expression of the fluorescent protein gene from a CMV-EF1\textsubscript{\alpha} hybrid promoter.

The sequence of the human Recombineering target gene of choice (\textit{eGFP}) indicated that the protein product was slightly different at the amino acid level than was the bacterial Recombineering target product GFP\textsubscript{mut3*}. Both proteins have the amino acid Tyrosine at position 66, which contributes to green fluorescence, but some of the amino acids surrounding residue 66 were different (\textbf{Figure 5. 2}). Closely related fluorescent protein genes with spectra different than eGFP show that there are a few amino acid
differences between the protein variants that could account for the different fluorescent spectra. I focused on the sequence differences in and near the chromophore region, as I previously did in bacteria and as suggested by Tsien (1998) and Zacharias (2006). Between the variants, a Cyan fluorescent protein gene \((cfp)\) encoded a Tryptophan at position 66, a Blue fluorescent protein gene \((EBFP2)\) had a Histidine at position 66, and a Yellow fluorescent protein had a Tyrosine instead of a Threonine at position 203, (Figure 5.2). The CFP protein chromophore region was similar to the Aqua protein created by us from \(gfpmut3^*\). The EBFP2 fluorescent protein differs from our Violeta fluorescent protein in that it has a Histidine at position 66 along with other residues in the chromophore. The yellow fluorescent protein was consistent, keeping the same chromophore region as the green, but changing the amino acid 203 from threonine to tyrosine.

The most convenient way to validate if it was possible to change the fluorescent properties of eGFP by targeting the chromophore regions was to use plasmid Recombineering in \(E. coli\). Unfortunately, pNL-eGFP/CEF did not express eGFP in bacteria and it was not possible to phenotypically evaluate new fluorescent proteins created from this plasmid in \(E. coli\). Therefore, a T7 promoter and Shine-Dalgarno sequence were added in between the mammalian CEF promoter and the \(egfp\) coding region to express eGFP and the variants in \(E. coli\). We called this mammalian/bacterial dual expression lentiviral vector pDual-egfp (Figure 5.1). Once it was possible to screen for fluorescent protein expression in \(E. coli\), different eGFP fluorescent protein variants from pDual-egfp were made in \(E. coli\). The \(egfp\) gene was then mutated at position Y66
to Histidine, to Tryptophan and to a stop codon for a dark version. The position T203 was also mutated to Tyrosine.

<table>
<thead>
<tr>
<th>Protein ID</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>His6-eGFP</td>
<td>HGSSH1HHHHHSSGLVPRGS1HMASMTGGQ0MGRGDPPVAT</td>
</tr>
<tr>
<td>GFPmut3</td>
<td>M4KG1E1L1G1P1V1L1V1E1D1G1L1D1</td>
</tr>
<tr>
<td>158327647 YFP</td>
<td>M4KG1E1L1G1P1V1L1V1E1D1G1L1D1</td>
</tr>
<tr>
<td>50365745 ICFP</td>
<td>M4KG1E1L1G1P1V1L1V1E1D1G1L1D1</td>
</tr>
<tr>
<td>240017714 BFP</td>
<td>M4KG1E1L1G1P1V1L1V1E1D1G1L1D1</td>
</tr>
</tbody>
</table>

Figure 5. Chromophore region determines the spectral properties of fluorescence.

Protein sequences of the closest variants to eGFP that have unique spectra were aligned in ClustalW. Amino acids that uniquely identify determinants of specific colors are shaded in the related color. Amino acid differences that do not uniquely identify color determinants are shaded in gray.

Mutagenesis of pDual-egfp was carried out by using both Recombineering in E. coli and by site-directed mutagenesis in vitro to compare both methods. While obtaining the mutants by site-directed mutagenesis was straightforward following the instructions of the kit manufacturer, Recombineering was more difficult. Site-directed mutagenesis selects for mutants by DpnI digestion while in Recombineering there was no selection for
recombinants. Creating these mutants by Recombineering also faced all the *caveats* of this method:

- Lentiviral plasmids have LTR repeats, which can recombine with each other and delete the fluorescent protein gene and its promoter (Lindley, Lu, Geffin and Myers, in prep.).

- The plasmid has a multicopy origin of replication; therefore, the pool of plasmids that must be screened to find recombinants was large.

- The plasmid origin of replication was unidirectional, and the oligos used to target the lagging strand template created heteroduplexes that were highly correctable by MMR.

To overcome all these limitations, Recombineering was performed in *E. coli* DH10B \([\lambda cI^{857} \Delta(cro-bioA) \leftrightarrow tetRA] recA^- mutS^-\). This strain transiently expresses the Recombineering enzymes upon a heat shock. It has a *recA* deletion to decrease LTR recombination and plasmid concatamerization. The *mutS* gene was also deleted to disable Mismatch Repair and stabilize heteroduplex recombination intermediates.

Mutagenesis was performed by co-electroporating a substoichiometric amount of pDual-egfp (< 1 plasmid molecule per cell) and saturating amounts of oligos. By co-electroporating the plasmid and oligos, Recombineering was performed under conditions that reduced the initial copy number of the plasmid (Thomason *et al.* 2007). The cells were grown overnight to allow recombinant plasmid strains to segregate from non-recombinant plasmid strains. Plasmids were isolated and transformed into the expression strain at substoichiometric amounts to obtain only one plasmid per cell, or one plasmid clone per colony. The resulting colonies were evaluated by loss of the green fluorescence by screening on the Dark Reader box. Darker colonies were isolated and genotyped.
Figure 5.3 Fluorescent proteins created from egfp by Recombineering and SDM have different spectra. Recombineering was performed by co-electroporating a substoichiometric amount of pDual-egfp and saturating amounts of oligos to RIK473 [E. coli DH10B [λ, clp32 Δ(cro-bioA) < tetRA] recA mutS]. Site-directed mutagenesis was performed using the QuickChange Lightning Site-Directed Mutagenesis Kit #210518-5 following the manufacturer’s instructions. Plasmids were isolated from the Recombineering strain or site-directed mutagenesis mix and transformed into the expression strain RIK276 at substoichiometric amounts. Protein extracts were collected using BugBuster Protein Extract Reagent as indicated by the manufacturer (70584, Novagen). The excitation and emission spectra of eGFP fluorescence and its variants in the extracts were determined with a PTI QuantumMaster spectrofluorometer. The data were collected using Felix32 software, and later analyzed using Microsoft Excel and GraphPad Prism. Each spectrum was normalized to its peak value and plotted.

Recombinants and site-directed mutagenesis mutants were validated by allele-specific PCR. Protein extract spectra were evaluated for each mutant (Figure 5.3). The H<sub>66</sub> variant shifted the spectra towards blue, the W<sub>66</sub> variant shifted the spectra towards cyan, the Stop<sub>66</sub> variant was completely dark and the Y<sub>203</sub> variant shifted the spectra towards
yellow. The new eGFP variants were named: Azure (H_{66}), Celeste (W_{66}), Dark (Stop_{66}) and Mostaza (Y_{203}). Unfortunately, the changes introduced in the eGFP gene compromised the fluorescence intensity of the variants, indicating that some of the other changes in the original eBFP and cfp genes were also important for brightness. The best of the variants was Mostaza, with 1/10 of the fluorescence intensity of eGFP. The DNA and protein sequence for each variant is illustrated in the Appendix B: Fluorescent Protein Sequences.

![Figure 5.4 eGFP variants at the flow cytometer](image)

Figure 5.4 eGFP variants at the flow cytometer

Cells were harvested, washed with PBS and resuspended in DMEM. Samples were analyzed at the BD LSR-Fortessa-HTS, 19-color, high-throughput, flow cytometric analyzer. Cells were excited with a UV laser of 350 nm and the blue laser of 488 nm. The data were analyzed using FACSDiva Version 6.1.3 software. Green fluorescence is represented on the X axis and blue fluorescence is represented in the Y axis. Using these settings, Azure cells can be separated from the other fluorescent cells.

The expression of the eGFP variants in human cells were tested by transient transfection in 293T cells and evaluated by microscopy and flow cytometry (Figure 5.4).
We then generated lentivirus particles expressing the different variants and transduced the human cells of choice for Human Recombineering.

**Reporter cell line**

Cell line of choice for human Recombineering

Following the lessons learned from bacterial Recombineering the ideal cell line to evaluate Human Recombineering should be actively replicating, easy to transfect and MMR deficient. In addition we sought a cell line that is genetically stable to decrease the noise in the data and unintended recombination events. Ideally, the innate immunity expressed in the human cell line should not be responsive to cytoplasmic DNA and to lentiviruses. We were also very interested in applying Recombineering to adult human stem cells for their potential to generate disease models and autologous stem cell therapy and to evaluate the potential of Recombineering to revert a monogenic disease mutation such as the *LMNA* allele that causes Hutchinson-Gilford Progeria Syndrome (HGPS, Progeria).

293T cells are ideal for evaluating Recombineering

The human cell line of choice that grouped most of the desired features for evaluating Recombineering was the immortal HEK293T cell line. These cells replicate every ~18 hours, are easy to transfect and are defective for both MMR and innate immunity. In our hands 293T cells showed high transfection efficiency, they usually recovered well from transfection, and had ~10% of cells at the G1/S transition during steady-state growth.

293T cells originated from a highly transfectable derivative of the human primary embryonic kidney cell line HEK293 (ACC 305) by stable co-transfection of a plasmid encoding a temperature sensitive mutant of the SV40 large T-antigen (tsA1609) and a
neomycin resistance conferring plasmid. The temperature sensitive SV40 large T-antigen is expressed best at 33 °C, well at 37 °C, but not at 40 °C (Pear et al. 1993; Rio et al. 1985; DuBridge et al. 1987). At 40 °C, this protein is defective both for autoactivation of T Antigen transcription and makes a T antigen protein that is unable to bind the SV40 origin and initiate DNA replication.

The SV40 large T-antigen binds to p53 and inhibits its function (Levine 1997). Therefore, 293T cells can accumulate high levels of p53 protein without killing the host (onc2010457x8, nature supplementary material). The p53 protein senses genotoxic stress such as DNA breaks and stalled replication forks in response to damage-dependent activation of DNA-dependent protein kinases ATM and ATR which phosphorylate p53, stimulates its ability to bind DNA and trigger cell-cycle arrest or programmed cell death by apoptosis (Murray-Zmijewski et al. 2006). The T antigen blocks the p53 DNA binding domain (Levine 1997), thereby inactivating most of its functions.

The T antigen also interferes with cell cycle regulator retinoblastoma protein RB1/pRb and TP53 and promotes the G1 to S phase transition. In S phase it promotes a second round of DNA synthesis before cells enter mitosis (Friedrich et al. 1992; Perry & J. M. Lehman 1998).

293T cells are MMR deficient since they lack hMLH1, hPMS2 and hMLH3 proteins. The promoter of hMLH1 and hMLH3 is epigenetically silenced due to hypermethylation. Lack of hMLH1 leads to proteolytic degradation of hPMS2 (Trojan et al. 2002). These cells are also deficient for several innate immunity pathways like h-cGAS and STING, failing to induce IFN-β (Sun et al. 2013). cGAS is a cytosolic DNA sensor that induces interferon by producing the second messenger cGMP-AMP dinucleotide, which binds to
and activates the adaptor protein STING. Overexpression of cGAS activates the transcription factor IRF3 and induces interferon-β in a STING-dependent manner.

All the preceding characteristics of this cell line make it ideal to test human Recombineering. As per ATCC, 293T cells have a near-triploid karyotype where 30% of the analyzed cells had 64 chromosomes and 4.2% cells had higher ploidy. There were 3 copies of the X chromosome and none of the Y chromosome, indicating these cells were obtained from a female fetus. The 293T cells used here were a gift from Dr. Priya Rai and the passage number is unknown. It has been reported that tumor aneuploid cells like the 293T cells adapt to the culture conditions and deviate from the original strain phenotype (Gillet et al. 2011). Since we have not evaluated the 293T karyotype or genotype, we don’t know how much our cells have deviated from the evaluation by ATCC. Unfortunately, polyploid cells could be more prone to suffer gross genome rearrangements when viral synaptases are overexpressed. Nevertheless, Recombineering experiments were performed in the 293T cell line and the results are discussed in this chapter.

MIAMI adult stem cells

The human stem cell line of choice was the Human marrow-isolated adult multilineage inducible (MIAMI) stem cells (D'Ippolito et al. 2004; D'Ippolito et al. 2006). MIAMI cells have enormous potential to establish improved models of human genetic diseases and for developing autologous stem cell therapy whereby a patient's own stem cells will be corrected ex vivo for a genetic defect and then returned to the patient. Adult stem cells are small reservoirs of self-renewing cells that replenish and maintain tissues, especially after damage. They are present in a variety of tissues and are highly
dependent on their niche, from which they receive signals that influence their fate (Jones & Wagers 2008). The MIAMI cells used in this study are a homogeneous subpopulation of human stromal cells isolated from bone marrow as immature cells characterized by a unique molecular profile. The immature status of these cells is indicated by the expression of human embryonic stem cell markers like SSEA-4, Oct-4, Nanog, Bmi-1, Rex-1 and hTeRT. Those embryonic stem cell markers are maintained in long-term cultures (>30 population doublings) when MIAMI cells are cultured at low oxygen tension. MIAMI cells are related to mesenchymal stem cells (MSC), but resemble primitive stem cells in their capacity to differentiate in vitro into mature cells found in tissues derived from all three germ-layers. Continuous growth of MIAMI cells for 10 months allows isolation of $10^{23}$ cells with only low levels of cell cycle inhibitor proteins p53, p21, and p16, and rare expression of β-galactosidase activity, indicating little or no senescence.

MIAMI cells replicate actively, although they divide each 36 hours at the established growth conditions, with a low percentage of the population in early S phase at each particular time (Curtis 2010). The slow replication rate could limit the amount of cells that could incorporate the oligo at the target sequence at any giving time, unless cells are synchronized in late G1 before oligo delivery. Flagler et al. (2008) demonstrated feasibility of gene conversion in MSC, obtaining rates of $10^3$ using modified end-protected oligos.

In these studies MIAMI cells were transduced with the lentiviral gene targeting reporters pDual-Mostaza and pDual-eGFP (Figure 5.5). ssDNA oligo Recombineering substrates were transfected using Lipofectamine 2000 (Figure 5.7). The initial source of
the Synaptase was from a plasmid that didn’t contain a selection marker, but the transient transfection rates and post-transfection viability levels were so low that Recombineering experiments were not feasible. In general, MIAMI cells were very hard to transfect with dsDNA, maybe because of the innate immune response to cytoplasmic dsDNA. The efficiencies were very low using both transfection reagents and electroporation. We then developed a lentiviral inducible source of synaptase to avoid dsDNA transfections. Creation of new transgenic inducible synaptase stem cell lines is underway, pending optimization in a simpler transformed cell line.

**Figure 5.5** MIAMI human stem cells expressing eGFP and Mostaza
MIAMI cells were transduced with lentiviral particles of pDual-Mostaza and pDual-eGFP. Cells were grown on coverslips, fixed and stained with DAPI. Images were taken in a Leica SP5 confocal microscope using a 63x oil-immersion 1.3 numerical aperture objective. Confocal acquisition parameters were determined at the beginning of the study and the same parameters (e.g., gains, slit aperture, laser intensity) were used for all the images. Confocal optical sections were 0.6 μm thick. Field selection was performed using the DAPI channel to identify MIAMI cells. Adobe Photoshop was used to merge the colors. In blue, DAPI staining the nuclei, in green, MIAMI cells transduced with pDual-eGFP and in Yellow, MIAMI cells transduced with pDual-Mostaza. The bar in white at the lower right corner corresponds to 50 μM.

While the MMR status of these primary cell isolates is unknown at this time, it is likely that they are MMR-proficient, which could decrease Recombineering rates unless care is taken to design oligos to make uncorrectable alleles. We therefore designed oligos to change the Mostaza to green variant using a 4 nucleotide mismatch or bubble that is not as affected by MMR as are single and double nucleotide mispairs (Dekker et al. 2006). Interestingly, Rodríguez-Jiménez et al. (2008) reported how MMR activity is
lower in hypoxic (1% O$_2$) than in normoxic conditions. Since MIAMI cells are cultured at 3% O$_2$ tension, their MMR might be not so active under these conditions. It would be worthwhile to evaluate MMR at low oxygen tension in MIAMI cells.

While primary stem cells are karyotypically stable, they can spontaneously differentiate to other cell lineages if the culture conditions are not “just right”, so the number of experiments an investigator can do with a batch of cells are limited. Therefore, we sought to first develop the human Recombineering protocol in HEK 293T cells and then extrapolate it to MIAMI cells after all the kinks were worked out.

HeLa progerin reporter cells

The monogenic human disease model cell line used in these studies was the HeLa Progerin promoter cell line provided to us by Dr. Thomas Misteli at NIH/NCI. We were interested in evaluating Recombineering to treat monogenic diseases like Progeria and this cell line has a fluorescent Progeria reporter built into it. This cell line replicates every ~24 hours, which provides opportunities for Recombineering at the replication fork. However, plasmid transfection of the Synaptase gene was suboptimal with an average transfection efficiency of 5 to 10% of the cells.

MMR has been reported to be proficient in HeLa cells (Thomas et al. 1991), which can decrease the apparent Recombineering frequencies. The modal chromosome number for HeLa is 82 with a range of 70 to 164 chromosomes, as reported by ATCC. Gross chromosomal rearrangements have been reported for this immortal cancer cell line (Landry et al. 2013). Therefore, overexpressing a synaptase in this cell line could promote chromosome rearrangements, obscuring the clarity of the data. This cell line’s
features and the Recombineering data using the Progeria reporter will be discussed in more detail in the next chapter.

Building a human reporter cell line

Lentiviral stocks of the engineered pDUAL-egfp variants were produced in 293T cells and used to transduce 293T and MIAMI cells. From all the variants, cells transduced with the Mostaza and eGFP were the brightest. Even though the Mostaza and eGFP spectra are very close, the cells could be distinguished using confocal microscopy (Figure 5.5) and by flow cytometry (Figure 5.6). Since Mostaza fluorescence is only 10% as bright as eGFP, the Recombineering assays were performed by reverting the Mostaza fluorescent protein gene to eGFP via oligo-dependent gene conversion. The gain of green fluorescence in Mostaza cells was recorded using the Accuri flow cytometer in our lab and the gates presented in Figure 5.6.

The cells used for Recombineering were an unselected mixture of transduced and non-transduced Mostaza-293T cells. The insertion sites are likely to be diverse and possibly different in most cells while the average copy number of the transgene target has not been determined.

Figure 5.6 MIAMI-egfp and MIAMI-Mostaza cells are distinguishable by flow cytometry
The Accuri C6 flow cytometer was used to quantify Mostaza and eGFP MIAMI cells. Fluorescent excitation was produced with the 488 nm laser and emission was collected using 510 ± 7.5 nm and 540 ± 10 nm filters. Mostaza and eGFP cells showed a differential angular distribution that allowed differentiation and quantification of cells without compensation. Gates were set using untransduced MIAMI cells (a), MIAMI-Mostaza transduced cells (b) and MIAMI-eGFP transduced cells (c). Transfected Mostaza and eGFP cells were mixed to evaluate the ability of the Accuri C6 flow to separate them (d-k). Untransduced cells were quantified in the R2 gate, eGFP cells were quantified on the P4 gate and Mostaza cells were quantified from the P5 gate. The expected percent of eGFP in mixtures of Mostaza and eGFP cells was plotted vs the observed percent (l).
**Oligo substrates**

**Oligo substrate nuclear delivery protocol**

For efficient Recombineering, high transfection efficiency of the substrate oligo is required. DNA does not passively diffuse into cells at high enough rates to be useful for genome engineering. In bacteria, oligos were delivered by creating holes in the plasma membrane by electroporation. Once oligos pass the cytoplasmic membrane they are in the same compartment with DNA and DNA metabolic proteins and Recombineering efficiencies can be quite high. Our studies made it evident that DNA uptake was rate-limiting for Recombineering in bacteria and we anticipated that the same would be true for mammalian genome editing. In mammalian cells, oligo substrates not only need to cross the plasma membrane, but also need to escape cytoplasmic obstacles to transfection and localization to the nucleus were the target genome is.

Delivery of oligos to the nucleus of mammalian cells was evaluated by Flagler *et al.* (2008). They showed that when Lipofectamine 2000 was used to transfect modified oligos that contain four 2′-O-methyl RNA residues incorporated at each end to protect them from nuclease degradation, oligo/Lipofectamine 2000 complexes could repair genes at 0.3 % frequencies in mouse MSC. However, several other studies have shown that DNA modifications of this type reduce viability (Rios *et al.* 2012; Olsen *et al.* 2005a; Aarts & te Riele 2011) and likely lead to lower gene targeting frequencies. We evaluated if MIAMI cells could be transfected with fluorescein-conjugated oligos (FL-oligos) using Lipofectamine 2000. Transfection was performed following the instructions of the manufacturer using 280 nM FL-oligos and cells grown on cover slips. Cells were incubated with the oligo/transfection reagent for 5 hours, then fixed and stained with DAPI. Oligo delivery was evaluated using confocal microscopy (Figure 5. 7). FL-oligos
(represented in green) were observed inside cells. Nuclear localization was evaluated by colocalization with the blue fluorescence from DAPI and Z-sections. Inspection of Z sections of confocal planes and the colors overlap (teal) shows oligos localizing to the nucleus. Inspection of different fields indicated that all MIAMI cells took up FL-oligos and that each cell localized oligo DNA to the nucleus. This result showed that Lipofectamine 2000 could deliver oligos to the nucleus.

![Image](image_url)

**Figure 5.** Lipofectamine 2000 delivers oligos to the nucleus of human cells

MIAMI cells grown on cover slips were transfected with 280 nM fluorescein-conjugated oligos using Lipofectamine 2000. Complexes and cells were incubated for five hours. Cells were then fixed and mounted with DAPI mounting medium. Fluorescein emits green fluorescence and DAPI stains the nucleus in blue. The image was obtained using a Zeiss confocal microscope with a 20 X objective. Nuclear localization of oligos was evaluated by examining colocalization of fluorescein with the blue fluorescence from DAPI (left) and by examining Z-sections of confocal planes (right). This image was processed using Adobe Photoshop.

Unfortunately, the experience in the lab is that Lipofectamine 2000 is toxic. In a recent study, plasmid DNA/Lipofectamine 2000 complexes were shown to produce gene expression changes related to oxidative stress response, protein misfolding response and DNA damage response pathways (Fiszer-Kierzkowska *et al.* 2011). Many gene targeting
studies comment on the viability issues and most use 0.1 to 0.5 % Lipofectamine 2000 in the media to deliver oligos, with concentrations ranging from 3 to 280 nM and short incubations times ranging from 1 to 6 hours (Andrieu-Soler et al. 2005; Flagler et al. 2008; Campbell 2007; Bertoni et al. 2009; Rios et al. 2012). The Lipofectamine 2000 manufacturer recommends that cells be plated at 90 % confluency for higher viability, however when performing Recombineering, cells must be replicating in the exponential phase for maximum rates (according to the Canonical Model for Recombineering). While 293T cell growth might not be affected by contact inhibition since it is an immortal cell line, primary stem cell growth might be more affected, so this must be taken into consideration.

Delivery of FL-oligo to 293T cells using Lipofectamine 2000 was quantified using flow cytometry (Figure 5.8). We found that FL-oligo transfection peaks at 90%, that greatest levels of transfection is detected when Lipofectamine 2000 is 0.4 % in the medium, that Lipofectamine 2000 compromises viability while oligos do not, and that saturation of transfection is detected when the oligo is at 200 nM in the medium. The best combination of transformation efficiency and viability was obtained when the transfection complex included 200 nM oligo and 0.3 % Lipofectamine 2000.

### Table 5.1 Oligo transfection frequencies in 293T cells using Lipofectamine 2000

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Oligo (nM)</th>
<th>Lipofectamine 2000 (%)</th>
<th>Oligo transfection efficiency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gene targeting as in (Andrieu-Soler et al. 2005)</td>
<td>4.13 nM</td>
<td>0.13 %</td>
<td>≤ 61 %</td>
</tr>
<tr>
<td>FL-oligo titration: best ratio of viability and efficiency</td>
<td>200 nM</td>
<td>0.27 %</td>
<td>87.3 %</td>
</tr>
<tr>
<td>Conditions used for transfecting MIAMI cells (Figure 4.7)</td>
<td>280 nM</td>
<td>0.40 %</td>
<td>91.6 %</td>
</tr>
</tbody>
</table>
Figure 5. Lipofectamine 2000-oligo complexes transfected 90% of 293T cells. 293T cells were seeded in 48 well plates one day before transfection. FL-oligo and Lipofectamine complexes were performed as recommended by the manufacturer. Complexes were incubated with the cells for 4 hours. Cells were then collected, and washed with PBS two times before flow. A) FL-oligos 293T transformation efficiency. B) Relative amount of the cells after FL-oligo/Lipofectamine 2000 treatment. C) Representation of the flow data with some of the used or recommended conditions. The best combination of transformation efficiency and viability was reached when the transfection complex included 200 nM oligo and 0.3% Lipofectamine 2000.
Oligo design

Substrates for human gene targeting are usually small oligos in the 20 to 80 nt range with sequence complementarity to the target region (Nickerson & Colledge 2003; Andrieu-Soler et al. 2005; Bertoni et al. 2009). The desired change is placed in the middle of the oligo, with perfectly complementary arms at both ends because efficiency decreases when the length of complementarity is shorter than 9 nucleotides at both ends (Flagler et al. 2008). The nucleotides modified to introduce the genetic change were selected for lower mismatch repair recognition and higher codon usage correlation in human cells. All the oligos designed and synthesized for fluorescent protein engineering in human cells are represented in Figure 5.9.

Oligos used to change Mostaza to Green

The Recombineering substrates used to revert Mostaza to eGFP were ssDNA oligos, without any modification. These oligos had the sequence of the parental eGFP gene. They resemble the Mostaza sequence differing in the middle of the oligo (see Mostaza oligos in Figure 5.9). The region of homology provided the complementary sequence for annealing, and the different sequence in the middle determined green fluorescence.

Oligos to change Mostaza to eGFP were designed to produce a 4-5 nt bubble when hybridized to the template (Figure 5.10). The bubble of unpaired nucleotides encodes a synonymous substitution at the S_{202} codon and T_{203} (the determinant of green fluorescence) in place of the Y_{203} codon (the determinant of yellow or Mostaza fluorescence). The synonymous substitution was added to create a bubble that avoid MMR as suggested by (Dekker et al. 2006).
Figure 5.9 Oligos for fluorescent protein engineering in human cells.

Oligos used to change between Mostaza and eGFP genes and to modify the chromophore region of eGFP and its variants. Oligos that introduce a change in the target gene were used to evaluate Recombineering and oligos without any change were used as no change control. The legend shows the oligo number as in the Materials and Methods section, the distinctive amino acid, the strand specificity and the length in nucleotides. eGFP variants were created by Recombineering using these oligos. eGFP variants were used to quantify Recombineering in human cells by reverting them to eGFP.

Viral synaptases: ICP8 and HumBeta

While ssDNA oligo gene targeting works in human cells (Rios et al. 2012), ssDNA oligo Recombineering works much better in E. coli. HR is an important DNA break repair pathway for all creatures. In E. coli, the host synaptase protein (RecA) has a minimum effective pairing length of ~50 nt, but is greatly stimulated by increased homology up to ~500 nt (Shen & Huang 1986). In contrast, phage λ Beta protein can...
efficiently pair smaller complementary regions (~20 nt) and is stimulated by additional homology up to ~100 nt (Swingle et al. 2010b). In this study we aimed to evaluate if a transient burst of a viral synaptase could increase current rates of gene targeting in human cells to the levels of bacterial Recombineering.

There are several lines of evidence suggesting that bacterial Recombineering is host specific (Datta et al. 2008; van Kessel & Hatfull 2008; van Kessel et al. 2008; van Kessel & Hatfull 2007; Swingle et al. 2010a; Lesic & Rahme 2008; Swingle et al. 2010b) and that SynExos coordinate their activity with host factors (Chapter 7). When SynExos are expressed in cells different than their hosts, Recombineering efficiency is reduced and the degree of reduction is roughly correlated with the phylogenetic distance between the viral host and the test bacterium. In one study, host specificity was mostly evident when the substrate was dsDNA (a reaction that requires both exonuclease and synaptase functions) and less evident when the substrate was ssDNA (a reaction that only requires the synaptase) (Datta et al. 2008).
Interestingly, Zhang et al. (2003) showed that Beta was able to catalyze gene targeting in mouse ES cells, albeit at much lower frequencies than in its host, *E. coli*. From this article, it is not known if Beta protein was well expressed or well localized to the nucleus. The human viral synaptase ICP8 is an early gene that it is well expressed and localizes within the nucleus to replication compartments (Taylor & Knipe 2003; Taylor et al. 2003; Quinlan et al. 1984). The lower frequencies of Recombineering in mouse ES cells compared to the native Beta host *E. coli* could be because Beta fails to coordinate with cellular DNA metabolic pathways to catalyze Recombineering. Alternatively, it is possible that Beta is able to catalyze synapsis between oligos and a human genomic target, but needs to be better expressed and properly localized to the nucleus were the DNA target is for high rates of Recombineering.

This study aimed to evaluate if a viral synaptase can increase gene targeting rates above the endogenous rates catalyzed by host recombination proteins. Recombineering was evaluated using a human viral synaptase and a humanized, nuclear targeted, *E. coli* bacteriophage synaptase. We attempted to mimic the bacterial Recombineering protocol, where the synaptases are expressed in a burst of induction just before adding the oligo substrate for recombination. We were also interested in creating a cell line that encoded the Mostaza fluorescent protein target gene and an inducible form of the synaptases. To avoid the negative response from the innate immune response to transient dsDNA transfection and to eliminate the transfection efficiency variable when comparing Recombineering rates, a stable and inducible cell line was anticipated to avoid the negative influence of the innate immunity response to transient dsDNA transfection and eliminate the transfection efficiency variable when comparing Recombineering rates. We
therefore, created inducible systems expressing the synaptases to create the
“Recombineering Reporter Cell line”.

Source of viral synaptases

Plasmid clones bearing the ICP8 gene were requested from Dr. David Knipe at Harvard, since we didn’t want to mess with infectious HHV1 viruses. Dr. Knipe kindly provided us with the plasmids pCMV-ICP8 and pCMV-ICP8-GFP. The coding sequence of ICP8 and ICP8-GFP were used to produce plasmids pNEBRX1-Hygro-ICP8, pNEBRX1-Hygro-ICP8-GFP, pSLIK/TREPitt::Crimson-P2A/ICP8 and pSLIK/TREPitt::ICP8-P2A/Crimson.

The bacteriophage synaptase Beta gene was redesigned to be expressed in human cells. The sequence was optimized using human codon usage optimization software from Genewiz. The optimized Beta gene was fused to an SV40 NLS to promote localization of the bacteriophage synaptase to the nucleus as represented by Kalderon (1984). Next to the NLS, an HA epitope was also added to follow expression and localization of the humanized Beta protein. The HumBeta gene was synthesized by Genewiz and cloned in the plasmid pUC57-Kan::NLS/HA-HumBeta. The coding sequence of HumBeta was used to create the plasmids pSLIK/TREPitt::Crimson-P2A/HumBeta and pSLIK/TREPitt:HumBeta-P2ACrimson.

Expression and localization of viral synaptases

ICP8 and ICP8-GFP expression and localization were evaluated in 293T, MIAMI and HeLa cells using transient transfection with plasmids pCMV-ICP8 and pCMV-ICP8-GFP. ICP8-GFP was expressed in all three cell lines and showed nuclear localization in punctate subnuclear structures (see Appendix C, for a representation of ICP8 expression
in the above cell lines). The steady-state levels of ICP8 were high enough to detect between 8 hours to 9 days post-transfection. Interestingly, some cells showed diffuse nuclear ICP8-GFP while in other cells, ICP8-GFP fluorescence appeared as bright punctuate spots (Figure 5.11). de Bruyn and Knipe (1988) reported that ICP8 spots were associated with pre-replication centers. Alternatively, ICP8 may be stored in subnuclear compartments like P bodies being prepared for degradation.

Figure 5.11 ICP8-GFP localizes to the nucleus in MIAMI cells. Cells were transfected with pCMV-ICP8-GFP. 16 hours later cells were fixed with 4% PFA, stained with DAPI and mounted with Prolong Gold (Invitrogen). Images were taken with a Delta Vision OMXV4 fluorescent microscope and processed with their deconvolution software. On the top left is an image collected with the DAP filter (435 nm). On the top right is an image collected with the FITC filter (523 nm). On the bottom right is the merge of both images were the DAPI image was colored in blue and the FITC image was colored in green.

ICP8 and ICP8-GFP expression were evaluated by transfecting 293T with pCMV-ICP8 and pCMV-ICP8-GFP. Both proteins were expressed at the expected sizes of 129 and 154 kDa respectively when 293T cell extracts were evaluated by Western blot with an antibody against ICP8 (Figure 5.12). ICP8 expression was also evaluated by
microscopy. Images were collected in the Delta Vision microscope and processed with their deconvolution software. In this image, ICP8-GFP is localizing to the nucleus as indicated by the DAPI control. In this particular cell ICP8-GFP is sublocalized within the nucleus in big aggregates (Appendix C).

Figure 5. 12 ICP8 expression in 293T cells detected by Western blot
293T cells transfected with pCMV-ICP8 and pCMV-ICP8-GFP were collected 24 and 48 hours after transfection and analyzed by Western blot. ICP8 and ICP8-GFP were detected with an anti-ICP8 antibody. As a loading control α-tubulin was also probed in the same membrane.

Inducible synaptase-Recombineering reporter cell lines

RheoSwitch system from NEB

ICP8 and ICP8-GFP were cloned into the RheoSwitch system from NEB as pNEBRX1-Hygro-ICP8 and pNEBRX1-Hygro-ICP8-GFP. This system consists of two plasmids: pNEBR-R1, which expresses a Ligand-responsive transcription factor; and pNEBRX1-Hygro, which expresses the gene of interest (ICP8 or ICP8-GFP) from a promoter activated by the transcription factor when the ligand is added to cells. The RSL1 ligand is a synthetic ecdysone agonist of a Drosophila hormone receptor fused to the Yeast Gal4 transcription activator. This bipartite regulatory system is both highly specific (it shows no effect on endogenous transcription in the mammalian cell line
tested) and allows up to 10,000-fold regulation of gene expression. After successful cloning, the inserts and flanking regions were validated by sequencing.

Expression of ICP8 and ICP8-GFP from these plasmids was evaluated in 293T cells by Western Blot, microscopy and flow cytometry after cells were cotransfected with the regulatory and expression plasmid and subsequently induced by the ligand. ICP8 and ICP8-GFP were recognized by anti ICP8 antibody by WB and ICC. The Western Blot showed bands at the expected size, as seen before for pCMV-ICP8 and pCMV-ICP8-GFP. The ICC showed both proteins localizing to the nucleus. In some cells ICP8 distribution was more homogeneous within the nucleus while in other cells ICP8 formed punctate intranuclear structures (Figure 5.13). Flow cytometry showed that both the percentage of cells expressing ICP8-GFP and the intensity of GFP correlated with ligand concentration (Figure 5.14).

![Figure 5.13 ICP8 expressed from inducible RheoSwitch plasmids](image)

293T were co-transfected with pNEBR-R1 and pNEBR-X1-ICP8. 24 hours later inducer (RSL) was added to cells in the top panel. After 48 hours incubation, cells were processed for immunofluorescence. ICP8 protein was detected using an anti-ICP8 antibody (red) and nuclei were stained with DAPI. Cells were imaged using a Nikon fluorescent microscope.
Figure 5. 14 ICP8-GFP expression correlates with ligand concentration  
Cell detection correlates with the increase in fluorescence intensity. 293T cells co-transfected with RheoSwitch plasmids encoding ICP8-GFP were selected with 200 µg/ml Hygromycin. A) Cells were incubated with RheoSwitch ligand (RSL). B) GFP fluorescence intensity correlated to the percentage of transfected and inducible cells. GFP fluorescent cells were measured using the Accuri flow cytometer.

We tried to establish 293T cell lines with the pNEB system. Unfortunately, 293T were already Neomycin resistant which was the selection for the regulatory plasmid. We tried co-transfection of the Regulatory and Expression plasmids. Cells were selected for Hygromycin resistance, which selects for the expression plasmid. Clonal isolation became complicated because of a lack of a selection for the correct clone. The goal for the inducible cell line was to express ICP8 for the Recombineering experiment only. In this case the only way to determine that ICP8 was inserted in the chromosome was to
induce ICP8 expression and fix the cells to detect ICP8 by ICC or WB, which turned out to be labor intensive.

We tried to establish MIAMI cells with the pNEB system, but the transfection efficiency for both plasmids was low and, although selection can be applied to establish the cell lines, MIAMI cells do not respond well to drug selection. In conclusion, while the RheoSwitch system looked like a very promising inducible system for synaptase expression, it was difficult to establish the cell lines. Therefore, we went back to the original design using lentiviral vectors for high transduction efficiency.

Lentiviral Tet inducible system

Viral synaptase genes were cloned into the pSLIK-Zeo lentivirus vector for high transduction efficiency, selection and inducibility (Shin et al. 2006). The pSLIK-Zeo lentivirus expresses the rTA transactivator from a ubiquitin promoter (Figure 5. 15A). The transactivator gene is followed by an IRES that drives the ble gene to confer zeocin resistance selection. pSLIK-Zeo uses the Gateway system for insertion of genes of interest, which facilitated cloning of the synaptases. The synaptases were first cloned into the pENTR2B/TREPitt donor vector under the expression of a minimal Cauliflower mosaic virus promoter (CMV) controlled by a Tet ON operator (Pluta et al. 2005). The pENTR2B/TREPitt vector also includes the attL recombination sites for the Invitrogen LR Clonase II recombination reaction used to transfer the Tet operator, CMV minimal promoter and gene of interest into pSLIK-Zeo. The Tet operator induces transcription in the presence of doxycycline and the rTA transactivator. Since the rTA transactivator is expressed constitutively in cis, expression of the synaptases is only limited by the diffusion rate of doxycycline from the growth medium into the nucleus.
To follow expression of the synaptases, their genes were fused to a red fluorescent gene, E2-Crimson (Strack et al. 2009), through a P2A linker in a single open reading frame. The P2A linker causes ribosome skipping to produce equimolar amounts of the upstream and downstream protein products (Szymczak-Workman et al. 2012).

Unfortunately, the P2A peptide leaves a proline residue at the N-terminal end (Nt) of the C-terminal (Ct) protein and an 18 amino acid peptide at the Ct of the Nt protein (Figure 5. 15B). Previous reports have shown that these synaptases are moderately defective when fused to reporter genes (Taylor et al. 2003; Poteete 2011). Since we didn’t know if any of these additions might affect the recombination activity of the proteins, E2-Crimson was cloned either upstream or downstream of the synaptases in separate lentiviral constructs.

Six pSLIK/TREPitt lentiviral plasmid constructions were created:
pSLIK/TREPitt::Crimson/P2A-ICP8, pSLIK/TREPitt::Crimson/P2A-NLS/Ha-HumBeta,
pSLIK/TREPitt::Crimson/P2A, pSLIK/TREPitt::ICP8-P2A/Crimson,
pSLIK/TREPitt::NLS/Ha-HumBeta -P2A/Crimson and pSLIK/TREPitt::P2A/Crimson, also known as pSLIK1-6, respectively (Table 5. 2). The cloning details are described in Material and Methods. Sequences and maps of every plasmid produced in this study are in the Sequences and Maps Appendix. Lentivirus particles produced from these lentiviral vectors are proficient to transduce any human cell line. Transduced cells could be isolated by zeocin selection. Synaptase expression can be followed by E2-Crimson expression.

Do pSLIK clones express the synaptases in human cells?

ICP8 expression from pSLIK1 and pSLIK4 was detected with an anti-ICP8 antibody and NLS/Ha-HumBeta expression from pSLIK2 and pSLIK5 was detected with an anti-
HA antibody as shown via Western Blot (Figure 5.16). Two possible protein products were predicted for each: unprocessed full-length polyprotein with the Crimson fluorescent protein fused to the synaptases and processed synaptases expressed as separate polypeptides without Crimson. Only one protein product was observed for each construction and it seems to be the smaller products predicted by efficient P2A cleavage to yield unfused ICP8 and HumBeta proteins. HumBeta is a new protein specially created to be expressed in human cells. From these analyses HumBeta is very well expressed. In these blots samples expressing only the Crimson controls were also run; these were not detected by the antibodies (as predicted) so to simplify the figure they were not indicated.

**Figure 5.15 Synaptases expressed from Tet inducible lentiviral system.**

**A)** Diagram of Lentiviral vectors expressing synaptases from TREPitt Doxycycline-inducible promoter regulated by the rTA transactivator. Crimson is a red fluorescent reporter gene fused to the synaptases via P2A. Zeocon resistance provides a means to select clones that constitutively express rTA, allowing maximally efficient Doxycycline-induction. Shown is an example of Crimson fused to the N-terminus of a viral synaptase but reciprocal constructions fuse Crimson to the C-terminus of each synaptase. Control vectors are identical but lack the synaptase genes. Six vectors were constructed, Crimson/P2A/ICP8, Crimson/P2A/HumBeta, Crimson/P2A, ICP8/P2A/Crimson, HumBeta/P2A/Crimson and P2A/Crimson. **B)** Synaptase-P2A/reporter genes produce two proteins from a single transcript. Picornaviruses use 2A peptides to mediate “cleavage” between two proteins between the glycine and the proline residue. We used the porcine teschovirus-1 sequence (P2A) because it was the shortest and most efficient of the 2A variants.

**Table 5.2 Six lentiviral constructions were created**

<table>
<thead>
<tr>
<th>pSLIK #</th>
<th>Insert</th>
<th>Lentivirus size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Crimson/P2A-ICP8</td>
<td>11254</td>
</tr>
<tr>
<td>2</td>
<td>Crimson/P2A-NLS/HA-HumBeta</td>
<td>8768</td>
</tr>
<tr>
<td>3</td>
<td>Crimson/P2A control</td>
<td>7863</td>
</tr>
<tr>
<td>4</td>
<td>ICP8-P2A/Crimson</td>
<td>11447</td>
</tr>
<tr>
<td>5</td>
<td>NLS/HA-HumBeta-P2A/Crimson</td>
<td>8763</td>
</tr>
<tr>
<td>6</td>
<td>P2A/Crimson control</td>
<td>7866</td>
</tr>
</tbody>
</table>
Figure 5. 16 ICP8 and HumBeta expressed from pSLIK transgenes in 293T cells.

Western Blot analysis of pSLIKs lentiviral constructions. 293T cells were transfected with each pSLIK plasmid and cells were grown in the presence of 1000 ng/ml Doxycycline. Samples were collected and prepared as indicated in Material and Methods. A) ICP8 expression from pSLIK1 and pSLIK4. B) HumBeta expression from pSLIK2 and pSLIK5. The table indicates the masses of predicted and actually detected protein products in these Western Blots.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Possible proteins</th>
<th>Molecular Weight</th>
<th>Observed?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crimson/P2A-ICP8 (pSLIK1)</td>
<td>Crimson-P2A-ICP8 (fusion protein)</td>
<td>157 kDa</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>PH-ICP8</td>
<td>136 kDa</td>
<td>Yes</td>
</tr>
<tr>
<td>ICP8/P2A-Crimson (pSLIK4)</td>
<td>ICP8-P2A-Crimson (fusion protein)</td>
<td>156 kDa</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>ICP8-P2A</td>
<td>137 kDa</td>
<td>Yes</td>
</tr>
<tr>
<td>Crimson/P2A-NLS/HumBeta (pSLIK2)</td>
<td>Crimson/P2A-NLS/HumBeta (fusion protein)</td>
<td>65 kDa</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>PH-NLS/HumBeta</td>
<td>36 kDa</td>
<td>Yes</td>
</tr>
<tr>
<td>NLS/HumBeta/P2A-Crimson (pSLIK5)</td>
<td>NLS/HumBeta/P2A-Crimson (fusion protein)</td>
<td>65 kDa</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>NLS/HumBeta/P2A</td>
<td>38 kDa</td>
<td>Yes</td>
</tr>
</tbody>
</table>

What is the localization of the synaptases expressed from the pSLIK plasmids?

ICP8 maintained the nuclear and subnuclear localization characterized by speckled structures (Figure 5.17, top panel) as seen previously by de Bruyn Kops & Knipe (1988). Representative data for ICP8 expression from pSLIK1 (shown) has a similar distribution that ICP8 from pSLIK4 (not shown). The P2A linker sequence leaves an 18 amino acid peptide at the Ct of the first protein. ICP8 nuclear localization signals are localized at the Ct (Gao & Knipe 1992), therefore we were concerned that this peptide might affect the localization. Nevertheless, ICP8 expression and localization looked normal.
NLS/HA-HumBeta appeared to localize to the nucleus. The protein distribution looked more homogeneous in the nucleus than was the ICP8 nuclear distribution, *i.e.* not speckled (Figure 5.17, middle panel). Some of the NLS/HA-HumBeta protein seemed to localize to the periphery of the nucleus, especially when NLS/HA-HumBeta was expressed from pSLIK5 (data not shown). One possible cause could be that HumBeta over-expression saturates the nuclear pore complexes. Another possible cause could be that the negatively charged HA epitope peptide that follows the positively charged SV40 nuclear localization signal peptide inhibits uptake. However, the anti-HA antibody recognized the peptide when ICC was performed, suggesting that if there was an interaction between the HA and NLS peptides it is not so tight as to occlude the HA epitope from the antibody. A third possibility could be that HumBeta self-oligomerized in
cells into large complexes that were not readily transported through nuclear pore complexes. This possibility is consistent with observations that Beta self-assembles in vitro into ~12 subunit toroids of about 350 KD (Passy et al. 1999), but as similar phenomena exist for ICP8 (Tolun et al. 2013), this too is unlikely.

Crimson was detected by its intrinsic fluorescence (Figure 5. 17, in all 3 panels). It seemed to distribute throughout the cells with a brighter signal in the cytoplasm, as predicted.

Recombineering cell lines created by pSLIK lentivirus transduction

Recombineering cell lines expressing the Recombineering target Mostaza gene and an inducible synaptases were created by transducing the Mostaza-293T and Mostaza-293 cell lines with lentiviral particles derived from the six pSLIK plasmids. After transduction some cells were incubated with doxycycline to evaluate transduction efficiency via flow cytometry by following Crimson expression (Figure 5. 18A). It is noticeable that the cell lines with the lowest transduction efficiency were the ones expressing the human herpes ICP8 synaptase. The herpes ICP8 synaptase gene is ~4 Kb, leading to a lentiviral transducing particle of 11284 bp for pSLIK1 and 11447 bp for pSLIK4. In contrast, the NLS/HA-HumBeta synaptase is ~1 Kb leading to lentiviruses of 8768 bp for pSLIK2 and 8763 bp for pSLIK5. The reporter control vectors are even smaller and produce lentiviruses of 7863 bp for pSLIK3 and 7866 bp for pSLIK6 as represented in Table 5. 1. In this case, differences in packaging efficiency resulting from different sized lentiviral constructs might have had an effect on the transduction efficiency. Alternatively, toxicity from herpes synaptase expression might compromise
cell viability or growth. Figure 5. 18B shows that Crimson expression is mainly due to doxycycline induction.

293 cells were also transduced to evaluate if T antigen-activated replication from the SV40 origin upstream of the Mostaza target transgene created the strand specificity for Recombineering (as discussed below). Although the 293T cell line is a derivative from HEK 293, they also differ in MMR (Trojan et al. 2002; Cejka et al. 2003) which could complicate the evaluation of the T antigen role in Recombineering strand specificity.

**Human cell Recombineering**

Gene targeting was evaluated in 293T-Mostaza cells. The ability of viral synaptases to promote gene targeting was evaluated from transient transfection of ICP8 from the pCMV-ICP8 and by induction of ICP8 or HumBeta from the pSLIK stable cell lines.

![Diagram](image.png)

**Figure 5. 18** 293T-Mostaza-pSLIK and 293-Mostaza-pSLIK cells transduction efficiency and inducibility. Transduced 293T-Mostaza-pSLIK and 293-Mostaza-pSLIK cells were grown in the presence of 1000 ng/ml doxycycline. Data was analyzed using the Accuri flow cytometer on the FL4 channel. A) Transduction efficiency reflecting the percent of cells that express Crimson prior to zeocin selection. B) Doxycycline induces transgene expression from a low background level.
293T cell gene targeting

Gene targeting frequencies in 293T using modified oligos (e.g. with uncleavable phosphorothioate linkages, polyU or O6-methyl-protected ends, or synthesized as LNA (locked nucleic acid-protected ssODNs with a 2'-oxygen 4'-carbon methylene bridge in ribose)) peaks at 0.2% while gene targeting frequencies using oligos without modifications peaks at only 0.02% (Andrieu-Soler et al. 2005; Disterer et al. 2012; Nickerson & Colledge 2003; Olsen et al. 2009). We evaluated gene targeting in the 293T-Mostaza cell line and found that sequence-optimized unmodified oligos that target the Mostaza gene and encode a change in fluorescence from Mostaza (yellow) to green produced a significant change in the frequency of green fluorescent cells. The total amount of recombinants within the population carrying the Mostaza gene was about 0.30 and 0.27 % for 2 independent experiments (Figure 5. 19). When subtracting the no oligo background the frequencies were 0.22 and 0.20 %. These frequencies are similar to what reported by others when using end protected modified oligos and ~10-fold higher than the average response with unmodified oligos (Andrieu-Soler et al. 2005; Disterer et al. 2012; Nickerson & Colledge 2003; Olsen et al. 2009).

A parallel transformation experiment was performed with a fluorescent oligo to determine the percent of cells that take up the oligo using flow cytometry to estimate transfection efficiency (Figure 5. 8 and Table 5. 1). For the transfection conditions used in these experiments (4 nM oligo and 3% Lipofectamine 2000) fewer than 61 % of the cells were transfected with the oligos. If gene targeting efficiency is recalculated by assuming that only cells that express the Mostaza target gene and that take up the oligo substrate can experience genome editing by recombination, the gene targeting frequencies peak at about 0.49 and 0.44 % for each individual experiment. After
subtracting the “no oligo” background of green fluorescence, the frequencies were 0.42 and 0.37%. Therefore the results of using our optimized oligo-mediated gene targeting protocol is \textit{\textasciitilde20-fold higher gene targeting rates than what has been previously reported in the literature.}

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure5}
\caption{19 293T cells gene targeting.}
\end{figure}

293T-Mostaza cells were incubated with 4.13 nM of oligo 85 in Lipofectamine 2000 for 23 hours (experiment 1) and for 26 hours (experiment 2). Conversion of cells from Mostaza to green was evaluated 4 days and 7 days after oligo treatment respectively, using the Accuri flow cytometry. The percent of green cells was determined as the frequency of green cells over the total Mostaza cells in the mixture of cells without treatment. Data were analyzed using Prism 5 software. The oligo effect was considered extremely significant with $P < 0.0001$ (****) when compared to no oligo experiments by 2 way ANOVA.

We surmise that gene targeting in 293T cells is facilitated by endogenous cellular homologous recombination functions like Rad52, a synaptase protein like the viral synaptases we have discussed previously (Shinohara et al. 1998). Would viral synaptases further augment cellular gene targeting frequencies using our optimized oligo-mediated recombination protocol? This question is evaluated below.

\textbf{Human Recombineering using transient transfection}

293T cells transduced with pDual-Mostaza-derived lentiviral particles expressed the Mostaza fluorescent protein target gene in 93% of the transductants. Subsequently, these cells were sequentially transfected with pCMV-ICP8 and then with oligos that convert Mostaza to green. The synaptase was transfected first and allowed to be expressed before
the oligos were transfected into cells. For this experiment, 293T-Mostaza cells were seeded in 6 well plates and transfected with pCMV-ICP8 using Fugene 6 at a ratio 6:2 (µl Fugene 6 : µg pCMV-ICP8) in triplicates. The next day, cells from each well were passed to a 24 well plate. The following day, media were changed to 0.75 ml of DMEM, 5 % FBS and cells were transfected with 4 nM oligos using Lipofectamine 2000 in triplicates, as done in Andrieu-Soler, (2005). 3 hours later 0.75 ml of DMEM, 15 % FBS was added to each well and 3 hours later the transfection mixture was removed and fresh medium was added to the cells. The next day, cells from each well were expanded to another well of a 6 well plate. Recombinants were quantified using flow cytometry. This protocol is also described in Table 5.3.

**Table 5.3 Human Recombineering protocol by synaptase transient transfection.**
The frequencies of Recombineering are limited by the cells expressing the target (93 %), the pCMV-ICP8 transfection efficiency (~ 60 %) and the oligo transfection efficiency (~ 57 %). Cell expressing the target gene and transfected with both DNA are 32 % of the total.

<table>
<thead>
<tr>
<th>Experimental days</th>
<th>Recombineering Outgrowth days</th>
<th>Experimental steps</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>Seed 293T/Mostaza (93%) in 6 well plates</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>Transfect with pCMV-ICP8 using Fugene 6 at a ratio 6:2 (~60% efficiency)</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>Split each well into a 24 well plates</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>4 nM oligo transfection with Lipofectamine 2000 (~57%)</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>Media were changed and cells from each well were transferred to 6 well plates.</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>Cells were split and run by flow</td>
</tr>
</tbody>
</table>

Does ICP8 expressed from pCMV-ICP8 catalyze oligo-mediated gene conversion?

To answer this question gene conversion was evaluated in the presence and absence of the human viral synaptase ICP8 (Figure 5.20). The data shows that there was a modest increase in recombination in cells expressing ICP8 (up to 25-fold above background). The presence of ICP8 stimulated recombination even for the 35 nucleotide oligos, which do not significantly impact recombination in the absence of ICP8 (unpaired Students t test). This increase in recombination could be due to the ability of ICP8 to
catalyze strand annealing and/or by protecting the oligos from degradation by cellular exonucleases.

![Figure 5. 20 ICP8 stimulates recombination.](image)

293T cells expressing the Mostaza recombination reporter were sequentially transfected with pCMV-ICP8 and with oligos that convert Mostaza to green. Oligo size is shown on the X axis. Oligo-as indicates that oligo sequence correspond to the antisense strand of egfp and Oligo-s indicates that oligo sequence correspond to the sense strand of egfp. Recombinants were quantified using flow cytometry. The no oligo data for 293T or 293T-pCMV-ICP8 was subtracted to each data point. Data were analyzed and plotted using Prism 5 software.

Does oligo size influence gene targeting frequencies?

The effect of oligo size on Recombineering was evaluating with oligos 35, 45, 55 and 65 nucleotides long (Figure 5. 20). One way ANOVA found significant differences between oligos of different sizes with a steady trend toward more recombination with longer oligos in both the presence and absence of ICP8. Longer oligos might promote the homology search by providing both a larger sequence for target identification and by increasing binding free energy when paired with the target. Each ICP8 monomer binds about 10 nucleotides (Makhov et al. 2009), but ICP8 binding to ssDNA is cooperative (Gourves et al. 2000). Synaptases are also predicted to form more stable oligo/synaptase
filaments with bigger oligos. It is worth noticing that Recombineering in *E. coli* is more efficient with oligos ~60 nucleotides long (Sawitzke *et al.* 2011).

Is gene targeting strand-specific?

It was previously seen that the lagging strand template in DNA synthesis provided the best target for oligo-mediated recombination and that transcription was not relevant (Aarts & te Riele 2010a; Ellis *et al.* 2001). In this experiment, “green” oligos with sequences otherwise identical to the sense and the antisense strands of the Mostaza target gene were transfected in separate experiments. The data showed that the antisense oligos were better recombination substrates than were the sense oligos. Since the reporter target gene was likely integrated in random orientation in individual cells with respect to chromosomally initiated replication forks, we cannot account for this bias based on proximity and orientation with respect to host origins of replication (as we saw previously in *E. coli* and in plasmids). Interestingly, the lentiviral vector pDual-Mostaza contains an SV40 origin/promoter. If the SV40 origin/promoter initiated replication with the help of the T antigen expressed in 293T cells, then we expect the antisense oligos to produce more recombinants as these oligos would hybridize to the lagging strand template and give the observed results. While MMR is a potent inhibitor of recombination, the mismatch for both oligo groups produce bubbles of the same size which should be repaired at similar rates. Furthermore, if the unpaired sequences were detected by MMR, the presence of a C.C mismatch formed by pairing sense oligos with their targets would be expected to decrease MMR and increase gene targeting rates, therefore MMR does not explain preference for the antisense oligos. The simplest
interpretation of these data is that replication initiating at the SV40 origin is responsible for the oligo preference.

In conclusion, gene targeting in 293T-Mostaza cells demonstrated that among oligos of sizes ranging from 35 to 65 nucleotides, longer oligos produced more recombinants, that oligos with the sequence of the antisense strand produced more recombinants (probably due to lagging strand template bias), and that ICP8 expression stimulated recombination. If Recombineering were to be normalized to cell transfection efficiencies, the Recombineering efficiency became 7%, a 35-fold increase in gene targeting efficiency. These data are consistent with promotion of human Recombineering by ICP8. We next tried to optimize the Recombineering rates by optimizing plasmid transfections.

As ICP8 was introduced in these experiments by transient transfection, 293T transfection efficiency was determined by transfecting cells with plasmids encoding the ICP8 fused to the gfp gene and evaluating the percent of cells expressing the green fluorescent protein by flow cytometry. As a control, two other plasmids were used. pMAX-GFP was used as a positive control because it is easy to transfec since it is an small plasmid (3486 bp), it has a CMV promoter like pCMV-ICP8, the maxGFP protein has a fast maturation time and it is not toxic to cells. pDual-eGFP was used as a control for the effect of plasmid size on transfection since it was bigger in size than pCMV-ICP8-GFP. GFP was expressed from pMAX-GFP and pDual-eGFP in 100% of the cells for at least four days and then the number of cells expressing the GFP protein started to decrease. However, pCMV-ICP8-GFP transfection efficiency was only ~60% and the
percent of cells expressing the green fluorescent protein started to drop two days after transfection.

We also experimented with several ratios of pCMV-ICP8-GFP and transfection reagent, but none of them produced the high transfection efficiencies of the plasmid controls. The lower transfection efficiency of pCMV-ICP8-GFP could have been due to ICP8 toxicity to cells leading to death of transfected cells, consistent with the higher amount of “floater” cells after pCMV-ICP8 transient transfection. Other hypotheses include that ICP8 expression promoted plasmid concatemer formation as it does in HHV1 and that concatemers may segregate inefficiently during cell division. Another possibility was that the flow cytometer detects cells expressing cytoplasmic fluorescent proteins with higher sensitivity than it detects cells expressing nuclear fluorescent proteins, but we did not perform that control in this study (Figure 5. 21).

pCMV-ICP8 Recombineering optimization

In order to optimize recombination frequencies, pCMV-ICP8 was transfected at different ratios of Fugene 6 to plasmid, and cells were incubated with different amounts of oligos (Figure 5. 22). Fugene (µl) to pCMV-ICP8 (µg) ratios included 3:2, 6:2 and 6:4. Two antisense oligos were evaluated, one with the sequence to change Mostaza to Green fluorescence (# 85) and the other identical in the sequence to Mostaza as a “selfing” oligo (# 131). The data were collected on days 3-4 (column A) and on day 9 (column B).
Figure 5. 21 pCMV-ICP8-GFP transfection is lower than for other plasmids. 293T cells were transfected with the indicated plasmids using Fugene 6 transfection reagent. Transfection efficiency was measured using the Accuri flow cytometer FL1 channel. Data were analyzed and plotted using Prism 5 software. The top graph compares pmax-GFP and pDual-GFP plasmids to pCMV-ICP8-GFP transfection efficiency and GFP loss kinetics. The bottom two graphs titrate Fugene 6 and pCMV-ICP8-GFP ratios to optimize transfection efficiency.
By the time of the first data collection, the reading for Green cells was consistent for each oligo type; “Green” oligos produced more Green cells and oligos with the Mostaza sequence did not deviate from background. Cells transfected with pCMV-ICP8 outperformed the endogenous Mostaza to Green conversion as shown in the previous experiment. By day 9, the frequency of Green cells was reduced close to background levels.

![Graphs showing recombinant phenotype decrease over time](image)

**Figure 5. 22 Recombinant phenotype decreases over time.**
Mostaza to green conversion was evaluated using transient transfection of pCMV-ICP8 using Fugene 6 at ratios indicated in the figure as Fugene (µl): pCMV-ICP8 (µg). Oligos 85 (65 nt as Green) and 131 (65 nt as Mostaza) were mixed with 2.5 µl Lipofectamine 2000 per well. In column A) data was collected on days 3-4 and in column B) data was collected on day 9. Transfection efficiency of pCMV-ICP8-GFP measured in a parallel experiment just before adding the oligo was 3:2=20.6%, 6:2=31.5% and 6:4=26.4%
pCMV-ICP8 transfection efficiency was estimated by transfecting in parallel pCMV-ICP8-GFP. Just before adding the oligos the parallel pCMV-ICP8-GFP transfection efficiency for 3:2 was 20.6 %, for 6:2 was 31.5 % and for 6:4 was 26.4 %. This experiment indicated that the plasmid transfection efficiency for this experiment was really low, even lower than shown in Figure 5. 21. Perhaps the stage of the cells or unknown aspects of the protocol make this step irreproducible. If the transfection efficiency of pCMV-ICP8-GFP correlates with the transfection efficiency of pCMV-ICP8, then we could conclude that the highest transformation efficiency was reached with the ratio 6:2. Nevertheless, there was not a significant correlation between transformation efficiency and Mostaza to Green conversion. Oligo titration from 50 to 400 nM didn’t show a significant stimulation with oligo concentration, but gene conversion mostly peaks by 200 nM oligo.

The reduction of cells expressing the Green fluorescent protein could be explained by the segregation of recombinant chromosomes as previously shown for bacterial Recombineering in Figure 4. 5 and 4. 6. However, the decrease of cells expressing Green fluorescence was more dramatic in cells transfected with ICP8, which could indicate that ICP8 is toxic to cells and recombinants might arrest or die.

It was noticed that cells transfected with pCMV-ICP8 produced a background of Green fluorescence. The data for experiments not transfected with oligos in Figure 5. 22 were plotted in Figure 5. 23. In this figure it is easier to see that plasmid transfection produced a significant increase in the background Green fluorescence. We associated the Green background to changes in cell physiology that increases green autofluorescence.
ICP8 expression produces green cells.

Mostaza to green conversion was evaluated using transient transfection of pCMV-ICP8 and Fugene 6 at ratios indicated in the X axes as Fugene (µl): pCMV-ICP8 (µg). Data were collected on day 9. Transfection efficiency of pCMV-ICP8 measured with a parallel experiment with pCMV-ICP8-GFP was 3:2 = 20.6 %, 6:2 = 31.5 % and 6:4 = 26.4 %

We then questioned if ICP8 could produce pleiotropic effects and evaluated if Mostaza reporter expression was affected by ICP8 and oligos. Figure 5. 24 compared the frequency of cells that are not in the Mostaza or Green channel (Dark).

Cells were transiently transfected with pCMV-ICP8 using Fugene 6 at ratios indicated in the top of the figures as Fugene (µl): pCMV-ICP8 (µg). Data were collected on day 9. Transfection efficiency of pCMV-ICP8 measured with a parallel experiment with pCMV-ICP8-GFP was 3:2 = 20.6%, 6:2 = 31.5% and 6:4 = 26.4%. Dark cells represent the cells that are not expressing Green or Mostaza fluorescence.
As per transfection with pCMV-ICP8-GFP the ratio 6:2 produced the highest transformation efficiency and subsequently the ratio 6:4. In this analysis, pCMV-ICP8 transformation efficiency correlated with the increase in the Dark cells that lost the ability to express the reporter gene. It is noticeable that the fraction of Dark cells is not affected by the green or Mostaza (self) oligo or by the oligo concentration. This graphs shows that ICP8 might induce pleiotropic effects like loop out of the direct LTR repeats that flank the target Mostaza gene or increase rates of random mutagenesis.

In general, we found that transfection efficiencies were irreproducible. We performed several recombination experiments where transfection efficiency fluctuated from 5 to 80% and the higher the transfection efficiency, the higher the cell death rate. Optimization of the ICP8 transfection efficiency was counterproductive, since cells died after transfection. It remains unknown if the cell death was due to the activity of the ICP8 synaptase or due to an innate immune response to the plasmid, though we note that other plasmids did not have this problem, suggesting that ICP8 expression is toxic. We then asked the question “Is transient ICP8 expression detrimental to cells?” Propidium iodide was used to evaluate cell death.

Propidium Iodide cannot pass through intact cell membranes but can freely enter cells with compromised membranes, which are assumed to be dead. Inside cells, Propidium Iodide binds to dsDNA and dsRNA and fluoresces red (630 nm, max). The red fluorescence can be recorded using Accuri flow cytometry channels FL2 or FL3. Red fluorescence is then an indicator of cell death.

To identify if ICP8 induced cell death, cells were transfected with pCMV-ICP8-GFP using Fugene 6. Cells were incubated with transfection mix and collected at different
times. Cells were then incubated with Propidium Iodide and run in the Accuri flow cytometer. GFP positive cells were counted when the fluorescence intensity was > 20000 in the FL1 channel (530 +/- 15 nm). Propidium Iodide positive cells were counted when the fluorescence intensity was > 10000 in the FL3 channel (670 nm LP). We then asked the question “Are the ICP8-GFP expressing cells dying more than the non-transfected cells?” To answer this question, the Propidium Iodide positive cells were plotted as a function of the flow data recording time after transfection (Figure 5. 25).

![Figure 5. 25 pCMV-ICP8-GFP transfection induces cell death.](image)

Transfected cells were incubated with Propidium Iodide (PI) at the indicated time after transfection. Fluorescence was analyzed using the Accuri flow cytometry FL1 channel for ICP8-GFP and FL3 channel for Propidium Iodide. Data was plotted using Prism 5 software. This figure represents the fraction of ICP8-GFP fluorescent cells dying and the fraction of non-ICP8-GFP fluorescent cell dying.

The data indicate that ICP8-GFP expression highly correlated with Propidium Iodide staining, with a P value of 0.0057, indicating that cells transfected by pCMV-ICP8-GFP were dying. This observation could explain the low transfection efficiency and the fast decay kinetics of cells expressing the ICP8-GFP protein. However, it could not be determined from these data if cell death was induced by ICP8-GFP protein.
overexpression or by pCMV-ICP8-GFP DNA inducing an innate immune response. Therefore, we developed a stable inducible source of ICP8 to transiently dial in the amount of ICP8 expression and eliminate the introduction of possibly “toxic” DNA.

**Recombineering using stable cell lines**

293T-Mostaza cells were transduced with the Synaptase pSLIK clones and selected for carrying the transgene using zeocin (**Figure 5. 26**). We then tested the ability of these cell lines to promote Recombineering.

![Figure 5. 26 pSLIK transgene expression after zeocin selection.](image)

293T-Mostaza cells were transduced with the pSLIK plasmids. Two days later zeocin was added to the media. Media were changed every 3 days, cells were passed when colonies became confluent. 293T-Mostaza nontransduced cells were treated in parallel with zeocin for each drug concentration to confirm that selection was working. Zeocin selection was completed in about a week.

Recombineering experiments were performed by inducing synaptase expression with 1000 ng/ml doxycycline or as indicated in the particular experiment. 8-24 hours
after induction, oligo-Lipofectamine 2000 complexes were added to cells. Conversion of cells from Mostaza to green fluorescence was measured via flow cytometry over time. The following questions were asked:

Do synaptases expressed in 293T-Mostaza-pSLIK cells catalyze oligo-mediated gene conversion?

The effect of ICP8 and HumBeta synaptase expression on gene targeted was evaluated in 293T-Mostaza-pSLIK cells (Figure 5. 27). ICP8 was evaluated using pSLIK1 and pSLIK4 cells and HumBeta was evaluated using pSLIK5. In general, stable pSLIK cell lines were treated with doxycycline for induction of the synaptases followed by oligo transfection.

In the transient 293T-Mostaza pCMV-ICP8 transfection experiments, the antisense oligos consistently outperformed the sense oligos (Figure 5. 20). The effect of the oligo strand was re-evaluated in these experiments. Oligo substrates like the sense and the antisense strands of the target gene were transfected in separate experiments (Figure 5. 27). A no change (“selfing”) control was also included in experiments where the oligo had the same sequence as the Mostaza gene. Green and Mostaza fluorescence were interrogated using the Accuri flow cytometer using the gates on Figure 5. 6.
**Figure 5.** Gene targeting promoted by ICP8 and antisense oligos, but not by HumBeta.

A, B, D and E) 293T cells transfected with pSLIK1 or pSLIK4 were selected for zeocin resistance with 100 µg/ml zeocin for a week and seeded on a 24 well plate in the presence of 1000 ng/ml doxycycline. 8 hours later the media were changed and 4.13 nM oligo 85, 84 and 131 were independently transfected using Lipofectamine 2000 and Plus reagent. The oligo transfection mix was incubated with the cells for 17 hours. During this time, ~ 26 % of pSLIK1 transduced cells and 60 % of pSLIK4 transduced cells expressed Crimson upon doxycycline induction, indicating that only ~26 % of pSLIK1 cells were expressing ICP8, while ~ 60 % of pSLIK4 cells expressed ICP8.

C and F) 293T-Mostaza cells transfected with pSLIK5 were selected for zeocin resistance with 300 µg/ml zeocin for a week and seeded on a 24 well plate in the presence of 100 ng/ml doxycycline. 19 hours later, doxycycline was washed out and 4.13 nM oligos 85, 84 and 131 were independently transfected using Lipofectamine 2000 and Plus reagent. The oligo transfection mix was incubated with the cells for 12 hours. During this time ~ 97 % cells were induced by doxycycline for HumBeta/Crimson expression, indicating that most of the cells expressed the HumBeta protein. Antisense (as) and sense (s) Green oligos had the target sequence, differing by 4 nucleotides that produce the Mostaza to Green conversion, while the Mostaza oligo had the exact sequence - “self” - of the target gene, and should not produce color change. Mostaza to Green conversion was recorded using Accuri flow cytometry. pSLIK1 (A, D) data was collected on day 5; pSLIK4 (B, E) data was collected on day 4 and pSLIK5 (C, F) data was collected on day 4. Data was analyzed using Prism 5 software. The asterisks (*) represent the degree of significant difference of Mostaza to green conversion between samples treated with that oligo and samples treated with no oligo by unpaired t test, two tailed distribution. The capped line with the asterisks (*) represents the degree of significant difference of Mostaza to green conversion between endogenous and synaptase expressing cells by unpaired t test, two tailed distribution. A-C) data normalized to the number of Mostaza target. D-F) data also normalized to oligo transfection efficiency as estimated by the FL-oligo in **Figure 5.** B and to synaptase inducibility, as determined by Crimson expression.
A summary of Recombineering data follows, organized by lentiviral vector employed. In general the antisense oligos were better recombination substrates than sense oligos in the presence and absence of the synaptases, consistent with the transient transfection results and with the idea that replication arising from the SV40 origin of replication, which is located upstream of the Mostaza gene in the Recombineering reporter construct, is a primary source of replication forks encountered by the oligos. ICP8 stimulated gene conversion at rates that exceeded the endogenous recombination rate but humanized Beta protein did not.

**pSLIK1 gene targeting analysis (Table 5.4)**

Table 5.4 Gene targeting frequencies and statistics for pSLIK1.

Data were normalized for Mostaza target gene expressed in 91 % of the cells (second column), for oligo transfection efficiency in 61 % of the cells and for Crimson-P2A/ICP8 expression upon 1000 ng/ml doxycycline induction in 26 % of the cells (third column). Data represented in Figure 5.27A (normalized for Mostaza target gene expression) and Figure 5.27D (normalized for Mostaza target gene expression, oligo transfection and Crimson-P2A/ICP8 expression). Gene targeting frequencies represent the average of three replicates in one experiment (3). P values were evaluated using unpaired t test, two-tailed distribution. NS stands for no significant effect. The statistics reflect the P value summary.

<table>
<thead>
<tr>
<th>pSLIK1</th>
<th>Normalized to cells expressing target gene</th>
<th>Normalized to cells expressing target gene, transfected with oligo, and induced for synaptase</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Gene targeting frequencies (%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Oligo</td>
<td>Endogenous</td>
</tr>
<tr>
<td>No oligo</td>
<td>0.26 ± 0.06 (3)</td>
<td>0.22 ± 0.05 (3)</td>
</tr>
<tr>
<td>Antisense</td>
<td>0.32 ± 0.07 (3)</td>
<td>0.36 ± 0.04 (3)</td>
</tr>
<tr>
<td>Sense</td>
<td>0.21 ± 0.08 (3)</td>
<td>0.17 ± 0.01 (3)</td>
</tr>
<tr>
<td>Self</td>
<td>0.15 ± 0.03 (3)</td>
<td>0.22 ± 0.02 (3)</td>
</tr>
</tbody>
</table>

**Do oligos matter?**

P values when comparing oligo vs to no oligo

<table>
<thead>
<tr>
<th>Oligo</th>
<th>P values when comparing oligo vs to no oligo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antisense</td>
<td>0.2770 (NS)</td>
</tr>
<tr>
<td>Sense</td>
<td>0.3925 (NS)</td>
</tr>
<tr>
<td>Self</td>
<td>0.0439 (Inhibits *)</td>
</tr>
</tbody>
</table>

**Does ICP8 matter?**

P values when comparing doxycycline treatment (Endogenous + ICP8) vs no doxycycline (Endogenous)

<table>
<thead>
<tr>
<th>Oligo</th>
<th>P values when comparing doxycycline treatment (Endogenous + ICP8) vs no doxycycline (Endogenous)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No oligo</td>
<td>0.4316 (NS)</td>
</tr>
<tr>
<td>Antisense</td>
<td>0.5284 (NS)</td>
</tr>
<tr>
<td>Sense</td>
<td>0.4415 (NS)</td>
</tr>
<tr>
<td>Self</td>
<td>0.0255 (Stimulates *)</td>
</tr>
</tbody>
</table>
Do oligos stimulate Mostaza → Green conversion using host synaptases?

Endogenous gene conversion was stimulated by the antisense Green oligo, but the stimulation was not statistically significant when evaluated using unpaired t test, two-tailed distribution (P = 0.2770). However, when data were adjusted for oligo transformation efficiency (normalized), the effect of the antisense oligo was consistent with stimulation of endogenous gene conversion (P = 0.0229).

The sense Green oligo had no effect on endogenous gene conversion (P = 0.3925), even when the data were normalized for oligo transformation efficiency (P = 0.3786).

Endogenous gene conversion was inhibited by a nonmutagenic antisense Mostaza oligo control that had the same sequence as did the target gene (“self”, P = 0.0439). However, when the data were adjusted for oligo transformation efficiency, the self oligo showed no effect on endogenous gene conversion (P = 0.8090).

Do oligos stimulate Mostaza → Green conversion in the presence of ICP8?

ICP8 and the antisense oligo stimulated gene conversion and did so to levels exceeding that of the host cell (P = 0.0220). When the data were normalized for oligo transformation efficiency and ICP8 inducibility, the effect of the antisense oligo was consistent with a greater stimulation of gene conversion (P = 0.00169).

The sense Green oligo had no effect on ICP8-mediated gene conversion (P = 0.1816) even when data were adjusted for oligo transformation efficiency and ICP8 inducibility (P = 0.0169).

Gene conversion under the influence of ICP8 was unaffected by the self oligo (P = 1.0000) unless the data were adjusted for oligo transformation efficiency and ICP8 inducibility (P = 0.0169).
Does ICP8 expression stimulate Mostaza → Green conversion at rates that exceed endogenous functions?

ICP8 expression alone did not stimulate oligo-independent Mostaza → Green conversion (P = 0.4316) unless the data were adjusted for ICP8 inducibility (P = 0.0095).

While ICP8 and the antisense Green oligo stimulated gene conversion when compared to no-oligo, this effect was not significant when compared to the endogenous antisense oligo gene conversion (P = 0.5284) unless the data were adjusted for oligo transformation efficiency and ICP8 inducibility (P = 0.0004).

ICP8 and the sense Green oligo did not stimulate Mostaza → Green conversion (P = 0.4415) unless the data were adjusted for oligo transformation efficiency and ICP8 inducibility (P = 0.0007).

While ICP8 and the self oligo did not stimulated gene conversion when compared to no-oligo, endogenous functions decreased the rate of gene conversion in the presence of self oligo. When the null effect of ICP8 expression is compared to the inhibition by the endogenous functions, self oligo gene conversion is significantly different between these two experiments (P = 0.0255). When data were adjusted for oligo transformation efficiency and ICP8 inducibility, the effect of doxycycline induction was consistent with a greater stimulation of Mostaza → Green conversion (P < 0.0001).

In conclusion, ICP8 expressed from pSLIK1 modestly but consistently outperformed cellular functions for gene targeting rates.
**pSLIK4 gene targeting analysis (Table 5.5)**

Table 5.5 Gene targeting frequencies and statistics for pSLIK4.
Data were normalized for Mostaza target gene expressed in 91% of the cells (second column), for 61% oligo transfection efficiency and for 26% of the cells showing ICP8-P2A/Crimson expression upon 1000 ng/ml doxycycline induction (third column). Data represented in Figure 5.27B (normalized for Mostaza target gene expression) and Figure 5.27E (normalized for Mostaza target gene expression, oligo transfection and ICP8-P2A/Crimson expression). Gene targeting frequencies represent the average of three replicates in one experiment (3). P values were evaluated using unpaired t test, two-tailed distribution. NS stands for no significant effect. The statistics reflect the P value summary.

<table>
<thead>
<tr>
<th>pSLIK4</th>
<th>Normalized to cells expressing target gene</th>
<th>Normalized to cells expressing target gene, transfected with oligo, and induced for synaptase</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Gene targeting frequencies (%)</td>
<td></td>
</tr>
<tr>
<td>Oligo</td>
<td>Endogenous</td>
<td>Endogenous + ICP8</td>
</tr>
<tr>
<td>No oligo</td>
<td>0.24 ± 0.10 (3)</td>
<td>0.39 ± 0.049 (3)</td>
</tr>
<tr>
<td>Antisense</td>
<td>0.31 ± 0.038 (3)</td>
<td>0.53 ± 0.058 (3)</td>
</tr>
<tr>
<td>Sense</td>
<td>0.26 ± 0.034 (3)</td>
<td>0.40 ± 0.075 (3)</td>
</tr>
<tr>
<td>Self</td>
<td>0.26 ± 0.022 (3)</td>
<td>0.44 ± 0.14 (3)</td>
</tr>
</tbody>
</table>

Do oligos matter?
P values when comparing oligo vs no oligo

<table>
<thead>
<tr>
<th>Oligo</th>
<th>0.3424 (NS)</th>
<th>0.0360 (Stimulates *)</th>
<th>0.0193 (Stimulates *)</th>
<th>0.0016 (Stimulates **)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antisense</td>
<td>0.7246 (NS)</td>
<td>0.8842 (NS)</td>
<td>0.0484 (Stimulates *)</td>
<td>0.0258 (Stimulates *)</td>
</tr>
<tr>
<td>Sense</td>
<td>0.7361 (NS)</td>
<td>0.5808 (NS)</td>
<td>0.0410 (Stimulates *)</td>
<td>0.0656 (NS)</td>
</tr>
</tbody>
</table>

Does ICP8 matter?
P values when comparing doxycycline treatment (Endogenous + ICP8) vs no doxycycline (Endogenous)

<table>
<thead>
<tr>
<th>Oligo</th>
<th>0.0798 (NS)</th>
<th>0.0055 (Stimulates **)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antisense</td>
<td>0.0053 (Stimulates **)</td>
<td>0.0007 (Stimulates ***)</td>
</tr>
<tr>
<td>Sense</td>
<td>0.0427 (Stimulates *)</td>
<td>0.0055 (Stimulates **)</td>
</tr>
<tr>
<td>Self</td>
<td>0.0835 (NS)</td>
<td>0.0223 (Stimulates *)</td>
</tr>
</tbody>
</table>

**Do oligos stimulate Mostaza → Green conversion using host synaptases?**

Endogenous gene conversion was stimulated by the antisense Green oligo but not in a significant way (P = 0.3424), unless the data were adjusted for oligo transformation efficiency (P = 0.0193).

Endogenous gene conversion was not stimulated by the sense Green oligo (P = 0.7361), unless the data were adjusted for oligo transformation efficiency (P = 0.0410).

Endogenous gene conversion was not stimulated by the nonmutagenic antisense Mostaza oligo that had the same sequence as did the target gene (“self”, P = 0.7361), unless the data were adjusted for oligo transformation efficiency (P = 0.0410).
**Do oligos stimulate Mostaza → Green conversion in the presence of ICP8?**

ICP8 and the antisense oligo stimulated gene conversion in a significant way (P = 0.0360). When the data were adjusted for oligo transformation efficiency and ICP8 inducibility, the effect of the antisense oligo became consistent with a greater stimulation of gene conversion (P = 0.0016).

ICP8 and the sense Green oligo had no effect on gene conversion (P = 0.8842), unless the data were adjusted for oligo transformation efficiency and ICP8 inducibility (P = 0.0656).

ICP8 and the self oligo (P = 0.5808) didn’t significantly stimulate gene conversion, even when the data were adjusted for oligo transformation efficiency and ICP8 inducibility (P = 0.0656).

**Does ICP8 expression stimulate Mostaza → Green conversion at rates that exceed endogenous functions?**

ICP8 expression alone produced an increase in green fluorescent cells, but this effect was not statistically significant (P = 0.0798), unless the data were adjusted for ICP8 inducibility (P = 0.0055).

ICP8 and the antisense Green oligo stimulated Mostaza → Green conversion (P = 0.0053), although the increase in green cells due to ICP8 expression obscure this result. When the data were normalized for oligo transformation efficiency and ICP8 inducibility, the effect of the antisense oligo became consistent with a greater stimulation of gene conversion (P = 0.00107).

ICP8 and the sense Green oligo stimulated Mostaza → Green conversion (P = 0.0427), although the increase in green cells due to ICP8 expression obscure this result.
When the data were normalized for oligo transformation efficiency and ICP8 the effect of the sense oligo was consistent with a greater stimulation of gene conversion ($P = 0.0055$).

ICP8 and the self oligo seems to stimulate Mostaza $\rightarrow$ Green conversion over endogenous functions, but this effect was not significantly different ($P = 0.0835$). This effect might be an artifact of the green cell stimulation produced by ICP8 expression and not contributed by the self oligo control. When the data were adjusted for ICP8 inducibility the effect was significant ($P = 0.02235$).

In conclusion, ICP8 expressed from pSLIK4 consistently outperformed cellular functions for gene targeting rates. However, doxycycline induction of pSLIK4 also produced an oligo independent increase of the green fluorescent cells.

The effect of ICP8 on gene conversion was weaker than the optimized Recombineering rates we obtained in bacteria with Beta protein. It might be possible that the additional two amino acids on the N terminus of ICP8 expressed form pSLIK1 and the 21 amino acid portion of the P2 peptide on the C terminus of ICP8 expressed from pSLIK4 could interfere with ICP8 synaptase function.

In conclusion, expression of ICP8 stimulated gene targeting in the pSLIK1 and pSLIK4 cell lines, but this stimulation may have an oligo-independent component that is additive to the overall rate of green fluorescent cell production. This is a surprising conclusion and may reflect an unknown effect of ICP8 expression.
**pSLIK5 gene targeting analysis (Table 5. 6)**

Table 5. 6 Gene targeting frequencies and statistics for pSLIK5.
Data were normalized for Mostaza target gene expressed in 91 % of the cells (second column), for 61 % oligo transfection efficiency and for 26 % of the cells showing HumBeta-P2A/Crimson expression upon 1000 ng/ml doxycycline induction (third column). Data represented in represented in Figure 5. 27C (normalized for Mostaza target gene expression) and Figure 5. 27F (normalized for Mostaza target gene expression, oligo transfection and HumBeta-P2A/Crimson expression). Gene targeting frequencies represent the average of three replicates in one experiment (3). P values were evaluated using unpaired t test, two-tailed distribution. NS stands for no significant effect. The statistics reflect the P value summary.

<table>
<thead>
<tr>
<th>Oligo</th>
<th>Normalized to cells expressing target gene</th>
<th>Normalized to cells expressing target gene, transfected with oligo, and induced for synaptase</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Endogenous</td>
<td>Endogenous + HumBeta</td>
</tr>
<tr>
<td>No oligo</td>
<td>0.10 ± 0.031 (3)</td>
<td>0.11 ± 0.010 (3)</td>
</tr>
<tr>
<td>Antisense</td>
<td>0.26 ± 0.086 (3)</td>
<td>0.21 ± 0.017 (3)</td>
</tr>
<tr>
<td>Sense</td>
<td>0.14 ± 0.022 (3)</td>
<td>0.10 ± 0.027 (3)</td>
</tr>
<tr>
<td>Self</td>
<td>0.086 ± 0.022 (3)</td>
<td>0.22 ± 0.055 (3)</td>
</tr>
</tbody>
</table>

Do oligos matter?
P values when comparing oligo vs to no oligo

<table>
<thead>
<tr>
<th>Oligo</th>
<th>P values when comparing oligo vs to no oligo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antisense</td>
<td>0.0435 (Stimulates *)</td>
</tr>
<tr>
<td>Sense</td>
<td>0.1468 (NS)</td>
</tr>
<tr>
<td>Self</td>
<td>0.4746 (NS)</td>
</tr>
</tbody>
</table>

Does HumBeta matter?
P values when comparing doxycycline treatment (Endogenous + HumBeta) vs no doxycycline (Endogenous)

<table>
<thead>
<tr>
<th>Oligo</th>
<th>P values when comparing doxycycline treatment (Endogenous + HumBeta) vs no doxycycline (Endogenous)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No oligo</td>
<td>0.5950 (NS)</td>
</tr>
<tr>
<td>Antisense</td>
<td>0.3653 (NS)</td>
</tr>
<tr>
<td>Sense</td>
<td>0.1247 (NS)</td>
</tr>
<tr>
<td>Self</td>
<td>0.0166 (Stimulates *)</td>
</tr>
</tbody>
</table>

Do oligos stimulate Mostaza → Green conversion using host recombinases?

Endogenous gene conversion was stimulated by the antisense Green oligo (P = 0.0435) and support for this conclusion was strengthened when the data were adjusted for oligo transformation efficiency (P = 0.0186).

Endogenous gene conversion was stimulated by the antisense oligo, but not in a statistically significant way (P = 0.1468), unless the data were adjusted for oligo transformation efficiency (P = 0.0092).
Endogenous gene conversion was not stimulated by the nonmutagenic antisense Mostaza oligo that had the same sequence as did the target gene (“self”, $P = 0.47461$), even when the data were adjusted for oligo transformation efficiency ($P = 0.2444$).

*Do oligos stimulate Mostaza $\rightarrow$ Green conversion in the presence of HumBeta?*

HumBeta and the antisense Green oligo stimulated gene conversion ($P=0.0014$), although at lower rates than in the absence of HumBeta. When the data were adjusted for oligo transformation efficiency and HumBeta inducibility, the effect of the antisense oligo became even more strongly supportive of the view that the antisense Green oligo stimulated gene conversion ($P = 0.0002$).

HumBeta and the sense Green oligo did not stimulate gene conversion ($P = 0.5704$), unless the data were adjusted for oligo transformation efficiency and HumBeta inducibility ($P = 0.0092$).

HumBeta and the self Mostaza oligo stimulated gene conversion ($P = 0.0297$). When the data were adjusted for oligo transformation efficiency and HumBeta inducibility, the self oligo data showed greater support for the conclusion ($P = 0.0092$).

*Does HumBeta expression stimulate Mostaza $\rightarrow$ Green conversion at rates that exceed endogenous functions?*

Unlike ICP8 expression, HumBeta by itself did not stimulate oligo-independent green fluorescent cell production ($P = 0.5950$), even when the data were adjusted for HumBeta inducibility ($P = 0.4844$).

HumBeta and the antisense Green oligo stimulate Mostaza $\rightarrow$ Green conversion at lower rates than the endogenous functions, as if HumBeta inhibits gene targeting.
Although the inhibition was not significantly different ($P = 0.3653$), even when the data were adjusted for oligo transformation efficiency and HumBeta inducibility ($P = 0.4202$).

Likewise, HumBeta and the sense Green oligo did not stimulate Mostaza $\rightarrow$ Green conversion ($P = 0.3653$), even when the data were adjusted for oligo transformation efficiency and HumBeta inducibility ($P = 0.1539$).

However, HumBeta and the nonmutagenic self oligo stimulated Mostaza $\rightarrow$ Green conversion ($P = 0.0166$) and this conclusion was further strengthened when the data were adjusted for HumBeta inducibility ($P < 0.0156$). How HumBeta promoted Mostaza $\rightarrow$ Green conversion with the self oligo escapes our understanding.

In conclusion, expression of HumBeta decreased the rates of gene conversion catalyzed by endogenous cellular functions, but was found to have no significant influence. HumBeta expression produced the same rate of Mostaza to green conversion when using antisense and self oligo, as if the effect of gene conversion is independent of the oligo sequence. *The simplest interpretation of these data is that HumBeta is not active as a Synaptase in human cells, but could inhibit the host synaptases in doing gene targeting.*

As with the ICP8 constructions, it might be possible that the additional 21 amino acid portion of the P2 peptide on the C terminus of HumBeta interferes with HumBeta synaptase function. HumBeta also has an HA tag and an NLS at the N terminus that conceivably could compromise synaptase function (though similar N-terminal tags have had no effect on Beta function *in vitro* and in bacteria; Tolun and Myers unpublished). It should also be remembered that the localization of HumBeta to the nucleus showed diffuse, nonorganized nuclear pattern and occasional accumulation at the nuclear
periphery (Figure 5.17), unlike the punctate organized nuclear patters of ICP8, suggesting that HumBeta function might be compromised.

Do changes in the fluorescence phenotype of putative recombinant cells correspond with the expected change in genotype?

Until now, gene conversion was measured by interrogating cells by their fluorescent phenotypic properties by flow cytometry. It was necessary to validate if the green cells were indeed recombinants and if the change in phenotype corresponded to the expected sequence introduced by oligo-mediated genome editing.

Since gene targeting appeared only in a small fraction of the cells in the experiments, recombinant cells appeared at a low frequency within the culture (Figure 5.28 left). It was necessary to enrich for the recombinant population for genotyping studies. To enrich for recombinants, cells from recombination experiment 82 (pSLIK1+Dox+oligo85) were sorted for Green fluorescence. During sorting, two populations were isolated, one expressing both green and Mostaza fluorescence and another cell population expressing only Green fluorescence. The sorted cells where both Mostaza and Green fluorescent proteins were expressed, proliferated in culture while the sorted cells that only expressed green fluorescence did not proliferate. The sorted experiment expressing both fluorescent proteins was expanded and re-evaluated using the Accuri flow cytometer. The sorting experiment enriched the Green population from 0.1 % to 39 % (Figure 5.28 right).
Figure 5. 28 Enrichment for recombinants by cell sorting.

Cells from experiment 82 (pSLIK1+doxycycline+oligo85) were expanded for about 3 weeks. Cells in the culture were enriched for green cells using BDFACS Aria IIu sorter and band pass filters 510/20 for green fluorescence and 550/30 for Mostaza fluorescence. The figure shows cells from this experiment before and after sorting evaluated by the Accuri flow cytometer. FL1-A measures green fluorescence intensity, FL2-A measures Mostaza fluorescence intensity. R1 gate corresponds to dark cells, P3 gate corresponds to Mostaza cells and P2 gate corresponds to Green cells. These gates were established with 293T-Mostaza, 293T-Green and 293T cells.

The genome of the sorted cells was extracted and the fluorescent protein target gene was amplified by PCR and sent for sequencing. The sequencing result showed that the Mostaza fluorescent allele sequence was well represented and that while the Green allele was also represented, it appeared in the sequencing trace close to the background levels (Figure 5. 29A).

The presence of the green allele was also evaluated using allele-specific PCR. Genomic DNA directly isolated from sorted cells and controls was amplified by PCR with one set of primers to amplify Green only and another set of primers to amplify Mostaza only. A common forward primer was used for both the Mostaza and Green-specific PCR reactions while the reverse primers differ in sequence at their 3' ends, reflecting the different complementary sequences to the two different allele templates. The allele specific PCR showed that a control cell line that expressed Mostaza produced a PCR
product with the Mostaza primers only, that a control cell line that expressed Green fluorescence produced a product with the Green primers only, that a cell line that was transfected with both Green and Mostaza targets produced products with both primers, and that a cell line that does not have any of the fluorescent transgene does not produce any amplified product. The experiment 82 sorted cell population produced products with both primer pair types.

**Figure 5.** 29 Recombinant green genotype detected by allele specific PCR and sequencing.
A recombinant cell mixture was sorted for cells with green fluorescence. Genomic DNA was isolated, the target gene was amplified by PCR and sequenced. **A)** Sequence of the Recombineering target region from experiment 82 shows peaks for Mostaza and also for oligo-introduced Green sequences. **B)** Allele-specific PCR shows evidence of genuine green recombinant allele sequences among nonrecombinant Mostaza genes in a mixed population. Genomic DNA isolated from sorted experiment 82 cells was amplified using primers that hybridize specifically to each allele.

*The sequencing and allele specific PCR confirmed that experiment 82 had both Mostaza and Green recombinant genotypes.* From these studies we can conclude that
oligo-mediated gene targeting is changing the genomic DNA at the target region from Mostaza to green as it is specified in the oligo sequence. This strongly suggests that at least some of the Green variants are true recombinants.

**Discussion**

Is the green cell (putative recombinants) phenotype stable?

When analyzing the recombination experiments over time, the percent of cells expressing Green fluorescence decreased over time (Figure 5.30). This decrease between recombinants has been observed in other studies (Aarts & te Riele 2011; Olsen et al. 2005a) and has been related by those investigators to cell cycle arrest due to activation of DNA damage response. In this study, the Green population seemed to decrease to background levels in the course of two weeks. Interestingly, the Green allele was detectable by allele specific PCR and sequencing of the Recombineered Mostaza gene ~3 weeks after the recombination experiment (Figure 5.29). The Green allele was detectable in sorted cells that fluoresce both Mostaza and Green but when sorted to isolate the cells expressing Green fluorescence only, Green cells didn’t proliferate. **This observation is consistent with green recombinant cells being in an arrested state** as observed by Aarts and te Riele (2011) and by Olsen et al. (2005a).

In *E. coli* Recombineering, after the substrate oligo annealed to its complementary target, there was segregation of recombinant from non-recombinant DNA strands which decreased the % of Green cells to at least 25 % of the original frequency of recombinants, assuming only one inserted target gene per genome, (Figure 4.5 and 4.6). Here we evaluated the recombinant population within the total treated cells over an extended
period of time and saw a decline of the recombinant frequencies as reported by us in bacteria (Valledor et al. 2012) and by others in mammals (Aarts & te Riele 2010b).

**Figure 5.30 Recombinant green cells population decreases over time.**
For the Recombination Exp3, 293T cells transfected with pSLIK4 and selected for zeocin resistance with 100 µg/ml zeocin for a week, were seeded in a 24 well plate in the presence 1000 ng/ml doxycycline and incubated overnight. The next day, 4.13 nM oligo 85 in Lipofectamine 2000 was added to cells (leaving doxycycline in the medium) and incubated for ~ 24 hours. During this time, ~ 87 % of pSLIK4 transduced cells expressed Crimson upon doxycycline induction. For the Recombination Exp2, 293T cells transfected with pSLIK1 or pSLIK4 and selected for zeocin resistance with 100 µg/ml zeocin for a week, were seeded in a 24 well plate in the presence of 1000 ng/ml doxycycline. 8 hours later, the medium was changed and 4.13 nM oligo 85 in transfection reagent was added to cells and incubated for 17 hours. During this time, ~ 26 % of pSLIK1 transduced cells and 60 % of pSLIK4 transduced cells expressed Crimson upon doxycycline induction. Conversion of cells from Mostaza to Green used transfection reagents Lipofectamine 2000 and plus reagent (Exp2/1), Lipofectamine 2000 (Exp2/2). Conversion of cells from Mostaza to Green was evaluated using the Accuri flow cytometer and analyzed with Prism 5 software.
Six Recombineering experiments monitored for 2 weeks produced an average decline of recombinants to 35% of the data recorded four or five days after the oligo treatment (Figure 5.31). Theoretically, DNA segregation should be completed within 2 replication cycles. If 293T cells replicated every 18 hours, it is expected that within 36 hours the strand segregation should be completed, before the first data recorded time point. Instead, the decline of green 293T recombinants continued over at least two weeks.

If *E. coli* divided every 30 minutes we could expect that strand segregation would be done in one hour. Instead, the observation was that it took around four hours to segregate...
all the recombinants from non-recombinants. If the same timing is applied to mammalian cells, four hours of *E. coli* post-Recombineering growth could be related to an ~ 6 days of 293T cell growth. Instead, the decay is observed for at least two weeks.

In *E. coli* Recombineering there is a cell cycle arrest for about one hour induced by some of the functions of the Lambda lysogen like Kil. It is also known that exponentially replicating *E. coli* fires a new replication fork before finishing the previous round of DNA synthesis. Multiple replication forks within a genome provide more possibility for the target sequence to be exposed and for the oligos to be delivered to the target. As the number of genomic DNA targets increase per cell, complete genomes will segregate into new cells over a longer time (one doubling time per target equivalent).

For 293T-Mostaza cells, the target gene is expected to be randomly integrated in a different genome location in each transduced cell, and the copy number per transduced cell is currently unknown. The integrated lentiviral sequence includes an SV40 origin of replication upstream of the Mostaza gene. Interestingly, the T antigen promotes the G1 to S transition and fires the SV40 origin at least twice during each S phase, promoting cells to form a tetraploid state (Wu 2004; Friedrich *et al.* 1992; Perry & Lehman 1998). Therefore, the genomic content of 293T may look like the *E. coli* genome in rapidly growing cells where a second round of genome synthesis is started before the first is done. We could expect then that segregation of recombinants could take at least 3 cycles, which should take (without factoring in any cell cycle arrest), ~ 54 hours. As mentioned previously, the decay is seen for about 2 weeks, therefore chromosomal segregation cannot completely explain the decrease of the frequency of recombinants over time.
It is possible that in addition to gene targeting ICP8 also promotes recombination between the repeated LTR sequences flanking the Mostaza target transgene. If this event happens after the gene targeting event, then the population of cells with the Green phenotype would vanish over time and newly formed nonfluorescent (dark) cells would arise in the population. Indeed, there was an ICP8-dependent increase in the dark cell population (see below) but this population displayed different kinetics than did the Green cells. The increase in dark cells stabilized by five days post-Recombineering while the green population kept decreasing after day five, suggesting that something else was happening, such as the cell arrest seen after sorting, or a decrease in GFP due to protein ruover.

In summary several hypotheses have been postulated for the decrease in the Green population:

- A cell cycle arrest (confirmed by sorting and impaired proliferation of recombinants).

- Chromosomal segregation of the recombinant strand from more than one duplicating chromosome.

- Target gene inactivation by gene silencing, mutagenesis or LTR recombination (confirmed by an increase in the dark population up to 10% of the total number of cells).

Does doxycycline-dependent Crimson expression produce green cells?

The increased frequency of Green fluorescent cells promoted by doxycycline treatment of pSLIK4 cells in the absence of oligos was surprising. An experiment was carried out to evaluate if the doxycycline-induced Green cells were just an artifact caused by expression of the far-red fluorescent protein Crimson. The frequency of conversion of
Mostaza cells to Green was evaluated for pSLIK4 (ICP8-P2A/Crimson) and pSLIK6 (P2A/Crimson) cells as a function of doxycycline concentration (Figure 5.32). For this experiment, 293T cells transfected with pSLIK4 or pSLIK6 were selected for zeocin resistance with 100 µg/ml zeocin for a week and seeded on a 24 well plate in the presence of 50, 100 and 1000 ng/ml doxycycline and incubated overnight. The next day, 4.13 nM oligo 85 (65 nt, antisense) mixed with Lipofectamine 2000 was added to cells (leaving doxycycline in the media) and incubated for ~ 24 hours. During this experiment, Crimson expression was induced for ~ 87 % of pSLIK4 and 98 % of pSLIK6 transduced cells. Conversion of cells from Mostaza to Green was evaluated using the Accuri flow cytometry and analyzed with Prism 5 software.

![Figure 5.32 ICP8 promotes green fluorescent cells.](image)

293T cells transfected with pSLIK4 or pSLIK6 were selected for zeocin resistance with 100 µg/ml zeocin for a week, seeded on a 24 well plate in the presence of 50, 100 and 1000 ng/ml doxycycline and incubated overnight. The next day, 4.13 nM oligo 85 (65 nt, antisense) mixed with Lipofectamine 2000 was added to cells (leaving doxycycline in the media) and incubated for ~ 24 hours. During this time, ~ 87 % of pSLIK4 transduced cells and 98 % of pSLIK6 transduced cells were induced by doxycycline for Crimson expression. Conversion of cells from Mostaza to Green was evaluated using the Accuri flow cytometry and analyzed with Prism 5 software. The asterisks reflect the difference between oligo and no oligo samples evaluated by one tailed t test. A) pSLIK4, data collected on day 3 B) pSLIK6, data collected on day 4.

Gene conversion was significant for both pSLIK4 (P = 0.0002) and pSLIK6 (P < 0.0001) when comparing the effect of oligo in gene targeting in the absence of doxycycline.
by unpaired t test, indicating that both cell lines had endogenous host functions that
catalyze gene conversion. In previous experiments addressing this same question,
endogenous gene conversion rates was stimulated by oligos, but variation in the data and
high oligo independent green background eliminated significance for results with pSLIK1
and pSLIK4 (Figure 5. 27).

Doxycycline promoted production of Green fluorescent cells in the absence of oligo
in pSLIK4 cells but not pSLIK6 cells. 2 way ANOVA showed that both oligo and
doxycycline treatments produce significant increases in the frequency of green pSLIK4 cells
(P < 0.0001 for both oligo and doxycycline). In contrast, for pSLIK6 cells, only the oligo
treatment produced a significant increase in Green fluorescent cells (P < 0.0001) while
doxycycline treatment was not significant (P 0.6301). Since doxycycline treatment did not
increase Green cell frequencies in pSLIK6 cells, we discard the possibility that Crimson
fluorescence or doxycycline itself could artfactually affect the scoring of green cells.
Therefore, the oligo independent, doxycycline-induced Green phenotype in pSLIK4 cells
is due to the expression of ICP8.

What is the origin of ICP8-dependent oligo-independent green cells?

ICP8-P2A/Crimson could promote formation of cells with the Green phenotype by
genetic changes or by physiological changes. For example, dying cells can exhibit green
autofluorescence (common knowledge of cell workers). If the oligo-independent
population of Green cells were due to death, the green fluorescent population of cells
would be predicted to vanish over time.

In bacteria, physiological changes that lead to increased cell size can have high
backgrounds of fluorescence that show up in the FITC channel of flow cytometers
(Renggli et al. 2013), although this has not been reported for human cell lines. If this were the case in our studies, we predict that the frequency of green cells would be higher in larger cells.

Alternatively, ICP8 might induce mutations, including ones that convert Mostaza to Green. As mutations are random, we predict that mutation should increase the frequency of nonfluorescent cells at a higher rate than mutation to form green cells because loss of function mutations are much more common than are specific gain of function mutations.

Could ICP8-induced cell death contribute to false positive Green cells that distort the data? We observed that cells detach, clump and round up after doxycycline induction, consistent with cells dying (or at least suffering). Furthermore, in the transient pCMV-ICP8 transfection experiments we observed increased rates of cell death (inferred from propidium iodide staining). However, the effect of doxycycline was consistent across all the pSLIK lines, including the non-synaptase controls and the death during transient transfection might well be due to induced innate immune responses as previously discussed.

On the other hand, it is evident in Figure 5.30 that the oligo-independent population of Green cells vanished over time, although at approximately the same rate as did the oligo-treated samples. If Green cells were the product of arrested cells in the cell cycle before mitosis, the kinetics of green fluorescent cell loss might be similar to that of arrested *bono fide* green fluorescent recombinant cell loss. However, in an analysis of cell cycle disturbances, no significant differences were detected between doxycycline-induced and non-induced cells (Figure 5.33). If Green cells were the product of enlarged cells arrested in the cell cycle before mitosis as reported in bacteria (Renggli et al. 2013) we
should see an enrichment in Green cells among those with greater cell scatter. For this purpose, data from recombination experiment in Figure 5. 32 were reanalyzed to examine correlations between cell size (SSC vs FSC) and frequency of Green cells; no such correlation was detected so this proposal was rejected (not shown).

![Figure 5. 33 Cell cycle analysis of cells expressing ICP8.](image)

Cells were incubated with doxycycline for 24 hours and then fixed with ethanol and stained with propidium iodide. DNA content was analyzed by flow cytometry as indicated in Figure 2. 1.

In conclusion, there was evidence that ICP8 and doxycycline were toxic although doxycycline did not induce autofluorescence as seen in pSLIK6. It is possible that ICP8 expression induced autofluorescence. The Green cells vanished over time, but there is no evidence of cell cycle arrest when the whole population was analyzed. This experiment needs to be refined, where the cell cycle is compared between Green and Mostaza cells.

Finally there is the possibility of ICP8-dependent mutagenesis. We saw that ICP8 expression increased the frequency of cells that do not fluoresce as Mostaza or Green
The frequency of dark cells upon ICP8 expression was about 10% while the frequency of Green cells was on the order of 0.1% to 0.3%. Therefore the rate of conversion of fluorescent to dark cells was almost 100 times greater than the conversion rate of Mostaza to Green cells. This observation is consistent with the proposal that ICP8 induced random mutations, with a higher frequency of nonfluorescent cells produced by loss of function mutations and a lower frequency of Green cells produced by specific gain of function missense mutations. This hypothesis could be validated by sequencing the target region of isolated dark and Green cell clones.

In conclusion, the origin of ICP8-dependent oligo-independent green cells is still unknown, but could be related to autofluorescence of dying cells or due to ICP8 off-target events.

How specific is ICP8 activity in human cells?

While there was an increase in a genuine Green recombinant population upon doxycycline and oligo treatment, there was also an ICP8-dependent but oligo-independent increase in Green cells, an increase in a nonfluorescent “Dark” population and a concomitant decrease in the Mostaza population. Furthermore, the population of Green cells vanished over time as explained above, which could indicate that many of these cells are dying and produce green autofluorescence (Bartova et al. 2008; Renggli et al. 2013). The dark cells were stimulated by the expression of ICP8 and their frequency increased with increasing doxycycline concentration (Figure 5.34). The Dark population increased over time and then levels off by four to five days after doxycycline induction, perhaps mirroring when ICP8 levels declined following removal of doxycycline from the media.
What then is the origin of the Dark variants? In bacteria, it is difficult to propagate lentiviral vectors in the presence of homologous recombination functions such as Recombineering enzymes due to deletion of the virus through recombination between the directly repeated LTR sequences (personal experience and Lindley, Lu, Geffin and Myers, in preparation). Perhaps as in bacteria, ICP8 could mediate recombination between the directly repeated LTR sequences of lentiviruses in human cells, thereby looping out (deleting) the fluorescent Mostaza recombination reporter gene and/or the Crimson gene in the pSLIK integrants. In another scenario, the SV40 origin of replication is just upstream of the Mostaza gene and could be a hot spot for dsDNA breaks near the Mostaza gene. ICP8 could compete with NHEJ during non-S phase repair of DNA breaks and anneal regions of micro-homology to create deletions that inactivate protein fluorescence. It would be interesting to sort the dark population and see if it is possible to amplify the Mostaza gene, capture those sequences and sequence the clones. In a third scenario, ICP8 could affect the integrity of cellular genomes. 293T cells are MMR deficient (Cannavo et al. 2005), which is known to permit recombination between partially homologous sequences. Maybe ICP8 pairs regions of partial homology in 293T causing genome rearrangements that negatively affect Mostaza expression. Alternatively, ICP8 might induce random mutations including ones that modify Mostaza. Finally, ICP8-mediated pairing could promote gene silencing similarly to what has been observed in fungi for inactivation of transposons (e.g. Cambareri et al. 1991; Barry et al. 1993; Galagan & Selker 2004).
293T cells transfected with pSLIK4 or pSLIK6 were selected for zeocin resistance with 100 µg/ml zeocin for a week, seeded on a 24 well plate in the presence of 50, 100 and 1000 ng/ml doxycycline and incubated overnight. The next day, 4.13 nM oligo 85 (65 nt, antisense) mixed with Lipofectamine 2000 was added to cells (leaving doxycycline in the media) and incubated for ~ 24 hours. During this time, ~ 87 % of pSLIK4 transduced cells and 98 % of pSLIK6 transduced cells were induced by doxycycline for Crimson expression. Dark cells were the percent of cells that do not expresses the target gene Mostaza or Green, evaluated using the Accuri flow cytometry with channels FL1 and FL2. Data were analyzed with Prism 5 software. Curves represent a non-linear fit of a one phase association function by the method of least squares.

In conclusion, there was an increase in both Green and Dark cell populations that could be associated with cell death and ICP8 off target effects. It is logical that this protein could produce such effects and that was one of the reasons to create an inducible system to minimize the unwanted effects of reprograming human DNA metabolism.

What are the frequencies of gene targeting?

The frequency of gene targeting is on the order of 0.1 to 0.2 % when the “no oligo” background of green fluorescence is subtracted from the experiments (Table 5.7). When comparing ICP8 to host gene targeting, there was a modest increase in recombination frequency. This increase in frequency was not always statistically significant and often
the no oligo background level was higher than for gene targeting by the endogenous host factors. More experiments need to be done to compare cellular recombinases to ICP8 for gene targeting to arrive to stronger conclusions. When comparing HumBeta to host gene targeting, there was a modest decrease in recombination frequency. This decrease in frequency was not statistically significant, therefore we cannot conclude that Beta inhibits gene targeting, but we can conclude that Beta has no beneficial effect on human Recombineering.

Table 5.7 Gene targeting efficiencies after subtracting oligo-independent events.

<table>
<thead>
<tr>
<th>Synaptase</th>
<th>Cell line</th>
<th>Rec Frequencies (%)</th>
<th>Rec Frequencies without no-oligo background (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cellular (Rad52?)</td>
<td>pSLIK6</td>
<td>0.30 ± 0.02 (3)</td>
<td>0.22</td>
</tr>
<tr>
<td>ICP8</td>
<td>pCMV-ICP8 transient</td>
<td>0.23 ± 0.04 (3)</td>
<td>0.22</td>
</tr>
<tr>
<td></td>
<td>pSLIK1</td>
<td>0.36 ± 0.04 (3)</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>pSLIK4</td>
<td>0.53 ± 0.06 (3)</td>
<td>0.14</td>
</tr>
<tr>
<td>HumBeta</td>
<td>pSLIK5</td>
<td>0.19 ± 0.03 (3)</td>
<td>0.08</td>
</tr>
</tbody>
</table>

In human cells, in addition to the *E. coli* RecA orthologs in the Rad51 family, there is a synaptase called Rad52, which could be responsible for much of the previously reported oligo-dependent gene targeting. In a recent study, Rad52 depletion decreased the rates of DNA break repair by homologous recombination (Schumacher *et al.* 2012), consistent with this view. These investigators also examined the role of ICP8 expressed from Herpes virus in infected cells and found that ICP8 failed to make a difference in cells expressing Rad52. Interestingly, ICP8 was also unable to promote DNA break repair in the absence of the activity of its associated SynExo nuclease, UL12. However in Rad52 depleted cells, UL12 plus ICP8 restored DNA break repair by recombination and did so using a ssDNA annealing mechanism similar to the model proposed in this thesis for
Rcombineering. While the substrates in this study are different from what were employed in this thesis (dsDNA in the paper, ssDNA here), and whereas it is well-known that DNA break repair by SSA requires 5’ → 3’ dsDNA resection to expose complementary ssDNA for pairing, it might also be true that UL12 is necessary for ICP8 to catalyze Rcombineering. This remains to be tested.

ICP8 has been previously shown to interact with the human ssDNA binding protein RPA (Taylor & Knipe 2004a), possibly coordinating ICP8 action with host cell replication. Rad52 is also an RPA interacting protein (Plate et al. 2008; Sugiyama & Kantake 2009; Wu et al. 2008; Mer et al. 2000; Shinohara et al. 1998; Park et al. 1996) and this interaction is required for its activity in recombination. ICP8 has not been shown to interact with Rad52, however. Could Rad52 compete with ICP8 for RPA binding? If so, then endogenous Rad52 may inhibit ICP8-mediated recombination by decreasing ICP8 access to replication intermediate and DNA metabolic factors. Furthermore, several lines of evidence point to competition between SSA and strand invasion pathways of recombination, suggesting that the negative influence of Rad51 on SSA may have a similar effect on Rcombineering using ssDNA oligos (Sugiyama & Kantake 2009; Plate et al. 2008; Wu et al. 2008; Park et al. 1996). In the future, the effects of both Rad52 and Rad51 depletion should be examined in the Rcombineering reporter lines generated in this thesis.

Is Rcombineering host specific?

The observation that ICP8 promotes gene targeting and that Beta does not suggests that ICP8 could work with human cell machinery to promote gene targeting (upon protocol optimization) while Beta does not. It is interesting that ICP8 forms nuclear
speckles that might be a sign of intranuclear localization or interaction with specific proteins in “recombination centers” (de Bruyn Kops & Knipe 1988). HumBeta localization is more diffuse throughout the nucleus, accumulating mostly at the periphery. The periphery of the nucleus is mostly populated with heterochromatin bound to the nuclear lamina while euchromatin localizes to specific sites towards the inside of the nucleus (Bartova et al. 2008). Maybe HumBeta does not have access to open chromatin or replication forks. On the other hand, Beta was previously show to stimulate gene conversion in mouse ES cells (Zhang et al. 2003). We wonder what are the differences between mouse ES cells and human 293T cells that could promote or inhibit Beta function. It is possible that the P2A peptide at the Ct and the NLS and the HA epitope peptides at the Nt could affect Beta’s function (Poteete 2011). In future studies, Recombineering using pSLIK2 HumBeta expression, where all the modifications are at the N terminus (NLS and the HA epitope, proline and a histidine) should be tested.

The trends in these data support the hypothesis that Recombineering is host specific. Even when Beta has been engineered to be expressed well and is localized to the nucleus, it might also be important that synaptases coordinate with host DNA metabolism via subnuclear localization to replication compartments to be efficient in human cells. It is interesting to recall at this point that ICP8 protein has been shown to interact with human proteins involved in homologous recombination, replication, heterochromatin remodeling and DNA repair (Taylor & Knipe 2004a) while my studies (see chapters 4 and 8) indicate that Beta protein interacts with bacterial proteins involved in homologous recombination, replication, nucleoprotein complex remodeling and DNA repair. We propose that subnuclear localization promoted by allele-specific protein interactions is essential for
human Recombineering and that similar interactions are likewise required for bacterial Recombineering.

Model for ICP8-dependent mutagenesis

In addition to stimulating human cell Recombineering, ICP8 also produced a significant increase in apparent gene convertants (Green cells produced from treatment of Mostaza cells) using a nonmutagenic “self” oligo with the sequence of the Mostaza target allele and also in the absence of oligos. This effect was seen using two different inducible ICP8 transgenes (though not in the experiments where ICP8 was introduced via transient transfection). This is a surprising conclusion and may reflect a different mechanism for targeted in vivo mutagenesis different than what was postulated in our Model for Recombineering. What, then, is the source of these Green variants? One possibility is that in addition to targeted Recombineering, ICP8 increases the general rate of mutation. In that case, perhaps the Green variants that appear in the absence of Green oligos result from ICP8-dependent Mostaza \( \rightarrow \) Green mutation. This is expected to be a relatively rare type of mutation, as the target size for this gain of function mutation is quite small and the range of substitutions are quite restricted. One prediction of this model is that loss of function alleles of the target gene should be stimulated by ICP8 expression to a much greater extent than the formation of Green variants as the target size would be a substantial portion of the Mostaza gene and the variety of mutations that could inactivate fluorescence is large.

In Figure 5.35 it is evident that the frequency of Dark variants far exceeds that described previously for Green variants. Furthermore, this is stimulated by doxycycline induction in pSLIK4 cells (see also Figure 5.34), though not in pSLIK1 cells or pSLIK5
cells expressing HumBeta. The frequency of Dark cells was oligo-independent and the sequence of the oligos had no bearing on the outcome.

Figure 5. Mostaza fluorescence is impaired upon doxycycline induction in pSLIK4.
A and B) 293T cells transfected with pSLIK1 or pSLIK4 were selected for zeocin resistance with 100 µg/ml zeocin for a week and seeded on a 24 well plate in the presence of 1000 ng/ml doxycycline. 8 hours later the media were changed and 4.13 nM oligo 85, 84 and 131 were independently transfected using Lipofectamine 2000 and Plus reagent. The oligo transfection mix was incubated with the cells for 17 hours. During this time, ~26% of pSLIK1 transduced cells and 60% of pSLIK4 transduced cells expressed Crimson upon doxycycline induction, indicating that only ~26% of pSLIK1 cells were expressing ICP8, while ~60% of pSLIK4 cells expressed ICP8. C) 293T-Mostaza cells transfected with pSLIK5 were selected for zeocin resistance with 300 µg/ml zeocin for a week and seeded on a 24 well plate in the presence of 100 ng/ml doxycycline. 19 hours later, doxycycline was washed out and 4.13 nM oligos 85, 84 and 131 were independently transfected using Lipofectamine 2000 and Plus reagent. The oligo transfection mix was incubated with the cells for 12 hours. During this time ~97% cells were induced by doxycycline for HumBeta/Crimson expression, indicating that most of the cells expressed the HumBeta protein. Antisense (as) and sense (s) Green oligos had the target sequence, differing by 4 nt that produce the Mostaza to Green conversion, while the Mostaza oligo had the exact sequence “self” of the target gene, and should not produce color change. Dark cells were lack Mostaza or fluorescent when recorded using Accuri flow cytometry. pSLIK1 (A) data was collected on day 13; pSLIK4 (B) data was collected on day 12 and pSLIK5 (C) data was collected on day 7. Data was analyzed using Prism 5 software.

So what, then, is the source of the Dark variants? These cells might possess the Mostaza transgene, but be unable to express it as might occur if epigenetic gene silencing were at the source. Alternatively, Mostaza may be expressed but in mutant configuration that does not produce an active fluorescent protein product. Finally, the Mostaza transgene may be absent. This might occur if a deletion covering the Mostaza gene occurred. These possibilities may be tested by PCR and DNA sequencing. PCR is also expected to distinguish random deletion of Mostaza from a site-specific deletion of the
transgene, for example by homologous recombination between the directly duplicated LTRs of the lentiviral insert. Such deletions have been previously observed in our lab in plasmid-borne lentiviruses (Lindle, Geffin and Myers, in preparation).

If mutation is in some way stimulated by ICP8 expression, what might be its mechanism? ICP8 is expected to compete with the host single-strand DNA binding protein, RPA, for replication intermediates. It has been reported that inhibiting RPA function increases mutation rates (Santocanale et al. 1995; Hass et al. 2010). ICP8 binding in replication forks might also stall replication and promote dsDNA break formation which is expected to increase the chance of Mostaza deletion either by random non-homologous end-joining or by targeted homologous recombination of the LTRs.

**Conclusions**

1. 293T gene targeting using unmodified oligos was optimized to levels 20-fold higher than what has been previously reported. 293T cells might be more proficient for gene targeting because of their higher transfection efficiency and/or defective innate immune response, impaired MMR, decreased DNA damage response induced apoptosis, increased replication in the presence of the T antigen and/or greater amount of cells at the G1/S transition when oligo DNA substrates are delivered.

2. ICP8 is active in human cell Recombineering and stimulates recombination. This establishes ICP8 as a *bono fide* recombinase in the SynExo family of viral two component recombinases and further encourages the development of organism-specific SynExos for genome editing in other species.

3. Oligo-mediated Recombineering changes the target gene sequence as predicted by the sequence of the oligo and by the phenotypic switch from Mostaza to Green.
4. Control oligos identical in sequence to the target gene do not produce a significant oligo-dependent Mostaza \( \rightarrow \) Green change.

5. The current protocol for ICP8 induction seems to produce off-target effects as inferred from a transient oligo-independent increase in a green cell population that decays over time. ICP8 expression also produces an increase in nonfluorescent Dark variants that are stably maintained and that correlate with ICP8 expression levels. These observations might represent ICP8-dependent Lentiviral LTR loop-out recombinants.

6. Human-optimized enterobacteria phage lambda Beta (HumBeta) does not catalyze Recombineering in human cells despite being well-expressed and localized to the nucleus, supporting the hypothesis that Recombineering is host-specific. Nuclear staining of HumBeta indicates that it fails to form the subnuclear punctate staining domains seen to occur in ICP8 expressing cells. This, too, is consistent with a failure of HumBeta to form host-specific protein-protein interactions, as will be further illuminated in Chapter 7.
CHAPTER 6. REPAIRING PROGERIA BY RECOMBINEERING

A long-term goal of the project was to apply Recombineering to repair a monogenic disease in a human cell. We were interested in the monogenic disease Hutchinson-Gilford progeria syndrome (HGPS, hereafter referred to as Progeria), which leads to premature aging and death by an average age of 13 years. The disease is caused by a de novo heterozygous synonymous substitution mutation (C1824T) of the lamin A/C (LMNA) gene (Eriksson et al. 2003; De Sandre-Giovannoli et al. 2003). The mutation activates a cryptic donor splice site in exon 11 of LMNA that is used about 85 % of the time to join exon 11 to exon 12 of the LMNA mRNA (Reddel & Weiss 2004). A mutant lamin A protein (Progerin) with a 50 amino acid internal deletion is generated by this aberrant splicing (Figure 6. 1). The deleted region includes a cleavage site for an endoprotease (Zmpste24) that normally removes the farnesylated C-terminus of lamin A. As a consequence, Progerin is permanently carboxyfarnesylated (De Sandre-Giovannoli et al. 2003; Eriksson et al. 2003; Dechat et al. 2007). Progerin incorporates abnormally into the nuclear lamina, leading to a host of pleiotropic effects (Scaffidi & Misteli 2005; Dechat et al. 2007; Cao et al. 2007; Goldman et al. 2004).

**Figure 6. 1** Mutation in LMNA activates a cryptic splice site, leading to aberrant splicing of pre-lamin A message.

The resulting Progerin protein retains farnesylation and causes the mutant phenotype.
The tissues most affected by Progeria in diseased children are maintained by marrow stromal stem cells (Hennekam 2006). Current models for this disease include fibroblasts derived from patients (cells that do not manifest the full mutant phenotype) and induced pluripotent stem cells derived thereof (which likely retain much of the fibroblast gene expression profile), so the available human stem cell models do not resemble the natural disease. Other models employ Progerin expression from a transgene, however since Progerin is not expressed at the normal disease-associated levels from its normal locus, there is uncertainty about the fidelity of this model due to problems arising from Progerin over-expression and from the ectopic location of the transgene insertion. The best model to study Progeria will have the natural mutation in the affected stem cell population at the normal locus. It is possible to create such mutations by targeting small DNA oligos to modify LMNA with the disease-causing allele (C1824T), but the efficiency of this sort of alteration occurs $\leq 10^{-4}$ using current gene editing technologies. Therefore, there is a need to increase the recombination efficiency in human stem cells to create disease models. The approach we sought to employ is Recombineering.

**Progeria Recombineering reporter strain**

In order to study Recombineering in the repair of Progeria and to have a quantifiable output to initially optimize the reaction, we obtained a HeLa strain that contains a Progerin splicing reporter “mini-gene” construct that can be used to quantify recombination frequencies by changes in fluorescence linked to modifying LMNA sequences (Figure 6.2). This reporter was especially interesting because it is perfect for detecting repair of the Progeria-causing mutation LMNA(C1824T) *in vivo.*
Figure 6. 2 Progerin reporter HeLa cell line used to quantify Recombineering.
Progerin splicing reporter ectopically encodes a portion of the Progerin gene with a eGFP gene fused to LMNA exon 12, so that the fusion protein created by the Progerin splicing reaction is GFP+. A stop codon introduced between the C1824T allele and the native splice donor site is polar on GFP expression, making cells GFP- when the LMNA transcript is spliced properly in this reporter. A dsRed gene expressed from a CMV promoter is linked to the Progerin reporter construct. The Progerin reporter expresses Progerin-eGFP and dsRed proteins, but upon correction of T1824 to C1824, only dsRed is expressed.

Dr. Tom Misteli’s Lab created a Progerin splicing reporter that ectopically encodes a portion of the Progerin gene. The construction has the eGFP gene fused to LMNA exon 12, so that the fusion protein created by the Progerin splicing reaction is GFP+. In this construction, a stop codon was introduced between the C1824T allele and the native splice donor site between exon 11 and the intron. When the LMNA transcript is spliced properly in this reporter, the stop codon in exon 11 is polar on GFP expression, making cells GFP-. Therefore, it is possible to monitor replacement of T1824 with the “C” residue at that site by loss of GFP fluorescence. Since loss of fluorescence could be caused by trivial reasons not related to Recombineering, a dsRed gene expressed from a CMV promoter is linked tightly to the Progerin reporter construct. Therefore, the
Progerin reporter expresses Progerin-eGFP and dsRed proteins, making cells yellow but upon correction of T1824 to C1824 in vivo by Recombineering, only dsRed will be expressed resulting in red cells.

The spectra of the reporter proteins was determined in order to select a filter set for quantifying gene targeting and imaging the cells. The spectra of the HeLa Progerin reporter strain extracts were consistent with the GFP gene being eGFP and the RFP gene being dsRed2 (Figure 6.3). The spectra of these proteins show significant excitation overlap. From this analysis, we learned that in order to differentiate green from red fluorescence, a narrow band pass filter for eGFP was needed to differentiate eGFP+ cells (non-recombinants) from eGFP- cells (recombinants).

![Figure 6.3](image)

**Figure 6.3** The Progerin-GFP excitation spectrum partially overlaps the dsRed excitation spectrum.
The pPRO1.1go HeLa cell line carrying the Progerin Reporter expressed both dsRed and eGFP. Protein was extracted from HeLa cells in 1 % deoxycholic acid, 1 % NP40, 25 mM Tris–HCl (pH 8), 50 mM NaCl and protease inhibitor cocktail from Sigma (P8340). The cells were incubated at 4 °C for 30 minutes on a shaker and scraped from the dish. Samples were sonicated (at power 7, 1 minute on and 1 minute off for 3 pulses) and then centrifuged at 16 000 g for 5 minutes. The excitation and emission spectra of eGFP and dsRed fluorescence in the extracts were determined with our spectrofluorometer (PTI QuantumMaster) with 2 nm slits at a slew rate of 1 nm/second (the average integrated one second samples collected ~10⁴ - 10⁶ emitted photons). To collect eGFP excitation spectra, the excitation monochromator scanned from 430 to 530 nm and emission was monitored at 540 nm. To collect the GFP emission spectra, samples were excited at 482 nm, and the emission spectra were collected from 490 to 690 nm. To collect dsRed excitation spectra, the excitation monochromator scanned from 430 to 590 nm and emission was monitored at 650 nm. To collect the dsRed emission spectra, samples were excited at 530 nm, and the emission spectra were collected from 540 to 700 nm. The data were collected using the Felix software, and later analyzed using Microsoft Excel and GraphPad Prism. Each spectrum was normalized to its peak value and plotted.
The Progerin reporter cell line was imaged using the Celigo (Figure 6. 4). All the analyzed clones emitted red and green fluorescence, but with a range of fluorescence intensity for both colors between cells. It also showed some cells within the clones that were not fluorescent, which may be a source of concern when trying to identify recombinants (I will address this, below).

![Figure 6. 4 Progerin reporter HeLa cell line fluoresce green and red.](image)

Cells were grown in flat bottom 96 well plates. Just before imaging, media was changed to HBSS. Images were taken using a Celigo adherent cell cytometer with live cells.

**Developing recombination quantification assay**

After recombination assays, cells were scored for loss of green fluorescence with retention of red fluorescence using flow cytometry (Figure 6. 5). The data could be better if we had a non-fluorescent control of the same cell line so we could analyze autofluorescence and scatter from the same cell line. The data could also be improved if single-color fluorescence HeLa cell line control were available to adjust compensation for each fluorescent protein. A Green- Red+ clone was isolated for this purpose from a Recombineering experiment with the parental Progerin reporter Green+ Red+ strain (Figure 6. 5). When using this control for compensation, the data look much better.
Unfortunately, in this clone the Red signal looks like it also decreased a bit, as if both fluorescent genes were less expressed. Also, the dsRed signal is very strong, approaching the limits of the photomultiplier used to quantify fluorescence intensity. By using the Green- Red+ strain, the sorter settings were compensated to better evaluate the data, but the Red signal remained very strong. To reduce the big signal of dsRed maybe it would be better to excite both fluorescent proteins with the 488 laser, which is not the biggest excitation peak for dsRed, and to then record fluorescence with different filters. This is predicted to decrease the emitted red fluorescence intensity to better balance the green and red outputs.

![Figure 6. 5 Recombination assay: loss of green fluorescence.](image)

Recombinants were also quantified using the Celigo. A mock experiment was performed to determine if the Celigo could identify putative recombinants from the original strain. For this purpose, Green- Red+ cells and Green+ Red+ cells were mixed.
and plated covering a range of 0 %, 0.1 %, 0.5 %, 1 %, 5 %, 10 %, 50 % and 100 % Green- Red+ among total cells plated. Each of these mixes were plated in 12 replicate wells for improved statistical analysis and grown for 4 days before collecting the data. The % of green fluorescent cells vs the total amount of red fluorescent cells per well was recorded by the instrument. The % of Green- Red+ cells recorded was plotted against the % of Green-Red+ cells plated 4 days earlier (Figure 6.6). This plot depicts what we could expect in a recombination experiment: loss of green fluorescence with retention of red fluorescence. The plot produced a straight line with $R^2 = 0.996$ and equation $y = 1.303 + 0.9728x$. The $R^2$ indicates that there is a one to one correspondence between the fraction of green cells in the mixture and what is observed by the Celigo; in other words this instrument is quantitative. An intercept of 1.3 indicates that there is some background green fluorescence that will need to be subtracted to make the line pass through the origin.

![Graph](image)

**Figure 6.6** Quantification of mock recombinants with Celigo.
Progerin reporter cells Green+ Red+ (G'R+) were mixed with red florescent cells Green- Red+ (G'R-) at different ratios. Green and red fluorescent cells were quantified using the Celigo. The plated % of G'R/ total was plotted in the X axes and the observed % of G'R/total was plotted in the Y axes.
The Celigo had advantages over flow cytometry since it was possible to record pictures in addition to obtain quantitative data. The Celigo was less invasive, since cells were analyzed while attached to the plates used to grow them. The Celigo was especially good for clonal analysis, where the clones could have as little as 4 cells.

Flow cytometry had advantages over the Celigo, too, as it was possible to analyze a large number of cells and sort them to study genotypes. As recombination frequencies could be very low, flow cytometry was the preferred technology to quantify recombinants vs non recombinants and to isolate putative recombinants by cell sorting. These putative recombinants were then plated and verified using the Celigo, as describe below.

**Recombination in Progerin HeLa cells**

Gene targeting was evaluated using 99 nucleotide long oligos complementary to the Progerin reporter mini-gene. The oligos carried a single nucleotide change to revert the splicing defect caused by the T1824 allele in the Progerin reporter (Figure 6. 7). Oligos were designed to target both DNA strands since Recombineering is more efficient when targeting the lagging strand template (Costantino & Court 2003a; Valledor et al. 2012), but we had no information about the location of the Progerin reporter in the human genome, and therefore we had no idea which was the direction of the replication fork over the Progerin reporter. In Chapter 5, we noted that the SV40 origin embedded in the promoter/enhancer region could be activated by the SV40 large T antigen integrated in the HEK 293T cell line. In the HeLa Progerin reporter cell line, there is an SV40 promoter/enhancer driving expression of dsRed2. We do not expect that the T antigen is present in the HeLa, so we do not anticipate that the SV40 origin is active in this reporter (NB it is also not understood how dsRed2 transcription originates from the SV40
promoter/enhancer without T antigen…). Oligos targeted to either strand were used in separate trials. Oligo LMNA1824C was designed to hybridize to the antisense strand and other three oligos were designed to hybridize to the sense strand. Oligos sense LMNA1824C and antisense LMNA1824C were designed to revert the mutation T1824 to C1824. Unfortunately, both of these oligos created highly reparable mismatches when annealed to the Progerin reporter and MMR is active in HeLa cells (Cannavo et al. 2007). A synonymous substitution that avoided MMR was pursued to design oligos that avoid the best MMR. The Progeria allele is a synonymous substitution in a glycine (Gly) codon in exon 11 of LMNA. An analysis of Gly codons showed that A1824 and G1824 will also encode for Gly and disrupt the Progerin splice donor site. From the four possibilities of targeting leading and lagging strand for each of these substitutions, the mismatches corresponding to the oligos hybridizing the sense strand were the less recognizable by the MMR as per (Su et al. 1988) and were therefore selected for these studies. Both oligos LMNA1824A, and LMNA1824G, hybridize to the sense strand.

Figure 6.7 Oligos to target the Progerin minigene. LMNA oligos designed as the target region with an indicated change to eliminate the cryptic splicing site. Progerin oligo designed as a gene targeting control and as a source to create a Progerin disease model cell line. Sense oligos were represented by “s” and antisense oligos were represented as “as”.

The source of the synaptase was pCMV-ICP8 transient transfection, using Amaxa electroporation. Expression of ICP8 in the Progerin reporter was evaluated by Western
Blot (Figure 6. 8) and confocal microscopy (Figure 6. 9). The Western blot analysis showed that ICP8 is expressed at the expected size of 129 KDa.

![Western Blot analysis of ICP8 expression.](image)

Progerin reporter cells were transfected with pCMV-ICP8 and probed with an anti-ICP8 antibody.

Confocal microscopy was performed in a Green- Red+ progerin line. ICP8 expression was evaluated from the pCMV-ICP8-GFP plasmid. ICP8-GFP showed nuclear localization as reported previously. Some cells showed diffuse nuclear ICP8-GFP while in other cells, ICP8-GFP fluorescence appeared as bright punctuate spots (Taylor & Knipe 2004b). (de Bruyn Kops & Knipe 1988) reported that ICP8 spots are associated with pre-replication centers where the host cell DNA replication apparatus is recruited by ICP8. ICP8 may be associated with subnuclear PML bodies where proteins are prepared for degradation unless ICP8 disrupts PML body activity in the creation of pre-replicative compartments (Burkham et al. 1998).
Figure 6. 9 ICP8-GFP localizes to the nucleus in the progerin reporter cell line.
The cell sorter was used to isolate a derivative of the HeLa Progerin reporter following ICP8-mediated Recombineering that expresses DsRed but not GFP. This strain was subsequently transfected with pCMV-ICP8-GFP and live cells were visualized using the Zeiss LSM 780 laser scanning confocal microscope live in HBSS, 8 hours after transfection.

Table 6. 1 pCMV-ICP8-GFP transfection efficiency
Cells were transfected with the indicated amount of plasmid and imaged 8 hours later live in a Zeiss confocal.

<table>
<thead>
<tr>
<th>Progerin Reporter Cells</th>
<th>5 µg (3 Z stacks)</th>
<th>2 µg (2 Pictures)</th>
<th>1 µg (2 Pictures)</th>
<th>1 µg (5x5 tile)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Green</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>27</td>
</tr>
<tr>
<td>Total</td>
<td>25</td>
<td>17</td>
<td>24</td>
<td>530</td>
</tr>
<tr>
<td>% (green/total)</td>
<td>24%</td>
<td>35%</td>
<td>25 %</td>
<td>6 %</td>
</tr>
</tbody>
</table>
In an attempt to determine if the transfection efficiency correlates with the plasmid concentration, cells were transfected with 1, 2 & 5 µg of pCMV-ICP8-GFP as shown in Figure 6.9. The transfection efficiency analyzed from the pictures were between 24 and 35 % (Table 6.1). Chi squared analysis shows no significant difference between the samples, indicating that something else in the transfection protocol other than plasmid concentration is the limiting step. Unfortunately, these microscopic fields of view were selected by the appearance of green fluorescent cells, and not randomly. To overcome bias caused by eye selection, a tile of 1 x 1 mm was evaluated for the 1 µg of pCMV-ICP8-GFP chamber. In the tile, cells that were not in the same depth of field were excluded. Transformation efficiency was estimated as the fraction of green nuclei among total nuclein detected with DAPI. The analysis of the tile showed a transfection efficiency of 6 % which is more reliable because they were not biased by pre-selection for green nuclei (Table 6.1).

The first Recombineering experiment was evaluated by cotransformation of 10 µg pCMV-ICP8 and 1 µg of oligo using Amaxa electroporation. The cells were grown in 5 % serum pre-electroporation and in 10 % serum post-electroporation. A caveat of this experiment was that the cells grown for this experiment suffered an unplanned 41.5 °C heat shock (courtesy of an incubator thermostat failure) for an indeterminate period of time that could have been as long as 48 hours. The high temperatures in the incubator were returned to normal one day before the electroporation. After electroporation, gene conversion was evaluated using flow cytometry.
The signal for Progerin gene conversion was loss of green fluorescence, therefore the outgrowth time was performed long enough to allow segregation between recombinants and non-recombinant cells and GFP protein decay within the non-recombinants. If only one copy of the Progerin reporter was inserted in the HeLa genome, then upon “T1824 correction” in one strand (heteroduplex), cells will need to divide and replicate again for the change to be incorporated in both strands (homoduplex). HeLa cells divide approximately once a day. Therefore, we predicted that if transfection does not affect cell division rate, recombinant DNA strands should segregate within two days after transfection. However, these clones were expected to retain about half the eGFP protein of a non-recombinant cell since eGFP has $t_{1/2}$ of 1 day (Verkhusha et al. 2003). Therefore recombinants were analyzed 3 days after the DNA transfection.

Loss of green fluorescence among red fluorescent cells was monitored using flow cytometry as described above (Figure 6. 5). Red only cells from quadrant 1 (Q1) were recovered using the Aria II sorter. Those cells were plated and subsequently analyzed using the Celigo. Green- Red+ clones were expanded and analyzed by PCR and sequencing.

The first Recombineering experiment produced 11.65 % ± 1.45 % Green- Red+ putative recombinant cells after subtracting the background of 5 % pre-existing Green-non-recombinant cells in the culture (Figure 6. 10). These frequencies were higher than reported before by other studies, but the background frequency of no treated cells was really high too. It is possible that the high temperature on this experiment stimulated spontaneous loss of green fluorescence. Pretreatment of human cells with a heat shock has a parallel with the induction conditions we use with the E. coli Recombineering
system in which Recombinases are induced by a heat shock (and in which Recombineering is very efficient).

Figure 6. 10 Quantification of putative recombinants by flow cytometry, exp 1.
Cells were grown in 5 % serum to 70–80% confluency for a total of 3 days. At about, 48 hours of growth, the incubator was found at 41.5 °C. Cells were collected and transfected simultaneously with 10 µg pCMV-ICP8 and 1 µg oligo using Nucleofector II and the program Q-013. Cells were grown until they reached $10^7$. Loss of green fluorescence was analyzed by flow cytometry.

Because the first experiment was affected by the unplanned heat shock, the experiment was repeated. Cells were grown to 70 – 80 % confluency at 21 % or 3 % O₂ tension as indicated. Cells were collected and transfected with pCMV-ICP8 and oligo using Nucleofector II and the program Q-013. Cells were grown until they reached $10^7$ and were sorted for loss of green fluorescence. This transfection produced lots of cells floating in the culture. Cells were grown at 21 % and 3 % O₂ tension to evaluate the effect
of oxygen tension in MMR, since a highly reparable mismatch needs to be created.
(Rodríguez-Jiménez et al. 2008), reported a down regulation of mismatch repair genes at low O2 tensions. For this second Recombineering experiment, cells were transfected with 1) no oligo, 2) oligo alone, 3) cotransformation of oligo and 10 µg of pCMV-ICP8, and 4) sequential transformation of 10 µg of pCMV-ICP8 followed with oligo 72 hours later to allow recombinase expression prior to introducing the recombination substrate.

The results for this second Recombineering experiment (Figure 6.11) were different than the results from the first experiment. This time the co-electroporation of oligo and recombinase did not produce a change in fluorescence in most of the cases except for two of the trials where cells were grown at 3 % O2 (Figure 6.11, green bars).

**Figure 6.11 Quantification of putative recombinants by flow cytometry, exp 2.**
Cells were grown to 70 – 80 % confluency at 21 % or 3 % O2 tension. Cells were collected and transfected simultaneously with 10 µg pCMV-ICP8 and 1 µg oligo using Nucleofector II and the program Q-013. Cells were grown until they reached 10⁷. Loss of green fluorescence was analyzed by flow cytometry.
Interestingly, oligos alone produced putative recombinants at a rate of about 2.28 \% \pm 0.26 in cells that were grown at 21 \% O_2. The rate of green loss between the different oligos was contrary to the expectations. Targeting the leading or lagging strand did not make a significant difference in the results. Furthermore, MMR may not have been a factor, since each of the oligos worked equally well despite being predicted to make mismatches in heteroduplex recombination intermediates that are corrected with different efficiency by MMR. These results were strange since the strand bias and the MMR effect has been recognized as a conserved mechanism across taxa (E. coli (Costantino & Court 2003a); murine ES cells (Dekker et al. 2006); human hepatocytes (Igoucheva et al. 2008)). However, lack of strand specificity might be a consequence of working with a nonclonal cell line with random integration of the Progerin reporter (this is not known at this time). 21 \% O_2 tension produced more putative recombinants for oligo alone.

Addition of pCMV-ICP8 decreased the frequencies of green loss indicating that ICP8 did not help oligo-mediated recombination. Chi square analysis showed significant inhibition, P < 0.05, for three of eight trials when pCMV-ICP8 was electroporated 72 hours before the oligo. Our initial hypothesis was that ICP8 would catalyze homologous recombination and that we would see an increase in the formation of recombinants in cells expressing the recombinase. It is possible that constitutive high-level expression of ICP8 is suboptimal for recombination. Since there was so much cell death when transfecting 10 \mu g pCMV-ICP8, a strong possibility was that the 10 \mu g of pCMV-ICP8 DNA activated the innate immune response. Innate immunity might destabilize the oligos and/or induce plasmid-transfected cells to die (Hornung & Latz 2010).
Validating recombinants

The next question was to determine if the GFP-putative recombinants was accompanied by a corresponding change in the genotype. Putative recombinants were isolated by cell sorting and single clones were expanded as shown in Figure 6.12. As is shown in the top of Figure 6.12, some clones expressed eGFP while others lost the green fluorescence while maintaining red fluorescence. The images in the bottom of Figure 6.12 are interesting, since they resemble a sectored clone arising from segregation of heteroduplex recombination intermediates previously observed while Recombineering in *E. coli*.

![Image](image.png)

**Figure 6.12 Isolating Green+Red+ clones with Celigo.**
Cells electroporated with oligos were sorted for loss of green fluorescence and maintenance of red fluorescence. They were seeded in plates at low dilutions to isolate single clones. Clones were visualized on the Celigo in HBSS.
It is notable that there was a fraction of non-DNA transfected Progerin reporter cells that had low-level GFP expression as if the green phenotype was not 100% even in the absence of treatment.

Green- Red+ clones were expanded to verify putative recombinants by PCR and sequencing. Genomic DNA was isolated from 30 clones. PCR was performed in order to amplify the splicing site region with primers 1 and 8 as shown in Figure 6.13 and in Table 6.2. Primers were designed such that a fragment of 1173 bp containing the Progerin splicing site in the middle of the product was expected, but only the non-treated control and one of the 30 clones produced a product of the expected size (Figure 6.14A). The PCR products from control cells and clone 2-7 were sent to be sequenced. Both sequences were identical, indicating that clone 2-7 was not a recombinant despite being Green- Red+.

A genomic DNA sample from Green- Red+ clone 2-4 failed to produce the expected 1173 bp PCR product, so it was subjected to PCR using 7 other primer pairs that hybridize within the reporter region to see if any of the Progerin reporter sequences were present. The non-treated control produced the expected product for each of the 7 primer pairs but none of these additional PCR reactions produced a product from clone 2-4 (figures not shown). To control for the quality of the genomic DNA, PCR was performed to amplify a fragment in the native LMNA locus using another primer pair 3+7 (Figure 6.14B). Four putative recombinant clones (including 2-4) and the non-treated control were tested and all of them produced the expected LMNA product, indicating that all the samples had genomic DNA suitable for PCR.
The Progerin reporter is constitutively expressed from the CMV promoter while the closely linked dsRed2 control reporter is constitutively expressed from the SV40 promoter. The products of Progerin splicing and normal Lamin A splicing are represented, along with the associated phenotype. PCR was performed on DNA isolated from sorted Green- Red+ cells for genotyping. The primers used to amplify portions of this region by PCR are represented by the arrows and numbers. Table 6.2 details the name, sequence and positions of each primer in the reporter.

### Table 6.2 Primers used to amplify the Progerin reporter.

The oligo number correspond the number indicated in Figure 6.13. The oligo name correspond to the number and name indicated in the Material and Methods Chapter.

<table>
<thead>
<tr>
<th>Oligo #</th>
<th>Oligo name</th>
<th>Sequence</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>28.Misteli Forward</td>
<td>TCGTGAGGCTGGGCGAGGT</td>
<td>872-891</td>
</tr>
<tr>
<td>2</td>
<td>30.Misteli Plus</td>
<td>ACAGGTTGCTCCACTCCCCAGTCA</td>
<td>1019-1040</td>
</tr>
<tr>
<td>3</td>
<td>38.T1824_Progerin</td>
<td>GCTCAGGAGCCCAGGAGGTGGGT</td>
<td>1338-1357</td>
</tr>
<tr>
<td></td>
<td>39.C1824_LMNA</td>
<td>GCTCAGGAGCCCAGGAGGTGCC</td>
<td>1338-1357</td>
</tr>
<tr>
<td>4</td>
<td>40.T1824_anti_Progerin</td>
<td>GAGCCAGAGAGATGTTGCCCA</td>
<td>1357-1377</td>
</tr>
<tr>
<td></td>
<td>41.C1824_anti_LMNA</td>
<td>GAGCCAGAGAGATGTTGCCG</td>
<td>1357-1377</td>
</tr>
<tr>
<td></td>
<td>42.A1824_anti_LMNA</td>
<td>GAGCCAGAGAGATGTTGTCCT</td>
<td>1357-1377</td>
</tr>
<tr>
<td></td>
<td>43.G1824_anti_LMNA</td>
<td>GAGCCAGAGAGATGTTGTCCC</td>
<td>1357-1377</td>
</tr>
<tr>
<td>5</td>
<td>31.Misteli Minus</td>
<td>ACATGATGCTGCAGTTCTGGGG</td>
<td>1827-1848</td>
</tr>
<tr>
<td>6</td>
<td>29.Misteli Reverse</td>
<td>GGCACACCCCCGGTGAACAG</td>
<td>2025-2044</td>
</tr>
<tr>
<td>7</td>
<td>76.dsRed_Fw</td>
<td>CCAGTCCAGTACGGCTCCAAGGT</td>
<td>4357-4380</td>
</tr>
<tr>
<td>8</td>
<td>77.dsRed_Rv</td>
<td>GAGCTCGAGATCTCAGAAGCTGGGT</td>
<td>4833-4856</td>
</tr>
<tr>
<td>9</td>
<td>126.Progerin Fw</td>
<td>TCCACAGGTGTCCACTCCCCAGTTC</td>
<td>1016-1039</td>
</tr>
<tr>
<td></td>
<td>127.dsRed Rev</td>
<td>ATGTCGCGGAGGTGCTTTCAC</td>
<td>4385-4404</td>
</tr>
</tbody>
</table>

As all these clones remain Red+, any putative deletion of the reporter construct was expected to have left the dsRed2 gene intact. Another PCR was performed from dsRed2 using primers 1 and 8 in the figure, but only the Progerin control produced a product.
(Figure 6. 14C). Since the clones were red upon sample collection, this result might indicate that the region from primer 1 is altered. In light of this result, the inability to amplify the Progerin reporter region in clone 2-4 suggests that this region may have been deleted in 2-4 and possibly in other putative recombinant clones.

![PCR amplification from genomic DNA of isolated clones.]

After recombination experiment, genomic DNA from clones was isolated. The target region was amplify by PCR. A) PCR amplification of splicing region with primer pair 1+8 as shown in Figure 5. 11. The expected size for this product was 1173 bp. B) PCR amplification of LMNA locus with primer pairs 3+7. C) PCR amplification from dsRed locus with primer pairs 1+8.

HeLa cell lines are notoriously unstable and have about 82 chromosomes with many examples of genome rearrangement (Kadaja et al. 2009; Macville et al. 1999). As a consequence, we might expect that genomic rearrangements could influence our ability to unambiguously identify recombinants in this reporter strain. Nevertheless, it remains intriguing that introduction of the oligos (in the absence of ICP8) targeted to the Progerin splicing allele led to significant (P<0.001) enhancement of loss of green fluorescence with retention of red fluorescence. An interesting question is: Are the oligos inducing
genomic rearrangements in this HeLa cell line? This question could be tested by using 4 oligos to test transformation-dependent and homology dependent deletion of the target site: 1) an oligo to repair the Progerin splicing site; 2) an oligo identical to the Progerin splicing site; 3) an oligo to create a stop codon in dsRed2; and 4) an oligo identical to a portion of dsRed. This experiment should tell us the effect of oligo treatment in HeLa genome stability.

The inability to amplify the splicing region after the Recombineering experiments may discourage the use of HeLa cells to study recombination. Analysis of the Progerin reporter plasmid showed co-oriented 42 bp repeats of the multiple cloning site flanking the reporter region (5’ to Progerin and 3’ to dsRed2, Figure 6. 15). Cells were red, therefore, it is expected that either the dsRed2 locus is present in the clones or the dsRed2 protein half-life is longer than that of eGFP and that dsRed2 fluorescence is maintained for a period after eGFP fluorescence is no longer measurable. Indeed, dsRed2 protein can persist for more than a month with a half life of 4.6 days, while eGFP disappears within days with half life of one day (Verkhusha et al. 2003). Interestingly, the frequencies of gene targeting from rec1 was similar to appearance of dark cells from the Mostaza reporter. Perhaps ICP8 promote recombination between co-oriented repeats that flank the reporter construct. This might generate cells that are Green- Red+ cells if assayed before dsRed2 protein is degraded.

CMV Promoter/enhancer
Progerin
eGFP
MCS-SV40 polyA
SV40 Promoter/enhancer
dsRed2
Synthetic polyA-AmpR-(likely an E. coli origin in here, too)
Directly repeated sequence flanking eGFP-5’ and dsRed2-3’. Loop out Green- Red-
Conclusions

Preliminary human Recombineering experiments were performed in a Progeria reporter model cell line. Phenotypic data suggested that oligo-dependent Recombineering was successful. A heat shock experiment produced 12% more putative recombinants than the no-DNA control. A second experiment without heat shock produced lower rates of gene conversion. Oligo alone produced a 2 fold increase in the putative recombinants at 21% O₂ tension and 0.5 fold at 3% O₂ tension. With transient transfection by pCMV-ICP8, an inhibitory effect was seen at 21% O₂ tension and no effect was seen at 3% O₂ tension. The caveat for this experiment is that many more cells were floating (possibly dead) after transfection. The floating cells could indicate that just the DNA transfection could have affected the viability of the transfected cells. Therefore, the effect of ICP8 in this second experiment is inconclusive. A limited amount of follow-up genotypic data have not been supportive, but suggest that the Progerin splicing region has been deleted in an oligo-dependent reaction.
CHAPTER 7. SYNEXO ALLELE SPECIFIC INTERACTIONS

Functional interactions between viral recombination proteins and host proteins

Understanding the mechanism of Recombineering requires the knowledge if viral recombinases function as independent units or if they coordinate their activity with host proteins through allele-specific protein interactions. The working hypothesis of this study was that viral recombinases coordinate their activity with specific host proteins through protein-protein and functional interactions. The prediction was that viral recombination proteins recruit host DNA binding proteins that increase recombination efficiency by: protecting ssDNA recombination intermediates from host exonucleases needed to promote homologous DNA pairing, by resolving recombination intermediates, by coordinating recombination with host DNA replication, and by regulating completion of recombination via disassembly of proteins that remain bound to recombination products.

The Recombineering mechanism was studied by elucidating the proteins physical interactions. Protein interactions were studied by affinity purification of enterobacteria phage λ SynExo-associated proteins assembled in vitro from purified components and evaluated by SDS PAGE and in vivo in E. coli cells expressing SynExo recombinase subunits with mass spectrometric (MS) analysis and associated bioinformatics to identify putative interaction partners. Functional interactions between the phage λ SynExo and E. coli host factors were studied using genetic approaches in vivo and the ReDa exonuclease assay in vitro. The results were then evaluated in light of complementary studies performed by others using the orthologous SynExo recombinase from Human Herpes
Virus type 1 expressed in human cells (Zhou & Knipe 2002; Taylor & Knipe 2004a; Balasubramanian et al. 2010; Mohni et al. 2011).

**Physical interaction studies using pull downs**

A pull down assay was performed to study if *E. coli* host ssDNA binding proteins physically interact with phage λ recombinases. The pull down assay was performed using Ni-NTA beads and His6-tagged Beta as the bait. Interaction of His-Beta with λ Exo, RecA (host) and SSB (host) was examined. The protein gels with bound and unbound protein fractions are shown in **Figure 7.1**.

![Protein gels](image)

**Figure 7.1 Beta forms a complex with λExo that disappears in the presence of RecA.**

His-Beta was used as the bait to examine the interaction with λ Exo, RecA and SSB using Ni-NTA beads. Protein complexes are indicated on the top of the gel lane. The upper gel shows the unbound fractions and the lower gel the bound fractions. Proteins bands were identified on the left side of the gel. The inset in red shows the effect of host protein in Beta and λ Exo interactions.

The bound fraction confirmed the previous observations that Beta and λ Exo form a complex (Radding & Shreffler 1966; Tolun 2008). Host proteins SSB and RecA were only visible in the unbound fraction, indicating no interaction with Beta. While SSB host protein did not interfere with Beta and λ Exo complex, RecA did. λ Exo and Beta
interaction was reduced by RecA suggesting competition between RecA and Beta for λ Exo. This result is consistent with known properties of related recombination proteins in yeast and E. coli (Beernink & Morrical 1999).

**Physical interaction studies using pull downs and mass spec**

Host and viral protein interactions were also studied from cell lysates to elucidate new interactions. For these studies, the genes for the λ SynExo recombination proteins (Figure 7.2) were cloned into the Flexi vectors from Promega fused to a Halo tag that facilitated affinity purification from cell lysates. Full-length SynExo components genes for Beta and λ Exo were cloned in addition to truncated version of those genes: Beta (1-177) and λ Exo (1-189), which were designed to test interactions within the SynExo complex.

![Diagram](image)

**Figure 7.2** λ SynExo genes were clone fused to a Halo tag for MS analysis. Each SynExo gene and their deletion derivatives were cloned independently and fused with a HaloTag to facilitate affinity purification.

The SynExo proteins were then expressed independently in E. coli strain BL21(DE3) that had been induced to express lambda proteins from the DE3 prophage and purified as complexes with E. coli host factors (Figure 7.3). Proteins co-purifying with SynExo
recombination proteins were fractionated by SDS PAGE, excised from gels, digested by trypsin, separated by HPLC and analyzed by electrospray mass spectrometry (Figure 7.4).

![Image of SDS-PAGE gel with bands labeled Beta, Beta(1-177), Exo, and Exo(1-189).]

**Figure 7.3** Host proteins co-purifying with λ SynExo proteins to be identified by MS. BL21(DE3) [pFN22K::SynExo] strains were induced for 3 hours, and clear lysates were made. Lysates were incubated with HaloLink resin to covalently link tagged protein complexes to the support. After extensive washing, proteins were eluted by using TEV Protease to cleave the HaloTag off the SynExo subunit. His-TEV protease was then removed by IMAC.

Putative host proteins interacting with λ SynExo proteins were identified from a MASCOT query of the BL21(DE3) peptide mass/charge predication database (http://www.matrixscience.com/search_form_select.html) and follow-up using MS tools at EXPASY (http://expasy.org/tools/). Proteins found in complexes with one or more SynExo proteins were evaluated in the STRING database of protein interactions in BL21(DE3) (http://string-db.org/) and by text mining in PUBMED (http://www.ncbi.nlm.nih.gov/pubmed) including genetic and enzymological studies. Proteins so identified suggest mechanisms for viral recombination coupled to host cellular replication and recombination machinery and how Recombineering complexes might be regulated by proteolysis. The results are summarized, below.
Beta interaction with λ Exo was confirmed

There is extensive evidence that Beta and λ Exo form a complex. For example: λ Exo co-purifies with Beta in a 1:1 complex (Carter & Radding 1971; Tolun 2008); λ Exo stimulates Beta and Beta stimulates λ Exo activity in vitro (Tolun 2008); viral SynExo-mediated dsDNA recombination requires homospecific Exo/Synaptase pairs that work together as a functional module (Muyrers et al. 2000; Datta et al. 2008) and that they co-purify when performing pull-downs (Muyrers 2000). The Mass Spec data in this work confirmed previous observations that Beta and λ Exo form a complex.

λ Exo Ct α helix is not required for Beta interactions

Quite frequently, DNA metabolic factors interact via unstructured C-terminal (Ct) protein interaction domains. A truncated version of the λ Exo (1-189) gene was designed to test if the Ct domain α helix, represented in Figure 7. 5, was involved in protein-
protein interactions. The MS data showed that Beta expressed from the DE3 prophage copurified with HaloTag-\(\lambda\) Exo (1-189), indicating that the \(\lambda\) Exo Ct domain was not required for \(\lambda\) Exo-Beta interactions. Our result is consistent with the prediction that \(\lambda\) Exo-Beta interaction surface is not C-terminal but might include a helix between amino acids 75-95 (Kovall & Matthews 1998). That remains to be tested.

**Beta Ct domain seems to be required for interactions with \(\lambda\) Exo**

The Ct acidic unstructured domain of Beta decreases Beta affinity for ssDNA (Wu et al. 2006). Therefore, we hypothesized that this unstructured domain might be involved in protein-protein interactions. The MS data showed that \(\lambda\) Exo expressed from the DE3 prophage copurified with HaloTag-Beta but not with HaloTag-Beta (1-177). Indicating that this Ct domain might be at the interaction surface between \(\lambda\) Exo and Beta. These data need to be validated by reciprocal pull-downs especially in the light that the
truncated version of Beta interacted with the analogous RecE SynExo nuclease during this Mass Spec experiment.

λ Exo and Beta host proteins partners

The MS analysis showed that λ Exo interacted with Lon, OmpF, YegS, MviN, and ClpP, while Beta interacted with Lon, OmpF, YegS, MviN, RecE, and UvrD. Binding to Lon, OmpF, YegS and MviN proteins was common to both λ Exo and Beta.

The λ Exo Ct domain was required for most host protein interactions. λ Exo (1-189) only interacted with OmpF. In contrast, the Beta Ct domain didn’t show evidence of being required for host protein interactions. Beta (1-179) interacted with more proteins than full-length Beta. A list of the λ SynExo interacting partners is represented in Table 7.1 and in Figure 7.6.

Proteins so identified suggest mechanisms for viral recombination coupled to host cellular replication and recombination machinery and how Recombineering complexes might be regulated by proteolysis. The results are summarized, below.

<table>
<thead>
<tr>
<th>Candidates</th>
<th>Interesting from</th>
<th>Interactions with</th>
<th>Cellular process</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lon</td>
<td>MS</td>
<td>Beta 1-177</td>
<td>λ Exo</td>
<td>ATP-dependent protease/chaperone</td>
</tr>
<tr>
<td>ClpP</td>
<td>MS, Functional association with Lon</td>
<td>λ Exo</td>
<td>ATP-dependent protease/chaperone</td>
<td>Remodels protein/DNA complexes in switch from recombination to Replication (Mu transposition)</td>
</tr>
<tr>
<td>ClpX</td>
<td>Interactome of ClpP, Functional association with Lon</td>
<td></td>
<td>ATP-dependent protease/chaperone</td>
<td>Remodels protein/DNA complexes in switch from recombination to Replication (Mu transposition)</td>
</tr>
<tr>
<td>HslU</td>
<td>Functional association with Lon</td>
<td></td>
<td>ATP-dependent protease/chaperone</td>
<td>Exoribonuclease regulator</td>
</tr>
<tr>
<td>HslV</td>
<td>Functional association with Lon</td>
<td></td>
<td>ATP-dependent protease/chaperone</td>
<td>Exoribonuclease regulator</td>
</tr>
<tr>
<td>SulA</td>
<td>Interactome of Lon (substrate of Lon and HslUV)</td>
<td></td>
<td></td>
<td>Cell division inhibitor, degraded by Lon, stabilized by phage lambda Kil protein</td>
</tr>
<tr>
<td>OmpF</td>
<td>MS</td>
<td>Beta 1-177</td>
<td>λ Exo 1-189</td>
<td>Osmotic pressure</td>
</tr>
</tbody>
</table>

Table 7.1 λ SynExo interacting partners identified by MS or by functional assays.
<table>
<thead>
<tr>
<th>Candidates</th>
<th>Interesting from</th>
<th>Interactions with</th>
<th>Cellular process</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>UvrD</td>
<td>MS, Interactome of OmpF</td>
<td>Beta 1-177</td>
<td>DNA metabolism</td>
<td>Helicase, replication restart factor (rec-dependent replication)</td>
</tr>
<tr>
<td>DnaT</td>
<td>Interactome of OmpF</td>
<td></td>
<td>DNA Replication/Repair/Recombination</td>
<td>Primosome assembly</td>
</tr>
<tr>
<td>PriA</td>
<td>Interactome of OmpF</td>
<td></td>
<td>Recombination-dependent DNA Replication</td>
<td>Primosome, Replication restart</td>
</tr>
<tr>
<td>PriB</td>
<td>Interactome of OmpF</td>
<td></td>
<td>Recombination-dependent DNA Replication</td>
<td>Primosome, Replication restart</td>
</tr>
<tr>
<td>PriC</td>
<td>Interactome of OmpF</td>
<td></td>
<td>Recombination-dependent DNA Replication</td>
<td>Primosome, Replication restart</td>
</tr>
<tr>
<td>DnaB</td>
<td>Interactome of OmpF</td>
<td></td>
<td>DNA Replication</td>
<td>Main replication helicase</td>
</tr>
<tr>
<td>DnaC</td>
<td>Interactome of OmpF</td>
<td></td>
<td>DNA Replication</td>
<td>DNAB helicase loader/unloader</td>
</tr>
<tr>
<td>YegS</td>
<td>MS</td>
<td>Beta 1-177 λ Exo</td>
<td>Lipid metabolism</td>
<td></td>
</tr>
<tr>
<td>RuvC</td>
<td>Interactome of YegS</td>
<td></td>
<td>DNA Replication restart/Repair/Recombination</td>
<td>HJ, reduces the frequency of lambda recombination</td>
</tr>
<tr>
<td>RuvA</td>
<td>Interactome of YegS</td>
<td></td>
<td>DNA Replication restart/Repair/Recombination</td>
<td>Holliday Junction helicase subunit</td>
</tr>
<tr>
<td>RuvB</td>
<td>Interactome of YegS</td>
<td></td>
<td>DNA Replication restart/Repair/Recombination</td>
<td>Holliday Junction helicase subunit</td>
</tr>
<tr>
<td>MviN</td>
<td>MS</td>
<td>Beta 1-177 λ Exo</td>
<td>“Virulence factor”</td>
<td></td>
</tr>
<tr>
<td>UvrC</td>
<td>Interactome of MviN</td>
<td></td>
<td>DNA Repair (NER)</td>
<td>Endonuclease, cuts DNA distortions</td>
</tr>
<tr>
<td>RecA</td>
<td>ReDa, λ Exo pull-downs, increases Recombineering frequencies</td>
<td>Beta λ Exo</td>
<td>DNA Replication/Repair/Recombination</td>
<td>Homologous DNA pairing, SOS inducer, repair factor co-protease</td>
</tr>
<tr>
<td>RecG</td>
<td>Conditionally inhibits Recombineering</td>
<td></td>
<td>DNA Repair/Recombination</td>
<td>Holliday Junction helicase with opposite polarity of RuvAB</td>
</tr>
<tr>
<td>Ssb</td>
<td>Binds ssDNA and Exol, ReDa assays</td>
<td>Beta</td>
<td>DNA Replication/Repair/Recombination</td>
<td>Main ssDNA binding protein</td>
</tr>
<tr>
<td>ExoI</td>
<td>ReDa, Beta and RecA prevent Exol from resecting nascent ssDNA</td>
<td>Beta λ Exo</td>
<td>DNA Repair/Recombination</td>
<td>3’ ssDNA exonuclease</td>
</tr>
<tr>
<td>MutS</td>
<td>Inhibits Recombineering, MMR</td>
<td></td>
<td>DNA Repair/Recombination</td>
<td>MMR</td>
</tr>
<tr>
<td>RecE</td>
<td>MS, Functional analog of λ Exo</td>
<td>Beta 1-177</td>
<td>DNA Recombination</td>
<td>RecET Recombineering system</td>
</tr>
<tr>
<td>LigA</td>
<td>λ Exo analog RecE binds Ligase</td>
<td></td>
<td>DNA Replication/Repair/Recombination</td>
<td>Replication</td>
</tr>
</tbody>
</table>
Coupling of recombination to replication via the primosome

Recombination is known to create substrates for origin-independent DNA replication (Asai et al. 1994; Kogoma et al. 1996), sometimes referred to as “replication restart” (Liu et al. 1999; Marians 2000; Sandler 2000). Recombineering is mechanistically linked to replication and replication is conditionally essential for phage lambda-mediated recombination, depending on the substrates and DNA targets (Stahl et al. 1997; Ellis et al. 2001; Li et al. 2003; Poteete 2008; Lim et al. 2008). Primosomes are nucleoprotein
assemblies that recruit the replicative helicase onto primer/template complexes at stalled or newly restored DNA replication forks. In *E. coli*, binding of PriA protein to branched DNA (as is found in many recombination intermediates) triggers primosome assembly. Six preprimosomal proteins, PriA, PriB, PriC, DnaB, DnaC, and DnaT bind at branched structures comprised of dsDNA adjacent to SSB coated single-stranded DNA. Unlike some phage replication origins, addition of the DnaG primase is not required at recombination intermediates as the 3’ hydroxyl terminus of the invading strand is already hybridized to its target, creating a primer/template context suitable for re-initiating DNA synthesis with the assistance of the primosome. In this way, DNA replication is physically and mechanistically linked to recombination. Mutants that disrupt the primosome disrupt HR and eliminate replication that is origin independent.

All four λ SynExo proteins bind OmpF, whose interactome includes primosome assembly factor DnaT, the PriABC primosome complex, the replicative helicase DnaB and the helicase loader DnaC (**Figure 7.7**). Beta (1-177) binds UvrD. The UvrD interactome also includes PriABC connecting the interaction network to SynExos through two nodes (**Figure 7.8**). UvrD is a Helicase that binds nick DNA and unwinds it. It interacts with the primosome and is required for origin independent replication. Recombination-dependent replication requires UvrD and recombination creates UvrD substrates. UvrD also interacts with MMR, which modulates Recombineering rates by destabilizing heteroduplex intermediates, and DNA ligase, which completes the Recombineering reaction by sealing ssDNA nicks.
Figure 7.7 λ SynExo indirect interactions with DnaT (PriABC, DnaBC) indicates initiation of recombination-dependent DNA replication by primosome assembly.

Interaction inferred through binding to OmpF (a membrane protein). Omp interacts with Beta, Beta 1-177, λ Exo and λ Exo 1-189. The primosome is involved in initiation of recombination-dependent DNA replication. Figure was created using STRING (http://string-db.org/), IntAct (http://www.ebi.ac.uk/intact/) and primary publications.

**Processing of recombination intermediates via resolvases**

Recombination sometimes creates branched DNA intermediates called Holliday Junctions (HJs). Completion of dsDNA Recombineering is partially dependent on RuvC resolvase (Poteete et al. 1999; Poteete & Fenton 2000), though this has not been tested for ssDNA oligos yet. Two λ SynExo proteins bind YegS, whose interactome includes HJ helicase RuvAB and RuvC HJ resolvase (Figure 7.9). Three of the four viral λ SynExo recombinases bind MviN, whose interactome includes UvrC, another potential HJ resolvase (Figure 7.10). UvrC creates a nicked DNA substrate for UvrD helicase,
UvrBC recruits UvrD to the nicked DNA and UvrD (Figure 7.8) subsequently unwinds the nicked DNA and displaces UvrC in preparation for completion of repair of the nicked DNA. Together these data suggest that λ SynExo-mediated Recombineering might produce branched DNA structures that require host proteins to complete recombination.

Figure 7.8 λ SynExo direct interactions with UvrD helicase (MutU) indicates initiation of recombination by primosome assembly and resolution by nucleotide excision repair and DNA unwinding. UvrD interacts with Beta(1-177). UvrD binds nick DNA and unwinds it. It interacts with the primosome and is required for origin independent replication. Figure was created using STRING (http://string-db.org/), IntAct (http://www.ebi.ac.uk/intact/) and primary publications.
295

Figure 7. 9  λ SynExo indirect interactions with RuvABC and LigA indicate resolution of recombination by Holliday junction cleavage and sealing of residual ssDNA nicks following resolution Interaction of RuvB inferred through binding to YegS (a lipid kinase). YegS interacts with Beta 1-177 and λ Exo. RuvAB helicase catalyzes branch migration to produce a Holliday junction that is cleaved by the Holliday junction endonuclease RuvC. Figure was created using STRING (http://string-db.org), IntAct (http://www.ebi.ac.uk/intact/) and primary publications.

Regulation/disassembly via proteases

The toroidal homotrimeric λ Exo complex becomes topologically linked to its DNA substrate during DNA processing, contributing to the high processivity of DNA resection (Subramanian et al. 2003; Zhang et al. 2011a). Other highly processive DNA metabolism proteins rely on accessory factors (e.g. DnaB helicase toroids are opened by ATP-bound DnaC to facilitate helicase loading and unloading on DNA (Davey et al. 2002)) or by intrinsic nucleotide-dependent modulation of inter-subunit contacts (e.g. MutS/ATP encircles DNA to greatly decrease the dissociation rate while MutS/ADP has a much shorter dwell time and releases DNA (Cho et al. 2012). No such activities are known for
λ Exo and while binding to dsDNA ends and subsequent DNA resection is the likely mechanism of threading the λ Exo onto DNA, removal of λ Exo from DNA following recombination remains an important outstanding issue. What removes λ Exo when recombination is complete? Furthermore, Beta protein binds more tightly to nascent dsDNA formed by Beta-mediated annealing of complementary ssDNA chains during recombination than it does to ssDNA substrates (pre-existing dsDNA is not bound well). This is a similar problem as stated above for λ Exo: What disassembles Beta/DNA complexes when pairing is complete?

![Diagram](image)

**Figure 7.** 10 λ SynExo indirect interactions with UvrC and RuvC indicate resolution of recombination by nucleotide excision repair and Holliday Junction cleavage, initiation of recombination by strand invasion via RecA, and DNA replication via the DnaN sliding clamp component of the replisome. Interaction of UvrC inferred through binding to MviN (a virulence factor). MviN interacts with Beta, Beta (1-177) and λ Exo. UvrC nicks DNA at either side of an intra-strand cross link. Figure was created using STRING ([http://string-db.org/](http://string-db.org/)), IntAct ([http://www.ebi.ac.uk/intact/](http://www.ebi.ac.uk/intact/)) and primary publications.

We identified ATP-dependent proteases Lon and ClpPX as λ SynExo interacting proteins with additional support from data mining. These proteases (or chaperones as in
the case of ClpX acting singly, (Levchenko et al. 1995)) are known to regulate important aspects of DNA metabolism and cell division (Figure 7. 11), including the transition from recombination to replication (Sasakawa et al. 1987; Levchenko et al. 1997; Servant et al. 2007; Abdelhakim et al. 2008; Ambro et al. 2012), suggesting that they could mediate the disassembly of DNA-recombinase complexes once homologous DNA pairing is complete.

**Figure 7. 11** λ SynExo directly interacts with Lon protease while indirect interaction with other proteases and chaperones indicate λ SynExo regulation or disassembly from DNA via proteases. Lon binds with Beta(1-177) and λ Exo. Lon functionally interacts with other ATP-dependent proteases such as ClpPX. And HSLUV and with molecular chaperones involved in complex assembly and disassembly. One of the targets of Lon, SulA is a cell division inhibitor and inhibition of cell division is a feature of induction of Recombineering in our *E. coli* system. Figure was created using STRING (http://string-db.org/), IntAct (http://www.ebi.ac.uk/intact/) and primary publications.

This provocative possibility could solve a major mystery in all recombination pathways (high nuclease processivity and tight DNA binding drives recombination but must stop by an unknown mechanism to complete DNA repair). Alternatively, interaction
with these chaperone-type activities might be an artifact of protein misfolding due to overexpression, so it will be important to measure protein half lives in the presence and absence of these proteases during steady-state expression and in the context of recombination.

**Reverse pull-down experiments to validate MS data**

To confirm interactions identified from cell extracts and MS, reverse pull-down experiments are on the way. Two strategies have been evaluated. In one, *E. coli* has been co-transfected with plasmids expressing the λ SynExo genes and with plasmids from the ASKA collection (Kitagawa et al. 2005) expressing putative interacting partners (Table 7.1). The protocol is in the process of optimization, but we have learned that saturating amount of both plasmids should be use for cotransfection (>100 ng of each), that expression of both proteins varies between clones, that induction should be performed in cultures larger than 30 ml for enough material for pull-downs, and that expression varies between clones so that multiple clones should be screened before pull-downs are performed. In the second strategy, each protein is purified (λ SynExos using the HaloTag and ASKA clones using nickel beads), mixed in different amounts to identify interacting proteins and to estimate their binding kinetics. One confounding issue is that most of the proteins in the SynExo Interactome are toxic when overexpressed, so we are using the bacterial two-hybrid system in parallel to hedge our bets.

**Genetic approaches to validate MS data**

In a complementary genetic approach, the functional significance of these putative host factors for viral recombination can be evaluated when the gene of interest is deleted (save *dnaB* and *dnaC* which are essential) or overexpressed (though see comment about
overexpression toxicity). Deletions (*hypomorphic alleles*) were created by phage P1 transduction from the KEIO collection (Baba *et al.* 2006) into Recombineering hosts. Gene overexpression (*hypermorphic alleles*) were studied using plasmids isolated from the ASKA collection (Kitagawa *et al.* 2005) and transformed into Recombineering hosts. To evaluate the consequence of hypomorphic and hypermorphic alleles on both ssDNA Recombineering and dsDNA Recombineering, a new reporter strain was created that coupled GFP fluorescence to acquisition of resistance to the antibiotic kanamycin. This reporter should allow us to quantify recombination rates from 0.000001% to 100%. Recombinants will acquire kanamycin resistance, which allows one to quantify low frequency recombination for the strains with genetic backgrounds that impair Recombineering. Recombinants also will acquire GFP fluorescence upon repair, which allows quantification with greater statistical support as Recombineering can be directly assessed independent of separate plating for total viable cells and selected recombinants. This unique recombination reporter system should provide a clear indication of the extent to which the λ SynExo Interactome contributes to viral recombination. A diagram of this reporter is shown in Figure 7. 12.
Recombineering reporter to evaluate hypomorphic and hypermorphic alleles of λ SynExo
Interactome.
Nonsense or frameshift alleles targeted by ssDNA Recombineering to the first half of the gfpmut3* gene in the E. coli
genome a tyrosine codon toggle cells from Green and kanamycin resistant to Dark and kanamycin sensitive. By using
another ssDNA oligo, Recombineering can restore Green fluorescence and kanamycin resistance simultaneously. The
reporter allows recombination efficiencies to be measured either without selection (1,000 fold range, excellent
statistics) or with selection (100,000,000 fold range, statistical fluctuation is an issue).

Functional interactions between host proteins and viral recombination
proteins

λ Exo activity was studied using the ReDA assay (Tolun & Myers 2003). Briefly,
linear dsDNA and PicoGreen®, a fluorescent dye that binds preferentially to dsDNA,
were mixed with λ Exo. After sufficient time for λ Exo to bind to the ends of the dsDNA,
DNA digestion was initiated by addition of Mg²⁺. DNA digestion proceeds 5′ → 3′,
producing 3′ ssDNA overhangs and free dNMPs. During digestion, PicoGreen® is
displaced and the fluorescence of the sample decreases. The fluorescence emission was
recorded in a PTI QuantumMaster spectrofluorometer by integrating one second samples...
representing typically $10^4$ - $10^6$ photons emitted in each second. Samples were recorded for 600-1500 seconds, depending on the substrate length and the time to achieve complete digestion of the sample, and the fluorescence output was converted to nucleotides released as described in Materials and Methods. Data were plotted as a function of reaction time and are shown in the following figures. In these time course experiments, DNA digestion by $\lambda$ Exo revealed a multi-exponential progress curve that indicated fast initiation with a subsequent a decay in activity towards the end of the timecourse (Figure 7.13).

\begin{figure}[h]
\centering
\includegraphics[width=0.8\textwidth]{figure7_13.png}
\caption{ReDa assay: $\lambda$ Exo resects dsDNA 5' $\rightarrow$ 3' producing 3' ssDNA overhangs and free dNMPs.}
\end{figure}

$\lambda$ Exo activity is inhibited by nascent ssDNA

When we examined $\lambda$ Exo progress curves, two distinct regions were evident Figure 7.13. First there was a linear rate and then there was a downward deflection. The extent of digestion was not 100% of the substrate. The downward deflection suggests that the enzyme had become inactivated or that the substrate was no longer digestible. Enzyme
inactivation was ruled out, as addition of fresh substrate restored digestion to the same initial rate (Tolun & Myers 2003). One could also argue that the failure to completely digest the substrate was not due to the limit of processivity because when we added Heparin, a trap for λ Exo, the reaction stopped earlier (Figure 7.14). We suspect that something has rendered the DNA undigestible as addition of ExoI, a 3’ → 5’ ssDNA exonuclease, at sufficient concentration to exceed the rate of production of nascent 3’ ssDNA by λ Exo, removed this limitation and permitted complete digestion of the substrate (Figure 7.14). From these observations we concluded that ssDNA is somehow inhibiting λ Exo function.

Figure 7.14 Nascent ssDNA inhibit λ Exo activity but increases λ Exo dwell time.

**ssDNA annealing proteins overcome the inhibition of λ Exo by ssDNA**

If a ssDNA structure is inhibiting λ Exo function, we reasoned that ssDNA binding proteins might relieve this inhibition by melting this structure. Previous results from (Tolun 2008) showed that Red Beta protein might have this effect on λ Exo but that *E.
coli ssDNA binding protein (SSB) did not. The effects of Beta were re-examined as well as the influence of a host homologous recombination protein, RecA, known to cooperate with the Red system in recombination. To test if ssDNA annealing proteins overcame the inhibition of λ Exo by ssDNA, we added these proteins to λ Exo reactions and looked at the effect on the extent of digestion by λ Exo on the substrate (Figure 7.15). The extents of digestion when these ssDNA annealing proteins were in the reaction were closer to complete digestion of the substrate than when λ Exo was alone (λ Exo digested 83% of the substrate, while λ Exo + Beta digested 92% and λ Exo + Rec A digested 93% of the substrate). From these results, and in light of previous observations (Tolun 2008), we concluded that ssDNA annealing proteins but not SSB appear to remove the obstruction to λ Exo activity that ssDNA introduces.

![Figure 7.15](image-url)

**Figure 7.15** ssDNA annealing proteins overcome the inhibition of λ Exo by ssDNA.

**ssDNA increases λ Exo processivity but blocks reinitiation of DNA digestion by λ Exo upon dissociation from partially degraded DNA**

Based on the data in Figure 7.14, λ Exo dissociated from partially digested DNA but could re-associate and continue digestion in the absence of a Heparin trap. We wondered
if the ssDNA feature described above inhibited λ Exo nuclease activity by decreasing its processivity. We measured processivity by conducting λ Exo reactions in the presence and absence of ssDNA annealing proteins in the presence of Heparin (Figure 7.16). The values of processivity for each of these protein mixes expressed as the fraction of the substrate digested in a single round of complex formation were λ Exo: 53 %, λ Exo+Beta: 89 % and λ Exo+RecA: 81 %. From these data we concluded that ssDNA annealing proteins increased the processivity of λ Exo, possibly by removing the inhibitory effect of ssDNA on λ Exo.

![Complete digestion graph](image)

**Figure 7.16** ssDNA increases λ Exo processivity but blocks reinitiation of DNA digestion by λ Exo upon dissociation from partially degraded DNA.

The role of ssDNA in altering λ Exo activity is complex. Tolun (2008) reported that ssDNA played an important role in λ Exo processivity by acting as a tether that stabilized λ Exo/DNA complexes. Our results confirmed this observation (Figure 7.16). It was apparent that if one removed the nascent 3’ ssDNA with ExoI in the presence of Heparin,
the number of nucleotides released is much less, consistent with a more distributive mode of λ Exo action. Another former student on the lab also saw a similar effect of removing phosphate from the 5’ end of dsDNA substrates or by introducing a mutation that weakened binding of λ Exo to the 5’-phosphoryl terminus of dsDNA (Subramanian et al. 2003). We could not rule out that RecA and Beta stimulated processivity by improving 5’-phosphoryl end binding during each catalytic cycle. Indeed, the current model for λ Exo activity is that it translocates by a combination of Brownian motion and 5’ phosphoryl end binding, making it a “Brownian Ratchet” (Subramanian et al. 2003). If these ssDNA annealing proteins stimulated forward translocation or end binding, or improved melting of the 5’ strand to allow positioning in the active site (van Oijen et al. 2003), we might also have obtained these results. Additional pre-steady-state experiments would be necessary to test these competing models.

While we do not yet understand the mechanism, these results are the first demonstration that homologous DNA pairing proteins stimulate the processivity of a recombination exonuclease. These data are consistent with the previously observed increase in the lifetime of λ Exo/DNA complexes (Tolun 2008).

In conclusion, we cannot simply treat nascent ssDNA as an inhibitor of λ Exo function because, as shown in Figure 7.16, removal of ssDNA by ExoI drastically decreases the processivity of λ Exo, consistent with the view that the toroidal structure of λ Exo forms a topological link with nascent ssDNA generated by dsDNA resection by λ Exo (Kovall & Matthews 1997). This result stands in stark contrast with the effect of ExoI on λ Exo evident in Figure 7.14. Therefore we conclude that ssDNA both blocks reassociation of λ Exo with partially resected substrates and also improves processivity
by decreasing dissociation of λ Exo from partially resected substrates. This provides a mechanism, then, for regulating λ Exo activity during homologous recombination: Blunt or nearly blunt dsDNA breaks are readily bound and resected by λ Exo and partially resected substrates permit extensive digestion by λ Exo until such a time as λ Exo dissociates OR IS REMOVED from DNA in a feedback regulatory mechanism. Removal of λ Exo would lead to a state where λ Exo has difficulty reassociating with the substrate, as previously shown in Tolun’s work. The previously mentioned ATP-dependent proteases and chaperones (Lon, ClpPX, HslUV) could remove λ Exo and/or Beta protein in vivo.

**Host RecA blocks ExoI from resecting nascent ssDNA produced by λ Exo and increases λ Exo processivity by increasing dwell time**

In the data plotted in Figure 7.15, RecA stimulated λ Exo activity since the nucleotide release increased when RecA was present in the reaction. In Figure 7.16, RecA increased processivity in the presence of the λ Exo trap, Heparin, indicating that RecA might have stimulated processivity by increasing λ Exo dwell time. When these reactions were then incubated with ExoI, the stimulation by RecA was reduced, but some nascent ssDNA remained protected from digestion by the host exonuclease ExoI (Figure 7.17).
RecA increases processivity and dwell time and protects nascent ssDNA recombination substrate from ExoI resection.

**λ SynExo Beta blocks ExoI from resecting nascent ssDNA and increases processivity by increasing dwell time**

In the data plotted in **Figure 7. 15**, Beta appeared to stimulate λ Exo, since the nucleotide release increased when Beta was present in the reaction. In **Figure 7. 16**, Beta also evidently increased λ Exo processivity in the presence of Heparin, indicating that Beta might have stimulated processivity by increasing λ Exo dwell time. When these reactions were then incubated with ExoI, ExoI had no effect on Beta-stimulation of λ Exo, as if Beta protected the nascent ssDNA substrate from the host exonuclease ExoI (Figure 7. 18). The effects of Beta on λ Exo processivity and its ability to overcome inhibition of processivity by ExoI is very clearly demonstrated in **Figure 7. 19**.
To measure processivity, we added Heparin 100 seconds after the reaction was being recorded and 80 seconds after the reaction was started with MgCl$_2$. The arbitrary fluorescence units were converted to nucleotides released and normalized after the signal
settled down after injection of Heparin (this occurs at 111 seconds). Processivity was estimated by taking the average of the signal over 50-100 seconds at the end of the progress curve. The post-Heparin progress curves look like they are dominated by a simple single exponential process, probably dissociation of λ Exo from the DNA with subsequent trapping on the Heparin, which keeps the data simpler than in the multi-turnover condition in the absence of Heparin. The post-heparin progress curve was fit to a single exponential function; the value for the plateau (the asymptote that the curve approaches) defines processivity:

**Equation 7. 1 Processivity.**
This is an exponential rate equation corrected for background. $P_0$ is background, $k$ is the rate of nucleotide release and $t$ is the elapsed time.

$$P = P_0 + (\text{Plateau} - P_0)(1 - e^{-kt})$$

The rate was calculated as the first derivative of the post-Heparin progress curve, and dwell time was calculated from:

**Equation 7. 2 Dwell Time.**

$$\text{Dwell Time (s)} = \frac{\text{Processivity (nts released)}}{\text{rate (nts released/second)}}$$

In calculating the dwell time, we found that when Beta was in the reaction, dwell time increased by 30 % (see table below). This indicated that the increase in processivity observed when Beta was in the reaction may have been due to Beta having a positive effect on maintaining the λ Exo-DNA complex in the active form. Alternatively, Beta filaments might enforce an allosteric effect on the enzyme-substrate complex or might prevent λ Exo from falling from the DNA.
Table 7.2 Beta increases processivity by increasing dwell time.
All the reactions incubated with Heparin. The number of reaction is indicated in parenthesis and correspond to the number in the figures. Rate processivity and dwell time were normalized to $\lambda$ Exo values.

<table>
<thead>
<tr>
<th>Proteins</th>
<th>Rate (nt/s)</th>
<th>Processivity (nt)</th>
<th>Dwell Time (s)</th>
<th>Relative Rate</th>
<th>Relative Processivity</th>
<th>Relative Dwell Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda$ Exo (16)</td>
<td>4.4</td>
<td>740</td>
<td>170</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$\lambda$ Exo + Beta (18)</td>
<td>4.4</td>
<td>1000</td>
<td>230</td>
<td>1.0</td>
<td>1.3</td>
<td>1.3</td>
</tr>
<tr>
<td>$\lambda$ Exo + ExoI (19)</td>
<td>2.9</td>
<td>510</td>
<td>180</td>
<td>0.7</td>
<td>0.7</td>
<td>1.0</td>
</tr>
<tr>
<td>$\lambda$ Exo + Beta + ExoI (21)</td>
<td>4.6</td>
<td>1200</td>
<td>250</td>
<td>1.1</td>
<td>1.6</td>
<td>1.5</td>
</tr>
</tbody>
</table>

**Lags in progress curves**

When Beta or RecA or ExoI were added to $\lambda$ Exo reactions, a lag before the linear phase of the reaction was obtained (see many of the preceding figures). This may be an artifact of the assay system being used or may indicate interesting information about the enzyme-catalyzed reaction itself. Here are some things we considered:

**Inadequate temperature control.** Magnesium and Heparin solutions were used at room temperature as was the reaction mix in the cuvette. Only the enzymes were sitting on ice. The volume of enzymes varied depending of the reaction mix in the range of 1 to 40 µl when several enzymes were mixed. Since the total volume was 2 ml and the reaction was well mixed before the start, and since the magnitude of the lags were very reproducible through all the different mixture volumes, we don’t think this could be a cause for the lag observed when other proteins were added to the reaction.

**Pre-steady-state transients.** A lag phase in the time-course of product formation could be due to the time taken for the concentrations of the intermediate enzyme-substrate and enzyme-product complexes to rise to their steady-state levels. However, we were evaluating the enzyme as the E-S complex itself, treating DNA as part of the enzyme, and product release is not DNA release, it is release of the 5’ phosphoryl terminus and the need to wiggle forward to find it again and re-establish the active E-S complex. So: $E + S \rightarrow ES$ inactive $\rightarrow ES$ active $\rightarrow ESP \rightarrow ES$(inactive) $\rightarrow P \rightarrow ES$
active etc. We propose that RecA and Beta promote ES (active) formation, perhaps by restricting reverse translocation of \( \lambda \) Exo from the 5’ phosphoryl terminus. The initial rate is not stimulated because filament formation on nascent ssDNA produced by \( \lambda \) Exo is required. In this case we are inclined to think that when complexes of \( \lambda \) Exo-RecA or \( \lambda \) Exo-Beta form, substrate binding or initial hydrolysis becomes rate limiting. Once this is formed, then RecA or Beta may increase the processivity of the enzyme by stabilizing the complex.

**Slow Magnesium binding.** The reactions were initiated with magnesium. If any of the proteins decreased the rate of magnesium association with \( \lambda \) Exo, there could be a lag in the progress curve. In the simplest model, the rate would be less than \( \lambda \) Exo alone until \( \lambda \) Exo was at equilibrium with the metal (magnesium is catalytic, not consumed). MgCl\(_2\) concentration in the reactions was 5 mM. The table below shows protein concentrations and the percent of magnesium that each would bind. The sum of the amount needed when we mixed these proteins was less than 0.1 % of the total amount of magnesium, therefore we don’t think that magnesium concentration was limiting in the reaction. Nevertheless this hypothesis is testable if we fit an equation to additional data at different magnesium concentrations.

<table>
<thead>
<tr>
<th>Proteins</th>
<th>Protein (nM)</th>
<th>Mg(^{2+}) bound by each protein (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \lambda ) Exo</td>
<td>13</td>
<td>0.0003</td>
</tr>
<tr>
<td>Beta</td>
<td>130</td>
<td>0.0026</td>
</tr>
<tr>
<td>RecA</td>
<td>624</td>
<td>0.0125</td>
</tr>
<tr>
<td>Exol</td>
<td>14.6</td>
<td>0.0003</td>
</tr>
</tbody>
</table>

**Hysteresis.** This refers to allosteric effects caused by slow structural isomerization of an enzyme. Frieden (1979) believed this to be an important feature of enzyme activity.
regulation. One way to incorporate hysteresis into this model would be to acknowledge that Beta and RecA polymerize to form filaments and to propose that the filaments alter \( \lambda \) Exo activity in ways different than the nonpolymerized Beta and RecA proteins. The thing is, hysteresis usually relates to ENZYME isomerization, not AFFECTOR isomerization. But if we consider these COMPLEXES to be the ENZYME, then there is a meaningful overlap. This line of thinking could be developed further, if one were so-inclined.

**Conclusions**

We confirmed that Beta binds \( \lambda \) Exo to form a complex and saw that Beta does not bind SSB or RecA. Some evidence was obtained suggesting that RecA interferes with \( \lambda \) Exo/Beta interaction, possibly through interaction with \( \lambda \) Exo, which was consistent with functional association of RecA and \( \lambda \) Exo through increased processivity and protection of the nascently formed ssDNA produced by resection of dsDNA by \( \lambda \) Exo. The mass spec studies confirmed that Beta and \( \lambda \) Exo formed a complex and that the Ct terminus domain of Beta was required for binding to \( \lambda \) Exo but that the Ct domain of \( \lambda \) Exo was not required to bind to Beta. The interacting host proteins suggest that \( \lambda \) SynExo activity could be regulated by or disassembled from DNA complexes via protease or chaperones, that Recombineering could be coupled to replication via facilitated assembly of the primosome on recombination intermediates, and that the resolution of Recombineering intermediates might involve Holliday Junction resolvases. While we cannot yet rule out expression, folding, stability, or localization issues, our studies and those of others point to allele-specific interactions between viral and host proteins that coordinate viral recombination with host metabolism.
We confirmed previous observations that Beta alters λ Exo activity and protected nascent ssDNA products of λ Exo resection from an *E. coli* ssDNA exonuclease, ExoI. We discovered that Beta improved λ Exo processivity, apparently by increasing the dwell time of λ Exo on DNA. The host *E. coli* recombinase RecA also modulated λ Exo activity by increasing processivity and dwell time and modestly protecting nascent ssDNA from ExoI.

These results are consistent with a model for coupling recombination protein activities in which λ Exo makes direct contacts with two homologous DNA pairing proteins, the λ SynExo synaptase (Beta) and a host recombinase (RecA). Both proteins form a stable nucleoprotein filament complex that holds λ Exo onto DNA during resection. Both the λ SynExo Beta synaptase and the host RecA protect the nascent ssDNA recombination substrate from exonucleases, which preserves fragile recombination intermediates. The host protein RecA competes with Beta to bind to λ Exo, perhaps displacing Beta on the emerging nucleoprotein filament to promote homologous DNA pairing. This “hand off” reaction is reminiscent of what has been proposed for bacterial, yeast and human recombination (Beernink & Morrical 1999; Maher *et al.* 2011).

Comparison of the interaction maps for the enterobacteria phage λ SynExo and the Human Herpes virus type 1 SynExo shows significant parallels (*Table 7. 4*).
Table 7. SynExo from λ phage and HHV1 show significant protein interaction parallels.

### Virus-Virus protein interactions

<table>
<thead>
<tr>
<th>Function</th>
<th>Herpes SynExo</th>
<th>Lambda SynExo</th>
<th>Herpes Interaction</th>
<th>Lambda Interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recombination</td>
<td>ICP8</td>
<td>UL12</td>
<td>ICP8, UL12</td>
<td>UL12, ICP8</td>
</tr>
<tr>
<td></td>
<td>Beta, λ Exo</td>
<td>λ Exo, Beta</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Replication</td>
<td>ICP8</td>
<td>UL5, UL8, UL42</td>
<td>Beta</td>
<td>P</td>
</tr>
<tr>
<td>Transcription</td>
<td>ICP8</td>
<td>ICP27</td>
<td>Beta</td>
<td>N</td>
</tr>
</tbody>
</table>

### Virus-Host protein interactions

<table>
<thead>
<tr>
<th>Function</th>
<th>Herpes SynExo</th>
<th>Lambda SynExo</th>
<th>Herpes Interaction</th>
<th>Lambda Interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recombination</td>
<td>ICP8</td>
<td>UL12</td>
<td>Rad50, Mre11, DNA-PKcs, Ku70, Ku86, XRCC4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Beta</td>
<td>λ Exo</td>
<td>RuvA, RuvB, RuvC, RecA, RecG, RecE</td>
<td></td>
</tr>
<tr>
<td>NHEJ</td>
<td>ICP8</td>
<td>UL12</td>
<td>Rad50, Mre11, Bu1</td>
<td></td>
</tr>
<tr>
<td>Repair</td>
<td>ICP8</td>
<td>UL12</td>
<td>hMsh2, hMsh3, hMsh6, PARP-1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Beta</td>
<td>λ Exo</td>
<td>UvrA, UvrB, UvrC,</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>UvrD</td>
<td></td>
</tr>
<tr>
<td>Disruption of Protein-DNA complexes</td>
<td>ICP8</td>
<td>UL12</td>
<td>hSNF2L, hSNF2H, BRM, BRG1, BAF57, BAF155, BAF170</td>
<td></td>
</tr>
<tr>
<td>Replication</td>
<td>ICP8</td>
<td></td>
<td>PCNA, Mcm2, RPA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Beta</td>
<td>λ Exo</td>
<td>DnaT, PriA, PriB, PriC, DnaB, DnaC, LigA</td>
<td></td>
</tr>
<tr>
<td>Transcription</td>
<td>ICP8</td>
<td></td>
<td>DEAD box p68</td>
<td></td>
</tr>
<tr>
<td>Cell cycle</td>
<td>ICP8</td>
<td></td>
<td>Beta</td>
<td></td>
</tr>
<tr>
<td></td>
<td>λ Exo</td>
<td></td>
<td>SulA</td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER 8. CONCLUSIONS AND FUTURE STUDIES

Conclusions

Fluorescent protein genes are sensitive and quantitative reporters for gene targeting. We developed a protocol for altering the spectra of fluorescent protein genes by Recombineering or site directed mutagenesis at the chromophore region. Several new fluorescent proteins were created. Mostaza, Azure and Celeste were created from the eGFP gene and Bronze, Amarillo, Violeta, Aqua and Mar were created from the GFPmut3* gene.

Fluorescent protein engineering with oligos in bacteria corroborated strand bias of Recombineering, where the synaptase Beta catalyzed the annealing of oligos targeting the lagging strand template at higher frequencies than oligos targeting the leading strand template. Changing color of fluorescent proteins involving different heteroduplex intermediates corroborated that MMR is a potent inhibitor of Recombineering. Different mismatched oligos were rejected from Recombineering intermediates with different frequencies while C.C mismatches escaped MMR. When MMR was avoided, single nucleotide changes were the most efficient substrates and longer nucleotide changes were less efficient. Finally, if the heteroduplex was not repaired, the strand that incorporated the oligo segregated during cell division to produce a mixed population of recombinant and non-recombinant cells. Recombinant segregation reduced the frequency of recombinant clones to about 1/4 - 1/8 of the initial frequency (as detected in “sectored” colonies of mixed clones) in experiments performed in exponentially growing E. coli.

When all these parameters were optimized, the efficiency of in vivo fluorescent protein engineering by Recombineering in bacteria could reach ~ 40 % of transformed cells.
Fluorescent protein engineering with oligos in mammalian cells was used to optimize gene targeting by unmodified ssDNA oligos. Guided by the bacterial studies, gene targeting was optimized to exceed the highest published rates using endogenous recombination functions. ICP8-dependent Recombineering were quantified by changes in phenotype and confirmed by showing that the recombinant phenotype corresponded to the predicted change in genotype. However, ICP8 expression was toxic to cells and caused loss of target gene expression. Human-optimized phage lambda Beta protein did not stimulate gene targeting, indicating that Recombineering is host specific. The stimulation of gene conversion by ICP8 in human cells was weaker than the optimized Recombineering rates we obtained in bacteria with Beta protein so there remains room for further improvement.

Nuclear oligo delivery was achieved using Lipofectamine 2000. Oligo transfection peaked at 90% when the final concentration of oligo/Lipofectamine 2000 complexes was 0.4% in growth medium. However, cell viability was really compromised by this condition. The best ratio of viability and oligo transfection efficiency (87%) was reached using a final concentration of 0.27% lipofectamine 2000. Oligos were found to not affect viability and were saturating for transfection by 200 nM.

Recombinase expression was confirmed in mammalian cells. Both ICP8 and the humanized Beta protein (HumBeta) were well-expressed and localized to the nucleus. They both produced cells with homogenous nuclear distribution of the proteins, but some ICP8 expressing cells show speckled or punctate subnuclear structures previously associated with replication compartments while HumBeta did not form speckles, consistent with HumBeta failing to coordinate its activity with host proteins, unlike ICP8.
Some HumBeta expressing cells also showed more protein at the nuclear periphery, unlike ICP8.

Longer oligos produced more recombinants both in the presence and absence of ICP8 with a significant steady increase between oligos 35 to 65 nucleotides long. It is worth noticing that Recombineering in *E. coli* is more efficient with oligos ~ 60 nucleotides long (Sawitzke *et al.* 2011). The similarity in oligo length might be related to the ability of the synaptases to form stable nucleoprotein filaments with oligos of this size.

Oligos hybridizing to the lagging strand template also produced significantly more recombinants than the oligos hybridizing to the leading strand template. This observation supports the strand bias seen in *E. coli* Recombineering, where the phage SynExo synaptase hybridizes oligos between Okazaki fragments.

There was a steady decrease in the frequency of ICP8-mediated human cell recombinants over time as described by other groups using endogenous functions. The decrease is likely due in part to chromosome segregation as was seen in bacteria, but it could also be due to cell cycle arrest or senescence. Nevertheless, gene conversion was validated by sequencing and allele-specific PCR, demonstrating that the change in the phenotype in recombinant cells corresponded to the change in the genotype.

ICP8 transient transfection and expression was toxic to cells, producing an increase in cell death and cell detachment from plates. In some experiments, ICP8 expression produced an increase of cells detected in the green channel of the flow cytometer, mimicking the recombination read out. We also observed an ICP8-dependent loss of reporter target gene expression. The simplest explanation for oligo-independent but ICP8-dependent green fluorescence is that ICP8 overexpression changes cell physiology.
somehow to increase background autofluorescence to levels that draw these cells into the gates used in the Accuri cytometer to capture the recombinant population. The simplest explanation for loss of Mostaza is that ICP8 induces loop out of the Mostaza transgene through crossing-over the direct repeats LTRs flanking the lentiviral transgene as has been seen in lentiviral plasmids propagated in *E. coli* expressing the phage lambda SynExo. Another possibility is that ICP8 expression increases the rate of random mutations, possibly by interfering with human RPA functions known to be required for normal DNA metabolism and genome maintenance.

Gene targeting in a HeLa Progeria reporter cell line suggested that oligo-dependent Recombineering was successful. A heat shock experiment produced three times as many putative recombinants as did the no-DNA control. A second experiment without heat shock produced lower rates of gene conversion and no stimulation by ICP8. Oligo alone produced a two-fold increase in the putative recombinants at 21 % O₂ tension and 0.5-fold at 3 % O₂ tension. Genotypic data suggest that the Progerin splicing region has been deleted in the recombinant clones.

A human codon optimized version of the phage lambda Beta protein fused to a nuclear localization signal (HumBeta) was expressed in human cells at high levels and efficiently targeted to the nucleus. Even so, HumBeta did not stimulate gene conversion above levels catalyzed by endogenous cellular functions. These results are consistent with the hypothesis that Recombineering is host specific due to failure to coordinate viral recombinases with host proteins. Consistent with this interpretation, HumBeta did not localize to replication centers in the human nucleus while the human viral synaptase ICP8 did so.
Additional evidence for host-specific protein interactions were obtained using proteomics and \textit{in vitro} functional assays. We confirmed that Beta binds \( \lambda \) Exo to form a complex and saw that Beta does not bind to host proteins SSB or RecA. Some evidence was obtained suggesting that RecA interferes with \( \lambda \) Exo/Beta interaction, possibly through interaction with \( \lambda \) Exo, which was consistent with functional association of RecA and \( \lambda \) Exo through increased processivity and protection of the nascently formed ssDNA produced by resection of dsDNA by \( \lambda \) Exo. Mass spectrometry studies of complexes formed between phage lambda SynExo proteins and bacterial proteins confirmed that Beta and \( \lambda \) Exo formed a complex and that the C terminal domain of Beta was required for binding to \( \lambda \) Exo but that the C terminal domain of \( \lambda \) Exo was not required to bind to Beta. The list of interacting host proteins suggest that \( \lambda \) SynExo activity could be regulated by, or disassembled from DNA complexes, via protease or chaperone activities. This observation is worth pursuing as it could answer some long-standing questions about termination of recombination and disassembly of highly processive enzyme complexes. The results also suggest that Recombineering is coupled to replication via facilitated assembly of the primosome on recombination intermediates. This is consistent with observations from other bacterial recombination pathways and with the protein interaction network for the Herpes SynExo in human cells. Finally, some evidence was obtained suggesting that resolution of Recombineering intermediates might involve Holliday Junction resolvases as has been seen in other bacterial recombination pathways. While we cannot yet rule out artifacts of protein overexpression, folding, stability, or localization issues, our studies and those of others point to allele-specific interactions between viral and host proteins that coordinate viral recombination with host metabolism.
This encourages the development of host-specific SynExo recombinases for genome engineering in other organisms (such as plants).

We confirmed previous observations that Beta altered λ Exo activity and protected nascent ssDNA products of λ Exo resection from an E. coli ssDNA exonuclease, ExoI. We discovered that Beta improved λ Exo processivity, apparently by increasing the dwell time of λ Exo on DNA. The host E. coli recombinase RecA also modulated λ Exo activity by increasing processivity and dwell time and modestly protected nascent ssDNA from ExoI.

These results are consistent with a model for coupling recombination protein activities in which λ Exo makes direct contacts with two homologous DNA pairing proteins, the λ SynExo synaptase (Beta) and a host recombinase (RecA). Both proteins form a stable nucleoprotein filament complex that holds λ Exo onto DNA during resection. Both the λ SynExo Beta synaptase and the host RecA protein protect nascently exposed ssDNA from exonucleases, which preserves fragile recombination intermediates. The host protein RecA competes with Beta to bind to λ Exo, perhaps displacing Beta on the emerging nucleoprotein filament to promote recombination via a strand invasion mechanism using homologous dsDNA and concomitantly inhibiting ssDNA annealing pathways that are inherently slower and less conservative. This “hand off” reaction is reminiscent of what has been proposed for the function of “mediator” proteins in bacterial, yeast and human recombination (Beernink & Morrical 1999; Maher et al. 2011).

The results of protein interaction studies suggest that viral SynExo-mediated Recombineering is directly coordinated with host DNA metabolism. Comparison of the
interaction maps for the enterobacteria phage λ SynExo and the Human Herpes virus type 1 SynExo shows significant parallels. Viral SynExo recombinases are host specific but not limited to bacteria. Organism-specific viral synaptases can be pressed into service to engineer human genomes for creation of disease models or for autologous gene therapy. Additional improvements are anticipated if the mechanism of coordination of host and viral proteins is further illuminated.

**Future directions**

**Human Recombineering**

Evaluate ICP8 specificity

ICP8 expression was correlated with a decrease of expression of the target transgene. Some hypotheses included that ICP8 promoted loop out of the transgene by crossing-over of direct repeated LTR sequences flanking the insert, or that ICP8 promoted random mutagenesis in some way or that ICP8 promoted gene silencing. We suggest to future genotyping of 293T, 293, HeLa and MIAMI cell DNA target regions upon ICP8 expression. ICP8 expression was also correlated with an increase of green fluorescence. Some hypotheses for the green fluorescence were that ICP8 induced cell physiology changes conducive to autofluorescence or mis-scoring by the flow cytometer gates or that ICP8 induced global mutagenesis rates as if it was affecting RPA functions. We suggest evaluating if ICP8 induces the green cell phenotype in cells that do not express the fluorescent reporter genes like Mostaza.
Role of T antigen in stimulating Recombineering due to activation of the SV40 origin of replication

As mentioned above, strand-specificity in Recombineering is primarily a property of the direction of replication fork travel across the target gene. We previously proposed that the consistent difference in recombination efficiencies observed was due to replication arising from the SV40 origin of replication which is located upstream of the Mostaza gene in the Recombineering reporter construct. This origin is predicted to be activated by the temperature sensitive allele of the SV40 large T antigen expressed in 293T cells at 33 °C, but not at 40 °C (May et al. 2005; DuBridge et al. 1987). We recommend exploring Recombineering at both temperatures and also compare Recombineering in HEK293 to 293T cells to see if strand specificity is abrogated.

Role of the cell cycle in human Recombineering

Homologous Recombination functions are active mainly during S phase while competing nonhomologous end joining mechanisms predominate in G1. We recommend expressing the viral recombinase and then delivering oligos to cells during early S phase, using either mechanical separation of cells, serum starvation or a double thymidine pulse to obtain synchronized cells.

Role of host factors such as Rad52 on human Recombineering

293T cells are proficient to do oligo gene targeting that is independent of ICP8 function. A recent study (Schumacher et al. 2012) reported that Rad52 was mainly responsible for single-strand annealing (SSA) mediated repair of dsDNA breaks even in the presence of ICP8. However that study also showed that knocking down Rad52 made ICP8 necessary for HR via SSA. We propose to validate if Rad52 is responsible for the
ICP8-independent recombination we are seeing by knocking down Rad52 in our reporter cell lines.

Evaluate Recombineering in embryonic and adult stem cells

Current assays have been developed in 293T because these cells are easy to transfect and to propagate. This cell line is defective in MMR (Cannavo et al. 2007) and this might contribute to the higher rates of oligo targeting. 293T is an immortal cell line with an aberrant number of chromosomes (http://en.wikipedia.org/wiki/HEK_cell). The multiple copies of chromosomes could make this cell line prone to off target recombination and produce gross genome rearrangements. While the potential of engineering 293T for genetic studies is great, even better would be to engineer primary cells, especially stem cells. Adult stem cells show great potential for autologous regenerative medicine. Autologous gene therapy using patient-derived stem cells could revolutionize medicine. The duplication time of these cells is slow, but we could coordinate Recombineering to the cell cycle via synchronization to improve the process. Embryonic stem cells have the biggest initial potential for Recombineering because they divide so rapidly and can be used to create many disease models.

Evaluate Recombineering of disease-causing alleles

Our studies currently employ fluorescent protein reporter genes to facilitate quantification of Recombineering to assist optimization of our protocols. To generalize this work, we suggest to target monogenic disease mutations, such as mutations in LMNA that cause an accelerated aging disease called HGPS or to modify the receptor and co-receptors responsible for HIV infection. Additionally, if ICP8 is found to be
looping out direct repeated sequences, we suggest evaluating ICP8 loop out of the extra repeated sequences within the *Huntingtin* gene.

**Improve Recombineering protocol**

ICP8 expression was found toxic to cells, we suggest evaluating human Recombineering with ICP8 fused to an inducible destabilizing domain. ICP8 expression from stable pSLIK constructions did not reach the gene targeting rates seen when ICP8 was expressed transiently. Tolun *et al.* (2013) suggested that both Nt and Ct of ICP8 are involved in the ultrastructure configuration for annealing of ssDNA. It is possible that the small tags of the P2 peptide could affect the protein-protein interactions of ICP8 and therefore, disrupt the annealing catalysis. We recommend developing a pSLIK-ICP8 without Nt or Ct modifications.

Gene targeting in pSLIK cell lines was performed with subsaturating oligo transfection efficiency. During transient transfection of pCMV-ICP8, Recombineering was only improved when oligo concentrations were increased from 50 to 200 nM, while oligos between 200 to 400 nM had no extra effect on endogenous gene targeting. It is possible that transient transfection of pCMV-ICP8 and oligos increased the innate immune response and compromised cell viability. Therefore, we re-evaluating Recombineering efficiency by titrating oligo delivery in inducible ICP8-expressing pSLIK cell lines comparing:

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Oligo (nM)</th>
<th>Lipofectamine 2000 (%)</th>
<th>FL-oligo transfection efficiency in 293T(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Known to work Recombineering (+ control)</td>
<td>4.13</td>
<td>0.13</td>
<td>≤ 61</td>
</tr>
<tr>
<td>FL-oligo titration: best ratio of viability and efficiency</td>
<td>200</td>
<td>0.27</td>
<td>87.3</td>
</tr>
<tr>
<td>Highest transformation efficiency</td>
<td>200</td>
<td>0.40</td>
<td>91.6</td>
</tr>
</tbody>
</table>
Doxycycline titrations performed for 48 hours in pSLIK4 cells showed a positive correlation between doxycycline concentration and the number of oligo-independent green cells and nonfluorescent cells that lost the transgene expression. We recommend validating if the same effect occurs for pSLIK1. ICP8 dependent green cells interfered with quantification of Recombineering results. We suggest considering another fluorescent reporter change that does not overlap cellular autofluorescence, like Green to Mostaza or further towards the red spectra. ICP8 dependent loss of target gene interfered with quantification of Recombineering results and possibly led to an underestimate of Recombineering efficiency. If loss of target gene is due to direct repeat transgene loop out, we suggest considering a non-viral reporter transgene, like a plasmid.

**Allele specific interactions**

The results of co-purification and mass spectrometry suggest a mechanism for viral recombination coupled to host cellular replication and recombination machinery and how Recombineering complexes might be regulated by proteolysis. We recommend validating the Mass Spec data by reverse pull downs and functional characterizations of Recombineering efficiency in strains modified to have hypomorphic (deletion) and hypermorphic (overexpression) alleles of the identified factors.
APPENDIX A. PLASMID SEQUENCES AND MAPS

Acquired plasmids for this study

>pNL-EGFP/CEF

This plasmid was a gift from Dr. Reiser. The sequence data was obtained from genewiz. The plasmid sequence is incomplete. The sequence includes the cef promoter, the egfp gene plus some additional HIV sequences downstream of egfp.

>PNL-EGFP/CEF

TCTTTTTTCGCAACGGGTGGTTTTCGCGCCAGAACAAGTAAGTGCCCGCTGTTGATT
CCCGGGCGCTGGCCTCTTTACGGGTATGGCCCTTGCGTGCCTTGAATTACT
TCCACGCCGCCCTGGCTGCACTACGTGATTCTTGTATCCCGAGCTTTCGGGTTGGAA
GTGGGTGGAGAGTTTCGAGGGCTTCGCTTTAGGAGGAGGAGGCCCTTTCGCTCCTGTC
TTGAGTTGAGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGG
CACCTTCGCGCTGTCTCGCTGCTTTTCGATAAATCTCTAGCCATTTAAAAATT
TGATGACCTGCCTGCAGAGCTCTTTTTTCTGGCAAGATAGTCTTGTAAATGCGGG
CCAAGATCTGACACTGTTATTTTGTTTGTGCTGGGGGCGCCGCGGCGGCGGCGGACGG
GCCCGTGCGTCCAGCCACATGTTTCGCGAGGCGGGCGGCGGCGGCGGCGGCGGCGGCGG
CACCGAGAACGGGAGGGGTAGTCTCAAGCTGGCCGGCTCCGGTGCC
TGGCCTCGCCGCCGCCGCTGTAATCCGGCCCGGCTGGGCGCAAGGGCTGGCCTGTCGCC
TCGCCACACCAGTGCTGAGCCGAGAAGATGGCCGCTCCGGGCCCTGCTGCTGCA
GGAGCTCAAAATGGAGGACCGGCGCTCGGAGAGGAGGCGGCGGCGGCGGCGGATGAC
CCACACAAAGGAAAAAGGGCCTTCGGCTCCTGCTCAGCCGGTGCTTTCATGACT
CACGGAAGTACCGGCGCCGCTCAGGCAACCTGCATTAGTTTCGAGCTTTTGG
AAAGAAGACAAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTACTT
CCCTGATTGGCCAGAACTACACACCAGGGCCAGGGTCAGATATCCACTGACC
TTTGGATGGTGCTACAAGCTAGTACCAGTTGAGGCCAGAT

>pET28a

Plasmid obtained from Novagen.

TGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTG
GTACGCGACGCGTGGACCGCTACACTTTGACCAGCAGGGCCTAGCGGCCGCTCTTTT
CGCTTTTCTCTCCCTTTTTTCTCGCCACGTCCGGCTTTCCCCGTCAGCTCT
AAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACC
CCAAAAAACTTGATTAGGGTGATGGTTCACGTAGGGCAAGTTGCTTTTAATATGTTGCCAGGACTCTTT
CACTCCAGTGCGCCTCCCCGTTCCGCTATCGGCTGAATTGTGATTGCAGAGTGA
TATTTATGCCAGCCAGCCAGACGCAGACGCAGCCCGAGACAGAACTTAATG
TATGACAGCATCAAGAAATAACGCCGGAACACATTAAGTCAGCGAGCTTCAACAG
CAATGGGCATCTCTGGTCATCCAGCGGACGATGTTAATGATCAGCGCTCCACAG
TTGCAGCGAGAAAGATTGTGCACCAGCAGCCTTTTTACAGGCTTTCGAAGCGTGCTT
CTACCATCGACACCACCACGCTGGACACCCAGTTGATTCGAGGCGAGATTTAAT
CGCCCGCAGCAATTTCGACGGCGGCAGTGCAGGAGGCTGAGGTGCAAC
GCCAATCAGCAACGACTGTTGTGCCGCCAGTTGTGGTCACGACCAGCTTGGGA
ATGGAATTTCAGCTCCGCCATCGGCCTTTTTTTTCGCTTTTCGAGAG
AATGTGGCTGGGCTTGGTACTCCACCACGCGGAAACGGGCTGACTGTGATAAGAGACACCG
GCATCACTCTGCGACATCTGATAACGTTACTGTTTACACTACACCACCTCTGAA
TTGACTCTCTCTCAGCGCTATACATGCCCATAACGCGAAAGGTTTTTGCGCATT
CGATGGGTGTCGGGAGATCAGCAGTCGCTTTGGCAGTCTCGGAGTTCCTGCA
GCAGCCCAGTAGTAGGTTGAGGCCGTTGAGCACCGGCCGCCGCCAGAGGAATGGT
GCATGCAAGGAGATGGCGCCCAACAGTCCCCCGGCCACGGGGCCTGCCACCA
TACCCACGCGAAACAAGCGCTCATGAGCGCGGGAATGGCGAGCGCGATCTTC
CCCATCGGAGTGTGGCGGATATAGCGCCAGCAACCGACCTGTGGCGCCGG
GTGATGCCGCGCCAGTAGTGCGCTCGGCGGTAGCGAGATCGAGATCGATCCGC
GAAATTAATACGACTCATATAGGGGAATTTGTGACGGGATAACAATTCCCT
CTGAAATAATTTTGTAAAATCTTAAAGAAGGAGATATACAGGAGATACATGCGGAGCAGC
CATCATCATCATCATCACAGCGGCCTGGGTCCGCCGCAGCCAGCATATGG
CTAGCATGACTGGTGACAGCAAATGGGTCTCGGGATCCGAATTGAGCTCCG
TCGACAAGCTTGCGGCCCAGTCACTCGAGCACCACCACCACCACCACGAGATCC
GGCTGCTAAAGAGGCCAAAAAGAGGACGTGATTGCTGCTGCTTGGCCACCGCTGAG
CAATAACTAGCATACATGCTTCCTTGGGCCTCTAAACGTTGAGGGGTTTTTT
GCTGAAAGGAGGAACTATATCCGGAT

>pFN22K

Promega Flexi vector for creating N-terminal HaloTag fusions. Expressed in BL21(DE3) and in mammalian cells.

>gi|193870792|gb|EU621377.1| CMVd1 Flexi Vector pFN22K (HaloTag 7)

CCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACG
CAAATGGGCGGTAGGCGTGTACGGTGAGGTCTATATAAGCAGAGCTGGTT
ATCGGCGGAGACATCTTCTCAAAACAGGGAAGGCAAACCTCCGGGCAAAAGC
GGACGAACATGGCTGAAGCGGATATTACCTGACTATACATCAGGGCTTCAGAAATT
CAGACCGGATTCTTTACTCAAAGGCAGACTTGCTGATTACCAAAAAACACGGACC
TTATCAAGACCTTTACAAAAATCAGATAATGGTTAAAACGAATTGCGCTCGGCTCA
CCCGGGGATCTCCTCAGAGCTGACCTGCAGGCATGAAGCTGATGGGCTGCTGAAGC
AACAAAGGCGAAAGGAAGCTGAGCTACATATAAAATGCTTTATTT
GTGAAATTGTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAAA
CAAGTTAACAAACAACATTGCATTCATTTTTATGTTTCAGGGTCGGGAGAT
GTGGGAGGTTTTTTTTAAGCAAGTAAAAACCTCTACAAAATGGTGAATAATCGAA
TTCTAATGGATCCTCTTTGCGCTTGCTTTCCCTGTTGAGATAGCCAGAGCCAGTAG
CTGACATTCATCCGGGTCAGCACCGCTTTCTCGGGAAGTGGCTTTCTACGTGTT
CCGCTTCCTTTAGCAAGCCCTTGCGCCCTGAGTGCTGGCCAGCGTGAGGCTTC
AAAAGAATTGCCAGCTGGGGCGCCCTCTGGTAAGGTTGGGAAGCCCTGCAAA
GTAAACTGGATGGCTTTCTTGCCGCCAAGGATCTGATGGCGCAGGGGATCAA
GATCTGATCAAGAGACAGGATGACGGTCGTTTCGCATGCTTGAACAAGATGG
ATTGCACGCAGTTCTCCGGGCGGCTTGGTGAGAGGCTATTCCGGCTATGACT
GGGCAACAACAGACAATCGGGCTCTGCTCTGATCGACCTGCTGGCCGGCCTGAACCG
GCAGGGGCGCCGGGTCTTTTTGTCAAGACCGACCTGCTCCGGTGCCCTGAATG
AACTGCAGGACAGCAGCGGCGGCTATCGTGCTGGGCTGGGACAGCAGGGGCGTTC
TTGCGCAGCTGCTGCTGACGTGTTGGTCAACTGAAGCGGAAGGGAAGCTGGCGCTA
TTGGGCGAAGTGCGGCGGGGCAGGATCTCCTGCTCATCTACCTTGCTCTGGCGGA
GAAAGTGATCCATCTGGCTGATGCAATGCAGGCTGCATACGCTTGATGGGCTACCTGCACCATTCGACCACCAAGCGAACAATCGCATCGAGCGAGCGACGCACGCACTCGGATGGAAGCCGGTCTTGTCGATCAGGATGATCTGGACGAAGAGCATCGGGGGCTCGCGCCAGCCGAACTGTTCGCCAGGCTCAAGGCGCGTATGCCGGATGGTGAGGATCTCGTCGATGACTCATGGCGATGCCTGCTTGCCGAATATCATGGTGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGAGCGGGACTGGGGGTTCGAAATGACCGACCAAGCGACGCCCAACCGGTATCAGCTCACTCAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAAATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACTGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGTGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAACTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTGTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTGCAAGCAGCAGATTACG
CGCAGAAAAAAAGGATTTCACAAGAAGATCCTTTTGATCTTTTCTACGGGGTCTG
ACGCTCAGTGGAACGAAACTCACGTAAAGGATTTTGGTCATGAGATTATC
AAAAAGGATCTTCACCTAGATCCTTTTATAGTCCGGAAATACAGGAACGCAC
GCTGGATGCGCCCCCTGCTGGGATGGTGAAACCATGAAAATGCCAGCTTCAG
TGATTAAGTGCTGGGGTAATGTGGCGCTCCTACGTGTCATAGGTTATTCAT
ACGGTTAAAATATTATCACCGCGCGATTGCGGCAAGTTCAGTTTTTCTCGGTGTTTGTGC
CATTTCACCTGTCTCCTGCCCGTGATCGCGCTGAACCGCTGTAGGTTTATAGCGGTGCT
ACAATTAAGGGGATTATGGTAAAATCCACTTACTGTCTGCTCCCCCTCGTAGGCATCGA
GATAAACC CGCAGTACTCCGCGCAGATGCCTCGGCGTAGAGGATCGAGATC
T

>pFN24K

Promega N-terminal HaloTag Flexi vector.
>gi|193870804|gb|EU621381.1| CMVd3 Flexi Vector pFN24K (HaloTag 7)

ATGGGCCGTTAGCCGTGTACGGGGAGGTCTATATAAGCAGAGCTGGTTAG
TGAACCGTCAGATCATAGAAGCTTATACGACTCGACTATAGGGTACGGATT
AGGTGACACTATAGAATAAGCCAATATCATCATGGCAGAAATCGGTACTGGC
TTTCCATTTGACCCCCCATTTATGTGGAAAGTCCTGGGCGAGCCAGATGACTACGT
CGATGTGGTGGTCGCGGAGATGGCACCACCTGTTGCTGCTCTGCTGACGGTAAACC
CCTCCTCCTACGTTGCGCAATCCATCCGGCATGTTGCACTGACCCATTCGC
TGCATTGTCCTGACCGTGATCGTGATAGGGCAATCCGACAAACAGACCTGG
GTATTTCTTCGACGACCCACGTCGCCCTCATGGGATGCTCTTATCGAAAGCCCTG
GGTCTGGAAAGGTTGGTTGTCGCTGCTTGCATATCAGACTGAGGTTCGCTCCGCTT
CCACCCGCAAGCGCAATCCAGAGCGCTCAAAGGTATTGCATTTATGGAG
TTCATCCCGCCCTATCCCGACCTGGGACGATGAGCCGACAAATTCGCTGG
CCTCCAGGCTTTCCGCAACCCGACGTGGGCGCGAAGCTGATCATCGATCA
GAACGTTTTTTATCGAGGGCGTACGCTGGCGATGGGGTAACGGGTCACTGGACTG
AAGTCGAGATGGACCATTACCGCGAGCCGTTCTGATATCATGGCCGCGA
GCCACTGTGGCGCTTCCCAAACGAGCTGCAATCCCGCGGTTAGGACCGACAC
ATCGTCGCGTGGTGAGAATACATGGACTGCTGGCTGCACCAGTCCCCTGCG
CGAAGCTGTGGCTTCTGGGCAACCCAGCGTTCTGATCCACCCCGGCGAAGC
CGCTCGCCCTGGCACAAGCGCTGAATCCGCTGGGATGAGCCGAGCCGAAGC
GGTCTGAATCTGCTGCAAGAAGACCCGACCTGTGACCGAGCGAGATCG
CGCGCTGGCTGTCGACGCTCGAGATTTCCCGAGCCCAACCATGCGGACCGAGATCG
GTACTTCAGAGCCGATAACCAGCGATCGGCCATGGAATAGTAAGGGAATCCACAT
GGCACAGGTTATCAACACGTTTGGACCGGGGTTGCGGATTATCTTCAGACATATC
ATAAGCTACCTGATAATTACATTACAAATTCAAGAAGCACAAGCCCTCGGCTG
GGTGGGATCAAAAAGGGAACCTTGCAAGACGCAGCTCGCTCCCGGGAAAAAGCATCGG
CGGAGACATCTTCTCAACACAGGGAAAGCCAAAACCTCCCGGGCAAAAGCCGACCATTATCA
GACCTTTACAAAAATCAGATAATGTGTAAAACGAATTTCGGGCTCGGTACCCGG
GGATCCTCTAGAGTCGACCTGACCGACAGTCAAGGCTATTGCTGCTCAACACCCGCTGAGCAATAACTAGCA
TGACTCTTTGGGCGGCGCTCGACGACATGATATTGATGCTTGGCTGAGCAATAACTAGCA
TTTGCAAAACAAACTAGATAGCTGAAAAATGCTTTTATTTTGTGAAA
TTTGTGATGCTATTGCTTTATTTGTGAACCATTATAAAGCTGCAATAAAGTT
AACAAACAAACTGACATCTTTATGTTTCAGGATGAAAAATGCTTTTATTTTGTGAAA
AGGGTTTTTATAGCAAGTAAACACTACTACAAATGTGGTAAAATCGAATTCTAAA
TGCTCTCTTTTGCAGCCTTCCTTGGCTGCTTGTCCGGCTCGGCGCCGCTTGCTTTCGCTTGGCG
CTCTTTAGCAGCCCTTGCGGCCCTGAGTGCTTGCGGCCAGCGTGAGCTAAGAGAAAATGCTTTATTTGTGAAA
CTGATGCGTTTCTTGCCGCCAAGGATCTGAGACGGGCAGGGGATCAAGATCT
GATCAGAGACAGATAGCTGCTAGCAGCTTGAGAAGATGGGAGCTTGGATTTCTCAGG
ACGACGGTTCTCCCGGGGTGGGTGGAGAGGCTATTCGGCTATGACTGGGCAACAGCAAATCGGCTGCTCTGATGCCGCTGTTCCGGCTGTCAGCGCAG
GGGCGCCCGTCTTTTTGTCAGCGACCTGTCCGGTGCCCTGAATGAACT
GCAGGACGAGGGACGCGCGGCTATCGTGGCGGCTGGCCACGACGCGGGCCTTCTTTCG
GCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAA
CAAACCCACCCTGGGTAGCCTTGGTGGTTTTTTTTTTTGGTCAAGCAGCGAGATTACCGC
CAGAAAAAAAGGATTTCAAGAACAGATCCTTTGATCTTTTCTACGGGGTCTGAC
GCTCAGTGGAAACGAAAATCTACAGTGAAGGATTTGGTGTAGGATTATCATTAA
AAAGGATCTTACCTAGATCCTTTTATAGTGCTGGGAAATACAGGAAACGCACGC
TGGATGGCCCTTTCGCTGGGATGCTGAAACCATGAACAAATATTGCAGCTTCAGTG
GATTAAGTGGGGATATGTCCTGCTCTCTCTGGTTCATAGGTATTTCATAC
GGTAAAATTTATATCAGGCAGGATGCTGGGGCATTTTCTCGGTGGTTTGTGCCCA
TTTTACCTGTCTGCTGCGCTGTACGCTGAGCTGAACGCTGAGTTTACGCGGTGAC
AATTAAGGGATTATGGTAATCCTACTTGACTGTGACTCCCTCGTAGCCATCGAGA
AAAACCGCAGTACTCCGGCCACGATGCGTCCCGGCGTAGAGGATTAGGATCGAGATCT
>pFC15K

Promega C-terminal HaloTag Flexi vector

>gi|163943735|gb|EU332338.1| CMVd1 Flexi Vector pFC15K (HaloTag 7), complete sequence

CCAAAAATCAACGGGACTTTTCCAAAAATGTCGTATAATAACCCCGCCCCGTTGACG
CAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTGGTT
TAGTGAACCGTCAGATCAGAGCTTAATACGACTCACTATAGGGTACGA
TTTAGGTGACACTATAGAATAAGGAGCGATCGCCATGGAATAAGTAAGGAAT
CCACATGGCACAAGGTATCTAGTTGACGGGGTTGCGGATTATCTTCAG
ACATATCATAAGCTACTGATAATTACATTACAAAATCAGAAGC
ACAAGCCC
TCGGCTGGGTGGCATCAAAAGGGAACCTTGCAGACGTCGCTCCGGGGAAAAG
CATCGGCGGAGACATCTTCTCAAACACTAGGAAAGGCAAACTCCCGGGCAAAAG
CGGACGAACATGGCGTGAAAGCGGATATTAACTATACATCGGCTTCAGAAAT
TCAGACCGGATTCTTTACTCAAGCGACTGGCTGATTTCACAAACAAACCGGACC
ATTATCAGACCTTTTACAAAAATCAGATAATGTTTAATGACCCCGTGTCGAGCT
CTCGAGCCAACCACCTGAGATCTGTACTTTCAGAGCGATAACGATGGATCCG
AAATCGGTACTGCGCTTTCCATTCGACCCCCATTATGTGGAAGTCTCTGGGCGAG
CGCATGCACTACGTGATGTGTTGGGTCCGCGGATGGCACCCCCCTGTGCTGTTTCT
GCACGGTAACCCGCACCTCCTCTTCTACGTGTTGGGCAACATCATCCCGCATGTTG
CACCGACCCATCGCTGCATTGCTCCAGACCTGATCGGATATGGAACCAATCCGA
CAAAACCAGACCTGGTTATTTCTCAGACGACACCAGTCCGCTTCTCATGGATGCCT
TCATCGAAGCCCTGGGTTCTGGAAAGAGGTTCGTCTCTGGTCAATTCACGACTGGG
CTCCGCTCTGGGTTTCCACTGGGCAAGCGGTAACCCGACCTCCTACGTGTGGCGCAACATCATCCCGCATGTTG
CACCGACCCATCGCTGCATTGCTCCAGACCTGATCGGATATGGAACCAATCCGA
CAAAACCAGACCTGGTTATTTCTCAGACGACACCAGTCCGCTTCTCATGGATGCCT
TCATCGAAGCCCTGGGTTCTGGAAAGAGGTTCGTCTCTGGTCAATTCACGACTGGG
CTCCGCTCTGGGTTTCCACTGGGCAAGCGGTAACCCGACCTCCTACGTGTGGCGCAACATCATCCCGCATGTTG
ATTGCATTATGGAGGTTATCCCATCCGCCCTATCCCCGACCTGGGACGAATGGGCCAGA
ATTGCCCCGCAGACCTTCCAGGCCCTTCCGCACCCACCAGACGTCCGCGCCGCAAG
CTGATCATCGATCGAAGCTGTTTATCGAGGGTACGCTGCCGATGGGTGTCGT
CCGCCCGCTGACTGAAATCGGAGATGACCATTACCGCAGCCCTTCCTGAAT
CCTGTGACCAGCAGCCACTGTTGGCGCTTCCAAAACGAGCTGCCAATCGCCG
GTGAGCCAGCGACATCGTCCGCTGGTTCGGAAGAATACATGGAATGGCGTCGCA
CCAGCTCCCTGTCCCAGAAGCTGCTTCTGGGCGACCCCCAGGCGTTCTGATCC
CACCAGCGCCGCTCCCGCTTGGCAAAGCTGCCTAATCTGCAAGGCTGT
GGACATCGGGCCCAGGTCTGAATCTGCTGAAAGAAGACAACCGGACCTGATC
GGCAGCGAGATCGCGCGCTGCTGTCTACTCTGGAGATTTCCGGTAAATAGA
ATTCTAGAGTCCGCACCTGCAGGCGATGCGAAGCTGATCCGGCGCTGCTAACAAGGCC
CGAAAGGAAGCTGAGTGGGCTCCTGCGCCACCGGCTGAGCAATAACCTAGCATAAC
CCCTTGCGCGCCGCTCCAGACGACATGAAAGATACATGGAGGTTTG
GACAAACCACAAGTGCAATGCGAAAATGCTTTATTATTGAAAAATG
TGATGTCTATTGCTTTATTGTAACCATTATAAGCTGCAATAAAAGAT
ACAACATTTGCAATCCATTTTTATGTTTCAGGTTCAGGGGAGATGGGAGGTT
TTTTAAGCAAGTAAAACCTCTCAAAATGTGGTTAAAATTGAAATGAACTG
CTCTGGTGCTTGCTTTTTCCCTTGTCAGATAGCCCATGACTGCAATTCTCCTTC
CCGGGTTCAGACACCGTTTCTGCGGAACTGGGTTTTCTACTGTTTCCCGCCTTTCTTT
AGCAGGCCCTTTCGCACCCATGGAAGTCTTGGCGGAGCTTCAAAGATGTTG
CCAGCTGGGCGCCACCTCTGTGTAAGCGGAGGTATGAGTACTGGGCAAA
AGACAGGAAGATGACGGTCTGGTTTCAGATGCTATTGCAATCAGATCTGACT
GGTCTCCTCCGCGCCGCTTGGTGAGAGGTATTTCGCTATGACTGGGACACAC
AGACAATCGGCTGCTCTGATGCGGCGGTCTTCCGGCTGTCCAGCGCAGGGCAGG
CCGGGTTCCTTTTTTGTCAAGACCCGACCTGTCGCCGTGCCCTGAATGAAACTCAGG
ACGAGGCACGCGCGGCTATCTGTGGCTGGCCACGACGGGCGTTCTTTCGCCGAGC
TGTCGTCGACGTCTGTCAAGGCGGAAGGGGACTGGGCTATGCTATTGGCGAA
GTGCGCGGGCGAGGATCTCCTCTCAGCGTGACCGCCAGAAAGTATCCT
CATCATGGCTGATGCAATGCGGCGGCATACGCTGTCAGGGCTACCTGC
CCATTCGACCCACAGCGAAACATCGCAGTGCAGCGACGCACGGCAGCTCGGATGG
AAGGCCGCTTTTCGATCAGGATGATCTGGACAGAAGGCAGCATACGGGGGCTCGC
GCCAGCGGACTGTCTGCGACCGGCGTATGCCGGGATGGGAGGAT
CTCGTCGACTCATGCGATGCGCTGGCTTGCCGAATATCAGTGGGAAAATG
GCCGCTTTTCTGGAATTCATCGACTGTTGGCCGGCTGGGTTGTCGCGGACCGCTAT
CAGGACATAGCGGCTACCGTGAATATTGCTGAAGAGC
GGGCTGACCAGCTCTCTCCTGCTTTACGGTGATTCGCAGCGCTCCCGATTCGCAGCGCT
ATCGCCTTTCTATCGCCTTCCTTCTGAGTTCTCCTGTAGCTGTGGTCGCTGTT
GAAATGACCGACCGACGCGACCGTTCCCGGATCCAGCTCAGCTCCAAAAGGCGG
ACTTACAGTTCAGGCTGTTCCATAGGCTCCGCCCTGCCGCCGCTGCTCTCCTG
TTCCATAGGCTCCGCCCTCTGAGACGACATCACAAATCGGACGCTCAAGT
CAGAGGTTGGCAGAAACCCGAGCAGGACTATAAAAGATACCAGGCGCTTTCCTCCCCTG
GAAGCTCCCTCTGCTGCCTCTCTGTTCTCCGACCCCTGCCGCTTACGGCAGAAG
TTTCCATAGGCTCCGCCCTGCCGCCGCTGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTG
GTATCTCAGTTGTCCTGTCCTGCTCTGCTCCAAGCTGGGCTGTTGCACGAAAC
CCCCCGTTCAGCCCGACCCTGCTGCCTCTTATCCGGTAACATCGCTCTTGAGTCC
TTAGCAGAGCGAGGTATGATGAGGCTGCTACAGAGTTCTTGAAAGTGTTGGCC
CAGGTTTTCTCCGAAAAAGAGTTGGTGCTACCTTGGATCAGCGCATCCGGAAAA
CAGCTGGTGTTGCCGAGTTTTTTTGGTGCAAGCACAGCAGATTACGCGGCAAGAA
AAAAGGATTTCTTGAAAGATCCCTTTTGTCTCTTCTACGGGGTCTGACGCTCAGT
GGGACGAAAACTCAGCCAAGGATTTTGGTCATGAGATTATCAAAAAGGAT
CTTCACCTAGATCCTTTTATAGTCCGGAAAAATACAGGAAACGCACGCTGGATGG
CCCTTCGCTGGGTAGTGTTGAAACCATGAAAATGCGACCTCGAGCTGAGTAAAG
TGGGGGTAATGTGAGGCTCGGCTACCTGCTGGTTGCAAGAGGTTATCCATACGGGTTAAA
ATTATCAGGGCGGATTGCGGCAGTTTTTTTCCGGTGTTTTGTGCTCCCATTATTACC
TGTCATGCTCGCCGTGATCGCGCCTGAAACGCGTTTTTGGCGGTGTCACAATTAAAGG
>pCMV-ICP8

Plasmid expresses HHV1 ICP8 from CMV promoter. This plasmid was a gift from Dr. Knipe.

> pCMV-ICP8

TCAATATTGGCCATTAGCCCATATTATTICATTGGTTATATAGCATATAATGACCGCCATGTTGGCATTGATTATTGACTAGTTATTAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCAGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGA
CTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCA
GTACATCAAGTGTATCATATGCGCCACGCTCAGCTCAGCTTACCATGGTGATGCGGTTT
TGCGAGTACATCAATGGGAGTTTGTTTTGGCAACCCAAAATCAACGG
GACTTTCCAAATGTGTAATAAACCCTGGCTGCAGCTGAACATGGAATGGCGGTA
GGCGTGTAACGGTGGAGGTATATGAGCTGCTGTAGGTCAGGCGTA
GATCAGTACAGCTTTATTGCGGTAGTTTATGACGTAAATTGCTAAGCGCAG
TCAGTGCTTCTGACAAACAGGCTCAGAATGCTCCGGGGTTTGCTGCGTGAAGGGTGGG
GATTCATTCCCCAACCCTGGCTCTCCCCACCTCCCCCTAGGCCGTCCGCCTGGGC
CCCCCCTGATTACGGCTCCGGGGAGCACTACCTCCTCGGCTGCTCTGCTGTCG
GCCGGGATTACCTCCGTTTGCGGACGGATTGCCAGGAGACATGGAGACAAA
GCCCAAGACGGGAAACCACCACTCAAGGTCCTCCCCCCGGGCTGGGCTGGG
TGAGCGCCTCGCCGCTGCCTCGACAGCATCGAGCTCTCGTGGCAGCTGCAGC
GCGAGCGGCTCAGGCTCAGGCTGCGCCTGCTGGCGCCCTGGCCTGGTC
GTCGCTCGCCGCTGCTCCTCGGAGACGTAGCGCTGACGCTGACGCTG
GCGCAGCGGCGATGCGACGTCGTCCGGCCTGCTGGCGCCCTGGGCTGGG
CTGTCCGTGTCGCTTCTCAACGGCGGCACCTGGAGCCACACCCACAG
GCCCCAACCCTGACGCGACTCTCGGAGCGCGCGACGCGCCTTGGGCTTTGC
GGACTCACCCCCGCGCAGCGACCTCAACAGGAGACCAAGGCGCAGGC
GCTGTGAGCGCCCTGGGACCCGAGACGGCGCCTCTTGATCTGGTCG
TTACCGAGGGCTTTCAAGGAGGCCGTGTGACATCAACAACACCTTTTCTGCACCT
GGAGGCTTCCGACAAGTAAACCATAAGGGGGGAGGAGGTCGACACCCTATACC
CGTGATTCGGCTGCAGCTTTCACTGCGGATTTTAGCCGGGTTCGATCGGCCGAC
CGTTCAACGCCAACCCACCAGTGATCGTCCGAAAGGAGGTCGACCTGCCCTTT
GGGAGGCTCGGACAAGGTAACCATAGGCGGGGCGGAGGTGCACCGCATACC
CGTTCAACGCCAACCCACCAGTGATCGTCCGAAAGGAGGTCGACCTGCCCTTT
TTTTTTTAACCCGCCCCCTCAACCCGCTCTCTGTTTCAGGCGGTCGTCGGGAAC
CCGCCGCAACTGCATGCGAGCCGAAACCTGGACCGCCTGCGGCCGGGCGGCCG
CCACCTGGCGGGTTGCAAGAAAACCAGAGGGCGCCGCGCT CCTCCCGCCGACATT
ACGTTCACGGCCTTCAAGGAGGCCAGCCGCTCTTCGAGGCGGTCGTCGGGACCC
GCACCTGGCGTTTACAGTGGTAGTGAAATATTGAGCAAGGGATTCCCCTGCCAAG
ACGGCGGCCAGCAAGCCCCGGCGGGGCGGCTGCCTACTGGACGCCTGGCTCCG
CTTCGAGCAGCCGCCCCACCAGCAGCCTCCGAGCTGCTGCGAGGGCCAGG
ACGGCGGCCAGCCGCTCTTCGAGGCGGTCGTCGGGACCCAAAGGACCACAGCAACC
CTCTCTCGGCAAGGCCCAAGGACAGCCCTCCGAGGGAACTGGAGCGCTGCGA
CGAGCTGGGGCTCAGCGGTCCGGCGGACCCAAAGGACCACAGCAACCGCATG
GACAGCGACTGCGACGTGGGAACTACGCCGCTCTCGGCCCTGAAGC
TGGACGTCCATGGTCGCCAGCTGCAACCTGCTGCGGCCCGTCATGGGC CGGC
GCCCCATGGTGCTGTTGGGTTGAGCATCAGCAAATACTACGGCATGGCCGG CG
CAACGACCCTGTGTTCCAGGCGGGGAACCTGGCCAGCTGCAACCTGCTG
CGGCCCGTCATGGCGGCGC
TCCCCGGGCCCCTTCTATTTTTGAGCTCTGGCTCTCTGGCTCTCTGGCTCT
GAAAGCTCGCTGTGCGAGCAGCTCCGGG
GCATTATCTCCGAGGGCGGGGCGG
CCGTCGCCAGTAGCTGTTCGTGGCGACCGTGAAAAGCCTGGGGCCCCGCAC
CCAGCAGCTCGACATTGCGAGCAGCTCGACCTCCGGGGCCATTATCTCCGAGGCGGACAGGTCGC
GAGTGGTCGACGGACGCGGCCCCTTGAGGTTGGCAGCCAGCGGCAGGCAGGCAGGC
GACGAGGAGATGATGAGGCTGACCCTGGCTGGCGCTCCTGGAGGACGAGTACCTA
AGCGAGGAGATGATGAGGCTGACCCTGGCTGGCGCTCCTGGAGGACGAGTACCTA
AGCGAGGAGATGATGAGGCTGACCCTGGCTGGCGCTCCTGGAGGACGAGTACCTA
GTCAGCCCAACTCCGGCAACGCGCCGGAGGTGTTAAGCTTTTGAGATTTTGAGCT
GCGAGGACGACAACGCAGCCGCCCCCCCTCGAGGTGGCGACCCAGGCGAGGCC CGGC
CATTGGCGCCCGCACACGGCGCTCCAGGGATGACCCGCCTTTGGGGAGGGG
GGGTTGTATGATGAGTTTGGACAACCACAACTAGAATGCAGTGAAAAAA
ATGCTTTTATTTTGAGATGCTATTTGCTTTTTATTTGTAACCATTATAAG
CTGCAATAAAACAAGTTACCAACAACAAATTCATCATATTTTGATTTTGCTCTGC
AGGGGGAGATGATGAGGCTTTTTTTAAAGCAAGTTAAAAACCTCTACAATGTGG
TAAAATCGATAAGGATCCGGCTGCTTTAGCAATGCGAGGAGGCCCAGCCG
CGCCCTTCCCAACAGTTAGCAGCAGCTGAAATGGCGAATGGCAATGGCGCGGG
CGGCGCATTAAGCGCGCGGGGGTTGTTGTTGTTACGCGACGCGACGCGG
CTTGCCAGCGCCCTAGCGCCCGCTCTTTTGCTTTTTTCTCCTGCGCC
CGCTTTTTTGCAACACATGGGGGATCATGTAACCTCGCTTTGATCGTTGGGAAC
CGGAGCTGAAATGAAGCCATACCAAAACGACGAGCAGTGACACCACAGATGCCTGT
AGCAATGGCCAACAAACGTGGCGAAAACACTTTAAAATCTGGGCAAACACTTACTCTA
GCTTCCCCGGCAACAATAATAAGACTGGATGAGGGAGCGGATATAAGTTGCAGGAC
CACTTCTGCAGCTCGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAAATCTGGGA
GCCGCTGAGCGTGGTCTCGCCGTATCATATTGCACGACTGGGCGCCAGATGGTA
AGCCCTCCGTTATCGTATGTTATATACACAGACGCGGAGTCAGGCAACTATGGA
TGAACGAAATAGACAGATCGCTGAGATAGTGCTCCTCAGTAGTTACGCATATGGA
TAACTGTCAGACCAAGTTTACTCATATATATCCTTTAATGGATTTAAAACTTCA
TTTTTAAATTAAAAAGATCTAGGTGAAGATCTTTTTGATAATCTCATGACCA
AAATCCCCATAGTGGAGTTTTCTGTTCCACTTACGAGCAGACCGCCAACCGTAAAA
GATCAAAAGGATCTTTCTTGAGATCCTTTTTTCTCGCCGTAATCTGCTGCTTGCA
AACAAAAAACCCACCGCTACCAGCGGTGTTGTTTGGCCGGATCAAGAGCTAA
CCACTCTTTTTCCGAAGGGTAACCTGCGCTCCACTCAGCAGCAGCAGGATACCAAAT
CTGTTCTTCTAGTGTAGCGCAGTCTAGTGCCACCACCATTCAAGAAGCTCTGTAGCA
CCGCCCTACCATTACCTCGCTCTGCTAATCTGTTACAGTAGTGCTGCTGCGAGTG
CGATAAGTCGTTCTACCCGGGTTTGACTCAAGACGATAGTACCAGGAAG
GCGCAGCGGTCGGGCTGAACCGGAGGGGCTTCGTCGACACAGCCAGCTTGGAGGC
GAACGACCTACACCGAGATACCTACAGCAGGATGCTATGAGGAAAAGCG
CCACGCTTCCCCAGGAGAAGGCGGACAGGTATCCCAGTAAACGGCGAGG
TCGGAAACAGGAGACGCAGGAGGAGCTTCCAGGGGAAACGCGCTGTATC
TTTATAGTCTGCGGGTTTCCGACCCTCAGACTTTGGAGCGTCGATTGGTGTAT
GCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTT
ACGGTTTCTGGCCTTTTTGCTGGCCTTTTTGCTCACATGGCTCGACAGATCT

>pCMV-ICP8-GFP

Plasmid expresses ICP8-GFP fusion from the CMV promoter. ICP8 is partially functional. This plasmid was a gift from Dr. Knipe.

>pCMV-ICP8-GFP

TCAATATTGGCCATTAGCCATATTATTTICATTGGTTATATAGCATAAACATAAATCAATA
TTGGCTATTGGCCATTACGATTCTATCTATATGACATTATGTACATTACATTATA
TTGGCTCATGTCAATATGACCGCCATGTTGGCATTGATTATTGACTAGTTAT
TAATAGTAATCATTACGGGGTCAATTAGTTCATTAGTTCATAGTACATTACATTAT
TAATAGTAATCATTACGGGGTCAATTAGTTCATTAGTTCATAGTACATTACATTAT
TAATAGTAATCATTACGGGGTCAATTAGTTCATTAGTTCATAGTACATTACATTAT
TAATAGTAATCATTACGGGGTCAATTAGTTCATTAGTTCATAGTACATTACATTAT
TAATAGTAATCATTACGGGGTCAATTAGTTCATTAGTTCATAGTACATTACATTAT
CTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCA
GTACATCAAGTGATATCATATGCAAGTCCGCCCCCTATTGACGTCAATGACGG
TAAATGGCCCCGCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTA
TTTGACTCACGGGGATTTCCAA
GTCTCCACCCCCATGCAATGGGAGTTTGTTTTGGCACCAAAATCAACGG
GACTTTCCAAAATGTCGTAATAAACCCCGCCCGCTGACGCAAATGGGCGGTA
GGCGGTGTACGGGTGGGAGGTCTATATAAGCAGAGCTGTTAGTGAACCGTCA
GATCAGTGAAGCTTTATTGCAGATCAGCTTTATACAGTAAAATTGTCGAACGCAG
TCAGTGGTTCTGACACAACAGTGCTCGAATTTAAGGGCTAGAGGATTTAACG
ACTCACTATAGGCTACCTCGAGAATTCGCGGGGGTTTGTACGGTGAAGGGTAGG
GATTCATTCCCCAACCCGGGTCTCCACACCCCTCCCTCCCTGACGCGCGCCCCC
CCCGGGATTTTGACGTCGCGGGAGACATACACCTCGTCGCGGCCTCGTCGTG
GCCGCGGATTACCTCCGATTGCGGACGAGTAGTTGACGGGACGAGCAGAAA
GCCCAAGACGGCAACCACCATCAAGGTCTCCCCGCCGGGCCCTGGGATACGTG
TACGCTGCAGCGTGCTGCTGCAAGAGCATCGAGCTTCTGCGCTGTTAAGGCTCGGC
GCGCAGCGGCAGCGTTGCGCCGCGTGCCGCGGCCCTGGGCTGGCCTGACC
GTGGAGAGCGGCTTTGAGGCAACAGTAGCGCGGTGCAGTGGTGGGTCTCGCAACGA
CGGGGGCTCGGGGGGATCCGCGGTGTCACCTGGAAACGCGCAGCAGCTACACG
CTGTCGCCGAGTGCTTTCTCACGGCGGCGCCGACCTGGGAACCCAGCACCAG
GCCCAACCTGACGCGACTCTGCGAGCGGCGACGCGCCGATTTTGGCTTTTC
GGACTACACCCCGGCGCGCGAGCTCAACACGAGACGAGCGGGGAGGC
GCTGTTGAGCGCCTCGGGCTGACCGGACCAGCAGCGCCCTCCTGTTATCTGAGTCG
GCGCGGACGGATCCGAGACCGCAGCCCGGACCACATCATGCAAGGAGACGTACCAGCG
CGGCGACGGAGCGCTCATGGCGGAAACTCGAGACCGCTGGGAGACCAGCTACGTTGGGAC
GCCCGTCCCAAGCCCTGATGGTGTAACACACGCTAGGCGGTGGTGGAGGGGAGAATTTACTGGC
TCTGGGGCGAGGACCCAGGCCCCATTCCGGATGCCAGGCTGGAGACCATCAGTCGGTACCTG
GGCCCATGCCCCCTGGCTTTCAGGATCGGCGGCAGCCTCAGGCGCGGCTGGAGACCATCATC
ACCAACCGAG
GCCCTGCATACGCTATGAGCGGATGCTGGAGGAGGGAACCTCAAAGTTCGCAGCAGG
AGCAGCTGATGCGCAACCTGGTGGAGGGGAGGAACTTCAAGTTTCGCGACGGGTCTGGGCGAGGCCAACCACGCCATGTCCCTGACGCTGGACCCGTACGCGTGCGGGGCACTGCCCCCTGCTTCAGCTTCTCGGGCGGCGATCCAACCTCGCCGTGTA
CTCGGAGGGGGGCGTCCGAGGCCGCCAAGGCGGGTCGCCAGCCTCCAGAGCGCCTACCAG
AAGCCCGACAACGCAGCTGGACATCTCTTCTCTCGGACGGCAAGCAAGCCCCCGGGGTCACAAGCA
GCCGAACCCGCAGCTGTCGGGCTTTCTGCTGAGCAGTTCCACGCGGCCATCTTCCCCAACGTGAC
GAGCTGGCGATGTACTACATGGGAAACCAGCTTCTGCAGGTACTGGGTAGTACCT
CGACATACTTCAACACCCCTCACGCGGCAATCGCAGGGTCCCGCCGCTCCC
CCACGCTGCAGGGCCGGCGCGGGCGCTGGTGACGCGCAGGGCGGGCCTG
GAGCGCCGGGCCCCGGCCTCTGATAGGACGCCTTGAGACGCGCAATCCGGGGCCCGG
ATACAACTATAACTCCACACATATGTATACATCATGCGGACAAGCAAAAGAAT
GGCATCAAGGTCTTAAAGGATACAGACACACACATTTGAGGTAGGATCCGTGC
AGCTGGCCGACCATTATCAACAGAACCCTCCAAATCGGGGACGGCCCTGTGCT
CCTCCCAGACAAACCATTACCTGTCCACCCAGTCTGGCCCTGTCTAAAGATCCCA
ACGAAAAAGAGAGACCCATGGTCTCCTGCTGGAGTTTGTGACCCGCTGTGGGAT
CACACATGGCATGGACGAGCTGTACAAGTGAGCGGCCGCTTCGAGCAGCAGACAT
GATAAGATACATTGATGAGATTGGAACAAAAACCAACTAGAATGCGATGTAAAA
AAATGCTTTTATTTGTGAAATTTGTGATGCTATTGTCTTTATTTTGTAACCATTATA
AGCTGCAATAAAACAAGTTAACAACACAATTTGCAATCTTTATTTATGTTTCAGGT
TCAGGGGGAGATGTGGAGGGAGTTTTTTAACAAAGTAAAAACCTCTACAAATGT
GGTAAAAATCGATAAAGGTACTGGCGTAAATAGCGAAGAGGGCAGCCACC
GATCGCCCTTCACCAACAGTTGCGAGCCTGAATGGCGAATGGACGCGCCCTG
TAGCGCGCATTAAGCAGCGCCGGGTTGTGTTGCTACGCAGCGGTACGGCCT
ACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTCTCCTTCTTC
GCCACGTTCCGGGGCTTTCCCCGTCAAGGTCTAAATCGGGGGGCTCCCTTTAGG
GTTCGATTTAGTGCTTTACGGACACCTCGACCCCCAAAAACTGATAGGGGTG
ATGGTTCACGTAGTGCGGCTATCGGCGGCTCGATAGACGGTTTTTCGCCCTTTGACG
TTGGAGTGCTCCTTTAATAGTGGAAGCTCTTGTGTTCAAACTGGAACAACACT
CAACCCCTATCTCGGTCTATCTTTTTGATTATAAGGGATTTTGCGGATTTCGGC
CTATTGGTTTTAAAAATGAGCTGATTAAAAAAATTAAACGGAATTAAAAAC
AAAATATTAACGCTTACAAATTTCTGTAGCGGTATTCTTCCTCTGAGCATCTGT
GCGGTATTCTCACACCCGATATGGGTGACTCTTCAGGTCACAAATCTGCTCTGATGCC
GCATAGTTAAGCGGCCCCGACACCCCCGGAACACACCCGCTGAGCGGCCCCGTAG
TTCATTTTTAAATTAAAAGGATCTAGGTGAAGATCTTTTTGATAATCTCATG
ACCAAAATCCCTAAACGAGTTTTTCGTTCCAAGACGGCAGACCCCCGTAAGA
AAAGATCAAAGGATCTTCTTGAAGATCTTTTTTCTGCGGTATATCCTGCTGCT
TGCAAAACAAAAAAACCACCGCTACCGAGCGTGGTTGGTTTGGCCGGATCAAAGA
GCTACCAACTCTTTTTCCGAAGGTAACTGCTGCTACAGCAGAGCGCAGATACCA
AATACGTGTCTTTCTAGTGAGCGCAGGTAATGGCTCCAGCAGACGGCAGATACCA
AGCACCAGCCTACATAACTCCTGCTCAGCTAATCCTGTTACCAAGTGGCTGCTGCA
GTGGCAGATAAGTCTGTGCTTTACCAGGTGTTGACTCAAGACGAAGATGTACCGGA
TAAGGCCAGCAGCCGTCGCTGTAACCGGGGGGTTTCGTCGACACACAGCCAGCTTG
GAGCGAACGACCTACACCAGAAGATACGATACGCTAGCTATGATAGAA
AGCGCCACGCTCCCGAAGGGGAAAGGCCAGTCAGATCTCGGTAAGCGGC
AGGGTCAAGAAGGAGACGCAGCGGAGGAGCTTCCAGGGGAAAAGCGCTGG
TATCTTTATAGTCTCTGCTCGTTTCGCCACCTCCTGACTTGAGCTGATTTTTG
TGATGCTCGTCAGGGGGGCGGAGCCTATGAAAAACGCCAGCAACGCAGGGCC
TTTTACGTTCCCTGGGCTTTTGGCTGCTGCTTTTTGCTCACATGGCTCGACAGATC
T
>pTEC19

Plasmid source of E2-Crimson. From Dr. Ramakrishnan via Addgene (30178).

>pTEC19

CGCGGGCTGCAGGATATCGATACGTAACGTCCGACCTCGAGGG
GGGGGCCCGGTACCAGATCTTTAAA
TCTAGATTTAAGAAGGAGATATACATA
TGGATAGCAGTGAACGTCAAGGCCTTTCATGCGCTTCAAGGGTCACAT
GGAGGGCTCCGTAACGGGCAACCACTTGTACGAGGCTGGGCCTGACAGG
CAAGCCCTACGAGGCGCCACCCAGACCCCAAGCTGAATGAACGCAAGGGCGG
CCCCCTGCCCTTCGCTGGACATCCAGTGTCCCTTCTACGGCTCCA
AGGCGTACATCAAGCACCCCAGCGACCATCCTCATGCGACGACGCGCTT
CCCCGAGGGCTTCAAGTGGGAGCGCTGGGCGTGAGCGTGAGCGTGACCGA
CTCCCGAGGAGCGGCGTGGATGAACTTCCAGGGCGTGACCGTGACCCAGGACTCCTCCCTGACATCAGGCGCGCGTTG
GTGACCGTGACCCAGGACTCTCCCTGACAGGAGCGGCGACCCCTCTACCTACACCAG
TGAAGTTTCATCGGCGTGAACCTCCCCCTCCGACGGCCCCGTAATGCAGAAGAA
GACTCTGGGCTGGGAGCCCTCCACTGAGCGCAACTACCCCCGCGACGGCGTGC
CTGAAGGGCAGAACATGCGCTGAAGCTGAAGGGCGGCGGCCACTAC
CTGTGCTGAGTTCAAGTCCATCTACATGGCGCAAGAAGGCCTGCGAAGCTGCTCCCT
GCTACCACTACGTGGACTAACAAGCTCGACATCCTCCCCACAACGAGGACTA
CACCCTGGTGGAGCAGTACGAGCGCGCCGAGGCCCGCCACCACCTGTTCAG
TAGGCTAGCCACCTGACGTCGGGGGGAAGGGGGAAGGCATACGTTGTGCTCTAA
AATCTCTGATGTTACATTGCACAAGATAAAAAATATATCATCATGAAACAATAA
AACTGTCTGCTTACATAAACAGTAATACAAGGGGTTATGTAGGCATATTCA
ACGGGAAACGTCTTGCTCGAGGCCCGCGATTTAAAAATCCACATGGGATGTGAT
TTATATGGGTATAATAAGGGCTCNGCGATAATGTCCGGGAACTCAGGGCAAA
TCTATCCTGGTGATGGGAAGCCCATGCGCCAGTTGGTTTCTGAAACATGGGC
AAAGGTGCTGCTGCTGGAAATGTATGTACAGATGAGATGGTACAGACTAAACTTGGC
TGACGGAATTATATGCCTCTTCCGACCCTAGCATTTTATCCGTACTCTCTGAT
GACTAGAAGCATGAGATTATCAAAAAAGGATCTTCCACCTAGATCCCTTTAAAAATTA
AAAATGAAAGTTTTAAATCAATCTAAAGGTATATATGAGTAAACTTGGGTCTGAC
AGTTACCAATGTCCAATATGGGACCACCTATCTCAGCGATCTGCTATTTTCG
TTATCATCCATAGTTGCCCTGACTCCCGCAGCAGTGCGCCGACCAGCCCCTCATCGTGCA
ACGCTGACCTCCGGTGACAGGCGGTGTCGCGACCCGGCGCCGGATTT
AAGCGGGCCCATTACCTGTGATAGGGATGTCGCTGCTGACACAAGAATCCCTTG
TTACTTCTCGACCCTATTGATTGCTGGATATCTCTACGCGAGCTGGGAACGA
CCAGGAATTCTGGGAGGCCGGCTGGCCGCCGAGGCTGGGAAGCTGCCTGC
CCGGTGCAGCCGGGTGCTGCGGGCTCCCGGAGCAGCAACAAAAACCTGACTTG
<table>
<thead>
<tr>
<th>TCAGCGAGCCCGGCGGCGGTGATCAAGCTGTTCGCGGAGCAGTGCTGCTGCTCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>GGAGAGGCTCTCGCTGAGTCCGAGGCTAGCTACGCTCTGCCTGCTGCCG</td>
</tr>
<tr>
<td>GGTGCCGCTGGCCCTACCTGCTGGATGAGCCGGATGACCGGGCACCACCTTGGC</td>
</tr>
<tr>
<td>GTCCGCACTGGACGCACGACCACAGAGCGCTGACGCCGCTGCAGCCGGAACA</td>
</tr>
<tr>
<td>CCGTGCTCACCCCCCCATTCCGAGGCTTCTCCCCGAGACCTGCTGGAGGAGC</td>
</tr>
<tr>
<td>CGCGGCGCAACGTCCAGGGACCACCCGCGGGTGGGCTACGCTCTGGGCGGCTG</td>
</tr>
<tr>
<td>CTGGAACCGCCTGGAGGACTGGCTGCGCGGGAGCCTGGGACACGCTGTGGCGC</td>
</tr>
<tr>
<td>GCCAACCCTGGGTCGTCACGGCGACCGTACCCAGGACACGCTCTCGTGAGA</td>
</tr>
<tr>
<td>CCTGGGGCGAGCCGGACTTGTCACCAGGGATCGTGACTTCACCGAGCCTATGG</td>
</tr>
<tr>
<td>GGAGACTCCGCTCTACAGCCTGGTGAACCTGACTCTCAACGCCTTCGGGGCG</td>
</tr>
<tr>
<td>ACCGCGAGATCTCATTGCCCAGGCGTCTCGACGGGCAGGGGAATGGAACCGAC</td>
</tr>
<tr>
<td>AGGACTCTCCGCCGGAACCTGCTCGCCTTACCTTCCATCGCAGACTTCGAGGT</td>
</tr>
<tr>
<td>TCTGAGGAGACCACCGCTGGACCTCTCCGGGCTTCCAGCGATCCGGAGAAGCT</td>
</tr>
<tr>
<td>CAGTGTTCTCTGGGGCGGCCTTCCCGGACACCAGCCCCGCCGCGCCTGACGG</td>
</tr>
<tr>
<td>CTTAATCTTTGCTCATTCTCAGGGAATTCCAAGTGTTTATGTTGAGATTGTT</td>
</tr>
<tr>
<td>GGACGAGTCGGAATCGCAGACCAGATCCAGCTTGTGCAATCTATGGAACCTT</td>
</tr>
<tr>
<td>GCTGGGTGAGTTTTCTCTCTCCATTACAGGAAACGGCCTTTTTCTAAAAAAATAGG</td>
</tr>
<tr>
<td>ATTTGAAATCCTGTATGAAATAATTGCAGTTTCAATTGATGCTGAGTTGTT</td>
</tr>
<tr>
<td>TTTCTAATCAGAATTGTTATTGTGTAACACTGGCAGACGCCATACTGCCGTA</td>
</tr>
<tr>
<td>CTTGACGGGACGGCCTTGGATTAAATCGAAGTTTTGCTGAGTTGAGAGG</td>
</tr>
</tbody>
</table>
GATCAGATCACGCACTTTCGGACAGACCGGCTTCTTCGTGGCAAAGCAGAA
AGTTCAAAAATCACCAACTGCTCCACCTACAACAAAGCTTCATCAACCAGTG
CTCCCTCACTTTCTGGCTGGATGATGGCGGATTCAGGCTGGTAGAGTCACG
CAACACCTTCTTCACGAGGAGCAGACCTACCAGCAGGTGCTGGTATGAGTCACAG
CGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCCGTAATTCT
GCTGCTTGCAAAACAAAAAACAAGCAGCTACCAGCGGTGCTGGTAGTTGGTTTGCCGGA
TCAAGAGCTACCACACTCTTTTCCGAGGTAACCTAAGAGAGGAGGACAGAGCTAG
ATACCAAATACCTGCTCTTCTAGTGTAACCCGTAGTTAGGCAACCACACTTCAAGAA
CTCTGCTAGCACCGCTACATACCTCGCTCTGCTAATCTTTGGTACAGAGGCTG
CTGCCAGTGGCAGTCAAGCTCTACCGGGTTGAGACTCAAGACGATAGTT
ACCAGGATAAGGGCAGACGGTGCTGGGAACCGGGGCTGCAGAAGGGGGGATTGTGCAACACAGCC
CAGCTTGGAGCGAACGACCTACACCGGAACTGAGATACCTACAGCGTGAGCAT
TGAGAAAGCGCCACCGCTTTCCCAGGAGAGGAAGGCGGACAGGATATTCCGGAAC
AGCGGCAGGGTGCGAACAGGAGAGCCGAGCCAGGAGGCTCCAGGGGGAAAC
GCCTGGTAGCTTATAGTGCTCGTGGGTCGCCACCTCTTGACTTGGACGCTCG
ATTTTTGTGATGCTCGTCAGGGGCGGAGCTGGCTATGGAAAAACGCCAGCAAC
GCWGCGCTTTTTACGGTTTGGCCTGCTGGTTTGGTTTGGGTGCTCAGCTGGTGTT
CCTGCTGTTATCCCTGATTCTGTGAGTAACCCGTATTACCGCTTTTGAGTGAGC
TGATACCGCTCGCCGAGGCAGGAGACCGGAGCCAGCAGCGTGGACCGCCCCACACAG
CTCCGTAAGTTCCGCTGCTGCGTGGCTGTACCCCGCACTTCAGGAGCCAGG
GGGCTACAGGGGTCAAGGCGCGTGACAGGCCACAGCGTGCTCTAGCG
GCCCGGAAGCATCCACGACGCAGCTGATGTTCTCCTCCTGGTTTTCATCATCAGGGGCTCGAGCGG
GAGCGGGGGGAGTGTGCAGTTGTGGGGTGACCCTACAGCGAAATATCTGACTTGAGCTCGTGTCGGACCATACACCGGTGATTAATCGTGGTTTATTATCAAGCG
TGAGCCACGTCCGCGACGAATTTGAGCAGCTCTGGCTGCCGTACTGGTGCTCCCT
GGCAAGCGACGATCTGCTCGAGGGGATCTACCGCCAAACAGCGCTGGCAAAACTCTGCTGG
TCGTTGGACGTAGACCATCACCAGACGCACGCTCCGAGCGCTCAGCGCCCGGGG
GTCCCCATCGCTGCCGAACCGGATCTGTGGGCAATCGCCGATACACCGCGCC
CACGCAGTGTGGGCACTCAACGCCCCTGGCACGCACCGAATACGCGCGGC
GTAAGCCGCTCGCATACATGCGCAGCGCTCCGAGCGCTCAGCGCCCGGG
GTCCCATCGCTGCCCAACCGCAGATCTGTGGGCAATCGCCGATACACCGCGCC
GATGGGCGACCGCAGTTACTCAGGCGCTCATGACCAAAAACCCCGGACATC
GCCCTGGAAAAAGGAATGGCTCCACTCGAGATCTCAGTACACACTCAGCCGACATCG
AGGCCGAGCTCGCGCGCCGCAACATGCCACCAGCGCGGTGGCCGTACAGCAGACCCAC
GTACAAAGCGGCTCGCGCCACCCTAGGGCGAAATTCGCACTGTTCTGATTCC
GTCAGGTTGTGGCGCTATCTTCCCAGCCCTCTATCGCGATCTACCTCGGACC
GAACGTGGACCGAATCGCCCCGACGCGATCATATGCGGAAGCCGACCCAC
GCCGAATTTCCGTGCAACCGACGTGTGCTCCGGACCCGCTACCGGAGCAAGG
TCCGCGCCATCGCCCAACAGCATTTGGCGTTGGATCACAAAACCAATCGCCGAC
TTGGCGGACGGATCGTGTTACAGGCGCCACACTCAGTGCGCCATGCG
GCCATCTCGCGGAAGGCCAGCGCAGCGCCACGGCGGCGAGCACAGTTGCG
GGGCGCGAAAGGCGCCGCTAGGAGGCAATGGCTGTCTGATCGCCGACGGCTA
CAGCGACGGCTACAGCGACGGCTACAACTGGCAGCCGACTGTCCGCAAAAA
GCGGCGGTGACCGCCGCTCGAATCAGCGACGCAGCGCTGGTGGTCCCGGAGCAGG
CACGTCCGCTCGGTGCGCAGGAACGCAGCGAGTGGTTCCGCGGACGAGG
CTGCACGCAGCAGCAGCAGCATCCGCCTATCATCAGACGACGAGGGCCACTCTTG
GCCGCAAGCCGCCAAAAACATTTCGGGGCTGCATCTGGACACCGTTAAGCGACTC
GGCTATCGGGGGGTTCCGCGTGGGGGGTTCCGTTGCAACGGGTCGGACAGGTAAA
AGTCTCTGCTAGACGTAGTTTCTGTGTTTGGCAGCATGCTTCTCCTGTTGCGTG
TTTCTGGTTCGTCCCTTTTTGAATACACAGCCACGAGACCGAGGCTGGTCTACGAATC
TTGGTCGATACCAAGCCCATTTCCGCTGAATATCGTGGAGCTCACCGCCAGAAT
CCGTGGCTGCTAGTACGCTGCGAACTCGTTTGTAGTCTGTTTTGTGGGCA
TCCGTGCGCGGCGGCTCTAGAAGCTAGTGGATCTGACCCGCTCCACAACGGC
TGAGCAGGGTGCGACTGACGACGAGGCGACCGATACGAGGCGGATTTTGAACAGG
TTCACTCGAAGCCTCCGCTGGGTTCGGAGCATCGTCGATCACATGAGGGTGCGT
CTGACACCGATTCGACTATGCACAGTAGTAGGTAGCCGCCGCG
TCGGCGCCGTCGTGGCGTCGCGATGGTTGGCGC
CAGACCCGCCTGAGATCGATCGATCGATCGAGGCATCGGCAG
GACGGGCATTCCGCTCCTGGCCGCGCCTGCTGCGAGCTGAG
GACGTCGCGCGCGCTCCGGGAGCATCGTGAGTCCTGGCGTCCG
CAGACCCGCCTGAGATCGATCGATCGATCGAGGCATCGGCAG
GCTTGGTCCAGCCGTAGTGAGACGGGATGTCTTCCCGGTTGTCGGGTGGGTG
GGGCGGTGACATGTGTCGTGGCCTGGGGGTGTCGGGGGTGTCGAGCTGCTCG
GGCAGCAACCGGCTGGACGATGCACGTCGCTTGCACCTCCTGTGAGGGCGGG
TTGAGCCGCATGCGCCCTGACCTGCACCGTTGGACCGCAGTCGTTAGGGGG
CGTGTAGCATGGAGTGACGAAAGGGGGCCTAAATCAGATGATCC
>pENTR2B/TREPitt

Gateway entry vector with optimized Tet-On cis-acting sites. This was a gift from Dr. Reiser.

CTTTCCCTGCATTATCCCCTGATTCTGTGGATAAACCGTATTACCGCTAGCATGG
ATCTCGGGGGACGTCTAACTAAGCGAGAGTAGAGGAACTGCCAGGCATCAA
ATAAAAACGAAAGGCTCAGTCGGAAGACTGGGCCTTTCGTTTTATCTGTGTTT
GTCGGTGAACGCTCTCCTGAGTAGGACAAATCCGCGGGAGCGGATTTGAAC
GTTGTGAAGCAAGCAGGGGAGGCTGCACGAGCGGCCGCGCCATCAA
ATACCAAAACTAAGCAGAAAGGCCATCTCAGTCGGGACCCCGCCGACCGGACCTGC
TTCTACCAACTCTTCCTGTGGTATGTTAGGTTACTCTGTTACACCGGAGTCG
ATTTTTTGTATAGTAGTGACCTGTTCCTGCTTGCAAACAATGTAAGTACG
CTTTTTAAATGCGAAACTTTGTAACAAAAAGGACGGCTGGCGCCGCGGACCA
GTGCGACAATCTATCGCTTTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTCTG
AAACATGGCAAAAGGTAGCGTTGGCCAATGATGTTACAGATGAGATGGTCAGAC
TAAACTGGCTGACGGAATTTATGCTCTTTCCGACCATCAAGCATTTTATCCG
ACTCCTGATGATGCTGTTACTCACCACCTGCGATCCCCGGAAAAACAGCAT
TCCAGGTATTAGAAGAAATATCTCTGATTACAGGTGAAATATTGTTGATGCGCTG
GCAGTGTTTCTGCGCCGGTTGCATTCTCAGTGTTTGTATATTGCCTTTTTAAT
AGCGATCGCTGTATTTTCGAGGTATTTTGATGACGAGCGTAATGGCTGGCCTG
GTTGAACAA
GTCTGGAAAGAAATGCATAAACTTTTGCATTCTCACCAGGATTACGTGTAAC
TCATGGTGATTTCACCTTGAAACCTTTATTTTTGACGAGGGGAAAATTAATAG
GTGTTATTGATGTTGAGCGTGATCGGAATACGAGAGCGATACCAGGATCTTGC
CATCCTATGGAACCTGCCTGGTGAGTTTTTTCATCTTCCATTACAGAAACCAGCTTT
TTAAAAATATGGATATTTGATAATCTCTGATATGAAAAATTTGGCAGTTTTCTAT
ATGCTCGATGAGTTTTTCTAATCAGAATTGGTTAATTGGTTGTAACATTATTC
AGATTGGGCCCCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGA
TCTTCTTGAGATCCTTTTCTGCGCGTAATCTGCTGCTTGAACCAACAAAAAA
ACCACCGCTACCAACCGGTGGTTTGGTGCCCAGATCAAGAGCTACCAACTCTTT
TTCCGAAGTTAACTGGCTTACGACAGCGACGATACCAAATTACCAACTGTTTTCTA
GTGTAAGCCGTAGATTAGGCCACCCCACTTCAAGAAAACACTCTGTAACCCGCTAC
ACCTGCCTCCTGCAATCTCTGTAACCTGCGCGCTGGCTGAGCTGACACCG
TGCTTTACCGGGTGTTGACCTAAAGAGGCATACGTTACCGGATAAGGCAGCGGCT
CGGGGTGAACGGGGGGTTCGTGCACACAGCCAGCTTGGAGCGAACGACCTA
CACCGAAGTGGATACCTACACGCGTGAGCTATGAGAAAGCAGCCACGCTTTCC

GAAGGGGAGAAAGGCGGAGCTCTTAGGGATCGGAGGTCGGAACAGGA
GAGGACCACGAGGAGCTTGACGACGGGAAACGCCTGATCTTTATAGTCCTG
TCGGGTTTCGACCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGG
GGGCGGAGCCTATGGAAAAACGCGCCAGCAACGCGGCTTTTTACGGTCTCTG
AGCCTTTTGCTGGCCTTTTTGCTACATGTT

>pSLIK-zeocin

From Dr. Fraser via Addgene (25736).

GTGACGCGATCGGGAGATCTCCGGATCCCCTATGCTGACATTCACTCTAGTAC
CTGCTCTGATGCACTCGCTTCTTGTGTTG
AGGTCGCTGAGTATGCGCGAGCAAAATTTAAGCTACACAAAGGCAAGGCTT
GACCGACAATTGCATGAAGAATCTGCTTAGGTTAGGCGTTTTGCGCTGCTTC
GCGATGTACCGGGCCAGATATACGCCTTGACATTGATTATTGACTAGTTATTAA
TAGTAATCAATTACGGGGTCATTAGTCTACAGCCATATATGGAGTTCCGCGT
TACATAACTTACGGTTAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGC
CCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTT
CCATTGACGTCAATGGGGTGGAGTATTTACGGTTAAGCTGCCCAACTGCCCCTGGCGCTG
CATCAAGTGTATCATATGGCAATCGCCCCCTATTGACGTCAATGACGGTA
AATGGCCCCGCTGACATTATGGCCAGTACATAGCCTTTATGGGACTTTTCCCTACT
TGGCAGTACATCATTACGTAATTAGTCATCGCTATTACCATGGGTAGTGACGGTTTTG
GCAGTACATCAATGAGGGCGTTAGCAGGTTTTGACACTACGGGGATTTCACACT
CTCCCCCATTTGACGTCAATGGGAGTTTTTGCTGACCAAAATCAACGGG
ACTTTCCAATATGTCGTAACACACTCCGCACCCATTGACGCAAATGGGCGG
GCAGTGTCGCCGGTAGGTAACTAGCTATTACATGACGGGATTTCACACT
CTCTGTTAGACCAGATCTGAGCCTGGGAGCTCTGCTACTAGGGAACC
CAGTCTTAAGCCTCAATAAAAAGCTTGGCCTTGAGTGCTTCAAGTAGTGTGC
CGTCTGTTCGTCAGACTCTGGTAACTAGAGATTCGCGGGATTTCCAAGTG
GCAGCTGTTAATCTGGTCAGGTTGCTGGTTGCTACTAGTGCACC
GAGGGCTAGAGGAGATTCGCGGGATTTCCAAGTAGTGCACC
GAAATTAGATCGCGATGGAATCAGGGTTATCGGTTTAAAGCCATAGGGAAGG
AAACCAGAGGAGCTCTCTCGACGAGACTCGGGGTGGTGAACGCGACG
CAAGAGGGCGAGGCGACTGCTGGAACGCGGTGATGTACAGCCAAATTTTGAAC
GAGGCTAGAGGAGAGATGGTGGCGAGAGCGTCACTATTAAGCGG
GAGAATTAGATCGCGATGGGAAAATTCTGGTAAAGCAGGGAAAGG
AAATATATATAATAGTATGGGCAAGCGAGAGATATTTTGACTACGT
AGTTAATCTGGGCCTGTTAGAAACATCAGAAGGCTGTAAGCACAAATTACTGG
AAAATTTTCGGTTTTATTACAGGGACAGCAGAGATCCAGTTTGGTTAATTCTA
GACAGTTTATAACAAACAGTTTGTACAAAAAGCTGAACGAGAAACGTAAAATG
ATATAAATATCAATATATTTAAGATATTTTCGATAAAAAACAGACTACAATA
ATACTGTAAAAACACACATATCCAGTCAGCAGGAGGGGCGCATTAGGCACCC
CAGGGCTTTACACTTTATGCTTCCGGCTCGTATAATGTGTGGGATTTTGAGTTAG
GATCCGGCGGAGATTTTCAGGCTAAAGGAGCTAAATGGGAAAAAATGAC
ACTGGATATACCACCCTGATAATATCCAAATGGCAGTAAAAGCATATTG
GGCATTTTCCAGTGCTGCTAAATGCATTATACGAGCAGGTTCAGCTGATA
TTACGGGCTTTTTAAAAAGACCGTAAAAGGAAGAAAAAAGCACAAGTTTTTATCCGGG
CTTTATTCCACTTTTGGCCCGCTGATGAATGCTCATCCGGAAAATCTCGGTATGG
CAATGAAAAAGCAGGGAGCTGTTGAGGATATGGGAGATGTCACCCCTTGTACAC
CGTTTTCCATGAGCAAACGCTAATTTTTCTGCTCTGGAGTAAGATACCACG
ACGATTTCGCGCAGTTTCTACACATATATTCGCAAGATGTGCTGCTTGTTACGT
GAAAAACCTGGCTTATTTCTCATTAAAAAGGGTTTATTGAGAAATATGTGTTTTCGCTC
AGCCAATGCCCTGGAAGATTTTCCAGATTGTTGATGGGCTAATGTGCAG
ACAAACTTCTTCCGCCCCGTTTTACCCATGGGCAAATATATTACGCAAGGCGAC
AAGGTGCTGATGGCGCGCTGGCGATTTGTCATCATGCTGGCTGTGGCTTTT
CCATGTCCGCGAGAATGTTAAAATGAGATACAGTACTGCGAGATGGGCAAG
GGCGGGGCGTAAAAGATCTGGATCCGGCTTACTAAAAGCCAGATAACAGT
TGACAGCGACAGCTATCAGGCTGCAAGGCGATATGAGTGCTAGATGTGCAAAATG
TGACAGCGACAGCTATCAGGCTGCAAGGCGATATGAGTGCTAGATGTGCAAAATG
TGACAGCGACAGCTATCAGGCTGCAAGGCGATATGAGTGCTAGATGTGCAAAATG
TGACAGCGACAGCTATCAGGCTGCAAGGCGATATGAGTGCTAGATGTGCAAAATG
CTGGAAAGCGGAAAATCAGGAAGGGATGGCTGAGGTCGCCCGGTTTATTGAA
ATGAACGGCTCTTTTGCTGACGAGAACACAGGGACTGGGTAAATGCAAGTTTAAG
GTTTACACCTATAAAAGAGAGAGCCGTTATATCTGCTCTGGGGATGATGGTACAG
GTGATATTATTGACACGCCTTCGCTGACGAGAACAGGGACTGGTGAAATGCAGTTTAAG
GTTTACACCTATAAAAGAGAGAGCCGTTATCGTCTGTTTGTGGATGTACAGAT
ACGTCTGCTGTCAGATAAAAGTCTCCCTGTTAAGTTTACCCGGTGGTGTCATATCG
GGGATGCAAAGCTGGCGCATGATGACCACCGATATGGCCAGTGTGCCGGTCTC
CGTTATCGGGAAAGAAGTGCTGCTGCTCTCAAATGACATCAAA
AACGCCATTAACCTGATGTGTTCTGGGGAAATATAAAATGTCAGGCTCCCCTTTATACA
CAGCCAGTCTGCAAGTCAGCACCATAGTGCAGTGCATATGTTGTGTTTTACAGATAT
TATGCTGCTGTTTTTAATGCAAATCTAAATTATATATTGATATTATATAC
TTTTACGTTTCTCCTCAGCTTTCTTGTACAAGATGTTGTTGATAACATCTCGAGC
CATAAAGATGGTTAAATACACCCACCAAGGATCTGCGCTCCCAGCAGGCTGC
GCGCCTCCGCAGGCGGCGCCCCCTCCTACGGCAGCGCTGGGCACGAGCG
AAGGGCGACAGCAGCGCTCTGATCCTCTCCGGGCCAGGACGCTACAGAGCGGC
CCGCTGCTCATAAGACTCGGCCTTTTGAGAAGCCAGAGAGAGAGAGACGCTG
TTAGGACGGACTTGAGTTGACTCTAGGGCAGTGGTTTTCTTTCCAGAGAGCG
GAACAGGCCAGGAAAAGTACTCCTCTTCTGGAGATCTGCGGAGGGATCTCC
GTGAGGCGCTGAAACCGGGATTATATATAAGGAGACGCAGCCGGTTGTCGACAG
CTAGTTCGCCAGGCAGGATTTTGAGGTCCGGCTCTGGGGCTTCTCTTGATCGCTG
TGATCGTGACTTGGGTTGAGTACGGGGCTGGCTGGGCTGCTGGCAGAGACG
CCGCCGGCCTGCTGGAGGGTACGGGAAGCGTGTTGAAACTGGGAGGTTGGGAGCG
GTAGTCTGGGTCCCGAGCAAGGTTGCCCTGAACTGGGGGTTGGGGAGCG
CAGCAAAAATGCGGGCTGTCTCCCGAGTCTTGAATGGAAGACGCTTTGTGAGGGCG
ACCAGTTGAGCAAGAGAAGGTAGAAGAAGCCAATGAAGGAGAGAACAC
CTTGTTACACCCCTGTGAGCCTGACATGGGATGGATGACCCGGAGAGAGAAGTA
TTAGAGTGGAGGTTTGACAGCAGCCCTAGCATTTTCACTACATGGGCGAGAGGC
TGATCCGGACACTGTACTGGCTCTCTCTGTGTTAGACCAGATCTGAGCCTGAG
CTCTCTGGCTAAGCTTAAAGCCTCAATAAAGCTTGCTTTG
AGTGCATTCAAGTAGTGTTGAGCCTGAGCCTGAGAGAAGTA
CCCTCAGACCCTTTTTAGTCAGTGTTGAGAAGATCTCAGAGGCCAGCCGTTAACC
CCGCTGATCAGCCTCGACTGTGCCCTTCTCTAGTTGAGAGCCATCTGTTGTTGGCC
CCTCCCCCGTGCCTTCCTTGACCCTGGAAAGGTGCGACCTCCCACCCTGTCTTCTCT
AATAAAAATGAGGAAATTTGCATCGCATTGCTGTTAGGTGTCATTCTATTCTG
GGGAGGTCGGGTTGGGCAGACAGCAAGGGAGATTGAGAAAGCAATAGC
AGGCATGCTGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACAG
GCTGGGGCTCTTGTAGGGGTATCCATTGCTGAGGCGGAAAGACAAC
GCGGCCCTGCCTTTCTGGGGCTCTCTCCCTTCTCTCCTCTCCTTCTCCTCCAGTTCGCGCTTT
CCCCGTCAAGCTCTAAATCGGGGCTCCCTTTTAGGTTCCGATTTAGTGTCTTT
ACGCGCAGCTCGACCCCCCCTTTTTTTCTCCCTTTGGGCTTGGAGTCTCAGTCTTT
CCATCGCCCTGATAGACGGAGTTCTTCTCCTTCTGGGAGGTAGTGGAGG
CCATCGCCCTGATAGACGGAGTTCTTCTCCTTCTGGGAGGTAGTGGAGG
AATAGTTGGACTCTTGTCTTCCAACACTGGAACACACTCAAACCTATTCGGTCTA
TTCTCTTCTTTTTATAGGGGATTTTTGCCGATTTTGGGCTATTGTTAAAATGA
AGCTGATTTAACAAAAATTTAACGCAGAAATTAATTCTGTTGGAATGTTGTCAGT
TAGGTTGGAAAGTGCCCCAGGCTCCCGAGCAGCAAGATGATGCAAAGCAT
GCATCTCAATTAGTCAGCAACAGGTTGTTGGAAGGTCCCGCTCCCCAGCA
GGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCGCC
CCCTAACTCCGCCCATCCCCCCTAACTCCGCCCATTTCCGCCCATTCTCCG
CCCCATGGCTGAC
CTCTGAGCTATTTCCAGAAGTAGTGAGGAGGCTTTTTGGAGGCCTAGGCC
TTGACAAATATCATCGGCATAGTAGATATACGACAGG
GAGGACTAAACCATGGCCAAGTGTGACCAGTGCCGTTCGGTGCTACCCCGC
CGCGACGTGGCCGGAGTGGAGTGCTGGACCGAGCAGTTTCTGCTGCCGAGCG
GGGACCTTCTGTGAGGACTTTCGCCGGCTGGTGTGGGAGGACGACCCT
GTTTATCATGACGCCGGTCCAGACGACCAGTGGTGGTACGCAGCCGC
CCCGCCCTTCTATGAAAGGGTGGGCTTCGGAATCGTTTTCCGGGACGCCGGCT
ATGGATCTTCTCAGCAGCGGGATCTCATGCTGGAGTTTCTTGGCCACCCCACCTT
GTTTATTCAGCTTTATAATGGTTAACAAATAAAGCAATAGCATCAAAAAATTCA
CAAAATAAGCTTTTTTTACTGCCATTCTAGTTGTGGTGGTCTCAAACTATGC
AATGTATCTTTATCATGTCTGTATACCGTCGACCTCTGCTAGCTAGCTTGGCGTA
ATCATGTCATAGCTGTTCCTCTGTGAAATTGTTATCCGCTACAATTCCAC
ACAACATACGAGCCGGAAGCATAAAGTGAAAGCCTGGGGCTGCTAAGGAGGT
GAGCTAACATCAAATTGCGTTGCCACTGCCCTCCCCTTTCGAGTGCCGGAA
ACCTGTCGTGCCAGCTGCAATTAATGAAATCGGCCAACGCGGAGGAGCGG
TTTGCAGTTTGGGGCGCTCTTCCGCTTCCTGCTCACTGACTGCCTGCGCTCGGT
CGTTCGGCTGCGGCAGCGGATACAGCTCACTCAAGGCGGTAATACGGGTAT
TCCACAGAATCAAGGGATAACGCAGGAAAAACATGGTGAACAAAAAGGCCAG
CAAACCGAAGGACTATAAAAGATACCAAGGGCGTTTCCCCCTGGAAGCTCCCTC
GTGCCGCTCCTCCTGTGGTACCCGCTGCCAATACCTGCTCAACTTGCCGCTTTTCTC
CCTTCGGGAAAGCGTGCCTGCTTTTCTCATAGCTCAGCTCAGCTGTAAGGTGATCTCAGTTTC
GGTGTAGGTCGTTGCTCCAAGCTGGGCGTGTGTCGAAGAACCCCTGGGTTTAC
CCCCGACCCTGGCCGCTTTATCAGGTAATACTACGTCTTGTAGCATCCAAACCCGGAAG
ACACGACTTATCGCCACTGGCAAGCCACTTGGAACAAGGATTACGAGGACGC
AGGTATGTAGCCGGCTGCTACAGAGTTCAGGAAGTGTTGGGCCTAAGACTACGGCT
ACACTAGAACAGATATTTTTGATATCTGCGCTCTGCTGAAGGCCATACCTTCC
GGAAGACAGTTGTTGAGCTCTTGTGCCGCTGAAGCTGGGCAACAAAAACAGGCTGGTAGCG
CTGTTTTTGTGGTCAAGCAGAGTTACGCGCAAAAGGATCTCA
AGAAGATCCTTTTGTATCCTTTTCTACGGGAGTCGCTCAGTGGAAACGAAAAC
CAGTTAAGGGATTTTGGTCATGAGATTATCAGAAAAGGATCTTCTACCTAGATC
CTTTAATTTTAAATGAAAGTTTTAAAAATCAATCTAAAGTATATATGTGTAACAC
TTGGTGCTGACAGTTACCAATCATTACATGAGGCACCTATCAGCTACCGAGATC
GTCTATTCTCGCATCAGTGGCTGTGGACTTCCCCGTGCTGGTAGATAACTACG
ATACCGGAGGGCTTACCATCTGGCCAGCTGCTAATGATTACCCGGCGAGACC
CACGTCACCAGGGCTCCAGATTTATCAGCAATAAAAACCAGCCAGCCGGAAAGGGC
CGAGCGCGAGAAGGTGCTCTGCAACTTTATCCGCTCTCCATCCATGCTTATAATT
GTTGCCGGGAAGCTAGAGTAAGTAGTTTCGCCAGTTAATAGTTTTCGCAACG
TGTTGCCCATTTGTACAGGCATCGTGGTGTCACGCTCTCGTTGGGATTGGCTT
CATTCACTCCGGTTCCAAACGATCAAGGGCAGTTACATGATCCCCCATGTTG
TGCAAAAAAGCGGTAGCTCCTCGTGCTCCGATCCTTGGGATTCAAAAGTCTA
GAGCTGTTATACACTCATGGTTATGGCAGCACTGCATAATTCTCTACTG
TCATGCCATCCGTAAGATGCTTTTCTGTGACTGTTGAGTACTCAACCAAGTCA
TTCTAGAAATAGTGTATGCGGCG ACCGAGTTGCTCTGGCCTGGGCGTCAATACG
GGATAAATACCGCGCCACATAGCAGAACCTTTAAAAGTGCTCATCATTTGAAA
CGTTCTTCGGGGGAGAAAAACTCTCAAAGGATCTTACCCTGCTGTGAGATCCAGTTC
GATGTAACCCACTCGTGCACCCAACCTGATCTTACGCACTTTTTACTTTCAACCA
GCGTTTCTGGGTGAGCAAAAAACAGGAAAGGCAATGCCGCAAAAAAGGGAA
TAAGGCGACACGGAATGTTGAATACCTCAACACTCTCTTCTTTTTCAATATTAT
TGAAGCATTTATCAAGGTTATTGTCTCATGAGCGGATACATATTTGAAATGTAT
TTAGAAAATAAAACAAATAGGGGGTTCCGGCAGCATTTCCCGAAAAAGTGCCA
CCTGAC
>pHR’ CMV 8.2ΔR

Helper plasmid for packaging lentiviral clones in HEK293T. This plasmid was a gift from Dr. Rai.

GGGCTGCAGGAATTTCAGCTCGCCCGACATTGATTATTGACTAGTTATTAATA
GTAATCAATTACGGGGTCATTAGTTCCATAGCCCATATATGGAGTTCCGGTTA
CATAACTTACGTTAAATGCGCGCCTGGGCTGACGCTGACCCCCGGCC
ATTGACGTCAATAATGGAGTGTGGAGTTACGGTAAAGCTTGGCAGTACAT
TCAAGTGTATCATGACATGCGTTATGTGACTTGGCAGTACATGACGT
GGCCCGCTGGCATTATGCCCAGTTATGCCAGTACATGACCTTATGGGACTTTCC
CAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTGGCA
GTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCACCCCAT
TGACGTGTGGAGGTCTATATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCTTG
GAGACGCCATCCACGCTGTTTTGACCTCCATAGAAGACACCAGGGACCG
ATCCAGCCTCCGCGGCGGGAAACGCTGCATTTGAAACGCGGATTCCCCGTGCC
AAGAGTGACGTAGTACCAGCTATAGAGTCTATAGGCCACCCACCCCTTGGCTTCTTATGC
ACGTAGACAAATACTGGGACAGCTACAACCATCCCTCCAGACAGGTGAG
GCTGTAGACAAATACTGGGACAGCTACAACCATCCCTCCAGACAGGTGAG
AGAAGCTTAGATCATTATAATACAGTAGAACCCTCTATITGTGTGAGATTAAAA
GGATAGAGATAAAAGACACCAAGGAAGCTTTAGACAAGATAGAGGAAGAGCA
AAAACAAAAGTAAGAAAAAGCACACAGCAAGCAGCAGCTGACACAGGACAC
AGCAATCAGGTTCAGCCAAAAATTACCTATAGTCAGAACATCCAGGGGCAAA
TGTTACATCAGGCCATATCACCTAGAAACCTTTAATGATGGGCTAAAAAGTAGT
AGAAGAGAAAGGCTTTTCAGCCCAGAAGTGATACCCCATGTTTTTCAGCATTATCA
GAAAGAGCCACCCCCAAGATTTAAACACCATGCTAAACACAGTGGGGGGA
CATCAAGCAGCCATGCAATGTGTTAAAAAGAGACCAATCAATGAGGAAGGTCA
GAATGGGATAGAGTGCAATCCAGTGCAATCGCAAGGGCTATTGCACCAGGAGAGA
TGAGAGAAACCAAGGGGAAGTGACATAGCAGGAACTACTAGTACCCCTTCAGG
AACAAATAGGATGGATGACACATAATCCACCTATCCCCAGTGGAGAAATCTA
TAAAGAGATGGATAATCTCTGGGATTTAAATAATAGTAAGAATGTATAGCCCT
ACCAGCATTCTGGACATAAGACAAGGACCACAAAGGAAACCCCTTTAGAGACTATG
TAGACCGATTCTATAAAACTCTAAGAGGAGCGACAAGCATAAGAATACATAGCAG
ACTTTTTTTAAAAGCATTTGGGACCAGGAGCGACACTAGAAGAAATAGTGGACAG
CATGTCAAGGGAGTGGGGGCAACCAGCCATAAAAGCAAGAGTTTTTGCTGAAAGC
AATGAGCCAAGTAAACAATCTAGCTACCATAATGATACAGAAATGCTAACAGAG
ATGCGCGAGAAAGACTGTAAGTGTCTTCAATTGTTGCAAAAGAGGACAC
ATAGCCAAAAATTCGAGGGCCCTAGGAATAAGGGCTGGTGGAAATGTGGA
AAGGAAGGACACCAATGAAAGATTGTACTGAGACAGAGCTAACTTTTTTAG
GGAAGATCTGGCCCTCCCACAAGGGAAGGGAAGCCAGGGAATTTCCTCTCAGAGCAG
ACCAGAGCACAAGCCCCACAGAGGAGAGCAGTTAGGGGGAAGAGAC
AACAACTCCTCTCAAGAAGGAGCAGCAGGAGATGATACAGATTTTA
GCTTCCCTCAGATCAGCTTGGCAGGCACCCCTCGTCACAATAAGATAGG
GGGGCAATTAAGGGAAGAGCTCTATTAGATACAGGAGAGAGATGATACAGATTA
GAAGAAATGAAATGGCCAGGAAGATGGAAACAAATAAGGATAGGGGGAATT
GGAGGTITTTATCAAAGTGGACAGTATGGATACAGATACCTATAGAATATCTGC
GACATAAAGCTATAGGTATCATAGGACCTACACCTGTGACAAACATAT
TGGAAGAAATCGTTGACTAGATTTGTGCTGCACCTTTAAAATTTCCTACTAGTC
CTATTGGAGACTGTCAGATTAAATTAAGGAGGAGATGATGAGCCAAAAGT
TAATAGGAGCCATGGACAGAAGAAAAAAAAATAAGGACATTAAGTAGAAATTGT
ACAGAAATGGAAAAAGGAAAGGAAAAATTTCAAAAAATGGGCGCTGAAATCCA
TACAATACCCAGTATTTGCCATAAAGAAAAAAGACAGTACTAAATGGAGAA
AATTAGTAGATTTCCAGAGAACTTAATAAAGAGAACTCAAGATTCTGGAAGTG
TCAATTAGGAATAACCATCCTGCAAGGGTTAAAAACAGAAAAATCAGTAAACA
GTACTGGATGTGGGGGAGATGCTATTATTTCAGTCTCCCCTTAGATAAAAGACTTCAG
GAAGTATACTGCATTACATACCTAGTATAAATAATGAGACACCAGGGATT
AGATATCAGTACAATGTGCTTCCACAGGGATGGAAGAGATCACCAGCATAAT
TCCAGTGTAGCATGACAAAAATCTTTAGAGCCTTTTAGAAAAACAAATCCAGA
CATAGTCATCTATCAATACTGGATGTATTTGTATGTAGATCTGACTTTAGAAA
TAGGGCAGCATAGAAACAAAAATAGAGGAACTGAACATACATCTGTTGAGGT
GGGGATTTCACCACACCAGACAAAAACATCAGAAAAAGACCTCCATTCCCTTG
GATGGGGTTATGAAAACTCCATCCTCTGATAAAAAAGGAGATCAGTACAGCCTATAGTGCTG
CCAGAAAAAGGACAGCTGGACTGTCAATGACATACAGAAATATTAGTGGGAAA
TGAATTGGGCAAGTCAGATTATGAGCAGGATAAAAGTAAGGCAAATATATGTA
AACCTCTTAGGGGAACAAAGCACTAACAGAAGTACTGACACTAACAGAAG
AAGCAGAGCCTAGAACTGGCAGAAAAACAGGGAGATTCTAAAAAAGAACCGGTAC
ATGGAGTGTATATGACCCCATCCAAAAGACTTAATAGCAGAAAAATACAGAAGCA
GGGGCAAGGGCGATGGACTTCAATTTATTTATCAAGAGCCATTTAAAAATCTG
AAAACAGAAAAATATGCAAAGATGAGGGTGCCCACACTAAATGATGTGAAA
CAATTAACAGAGGCGATACAIAAAAATAGCCACAGGAAAGCATAATAGG
GGAAAGACTCCTAAATTTAATTTACCACATAAAAGGAAAACATGGGAAGCAT
GGTGGACAGAGTATTGGCAAGGCAAGCTGGAACTGAGTCTGGAGTGGGAGTTTGTCAA
TACCCCTCCTTAGTGAAGTTATGGTACCAGTTAGAAAGGAAAACCCATAATA
GGAGGCAAAACTTTCTATGAGATGGGCGAGCCAAATAGGGAAAATCTAATTAG
GAAAAGCAGGATATGTAACTGACAGAGGAAGACAAAAAGTTGTCCCCCTAA
CGGACACAACAAATCAGAAGACTGAGTTACAAGCAATTTCATCTAGGTTTGCA
GGATTCGGGATTAGAAGTAAACATAGTGACAGACTCACAATATGCATTGGGA
ATCATTCAAGCACAACCAGATAAAGATGGAATCAGAGTTTAGTCAGTCAAATAA
TAGAGCAGTTAAPAAATAGCTGCACAGCTGAGGTACAGTCAAATAA
TAGAGCAGTTAAPAAATAGCTGCACAGCTGAGGTACAGTCAAATAA
AAAGGAAAAAGTCTACCTGGCATGGGTACCAGCAC
ACAAAGGAATTGGAGGAAATGAACAAGTAGATGGGTTGGTCAGTGCTGGAA
TCAGGAAGTGCTACTATTTTTTAGATGGAATAGATAAGGCCCCAAGAAGACATGA
GAAATATCAGTAATTGGGAGAGCAATGGCTAGTGATTATTTAACCCTACCACCT
GTAGTAGCAAAGAAATAGTGGCCAGCTGATAAAATGTCAGCTAAAAGGG
GAAGCCATGCAATGGCAAGATAGACTGTAGCCCAAGGAATATGGCAGCTAGATT
GTACACATTAGAAGGAAAAAGTTATCTTGAGGCAGTTTCATGTAAGGAGGTG
ATATATAGAAGCAAGAATATCCAGCAGCAGAAGGGCAAGAGAAACAGCATA
CTTCTCTTTAAAAATTAGCAGGGAAGATGGCCAGCTAAAACAGATACATACAGAC
AATGGCAGCAATTTCACCAGTACTACAGTTAGCTAAAGGCGCTTTTGTTGGGCGG
GGATCAAGCAGGAAATTTGGGACTCCCTACAAATCCCAGATGCAAGGAGTAAT
AGAATCTATGAATAAAGAAATTAAGAAATATAGGACAGGTAAGGAGATCA
GGCTGAACATCTTTAGACAGCAGTCTACAAATGAGGACAGTTATTCTTCCACAAT
AAAAGAAAAGGGGGATTTGGGGGTACAGTGCAGGGGAAGAATAGTAGAC
ATAATAGCAACAGCACATACAAACTAAGAATTACAAAAAACAAATTTACAAAA
ATTAAAAATTTTCCGGTTTATCAAGGCACACAGCATTCAGTTTGGAAG
GACCAGCAAGCTCTCTGGAAGGTTGAAGGGGCAGTGTAGTAATACAAGATA
ATAGTGACATAAAAAGGTAGTGCAAGAAGAAGGAAGAGTACATACAGGGATT
ATGGAAAAACAGATGGCAAGGTGATGATTGTGTTGGCAAGTACAGGATGAGG
ATTAACACATGGAAAAAGATTAGTAAAAACACCATATGTATATTTCAGGAAAG
CTAAGGACTGTGTTTTATAGACATCATACTATGAAAGTACTAATCCAAAAATAAG
TTCAGAAGTGACACATCCCACTAGGGGATGCTAAATTAGTAATAAACAACATAT
TGGGGTCTCGCATACAGGAGAAGAGAGACTGGCATTGGGTCAAGGAGTCTCCA
TAGAATGGAGGAAAAAGAGATATAGCACAACAAAGTAGACCCCTGACCTAGCAG
ACCAACCTATTAGCTCGACTATTACATTGTTTTCTCAGAATCTGCTATAAGA
AATACCATATTAGGCAGTATAGTTAAGTCCTTGAAGTGAATATCAAGCAGGAC
ATAACAAGGTAGCATCTCTACAGTACTTGGCACTAGCAGCATTAATAACCC
AAAACAGATAAAGCCACCTTTTGCCTAGTGTTAGGAAACTGACAGAGGACAGA
TGGAAACAAGCCCCAGAAAGACCAAGGGCCACAGAGGGAGCCATACAATGAAT
GGACACTAGAGCTTTTAGAGGAACCTTAAAGATGGAAGCTGTTAGACATTTTCC
TAGGATATGGCTCCATAACTTAGGACAACATATCTATGAAACTTACGGGGAT
ACTTGGGCAAGATGGAACCCATTATAAAGATCTGCAAACAACGTCGTTTTA
TCCATTTCAGAAATTGGGTGCGACATACAGAATAGGCGTACTCGACAGAG
GAGAGCAAGAAATGGAGCCAGCTAGATCCTAGACTAGGACCCTTGGGAAGCATC
CAGGAAGTCAGCTAAAACCTGCTTGTACCAATTGCTATTGTAAAAAGTGTTG
CTTTCATTGGCCAAGGTGTTTCTAGACAAAAACGCCTTAGGCATCCTCTATGGCA
GGAAGGAAGCCGAGACAGCGACGAAGCTCACTCAGAACAGTCAGACTCATC
AAGCTTCTCTATCAAAAGCAAGTAAGTAGCTACATGTAATGCAACCTATATAAGT
AGCAATAGTCATTAGTAGAATAATAATAGCAATAGTTGTGTTGCTCC
ATAAGAATCTAGAATATAGGAAAAATATTAGGACAAAAGAAAAATAGACAGG
TTAATTGATAGACCTATAGAAAGAGCAGAAGACAGTGCGCAATGAGAGGTGAA
GGAGAAAGTCAGCAGCTTGTGGGAGATGGGGGTGGGAATTGGGCACCATTGCTC
AGGATTTTGCTATAAGATGGGTGGCAAGTGGTCAAAAAGTAGTGTGATTGGA
TGGCCTACTGTAAGGGAAGATGAGACGAGCTGAGCCAGCAGCAGATGGG
GTGGGAGCAGTATCTCGAGACCTAGAAAACATGGGAGCAATCACAAARGTAC
AATACAGCAGCTACCAATGCTGATTGTCCTGGCTAGAAGGACAAGAGGAG
AGGAGGTGGGTTTTCCAGTCACACCTCAG
AAGCTTAATTTCACTCCAACGAAGACAAGATATCTTTGATCTCTGTGATCTACCAC
ACACAAGGCTACTTCTCCGTAGGCGAGAAACCTACACACCAGGGCCAGGGATCA
GATATCCACTGACCTTTGGATGGTGCTACAAGCTAGTACCAGTTGAGCAAGA
GAAGGTAGAAAGGAAGCCAATGAAGGAGGGAGAACACCGCTTTGGTACACCCCTGT
GAGCCTGATGAGTGGATGGCCAGAGAGAAGTCACCTCAGAGAGTGAGGTTT
GACAGCCGCCTAGCATTTTCATCACATGGCCGCAGAGCTGCATCCGGAGTAGCT
TCAAGAAGCTGTGAGGGCCGGCAGCCGGGCTGACCTTCAGACCTTTGGCACTGGA
GTGGCCCAGCAGACGCGGCATCGTGAGTCAGTGCACGACCAGCATCTGCT
CTCTCAACCTGGAAGAACACTACTGCAACTAGGCCCACCACTACCCCTGTCCAC
CCCTCTGCAATGAAATAAAACCCTTTGAAAGGACACTACAAGTTGTGTGTACAT
GCAGTGACTGTGCTATGTTGGC GGGA AACATGAGTGGGGCTGGCTGGAGTGGCGATGATAAGCTGTCAAACATGAGAATTCTTGAAGACGAAAGGGGCTCG
TGATACGGCTATTTTTTATAGGTTAATGTCACTGATAAAATGTTTTTCTTAGTCA
GAATTAATTCCGTTATTTCTCTAGTGTACCTAAATCGTATGTTGTTGATGATACA
TAAGGTTATGTTATTAATTGTAACGCGGCGTTCTAACGACAATATGTACAAAGCTTA
ATTGTTGACTCCTGGCTTACGTAACGACACCCTATCATCTCTCTCGTAACACT
GCCGTAGAGCGGTTTGGATGGGACGAACCTTTCTGAGTTTTCTGGAACGCGGT
TCCGACCCCGGAAATGGTCAGCGAACCAATCAGCAGGGTCATCGCTAGCCA
GATCCTCTACGCGACGTTCGTTGGCCGCATCACCAGGCCGACAGGTTCG
GTTGCTGGCGCCCTATATCGCGACATCACCGATGGGGAAGATCGGGCTCGCC
ACTTCCGGTCTCATGAGCGCTTTTGGTCGCGTGGGTATGGTGGCAGGCCCCGTG
GCCGCGGGGACTGTTGGGCGCCATCTCCTTGCATGACCATTCCTTTGCGGCCG
GGTCTCAGGCGCTAACCCTACTACTGGGCTGCTTCCTAATGCAAGGAGTCG
CTACCTGATAGGAGTGGTCTACTCTCAGTACAATCTGCTCTGATGC
CGCATAGTTAAGCCAGCCCGGACACCGCGGCAACACCCGCTGACGCGCCCTGA
CGGGCTTGTCTGTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGG
GAGCTGCATGTGTACAGAGGTTTTTCACCCTGCTACCCGAACACGCAGCACGA
AAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCTAGTATAATAATGGT
TTCTTAGACGTCAGGTGGCACTTTTTCCGGGAAATGTCGCGGAAACCCCTATTT
GTTTATTATGCTAAATACATTTAAAAATATGTATCCGTATGAGACAAATACCCC
TGAAAAATGCTTCAAATAATTTAAAAAGGAAGATGATGAGATTACACATT
TCCGTGTCCCTTTATTTCCCTTTTGCGGCATTTTGCTCTCTGTTTTTGCTCA
CCAGAAAAAGCTGGTGAAAGTAAAAGATGCTGAAAGATCAGTGGGTGCAGC
AGTGGGTACATCGAAGCTCGAAGGCTCAGCTTGGGATCATGTAACTCCTGGA
AGTGGGTACCCTCAGGCTGCTCAGCAGTCACGTTGGCGCCGATA
CACTATTCTCAGAATGGACTTTGTGGAGTACTCACCAGTCAAAAAGGACATCT
TACGGGATGGCATGAGTAAAGAATTATGCGACTGCTGCCATAAACATGAGT
GATAAACYCCTGACCCGCAACTTACTTCTGGAACACGATCGGAGGAGCAGGAGC
TAACCGCTTTTTTGCAACACATGGGGGATCATGTAACCTCGCCCTTGATCGTTG
GAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATG
CCTGTAGCAATGGCAACAACGCTGCGCAAACCTATTTAATCCTGTGAGTAAAT
AGGACCACTTCTGCAGTGACCTGCGGTTCTCCGGCTGCTGGATTTATGGCTGATTAAT
CTGGAGCCGCTGCTGAGATAGGTGATCTCCACTGATTAAGCC
ATTGGTTAATCTGAGACAACTTACTCATATATACTTTGAGTTGATTTAAAA
CTTCATTTTTAATTTAAAGGATCTAGGTGAAGATCTTTTTTGGATAATCTCAT
GACAAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTGACACCCGGTAG
AAAAGATCAAAGGATCCTTTTGGAGATCCCTTTTTTCTGCGCTGTAATCTGCTGC
TTGCAAACAAAAAACCACCAGCTACCAGCGGTGGTGGTGGTGGTGCACGATCAAG
AGCTACCAACTCTTTTTCCGAAGGTTAACTGGCTTTCCAGCACAGGCGAGATACC
AAATACTGTCCCTTTCTAGTGTACGCTAGTTGGCCACCACCCTTCAAGAACTCTG
TAGCACCAGCCTACATACCTCGCTCTGCTAATCTCCTGTATCAGTGGCTGCTGCC
AGTGGGCTAAGTGTCGTCTTTACCAGGTGGACTCAAGACGATAGTTACCGG
ATAAGGCGACACCGTGCGGCTG ATAAGGCGACACCGTGCGGCTG
GGAGCGAACCACCTACACCCGAACGTAGATACCTACAGCGTGAACGTGATGAGA
AAGCGCCACGCTTCGCCAAAGGGAAAGCGCACAGGTATCAGGCTGTAAGCGG
CAGGGTCTCAAGAAGGAGACGCAACGGAGAGGCTCTCCAGGGGGAAACGCCTG
GTATCTTTTATAGTCTCTGCGGGTTTTCGCCACCTCAGCTCTTGGAGCGTGGACTTTTT
GTGATGCTCGTCAGGGGGGGCCAGCTCATGAAACACACGCAACGCAGGC
CTTTTTACGGTTCTGGGCTTGGCTGTGCTACATGTTCTCTTCCCTGC
GTTATCCCCCTGATTCTGGATAATCCGATTATTACCGCCTTTTGAGTGAGCTGATA
CCGCTCGCCGCAGCCGAACGACCAGCCGAGGTCAGATGAGCGAGGAGGAAG
CGGAAGAGCGCCAATACGCAAACCGCTCTCCCGCCGCTGGGCGCGATTCA
TTAATGCAGCTGTGGAATGTGTGCTAGTTAGGGTGTTGGAAGATGCACAGGCAGCT
CCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAG
GTGTGGAAAGTCCCCAGGCTCCCGAGCAGGAAGATGAGGCGAGGCAAAACTCAGC
CTCAATTAGTCAGCAACCATAGTCCCGCCTCCCCTAAACTCAGCCCCTCCGCCCT
AACTCCGCCCCAGTTCCGCCCCATTCTCCGCCCCCATGGCTGACTATTCTTATAT
TTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATTCCAGAA
GGAGGCTTTTTTGAGGCCTAGGGCTTTTGCAAAAAGCTTGGACACAAGACAG
GCTTGCGAGATATGTTTTGAGAAATACCCACTTTATCCCCGCTCAGGGAGAGGCA
GTGCGTAAAAAGACGCGGACTCATGTGAAATACTGGTTTTAGTGCGCCAGA
TCTCTATAATATCCGCGCAACCTATTTTTCTCCGCAACACTTTTATAGCCGTA
ATAAACAGGCTGGGACACTTCACATGAGCAGAAAAATACATCGTCACCTGGGA
CATGTTCAGATCCATGCACGTAAACTCGCAAGCCGACTGCTGCTTCTGAA
CAATGGAAAGGCATTATTGCCGTAAGCCGTGGCGGTCTGGTACCGGGTGCT
TACTGGCGCTGTAACTTGGATTCTGTCATCTGCATACCGTTCCTTGTATTCTCAGC
TACGATCAGCAACACAGCGCAGGCTTAAGTGCTGAAACGCGCAGAAAGGC
GATGGCGAAGGCTTTCATCCTGTTATTTGATGACCTGGTTGGAATACCGGATGGTACTG
CGGATTGTCAATATCGTAAATGTATCCAAAAAGCAGCACTTGTCCCAACTCTTTCGCA
AAACCGGCTGGTCTCCGCTGTTGGATGACTATGGTTGATATGCCGCAAGAA
TACCTGGATTTGAAACAGCCGTTGGGATAATGGGCTGCTATTCCGTCCCGCAATCT
CCGCTCGCTAAATCTTTTAAACGCCTGGCAGTCCGGGCGTTGTTCTTTAACT
TCAGGCGGGTTACAATAGTTTCCAGTAAAGTATTTCTGGAGGCTGCATCCATGAC
ACAGGCAAACCTGAGCGAAACCCCTGTGTTCAAACCCCCTTTAAACATCCTGAA
ACCTCGACGCTAGTCCGCCGCTTTTAATCACGGCGCACAACCCGCCTGTGCAGT
CGGCCCTTGTGATTAAACCATCCCTCCTACTGGTATCGCATGATTAACGGCTGG
ATGTGGATCTGGGCACGGCATTGACCCACGCAAAATCTCGACGCAGGCCAC
GTATTGTGATTGAGCGATGACGCCGAACCTGACGCAGATGATTTATGACATCAGA
GATTGGCTACGGTGCCGCAACTGGGATTTATGAGTGAGGGCCCGGATCTTTGG
AAGGAACCTTTACTTCTGTGGTTGTAACATGACAAATTGGACAAACTACCTACAGA
TTTAAAGCTCTAAGGTAAAATATAAAATTTTTTAAAGTGTAATAATGTTAAACTA
CTGATTCTAAATTGTGTTGTATTATTAGATTTCCACCTATGGAAACTGATAATG
GGAGCAGTGGTGAATGCTTTTAATGAGGAAAACCTGTTTTTGCACCAAGAA
ATGCCATCTAGTGTGATGAGGCTACTGCTGACTCTCAACATTCTACTCTCC
AAAAAAGAAGAGAAAGGTGAAGACCCCAAGGACTTTCCTTCAGAATTGCT
AAGTTTTTGTGATCATGCTGTGGTTTAGTAAATAGAACCTTCTGCTTGCTTTTCAT
TTACACCACAAAGGAAAAGCTGCACACTGCTATACAAAAAGGAAATTATGGAAAA
ATATTCTGTAAACCTTTATAAGTGGCATAAACAGTTATAATCATACATACCTGT
TTTTTCTACTCCACACAGGCATAGTGCTGTCTGCTATTAATACATATGCTCAA
AAATTGTGACCTTTTAGCTTTTTTAAATTTTGTAAGGGGTTAAATAGGAATTTT
GATGTGATGAGGCCTTTGACGTAGGATGATAATCAGCCATACACATTGGTAAG
GGTTTTACTTGGCTTTAAAAACCTCCCCACACCTCCCCTGAACCTGAAACATA
AAATGAAATGCAAATTTGGTTGTT
>pCMV-VSV-G

Helper plasmid for packaging lentiviral clones in HEK293T. This plasmid was a gift from Dr. Rai.

GAGCTTGGCCCCATTCATACGGTATCCATATCATATGTACATTTATAT
TGGCTCATGTCCAAACATTACCAGCCTGTTGACATTGATTATTGACTAGTTATT
AATAGTAATCAATTACGCGGTACTTTAGTTTATAGCCCATATATGGAGTTCCGC
GTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGCCGACCTAC
GCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGAC
TTTCCATTGACGTCAAATGACGTATGTTCCCATAGTAACGCCAATAGGGAC
TACATCAAGTGTATCATATGGCCCAAGTACGCCCCCTATTGACGTCAATGACCGGT
AAATGCCCACCTGGCATTATGCCCAAGTACGCCCCCTATTGACGTCAATGACCGGT
TGTCATGCTTCCAAATGGGTCACTACTTTGTGATTCCGCTGGTATGGACCGAA
GTATATAACACATTCATCCGATCTTTCACTCCATCTGTAGAACAATGCAAGG
AAAGCATTGAAACACACGAAACAAAGGAACTTTGGCTGAATCCAGGCTTCCCT
TCAAAGTTGTTGGATATGCAACTGTGACGGATGCCGAAGCAGTGATTGTCCAG
GTGACTCCTCACCATGTGCTGGTTGATGAATACACAGGAGAATGGG
TTGATT
CACAGTTCTCAACGAAAATGCAGCAATTACATATGCCCCACTGTCCATAA
CTCTACAACCTGGCATTCTGACTATAAGGCTAAGGGCTATATGTGATTCTAAC
TCATTTCATGGACATCACCCTTCTCTCAGAGGACGGAGAGCTATCATCCCTGG
GGAAAGGAGGGCAGAGGTCCAGAAAGTAACTACTTTGGCTTATGAAACTGGAG
GCAAGGCCCTGAAGATGCAACTGCAAGCATTGGGGAGT
CAGACTCCCATC
AGGTGTCTGGTTCGAGATGGCTGATAAGGATCTCTTTCGACTGAGGCAGATTCC
CTGAATGCCAGAGGTCAGTATCTCTTGCTGCTCCACAGCAGTGTGGAT
GTAAGTCTAATTCCAGGACCTGGAGAGATCTGTGATTTGGATTATTCCTCTC
CTGGCAAGA
AACCTGGAGCAAAATACGAGCGGGTCCTTCCAAATCTTCCAGTGGATCTCAGGC
TATCTTGCTCCTAAAACCCAGGAACCGGTCTCTGTCCACCAATATAATCAATGG
TACCCCTAATACTTTGGAGCCAGATACATCAGAGTCTGATATTGTGGCTCCA
ATCCTCTCAGAATGTCGAATGCAGTGGAGTCTCAGACG
GTGGGATGACTGGGCACCATATGAAGACGTGGAATTTGGACCAATGGAGTT
CTGAGGACCAGTTCAGGATATAAAGTTCTTTACATGATTGGACATGGTAT
GTGGGACTCCGCATTTCCATCTTTACAGCTCAAAAGGGTCAGGTTGCAACATCCTC
ACATTAAGACGCTCTCCGAACCTTCTCGTAGATGAGAGTTTATTTTTTGT
GATAGCAGGGCTATCCAAAATCCAAATCGAGCTTGTAGAAGGTTGGCTCAGTA
GTTGGAAAGGCTCTATTTGCTCTTTTTTCTTTATCATAGGGTTAATCATGGGAC
TATTCTTGGTCTCCGAGTTGGTATCCATCTTTCGATTAATTAAGCAACACCC
AAAGAAAGACAGATTTTATACACAGATGAGATGAAACCGACTTGGAAAGTGA
TAAGATCCGTCAAGGAAATTACTCCTCTCAGTGATCAGGCTCACCATTACAGGA
TGGTGTCGGTGGTGTCGCGATGCCACTGACATGAGATGAACTGATTCGGATTTT
CCCTCCTGGCCAAAATATTATGGGACATCATGAAAGGCCCCTTGAGCATCTGACTTC
TGGCTAATAAAAGGAATTTATTTTCATTGAATATGTGTGTTGGAATTGTTTTGT
GTCTCTCACTCGGAAGGACATATGGAAGGGCAGAAATTACATATAAAAAACACAGAA
TGAGTAAATTTTGATTAGTTGGCAACATATGCCCATATGCTGCGTGCCATGA
ACAAAGGTTCTGATCATAAAGAGGCAGTCATCAGGTATATGAAACGCCCCTGTCTGT
CCATTCTTATTTCCATAGAAAAGGCTTGAGCTTGGAGATTTTTTTATATT
TTGTTTTTGTTTATTATTTCATTAAATTTTCCTTACATGTTTTATAGCCAGATTTTTCCTCCTGACTACTCCCAGTCATAGCTCTTATGGAGATCCCTCGACGGATCGGCCGCAATTCGTAATCATGTCATAGCTGTTTCCTGTGAAATTGTTATCCGCTCACAATTCCACACACATACAGCC
GGAAGCATAAAGTGAATGGGCTATAAATGGAGCTATCCGCTGCTCCAGTCCTCCGCTCCTCGAGCGGTATCAGCTCACTCAAAGGCCTTAATACGGTGCCAGAATCATGTAACACAT
GAGCTAATAATGCAATTGCAACTCCGCGGCAATTCGTAATCATGTCATAGC
GAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGAC
CTATAAAGATACCAGGCGTTTCCCCCGGGAAGCTCCCTCGTGCTCTCCGAGC
CTGCAATTAAATGAAATCCGCGCAACGCGCGGGGAAGGCAGGGGTCTTGCATATTGGGC
GCTCTGCTTCCTCCTCAGGACTCGCTGCGTTCGCTCCTCGTGCGTGCGG
CGAGCGGTATCAGCTCAGTTGGCAATTACCGTTATCCACAGAAATCAG
GGGATAACGCGAAGAAGACATGTGAGCAAAAGGCACAAGCCAGCAAAGGCGCAGGA
ACCCTAAAGAGCGCGCTTTGGCTGCTGTTTTTCCATAGGCTCCGCCCCCTTGAC
GAGCATAAACAAAATCGACGCTCAAGTCAGAGTGGCGAAGGACACAGGAGG
CTATAAAGATACCAGGCGTTTCCCCCCTGGGAAGCTCCTCGTGCCTCTCCTGT
TCCGACCCTGCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCG
TGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTT
CGCTCAAAGCTGGGCTGTGTCACGAACCCCCCTTCAGCCCGACCGCTGCG
CCTTATCCGGTAACTATCGTCTTTGAGTCCAACCCCGGTAAGACACGACCTTATCG
CCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCG
GTGCTACAGAGTTCTTGAAAGTGGTGGCCTAACTACGGCTACACTAGAAGAAC
AGTATTTGCTATCTGCGCTCTGCTGAAGCCAGTTACCTTCCGAAAAAGAGTTG
GTAGCTCTTGTACCCGGAACAAAAACCCACCCCCGTGGTACGGGTGGTTTTTTTTGTT
TGCAAGCAGCAGATTACGCGCAGAAAAAAAAGGATCTCAAGAAGATCTTTTAGGA
TCTTCTACCGGGGTCTGACGCTACGGACGAAAACCTACAGTTAAGGGATT
TTGGTCATGAGATTATCAAAAAAGGATCTTACCTAGATTTCTTTAAATAAAA
ATGAAGTTTTAAATCAATCTAAGATATATATGAGTAAAACGTGCTAGACAGTT
ACCAATGCTTAAATCAGTGAGGCACCTATCTCACGCACTGCTATCTATTTCTTCA
TCCATAGTTGCGCTCCTCCGTCGTTGTAAGATAACTACGATACCCGGAGGGTTT
ACCATTGGCCCCAGTGTGGTCAATGATACCCGGGAGCCCCACGCTACCGGCT
CCAGATTTATCGCAATAAAAAACCCAGCCAGCCGGAAAGGCGAGGCAGAAGT
GGTCTCTGCAAACCTATCCGCCCTCATCCAGTCTATAATTGTGGCAGGGAAGC
TAGAGTAAAGTATTTTGCAGTTAATAGTTTGGAGCAACGTTTGGCCATTGCTA
CAGGCATCGTGGTGCACGCTCGTCTGTTGATATGGCTTCCATTCAGCCTCGGCT
TCCCCGACGATCAAGGCAGGATTTACATGATCCCCCATTGGTGGCAAAAAAGCGG
TTAGCTCCTTCCGCTCCTCCGATCGTTTGGTVGAAGATAGTTGGCGCACGAGTCTA
TCACTACATGCTTTGCGATACCTGCTGCTGGAAGACCAGATTTCTGATATCGCCTCGT
AAGATGCTTTTCTGTGACTGGTATCTCAGAATCTCACGACTGCTTACCGAGGATAGT
GTATGCGGCGACCAGTTGCTTTGCCGCCGCTCAATACGGGATAATACCGC
GCCACATAGCAGAACCTTTAAAAATGTGTCTCACTATTGGAAACGTTCTTCGGGG
CGAAAACTCTCAAAGGATCTTTACCGCTGTTGGAGATCCAGGTTCTTTCTGGTG
AGCAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACAC
GGAAATGTGGAATACTCATACATCTTTCTTTTTCTAAATATTGGAAGCATTATTC
AGGTTATTGTCTCATGACGGGATAcatATTTGAATGTATTAGAAAAATAA
ACAAATAGGGGTTCGCGACACTTCCCCGAAAAGTGCCACCTAAATTTGA
GCGTTAAATATTTGTAAAATTCGCGTTAATTTTTGTAAAACTAGCTCATTTT
TTAACCAATAGGCCGAAAACTCGCGAAAAATCCCTTATAAATAAAAAGATAGAC
CGAGATAGGGTGTAGTTGTTGTTCCAGTTGGAAACAAGAGTCCACTTAAAG
ACGTGGACTCCAACGTTTTGTGTATCGCCACTATTAAAG
GGCCGCGACGTGCGACTTAAATCCCTGACAGCCTAAAAG
CCACTACGTGAACCATCACCCTTAAATCAAGGTTTTTTGGGGTGAGGTGC
AGCACTAAATCGGAACCCCTAAAAGGGAGCCCCCAGTTTAGAGCTTGACGGGGA
AAGCCCGCAGCAGTGCGAGAAAAGGAAGGGAGAAAGGGGAAGAAAAGGAGCGG
CGCTAGGGCCTGGCAAGTGTAGTGCATCGCTGCGCTAACACCACCAACC
GCCGCGCTTAATAGCGCCGGCTACAGGGCGCTCCCATTCGCCATTCAGGCTG
GCAACCTGTGTTGGAGGGCGATCGGTGCGGGCCTCTTCGCTATTGCCAGCT
GGCGAAAGGGGATGTGCAGTGCAAGGCAGTTAAAGTTGGGTAAACGCCAGGTTT
TCCAGTCACGACGTGTAAACGACGCGCCAGTGAGCGCGCTAAATACGACT
CACTATAGGGCGAAATTTGGAGCTCACCACGGGTTGCGGCGCGTCTCAGA
>pmaxFP-Green-N

Also known as pmaxGFP from Amaxa. This GFP is really bright and quick-folding variant that expresses in bacteria and mammalian cells.

```
TAGTTATTAATAGTAATCAATTACGGGTCATTAGTTCAATAGCCCATATATGG
AGTTCCCGCTTACATACATTCCGCTAAATGCGCCTCGGTGACCAGCCA
CGACCCCCGCCCCGATCGTCAATAATGACGTATGTTCCATAGTAACGCA
ATAGGCCGCTTACATAGGCTGAGTATTTACGCTAAACTGGCTG
ACTTGGCAGTATCATATGGTAATGCGTGGATAGCGTTGTAGCTACGGG
```
GGCCACGACGGGCGTTCCTTGCGCAGCTGTGCTCGACGTTGTCACTGAAGCG
GGAAGGGGACTGCGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCAT
CTCACCTTGCTCCTGGCCAGAAAGATATCCCATCATGGCTGATGCAATGCGGCG
GCTGCATACGCTTTGATCGGGCTACCTGCTGCCATATTCCCGACCCAAAGCGAAGAT
CGCATCGACGACACGTACCTCGGATGGAAGCCGGTTCTTGTCGATCAGGATG
ATCTGGACGAAGACATCAGGGGCTCGGCAGCCAGCAGCAACTGTCCTCGCCAGGCT
CAAGGCGAGCATGCCATGCCCCACGGGCGAGATCTCGTCTGACCCATGGGCATGCC
TGCTTGCCGAATATATCATGTTGGAAATATGGCCGTTTTCTGAGATTTCACTCGACTG
TGCGCCGCTGGGTGGTGGCGGCTACCCCGCATCCGGTGGTGGCTACCCTGT
GATATTGCTGAAAGAAGCTTGCGGCGCAATGGGCTGACCGCTTCCTCGTGCTTTAT
CGGTATCGCCGCTCCTCGGATTGCAGCGCATCGCCTTTCTATCGGCTTTCTTACG
AGTTCTTCTGAGCGGGACTCTGGGTTGGAATAGGCCAAGACGAGCAGCAGCAGC
AACCTGCCCATACGAGATTTTCGATTTCCACCGCCTCTCTCTATGAAAGGTTGGG
CTTGGGATATGGTTTTCCCGGACGGCGCTGGGAGATGATCCTCCAGCGCGGGGAT
CTCATGCTGAGTTTCTTCGCCCCAACCTAGGGGAGGCTAACTGAAACACGGA
AGGAGACAATACCAGGAAGGAAACCGCGCTATGACCCAATAAAGAGACAGA
ATAAAAACGCACGGGTGTTGGGTGTTTGTCTAAACGCGGGGTTGTCGCTCCA
GGGCTGGCAGCCTGCTCGATACCACCGAGACCACATGGGGGCAAATACGCC
CGCGTTTCTTTCTTTCTTTCCCACCACCCCCCAAGTGGGAGGCTGAAAGGCGCAG
GCTCGCGAAGCGCTCGGCGGCGACGCGGCGCGCCCTGACATGCGTACGTTACTCA
TATATACCTTTAGATTTAAAAATTTCAATTAAAAAGGATCTAGGT
GAAGATCCTTTTGGATAATCTCATGACCAAAATCCTTTAAGCTGAGTTTTCGT
TCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCC
TTTTTTTCTGCAGCATATCTGCTGTTGCAAACAAAAACACCAGCTACAG
CGGTGGTTGTGTTGAGCACGCTACAAACTCTTTCTCGGAGGGTAAC
CGCTTGAGCAGCAGACTCAGTACCTGCTACCATCTCGGTCTCTGCTAA
TCCTGTTACAGTGGGTGGGTGCTGAGTTGAAGTCGATTTTGAGTTG
GACTCAAGAGCATAGTTACCCGATAAGCCAGCGAGCGTGGGCTGAAACGG
GGTTCGTGACACAAGACAGCCAGCTTGGGAGCAGCAGACCTACACCCGAGAT
ACCTACAGCGTGAGCTATGAGAAACGGCCAGCTTCCGGGAAGGAGAAAGG
CGGACAGGTATCCGGTAAGCGGAGGGGTGGGAAGCAGAGGACGCACGAGGG
AGCTTCCAGGGGAAACGCTTGGTATCTTTATATGTCTGTGGGGTTTCGCCAC
CTCTGACTTGAAGCCGCTGATTCTTTCTGTAGTGCTGTCAGGGGCGGAGCGGCTATG
GAAAAACGCCAGCAACGCGGCTTTTTTACGTTTCGCTGCCCTTTGCTGCTT
TTGCTCAGTGTTCCTCTCGTTATCCCGTATTGCTGGAATACCCTATTACCAGGACGCTGAT
Phage lambda Red recombination system expressed from $P_L$ under $cI^{857}$ control in a temperature sensitive replicon. This plasmid was a gift from Dr. Court.

> pSIM5

CAGCTGTCTCTTATACACATCTCCGCTGTGCTTTTCAGTGAGATTTCGGATAACA
GAAAGGCCGGAATACACCAGCCTCGCTTTGTAACGGAGTAGACGAAAGTG
ATTGCAGCCTACCCGGATATTATCGTGAGGATGCGTCATCGCCATTGCTCCCCA
AATACAAAAACCAATTTACGCCAGTGCCTCGCTCCATTTTTTCGATGAAACTCCGG
CACGATCTCCTCAAAACTCGCCATGTACTTTTCATCCCCGCTCAAATCAGACAT
AATGCAGGCCTTCACGCTTACATACGCCGCGGTTCATAGTTGGCAAAAGTGACCAGGC
ATTTTTTCGCAGTCAACCTGCTGTACTGCATCGACCTGCACTGACCTTGAAAGCTGACT
TTATGCGCTGAAACCACCGAGCCGGAACCTTCATGAAATCCCGGGAGGTAAA
CGGGCATTTCAGTTCAAGGCCGTTGCCGTCACTGCATAAACCATCGGGGAGAG
CAGGCGGTACGCATACTTTTCGTCGCGATAGATGATCGGGGATTCAGTAACAT
TCACGCCGGAAAGTGAacTCAAAACAGGGGTCTGGCGTCGTTCTCGTACTGTATT
CCCCAGGGCAGTCTTTTAGCTTAACCTCCCGGAGCACAACCGGTGCAAAACCT
CAGCAAGCAGGGGTGTGGAAGTAGGACATTTTCATGTCAGGCACCAGCTTTCTCC
GGAGCGGGGGTTTTTGCTATCAGGTTTGTAACCTCTGAAAGCGGGGTGATGAGCAGCG
AGCCGTAATTTGTGCTCCAGCAATCATCTATCCCCCTCTGCCAGACAGCTCTACATCGAT
CCCCTACGCTGCAGGATAATGTCCGGTGTCATGCTGCCACCTTCTGCTCTGC
GGCTTTTCGTGTTTCAAGGAATCCAAAGAGCTTTACTGCTTCCAGCTGTGTCAGTT
CTGACGATGCACGAAATGTCGCGGCGAAATCTGGAACAGACGGGAATAAATCCTGGGAAACAGAGCGGGAATA
AGTCGTCATCCCATGTTTTATCCAGGGCGCAGACGCAAGAGTGTAATCTCCTTCG
ATGGTTTCATCGTTAACCAGGAGTGATGTCGCGTCCGGGCTGACGTTCTCGAGT
GTATGCAATTTTTCGAGAAATGCGCTCGGCTTTGCTATCAGGATACCAG
CAAATCCGAAGGGCAGACGGGACACTGGAATCATGGCTTTATGACGTAACAT
CCGTTTGGGATGCGACTGCCCAGGCCCCTCTGATTTCTCCTCGCTCTGGAAGTTT
TGAATGGTTCGCGGCAGCATTCCATCCATCCATTCCGTAACGCAGATCGGATG
ATTACGCTCTTCGCCTGAATCCCGCATGTACAGGATTCATTGTCTCTGCTCAA
AGTCCATGCACTAAACTCTGTGGTTTTCATTGATGAGTCGGGACAGCCAGCATCA
ACGCCCACCACCAGGGAACGATGCCATTCTGCTTTATCAGGAAAAGGCGTAATTT
CTTTGCAGCTCAGGATTAAGGGCGTACGCTGTGGGCAACGATCAGTAATGCGAT
GAACCTGCGCATCGCTGGCCATCACCTTTAAATGCGGTCTGGCGAGAGTGGTG
ATCAGTTTCTGTGGGTGCAGACGAATCCATGCGGACACGTTACTCAGGCAGGCTTCCC
AGCCAGCGTTGCAGGTGACATACTCAATTTCGTTTATACCTCTGAAATCAATATC
ACCGTCCGTCTTTTGCATTGCTGGTTTTCAATCGTGAGGTTGAGTACTTGCAC
ACAGCTAAATTAGCTGGTTTCAATGATCTCCGTTAAATCGACTACAATCG
CTAGTCTTTGCTTTGAAACAAACTAATCTCAGACATACATCTCAATTTGGTCT
AGGTGATTATAAACACTATAACAAATGAGATTGGGCTAGTCAATGATAATTACT
AGTCCTTTTCCTTTGAGGTTGGATCTGAATTCTGCTAATTTCTGACCTTTTGCTGG
AAAACCTTGAAATCTCCTAGACCCCTCCTGTGAAATTCCGCTAGACCTTTTGCTGTG
TTTTTTTTTTATATATCAGTCAGTTGATTATAATTTATAGAAATAAGAAAGAAATAAA
AAAAAGATAAAAAGATACTCCACGCTGGTATAACTCCTACTACTTTGAGT
CAGTTCGCAGTATTAAAGAAGATGTCGCAAACGCTTTGGCTCCTCTCTACAA
AACAGACCTTTAAAAACCTAAAGGCTTTAAGTCACCCCTCGAAAGCTCGGCA
AATCGCTGAATATTTTGGTGCTCGGACCCTACAGGCACCTGAGTCGCTGTCTT
TTTCGTCGCGATTTCAGTTTCTCGCTGCTCAAGGTCTGCTGCTTGGCAGTGAATGGGGTGTA
AATGGCACTACAGGCAGCTTTATGGATTCTAGCAGAAGAAACTACCCATAAT
ACAGAGAGGCCCCGTCAGGGCGTTTCAGGGCGTGTTGATGGCCGGGTCTGCTA
TGTTGGCTATCTGACCTTTGCTGTTCAGCTTTGCTGCTGCTGCTGTATTTTCCA
GTCTGACCTTTCCGATTATCCCGTGACAGGTCATTACAGACTGGGTATCGAC
CCAGTAAGGCGAGCGGTATCATCAACAGGCTTACCCGCTTTACTGTC
Phage lambda Red recombination system expressed from P_L under cI$^{857}$ control in a temperature sensitive replicon. This plasmid was a gift from Dr. Court.

CAGCTGTCTCTTTATACACATCTCCGCTGTGCTTTTCAGTGATTTCGGATAACA
GAAAGGCCGGGAATACCCACGCTCGCTTTGTAAACGGAGTAGACGAAAGTG
ATTGCGCCTACCGGATATTATCGTGAGGATGCGTCATCGCCATTGCTCCCCA
AATACAAAAACCATTTTCAGCCAGTGCCCTCGCCCATTTTTTCGATGAACTCCGG
CACGATCTCGTAAAAACTCGCCCATGTACTTTTTTCGATGAACTCCGG
AATGCGAAGCCCTTTCACGCCCTACGCGGCTATGAGTTGGCAGGAAAGTACCCAGGC
ATTGTTTCGCACCATGCGTTGACAATGATGTAAGAATTCGGCG
TTATGCACTCGGAAACCAACCGAGCGGGAACTTCATGAAATCCCGGGAGGTAAA
AACCTGGTGAGCAATGGTTTCAACCATGTACCGGATGTGTTCTGCCATGC
GCTCCTGAAAACATCGTCATCAAAACGCACCGGTAAATGGGATTTTTTGCTG
GCCGGTGGCGTTGAAAATGATCGATGCATAGCGATTTCAACACAGGTTGCTGGG
GCAGGCGCTTTTCCCATGTCGTCTGCCAGTTCTGCCTTTTCTTCTTCAACGGCCGA
GCTGCTGGTGATTAGTGACGCAGCAGCTCGAGCCTCAAGACGATCCTGAATGTA
ATAAGCGTCTCATGGCTGAACCTCTGGAATAGCTGTGAAAAATATCGCCGCGGA
AATGCCGAGGCTGATTAGTAATCCGGAATCGCAGTACTACGGCCATGCTTCTGTT
CGTATCACACACCCCCAAGCCTTTCTGCTTTTGAATGCTGCCCTTTCTTTCAGGGCT
TAAATTTTAAGAGCGTCACCTCATGGGTGTCAGTGCCTCCTGCTGATGTCGT
CAGTATCACGCGCCAGTGGTATTTATGTCACAACACCGCCAGAGATAATTATACAC
CGCAGATGTTATCTGTATGTTTTTTATATGAATTATTTTTTGCAAGGGGGGC
ATTGTTTTGAGGTAGGTGAGAGATCTGAATTGTGTTTTAGGTAGTTGTATCTAT
TTATTTTTCAATAAATAACCAATTGTTATGTTTTTTGCGGGGGGCA
AAGAAAACCCGGCGCTGAGGCCGCGAGCTGTAAGAGTTTGGGTAGCTCTGATCCGGC
AAAAAACCCCGCGTGGTTGCTAAGGGGTGTTTTTTTGTTTGCAAGCAGCAGATAC
GCGCAGAAAAAAAGATCTCAAGAAGATCCTTTTGATCTTTTTCTACGGGGTCT
GACGCTCAGTGGAAACGGAAAAACTCAGTTAAGGGGATTTTTGGGTAGMAGATTAT
CAAAAGGATCTTCCCATGATCCTTTTTAAATAAAAATGAAAGTTTTAAATCA
ATCTAAAGTATATATGAGTTAAACTTGGTGCTGACAGTTAACAATGCTTATACAG
TGAGGCCACCTATCTCAGCAGTCTCTATTTCTCGTCATCCATAGTTGGCTGAC
TCCCCGTCGTGTAGATAACTCAGTACGGAGGGGCTATTACGCATCTGGGCCCAG
TGCTGCAAATGATACCGCAGACCACGCTCACCAGGTTCAGTTAATACAGCA
ATAAAACCAGCCAGCCGGAAGGGCCCGAGCGCAGAAGTGGTCCCAGACTTTAT
AAACAGGGTACTCATACTCATTCTAAGTGGACTGCATAACTAACCCTTTCA
TACATCTCGTAgATTTTCTCTGGCGATGAAAGGCTAATAATTTCTCACAAGGCTAACA
TTTGAGAATTCTTGAAAGCTGATTGCTACTGCCCCCATTCCCATCTTGCTCGACA
GATTCCTGGGATAAGCGCAAAGTTCTCATTCTTTTTTCCATTAAATGCTTTAAGG
CGACGTGCGTCCTCAAGCTGCTCTTGTTGTAAATGGTTTCTTTTTTTGTGCTCATA
CGTTAAATCTATCACCGCAAGGGATAAAATATCTAACAACCGTGCGTGTTGACT
ATTTTACCTCTGGCGGGTATAATGTTGCGATGTAACAAAGGAGGTGTTGTATGGA
CAAACGAGATGTGTAAGAGACAGCTGACCGGTTTGGCTGCCCGAACCAGGG
CTGTTCTGGGTGTTGCATTATTTTGGTATCACGAATCAGATCCGGCTACAGTTT
GCCGGCTGAAAGCGCTATTCTTTCCCAGAAATTGGCCATGATTTTTTTCCCCAACAGGG
AGGCAGTCACTGGCTCCCCGTGTTCTGGCGAGCTTTTGATTCGATAAGCAGCATCG
CCTTTTCAGGCTGTCTATGTGTAAGCTTGAGCTGTAACAAGTGTCTCAGG
TGTTCAATTTCATGTTCATTGCTTTTATTTTACTGGTTTTTTCACCTGTTTCTATTAG
GTGTTACATGCTGTTGATCTGTGTTACATTGTCGATCTGTTCATGGTGAAACAGCTT
TGAATGCAACAAAAATCGTAAAGCCTCTGATGTATCTATACCTATCTTTTTTACACCG
TTTTCATCTGTCATATTGACAGCTTTCTTTCTTTGATATGTAACCGTGTTTTTGCGTGAGCCATGAGAACGAACCATTGAGATCATACTTACTTTGCAT
TGTTTTTGCAGGCATGAGAACGAAACCATTGAGATCTACTTTACTTTGGCAT
GTCACTCAAAAAATTTTGCTCAAAACTGCTGAGCTCAGTACTAGATAAAAGAAGCCAT
AAGAACCCTCAGATCTCCGTTATTAGCCAGTATGTCTCTAGTGTGTGCTGCTG
TGTTTTTGCAGCTGAGACAGTAAACCATTGAGATCTACTTTACTTTGGCAT
GTCACCTCAAAAAATTTTGCTCAAAACTGCTGAGCTCAGTACTAGATAAAAGAAGCCAT
CATCGTGTAGTGTATTTCTTAGTGTTATGAGGTAGGAAATCTGATGTAAATG
GTGTTTTGGAATTCTGCACTTTTATTATCTGTTTGTTCTCAGTGTTCTCGGTT
ACGAGATCCATTGTCTATCTAGTTCAACTTGAGAAATCAACGTATCAGTGCG
GCGGCCTCGTTATCAACCACAAATTTCATATTGCTGTAAGTGTAAAATCTTT
TACTTTATTGTTTTCAAAAAACCATTGGTTAAGCCTTTTAAACTCATGGTATGTA
TTTTCAAGCATAAATGAAATCTCATCAAGGCTAATCTCTATATTTGCT
CTTGTGAGTTTTCTTTTGTTAGTTTCTTTTAAATAACCACCTCATAAAATCCTCAT
AGAGTATTTTTTTCAAAAAAGACCTAAACATGTCAGGCCAGTTATATTATTTATGAATT
TTTTTAACTGGAAGATAGAGGCAATATATCTCTTTCACTAAACTAAATCTCAAT
TTTTCCCTTGGAGAATTTGCGTGGTATACTGGTTAAGCCTTTTAAACTCATGGTATA
TTTTTGTTTTGTTTTCAATAAGTTGAGTTCCAGATTATATTTTATGAATT
TTTTTATACCCATAAGCATTTCCCCTACTGTGTTTCATCATCTAGCGTAT
TGGTTTATAATGAAACGATACCCGCTCGGCTTTCTTTCTGTAGGGTTTTTACAATCGT
GGGGTTGAGTGTTGCCACACACAGCATAAAATAGCTTGGTTTCATGCTCCGTTA
AGTCATAGCAGACTAATCGCTAGTTTCTATTGCCCTTTTGAAAACAACACTAACATCAGAC
ATACATCTCAATTGGTCTAGGTTATTAAATACTATAACAAATGGAGATGGGC
TAGTCAATGATAATTACTAGTCCTTTTCCTTTGAGTTGGTATTCTGTAAATT
CTGCTAGACCTTTGCTGGAAAAACTTGAAAATTCTGCTAGACCTCTCTGTAAATT
CCGCTAGACCTTTGGTGTGTTTTTTTTTTTGTGTATTTATTCAAGTGGGTATAATTAT
GAAATAAGAAAGAATAAAAAAGATAAAAAAGATAGATCCAGCCCTGGTGT
ATAAATCCTACTACCTTTAGTCAATCCGCTAATTACAAAGGATGTCGCAAAC
GCTTTTGCTCTCTACAAACAGACCTTTAAACCTCAAGGCTTAAGTGACTACG
ACCCTCGGAGCTCGGGCAATCTCGGAATATTCCCTTTHTGCTCCGACCACATCA
GGCACCCTGAGTCGCTGTCTTTTGCTGACATTCCAGTTCGGCTGCTACGGCT
CTGGCAGTGGAATGGGGGTAATATGGCACTACAGGCCTTTTATGGAATTCAATG
CAAGGAAACTACCCAATAATACCAAGAAAAGCAGGGTTCGCTGCAGGGC
TTTTATGGCGGGTCTGCTATGTGGTCTATCTGACTTTTTGCTGTTCAGCT
TTCAGACTGGCTAATGCAACCAGTAAGGCAGCGGTATCATCAACAGGGCTTAC
CCGTCTTACTGTC

>pKD4

Kanamycin resistance gene flanked by FRT sites in a temperature sensitive replicon. Used as a PCR template for knocking out genes. Non-polar on neighboring genes due to introduced Shine-Dalgarno site outside one of the FRT sites. This plasmid was a gift from Dr. Wanner.
>pKD4
AGATTGCAGCATTACACGTCTTTGAGCGATTGTGTAGGCTGGAGCTGCTTCGA
AGTTCCCTATACTTTCTAGAGAATAGGAACCTTCCGAATAGGAACCTTTCAAGATC
CCCTCACGCTGCCGCAAGCAGCACTCAGGCGCAAGGGCCTGCTAAGGAACGGGAG
AACACGTAGAAAGCCAG
TCCGCAGAAACGGTGCTGACCCCGGATGAATGTC
AGCTACTGGGCTATCTGGACAAAGGGAAAAACGCAAGCGCAAAGAGAAACGAG
GTAGCTTGCACTGGGCTATACATGGCAGATAGCTAGACTGGGCGGTTTTATGGA
CAGCAAGCGAAACGGGAAATTGCCAGCTGGGGCGCCCTCTGCTGTAAGGGTGGAA
GCCCTGCAAAAGTAAACTGGGATGGCTTCTTCTTGCCGCAAGGATCTGATGCGGC
AGGGGATCAAGATCTGATCAGAGACAGGATGAGGATCGTTTCGCATGATTG
AACAAGATGGATTGCAACGCAAGTTCTCCGCGGCTTTTGGTTGGAAGAGGCTATT
CGGCTATGACTGGGCACAAACAGCATACTGGCTCTGGATGCGCTGCTTCCAC
CAGCAATATCATGTCGAAAGCCAGCTCTGGATGGCTTCTCTAGGGAGCTCTG
ACGCTTGAATCCGACACCATCTGGGTGCTTGGCCGCTTGGCCACGACGGGCCT
CTTGCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGG
GACTGGGCTGCTATTGGGCGAAGTGCCGGGGCAGGTCAATCCATCATCACTGGCCAC
ACGCGGCGGGGCGGGGCGGCTCTGGGCTCTGGGCTTCTGCGGCTGACTG
AGCGAGCAGCAAATCCGAGGAGCCAGCTCTGGCTTCACGAGATCAGCTG
CGAAGAGCACATCGGATGGAAGCCGCGTCTCCGCGCAAGCGAAGGCTGCTTCAGG
CGCATGCCGGACGGCGAGGATCAGGCGATTGACAGCTGACGCCTGGTCC
CGAATATCATTGTTGAAAAATGGCGCTTCTCTGGATTTACATCGACTGGCCGG
CTGGGCTCCGGGACCGCTATCAGGACATAGCGTTGGCTACCCCGTAGGTTTG
CTGAAGAGCTTGCGGCGAATGGGCTGACCGCTTCCTGCTGTCTACGGTATC
GCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTGACGAGTTCTT
CTGAGCGGGACTCTGGGGTTCGAAATGACCGACCAAGCGACGCCCAACCTGC
CATCACGAGATTTCCGATTCACCGCCGCTTTCTATGAAAGGTTGCTTCGGA
ATCGTTTTCCGGGACGCGGGCTGGATGATCCTCCACAGCGCCGGGATCTCATGCT
GGAGTCTCTCCGCACCCACCGCTCCTCTAAAGGCTCTGCTCTAGTTCTATACCTTTC
TAGAGAATAGGAATTCTCGGAATAGGAATAGGAGGATATTCATATGGACCA
TGGCTAAATCCCATGTGACGCCCTAAGGTTCTCTGTGTGCTACGAAAATGCTT
TGAGAGGCTCTTAAGGCCTTCTAGTGCCTACATCCCTGCTGGCTGGTGTCCACA
ACCGTTAAACCTAAAAGCTTTTAAAAGCCCTATATATTTTTTTTTTCTTATAA
AACTTTAAACCTTAAAGGCTATATTAAGGTTGCTTTATATATATTGTT
CAAAACATGAGAGCTTAAGTACGTGAAACATGAGAGCTTAGTACGTAGGGTTTAGTTCGTTAAACATGAGAGCTTAGTACGTAGGCTTAAACATGAGAGCTTA
GTACGTAAACATGAGAGCTTATAGTACGTGAAACATGAGAGCTTAGTACGTAGGCTTAAACATGAGAGCTTAGTACGTAGGCTTAAACATGAGAGCTTAAACATGAGAGCTTAGTACGTAGGCTTAAACATGAGAGCTTAGTACGTTAGCCCATG
AGAGCTTAGTACGTAGGCTGTTAGGGAAGTTAGTGCAGTTACGTTACGCTGAAG
TATCCAACAGGTGGAACTGCGGATCTTGCGGCCGCAAAAATAAAAATGAAGTT
TTAAATCAATCTAAAGTATATATGAGTAACATTGCTGACAGTTACCAATAG
CTTAATCAGTGAGGCAACTTCTACGTGACTCTGCTGCTGCTATATATCCATCCATAG
TTGCTGACTCCCGCTGCTGCTAGATAACTACGATACGGGAGGGCTTTACCATCT
GGCCCCAGTGCTGCAATGACGCGAGACCCACGCTACCCGGCTCCAGATT
TATCAGCAATAAACCAGCCAGCAGCGAAGGGCGAGCGCAGAATGTTTGCTCG
CAACTTTATCCCGCTTCTACATGCTATATATGTGTTCCGGAAGGCTAGAGTA
AGTAGGTCGCAATAGGTTGCGACAAGTGTTGGCCATTGCTACAGGCAT
CGTGGTGTCAGGCGCTCGTTGTTGTTATGGCTTCCCTCACGCTCGGTTCCCAAC
GATCAAGGCGAGTTACATGATCCCCCATGTTGCAAAAAAGCGGTTAGCTC
CTTCGGTCTCTCGATCGTGTGTCAGAAGTAAGTTGGCCGCAGTGTATCCTCA
TGGTTATGGCAGCAGCTGAGTAATTCTCTTTACTGTGTATGCCATCCCGTAAGATGC
TTTTCTGTGACTGAGTGAGTACTCAACCAAGTCAATTCTGAGAATAGTGATATGCG
GCGACCAGGTTGGCTCTTGCCCAGGCGTCAATACGGGATAATACCCGCCGCACAT
AGCAGAACTTTAAAAGTGCTCATCATTGGAAAAACGTTCTTTGCGGGCGAAAAAC
TCTCAAGGATCTTTACCAGCTTGGAGATCCAGTTCGATGTAACCCACTCGTGCA
CCCACCTGATCTTCAGCATCTTTTTACTTTCACCAGCGTTTCTGTTGAGCAAAA
AACAGGAAGGCAAAAATGCCGCAAAAAAGGGAAATAAGGGCGACACCGGAAAT
GTTGAATACTCATACTCCTTCTTTTCAATATTATTGAAGCATTATTTACAGGGTT
ATTGTCTCATGAGCGGAATACATATTTGAAATGTATTTTAGAAAAATAAAAAAT
AGGGGTTCGCCGACATTTTCCCGAAAAAGTGCACCTGCACTCGATGCCCCC
CGATGGTAGTGTGGGGTCTTCCCATGCAGAGTTGAACTGCCAGGCATCA
AATAAAACGAAAGGCTAGTCCAGAAAGACTGGGCTTTTCAGTTTTATCTGTGTG
TGTCGGTGAACGCTCTCTGAGTAGGACAAATCCGCGGGAGCGGAGTTTGA
CGTTGCGAAGGCAACCGGCCAGGGTGTCGGCAGGCAGCCCGCCATAAAC
TGCCAGGCATCAAAATTAAGCAGAAGGCCATCCCTGACGGGTGGCCTTTTTTCG
TGGCCAGTGGCAAGCTTGCATGCG
>pNEBR-R1

Plasmid obtained from NEB.

CACCTGACGCGCCCTGTAGCGCGCGCATTAAGCGGCAGGCGGTGTTGGTGTTTAC
GGGCAGCGTGACCGCTACAACCTTGCCACGCGCCCCCTAGCGCCCAGCTCTTTTCGTTT
TCTTCCCTTCCTTTCTCGCCACGTTGCGCGCTTTTTCCCGTCAAGCTCTAAATC
GGGGGCTCCCCTTTTACGGTTCCCGATTTTGTGCTTTACGGCACCTCGACCCCCAAA
AAACTTGATTAGGGTGTGATGTTTCAGTGTGGGCATCGCCCTGATAGACGG
TTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCC
AAACTGGAACAACACTCAACCCCTATCTCGGTCTATTCTTTTGATTTTATAAGGG
ATTTTGCGATTTCGCCCTTTATGGTTTAAAAATGAGCTGATTTAACAAAAATT
TAACGCGAATTCTTAAACAAAAATATTACGCTTACAATTTCCATTCCGACACAG
GCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGC
AAGCCTTAAGCAGCGTAATTTGATCAGGCTGCGAGATCTTGGCGAACCCTGG
CAATCCCTTCCGCAAACTCCACGATAAGTTGGACCCTGAGGATAGTCATCTCTG
TGATCTGGCGGAAGGAAGGTGTCCGACACTTTCTTTATTCTACGCTCCGCTTGGCTGC
CACGTCTCGTAATCTTCTTTCAAATCTTCTACGAGAAGCTGGATGTCACCTGCTCTCTCTC
CTGGATCTGGCGGAAGGGAGTGTCCGACTCTTCGTTTTCATCGTCCGCTTGCTGC
CACGTCTGCGTAATCCTCTTCAAATCTTCATCAGAAGGCTGCTCGTACCCGTCC
CTGGTACCAGATGAGCCCTGGCGATAAGGAACTGCTGGTTGGGTGGTGTCACACTTG
GGATGGTCTTCTCGGCAGTTTTGCACCAACAGCTTTTGTCCGAGAGAAAACCTTG
GACCACCTTCGTGAATCCTTTGCTGGCTTCACGAGGAGTGGAAGGTTCACACTGCATAA
TGGGCAGCATGTGGTGCTCCACCCTGCTCGGTCTGACAGGCAAGTTTTTGTCCTTC
TCCTCTGTCCTTTCTTCCTTCGCCGTTCACGACTGAGTCTTCGGGTACT
ACGCACTCAGCCAGATCCCCCGGAATATCCCGCGATACAGTCAACTGTCTTT
GACCTTTGTTACTACTCTTCTCTCCGATGATGAGTGGCCTCCACTTATCTATGCTGTCTCAATGTTAGAGGCATATCAGTCTCCACTGAAGCCAATCTATCTGTGACGGCTCTTTATTCACATTATCTTCTTGTACAAATAATCCTTGTTAACAATATGCTTTTTATATCC
TGTAAGAATCCATTTCATTTTCAAAAAATCATGTCAAGGCTCTTCTCAGGAAAAATTCA
GTAGAAATAGCTGTTCCTTCAGCTTTCTAGCCTTTGATTCCACTTTCTTCAGATGT
GCCCTAGTCAGCGGAGACCTTTTGGGTTGGGAGTGAGTCGGCAACTCCCACTTT
GTTCTTCAGACACTGGGCGCACTTTCGGTTTTTTCTTTGGGAGCATTGGCTTTTT
AAGTGGCAATATATCGCATTTGCTTGCATGAAGACAGTAAGCTTCCATGGTG
GCTCAGCTAGCCACTAAGATCTGCAATTTTGACCTAGTAAAAGTAAAACATA
AGCGTGTCACCTGAGATCATTGTATTCTCTTGCTTTTACACCTATATATTCTTCA
AAATTAGCAACTATCTTCTTTCAAAAAATGCCTTAATACGTTAAATATCCGCAAAG
CTCCCTCAATTTTACAAGGCTAAGTCACTTAAATACGAAAATTCTGCATTCAAG
GTTTCAACGTCAAACAGCGGAAAAATCCAGAGCGCTCTCTTTACCAAGCGTCCAC
AACTATGCGCTCGGGGTTGCGAGTGTGTTTTGTGAAGTTTTTTAGGCACCTT
TTGAAATGTAATCATTTGGGTCAATATGTAATTTTCAGTGTTAGACTAGTAAA
TTGTCCGCTAAATTTCTGGCCGTTTTTTGCTTTTTTTTTTTAGACGAAATCGCTAGCC
CCGCCACCATGGGCCCTAAAAGAAGACGTAAGAATCGCCCCCCCCCGACCCGATGT
CAGCCTGGGGGACGCAGCTCCACTTAGAGCAGGACGCTGGCGATGGCGCATGGCGG
TGCCGACGCGCTAGACGTTCGATCTGGAAGAATATGTTGGGGGACGGGGATTCC
CCGGGTCCGGGATTTACCCCCCACGACTCCGCCCCCTACGCGCTCTGGATAT
GGCCGACTCCGAGTTTGAGCAGAGATGTTTTACCGATGCCCCTTGGAATTGACGAGT
ACGTTGGGGGAATTCCAGATTGCTGTGGAGACAGATCTGGAGGCGAGAGCAGTTGC
TGTTGGAACAGAAGAGTGACAGGGCGTTGAGGGTCCTGGGGGAACCGGGGGT
TAGCGGCCAGCACGGGCAATGACCCCTGTGACTAACATCTGTCAGGCACTGGAC
AAACAGCTATTCCACGCTTTGTGAGTGCGGGGAAGAGGATCCCACACTTTTCTC
CTTGCCTCTGGATGATCGGCTCATATTGCTGCGGCGCAGGCTGAATGAACCTCC
TCATTGCCCTCTTTTCACACCGATCCATTGATGTTGAGATGTTTTACCGATGCCCTGGA
CCACAGGTCTTCTGCACGCTCCACCGCAACTCAGCCCATTCAGCAGGAGTAGGAGC
CATCTTTGATCGGGGTGCAGCAGACTAGTGTCAGAAATAGCGTGACAATGAGG
ATGGACAAGACAGACGCTGGCAGGGCAATCGGATTTCAATGGAACAGAGCTTCTAGTGA
AGGAGAGGGTTTGAAATCAGGCCCAAGGAGTTGAACCTTCTACTGTAATGGAATG
ATATGCGCTTTGGAAAGATATACTAGAAACACACATCCCCGATGAAAACAGGA
AGATTTGCAAAAACCTTGGTTCTCGCTGCCCCCTTTACGTTCCATAGGCCCTTAAG
TGTTTGGAGACATTTGTTTTCTTTGCTCATTGTGGAGATGTTCCAATTGATACG
TTCCTGATGGAGATGCTTGAATCACCCTCTGATTCTATGAACTGTGCAAAGAAGCT
CATGATAAGATACATGGATGAGTTTGGACAAACCCCCACACTGAGATGCAAGTA
AAAAATGCTTTATTTGTGAAATTGTGATGCTATTGCTTTATTTGTAAACCATT
ATAAGCTGCAATAAAACAAAGTTAACAACAACAAATTGCATTCATTTTATTTTCA
GGTTCAAGGGGGAGATGTTGGAGGTTTTTTAAAGCAAGTAACAACCTCTCACAA
TGTTGTAATAATCCGATAAAGCTCAGTACCTCTCTGAGGCGGAAAGAACCAGCTGTTG
AATGTGTGTCAAGGAGGTGTGGGAAGTCCCCAGGCTCCCCAGCAGGCGAGAA
GTATGCAAAGCATGCTATCTCAATTTAGTCAGCAACACCAGGATGTGGAAGGATCCCC
AGGCTCCCCAGCAGCGAAGATATGCAAAGCATGCTATCTCAATTTAGTCAGCA
ACCATAAGTCCCCGCCCCCTAACTCCGCCATCCCCGCCCTAACCCTCGCCAGTTC
CGCCCATCTCCCGCCCCATGGCTGACTAATTCTTTTTATTATGCAAGGCGGC
AGGCCGCCCTCGGCTCTGAGCTATTTCCAGAGTAGTGTGAGAGGCTTTTTGG
AGGCCCTAGGCTTTTTGCAAATAAGCTTGAATTCTTCTGACACAACAGTCTCGAACT
TAAGGCTAGAGCCACCATTGAAACAAGATGGATTGCGACGAGGGTTCTC
GCCGCTTGGTGGAGAGGCTATTTCCGCTATGACTGGGCGACACAGCAGAGGCG
GCTGCTCTGATGCGGCCGTGGTCAGCTGAGCGGAGGCGCCGGTCTCTT
TTTGTAAGACCCACCTGTGCCGTGCTCGCCGAATATGCAACTGCAGGACCGAGAC
CGCCTTAGTCGTGGCTGGCCACGACTGGCGCTTCCTGCAGCTGTGCTCGAG
GTTGTCATCGAGGGAAAGGGACTTGCGTCTTTGGGGCAAGTGCAGGGGG
CAGGATCTCTCTGTCTACCTCCTACCTTGCTCTGGCGAGAAAGATGATACCATCATGGC
TGATGCAATGCCGGCGCTGCATAAGCTTTGATCCCGCTACCTGCCCATTCGACC
ACCAAGCGAAACATCGCATCGAGCGAGCACGACTCGGATGGAAGCC
GGTCT
TGTCGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAA
CTGTTCGCCAGGCTCAAGGCGCGCATGCCGAGGGCGAGGATCTCGTCGTGA
CCCATGGCGATGCTTGCTTGCCGAATATCATGGGTAAGGAAATGGCCGCTTTTCT
GGATTCATCGACTGTGGCGGCTGGGTGTGGCGGACCGCTATCAGGACATAG
CTTGGCTACCCCTGATATTGCTGAAGAGCTTGGCGGGCAAT
GGGCTGACCG
CTTCCTCGTGTCTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCT
ATCGCCTTCTTGACGAGGTTTCTTCTGAGCGGGGACTCTGGGGTTCGAAATGACCG
ACCAAGCGACGCCAACCTGCCCATACGATGGCCGCAATAAAATATCTTTAT
TTTCATTACATCTGTGTGTGGTTTTTTGTGTGAACGTACAGTGCAGGGGCT
CCGTAATCATGGTCATAGCTGTTCCTCTGTTCTGTAAGTAAATCTCCGACTC
GGCAGTGAACGTAACTCAACTCATATGCGTGTGCCTCAGTCCCGCTTTCCAGT
CGGGAAACCTGCTGTCGAGCTGCACTATGAAATCGGCCAACCGCAGGGGAG
AGGCCGTTTTCGTTTTGCGCTCTTCCGCTCCGCTCAGGACTGACTGGCGCTGC
GCTCGGTCTCAGGCAGCGAGCAGGTATACGTGCTCAGCTACAAGGCGTAAT
ACGGTTATCCACAGAAATCGGGGTAAACGAGGGAAGAACATGTGAGCAAA
AGGCCAGCAGAAAAAGGCAGAGAAGCGTACAAAAGGGCCCGGTGGTGGCGTTTTC
CATAGGCCTCCGCCCGGCTGACGAGCATCACAAGAAATCGACGCAAGTACGA
GGTGCCGAAACCCGACAGGACTATAAGGATACACGAGCGTTTCCCCCTGGAAG
CTCCCTCGTGCAGCTCTCCGTTCGGCTCCGCGTGATACACCTGTCCG
CCTTTCTCCCTCGGGAAGCGTGGCGCTTTTCCTCATAGCTCAGCTGTAGGTAT
CTCAGTTTCGATTGGCTGTTCTCGCTCAAAGCTGGGGCTGTGTCAGGAACCCCC
CGTTAGCCGCCAGCGCTGCGCTTTATCCGGTAACGTGTTTGGAGCTCAACC
CGGTAAGACAGCATTATCGCCAGCGCAGCCAAGCTGTGATACAGGAATTAG
CAGAGCGAGGTATGTAGGCAGTTCTGCTACAGAGTTTTGAAGTTGGTGCGCTA
TACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAG
TTACCTTCGGAAAAAGAGTTGGTATGCCTTTGATCCGGCAAACAAACCACCGC
TGTTAGCGGTTTTTTTTGTGGCAAGCAGCAGATTACGCGCAGAAAAAAA
GGATCTCAAGAAGATCCTTTGATCTTTTCTAAGGGGCTCTGACGCTCAGTGGAA
CGAAAACTCAGTTAAAGGGATTTTGGTCATGAGATTATCAAAAGGATCTTC
ACCTAGATCCCCAAAAATGAGGTTTAAAATCAAATCTAAAAGTATAA
TGAGTAAAACCTGGTCTGACAGTTACCAATGCTTAAATGAGTGACGACCTATCT
CAGCGATCTGTCTATTCGTTTACATCCATAGTGCTGACTCCCCGTCGTGAG
ATAACTACGATACGGGAGGGCCTACCATCTGGGCCATGCATGCAATGATAC
CGCGAGACCACGCCTCCAGGCTCCAGATTTATCAGCAATAAAAAAAAAACGCGAC
CGGAAGGGCCGAGGCAGAACAGTGCTCTACAAGAGGGAGTTTTTGCAAGCTCCAG
TCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTT
GCACAACGTGGTTGCCCATTGCTACAGGCACTGTGTTGCACGCTGCTCGTTG
GTATGGCTTCATTCCAGCTCCGGTTCTCCCAACGATCAAGGCGAGTTACATGATCC
CCCATGTTTGCTCAATAGCCTCTCCCAGATCAAGGCGAGTTACATGATCC
AAGTAAAGTTGGCCGCAGTGTTATCACTCATGTTATGGCAGCAGCTGATAATT
CTCTTACTGTCATGCCATCCGTAAGAGTATTTTCTGACTGTTGAGTACTCA
ACCAAGTCATTTCTGAGAATAGTTGATGCGGCGACGCGATGTGCTTGGCCCGG
CGTCAATAACGGGATAATACCCGCGCCACATAGCAGAATTAAAAATGCTCAT
CATTGGAAAAACGTTCCTTCGGGCGGAAAACCTCTAAAGGATCTTTTCCCCT
AGATCCAGTTGATCCACCCACACTGTCGACCCAACTGATCTTCAGCATCTTT
TACTTTCACCAGCTTTCTGAGGCTGAGCAAAAAACAGGAAAGCCAAATGGCCGCA
AAAAAGGGAATAAAGGGCCGACACGGGAAAATGTTGGAATACTCATACTCTTT
TTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATA
TTTGAATGTATTTAGAAAAATACAAAAATAGGGGTCCCGCGACATTTCCCG
GAAAGTGC

> pNEBR-X1Gluc

Plasmid obtained from NEB.

> pNEBR-X1Gluc

GTTCTAGAGGATCTTTGTGAAGGAACCTTACTTCTGTGAGTACTAAATTTG
ACAAAACCTACCTACAGAGATTTAAGCCTTAAGGTAATATAAAAAATTTTAA
TGTATAATGTGTTAACACTACTGATTTCTAATTGTGTTATTTTAGATTTCAAC
CTATggGACTTTATGGAATGGGAGCAGTGATGATGAGGCTACTGCTGACTC
TCAACATTCTACTCTCAGAAAAGAAGGATAGAGACCCAGGACTTTCCCT
TCAGAATTGCTAAGTTTTTTGAGTCATGCTGTGTTTAGTAATAGAACTCTTTGC
TTGCTTTGCTATTACAAACCACAAAGGAAAAGCTGCACTGCTTATACACAGGAA
AATTATGGAAAAATATTCTGTAAACCTTTTAATAAGTGGGCATAAACAGTTGATA
ATAACATACCTTTTTCTTACTCACCACAGGACATAGTGTCTGTCTATTAAT
AACTATGCTCAAAAAATTGTGTACCTTTTAGCTTTTTTAATTTGTAAGGGGTTAA
TAAGGAATATTTTGATGTATAGTGCTTTGACTAGAGATCATATAATCACAGCCATA
ACATTTTGAGAGGTATTTTACTGTCTTTAAAAACCTCCCCACACCTCCCCCCTGAA
CCTGAAAACATAAATGGAATGCAAATGGTGGTTGTTAATTTGTCTTTATTGCAGCTT
ATAATGTTACAAATAAAGCAATAGCATCACAATTTTCACAAAAATAAAGCATT
TTTTTCACGTGATTCTAGTTGTGTTGTCTCCAAACTCATCAATGTAATTTATCA
TGTCTGGATCAACTACGTCAGGTGGCACTTTCCGGGGAATAGTGCGCGGAAC
CCCTATTGTCTTTATTTTTCTAAATACATTCAATATGTATCCGCTCATGAGACA
ATAACCCTGATAAAATGCTTCAATAATATTGAAAAAGGAAGAGATAGTATGAGATT
CAACATTTCGCTGTCGCCCCTATTCCCTTTTTTGCCTTTTGGCTCTCTGGTTT
TTGCTCACCAGAAACGCCTGGGTAAAGTAAAGATGCTGAAGATCAGTTGGG
TGCACGAGTGTTTACATCGAAACTGGATCTCAACACGCGGTAAGATGCCCTGAG
AGTTTTTCGCCCAAGAACGCTTTCTCCCAATGATGAGCCTTTTTTAAAGTTGCTGCT
ATGTGGCGCGCGTTATTATCCCTGTTGACGCGCGGGCAAGAGCAACTCGGTGC
CGCATACACTATTTCGGAGATGACATTGAGTTGGACTACCTACACAGTCACAGAAA
AGCATCTTACGGATGGCATAGCAGTAAGAGAATTATGCGATGCTGCGCATAAC
CATGAGTGATAAACCATTGCGCCACATTACTTCTGACAACGATCGGAGGACC
AAGGAGCATAACCAGCTTTTTGCAACAAATCGGGGATCATGTAACCTCACGGCTT
ATCGTTGGGAACCGGAGCCTGAATGAAAGCCATACAAACGACGACGTCGACA
CCACGATGCTGTAGCAATGGCAACAAACGTTGCAGCAAAACTATTTAATCTGGC
ACTACTTTACTCTAGCTTTCCGGCCAAACAATTAATAGACTGGATGGAGGCCGA
AAAGTTTCAGGCCCTCTCGGTCTCGCTTGCCCTCTCCGCTGCTGGTTTATTGC
TGATAATCTGGAGGCCTGGTGAGCTGGGTCTGCGTGATCATATTGCCAGCACTG
GGGCCAGATGGTAAGCCCTCCTCCGTATCGTAGTTATACACGACGGGGAGAGTC
AGGCAACTATTTAGATGTAAGCCAATAGACAGATCGCTGAGATAGGCTCAAAGT
GGCAACTATTTTACCTTCAATTTAATTTTATATTTTTAGATTTTG
ATTACCCGTTTGAATAATCGGAAAGGCCCAAAAAACAGGAAGATTTGTAAG
GCAAATTAAAAATGTGAAAGTTATTATTATTTGTTTAAATTCGCGTTAAAATT
TTGGTAAATCAGCTCATTTTTTAAACCATAAGGCCAAATCGGCAAATCCCTT
ATAAATCAAAAAAGATAGCCCGAGATAGGGTTGAGTTGTTTCCAGTTTTGGA
CAAGAGTCCACTATATTTAAAGAAGCTGGACTCCAACGTCAAAGGGCGAAAAAC
CGTCTATCAGGGCGATGGCCACTACGTGAACACCATCACCAAAAATCAATTTTT
TGGCGAGTGGCCGAAAAGCACTAAAATCGGAACCCTAAAGGGAGCCCCGGATTT
AGCTTTAGGCTTGACGGGGAACAGCGAGACTGGCAGAAAGAGGAGAACGTGC
GTTTGAAATCGCGCTATTTTCTGAGATCCTTTTTTCTGAGATCCTTTTTCT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
TGAGATCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTTTCTGAGATCCTTTT
GCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCT
TACCGGGTTGGAGCTGAAGCTATGAGAAAGCAGCCACGCTTCCGAAGG
GAGAAAGGCGGACAGGTATCCGGTAAAGCGCAGCGGTCGGAACAGGAGAGC
CACGAGGGAGGTTCCAGGGAAACGCTTGTTATCTTTATAGTCTCGGG
TTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGG
GAGCCTATGGAAAACGCACGCAAAGCGGCTTTTTACGGTTCCTGGC
GCTGGCCTTTTGCTCACATGTATCAGCCATATCACATTTGTAGTTTTACTTGCTTTAAAAAACCTCCCACACCTCCCCCTGAACCTGAAAATGAGT
CAATTGTTGTTAACCTTGTATTGAGCTTTATAATGGTTTAAATAAAGCAAT
AGCATCACAATAATTCACACAAATAAACGATTTTTCTACTGACATTCTAGTTGTT
TGTCGAAAACATCATCAATGTATTTTATCATGTCTGTGATCGCCATATACACATTT
GTAGAGGTTTACTTGTCTTTAAAAAACCTCCCACACCTCCCCCTGAACCTGAA
ACATAAAATGAATGCAATTGTTTTAATTGTTTTATTGCCAGCTTATAATGGTT
ACAATAAACGCAATAGCATCACAAATTTTCACACAAATAAACGCTTTTTTCAC
GCATTCTAGTTGTTTTGTCCAAACTCATCAATGTATCTTTATCATGTCTGCTC
TGTAATGTGAGTTAGCTACTCATATTAGGCAACCCAGGCTTTACACTTATGTCT
TCGGGCTCGTATGTTGTGGAATTGTGAGCGGATAACAATTTCACACAGGA
AACAGCTATGACCATTACGCCAAAGCTACGAGGCGGAGTACTGTCTCCTCCGA
GCGGAGTACTGTCTCCAGGGGAGTACTGTCTCCTCCAGGGGAGTACTGTCC
TCCGAGCGGAGTTCTGCTCTCCAGGGGAGTCTAGAGGGTATATAATGGG
GGCGGCTAGCTACTACCAGAGCTCATGGGCGCGGCGATAATGCGTGG
CGCCAAGCTTGGTACCGAGCTCGGATCCAGCCACCATGGGAGTCAAAGTTCTGGTGCCCTGATCTGCATCGCTGTGGCCGAGGCCAAGCCCACCGAGAACAACGAAGACTTCAACATCGTGGCCGTGGCCAGCAACTTCGCGACCACGGATCTCGATGCTGACCGCGGGAAGTTGCCCGGCAAGAAGCTGCCGCTGGAGGTGCTCAAAGAGATGGAAGCCAATGCCCGGAAAGCTGGCTGCACCAGGGGCTGTCTGATCTGCTCTGTCCCACATCAAGTGCACGCCCAAGATGAAGAAGTTCATCCCAGGACGCTGCCACACCTACGAGGCGACAAAGAGTCCGCACAGGGCGGCATAGGCGAGGCGATCGTCGACATTCCTGAGATTCCTGGGTTCAAGGACTTGGAGCCCATGGGAGCAGTTCATCGCACAGGTCGATCTGTGTGGACTGCACAACTGGCTGCCTCAAAGGGCTTGCCAACGTGCAGTGTTCTGACCTGCTCAAGAAGTGGCTGCAGCAACGCTGTGCGACCTTTGCCAGCAAGATCCAGGGCCAGGTGGACAAGATCAAGGGGGCCGGTGGTGACTAAGCGGCCGCTAGCGATCGCGGCGCGCCACGCCTCTCCGGATGTACAGGCATGCGTCGACCCTCTAGTCAAGGCCTATAGTGAGTCGTATTACGGGACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAACCCTGCGTTACCCAACTTAAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCTTCCCAACAGTTGCGCAAGGCTTCGGCGGGCTTTTTTTT
>pNEBR-X1Hygro

Plasmid obtained from NEB.
AACTATGCTCAAAAAATTGTGTACCTTTTAGCTTTTTAAATTGTAAAGGGTTA
TAAGGAATATTTTGTATGTAGTGCTTTGACTAGAGATCATAATCAGCCATACC
ACATTTGTAGAGGTCTTACTCGCTTTAAAAACCTCCCAACACTTCCCCTCTGAA
CCTGAAACATAAAATGAATGCAATTTGTTGTTGTTAATTTGTTATTCAGCTT
ATAATGGTTACAAATAAAGCAATAGCATCACAATTTCCACAAATAAAAGCATT
TTTTCACTGCATTCTAGTTGTGGTTGTCCAATCTCATCAATGTATCTTATCA
TGTCTGGATCAACTACGTCAAGTGGCAGGTAACCTGGGAAATGGTACCGGAAC
CCCTATTTGTTTTTATTTTCTAATAACATTCAATAATATTGAAAAAGGAAGAGATGAGTATGTATT
CAACATTTCCGTTGCTGCCTCTATTCCCTTTTTGCGGCATTTTGGCCTTCTCTT
TTGCTCACCCAGAAACGCTGGTGAAGTTAAAAAGATGCTGAAGACCATCTTGGG
TGCACGAGTGTTGTTACATCGAAACTGGAATCTCAACAGCGGTAAAGATCTTGGAG
AGTTTTTCGCCCCCAAGAAACGTCTCTCCTCAATGATGAGCACTTTTTAAGGTCTGCT
ATGTGGCGCCGTATTATCCCCGTGTTGACGCCTGGGAAGACGAAACTCCGTGC
CGCATACACTATTCTCAGAATGAATCTGGTTGAGTACTCACCAGTACAGAAA
AGCATCTTACGGATTGGCATAGCTAAAGAGAATTATGCAGTGCTGCCATAAC
CATGAGTGATAAACACTGCGCCAACACTTACTTCTGACAACGATCGGAGACCG
AAGGAGCTAACCCTTTTTGCCACAACATGGGGAATCATGTAACCTCGCCTTG
ATGGTTGGGAAACCGGAGCTGAAATGAAAGCCATACCAAAACGACGAGCATTGAC
CCACGATGCTTGTAGCAATGGCAAACCGTGTGCGCAAATATAACTATCGCGA
ACTACCTACTCTAGCTTCCCAGCAAATATTAGACTGGATGGAGGCAGGAT
AAAGTTGCAGGACCACCTTTCTCGCTCGGCCCTTCCGGCTTGCCTGGTTATTCG
TGATAAATCTGGAGCCCGGTGAGCGTGGCTCTCGCGGTATCTGATTGACGACTG
TATCAGGGCGATGGCCCACTAGTGAAACCATCACCACAAATCAAGTTTTTTGG
GGTCGAGGTGGCCGCTAAAGCAGCTAAATCGGAACCTAAGGAAGGGAAGGGGACT
TAGAGCTTGACGCGCCGCTAAATGCGGCCGCTACAGGGCCTGATTAAAGGAT
CTAGGTGAAGATCTCTTTTGATATACTTCTCATGACAAAAATCCCCAACTGCTAGT
TTTGTTCACGGAGCTGACAGCCCCGAGTAAAGAGATCAAAGGATCCTTCTTGA
GATCTTTTTTTTCTGCGCTAATCTGTGCTGTTGCAAAAAACACAAAAACCCGCT
ACCAGCGGCTGTTTGGGGATCAAGAGCTCTCAACTCTTTTTTCCGAAGG
TAACCTGGCTTCAAGCAGGGCAAGATACCAATATCTGTGCTGCTCTTACCCG
GGTTGGACTCAAGAGCATTACCGATTACGAGATTTTTGTGATGCTCGTCAGGGGGCGGAGCC
TATGGAAAAACGCCAGCAACGCCCTTTTACGGTTCCTGGCTTTTGCCCTTTGCTCACATGTATCAGCCATATCACATTTGTAGA
GGTTTTACTTGCTTTAAAAAACCTCCCACACCTCCCCCTGAACCTGAAACATAAAATGAATGCAATT
GTTGTTAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCAT
CACAAATTTCCACAAAATAAAGCATTTCACCTGACATTTGCTGTTTTGTCC
AAACTCATCAATGTATCTTATCATGTCTGGATCAGCCATATCACATTTTGTAGA
GGTTTTACTTGCTTTAAAAAACCTCCACACCTCCCCCTGAACCTGAACATA
AAATGAAATGCAATTGTTGTTAACTTTATTGCAGCTTTATAATGGGTACAAA
TAAAGCAATAGCATCACAATAATTTCAAAATAAAGCATTTTTTTTACTGCATTTC
TAGTTGTGGTTTGTCAAAAAATCATCAATGTATCTTATCATGTCTGCCCATGTAAT
GTGAGTTAGCTCACTCATTTAGGCACCCCCAGGTCTTTACACTTTATGCTTTCCGAC
TCGTATGTTGTGGTAATTTGTAGCGGATAACAATAATTCACACAGGAAACACGC
TATGACCATGATTCGCAAGCTACGGGCGGAGTACTGTGCTCTCCGACGGAGG
TACTGTTCCCGAGCGGAGTACTGTGCTCTCCGACGGGAGTACTGTGCTCTCCGAGC
GGAGTTGTTGCTTCCCGAGCGGAGACTCTAGAGGATATATAATGGGCCCCGGTC
AGCTACTACAGAGGCTCAGGCCCCGTTGCAATTGAAGCCGGCTGGGCAGCAAG
CTTCTCTGCAGGATATCTGGATCCACGAAATTCTTAATTAGCCGGGCTCTGAG
CGGCCGCTAGCGATCGCAGCGCGCCACGCGTCTCCGGATGTACAGGCATGCC
TCGACCCTCTCATGCAAAGGCTATATAGTGATCTGTAATACGGGACTGGCGTCTT
TTCAACGTCGACTGCGAGAAACCTTGCCGTACCACCCTTAAATCGCCTTGC
AGCACATCCCCCTTTCCGAGCTGGCGTAATAGCGGAAGAGGCCGCCAGCAG
CGCCCTTCTCCACAGTTTCGCGAGCTGATGCGGAAATGCGCGCTTCGCTTTGGT
AATAAGGCCGGCTTCCGGCGGGCTTTTTTTT
>pGEM-T

Obtained from Promega.

GGGCGAATTGGGCCGACGTGCATGCTCCCCGGCGCCCATGGCCGCGGGATA
TCACTAGTGCGCGCCGCTGCAAGTCGACCATTATGGGAGAGCTCCCAGCGGT
TGGATGCATAGCTTTAGGTATTCTATAGTGTCACCTAAATAGGCTTTGCGTAATC
ATGGTCATAAGCTTTCTCGTGAAAATTGTATCCGCTCACAATTCCACACA
ACATAAGCGGAGGATGAAAAGTGAAAGCCTGGAAGGTGCCTAATGAGTGA
GCTAACTCAGGATTTGCGTGGATGCTCAGCTTGGCCTTTCCAGTGGGAAAC
CTGTCGTGCCAGCTTATGAAATCGCGGCAAACGCAGCGGGAGAGGCGT
TGCGTATTGGCGCTCCGCTCTGACTCGCTGCTGCGCTC
GTTCGGCTGCGGCGAGCGGTATCAGCTCACTGATAAGACATGTGAGCGAAAAGAACATGTGAGCAAAAGGCCAGC
AAAAGGCCAGGAAACGTTGAGCTGCTGATGCTTGTCTTGTTCATAGGCT
CGCCCCCCTGACGAGCATCACAACAAATCGACGCTCAAAGTCAGAGGTGGCGA
AACCCGACAGGACTATAAAGATACCAGGGCTTTCCCCCTGGAAGCTCCCTCG
TGCCTCCTCTGTACACCGCTTACCCGGATATCTGTCCCGCCTTTTCTCC
CTTCGGGAAAGCGTGCCGCTTTTCTCATACTACGCTACGCTGTTAGGTATCTCAGTTCG
GTGTAGGTGCTTCCGCTCAAGCTGGGCTGTGTCAGCAAACGGTAAAGAC
CGACTTATCCCAACTGCGACGAGCAGACTTCTGGTAACAGGATTAGCAGAGCGAG
GTATGTAGGCGGTGCTACAGAGGTTTCTGGAAGTGGTGGCCTAACTACGGCTAC
ACTGAAAGAACAGTATTGGTATCTGCGCTTGTGAAAGCCAGTTACCTTTCGG
AAAAGAGTTGGTAGCTTCTCTGATCCGCCAACAACACACCAGCTGTTAGCGGT
GTTTTTTTGTGTGTAAGCGCGAGATTACGCGCAGAAGAAAAAGATCTCAAG
AAGATCTTTTGTATTCTTCTACCGGGGTCTGCTACGTAGGAAAGAAAACCTCA
CGTTAAGGGATTTCGGTCATGAGATTATCAAAAGGATCTTCACCTAGATCTCT
TTTAAATTAAAAATGAAGTTTTTAAATCATAATCTAAAGTATATATAGTAAAACCTT
GGTCGTGACAGTTACCAATGCTTAATCAGTGAGGCACCTCTACGTACGTCTGT
CTATTTCGTTCTACATCGGGTCTGACTCCCTCCCGTGCTGTAGATAACTACGAT
ACGGGAGGGGTTACCATCTGGGCCCCAGTGTGGAATGATACCAGCAGAACC
CGCTCACCCCGCTCACATTTATCCGCAAATACCCAGGCGCCAGCCACTGAGG
AGGCAGAAAGTGTCCTCTCGCAACTTTTATCCGCTCCATCCAGCTTTAATTGT
TGCCCCGGAAGGATAGAGTTAGTGTCGCCAGTTAATAGGTTGTCGCGGAACGGT
TTGCTACGGCATCCGTCGTTGTCAGCTCGTCGTTTGTGTTGATGCGCGCTTCA
TTCAGCTCCGGTTCCAAACGATCAAGGGCGAGTTACATGATCCCCCATGTGGT
ACGCCAGGGTTTTCCCATCAGACGGTTGTTAAAACGACGGCCAGTGAAATTGT
AATACGACTCACTATA

>pPRO1.4GO Progerin-gfp

Obtained as integrated in HeLa, gift from Dr. Misteli. Progerin Reporter integrated in HeLa.

>pPRO1.4GO Progerin-gfp

TCAATATTGGCCATTAGCCATATTATATCGATTGTTATATAGCATAAAATCAATA
TTGGCTATTGGCCATTGCATAGTTGTATCTATATCATAATATGTACATTTATA
TTGGCTCATGTCCAATATGACCCGCACTGTTGGCATTGATTATTTGACATTTAT
TAATAGTAATCAATTACGGGCGTCAATTAGTTCAATAGCCCATATATGGAATGTCG
CGTTACATAACTACGTTAAATGGCCCCGCTGGCTGACCGCCCAACGACCCCA
CGCCCATTTGACGTCAATAATGACGTATGTTCCCATAGCCCATATATGGAGTTCG
CTTCCATTGACGTCAATGGGATGATTTACGGGAATTTACGTTAAACTGCCCCACCTGGCA
GTCATCAAGTGTATCATATGGCAAGTCGCCCCCCTATTGACGTCAATGACGG
TAATGGCCCGCTGGCCATTATGCCCCAGTGACCATCTACGGGACTTTCCTA
CTTGCAGTACATCTACGTGATTAGTCAATGCTATCCATATCAGTTGACGTGTTT
TGGCAGTACCAACCATGCGTGTTAGCGGCGTTAGTCAGCTACCCGGGATTTCAA
GTCTCCACCCCCATGACGTCAATGGGAATTTGTTGCTGGCCACAAAAATCAACCG
GACTTTCCAAAATGTGCACGAAAATCTGCACGGCTCGCAGCCCGCCGCTTGACGCAAT
GGGCCGTTGGCTGCTACCTGGGAGGTTCTATATAAGCAGAGCTCGTTAGTG
AACGTCAGTCACTAGAAGCTTTATTTGCGTGATTTTATCAGTTAAATTGC
TAACGACGTCAGTGCTTCTGACACAACAGTCTCGAACTTAAGCT
TCTCTAAGGTAGCCCTTGCAGAAAGTTGGCTCGTGAGCAGTGGCAGGTAAGT
ATCAAGGTTACAAGACAGGTTTAAGGAGACCAATAGAAACTGGGCTTGTCGA
GACAGAGAAGACTCTTGCGTTTCTGATAGGACCTATTGGGTTACTGATCATC
CACTTTGCCTTTTCTCCACAGGTGTTCCACTCCCAGTCAATTACAGCTTTAA
GGCTAGAGTACTTAATACGACTCACTATAGGCTAGCGCTACCCGGACTCAGAT
CTCGAGGTACGGCTCTCATCAACTCCACTGGGGAAGAAGTGGCCATGCGCAAG
GCTGGTGCGCTCAGTGACTGTGGTTGAGGACGACGAGCATCAGGATGGAGAT
GACCTGCTCCATCCACCACCACGGCCTCCCACGACGCTCGGAGGGACCGAAG
CTGAGCTCAACCTGCCTGCAGGCTCAGGAGC
CTCTCTGCTTCTTCGCTCCAGGCCACCTGGCAGCTGCCAGTGTGTTGGGGCT
GGGGCTGCTGTCACCCAGCCGCTGCTGAGCTTTCTGCTCTTGCATCGGACAGG
GACGAGGAGAGCCACTGGCCTTCTGCAGGGGAGAGGGAGGAGCTGGGTCCAGGC
CCTGCTGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCAT
CCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGC
GAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAAGTTCATCTGCA
CCACCGGCAAGCTGACCCTGACCTGGCCCTTCTCGTGAACCG
CATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCA
ACATCCTGGGGGCAACAAGCTGAGTACAACCTACACAGCCACAACGTCTATAT
CATGGCCGACAAGCGAGAAGCCGCATCAAGGTGAACCTTCAAGATCAGCCCA
CAACATCGAGGAGGCCAGCAGCTGACGCTCGCCGACCCACGACTACGCAGAACAC
CCCCATCGGCGACGCGCCCGCCTCCGGCTGCTGGCTGCCCGACAACCACTACCTGAGCAC
CAGTCCGCCCCTGAGCAAAAGACCCCCAAGGAGCGGATCAATAGGTCTCTGC
TGGAGTTCTGTGACCACCGCGCAGGATCACCTCTCCGCCATGGAACGAGCTGTAACAA
GAAGCTTAGCCATGGCTTCCCAGGGAGATGGACCAGCTCACCTGCACGCT
GTGCTTCTGTAGGATCAATGTGATGCGGCGCCCGGCTGCCCTTCTTCTTATGAGGGTTA
ATGCTTTCGCAGACATGATAAGATACATTTGATGAGTTGACAGGGGAGATGTGGGAGGTAAAAAGAAATTTGATGAGTTGACAGG
AGAATTGCAAGAAAAATGCTTTTATTTTGTGAAATTTTGTGAGTCATTTGCTTT
ATTTTGACATTATAAGCTGCAAATAAAAGTTAAAACACAAAGTGCATT
CATTCTATGTTTGTCAGGGGAGATGTGAGGAGGTTTTTAAAAGCAAGTA
AAACCTCTACAATGTGGTAATAACCGATAAGGATCGATCCGGCTGGCGTA
ATAGCGAAGAGCGCCCGACAGCGATCGCCCTTTCCAAACAGTTGCGCAGCCTGAA
TGGCGAATGGAGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGG
TTACGCGCAGCGTGACCCTACACTTGCCAGGCCTTGAGCGCCCGCCTCCTTTC
GCTTTCTTCCCTTCTTTTGCTGCGGCTGCTTTCCCCGATAGCTCTTA
AATCGGGGGCTCCTTTTTAGGTGGTCCGATTAGGCTTTTACGACCTCGACCG
CAAAAAACTTGATTTGGGTGATGGTTCACGTAGTGGGCCATCGTCTAGTGGAGTCCACGTTCTTTAATAGTGGACTCTTTG
TTCCAAACTGGAACAAACACTCAACCCCTATCTCGGTCTATCTGGGTATTTTGAATTTATA
AGGGATTTTGCCGATTTTCGGCTATTGTTAAAAATAGGCGTATTTAAACAAA
TATTTAACGCGAAATTTTTAAAAATATTAACGTTTCAATTTCGCCCTGATGCG
GTATTTTTCTCTCTAGCAGCTCTGTGCAGGGTTATTTCAACACGCATACGCAGATCTG
CGCAGCACCATTGCCCTGAATAAATCCTCAAGAAAGAGGAAACTTGTTTAGGGTACC
TTCTGAGGCAGAAAAAGCAGCTGTTGGAATGTGTCAGTAGTGGGTGTTGAA
AGTCCCCAGGCTCCCCAGCAGGCCAAGATGATGCAAAAGCATGATCTCAATTA
GTCAGCAACCCAGGTTGGAAGTCCTCCCCAGCGCTCCCCAGCGAGCAAGTAGTATG
CAAAAGCATGCTCCTCAATTTAGTGAGCAACCATACTCCGGAATGGCTGCA
TCATCCCCGTCAACTCCGGCAGTTCCGGACTTCCGGGCTTGGAGGTACC
AGTCCCCAGGCTCCCCAGCAGGCCAAGTATGCGCATCTCAATTTAGTCCAGCCAGCCAGGCTGCA
CCAGAAGTGATGAGGGAGCTTTTTTGAGGTGGCCTAGGCTTTTGCAAAAAAGCTT
GCCCAAACCGGATCCACCCGATCAGCAGCTGGCTGGCTCTTCTCCGAGAAGCTC
TCACCCGAGTTCCATGCCTTCAAGGTCGCGTCAAGGTGGAGGACCCGTGAAACGGCCA
CGATTTGAGATCGAGGGCGAGGGCGAGGGCCGCCCAGTAGGGCCACAA
CACCGTGAAGCTGACCCATAGGGACCCAGGCGCCCGCTCAGGGCTGGGAGAC
ATCGTCTCCCCAGTTCAAGGTCGCTTCCAGGCTGAAGGTGTACGTGAAGCCTCGCC
CCGACATCCCCCGACTCAAGAAGGCTGCCTTCCCCAGGGCTTCAAAGGGGACG
GCGCTGATGAACTTCCAGGAGCGGCGGCGGTGCGGACCGTGACCAGGACTCC
GTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCC
GAAGAACCGTTTTTCAATGATGAGCAGCTTTTTAAAAGTTCTGCTATGTCGGGAGGT
ATTATCCCGTATTGACGCGCGGGAAGACAGCAACTCGGTCGCCGCATACACTATG
TCTCAGAATGACTTGGTTAGTACTCACCAGTCAGAGAAGGATCTTTACGG
ATGCGATGACAGTAAGGAGAATTATGCAGTGCCTGCCATAACCATGTAGTGATAAA
CAGTCGCGCCAAACTATTACTTCTGCAACAACGATCGGAGGACCGAAGGAGCTAACG
CAGCTTTTTTCCAACACATGGGGAATACATGTAACCTCGCCCTTGATCGTTGGGAACC
GGAGCTGAATGAACGCCATACAAACAGACGAGCGTGACACCAACGACTGCCCTA
GCTTCCCCCGCAAAATAAATAGACTTGAGTGAGGAGCCGGAATAAACGTTACAGGAC
CAGTCGCTGCTGGCCCTTCCGCGCCTGCGGCTTTTTATTGCTGATAAAATCTGGA
GCCGTTGAGCGTGTTTCTCGCGGTATCATTCGAGCAGCCTGGGCGCAGATGGTA
AGCCCTCCTTCCGTATCGTAGTTATCTACAGCGACGGGAGTTCAGGCAACTTGAA
TGAACTGAAATAGACAGATCGCTGAGATAGGTCCTCACCAGATTAAGCATTGG
TAACCTGTCAGACAAATGTTTACTCATATATACCTTTAGATTGGATTTAAAACCTCA
TTTTTAATTAAAAAGGATCTAGGTAAGATCCTTTTTTGATAATCTCATGACCA
AAATCCCTTAAACGTTGGTTTTTGCTTCCACTGAGCGCTACGCCGATGGAAAGG
GATCAAAGGATCTTTCTTGAGATCCTTTTTTTCTGCAGGTAATCTGCGTCTTGC
AACAAAAAACACCGCTACACAGCGGTGTGTTTGTGTTTGCAGGATCAAGAGCTA
CCAACCTTTTTCCGAAGGTAACCTGTGGGCTTACGAGAGCGACAGATACAAATA
CTGTCTTTTCTACAGTGGATGCTGAGTTAGGCGCCACCAGCCTTTCAAGAACACTTGAGCA
CCGCTACATACCTCGCTCTGCTAAATCCTGTACCGATGCTGGCTGCGCAGTGG
CGATAAGTCGTGCTTTACCGGGGACTCAAGACGATAGTTACCGGATAAG
Plasmids created in this study or by the lab

>pDual-EGFP

Lentiviral vector derived from pNL-EGFP/CEF with T7 promoter from pET28a inserted. Verified to direct expression of EGFP in both bacteria and mammalian cells.

>pDual-egfp
TCGGCACCAGTTTCGTGAGCCGGAAAGATGGCCGCTTCCCGGCCCTGCTGCAG
GGAGCTCAAAAATGGAGGACCGCAGCTCGGGAGAGCGGGCGGGTGAGTCAC
CCACACAAAGGAAAGGGCCTTTCCGTCCTCAGCCGCTTCATGTGACTCACG
CACGGAGTACCAGGGCAGCCAGCTCGAGCACTTCGATTAGTTCTCGAGCTTTTGG
AGTACGTCTCTTTAGGTTGGGGAGGGGAGGTTTTATGCATGGAGTTTCCCCCA
CACTGAGTGGGTGAGACTGAAGTTAGGCGACGCTTGAGACTGATGTTCTAATTAC
TCCTTGGAATTGGCCCTTTTTAGGGGTGATCTTGGTCTCTTCAAGCGCTTACG
ACAGTGTTCAAATTTTTTTTCTCTCAATTTCAGGCTGTTGCTGAGGAGATTCTGCA
GTCGACGGTACCAGGGCAGCCAATCCTCGATCCCGCAATCCCGAACATATAATACGACTCA
CTATAGGGGAATTGTGAGCCGATAAACAATCCCCTCTAGAAATAATTGGTTG
AACTTTAAAGAGGAGATATACCATGGGCAGCAGCCATCATCATCATCATCAC
AGCAGCGGCCTTGGTGCCGCGCGCCAGCCATATGGCTAGCATGACTGGTGGAC
AGCAAAATGGGTGCGCGGGATACCCATCGCACCACATGGTGAGCAAGGCCG
AGGAGCTGTTCACCAGGGGTTGGTGCCCATCTCTGGTGAGCTGGACGGCGACGT
AAACGGCCACAAGTTCAGCTGTCGGCGCGAGGGCGAGGGCGATGCCACCTA
CGGCAAGCTGACCCCTGAAGTTCCATCTGCAACCACCGCAAGCTGCCCGTGCCC
TGGCCACACCTCGTGACCACCCCTGACCTACCGGCGTGCAGTCTCTGAGCCGCTA
CCCCGACCACATGAAGACGACACGGGACCTTTCTTCAAGTCCGCGCATGCGCCGAAGGC
TCGTCGAGGACGCCACCATTTCTTCCAGGACCCGCAACTACAAAAGACCC
GCAGCGAGTTGAGGTGCTCGAGGCGCAACCCCTGTTGAAACCGCATCGAGCTGA
AGGGCATCGACTTCAAGAGGACGCGAACAATCTCTGGGCAACAAGCTGGAGTT
ACAACTACAACAGCCACACGTCTATATCATGCGCCGACAAGCGAAGAAGAC
GCATCAAGGTTGAACCTTCAAGATTCGCGCCACACATCGAGGAGGCCGAGCGTGC

GCTCGCCGACCACCTACCAGCAGAACAACCCCCCATCGGCGACGGCCCCGTGCTG
CTGCCCGACAAACCACCTCTGAGCACCACGTCGCCCTGAGCAAAGACCCCA
ACGAGAAGCGCGATCACATGGTCTCTGCTGGGATTTCTGACCCGCCGCCGGGAT
CACTCTCGGCATGGACGAGCTGTACAAGTCCGGGACTCAGATCCGACCTCGAG
ACCTAGAAAAACATGGAGCAATCAAAGTAGCAATACAGCAGCTAACAATG
CTGCCTTTGTGCTTGCTAGAAGCACAAGAGAGGAAGGAGGTGGGTTTTCCAGT
CACACCTCANGTACCTTTAAGACCAATGACTTACAAGGCTAGCTGATCTT
AGCCACTTTTTAAAAGAAAGGGGACTGGAAGGGCTAATTCCACACACACACAGCTACTTCCT
GATTGGCAGAATACACACACCAGGGCCAGGGGTCAGATATCCACTGACCTTTG
GATGGTGCTACAAGCTAGTACCAGTTGAGCCAGAT
Control vector derived from pCMV-ICP8 made by deleting ICP8.

```plaintext
>TCAATATTGGCCATTAGCCATATTATTTCTATTAGGTATATATAGCATAAATCAATA
TTGGCCTATTGGCCATTGCATACGGTGTTATCTATATCATATAATATGTACATTTATA
TTGGGCTCATGTCCAAATATGACCGCCATGTTGGCATTGATTATTTGACTAGTTAT
TAATAGTAAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCG
CGTTACATAACTTACGGTGAAATGGCCCGCCTGGCGTTACATGGGTAGTATTTACGGTAAATGCCCACTTGGCATACATCAAGTGTATCATATGCCAAGTC
CGCCCCCATTGACGTCATAAATGGAGGCTATGATCTTTCCCATAGTAACGCAATAGGGATATTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTT
```
TGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAA
 GTCTCCACCCCCATGACGTAATGGAAGTTTGTGGGACCAAAAATCAACGG
 GACTTTCCAAAAATGTCGTAATAACCCCCGGCCCGTTGACGAAATGGGCGGTA
 GGCCTGTAAGGTGGAGGAGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCA
 GATCAGTAAGGCTTTATTTGCGGTAGTTTATCACAAGTTAATTGCTAACGCA
 TCAGTGCTTCTGACACACAGTCTCGAACTTTAAGGCTAGAGTACTTAATACG
 ACTCAGTAAGGCTAGCTCGAGAATTCGCGGGGTTTGTGGAAGGTAGG
 GATTCATCCTCCCCACCCCCGTCTCCACCCCCCCTTGGACCCTGCCGCCCCC
 CCCCCGGATTGGACGCTCGGGAGACATACCTCAGTCGGGCGTCGGTCGTCG
 GCCGGGATTACCTCCGGTTGCGGACCAGTGGCAGAGAGCACAGACAA
 CACATGATAAGATACATTGATGAGTTTGGACAAACCACAAACTGAGATGCAG
 GAAAAAAATGCTTTTATTTTGTGAAATTGTGATGCTATTGTTTATTTGTAAACC
 ATTATAAGCTGCAATACAACAGTGAAGTTTGGGACAAAACACAAAACTGAGAT
 TCAGGTTCCAGGGGAGATGTGGAAGGTTTTTATTTAAGCAAGTAACACCTCTAC
 AAATGTGGTAAATCGATAAGGATCAGGCGGCTGGCTGTAATAGCGAAGAGGCC
 CGCAGCGATCGCCCTTCCCCCAACAGTGGCAGCCTGAAATGGGCAATGGACGC
 GCCCTGTAGCGGCGCATTAAGCGCAGGCGGCGGCTGTTGGGAGTGGGTGGTTTACGCGCAGCGT
 ACCGCTAACAATTGCCAGCCGCTGCTTCCGCTTCTTGTCTCCTTCTTCCTCTCTTC
 TTTCTCGCCACGTTGCCGCGTTTCCCCGTCAAGTCCTAAATCGGGGCTGCCCT
 TTTAGGGTGTTCCGATTTAGTGGCTTTACGCGACCTCGACCACAAAAACTTGATT
 AGGGTGATGGGTTACGTAGTGCACTGCCTTGTTTACGAGTACTGGTTTTTGCGCCT
 TTGACGTTGGAGCTCCAGTTTTATAGTGGACCTTTTGTTCCAAAACGGAAC
AACACTCAACCCTATCTCGGTCTATTTCTTTTGATTATAAGGGATTTTGCCGAT
TTCCGGCCTATTGGTTAAAAAATGAGCTGATTTAAACAAAAATTATACGCGAATT
TTAACAAAATATTAACGCTTACAATTTTCTGTAGTGCGGTATTTTTCTCCTTACG
CTATCTGTGCGGTATTTCACACCGCATATGGTGCACTCTCATACGCAAGCCTGACCTG
TCCGGGAGCTGCTGTCAGAGGTTCACACGTACCCGGAACCCGCGAAGCCTG
GACGAAAGGGCATCTCGATACGCCTGAATTTTAATTACGTACGGAATTTTTATAGGTTAATGTATCATGATAATAT
ATGTTTCTTAGACGTCAGGTGCCACTTTTCGGGGAAATGTGCGCGGAACCC
CTATTTTTATTTTCTAAATACATTCAATATGTATCCGCTCATGAGACAACT
AACCCTGATAAATGCTTCAATAAATATTGAAGAAAAAGAAGAGATGATGAGATTTT
ACATTTGCCGTGCTGCCCTTATTCCCTTTTTTGGCGGCAATTGTGCTTCTTGTTTT
GCTCACCCCCAGAAGCCTGGTTGAAAGTAAAAAGATGCTGAAGATCAGTTGGTG
CACGAGTGGGGTTACATCGAACTCGATCTCAACAGCGTGAAGATCAGTTGAGAG
TTTTTGCCCCGAAGAACGTCTCCATGGCGAGCTGCAGGTTTATAGTACGCTGACAGTG
TGAGCGCCTGATTATCTGGGCTATGGACTGACTCACCAGTCAAGAAAAG
CATACTATTCTCAGAATGACTTGTTGAGTACTCACCAGCAGCACAGAAGGCA
CATCTTACGGATGGCATGACAGTAAGAGAATATGCGATGCTGCTGCATAACCA
TGAGTGATAAACACTGCGCGCCAACCTTACTCTGCAGAACGATCGGAGGACCGAA
GGGAGCTAACCCTCTTTTTGCAACAACATGGGGGATCATGTAACTCGCCTTGATC
GTTGGGAACCCGAGCTGTAAGGCAATACCAACGACGCGCTGACACCA
CGATGCGCTGTAGCAATGGGCAAACACGTTGCGCAAACATTTACAACGTGAGGCGGATAAA
GTTGCAGGACACTTCTGCGCTCGGCCCCTTCCGGCTGGCTGTTTATTGCTGA
TAATCTGGAGCCGGTGGAGCGTGCTGGCTCTCGCGGTATCATTTGCACTGGGG
CCAGATGGTAAAGCCCTCCCCGTATCGTAGTTATCTACACGACGGGGAGGTCAGG
CAAATGTGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGAT
TAAGCATTTGGAAGTCACTGACAGCCAAGACTTATATTACATCTATATATTATGAT
TAAAACTTCATTTTTTTTTTTTTTAAAAGATCTAGGTGAAGATCCTTTTTTAT
CTCATGACCAAAATCTCCTTAAGCTGAGGTTTTTCTGGCCACTGAGCAGACCGG
CTGAGAAAAGATCAAAGGATCTTCTTTTTTGTACCTTTTTTTCTGCGCCTTATCT
GCTGCTTGGAAAACAAAAAACCACCGCTACCAGCGGTGTTTTGGTTTGGCCCGGA
TCAAGAGCTACAACACTCTTTTTTCCGAGGTAACTGGGGTTTTCCGACAGGAGC
ATAACAAATACGTGTTCTTCTGATGTAGGCAGATTTAGGCACACACTTCAAGAA
CTCTGTAGCACCAGCTACATACCTCGCTCTGCTCTAATCTCTGTACCAGTGGCTG
CTGCAGCTGGGATAAGTGCGTGTCTTACCAGGTTGGAACTCAAGACGATAGTT
ACCCGATAAGGCAGCAGCTCGGGCTGAACGGGGGGGTTGCGTCACGACAGCC
CAGCTTGGAGCGACACTACACCAGAAGTGAGATACCTACAGCGTGAGCTA
TGAGAAAAGCCACGCTTTCCGGAAGGAGAAAGCAGCGACAGGTATCCGGTA
AGCCGACAGGGTGCGGAACAGGAGACGCGACAGGGGAGCTTCCAGGGAACAC
GCCGATCTTCTTATAGTCTGCTGGGGTTTCGCCACCTCTGACTTGAGCGTGCTG
ATTTTGTGATGCTGCTCGAGGGGGGCGGACGCTATGGAAAAACGCACGGCAAC
GCAGCCCTTTTACGGGTTCTGCGCCTTTTGGCTGGGCTTTTTCGTCATGCTCGA
CAGATCT
>pNEBR-X1Hygro-UL12.5

GTTCTAGAGGATTTTTGTGAAGGAACCTTACTTCTGTGGGTGACATAATTGG
ACAAACTACCTACAGAGATTTAAAGCTCTAAGGTAAATATAAAATTTTTAAG
TGTATAATGTGTAAAACTACTGATTCTAATTGTTTGTGTATTTTAGATTTCCAC
CTATGGAAACTTATGAATGGGAGCAGTGGTGGA

ACGCCTTTAATGAGGAAAAC
CTGTTTTGCTCAGAAGAAATGCCATCTAGTGATGATGAGGCTACTGCTGACTC
TCAACATTCTACTCTCAAAGAGAGAGAAAGGTAGAGGCTACTGCTGACTC
TCAGAAATGCTAAGTTTTTTGAGTCATGCTGTGATTTTAGATTTTAGATTTGC
TTGCTTTGGCTATTAAACACACAAAGGGAAGGCTGCACACTGCTATACACAGAA
AATTATGGAAAAATATTCTGTAACCT

ATAACATACCTGTGTTTTTCTTACTCCACACACAGGCTAGGCTGCTATTAAT
GGGCCAGATGGTAAGCCCTCCGTATCGTAGTTATCTACACGACGGGAGTC
AGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACT
GATTAAGCATTGGTAACCTGACAGACCAAGTTTACTCATATATACTTTAGATTG
ATTTACCCCGGTGGATAATACAGAAAAACGCCCCAAAACAGGAAGATTGTATAA
GCAAAATATTGATTTTCTTTCTGCTGTATTCCTCTGATTCTGTGGAATAACCGTAT
TACCCGCTTGGAGTGAGTCGTGATACCGCTCGCCGCAGCCGAACGACCGAGCAGC
AGCGAGTCAGTGAGCCGAGGAAGCGGAAGAGCGCCTGATGCGGTATTTTCTCC
TTACGCATCTGTGCGGTATTTTCAACACGCACTCTCTCATACTGCGCCAGAAATC
CGCAGGTGTTTTTTGGGTCGGGGGTGTTTGGCAGCCACACACGCCCGGT
GTTCTGTGTCGCGCAGATACATGCGGTCCATGC
CCAGGCCATCCAAAAACCAT
GGGTCTGTCTGCTCAGTCCAGTCGTGGACCTGACCCCACGCAACGCCCAAAA
TAATAACCCCAACGAACCATAAAACCATTCCTCCCATGGGGGACCCCGTCCCTAA
CCCACGGGGCCAGTGCTATTGGCCCCAATGGGGTCTC
GGTGGGGTATCGACAGAGTGCCA
GCCCTGGGACCGAACCCCGCGTTTATGAACAAACGACCCAACACCCGTGCGT
TTTATTCTGCTTTTTATTTGGCGGTATAGCGCGGCGGTCTCTCGGTATTTGCTCT
CTTCCGTGTTTCAGTTAGCCTCCCCATCTCCCGATCCGGACGAGTGCTGGGG
CGTCGGTTTCCAATCTCGGCGAGTACTTTCTACACAGCCATCGGTCCAGACGGC
CGCGCTTCTGCGGGGCAGATTGTGTTACGCACCGACAGTCCCGGTGCAATCGG
ACGATTCGTCGATCGACCCTGCGCCCAAGCTGATCATCGAAATTTGCCGT
CAACAAAGCCTCTGATAAGGTTGGTCAGAACCAATGCGGAGCATATACGCCCG
GAGCCCGCGGCAATCTGCAAAGCTCCGGATGCCTCCGCTCGAAGGTAGCGCGTC
TGCTGCTCCATACAAGCCAACCACGGGCTCCAGAAGAAGATGTTGGCGACCT
CGTATTGGGAATCCCCGACATCGCCTCGCTCCAGTCAATGACCGCTGTTATG
CGGCCATTGTCCGCACTGACGGTGTCGTCACCTCATCAGCTCATCGAGAGCCG
TTCCCTTGCCGTCCGGAATGGGCGAACCCGCTCGCTCTGCTGCTGCTGCTGCT
AGCGATCGCATTCCATGGGCTCCGACCCGCTCGCTGTCGCTTCTTGCAAGATCGAAT
AGGTCAGGCAGCTCTTCGGAATCGCAGACGCTGACCCAGACCAGACGCTGACGCTG
CGCGGCGGATGCAAATGCCGATAAAACATAACGATCTTTTGAGAAGACCAG
GCGCAGCTTTTTACCCGCAGGACATATCCACGCTCCCTCTACATCGAAATACG
AAGCAGCAGATTCTCTGCCGCTCCGAGAGCTGAGCTGACCGGGAGAGCTGACG
GGGTCCGCTCGGTGGTCGACTCTCAAGCTATCAGCGCCATCGCTTGTGACGCTG
AGGGTGCTCGCTCGGTGGCGACACCGCGTCACCTTAATATCGGAAGTGGA
CCTCGGACCAGCGCGCCCGGCAGCTCAGTCTCGCTGCGTCTCGGAACTCCGAAATGAC
AAGACGGCTGGGCCGGGTTTGTGGTCATCATAGAAACTAAAGACATGCAAATATA
TTTCTTCAAGGGGACACCGGCCACCGGAACGCAACGCGGACCGGGACCGGGGATGA
AGCAGTAAAATTGTTAAACATATTATTTGTGGTAAAAATTCCATGTTAAAAATTTGG
TTAAAACTCAGCTTTTTTTAACCATAAGGCGCAAATCGGGAAAATCCCTTATA
AAATCAAAAAGAAATAGCCCCGAGATAGGGGTGGGATGTTTCTTCAAGTTTTGGGAAACA
GAGTCCACTTTAAAGAAGCGTGGAACCCAGTAAACAGGGCGAAAAACCGTC
TATCAGGGCGATGGCCCACTACGTGAACCATCACCAAAAAATCAAGTTTTTTTTG
GGTCCAGGTTGCGCTAAAGCCTAACTAGCGAACCCTAAAGGGAGCCCCATT
TAGAGCTTGACGCGGAGTGGCAAGTGACTACCGCTACGCTGCGCTGAAAGGAT
CTAGTGAAGATCTTTTTGATAATCTCATGACAAATCCCTAAGCTGAGCT
TTTCGTTCCACTGAGCGTCAGACACCCTTAGAAAAAGATCAAAAGGATCTTCTTGA
GATCCTTTTTTTCTGCGCTATAATCTGTGTGCTGTGCTGGCAAAAACACCAACCGCT
ACCAGCGGTGGTTTGCTGCGTGTTGAGTCCAGAGCTCCACACACTGCTTCTTCTCCCAAGG
TAACCTGGCTCTAGCAGAGCTACCACACTCCTCTGCTGCTTACCCCAAGG
TAGTTAGGCCACCACTTCAAGAATCTCTAGTACACCGCTACATACCTGCTCT
GCTAATCCTGTTACGAGTGGCTGCTGAGTGGAGATAAGTCGTGACTTTACCCG
GGTTAGCCTCAAGACGATAGTTACCGGATAAGGCAGCGGTGCTGGGTGTAAC
GGGGGTTCGTCGCACAGACGAGCGACCTTCTAGGAGCGAAGCCACCTACACCAGACTG
AGTACACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTCCCGAAGGGAGA
AAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACG
AGGGAGCTTCCAGGGGAACCGCTTGTATTTTATAGTCTCTGTCGTTTTCG
CCACCTCTGACTTGGAGCCTGCTGATTTTTGTGATGCTGCTGCTGAGGGGGGCGGAGCC
TATGAAAAAACGCCAGCAACGCAGCGCTTTTTACGGTTCCTGGCCTTTTGGA
CCTTTGCTCACATGTATCAGGATATCACAATTTTGTAGAGGTTTACTTGTCTT
AAAAACCTCCCCCCACCTCCCCCTGAACCTGAAAATGAATGCAATT
GTTGTTACTTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCAT
CACAAATTTTCACAAATAAAGCACATTAAAAACTGCAATATTGTTGATGTTTGTCG
CTCATGGACGGTCACACGAGGTGGGCGGCCGCTCTGGATATTCTCTG
GTCTCGGGGACATTCACGCTACCTGCGGCCCAGTCCCCAAGACCCCTG
CTTTTAGAGGTCAAATGCAGGGCCCAAAGTAGCTTTTCGCAACCCCATGGG
AGCGACCCACAGGCGCTCCGCACACGGAGGACTTTGCTACGGAGACTCTT
CGCGCCGGCGCCGCTCCCGGCCGAGAGGCTCTCG
TCACGCAAGACCAG
GCCTGGTCAGAGGCCACGCCTCGGCGAAAAAGGCGGTGCTCCGCCGCG
GATCGGGGCTTGGTGAATTAAATAGCGGCGTTGTTGCTCAGGAGGCTTTTCTG
TGGCGCCCCGCACCTCGGACGCCACACCATCTCCCGCTGTCTGGAGAGTCC
GGGATCTGGTCACGGCCGAGCCCGTCTTCGCAACCCCTGTCACCGAAGCTT
AAGCGAGATCTTTGGTGAGGGCTACGTGCTCGACAGCCACTTCCCCCGACTGCC
CCCCCACCCCGATCTGGTACGCTTTATCGGCAAGCGCACCACACGCAGCGG
GGAGGGCGTAACCGTTCGCGCTTTGGAGGACCGGCCGGCCGGCTCTCGGGGC
AGGACCAGCAAGGCCCGTCCATTCTCCCGAACAGGCGGTTCCGATCGCCCTG
ATCATTACCCCGGTCCCCGATCGAGCTGAGGTCGCAACGCGGTCTCTGC
CCGGGGCCCGGCCCTGCTGCTGCCGAAACAACGTCCTCATGCACGCGAA
GCAGCCCGCCTGGCATTCCGACACACGCTCGCGGAGCTATGCGGCTCTGTTCT
CCGGGGCCGGCCCTGCTGTGCGAAACAAAGTCCTCCTGCAACCGAAGGACGG
GGAGTGCGTCTCGGCTGACGCAGGCGCTAGCGATCGCCGGCCGGCCACACGCGTCTC
CGGATGTACAGGCATGCGTCGACCCTTAGTCAAGGCCTATAGTGGTAAGTTA
TTACCGGACTGCGCGCTCTTTACACGGCTCGTGACTGCGGGAAACCCCTGGC
TTA
CCCACTTAAATCGCATACCCCTGCAGCGAATCCCTCCCCCCTTGCGCAAGCTTGCGTAATAGC
GAAGAGGCCGCCGACCCGATCGCCCTCCGAACAGGCGCAGCTGAATGGCG
AATGCGCGCTTCGCTTGGTAATAAAGCCCGCTTCCGCGGCTT

463
>pNEBR-X1Hygro-UL12.5-HA

GTTCTAGAGGATCTTTTGTAAGGAACCTTACTTCTGTGGTGACATAATTGG
ACAAACTACCTACAGAGATTTAAGCTCTAAGGTAAATATAAAATTTTTAAG
TGTATAATGTGTAAAAACACTACTAGTTCTAATTGTGTATTTTAGATTCCAC
CTATGGGACTTATGGAATGGGAGCAGGTGGGTGAAATGCCTTTTAATGAGGAAAC
CTGTCTTGACTCAGAAGAAATGCCATCTAGTGATGAGGCTACTGCTGACTC
TCAACATTCTACTCTCAAAAGAGAGAGAGAGTAGAGACCCAGGACTTTTCC
TCAGAATTGCTAAGTCTTTTTGAGTCATGCTGTGTGTATGAATACGACTCTTC
TTGTGTCTATATTCACACACACGGAAGAAGCTGACATGCTATACAGAA
AATTATGGAAAAATATTCTGTAACCTTTTTAATAGTGGGATAACAGTTATAATC
ATAACATGACTGTGTTTTCTTACTCCACACAGGCATAGAGTGCTGTGCTATTAAT
AACTATGCTCAAAAAATTTGTTACCTTTTGAAGTTTTTAATTTTGTAAGGAGGTAA
TAAGGAATATTGTGATATAGTGCTCTTGA<CTAGA<ATCAATCAGCCATAACC
ACATTTGTAGAGGTGTTC<ACTTGCTT<TTTTAAAACCTCC<ACACCTC<CCCCTGAA
CCTGAAAACAATAAAATGAAATGCAATGTGTTGGTATTTAATTTGAATGAGCCTT
ATAATGGTTCAAATAAAAGCAATAGCATCAACAAATTTCAACAATAAAGCATT
TTTTCTACTGCATTCTAGTTGTGGTTGTCCAACACTCATCAATGTATCTTTATCA
TGTCCTGGATCAA<CTACGTAGGTGACACTTTTTCGGGAAATGTGCGGCGGAC
CCCTATTTGTTTATTTTTCTAATAATACATTCAAAATATGTATCCGCTCATGAGACA
ATAAACCTGATAAAATGCTTCAAATAATATGGAAAAGGAGAGATGAGATGATT
CAAACATTCCGTGTCGGCCCTTATTCCTTTTTTTCGCGGCAATTTTGCTTCTCTGTTT
TTGCTCAACCAAGAAACGCTGGTTGAAAGTAAAAGATGCTGAAGATCGTTGG
TGCACGAGTGGGTTATCAGAACTGGGATCTCAACACGGTGAAGATCCTTTGAG
AGTTTTTCGCCCCAAAGAAGCTTCTCAAATGAGCAGCATTCTTTAAAAGTGCTT
ATGTGGCCGCGGTATTATCCGTCGTGACGGCCGGCAAGAGAGCAACTCGTGTCGC
CGCATACACTATTTCTCAGAATGACTTTG<ATGATCTACACAGTCACACAGAAA
AGCATCTTACGGAACAGCAGAAGATGAGAATTATGCGATGCTGACCATAAAC
CATGAGTGATAAACACTGGCGCAACTTACTTCTGACAACGACTGGAGGACCG
AAGGAGCTAACC<CGCTTTTTTGCACAAACATGGGGGATCATGTGA<ACTCGCTTTG
ATCGTTGGGAACCGGAGCTGAATGAAGCCATACAAACACGACGAGCCTGACA
CCACGATGCTGTAGCAATGCGCAACACGTGGCAAAACTATTAACCTGGCGA
ACTACTTACTCTAGCTTCGGCAAAACATTAAGACTGGAGAGGGATGAGAG
AAATGGCCAGGACACACTTCTGCAGTCGCGGCTCCTCCGGCTGGCTGGCTTATTGC
TGATAAAATCTGGAAGCCTGAGCTGGCTGGCTGGCTTCGCGGTATCATTTGCGAGCATG
GGGCCAGAATGGTAAGCCCTCCGCTATCGTAGATTTATCTACACGACGGGAGTC
AGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACT
GATTAAGCATTGGTAACTTGTCAGACCAAGTTTACTCATATATATCTTTAGATTG
ATTACCCCGGTGGATAAATCAGAAAAAGCCCAAAAAAACAGGAAGATTGTATAA
GCAAATATTCAATTTCTTTCTGGCTTTATCCCCCTGATTCTGTGGATAACCGTAT
TACCCGCTTTTGAGCTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGC
AGCGAGTCAGTGACGCGAGGAAGCGGAAGAGCGCCGTGATGCGGTATTTTCTCC
TTACGCATCTGTGCGGTTATTTCACACACCGCAGTCTCTTCATATGCGCCAGAAATC
CGCGCGTGTTGGTTTTGGGGGTCGGGGGTGTTTGGCAGCCACAGACGCCCGGT
TTTCGTGTCCGCGCCAGTACATGCGGTTCCATGGCCAGCCCATCCAAACCAT
GGGTCTGTCTGCTCAGTCCAGTCGTCCTGACCCCACGGAACCCCGCACCCCAAAA
TAATAACCCCAAGCCACCATAAAACCATTCCCATTGGGGGACCCCCGTCCTA
CCACCGGGGCCAGTGGCTATGGCCAGGGGCCTGCGGCCCGACGTTGGCTGGA
GCCCTGGGCCTTCACCAGAACTTGGGGGGGTGGGGTGGGAAAAAGGAAAGAA
CGCGGGCGTTTTGCTCCCAATGGGGTCTCGGTGGGTATGACAGATGCCA
GCCCTGGGACCGAACCCCGCGTTTATGAACAAACGACCCAACCCGTCGCT
TTATTTCTGTCTTTTTATTGGCGGCATAGCGCGGGGTCTTTCTCCTGATATTGTCTC
CTTCCGGTTTTGATAGCCTCTCCCCATCTCCCGATCCGGAGCAAGTGGCTGGG
CGTCGGTTTCCACTATCGCGGAGTACTTTCTACACAGCCCATCGGTCGACGAGC
CGCGCTTCCTGCGGGGCGATTGGTGTACGCGCCGACGTCGCGCTCGGATCG
ACGATTGCGTGCATCGGACCCTTCTGCGCCCAAGCTGCATCATGAAATGCCGT
CAACCAAGCTCTGTGATAGATTGCTCAAGACCAATGCGGAGCATATACGGCCG
GAGCGCGCGGATCTTGCAAGACTCCGCGGATGCCTCCGGCTCGAAGTACGCGTC
TGCTGCTCCATAAGCCAAACCACCGGCTCCAGAAGAAGATGTGGCGACCT
CGTATTGGGAATCCCCGAACATCGCCTCGCTCCAGTCAATGACCGCTGTTATG
CGGCCATTGTCCGTCAGGACATTGTGGAGGCCGAATTCGCGTGCACGAGGT
GCCGGACTTCGGGGCAGTCCTCGGCCCAAAGCATCAGCTCATCGAGAGCC
GCCGGACTTCGGGGCAGTCCTCGGCCCAAAGCATCAGCTCATCGAGAGCC
CGCGACGGACGCACTGACGGTGTCGTCCATCACAGTTTGCCAGTGATACACA
TGTTTCAGGCAGGTCTTGCAACGTGACACCCCTGTGCACGAGCAGATGCAAT
AGGTCAGGCTCTCGCTGAATTCCCCAATGTCAAGCACTTCCGGAATCGGGAG
CGCGGCCGATGCAAAGTGCCGATAAACATAACGATCTTTGTAGAAACCATCG
GCCGAGCTATTATTTACCGCAGGACATATATCCACGCCTCTCTCCTACATCGAAACGCTGA
AAGCACGAGATTCTTCTCGCCCTCCGAGAGCTGATCAGGCTCGGAGACGCCTGTC
GAACCTTTTCGATCAGAAGACTTTCGCAACGACGCTCGGAGTGTTCACGGCTTTT
TCATATCAAGCTGATCTGCGGCACGCTGTTGACGCTGTTAAGCGGGTCGCTGC
AGGGTCGCTCGGTGTTCAGGCCACACCGGTCACCTTAATATGCAGAAGTGA
CCTCGGACCGCGCCGCCGCCACTGCATCTGCGGTCAATTCGCAATGCAC
AAGACGCTGCGGCGGGGTGGTTGTGTCATCATAGAACTAAAGACATGCAAAATATA
TTTCTTCAGGGGACACCGCAGCAAACCGGACGAAACCGGACCACGGGATGA
AGCAGTTAAAAATTGTAACGTTAATATTTTTGTTAAAAATTGCGTTAATTTTTG
TTAAATCGCTCATTTTTTTAACAAATAGGCCAAATCGGCAAAATCCCTTATA
AATCCCCAAGATAAGCCCGAGATAGGGTTGATAGGTGTTCAGGTTTGAACAA
GAGTCACATATTTAAAGAAGCTGGACTCCAACGTAAGGAGCGAAAAACCGTTC
TATCGGCGATTCGCGCCACTACGTGAACCATCACCACAAATCAAGTTTTTTG
GGTCGAGGTGCGTGTTAAAGCAGCTAAATCGGACACCCTTAAAGGGAGGAGGCCCGATT
TAGAGCTTGACGGGGAAAGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGAGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCACACCCCGCGCTTAATTCGCCGCTACAGGGCGCGTAAAAGGATTAGGTGAAGATCCTTTTTGATAATCTCATTGCAGCACAAAAATCCCTTAAACGTGAGT
TTTCGGTTCCACTGAGGCAGCAGACAGCCCCGGTAGAAAGATCAAAGGATCTTCTTGA
GATCCCTTTTTTCTGCAGTGTAATCTCGCTGTCTTGGCAAAAAAABAAAAACCACCGCT
ACCAGCGGTTGGTTTGTGTGCGAGTCAAGAGCTACCAACTCTTTTTCTCAGGAAGG
TAACTGCTTGCCACGACAGCAGATACCAAATACTGTTCTTTCTAGTGTTAGCCG
TAGTTAGGCCACCACTTTAAGAAACCTCTGTGAGCACCACGGCTTCATACCCCTCGCTCT
GCTAATCCTGGTTACAGTGCTGGTCTGGCAGTGGCGATAAGTCTGTGCTTTACCAG
GGTTGGACAGAGCAGATAAGTTACGGGATAAAGGGCGACGGCAGCGGTGCGGCTGAAC
GGGGGGTTCGTGGACACACAGCCCCAGCTTGGGAGCGAAGACCTACCCGAACGTG
AGATACCTACACGCGTGAGCTATGAGAAAGGCACCGCTTCCCGGAAGGGAGAGGCGGAGG
TTCCAGGGGGAAACGCCTGTATCTTTTATAGTCCTGTCGTTTCCCG
CCACCTCTCGAGCTGCTGATTTTTTGAGTCTTGCTGCTGAGGGGGGCGAGCC
TATGGAAAAACGCAGCGCAACGCGGCTTTTTACTCATTCTGCTGCTTCTGG
CCTTTTGCTCACATGTATCAGCCATATCACATTTTTGAGGTTTTACTTGCTT
AAAAAACCTCCACACCTCCCCCGAAACTGGAACAAAAAATGAATGCAATT
GGTTGTTAACCTGTGTTTGAGGTACGCTTTAATAATGTTACAAATAAAGCAATAGCAT
CACAAATTTTCACAAATAAGCATTTTTCACACTGCATTCTGTTGTTGCTCC
AAAACCTCATGATTTTATCATGATCTGGGATCAGCCCATATCACATTTTGAGA
GGTTTTACTTGCTTTAAAAACCTCCACACCTCCCCCTGAACCTGAAACATA
AAATGAATGCAATTGTTGTTAACTTGTTTATTGCAGCTTATAATGGTTACAAA
TAAAGCAATAGCATCACAAAATTTCAACAAAAAUGCATTTTTTTCACTGCATTC
TAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCGCTGCACTTG
GTGAGTTAGCTCACCATTAGAGAGCCCCAGGCTTTACACTTATGCTTCCGGC
TCGTATGTTGTTGTGGAATTGTGAGCGGATAAACAAATTTCAACACAGGAAACAGC
TATGACCATGATTACGCAAAGCTACGCGGAGGACTTGCTCTCCGAGCGGAG
TACTGTCCTCCGAGCGGAGGTACTGTCCTCCGAGCGGAGTACTGTCCTCCGAGC
GGAGTTCTGTCCTCCGAGCGGAGACTCTAGAGGTATATAATGGGGGCCGTC
AGCTACTACCAGAGCTCATGGGGCCGTGCAATTGAAAGCCGCTGGCGCCAAAG
CTTCTCTGCAAGGATatcgccgaccacatgtgateggctggctgatCTCCCCAAACGCGCTGCCCCTC
CCATATACTAGCCGAGACGTTGCAGCGCACCCTCTCGCAGGGTTGTGCTGCGGC
GTCCGCGCCCTCTCGGACCATCGGCTTCCCCTCTGGGCCCGCTGGATTATCTGTG
TTCCCTGGCCGTTGCTCTCGAGGAGGCGGATATGTGTTGGACCCGCGGACTCGGT
CGGACGCTATGGCCGCTGACGCGCCGCGGGCCCCCGGCCGCCGCCGCCGCCACGCG
TGCGGCCCCGCGCCCCTCATGGGGGTTTTACGAGGCGGACGCAAAACCGAGGC
CGACTGCGAGGCTGAGGCCCCCCCTGGCTCCGCGGCGGGCTCAGGACCGCATCCACC
CTCCGCTGGGGGCCCCAGGGCTGCTGTTTCTCCGCCCCAAGTGGCTGAAGACACA
ACGCGAGCGGTGGCCGGAATGTACAGTCTGCGCAGGTGATGTTCGCGGCGGGGT
GAACGAGGCAGGCAGGAGGCTGGCTGCTGTGGGCTGCTGATGGCCGCGCC
GACGACGCGCGCGAGGCGCGGCGCGCCGACACGCGCGCGCTTTTATCTTCACGAAA
CCAGCGACCTCGCCGAAGAGATCGCATACGTGTGGGGTGCTCTCATGGACCG
TCACACGGGATGTCGCGCGGCGTCCTCTGGATATTCTCGTCTGCTCTGGGACA
TTCACGGGCTACCTGGCCCCAGTCCCCAAGACCCCCCTGGCCTTTTACGAGGTC
AAATGCCGGGCAGTACGCTTCTCGACCCCCATGGAGCCACGCGACCCACCGG
CCTCCGCGTACGAGGACTTGATGGCACACCGGTCCCCGGAGGCGTTCCGGGC
ATTTATCCGGTGATCAGGGGCGATACGCTTTCGACCCCATGGACCCCAGCGAC
CCTCCGCGTACGAGGACTTGATGGCACACCGGTCCCCGGAGGCGTTCCGGGC
GGAGTTAAATAGCGGCGTTGTCTCGAGGTTGCTTCTGTITTGGCGCCCCCGAC
TCGGACGCCACACCATCTCCCGCGTTGCTCTGGAGGCTCCTCCGGGGATCGG
CGACCGACCACCAGGCTTCTCCCGAACCACCCGTCACCGGGGAACCTTAAGCA
AGATCTCTTGGA
GCAGGGCTACGTGCTCGACAGCCCACTTCCCCCGACTGCCCCCCCCACCGCAT
GTCCCCGGCCGAGGGAGGAGGCTCTCGGACGGGCTCTCGGGGCCGCAGGACCGA
CAAGGCTCCATTCTCCCGAACCAGGCCGTTCCGATCGCCCTGATCATTACCCCGT
CGCATCGATCCCCGAGATCTATAAGGCCATACCAGCAAGCAGCCGCCTGGCAT
TCGACGACACGCTCGCCGAGCTATGGGCCTCTCGTTCTCCGGGGCCCGGCTGGC
CTGCTGCCGAAACAACGTCTCCTCACCAGACGacggggaggtcgtctcgctacccctatgcgtccattcctcccgatcggtcgtctcattacccccctatgacgctcctgattacgcttgagcGGGCGCTAGCGATCGC
GGCGCGCCACGCGTCTCCGGATGTA
CAGGCATGCGTCGACCCCTCTAGTGCAAGGCTATAGGTGAGGCTGATATCCGGAC
TGCCGTCGTTTTTAAACCGTCGTAAGCTGGAAACCCCTTGCGGTACCCCAACTT
AATCGCCTTGGCAGCACATCCCCCTTTCCCGCCAGCTGCTGGTAATAGCGAGAG
CACCCGACCAGCTGCCCTTCCCAACAGTTGCAGCCTGAAATGCGAATGGCG
CTTCGCTTTGGTAAATAAGCCCGCTTCGGCGGGGCTTTTTTTTTT
>pNEBR-X1Hygro-MLS

GTTCTAGAGGATCTTTTGTGAAGGAACCTTACTTCTGTGTGTGTGACATAATTGG
ACAAACTACCTACAGAGATTTAAAGCTCTAAGGTAATATAAAAAATTTTTAAG
TGTATAATGTTAAACTACTGATTCTAATTGTGTGTATTTTTAGATTCCAAC
CTATGGAACCTTATGAATGGGAGCAGTGGTGGAATG

CCTTTAATGAGGAAAAC
CTGTTTTGCTCAGAAGAAATGCCATCTAGTGATGAGGCTACTGCTGACTC
TCAACATTCTACCTCAAAAAAGAGAGAAAGGTAGAGACCCAAGGACTTTCCT
GCAGAATTGCTAAAGCTTCTTTTGAGTCATGCTGTGTGTATTTTAGATAGACTCTGA
TCATGTTTTGCTATTTACAACCACAAAAAGGAAAAAAGTGCACCTGCTATACAAAGAA
AATTATGGAAAAATATTCTCTTCTTTATATAAGTACGATACACAGTTAATAC
ATAACATACTGTTTTTCTTACTCCACACAGGCATAGATGTCTGTATTTAAT
TGCTGCTCCATACAAGCCAACCACGGCCTCCAGAAGAAGATGTTGGCGACCT
CGTATTTGGGAATCCCGAACATCGCCTCGCTCCAGTCAATGACCGCTGTTATG
CGGCCATTGTCCGTCAGGACATTGTTGGAGAAGGCAACATCCGCGTGACACGAGGT
GCCGGACCTCCGGGACTGCTCGCTCCAGCACAATCTCACTGAGAGCTGAC
CGCGACGGACGCCACTCGACGAGACGCTCGCTCCAGTCAATGACCGCTGTTATG
CGGCCATTGTCCGTCAGGACATTGTTGGAGCCGAAATCCGCGTGCACGAGGT
GCCGGACTTCGGGGCAGTCCTCGGCCCAAAGCATCAGCTCATCGAGAGCCTG
CGCGACGGACGCACTGACGGTGTCGTCCATCACAGTTTGCCAGTGATACACA
TGGGGATCAGCAATCGGCATATGAAATCACGCCATGTAGTGTATTGACCGA
TTCTTTCGCTCGCTCCAGTAATGGGCGAATCCCGCTCGCTTGCTGCTGCTCAAG
AGCAGTCGATCCATCCATGGCCTCCGCGACCGGCTGCAGAACAGCGGGCAGTTCG
GTTTCAGGGCAGGTCTTTGCAACGTGACACCCTGTGCACCGGCAGGGGAGATGCAAT
AGGTCAAGGCTCTCTCGCTGAAATTCCCAATGTCAAGCACTTCCGGAATCGGGAG
CGCGGCCGATGCAAAGTGCCGATAAACATAACCAGATCTTTTGTAAGAAAACCATCG
GCCGCAGCTATTTTACCCGCAGGACATATCCACGCCCCTCCTACATCGAAGCTGA
AAGCACGAGATTCTTCTCGCCCTCCGAGAGCTCGCATCAGGTCAGCTGTC
GAACTTTTCGATCAGAAACTTTCGACAGACGTCGCGGTAGTTCAAGGCTTTTT
TCATATCAAGCTGATCTGCGGCACGCTGTTGACGCTGTTAAGCGGGTCGCTGC
AGGGTCGCTCGGTTGCCAGCCACACCGCTACCTTAATATGCAGAAAGTGA
CCTCGGACCCGCGCCGCCCCGACTGCATCTCGCTGTTGACACAGCTATTCACTGCAATGAC
AAGACGCTGGCGGCGGTTTGTGTCAATCATGAAACTAAAGACATGCAAATATA
TTTCTTCAGGGGACACGCGCAGCAACACGCGAGCAACGGGCGACGGGGATGA
AGCAGTAAAATTTGTTAAGCTATATTTTTGTTAAAAATTTCGCTTAATTTTTTG
TTAAATCAGCCTATTTTTAACCATAAGGCGAAAAATCGGAAAAATCCCTTATA
AATCAAAAAAATAGCCCGAGATAGGGTTTGAGTTGTTGTTCCAGTGTGTGGAAACAA
GAGTCCACTATTAAGAGGAGCTGGGACTCCGCAAACGATAAAGGGCGAAAACCGTC
TATCAGGGCGATGGCCCACTACGTGAACCATCACCCAAATCAAGTTTTTTGG
GGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCCCGATT
TAGAGCTTGACGGGGAAAGCGAACGTGGCGAGAAAGGAAGGGAAGAAA
GC
GAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGT
AACCACCACCCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTAAAAGGAT
CTAGGTGAAGATCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTAAGT
TTTCGTTCCACTGAGCTACGACACCCCGTAGAAAGATCAAAGGATCTTCTTGA
GATCCTTTTTTTCTGCCGCTAACTCTGCTGCTGTTCGCAAAACAAAAACCACC
CT
ACCAGCGGTGGTTGTTGCGGATCAAGAGCTACCAACTCTTTTTCCGAAGG
TAACGTGCTTCAGCAGCGCAGATACCAAATCTGGCTCCTACATACCTCGCTCT
GCTAATCCTGTACCAGTGCTGCTGCGCTAGTGCGATAAGTCTGCTGCTTACCG
GGTTGGACTCAAGACGATAGTTACCAGGATAAGCCGACGCGTCGGGGCTGAAC
GGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACCTACACCGAACTG
AGATACCTACAGCGTGAGCTATGAGAAGCGGACCCAGCCTTCCCCGAAGGA
AAACCGCAGATTCCGGTAAAGCGGCAAGGTGTCGAACAGAGAGAGCGCAG
AGGGAGCTTCCAGGGGAAAAGCCCTGTTATCTTTTATAGTCTGTGGTTTCG
CCACCTCTGACTTGAGCGCTCGATTTTTGATGCTGTCGGGGGCGGAGCC
TATGGAAAACGCCAGCAACGCGCCCTTTTTACGTTCTCCTGGCCTTTTGTGG
CCTTTGCTCATGTATCGCCATATCACACTTTGTAGAGTTTTTACTTGCTTT
AAAAAACCTCCACACCTCCCTCCCTGAACCTAAACATAAAAATGAAATGCAATT
GTTGTAACCTTTTTATTGAGCTTTAATAGTTACAAATAAAGCAATAGCAT
CACAAATTTCCACAAATAAAGCATTTTTTCACTGCATTCTAGTTGTGGTTGTCC
AAACTCATCAATGTATCTTATCATGTCTGGATCAGCCATATCACATTTTGTAGA
GGTTTTTACTTGCTTTTAAAAAACCCTCCACACCTCCCCCTGAAACCTGAAACATA
AAATGAAATGCAATTTGTGTGAATCTTATCTTATTTATATGTGTCTATGCAATTAC
TAATGCATGTTTTGTCAAAACATCATCAATGTATCTTTATATGTCTGCAATGTAAT
GTGAGTTAGCCTCACTCATTAAGGACCACCCAGCTTTTACACTTTATGCTTCCGAG
TCGTATGGTTGTGTGGAATTTGTGAGCGGATACAATTTTCACACAGGAACACAGC
TATGACCATGATATACGCAAGCTACGGGGACGATCTGTCTCGAGCGGAG
TACTGTCTCCGAGCGGAGTACTGTCTTTCCGACGACAGGAGTACTGTCTCCGAGC
GGAGTTCTGTCTCCGAGCGGAGACTCTAGGATATATGATATATGAGGCGTCGTC
AGCTACTACCAGAGCTCATGGGGCCGCCCAAGTATGTCCGCTGACTGAGCGGCTG
CTGCTGCGGGGCTGTGGCAGGCCTGCTTGGCGGGCTTTTTTTTT
TTAAGGCGCGCCCTCGAGCGGCGCGCTAGCGATCGCGCGCGCGCAGCCCGCTGTC
GGATGCTACAGGCATGCCTGACCCCTCTAGTAGGATTAGATGAGTCTGCTG
TACGGACTGGCCGCTGTTTACAAAGTCGACTGGGAAAAACCTGGCGTTA
CCCCACTTAAATCGCCTTGCAGCAGACTCCTCTCTGCCAAGCTGCGTAATAGC
GAAGAGGCCCCGACCCGATCGCCTTTTTTTCAACAGTGGCACGCTGACTTGAATGCG
AATGCGCGTTCGCTTTGGTAATAAAAGCCTCCTGCGGCGGGCTTTTTTTT
>pNEBR-X1Hygro-ICP8

GTTCTAGAGGATCTTTGTGAAGGAACCTTACTTCTGTGGGTGTGACATAATTGG
ACAAAACTACCTACAGAGATTTAAAGCTCTAAAGGTAATATAAATATTTTTAAAG
TGTATAATGTGTTAAAACCTACTGATTCTAAATTGTTTTGTGTATTTTTAGATTTCCAC
CTATGGAACCTTATGAATGGGAGCAGTGGTGGGAATGCCTTTAATGAGGAAAAC
CTGTTTTGCTCAGAAGAAATGCCATCTAGTGATGAGGCTACTGCTGACTC
TCAACATTCTACTCTCAAAGAAGAGGAGAGTAGAGACCCCAGGACATTTCCT
TCAGAATTGCTAAGTTTTGCTGACTTATAGAACTCTTGC
TTGCTTTGCTATTACAACACAAAGGAAGAAGCTGCAGCTGCTACAAAGAA
AATTATGGAACATATTCTGTAACCTTTATAAGTACGGCATACAGTTATAATC
ATAACATACTGTCTCTTACTCCACACACAGCATAGTGCTCTATTAAAT
AACTATGCTCAAAAATTTGTGTACCTTTTAGCTTTTTAATTTGTAAAGGGGTTAA
TAAGGAATATTTTGTAGTGATAGTGCGCTTTGACTAGAGATCATAATCAGCCATACC
ACATTTGTAGAGGGTTTTTACTTGGTTTTAAAAAACCTCCCAACACCTCCCCCCTGAA
CCTGAAAACATAAAATGAATGCAATTTGTTGTTTGAACCTTTTTATTTTATTGCAGCTT
ATAATGGTTACAAATAAGCAATAGCATCACAATTTTCAATAAAATAAACCAAGCATT
TTTTCACTGCATTCTAGTGTGTTTGTCCCAACTCATCAATGTATATCTTATCA
TGTCGGATCAAACCTACTACGTCAGTGGGCACTTTTTCCGGGAAATGTGCGCGGAAC
CCCTATTGTTTAAAATCTAAATAACATTCAAATATGTATCCGCTCATGAGACA
ATAACCCCTGATAAAATGCTTCAATAATATTGAAAAAAGGAAGAGTATGAGTATT
CAACATTTCCGTTGCACCCCTTTATATTCCCTTTTTTTCGGACATTGGCTCTCTTGT
TTGTCTACCCAGAAACGCTGGTGAAGTAAAAGATGCTGAAAGATCAGTGGG
TGACAGGAGTGATTACATCAGAATCGAGATCTCAACGAGTAAAGATCCTTGAAG
AGTTTTCGCCCGAAGAACGCTTCTCCTCAATTGATGAGCAGCACTTTTTAAGGTTCTGCT
ATGTGGCGCGTATTATCCCGTGTGACGCAGCCAAGGCGAAGCAGAATTATGCAGTGGCTGCCATAAC
CATGAGTGATAAACACTTGGCGCAACTTACTTTCTGACAACGCATACGAGAAGCCG
AAGGAGCTAAACCGCTITTTCGCAACACATGGGGGATCATGTAACCTCGCTTTG
ATCGTTGGAACCGAGCTGAATGAAGCCAACATACCCAAACGAGCAGCGTGACA
CCACGATGCCTGTAGCAATGCAACGTTGCGCAAACTATTACAAACTGCGCAG
ACTACTTACTCTAGCCTCCGGCAAAACTATTACGTAGGATGGAGGCGGAT
AAAGTTGCGAGACCACCTTCTCGCTCGCCCTTTCCGGCTTGGTTTATTTGC
TGATAAAATCTGGAGCCGGTGTGGACGTGGTGCTCGCGGTATCTGCTGACGACTG
GGGCCAGATGGTAAGCCCTCCGTATCGTAGTTATCTACACGACGGGGAGTC
AGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTCCTCACT
GATTAAGCATTGGTAACCTGTACAGACCCAAGTTTACTCATATATATTCTTAGATTG
ATTTACCCCGGTTTGATAAATACGAAAAAGCCCACAAAACAGGAAGATTGTATAA
GCAAATATTACGTCTTTTCTGTGTTATCCCTGATTTCTGTGGATAAAACGTAT
TACCAGCCTTTTGGATGAGCTGATACCAGCTCGCCGAGCAGCAGGAGGCCG
AGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCTGATGCGGTATTTTCTCC
TTACGCATCTGTGGGTATTTTCTACACCCGACTCTCTCCTATATGCGCCAGAAATC
CGCAGGTTGGTTTTTGGGGGTGAGGTGTGGGTGTGGGAGCCAGACAGCAGGCCC
GTTCTGCTGCGCAGCAGTATACTGCGGCTCATGCACGACATCGGCTTC
GGGTCTGTCTGCTCAGTCCAGTCGTGGACCTGACCCCACGCAACGCCCAAAA
TAATAAACCCCCACGAACCATAAAACCATTCCCCCATGGGGGACCCCGTCCCTAA
CCCACGGGCGCCGAGTGCGGCTATGGCAGGGCCTGCCCGCCGAGTGGGGCTGCA
GCCCTGGGACCGAACCGCGTTTATGACACCCGACCGGACGGGGAGGGAAAGAAGAA
CGCGGGCCATATTGGCCCCAATGGGGTCGCGGGGTATCGACAGAGTGTC
GCCCTGCGCGCCGCTGCTGCTGAGTACTTCTACACGCCATCGGCTCCGGAG
CGTCGGTTTCCACTATCGGCGAGTACTTCTACACGCCATCGGCTCCGGAGCGGC
CGCGCCTCTGCAGGGGCAGTTTGTGTACCGCCGACCTCGCCGAGCTCGG
ACGATTGTCGATCGACCCCTGCGCCGAGCTGCTACGAGCTCGAATCTG
CAAACCAAGCTCTGATAGAGTTGGTCAGACGAATGCGGAGCATATACGCCG
GAGCCGGGCGGAGTCCTGCAAGCTCGGATGCTCTTCGCTGAAATGAGCGCTC
TGCTGCTCCATACAAGCCAACCACGGCCTCCAGAAGAAGATGTTGGCGACCT
CGTATTGGGAATCCCGGACATCGCCTCGCCTCCAGGCTCAGTATGGAAGACGGCT
CGGCCATTGTCCGTCAGGACATTGTTGGAGCCGAAATCAGCGTGCAGACCCG
CGCGGACGGACGCAGCAGTGAGGTGTCGCTCCATACAGGTTTGCGCAAGTGATACACC
GAGCGTACGACTCCATGCGTCGCCCAATGCGGACACGGCTCGACGAACAGCGTCG
AGGCAGCTCATTTACCGCGGCTCCGACGGACGCACTGACGGTGTCGTCCATCACAGTTTGCCAGTGATACACATA
GTTTCAGGCAGGTCTTGCAACGTGACACCCGCGTCGCTGCTGCTGCTGCTGCTGCTG
GCCGGACTTCGGGGCAGTCCTCGGCCCAAAGCATCAGCTCATCGAGAGCCG
AGCGATCGCATCCTACTGCCCAATCTCCCGGACACGGCTCGAAGACAGCACGTC
GCGCAGCTATTTACCGCGGCTCCGACGGACACGGCTCGAAGACGCTATC
AGGTCAGGCTCTCCTCGTGAATTTCCCAATGTCAGCACTCTCCCGGAATCGGGACG
TCGAGCTGGGCGGGGTTTGTGTCATCATAGAACTAAAGACATGCAAATATA
TTTCTTCAGGGGACACGCCAGCAAAACGCAGCAACCGGCGACCCGGGAAATG
AGGCAGCTGGGCGGGGTTTGTGTCATACATAGAACTAAAGACACATGCAAATATA
TTTCTTCAGGGGACACGCCAGCAAAACGCAGCAACCGGCGACCCGGGAAATG
AGGCAGCTGGGCGGGGTTTGTGTCATACATAGAACTAAAGACACATGCAAATATA
TTTCTTCAGGGGACACGCCAGCAAAACGCAGCAACCGGCGACCCGGGAAATG
TATCAGGGCGATGGCCCACTACGTGAACCATCACCCAAATCAAGTTTTTTGG
GGTGCAGGTTGCCGTAAAGCACTAAATCGGAAACCCCTAAAGGGAGGCCCTTCT
TAGAGCTTGACGAGGGTGGCTGCAAGGCTGCGCGT
AACCACCACCCCCCGCGCTTTAATGCGGCCGCTACAGGCGCGTAAAGAGAT
CTAGGTGAGAGTCTTTTTGGATAATCTCATGACCAAAAATTCCTTTAAGGTAGT
TTTCGTTTCACGTAGCGTCAGACACCCCGTAAAGGATCAAAAAGGCTTCTTTTCA
GATCTTTTTTCTGCCTCTGCTGCTGCTGCTTTTGCAAAACAAAAACACCGCT
ACCAGCCGGTTGTTTTGGCCGATCAAGAGCTACACACTTTTTTTCCGAAAG
TAACCTGCTTCAGCAGAGCGCAGATACCCACATATACCTCTCTTGTAACCG
TAGTTAGGCACCACCTTCAAGAAACTCTGTAGCAACCGCCTACATACCTCGCTCT
GCTAATCCTGTACAGTGGGCTGCTGCCAGTGCCGATAAGTGGTGTCTTTACCG
GGTTGGACTCAAGACGATAGTTACCGGATAAAGCGCAGCGTCGGGCTGAAC
GGGGTTCGTGCACACAGCCCAGCTGGAGCGAACGACCTACACCGAACTG
AGATACCTACAGCGTACGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGA
AAGGCGGACAGGTATCCGGTAAAGCGGCAGGCTCGGAACAGGAGAGCGCAGC
AGGGAGCTTCCAGGGGAAACGCCTGGGATCTTTTATAGTCCTGCTCGGGGTTTCG
CCACCTCTGACTTGAGCGCTCGATTTTTGGTGATGCTCGTGAGGGGGCGGAGCC
TATGGAAAAACGCCAGCAACGCCTTTCTACGGTTCTGGCCTTTTGCTG
CCTTTTGCTACATGTATCAGCCATATCACAATTGTAGAGGTGTACTCTTTT
AAAAAACCTCCCACACTCCCCCTGAAACCTGAAACATAAAAATGAATGCAATT
GTTGTTAATTTTTTTATTGCAGCTTATATGGTTACAAATAAAAAGCAATACGAT
CACAATTTTCACAAAATAAAAAGCATTTTTTCACGTCTGTTGTGGTTGTCC
GTACACACGACGCTCATGCGCTCCGGGCAGCCATCCCAAGTTGCGCCAGCG
CGCCCGCGGAGGATCGCTTCCGGACCACAGCATGTTACAGCGA
CCGCCCGCGGGACCCATCGGCGTCTTCGGGACCATGAACAGCATGTACAGCGA
CTGCGACGTGCTGGGAAACTCGAGACCTCAGTACGTGGGACCAGGCGGCTCC
CCACGGCCATGGGGGGCGCTGAGACCACCATGACCGGATGAGGGTCCGTA
TACGAGTGGTAAACACGTCAGGCGAGGTGTGACGTGGACCAGGCGGAGGT
GATGCGCAACCTGTGGAGGGAGGAACCTTCAAGTTTCGCGACGATCTGGGC
GAGGCAACCCACGCCATGCTCCGTAGCGGACCCGTACGCGTGCAGGGCCAT
GCCCGCCGCTGCACCTCGGGCGGCATCAAACCTCAGGCTGTATCACAGGA
CTGGCCCTAGCCAGGCGGAGGACCTGAGGTGGCAGGGGACGGACCGCGGGCC
GCAACTTTCGCAATCTAATTCAACCGGTGCTGCGCCGCGCGGTGATGGACAT
GTTTAACCAAGGGTTTCTGTCCG GCCAAAAACGCTGACGTCGCTGCTCTCGGAG
GGGGCGGCTATCTGCCCTCCAGCCTACGCGCCCGCCAGACGGCCCGCCTC
AGAGCAGCTTTCGAAGGGCAGCTTGCGCCGCGGAGCAGCGCAACCGGCTCC
GCTGCGCGTCAAGAGCCCGCGTTGTTGTCGGCCGCGGAGCGCAAACCGCTCC
GAGGCCCGGCAAGGCGGGGTCGCGCGGCAGCCTCCAGACGCCTACCAGAGGG
GACAAGCGCGTGACATCCTCTCGGACCGCTGCTGCTCTGGAGCTGAGCAGT
TCCACGCGGCCATCTCTCCAAAACGGCAAGCCCGGCGGTCAAACCAGGCCAA
CCCGCAGTGGTTCTGGAGCGGCCCTCCAACGCGCTTCCCGCCCGGCTC
CTGTCGCGCGAGGACACATCGGAGCAGACATCGCGTTCATTTCAAAAGTTTTCCTGG
ACTACGGCGCGATAAAACTTTATTAACCGTGCCCCAACACGTCGAGCGGCT
GGCGATGTAATCAGTGGCAAAACGATTCTGCGGTACTCGCGATCACTCGACA
CAACGTCTAGGACTGGGAAAAACCTGGCGTTACCCAACTTAATCGCCTTGCAG
CACATCCCCCCTTTGCCCAGCTGGGCTAATAGCGAAGAGGCCGCCACCCGATCG
CCCTTCCCAACAGTTGCAGCAGCTGTAATGGCGAATGGCGCTTCGCTTGGTAAT
AAAGCCCGCCTCGCCCGGCTTTTTTTTT

>pNEBR-X1Hygro-ICP8-GFP

GTTCTAGAGGATCTTTGTGAAGGAACCTTACTTCTGTGATGATGAGCAGTGGTGGAATGCCTTTAATGAGGAAAAC
CTGTTTTGCTCAGAAGAAATGCCATCTAGTGATGATGAGGCCTACTGCTGACTC
TATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGG
CCTTTTGCTCACATGTATCAGCCATATATCACATTGTAGAGGTTTTACTTTGCTTT
AAAAACCTCCCCACACCTCCCCCCCTGAACCTGAAACATAAAAAATGAATGCAATT
GTTGTTAATTGTTATTTGCACTTTATAATGGTTACAAAATAAGCAATAGCAT
CACAATTTTCACAAAATAAGCATTTTTTCACACTGCATTTCTGACATTTTGAGA
AGTTTTACTTGCTTTAAAAACCTCCCCACACCTCCCCCCCTGAACCTGAAACATA
AAATGAATGCAATTGTTGTTAACTTTTGTTTATTGCACTTTATAATGGTTACAAA
TAAAGCAATAGCATCACAATTTTCACAAAATAAGCATTTTTTCACACTGCATTT
TAGTTGTGGTTTGTCCAAAACTCATCAATGTATCCTATCATGTGCTGACATGTAAT
GTGAGTTAGCTCACTCATTAGGCCACCCCCAGGCTTTACACTTTATGCTTCCGGC
TCGTATGTTGTGGAATTGTGAGCGGATAAAACATCTACACAGGAAACAGC
TATGACCATGATTACGCAAGCTACGCGGAGGTACGTGCCTCAGCGGAGAG
TACTGTCCCTCGAGCGGAGGTACTGTCCCTCGAGCGGAGGTACTGTCCCTCGAGG
GGAGTTCTGTCCCTCGAGCGGAGACTCTAGAGGGTATATAATGGGGCCGTC
AGCTACTACCAGAGCTCATGGGCTCTACATTGGAAGCGGCCTGGCCGAAG
CTTTATTCGCGGATGTTTATCACAGTTAAATTGCTAAAGCAGTCACTGGCTTTCTG
ACACAACAGTCTCGAACTTAAAGGCTAGAGTACTTAATACGACTCAGTATAGG
CTAGCCTCGAGAAATTCCGGGGGTGTGGAAGGTTGAGAATGGAATTCCATCTCCC
AACCCTCGGGGCTCCACCCTCTCCTGACGTGACGTGCGTGCCTCTCGCTTATTAC
TGAGGCTCGGGGAGACATACCTCGTGGGGCGTCCGTCCGTGCTGTCGCGGAGATTAC
CTCCGTTTGCGGACCGATTGCGAGAGCATGGAGACAAAGCAAGCCCAAGACG
GCAACCCACCATCAAGGTCCCCGGCTGGGGCTGGTCACGTCATGCGGC
TCCACGCGGACCATCTTTCCCAACGGCAAGCCCCCGGGGTCCAACCAGCCGAA
CCGCAGTGGTTCTGGACCGGCCCTCCAACGCAACCAGCTTCCGCCGCCCGCTC
CTGTCGCGCGAGGACATCGAGACCATCGCGTTCATTAAAAAGTTTTCCCTGG
ACTACGGGCGGATAAATTTTATTAACCTGGCCCCCAACAAACGTGAGCGAGCT
GGCGATGTACTACATGGCAAAACCAGATTCTGCCTGACTGCACTGACTCGACA
TACTTCATCAACACCCCTCACGGCCATCATCGCGGGGTCCCCGGCCGTCCCCCACG
CGTGCAGGGCCGCGGCGGCGGCGCTGTTGGACGCCGTGGACGCATCCGGCGCGTGGAC
GGGGGCCCGCGCGCTGATGGACGCCGTGGACGCGCATCCGGCGCGTGGAC
GTCCATGTTCGCAGCTCAACCTGCTGCTCGGGCCCGTCATGGCCCGCGCCGCC
ATGGCTGTGTGGGGTTGAGCATCAGCAAATAATACGGCATGGCCGGCAACGT
ACCGTGTGTTTACAGGCCGGAGCTGGGCGCAACCTGATGGGCGGPGGCAAAAACGC
GGGCCGCTGCCCTTTATTTGGCAACCGCAGGGCAAGCTTGCTCGCTGCTCCC
GGGGGCCGCTCTTTATTTTGGCAACCGCACGGCGCGAAGCGCGAACAGAAG
CTCGGTGTGGCACGAGCAGCTCCGGGCACTGTCTCCAGAGGCGGCGGCAGCGCT
GCCAGTACGGTTGTGGCCGAGCGGTGAGGCGGCGCGCCCTAGTGAC
CCACTCCGGCAACCGCGGAGGTGTTTAACTTTGGGGATTTTGGCTGCGAG
GACGAGACCGGGTTCCCTGAGGTGGCTCGCAGCGAGCGCAGGCGCTAGTG
GCCAGATCGGTGCTTGCCCACCGGTGACCTTGCTGGAGGTCGAC
GACGACCCCGGTCAGCGGCTGCGGGTCAGCGGAGCGGCGCGGCGGCGCTGTGAGCTG
CCGACGACCGGGGCCCCTCTGAGGTGGCTGCGGACGCGGAGCGGCAGGCGCTAGTG
GCCAGATCGGTGCTTGCCCACCGGTGACCTTGCTGGAGG
GGAACTGTTCACTGCGTGGTCCCAATTCTCGTGGAACTGGATGGCGATGTG
AATGGGCACAAATTCTGTGCAGCGGAGAGGGTGAAGGTGATGCCACATACG
GAAAGCTCACCCTGAATTTCTATCTGCAACACTGGAAAGGCTCCCTGTGCGCATG
GCCAACACTGGTCCTACCTTCACCTATGGCGTGCTAGTGGTTTTCCAGATACC
CAGACCATATGAAGCAGCATGACTTTTTTCAAGAGCGCCATGCCCCGAGGGCTTA
TGTGCAAGGAGAGAAACCACTTTTTTCAAAAGATGACCGGAAACTACAAAGACCCGC
GCTGAAGTCAAGTTCGAAGGTGACACCTGGTGGAATGGCGATGTTGAAGG
GCATTGACTTTAAGGAAGATGGAACATTTCTCGGCCACAAGGCTGGAATACAA
CTATAACTCCCCCACAATGTGTAATCATGAGCCGACAAGCACAAGAGCATTGACATC
AAGGTCAACTTTCAAGATCAGACACACAACATTGAGGATGGAGATCCGTGCACTGG
CCGACCATTATCAACAGAACACTCCCAATCGGCGACGCGCCCTGTGCTCCTCCC
AGACAACCATTTACCTGTCACCCAGTCTGGGCTTCCTAGGATGCTAAAGATCTCCAAGCA
AAGAGAGACCACATGGCCTCTGGTGGAGTTTTGTGACCGCTGCTGGGATCACAC
ATGGCATGGACGAGCTGTCAAAGTGAGCGCCGGCCTCCAGCCCATGGCATGAAA
GATAAATTGATGATTTGGAACAAAACCACAACCTAAGAATGCAGTGAAAAATG
CTTTATTTATGTGAATTGATGCTATTGCTTTATTTGTAAACTATAAAGCGT
CAATAAAACAAGTTATCTGAGTCCACCGAATTCTTTAATTAAGGCCGGCCTCGAG
CGGCGCAGCTACGATCCCGCAGCGCCACAGCTCCTCGGATGTACAGGCATGCG
TCGACCCTCAGTAAAGGCTATAGTGAGTCGTATTACGGAATGGCGCGTCT
TTACAAACGTCGACTGGGAAACCCCTGCGTACCACCAACTTAATCGCCTTGC
AGCACATCCCCCCTTTGCGCAGCTGGCTAATACGAAGAGCGCCGCCACCGAT
CGCCCTTCCCAACAGTTGCCGACAGCTGAATGGCGAATGGCGCTTCGGT
AATAAAGCCCCGCTTGCAGGGCTTTTTTTT
>pGEM-UL12

GGGCGAATTGGGCCCCGACGTGCAATGCTCCCCGGCCGACATGGGCGC GGATA
TCGCCATGGAGTGTCACGGTAGGGCCCAGCATGTCGCCGCGGGGACGCACCGTGAC
TAAGCGTCTGGGCCCTGGGACGAGCTCCCTCGTGGCCCCGAGCAGCTTTCTCGTCCC
GCCCGCCTCGACGTGATCAGCACGCCACCGCAGACGAAAAGCCCCGGGCCGCGTCGCCGCCACTTTCTGACGCCTCAGGGCCTCCGACCCCAGACAT
TCCGCTATCTCCTGGGGGCACCCAGCCCGACCCGGACGCCGATCCCGAC
TCCCCGGACCTTGACTCTATGTGGTCGGCGTCGGTGATCCCCAACGCGCTGCC
CTCCCATATACTAGCCGAGACGTTCGAGCGCCACCTGCGCGGGTTGCTGCGC
GGC

GTCCGCGCCCCCTGGCCATCGGTCCCCTCTGGGCGCCTGGATTATCT

> pGEM-UL12

GGGCGAATTGGGCCCCGACGTGCAATGCTCCCCGGCCGACATGGGCGC GGATA
TCGCCATGGAGTGTCACGGTAGGGCCCAGCATGTCGCCGCGGGGACGCACCGTGAC
TAAGCGTCTGGGCCCTGGGACGAGCTCCCTCGTGGCCCCGAGCAGCTTTCTCGTCCC
GCCCGCCTCGACGTGATCAGCACGCCACCGCAGACGAAAAGCCCCGGGCCGCGTCGCCGCCACTTTCTGACGCCTCAGGGCCTCCGACCCCAGACAT
TCCGCTATCTCCTGGGGGCACCCAGCCCGACCCGGACGCCGATCCCGAC
TCCCCGGACCTTGACTCTATGTGGTCGGCGTCGGTGATCCCCAACGCGCTGCC
CTCCCATATACTAGCCGAGACGTTCGAGCGCCACCTGCGCGGGTTGCTGCGC
GGC
TTTTTTGGTTGCAAGCGACAGATTACGCAGAAAAAGGATCTCACAG
GATCCTTTGATCTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACG
TTAAGGGGTGGTATGAGATTATCAAAAAAGGATCTTACCTAGATCTTTT
TAAATATATATATATAGTAAACTTTTG
GTCTGACAGTTACCAATGCTTATACAGTGAGGACCTATCTCAGCGATCTGTC
TATTTCGTTCATCCATAGTTGCTGACTCCCCGTGCTGTAGATAACTACGATA
CGGGAGGGCTTACCATTACCTGCCCAGTGCTGCAATGATACCGCGAGACCCAC
GCTCACCAGGCTCAGATTATATCGAATAAAAAACAGCCAGCCGGAAGGGCCGA
GCGCAGAAGTGGTCTGCAACTTTATCCGCCTCCATCCAGATCTATTATGGTT
GCCGGAAAGCTAGAGTAAGTGAGTTCGACCGATATAGATTTGCGCAACCGTTGT
TGCCATTGCTACGAGCATCGTGGTGTCAGCTCAGCTGTTTGGTATGGTCTTACAT
TCAGCTCCGGTTCCAAACGATCAAGGCGAGTTACATGATCCCCGATCTGTGC
AAAAAGCGGTTAGCTCCTTCGGGCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTC
ATGCCATCCGTAAGATGCTTTTCTGACTGGTAGTACTACTTAACCCAGTACTT
CTGAGAATAGTGATGCGCGCCACCCGAGGTGCTTTCGGGCCGCCTCATACCG
GATAATACCCGCAGCACTTAGAAGAACTTTAANAAGTGCTTCATCATGGAAAAAC
GTCTTTCGCGCGAAAATCTCAGGATCTTACCCGCTGGTAGATCCAGTTCG
ATGTAAACCACTCAGTGCACCAGAATCTGATCTTACAGCATCTTTTACTTTCAG
CGTTTTCTGGGTAGAAAAACAGGAAGGCGAAAATGCCGCAAAAGGGGAAT
AAGGGCGACACGAAATATGGAATACCTACATCTCTCCTTTTTCAATATTATT
GAAGCATTTATCAGGGTTATTGCTCAGCATGAGCAGACATATTTGAGATTT
TAGAAAAATAAACAAATAGGGGTTCGCCGCAGACATTTCCCGAAAAAGTGCCAC
CTGATGCAGGTTGAAAATACCACACAGATGGTAAGGAGAAAATACCCCATCA
GGAAATTGGAAGCTGAATATTATTTGTAAAAATTCTCGTTAAATTTTTGTAAA
TCAGCTCATTTTTAAACCCATAGGCGAAATCGGCAAAATCCCTATAATTAATCA
AAAGAATAGACCAGATAGGGTTGAGTGGTGTCCAGTGGTTAAGGAATAGGTC
CACTATAAAGAAGCTGAAGTGCAGCCTGAAAAGGCGAAACCCGATTTAGAG
GGGCGATGCGCCACTACGTGAAACCATACCCCTAAATCAAGTTTTTGGGCTCG
AGGTGCCTGAAAGCACTAAAACGGAACCCTAAAGGGAGCCCSCGATTAGAG
CTTGAAGGGGAAAGCGCGCGACACGCTGGCGAGAAAGGGAAGGGAAGGCG
AAAGGAGCGGCGCTAGGCGCCTGGCAAGTGTAAGCCGTCACGCTGCGCGTA
ACCACCACACCCCGCGCGCCTTTAAATGCGCCGCTACAGGGCGCGTCCATTCGCC
ATTCAAGGCTCGCAACTGTTGGGAAAGGCAGTCGCGGCTGCTCTCCTTCGCTTA
TTACGCGCAGCTCGGAAAGGGGATGTGCTGCAAGGCGATTTAGTGGGTAAG
CGCAGGGTTTTCCCAGTCGACAGTGTAAAACGACGGCCAGTGAAATTGTA
ATACGACTCAGTATA
>pGEM-UL12-HA

HAGGGCGAATTGGGCCCCGACGTGCGATGCTCCCCGGCCGCCCATG GCCCGCGGGA
TATCGCCAATGGAGTCCACGGCTAGGCCCAGCATGTCCGGCCGGGACGCACCGTG
ACTAAGCGTCCCTTGGGCCCTGGCGAGGACACCCCTCGTGGCCCCGACAGCC
CCCCCAAGCGCCCCCGCCCTAACGCTTCCGCTGACAACCACCTTCCGTCCC
CTGCCCCCCCCACCCAGCACATCAGACTGTGGACCCGAGCTCCCATTCGC
CCGTTAACCCCACTGATCAGCAACGCAACCGACACCCGACAGACGAAAAGCC
CCGGGCCGCTGCAGACTTTACGCTTGGACCCGAGCTCCCATTCGC
CTGCCCTTCTGCTGCCCCCTGGGGGCACCCACCGCGACCCGGACGCCGATCCCG
ACTCCCCGGGACCTTTGACTCTATGTGGTCGGCGTCGGTGATCCCCAACGCGCTG
CCCTCCCATATACTAGCCGAGACGTTCGAGCGCCACCTGCGCGGGTTGCTGC
GCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCC
GGCAAAAACACCACCGCTGAGCGGTGTTTTTTTTGTAATGCAAGCAGGAA
TTACGCGCAGAAAAAAGGATCTCAAGAAGACGTTTGAAGCTAGGCAATAGGAA
GTCTGAGCGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGA
TTATCAAAAAAGGATCTTCACCTAGATCTCCTTTAAAATAAAAATGAAAGTTTTAA
ATCAATCTAAAGTATATATGAGTAACCTTTGCTAGCTGACAGTTACCAATGCTTAA
TCAGTGAGGCACCTATTCAGGCAGATCTGCTATTTGTTCCATCCATAGTTGCC
TGACTCCCCGTCGTGTAAGATAACATGCATACGGGAGGGCTTACCATCTGGCC
CCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATC
AGCAATAAAAAACACGCCAGCGGAAGGCGCGAGCGAGAAGTGCTCTGCAAC
TTTATCCGACTTCCATACGTAATTAATTGGTTGCGGGAAGCTAGAGTAAGTA
GTTGAGGGCTTAAATAGTTTGGCCAGAAGCTTTGTTGCCATGACAGGGCATCGT
GTGTCACCGTCCGTTGGTATGCTTCTCATTGACTCGTCCGTTCCCAACGATC
AAGGCAGGTTACATGATCCCCATGGTAAGTTTGGCCAGGCTTTGCTACAGTCTTTC
GGTCCTCCGATCGTTGTCAGAAGTAAAGTTTGCGCGAGTGTATACACTCATGGT
TATGGCAGCAGCATAATTCTCTTACTGTACGATCCATCCGTAAAGATGCTTCTT
CTGTGACTGTTGAGTACTCAACCAAGTGCTATTGAGAATAGTGATATGCGCGG
ACCGAGTTGCTTGGCCCAGCGTCAATACGGGATAATACCGCGCCACATAGC
AGAATTAAAAAGTGCTCATCATTGGAAAAACGTTTCTCGGGGCGAAAACTCT
CAAGGATCTTTCCGCTGTTGAGATCCAGTTCGATGAAACCAACTCGTCGACCC
AACTGATCTTCAGCATCTTTTACTCCCACCAGCTTTCTGTTGGTGAGCAAAAAAC
AGGAAAGGCAAAATGCCGCCAAAAAGGAAATAAGGGCGACACCGGAAATGTTG
AATACCTCATACTCTTCTTTTTCAATATTATTGAAAGCATTTTATCCAGGTTATG

TCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAAACAAATAGGG
GTTCGCCGCACATTTCCCGAAAAAGTGCCACCTGATGCGGTGTGAAATACCG
CACAGATGCGTAAGGAGAAAATACCGCATCAGGAAATTGTAAGCGTTAATAT
TTTGTTAAAATTGCGGTTAATTTTTGTTAAATCAGCTCATTTTTTTAACCAATA
GGCCGAAATCGGAAAAATCCCTTTTATAAAATCAAAAAATAAGACCCGAGATAGG
GTGGAGTGTTGTTCCAGTTTTGGAACAAGAGTCCACTATTAAAGAACGTGGAC
TCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGCGATGCGCCACTACGTG
AACCATCACCCCTAATCAAGTTTTTTGGAAGTGGGTGCGTGCCAATAGCAATTAA
TCGGAACCCTAAAGGGAGCCCCCCTTTAGCTGGACGGAAGAAGCCGCG
GAACGTGGCGAGAAAGGAAGGGAAGAAGCAAGGAGACGGGCGGCTAGGG
CGCTGGCAAGTGTAGCGTCAGCTCGCTCGTAACCACCAACACCCGCACCGGC
TAAATGCGCCGCTACAGGGCGCTCCATTCGCCATTGCGTCGCAACTGTT
GGAAAGGGCGATCGGTTGCGCCGCTTTCTCTCGCTATTACCCAGCTGGCGAAAGG
GGGATGTCGTCGCAAGGCGATTAAGTTGGGTAACCGCCAGGTTTTTCCCCAGTCA
CGACGTGTTAAACGACGCCAGTGAATTGTAATACGACTCCTATA
>pGEM-UL12.5

GGGCGAATTGGGCCCCAGTGCATGCTCCTCCGGCCCGCCATGGCCGCGGGATA
TCGCCGCCACCATGTGGTGCCTCGGCTGATCCACAAACGGCGCTGCCCTCCCAT
ATACTAGCCGAGACGTCTCGAGCGCCACCTGCGCGGGTTGCTGCGCGGCTCC
GCGCCCTCCTGGCCATCGGTCCCTCTGGGCCCGCCTGGATTATCTGTGTTCC
CTGCGCTCTCGAGAAGCTGGGTTATGATGGGACCAGGACTCGGTCGGC
ACCTATGCGCCTGACGCGCCGCGGGCCTACGACCGC
ATCCACCCTCC
GCTGGGGCCCCCAGGGTCCGTGTTTCTCGCCCCAGTGGCTGAAGCACAACGC
CAGCCTGCGGCCGGATGTACAGTCTTCGGCGGTGATGTTCGGGCGGGTGAAC
GAGCCGACGGCCCGAAGCCTGCTGTTTCGCTACTGCGTGGGCCGCGGACG
507
ACGGCGGCGAGGCCGGCGCCGACACGCGGCGCTTTATCTTCCACGAACCCAG
CGACCTCGCCGAAGAGAACGTGCATACGTGTGGGGTCCTCATGGACGGTCAC
ACGGGGATGGTCGGGGCGTCCCTGGATATTCTCGTCTGTCCTCGGGACATTCA
CGGCTACCTGGCCCCAGTCCCCAAGACCCCCCTGGCCTTTTACGAGGTCAAAT
GCCGGGCCAAGTACGCTTTCGACCCCATGGACCCCAGCGACCCCACGGCCTC
CGCGTACGAGGACTTGATGGCACACCGGTCCCCGGAGGCGTTCCGGGCATTT
ATCCGGTCGATCCCGAAGCCCAGCGTGCGATACTTCGCGCCCGGGCGCGTCC
CCGGCCCGGAGGAGGCTCTCGTCACGCAAGACCAGGCCTGGTCAGAGGCCCA
CGCCTCGGGCGAAAAAAGGCGGTGCTCCGCCGCGGATCGGGCCTTGGTGGAG
TTAAATAGCGGCGTTGTCTCGGAGGTGCTTCTGTTTGGCGCCCCCGACCTCGG
ACGCCACACCATCTCCCCCGTGTCCTGGAGCTCCGGGGATCTGGTCCGCCGC
GAGCCCGTCTTCGCGAACCCCCGTCACCCGAACTTTAAGCAGATCTTGGTGC
AGGGCTACGTGCTCGACAGCCACTTCCCCGACTGCCCCCCCCACCCGCATCTG
GTGACGTTTATCGGCAGGCACCGCACCAGCGCGGAGGAGGGCGTAACGTTCC
GCCTGGAGGACGGCGCCGGGGCTCTCGGGGCCGCAGGACCCAGCAAGGCGT
CCATTCTCCCGAACCAGGCCGTTCCGATCGCCCTGATCATTACCCCCGTCCGC
ATCGATCCGGAGATCTATAAGGCCATCCAGCGAAGCAGCCGCCTGGCATTCG
ACGACACGCTCGCCGAGCTATGGGCCTCTCGTTCTCCGGGGCCCGGCCCTGCT
GCTGCCGAAACAACGTCCTCATCACCGACGACGGGGAGGTCGTCTCGCTGAG
CGGCCGCCTGCAGGTCGACCATATGGGAGAGCTCCCAACGCGTTGGATGCAT
AGCTTGAGTATTCTATAGTGTCACCTAAATAGCTTGGCGTAATCATGGTCATA
GCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAG
CCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCAC


ATTAATTGCCTGATGCTCACTGCCCCGCTTTTCCAGTCCAGCTGGAAAAACCTGTGCAG
AGCTGCTAAAATGAACTGCAGGAAGGGAAGGCGGTTGGATGCTGTTGGCAG
GCGCTCTTCCGCTTCGCTCACTGACTCGCTGCTCGTCTCGTCTCGTGCTC
GGCGAGCGGTATCAGCTCACCTAAAGGCGATAATACGTTATCCACAGAATC
AGGGGATAACGCAGGAAAGAACATGTGAGCAGAAGACGGAAGACCGTAAAAAAGGCCGCTG
GAACCGTAAAAGGCAGGGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTG
ACGAGCGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGGAACCCGACAG
GACTATAAAAGATACCAGCGGCTGGCTCTCCCTCTCGTGTCCGCTCTCCT
TTCCGACCGCTGCGTTTACCGGATACCTGTCCGCTCTTCCCTCCGCGGAAG
CGTGGCCGCTTTCTCATAGCTACGCTACGCTGTATGTTAGGTATCTCAGTCCGCTGTAGGTCG
TTGCTCCAAGCTGGGCTGTGTCAGCAACCAACCCGTTCCAGCCGACCGCTG
GCCTTATCCGGAACATCATCCTGCTTTCAGTCCCAACCCGCTAGAGACAGACTTATC
GCCACTGGCAGCGGCCCCTGGAATAACAGGATGACAGCGGAGGGTATGTAGGC
GGTGCTACAGAGTTCTTTGAAAGTGCTGCTGCTAGCTGAGGCTAAGCTAGAAGAA
CAGTATTTGGTAGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGTT
GGTAGCTCTTTGATCAGGGCAAAACCAACCCAGCTGTGAGCGGTGTTTTTTGT
TTGCAAGCCAGCAGATTTACGCCTAGAAAAGGATCTCAAGAGAGCTCCTT
ATCTTTTCTACGGGGCTGCTGCTGAGGCAAGAAACTCGTATAGGGGA
TTTTGGTCATGAGATTATCAAAAAGGATCCTCAGACGATCTCCTATATAATAAA
AAATGAAAGTTAAAATACTCAATAAGTATATATGAGTAAAACCTGGTGCTGACA
GTTACCATGCTTAATGAGGCTGACCTACTCTACGCTAAGCTTATTGCT
TCATCCATAGTTGCCTGACTCCCGTCTGCTGAGATAACTACGATACGGGAGG
GCTTACCATGCTGGCCACAGTGTGCAATGATACCGCGAGACCCACGCTCACC
GGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAG
AAGTGGTCCTGCAACCTTTATCCGCTCCATCCAGTCTATTAATTGGTTGCCG
AAGCTAGAGTAAGTAGTTGCACCAGTTAATAGTTTGGCAACGTTGTTGCAAAAA
GGTTTCCAAACGATCAAGGCCAGTTACATGATCCCCCATGTTTGTCGAAAAA
GCGGTTAGCTCCTCTCGTCTCCGATCGTTTGTCAGAAGTAAGTTGGCCGAGTA
GTTATCACTCATGGTTATGGCAGCAGTCTACATTCTTACTGTCATGCCAT
CCGTAAGATGCTTTTCTGTGACTGTGGAGTACTCAACCAAGTCATTCTGAGAA
TAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATA
CCGCAGCCACACATAGCAGAACTTTAAAAAGTGCTCATCATGATGGAAAAACGTTTCTTC
GGGGCGAAAAACTCTCAAGGATATTACGCTGTGAGATTCAGTTCGATGTAA
CCCACCTGTCACCCAACTGATCTTACGTCTTTTACTTTACTTTACCCAGCTTTCT
GGTGAGCAAAAACAGGAAGGCCAAAATGCCCAGAAAAAGGGAAAAAGGAATAAGGGC
GACACGGAAAATGGTAAATACCATACATCTCTCTTTTCAATATTATTAAGCA
TTTTACAGGGTTATTGTCATGAGCGGATACATATTTGAATGTATTTTAGAA
AATAAAAAATAGGGTTCGCGCAATTCCCCCAAGGATGCCACCTGTATG
CGGTGTGAAATAATCCGCAACAGATGCAGAAGGAAATACCAGTACAGGAAA
TTTGTAAGCGTTAATATTTTTTGTTAAAAATTCGCTTTAAATTGTTAAATCAGCT
CATTAAAAACAAATAGGGCGGAAATCGGAAAAATCCCTTTAATAATCAAAAAGA
ATAGACCGAGATAGGGTGAGTGTTGTTCCAGTTTGGAGAAAGAGGTCCACTA
TTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCGCGTTAAGGAGTGGGTCGAGGAGTG
ATGGCCCACACTGGAACCATCACCACTCATATTCAAGTTTTTGGGGTCGAGGGATGC
CGTAAAGCAGTAAATCGCGAACCTAAAGGGAGGCGCCGCAGCTTGGAC
GGGGAAAGCCGCGAAACGTGGCGAGAAAGGAAGGGGAAAGAAAGCGGAAAGGAAAGG
GCGGGCGCTAGGGGCGCTGGCAGGATGTAGCGGTCAGGTGCCTGCTGCTAACCACCA
CACCCGCAGCCTTTATGCGCCCCTACAGGCGCAGCAGCTTTCGCTATTACGCC
CTGCACAACCTGTGGGAAAGGGGCGATCGGTGCGGGGCTCTTCGCTATTACGCC
AGCTGGCGAAGGGGATGTGCTGCAAGGCGGATTTGGTTGGGTAACGCAGCAG
GGTTTTCCAGTCACGACGTTTGAAGGACGACGAGTGAATTGTAATACGA
CTCCTATA

>pGEM-UL12.5-HA

 >pGEM-Ul12.5-ha

GGGCGAATTGGGCCCAGTGCGCATGCTGGCCCAGGCGCAGCATGGCGGCGGGATA
TCGGCCGCCCACCAGTGGCTGGGCTGCTGATCCCAACGCCCTGCTGCCTCCCAT
ATACTAGCGGAGACGTTCGAGCACCACCTGCGCAGGGTTGCTGCAGCGGCGTCC
GCGCCCTCTGGCCATCGGTCCCCTCTGGGCCCGCCTGGATTATCTGTGTTCC
CTGGCGGTGGTCCTCGAGGAGGCGGGTATGGTGGACCGCGGACTCGGTCGGC
ACCTATGCCACACAGCCGCAGGGGCGGCCCCGCCCCGCCCCAGGTGAT
GCCAGCTATGGGCCTGCTCCGGCGGGGCCTCACGACCGCATCCACCCTCC
GCTGGGGGGCAGGTCCGTGTTTCTCGCCCCAGTGGCTGAAGCACAACGC
AGGAGAACGCGGCCGAAGGCTGATATCGTCTGCTTTGCCTACTGCGTG
GGGCCGCCGCGGAGGCCGGCGCCGACACGCGGCGCTTTATCTTCCACGA
CGACCTCGCCGAAGAAGACGTGCATACGTGTGGGTCTTCATGAGCCGTCAC
ACGGGGATGGTCGGGCCTCGCTATATTCTCGCTCGTCTCGGGACATTCA
CGGCTACCTGCGGCCCTTGATATTTCTCGTCTCGTCTCGGGACATTCA
GGCCCGCAGTGACGCTTTCGACCCCATGGACCCCAGCGACCCCACGGCCTC
CGCGTACGAGGACTTGATGGCAACACCGTGTCGCGCCGCTTTACGTG
ATCCGGTGATCAGCTCAGCAGCACTTTCGACCCCATGGACCCCAGCGAC
CGCCTCGGGCGAAAAAAGGCGGTGCTCCGCCGCGGATCGGGCCTTGGTGGAG
TTAAATAGCGGCGTTGTCGAGGAGGCTCTGCTCTTATGCTCAGGAC
CGGACGGACACACCATCTCCCCGTCTGGAGCGTCTCCGAGGATCTGGGTCCGC
GAGGCCGTCTTCCGCAAGCCCGAGCCGTCAACCGGAACTTTAAGCTGAG
AGGCTACGTGCTCGACAGCAACTTCGGCAGTCGTCCCACCCCGCATCTG
GTGACGGTCTCTCGGCAAGCCACCCGAGGAGGAGCCGCAGTCGATCTC
GCCTGGAGGACGGCGCCGGGTCTTCGCGGGCGGCAAGCCACAGCGGCT
TGAAGCCAGTTACCTTCGGAAGGAGATTGGTAGCTTTGATCCGGCAAAACA
AACCACCGCTGATGCGGTTGTTTTTTTGTGTTTGCAAGCAGCAGATTACGCGCA
GAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGGTCGTGACGC
TCAGTGGAACGAAAAACTCACGTTAAGGGGATTGTGGTCATGAGATTATCAAAA
AGGATCTTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCATTCTA
AAGTATATATGAGTAATTGGTCTGACAGTTACCAATGCTTAATCAGTGAG
GCACCTATCTCAGCGATCCTGCTATTTTCGCTTCATCCATAGTGGCCTGACTCCCC
GTCTGTGTAGATAACTCAGATACGGGAGGGCTTACCCTCCTGGCCCCAGTGCTG
CAATGATACCGAGAGCCACGCTCACCAGCTGCAGATTTATCAGCAATAAAA
CCAGCCAGCCGAAAGGGCCGAGCGCAGAACGGTGTCTCGAACTTTTATCCGCC
TCCATCCAGTCTTAATTGTGGCCGGGAAGCTAGAGTAAGTAGTTAGTGCCGAGT
TAATAGTTGGCGAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTTCAGCT
CGTCGTTCGTATGGCCTTCATCAGCTCCCGGTTCACCAACGTCAGCAAGCGAGTT
ACATGATCCCCCATGTTGTGGCAAAAAAGCGGTTAGCTCCTTCGCTGCTCCGAT
CGTTGTCAGAAAGTAAATGGCGCAGTGGTTATCTCTCTCATGTTTATGCGACGAC
TGCTATATCTCTTACTGTCACTGCCATCGGTAAGATGCTTCTGCTGTGACTGGT
AGTACTCAACCAAGTCTTCTGGAATAAGTGTATGCGGGCGACCGGAGTTGCTC
TTGCCGGGCGTCAATACGAGGATAATACCGCGCCACATAGCAGAACTTTAAAA
GTTGCTCATATTGGAAACGTTCTTCTGGGGCGAATACTCTCAAGGATCTTACC
GCTGTGAGATCCAGTTCGATGAAACCACTCTGTGCAACACAACTGTATCCTCAG
CATCTTTTACTTCTCACCAGCTTTCTGGGTAGCAAAAAAGAGGAAGGCAAAA
TGCCGCCCCAGGGAATTACGGCAGCAGCAGGTTATTGTCATGAGCGGAGA
>pFN22K::ICP8

N-terminal HaloTag fusion to ICP8.

CCAAAATGACCAGCACTTCCAAAATGTGGGATTAACCCCGCCCGTTGACG
CAAATGGCCGGTGAGGCGGTGAGGTGGTCTATATAAGCTAGATCTGGTT
TAGTGAACCCTCAGATCACTAGAAGCTTAATACGACTCACTATAGGGTACGA
TTAGGCTGACTATAGAATAAGGAGAATTACATCATGGCAGAAATCGGTACT
GGCTTTCATTCGACACCCCATATTATGGAAGTCCCTGGGGCGAGCGCATGACTA
CGTGCATGTTGGGTCGCAGCGATGCCACCCCTGCTGTGCTGCACCGAATACC
CGACCTCCTCCATCTGCTGCAGACACCTCAGATGGGACACCCCACCGATGCTG
CGCTGCATTGTCCTCCAGACCAGCAGTCATGGGATGCTCTGAGCTACAGGCG
TGGGTATTCTCTGACCAGACCCACGTTCCGCTTCATGGGATGCTCTCATCGAAGCC
CTGGGTCTGGGAAGAGGTCTGCTCTGGTCATTCACGACTGGGGCTCCGCTCTGG

GGAGTTCTCCGCTCTATCCCCGACTTGGGAGAATGGCCAGAATTTGGCCCGC

GAGACCTTCCAGGCCTTTCCGCAACCACCGACGTGGCCCGCAAGCTGATCATCG

ATCAGAACCAGTTTTTATCGAGGATCGCTGCTCCGATGGGTGTGTCGTCCCGCCTG

ACTGAAGTCGAGATGGACCATTACCGCGAGCCGGTCTGAATCTGCTGCAAGAAGACAACCCGGACCTGATCGGCAGCGAG

ATCGCGCGCTGGGCTGTCGACGCTCGAGATTTCCGGCGAGCCAACCACTGAGG

ATCTGTACTTTTCAAGAGCGATAACCGCGATCGCCATGGAGACAAAGCCCAAGACAACCATCAAAGGTCCCACCGGGCCCCTGGGATACGTGTACGCTCGCGTGTCCGTCCGAAGGCATCGAGCTTTGGCGTTACTGTCGGCGCGAGCG

GCGATGCCGACGTCGCCGTGGCGCCCCTGGTCGTGGGCCTGACCGTGGAGAGCGGCTTTGAGGCCAACGTAGCCGTGGTCGTGGGTTCTCGCACGACGGGGCTCGGGGGGTACCGCGGTGTCCCTGAAACTGACGCCATCGCACTACAGCTCGTCG

CGGCTTTGAGGCCAACGTAGCCGTGGTCGTGGGTTCTCGCACGACGGGGCTCGGGGGGTACCGCGGTGTCCCTGAAACTGACGCCATCGCACTACAGCTCGTCG

GGGCTTTCAAGAGGAGGCCTGGTGCATCAAAACAACCGGCTTCTGACCTGGGAGGC
TCGGACAAGGTAAACCATAAGCGGGGCCGAGGTCACCGCATACCCGTGTATC
CGTTTGAGCTGTTCATGCGCAGATTTAAGCCGAGTCATCGCCGAGCCGTTCACC
GCCAAACCGACGATCGATCGGGGAGAATTTTACCTACCCCGCCTCCGCTTAAAA
CGGCCCCCTCAACCGGCTCTCGGAGGCAGGCGTCGTGGCCGACAATACGCTGG
GCACTGCGATGCGAAACGATGGAGCCGTGGCCCGCCGCGCCGCGCGGACCTGG
CGTTTGACGAAACCGACGACGGCGCGCCGCGCTCCCGCGCAGACCTACGTTCCAC
GGGCTTCGAGCCACGCCGCAGGAACATACCCGCTTCGGTTTTTAA
CCGCCCCCTCAACCGCCTCCTGTTCGAGGCGGTCGTGGGACCCGCCGCCGTG
GCACTGCGATGCCGAAATGTGGACGCGCTGGCCCGCCGCGCCGCGGACCTGG
CCGCAAGGGGCCGCGCCGCCTAGTGGACGCTGCTGGCCAAGATGCTG
GGGATACAGGACGACGGCGGACCATCATGAGACGTACCGCGCGGCGAC
CGAGCGCGTCATGGCCGAACTCGAGACCCTGCACTACGTGGACCAGGCGGTC
CCACGCGCATTGGGGCGGCTGGAGACCATCATCAACCAACCAGCCAGGC
ATACGCTGGTAGAACAACGTCAGGCGAGGTTCGTCGAGCCCGAGGCTGGG
TGATGCACAACTTGGTAGGGAGGGAGGAACCACTTAAGGTCGACAGGCTTC
CGAGGGCAAACCACGCCATGTCCCTGACGCTGGACCGTACGCGTGCGGGCA
TGCCCCCTGACTTCAGTCTCAGCGGCGGCGATCCAACCTCAGCGGCTGCAT
CCTGGCCCTGAGCCAGTGCCACGGGGTGTTCGGCGGACAGTCGGTCGAGGG
CGCAACTTTGCAGATTAATCAAATCACAAACCGGTGAGCTGAGCTCGGGC
GGGGAGGAACTTCAAGTTTCGCGACGGTCTGGGAAAGG
CGAGGCCAACCACGCCATGTCCCTGACGCTGGACCGTACGCGTGCGGGCA
TGCCCCCTGACTTCAGTCTCAGCGGCGGCGATCCAACCTCAGCGGCTGCAT
CCTGGCCCTGAGCCAGTGCCACGGGGTGTTCGGCGGACAGTCGGTCGAGGG
CGCAACTTTGCAGATTAATCAAATCACAAACCGGTGAGCTGAGCTCGGGC
GGGGAGGAACTTCAAGTTTCGCGACGGTCTGGGAAAGG
CGAGGCCAACCACGCCATGTCCCTGACGCTGGACCGTACGCGTGCGGGCA
TGCCCCCTGACTTCAGTCTCAGCGGCGGCGATCCAACCTCAGCGGCTGCAT
CCTGGCCCTGAGCCAGTGCCACGGGGTGTTCGGCGGACAGTCGGTCGAGGG
CGCAACTTTGCAGATTAATCAAATCACAAACCGGTGAGCTGAGCTCGGGC
GGGGAGGAACTTCAAGTTTCGCGACGGTCTGGGAAAGG
CGAGGCCAACCACGCCATGTCCCTGACGCTGGACCGTACGCGTGCGGGCA
TGCCCCCTGACTTCAGTCTCAGCGGCGGCGATCCAACCTCAGCGGCTGCAT
CCTGGCCCTGAGCCAGTGCCACGGGGTGTTCGGCGGACAGTCGGTCGAGGG
CGCAACTTTGCAGATTAATCAAATCACAAACCGGTGAGCTGAGCTCGGGC
GGGGAGGAACTTCAAGTTTCGCGACGGTCTGGGAAAGG
CGAGGCCAACCACGCCATGTCCCTGACGCTGGACCGTACGCGTGCGGGCA
TGCCCCCTGACTTCAGTCTCAGCGGCGGCGATCCAACCTCAGCGGCTGCAT
CCTGGCCCTGAGCCAGTGCCACGGGGTGTTCGGCGGACAGTCGGTCGAGGG
CGCAACTTTGCAGATTAATCAAATCACAAACCGGTGAGCTGAGCTCGGGC
GGGGAGGAACTTCAAGTTTCGCGACGGTCTGGGAAAGG
CGAGGCCAACCACGCCATGTCCCTGACGCTGGACCGTACGCGTGCGGGCA
TGCCCCCTGACTTCAGTCTCAGCGGCGGCGATCCAACCTCAGCGGCTGCAT
CCTGGCCCTGAGCCAGTGCCACGGGGTGTTCGGCGGACAGTCGGTCGAGGG
CGCAACTTTGCAGATTAATCAAATCACAAACCGGTGAGCTGAGCTCGGGC
GGGGAGGAACTTCAAGTTTCGCGACGGTCTGGGAAAGG
CGAGGCCAACCACGCCATGTCCCTGACGCTGGACCGTACGCGTGCGGGCA
TGCCCCCTGACTTCAGTCTCAGCGGCGGCGATCCAACCTCAGCGGCTGCAT
CCTGGCCCTGAGCCAGTGCCACGGGGTGTTCGGCGGACAGTCGGTCGAGGG
CGCAACTTTGCAGATTAATCAAATCACAAACCGGTGAGCTGAGCTCGGGC
GGGGAGGAACTTCAAGTTTCGCGACGGTCTGGGAAAGG
CGAGGCCAACCACGCCATGTCCCTGACGCTGGACCGTACGCGTGCGGGCA
TGCCCCCTGACTTCAGTCTCAGCGGCGGCGATCCAACCTCAGCGGCTGCAT
CCTGGCCCTGAGCCAGTGCCACGGGGTGTTCGGCGGACAGTCGGTCGAGGG
CGCAACTTTGCAGATTAATCAAATCACAAACCGGTGAGCTGAGCTCGGGC
GGGGAGGAACTTCAAGTTTCGCGACGGTCTGGGAAAGG
CGAGGCCAACCACGCCATGTCCCTGACGCTGGACCGTACGCGTGCGGGCA
TGCCCCCTGACTTCAGTCTCAGCGGCGGCGATCCAACCTCAGCGGCTGCAT
CCTGGCCCTGAGCCAGTGCCACGGGGTGTTCGGCGGACAGTCGGTCGAGGG
CGCAACTTTGCAGATTAATCAAATCACAAACCGGTGAGCTGAGCTCGGGC
GGGGAGGAACTTCAAGTTTCGCGACGGTCTGGGAAAGG
CGAGGCCAACCACGCCATGTCCCTGACGCTGGACCGTACGCGTGCGGGCA
TGCCCCCTGACTTCAGTCTCAGCGGCGGCGATCCAACCTCAGCGGCTGCAT
CCTGGCCCTGAGCCAGTGCCACGGGGTGTTCGGCGGACAGTCGGTCGAGGG
CGCAACTTTGCAGATTAATCAAATCA
ATGGTCTGGTTGGGTTTAGACATCACGAAATACTACGGCATGGGCCGGCAACG
ACCGTGTGTTTCAGGCCGGGAACTGGGCCAGCCTGATGGGCGGCAAAAACGC
GTGCCCCTCCTCTATTATTGGACACACCCGCCAAGTTCTCCTGTGCCCTGCCC
GGGCCGGTTGTGTGCGCGGCCTCGAACCTCGGCGGCGGAGCGCACGAAAG
CTCGAGTAGCTGTTCGTCGACCCAGCTGAAAAAGCTGGGCCCAGCAGCAAGCA
AGCTGCAGATCGAGGACTGGCTGGCGCTCCTGGAGGACGAGTACCTAAGCGA
GGAGATGATGGAGCTGAGCTGACGGCAGCGAGCCGGCAACGGCGAGTG
GTCGACGGACCGCCGGTCTGAGAGGCTGGCCGACAGCCGGAGGCGGCTAGTCAG
CCAACCTCGGCAACGCCTGGGGGAGGTGTATACTTTGGGGATTTGGCTGAG
GACGACACGCGACGCCGGGCTCCTGGGGCCGGCCGGGAGGGACCCCACTTTG
CCGGCCGAAAACGGGCCCTCCACGGGGATGACCCGTTTGGGGAGGGGGGCCC
CGAAGGGGGAGAGCCTGACGTTGGATATGCTGGTTTAAAGCAATTCGGGC
TCGGTACCGGGGGATCTCTCTAAGTGCGACCTGACATCGAAGCTGATCACG
GCTGCAAAACGCGAAAAAGCTGGGAGGACGGTGTGTGCTGCCACCCCTGAGC
AATAACTAGCATACAAAACCCCTTGGGGGGCCGGGCTTCGAGCAGACATGATAAGAT
ACATTTGATGAGTTTGGAACAAAACCACAACTAGAAATGCAGTGAACATGCTTT
TATTGAGAATTTGTGATGCTATTGCTTTTATTGTAAACCATTATAAGCCTGAA
TAAACAAAGTTAACAACAAACTTGCAATTTCATTATTTATGTCTTTCCAGGGGG
AGATGTGGGAGGTTTTTTTAAGGCAAGTTAAACCCTACAAATGTGGTAAAT
CGAATTCTAATGGATCTCCTTGGCCTTGGTTTTCCCTTTGTCGACAGATAGCC
GTAGCTGACATTTCATTCCGGGGGTCAGACCCGTTTCTGCGGACTGGCTTTCTAG
TGTTCCCGCTCCTTTAGCAAGCCCTTGCAGCTAATGCTTTGCAGGCAGCGTGAG
CTTCAAAAGAATTGCCAGCTGGGGGC GCCCCTCTGGT AGTGGTTGGGA
CAAAGTAAACTGGATGGCTTTCTCCGCGCCGCTTTGGGGTGAGGAGGCTTATTCGCGCTA
TGACTGGGCAACAACAGACAATCGGCTGGCTCTGATGCGCCGGCGGCTGGCTGGCT
TCAGCGAGGGGGC GCCGTTCTTTTTTGGTCAAGACCAGACCCTGTCCGGTGCCCT
GAATGAAACTGCAGGACGAGGCAGC
CGGCTATCGTGGCTGGCCACGACGGG
CGTTCCTTGCGCAGCTGTGCTCGAGTTGTCACTGAAGCGGGAAGGGACTGG
CTGCTATTGGGCAGATCGGGGGCAGGATCTCCTGTCACTCTACCTTGGCTCC
TGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTT
GATCCCGCTACCTGCCCATCAGGCCACCAAGCGAAACATCGCATCGAGCGAG
CACGCACTCGGATGGAAGCCGGTTCTTGTCGATCAGGATGATCTGGACAAAGA
GCATCAGGGGCTCGCGCCAGCGCCTACTGTTCGCGAGCCAGCTCGATGATG
CCGGATGGTGAGGATCTCGTCGTGACTCATGGCGATGCCTGCTTGCAATAT
CATGGTGAAATAAGGCCGCTTTTCTGGATTTGATTCATCGACTGTGGCGCTGGGT
TGCGGGACCCGCTATCAGGACATAGCCTGCTACCCTGTGATATTGGCTGAAGA
GCTTGCGGCGAATGCGGTCGACCCTCTCTCTCGGCTTTACGATCTCGGCCGCTC
CCGATTCGACGCCTCGCCTCTATCGGCCCTTCTTACGAGGTCTTTCCTTGTAGCG
GGACTCTGGGGTTGGAATGACCGACCGAACGGACAGCCCACCCACTGTACAGCT
CACTCAAAAGCGCTATTACGGTATCCACAGAGAACGGGATAACGCAGG
AAGAACATGTTGACAAAAGGCCAGCAAAAGGGCCAGAAGCCGATAAAAAAGGC
GCGTTTGCTGGCGTTTTCCATAGGGCTCGCCGCCCTGACGAGCATGACACAAAA
TCGACGCCTCAAGTCAAGGTGGCACAACCCGACAGGACTATTTAAGATACCA
GGCGTTTCCCCCTGGGAAGCTCCCTCGTGCTCTCTGTTCCGACCCGCTGCCGC
TTACCGGATACCTGTCCGCTTTTTCCCTTCGGAAGCGTGCCTCTGGCTCTCTCAT
AGCTCACGCTGTAGGTATCTCATTCGAGTGCATGCTGCTTTCCGACTCCAGCTG
CTGTTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAAGTTATCTCGTCTTGAGTCCAACCCGGTAAGACACGACTTATGCCTACAGACCTTTTTGTATATCT
GCACGCTCTCTAGCCAGTTACCTTGACCGGAAAGAGGATTTGGGTAGCTCTTTGATCGCTGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGA
GGCAAACAAACCACCCGCTGGTAGCGGTGTTTTTTTGTTTGCAAGCAGCAG
TTACGCGCAGAAAAAGGATTTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTGCTAGCAGA
TTATCAAAAAAGGATGCTTCCTACCTAGATCTCTTTATAGTCGCAGAAATACAGGAAC
GCACGCTGGATGGCCCTTCGCTGGGATGGTGAAACCATGAAAAA
TGGCAGCT
TCAGTGGATTAAGTGGGGGTATGTAATGTGCGCTTGACCCTCTGTTCTGCTATAGGTAT
TCATACGTTAAAATTTATACGGGCGCAGTTGCAGCAGTTTTTTCGGGTGGTTTG
TTGCCATTTTTACCTGTCTCGTCCGCTGTAGCAGCTGCTGAACCGGCGTTTACGCGGT
GCCTACAAATAGGGATTATGTAATTCACCTACTGTCTGCTCCCTCGTAGCC
TCGAGATAAACCACGAGTACTCCGGCCACGATGCCTCAGGCGTAGAGGATCGA
GATCT
>pFN22K::ICP8-GFP

N-terminal HaloTag fusion to ICP8-GFP.

CCAAAAATCAACGGGACTTTCCAAAAATGTCGTAATAACCCCAGCCCCTTGACG
CAAATGGCAGTGTAAGGCGTGTGGAGGAGTCTATATAAGCAGAGCTGGTT
TAGTGAACCGTCAGATCAGTGAAGCTTAATACGACTCACTATAGGGTTACGA
TTAGGTGACACTATAGAATAAGGAAGATTCATCATCGAGAAATCTCGGTACT
GGCTTTCCTATCCACCCCACTATAGGGAAGTCTGGCGAGCGATCATGACTA
CGTCTGATGATTTGATCCCGCGATGCACCCCCCTATGCTGTCTTCTGCACGGTAAACC
CGACCTCCTCCTACGTGTTGCAGCAACATCACTCAGCCGATGTTGCAACCCCAT
CGCTGCTTGGTGTCCAGACCTGATCGAGTTGGCAATGGCAAATTCCAGACAAACCAGACC
TGGGTTATTTCCTCGACGGACACGTCCGCTTTCATGGAATGCTCTGATCGAAGCC
CTGGGTCTGGAAAGGCTGGTGCTCTGGCTATCAGACTGGGCTCCGCTCTGG
GTTCACGGAGGTACGCCTGCAGCTGGGATGGTGTGCTCAGCCGCTCTGG
GTTTCCACTGGGCCAAGCGCAATCCAGAGCGCTGGGCGCCGCAAGCTGATCATCG
ATCAGAAGCTGACTTTTATCGAGGCTGCCGAGCTGGTGCAGCTCATCG
ACTGAGTGGACAGAACTTACCACGGACAGCGTTTCAATCTGTTGACC
GCCGAGCTGGCTGCTGCAAGAATACATGAACTGGCTGCAAGCCAGCTG
GTCCGAAGCAGGCTGTGTCTTCTGGGCAACCCAGGGCCTTCTGATACCCAG
GACGCTGCAGCAGCAGAGCTGGGAGAGCTGGTGCACTGACGACGGGG
GGCAACCACATCAAGGTCCTCCCCCGGCGGCCCTGGGATACGTGATACGCTGC
GCGTGTCGGCTCCGAAGGCGATCGAGCTGGCGTTACTGTCGGCGCGCAGCG
GCGATGCCGACGTCGCCGACGTCGCCGACGTCGCCGACGTCGCCGACGTCGCCG
GGGCTTTGAGGCCACGTAGCCGTGGTCGTGGGTTCTCGCACGACGGGGCTC
GGGGGTACCGCGGTCCCTGAAACTGACGCCATCGCACTACAGCTCGTCCG
TGTACGTCTTTGACGCGTCATCAGGCCCAGCGCCGCTCCTGGTACCGAG
GGGCTTCAAGGAGGCGTGCTGATCAAAACACCCACCTTTTCTGCACCTGGGAGGC
CGACGGCGGCATGGCCGAACTCGAGACCCTGCACTACGTGGACCAGGCGGCGGTCC
CCCAGCAGCCATGGGCGGCGGTCGTGGAGGACCATCATCAACCAACCCGCAGAGGCCCTGC
ATACGGTGGTGAACAACGTGCAAGCAGGCTGGAGGCAGGAGTGGACGCTGGGG
TGATGCGCAACCTGTTGAGGGAGGGAGGAACCTTCAAGTTTTCGCAGCAGCTGGGACCGCGAGGTGGAGCAGCTGATGCGCAACCTGGTGGAGGGGAGGAACTTCAAGTTTCGCGACGGTCTGGG
CGAGGCCAACACCAGCAGCCATGTCCCTGACGCTGGACCCGTACGCGTGCGGGCCA
TGCCCCCTGCTTCAGTTCTCGGGCGGCGATCCAACCTCGCCGTGTATCAGGA
CCTGGCCCTGAGCAGGTGACACCGGGGGTTTGTTCGGGCGGCCGAGTGGAGCA
TGTTTAACAACCGGGTCTGTCGGCCAAAACGCTGACGGTGCAGTGCAGTCTCTCGGA
GGGGGGCGGCTATCTGCAGCCCCCAGCCTAACGCGCGCGAGACGCGCCGGCCGC
GAGAGCCACCTGCGAGGGCGACGTTGGCGCCGCTGACCCCTGGGGTTTTCACAAG
AGCTGCAGCGTCAAGGACGCGGCTGTGTTTTTCGGGGGCGGCCATCTCTGGGAACGCGAGCAGCTTCGAGGGCGACGTTGCCCGCGTGACCCTGGGGTTTCCCAAGG
AGCTGCGCGTCAGAGCCGCGTGTTGTTCGCGGGCGCGAGCGCCAACGCGTC
CGAGGCCGCAAAGGGCGGCTGCGCGAGCTCAGGCTCCCAGCAGGCGGCCGCGGCCG
ACCGCAGTGGTTCTGGACGCGCCCTCAACGCAACCCAGCTTCCCAGCCCGGCT
CCTGTCGCAGGGAGACATCGAGACCATCGCTGTCATATAAAAAGTTTTTTCTCGG
ACTACGGCGCATAACCTTATTAACCTGGGCCCCCAACACGTAAGCGAGCAGCT
GGCGATGTACTACATCGAAACGAGATTTCTCGGCTACTCGATCTACGTCA
TACTTCATCAACACCCTACAGGCACTACGCGGGTTGCTCCGCGCCTCCCCCAAG
CGTGCAGCGCGCGCTGAGGACGCACCTGCGGACCGGCAGGCGGGCGGGCGCTGGAGGC
CGGGGCCGCGGGCTCAGGGACGCAGCGCATCGCCGCGCCCCTGTCATGCGGGCGGCCCTCC
ATGGTCTGTTGTTGGTTGAGCATCAGCAAATACTACGGCATGGCCGGCAACG
ACCGTGTGTACCAGCCGGGAAGCTGGGAGCCGCTGATGCGGCGGCAACG
ATGCCCAGCCTCTATTATTTTGACCACCCACCCAATCTCCGTCCTGGGCTTCCCT
GGCCCGGGTTTGTGTGCGCGGCCTCGAACCCTCAGCGGCGGAGCCGACGAAAG
CTCGCTGAGCTGGCGCTTGGGACAGCTCTACGCAGCTTTTGGGAGGGGCGGGCGCC
GCCAGTAGCTGTGGTGCCAGCAAGATGAAAGCCTGGGGCTCAGGACCCACGAG
TGGAGATGAGCGCTAGACCCGCGTGCTGCTGGGAGCGCGCAGCAGCAGCCCTAGTAC
GCTGACGCGGACCGGCGCCCTGAGGTGCGCGACAGCGAGGCGGGACGGGCGATGT
GTCGACGACGCGGCCCTGAGGTTGCGCGACAGGCGAGGCGGCGGCTGACTAG
CCAACACTCGGCAACGCGGCGGGGAGATTATTTTGGGAGGATGCTGCGAG
GACGACAAGGCCAGCCGCGCTCCTGCGCCGCGGCGGCGGGCGGCGGCGGCGAAGCC
CCGCCCCGAAACCCGGCGTCTCAGGGATGACCCGTTGAGGAGGGGCTGCCCTG
CGACAAAACAGGAGCATGACGTGTTGGGATGATGCGAGGGTGGGAGGGGCTGCC
GGGAACCTAGAAGTGAGCTGACCAGGGGGAAGATTGTAAGGCTCAGGGCGAAGC
GGAACGTGCTTCACTGGCGCTGTGCCATCATCTCCACACTGAGACGCGGCTTAC
AATGGGCGCACAAATTTTCTGTCCAGCGGAGAGGTCAGGTAGGTGACACATACG
GAAAGCTCACCCTGAAATTCATCTGCAACACTGGAAGCTCCCTGCTGACCAG
GCCAACACTGGTCTACCTTCCACCTATGCGCAGTGCTGCTGGTCTGCCAGATACC
CAGACCATATGAAGCAGCATGACTTTTCTAAGAGCGCCATGCCCCGAGGGCTA
TGTGAGGGAGAGAACCACCTTTTCTAAGAGTAGCGGCCGCGGAGGGCTA
GCTGAAAGTGAGTGAAGCTGACCAGGGGGAAGATTGTAAGGCTCAGGGCGAAGC
GGAATTTAAGGAAGATGGGAAACATTCTCGGCAACAGCTGGAATACAA
CTATAACTCCCAAAAAATGTGATCATCGGCGCCGACAGCCAAAGAATGGGCATC
AAGGTCAACTTCAAGATCAGACACAAACTTTGAGGATGATCCGTGCAGCTGGG
CCGACCATTATCAACAGAACACTCCAATCGGGCGACGCCCTTGCTCTCTCCC
AGACAACCATTACCTGTCCACCCAGTCTTGCCCTGTCCTAAAGATTCACACAA
AAGAGAGACCACATGGTCCTGTGGGAGTTTGTGAAGACCCTGCTGGGATGACAC
ATGGCATGGACGAGCTGTACAAGGTTTAAACGAATTCGGGCTCGGTACCCGG
GGATCCTCTCAGAGTCGACCTGCAGGCATGCAAGCTGATCCGGCTGCTAACAA
AGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCCGCTGAGCAATACTAGCA
TAAACCCCTTGGGCGGCCGCTTGAAGCATGATAAAGTACATTGTACAGATGAG
TTTGGACAAACAAACTAGAAATGCAAGTGAAAAAAATGCTTTATTTTGAAAA
TTTGTGATGTGCTATTGCTTTTATTTGTAACCATTATAAGCGTCAATAAAACAGTT
AACAACAAACATGCTTCTTTATATGTTTTCAGTTTACGGGGGGAGATGTGGG
AGTTTTTTTTAAGCAAGTAAAACTCTACAATAATGTGTAATAATCAATCTAA
TGGATTCCTTTGCGCTTTGCTTTTCCCTTGTCAGATAGCGTAGCTGAC
ATTCACTCGGGGTACGACCCGGTTCTCGGACTGGCTTCTATCGTGGTTCGCT
TCCTTTAGCAGCCCTTGCGCCCTGAGTGCCGACGAGGCTTACGGCTTCAAAG
AATGCCAGCTCGGGCGCCTCCTGGAAGGTTTGGGAAGCCTGCAAAGTAAA
CTGGATGGCTTTTCTTGCCCGCAAGGATCTGATGGCAGGGGATCAAGATCT
GATCAAGACAGAGTGACGGGTCTTTGCGCATGCTTTGAAACAAAGATGATTGC
ACGCAAGGTTCTCGGCAGCCTGGGTGGAGAGGCTATTCGGCTATGACTGGGC
ACAACGACAAATCGGCTGCTCTGAGTGCCCGCTGTTCCGGCTGCTGACGCAG
GGGCGCCCGTTCTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACCT
GCAGGACGCAGGCGCGCTATCGTGGCTGGCAGCAGGCGCGTCCCTTGACG
GCAGCTTGCTGCTGACGTTGTCACATGGAAGCGGAAGGACTGCTGCTATTGG
GCGAAGTGCCCGGGCAGGATCTCGTACCTGGCTCTCTGCTCTGTTGGCCAGA
AGTATCCATCATGCTGATGCAATGCGGCCTGCATACGCTTGATCGGCT
ACCTGCCCATTGCCACCACCAAGCGAAACATCGCATCGAGCGAGCACGCACT
GGATGGAAAGCGGTCTTGTGATACGAGTAGATCTGGACGAGACGAGTAGG
GCTCGGCCCAGCCAACTGTTCGCCAGGCTCAAGGCGGTATGCGCGGATGCT
GAGGATCTCTCGTCTGTAACACGATGCTGCTCTGGCTTGCAGAATATATGCTG
AAAATGGGCGCTTTTCTTGAGACTCACTGACTGTTGGCAGCGCTGGGCT
CCGCTATACGACATAGCGTTGGCTACCCGTGATATTGCTGATAGGAGACTTG
GGCGAAATGGCGCTACGCTTCTCGTCTCTTACGGTTATCGTACGCGCCTCCCGATTC
GCAGCGCATCGCCTCTCTATCGCCCTCTTCTGAGAAGTTCTGTTGACGCGGACTCT
GGGGTTGCAGATGACGACCAACGACGACCGCCAACCCGCTATACAGCTCAGCTCA
AGGCGGTAATACGCTTATCCACAGAATCGAGGGGATAACGCAGGAAAGAAACA
TGTGAGCATAAGGGGCGAGCGAAAAAGGCCGAGCGTTGACTACGACGAGCTACA
TGCGCGGTTCCATAGGCTCGCCCGCCCGCCCTGACGAGCATCACAAAAATCGACG
CTCAAGTCAGAGGGCGGAAACCCGACAGGACTATAAAGATACCAAGGCCTTTT
CCCCCTGGAAGCTCCCTGACCGCTCCTGTTCCGGTCCGCTTACCGGATACCTGTCCG
GCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAA
TTGTTGAGAAGCTCGTTATGCAGGGGTTGTTGATGCTCGTCTCGTCTCAAGCTCAA
AGGCGGTAATACGCTTATCCACAGAATCGAGGGGATAACGCAGGAAAGAAACA
TGTGAGCATAAGGGGCGAGCGAAAAAGGCCGAGCGTTGACTACGACGAGCTACA
TGCGCGGTTCCATAGGCTCGCCCGCCCGCCCTGACGAGCATCACAAAAATCGACG
CTCAAGTCAGAGGGCGGAAACCCGACAGGACTATAAAGATACCAAGGCCTTTT
CCCCCTGGAAGCTCCCTGACCGCTCCTGTTCCGGTCCGCTTACCGGATACCTGTCCG
GCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAA

CAAACCCACCGCTGGTAGCGGTGTTTTTTGTGGCAAGCAGCAGATTACGCG
CAGAAAAAAGGATTTCAAGAAGATCTCTTGATCTTTTCTACGGGGTCTGAC
GCTCAGTGGAACGAAAACACACCAGTTGAGGATTTTTGGTCATGAGATTATCAA
AAAGGATCTTTACCTAGATCCCCCTATAGTCCCGAAATACAGGAACGCACGC
TGGATGCGCCCTTCGGCTGGGGATGTTGAAACCATGAAAAATGCGAGCTTCAGTG
GATTAAGTGCGGGAATGTTGGCCTGTACCCCTCTGTTGGCATAGGTATTCCATAAC
GGTTAAAAATTTATCAGGCGCGATTGCGGCAGTTTTTCGGGTTGGTTTGGTGGCA
TTTTTACCTGTCTCGCTGCCGTATCGCGCTGAACCGCTTTTTAGCGGGTGGGATAC
AATTAAGGGATTATGGTAAATCCACACTTACCTGTCTCCTGATTAGCCATCGAGA
TAAAACCGCAGTACTCCGGCCACGATGCGTCCCGTTAGTGGAGGATCGAGATCT

>pUC57-Kan::NLS/HA-HumBeta

NLS/HA fused to humanized Beta.
>pUC57-Kan::NLS/HA-HumBeta
TCGCGCGTTTCGGTGATGACGGTGAACCTCTTCATACATGCAGCTCCCGGA
GACGGTCACAGCTTTGCTCAGAAGCGATGCCCGGAGCGAGACAGACAGCAGCGCTCCCGGA
GGCGCGTCAGCGGAGCAGACAAGCCCGTCAG
GGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCTTAACTATGCGGCAT
CAGAGCAGATTGTACTGAGAGTGCAACCATATGCCTGTTGGAATACCCGCACAG
ATGCGTAAGAGAGAAAAATACCAGCAAGCAGCAGCATTCCGATTCGGTATTCGGC
AACTGTGGGGAAGGGCGATCGTGTCGAGGGCCTCTTGCTATTACGCAGAGCTGG
CGAAAGGGGGATGGTCTGCAAGGCGATTAAGTTGGTAACGCCAGGGGTTTTC
CCAGTCACGACGTTGTAAACGACCGAGAACGTGAATTGAGCGCTATTTGGGATG
CAATAAGGCTTTACGTACACCATGAGTGGCTCCTCCCCAAAGAGAAGAAAGAGGAT
AGGACACCAAGTACCCCTTATGACGTGCCCGATTACGCTGGGAGCAGGCAGGAGG
CTCCGCGGGAGAGAGGATCCGGAGGAGGCAGCATGTCCACAGCTCTCGCC
CACCCTGGGCGGAAAACTGGCGAGAGATGGCAGATGGGATGGGCTCCGGTGAGATCCC
CAGGAGCTGATCACCAACACTGAGGCAAACCGCCTTCAAGGGAGATGCTTCCG
ACGCCCAATTCAATCGCCCTCCTCATTCTCAGTGGCAGCAACCAGGTACGGCGCTAACCCT
TGGACCAAGGAGATTATCCGCTTCCCCGACAAGCAGAATGCTATGGTCTCCG
TGCTCGCGTGAGCGGCTGGGAGCAAGCAGATCTACAG
CGGCATGACTCGAGACAGCAACCGAGAGCTGCACATGCGGATGGATGAGCTAGA
GAAGGACAGAAACCACATCCTATCGCTCGTGACCAAGTGAGATGAGTGCGA
AGAGAACCCTTCCAGGGAGGCGAGGAGGAGGATCACAGGCGCAAGGGAGGAGG
TCCCCATCTAAGGAATGTCCAGACACAGAAGGCGCATGATCCAGCTCAG
TCGCTTTTGGCTTCGGGCTGCTACATGTAGGAGACTGCTGCAGCAGGATCGT
GGAAAATACCCGGCTACCCCGCTGAAAAGGCAGCCCGAAGGGGACATTACACC
CGTGAAACGAGACGACCCATGCAGGAGATCAATACCTGGCTCATTTGCCCTGGAC
AAAACCTGGGAGATGACCTCCTCCCCCTCCTGCGCAGCCAAATCTTCAGGAGGG
ACATCAGGGCCAGCTCCGAACTCACACAAGGCGAAGCTGTGAAGGCCCTCGG
ATTCCTCAAACAGAAAGGCCGCTAGCAGAAGGTCGCTCAGCATGGTAGTCTCAT
AGCTTTTCTCTGTTGAAATTGTTATTCGCTACAATTCACCACACADACGTA
GCCGGAAGCATAAAGTGTAAAGCCTGGGGCTCTGCTAATGAGTGAGCTAACTCA
CATTAATTCGGTTGCCTCCTCAGTGCCCCTTTCCACTGCGGGAAACCTGTCTGC
CAGCTGATTAATGAATCGGCAACGCCGCGGGAGAGGCCTTTCGCTATTG
GGCGCTCTTCCGCTTCTCGCTCAGTGACTGACTCCTGCAGCTCGGTCTCGGCTG
CCGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAAT
CAGGGGATAACGCAGGAAAGACATGTGAAGCAGAAAAGCCAGGAAAAGGCCA
GGAACCGTAAAAGGGCCGCAGTGAGCGCTTTCCTCCATAGGCTCCGCCCGCT
GACGAGCATCACAAAAATCGACGCTCAAGTGAGGTGGCGAAACC
CGACA
GGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCC
TGTTCCGACCCTGCCCCTTACCGCATACCTGTCCGCCTTTTCCCTCGGAA
GGCTGGCGCTCCCTTACATGCTGAGCTTCTTCTCTGTCAGTCGGTGTCTGC
GTTCGCCTCAAGCTGGGCTGTGCAGCAAGCGAACCCCCCGGTCCGCGACCGCT
GCAGCTCTTATCCGGGTAACTATCGTCTTGGAGTCTCAACCCGGTGAAGAAGACAGCTTA
TCGCCACTTGCCAGCACCACTGGTAACAGAGATTAGCAGAGCGAGGTAGTAG
GCAGCTCAGATCTGCTCTGCTGAGATAAGCCCATGTTACCTTCCGAAAAAGA
AACAGTATTTTGTATCGCTGCTGCTGCGAAGCCAGTTACCTTCCTCGGAAAAAGA
GGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGATTGTTTTT
TGTTTGCAAGCAGCAGATTACGCAGAGAAAAAGGATCTCAAGAAGATCCT
TGGATCTTTTCTACGGGGTCTGACGCTCATGGAACGAAAACCTACGTTAAGG
GATTTTGTCATGAGATTATCAAAGAGATCTTCACCTAGATCCTTTTTAAATT
AAAAATGAAGTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGGTCTGA
CAGTTAGAAAAACTCATCGAGACATCAAATGAAAACTGCAATTTATATTTCA
GGATTATCAAATACCATATTTTTGAAAAAGCCGGTTTTCTGAATGAGAAGAAG
ACTCACCAGAGGCAGTTCCCATAGGATGGCAAGATCTCTGGTATCGTCTGCGAT
TCCGACTCTGCTCAACATCAATACACCTATTTAATATTTTCCCCTCGTCAAAAAATAA
GGTTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAATGG
AAAAAGTTTATGCACTTTTCTTTCCAGACTTTTGAACCGCCAGCCATTACGCT
CGTCATCAAAATCACTCGCATCAACCAAAAACGTTATTTTCTCATCGTATTTGC
TGAGCGAGACGAATAACGCGATCGCTGTAAAAGGACAATTACAAACAGGA
ATCGAATGCAACCGGCAGGAACACTGAGGCAAGTTCCAGGATCAGCA
GTGTTGAGTAAACCATGCATCAGCATCAGGATACGGGATACGCCAATATGCTTCG
GAAGAGGCTACAATTTCTCCGTACGCGAAGTTTACGACTGACATCTCATCTGTAACA
TCATTGGCAACGCTACTTTTGGCCATGTGTTTCAAGAAACACTCGGGCATCGGG
CTTCCCATACAATCGATAGATTGTGTCGACCTGATTGGCCGACATTATCGCGAG
CCCCATTATACCATAAATCGCATCCATTGGGAATTTAATCGCGGCCTA
GAGCAAGACGGTTTCCCGTTGAATATGGCTCATACTCTTTTTCAATATTAT
TGAAGGCTATTATCGGTGGTTATTGTCTCATGAGCGATACATATTTGAAATGTAT
TTAGAAAAATAAACAATAGGGGTTCCGCGCACATTTTCCCCGAAAAGTGCCA
CCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGC
GTATCAGAGGCCCTTTTCGTC

>pUC57-Kan::NLS/Crimson-HumBeta

NLS/Crimson fused to humanized Beta.
AGAGCTGCACATGCAGGATCTACAGGAAGGACAGGAACCATCCTATCTGCGT
CACCGAGTGGATGGATGAGTGCAGAAGAGAAACCCCTCAAGACCAAGAGAGG
CAGAGAGATCACAGGACCTGGCAGTCCCATCCTAAGAGAATGCTCAGACAC
AAGGCCATGATCCAGTGCAGGAGCGTGGGAAAATACCGCCTACACCCGCTGAAAG
GCAGCCCGAAAGGGACATTACACCCGTGAACGACGAGACCATGCAGGAGAT
CAATACCCTGCTCATTGCCCCTGGACAAAAACCTGGGACGTGACCTCCTCCCCC
TCTGCAGCAAATCCTCAGGAGGACATACGGGCCAGCTCAAACCTACACA
AGCCGAAGCTTGAAAGGCGCTCGGTATTCCCTCAAAAGAAGGCGCTGAGCAG
AAGGCCATGATCCAGTGTGCCAGGCCTCGTTTTGCTTCGCCGGCATCTATGA
TAAGGACGAAGCCGAGAGGATCGTGGAAAATACCGCCTACACCCGTGAAAG
GCA
GCCCGAAAGGGACATTACACCCGTGAACGACGAGACCATGCAGGAGAT
CGAGCTTGGCTCGAGCATGGTCATAGCTGTTTCTGTGTGAAATTGTATCCG
CTCACAATTCACACACATACGGACGCGGAAGTAAAGCTGCCGCTCAGTCCTCAGAC
TTCCAGTCGGAAACCTTGCTGTGCAGCTCATTAATGAAACGCCAATCCTGCCC
CGGGGAGAGGGCGTCTTGCGCATGCTGGCGCACGCGTATCGCTACAAAGGG
CGTAAATACGGTATCTCCACAAGATACGGGGATAACGCAAGAAGGAACATGT
GAGCAAAGGCGCAAAGGCGGACAGGACCGTAAGAGGCGCGTTGCTGG
CGTTTTTCCATAGGCTCCCGCCTCCCTGAGCATACAAAGGATACCCCGCAGTTTCCCC
AGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATATTACCCGCG
CTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATAC
CTGTCCGCTTTTCTCCCTCGGAAGCGTGGCGCTTTCTCATAGCTCACGCTG
TAGGTATCTCAGTTCGGTAGGTGTCGTGCTCCTCAAGCTGGGCTGTCGACG
AACCCCCCGTTCAAGCCGACCGCTGCGCCTTATCCGGTAAACTATCGTCTTGGAG
TCCAAACCCGTAAGACACGACTTATCGCCACTGGCAGGAGCCACTTGGTAACA
GGATTAGCAGAGCGAGGTATGTAGAGCGGCTGCTACAGAGTTCTTGAAGTGGTG
GCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGA
AGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCCTTTCGATCCGGCAAACAAAAC
CACCAGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAA
AAAAAGGATCTCAAGAAAGATCCTTTGATCTTTTCTACGGGGGTCTGACGCTCAG
TGGAACGAAAACCTCACGTAAAGGGATTTTGGTCATGAGATTATCAAAAAAGGA
TCTTCACCTAGATCCTTTTTAAATTACAAATGAGTTTTAAATCAAATCTAAAGT
ATATATGAGTAAACTCTGGTCTGAGACTGAGTTAGAAAACCTCATCAGGACATCAAAT
GAAACTGCAATTATTTTATCTATCAGGATTATCAATACTATATTATTGAAAAAGC
CGTTTCTGTAATGAAGAGGAGAAACTCAACCGAGGCAGTTCCATAGGAAGTGCAA
GATCCTGGTATCGGTCTGCGACTTCCGACTCGTCCAAACATCAATAACACCTATT
AATTTCCCCCTCGTCAAAAAATAAGGTATCAAGTGAAATCATCCAGTAAGTGA
CGACTGAATCCGGTGAGAATGGCAGCTGGATTACACCTGCAACTGCAAAACACC
CGTTATTCATTCTGTAGTTGCGCCTGAGGACAGACGAAATACCGCATCGCTTGA
AAAGGCAATTACAAACAGGAAATCGAATGCACAACAGGGGGCAGAACACGCC
AGCGCATCAACAATAATTTCACCTGAATCAGGATATTTCTCTTCTAATACCTGAA
TGCTTTCCCGAGGATCGCAGTGGTGAGTAACCAGTACATCAGGAGTA
CGGATAAAATGGTAGGTCGGAAGGCGTAAACATTCGGTCAGCAGTTTTA
GTCTGACCATCTCAGTCAATCATCTTGGGAACGCTACCTTTTGCCATGTTTC
AGAAACAACCTCGGCGCATCAGGGGCTCCCATCAAAATCGATAGATTGTCGCAC
CTGATTGCCCCGACATTATCGCGAGCCCCATTTTATACCCATATAAATCATCGCTCC
ATGTTGGAATTTAATCAGCGGCTAGAGCAAGACGTTTTCCCCGTGAATATGGCT
CATACGTCTCTTTTTCAATATTATTGAAAGCATTATTCAGGGTTATTGTCTCAT
GAGCGGATACATATTGGAATGTTATTTAGAAAAATAAACAATAGGGATTCCG
CGCACATTCCCCGGAAAAAGTGCACCTGACGTCTAAGAAACCATTATTATCAT
GACATTACCTATAAAAATAGGGCTATCAGCGGCCCCCTTTTCGTCT

> pUC19 ΔKpnI

>GCGCCCAATACGCAAACCACCGCTCTCCCCGGCCGGTTGGCCCCATTATTAATGCG
AGCTGGACACGAGGTTCCTCCGACTGGAAAGCGGGCAGTGAGCGCAACGCA
ATTAATGTGAGTAGCTCAGCTCATTAGGCCACCCAGGCTTTACACTTTATGCT

> pUC19-ΔKpnI

<table>
<thead>
<tr>
<th>EcoRI (396)</th>
<th>HindIII (443)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3500</td>
<td>5600</td>
</tr>
<tr>
<td>2500</td>
<td>4500</td>
</tr>
<tr>
<td>1500</td>
<td>3500</td>
</tr>
<tr>
<td>1000</td>
<td>2500</td>
</tr>
</tbody>
</table>

pUC19(delta-KpnI) 2682 bp
CCTCGGTTATCCCTGATTCTGTGGATAAACCCTATTTTCTAGGTTAATGTCATGATAAT
ATGGTTTTCTTAGAAGTGGCAGGGATCGGGAAATGTCGTCGCGGAAAG
CTATTTTGCTATTCTAAATACATTCAATATGTTATCCGCTCATGAGCAAT
AACCTGATAAAATGCTTCAATAATATGAAAGAAGAGATGAGATAGTATTCA
ACATTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCTCTTCGTTTTT

>pUC19::NLS/HA-HumBeta

AKA pUC19ΔKpnI::NLS/HA-HumBeta. Expresses NLS/HA-HumBeta in bacteria.

>pUC19::NLS/HA-HumBeta

GACGAAAGGGCCTCGTGATACGCTATTTTTATAGGTTAATGTCATGATAATA
ATGGTTTTCTTAGAAGTGGCAGGGATCGGGAAATGTCGTCGCGGAAAG
CTATTTTGCTATTCTAAATACATTCAATATGTTATCCGCTCATGAGCAAT
AACCTGATAAAATGCTTCAATAATATGAAAGAAGAGATGAGATAGTATTCA
ACATTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCTCTTCGTTTTT
GCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTG
CACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAG
TTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGT
TCTGCTAT
GTGGCGCGGTATTATCCCGTATTGACGCCCAGGCGACAGACTGGCTGCCTCGCCCG
CATACACTATTTCAGAATGACTTTGGTGAGTACTCCAGTCACTACAGAAAAG
CATCTTACGGGATGGCAGTACAGTAAAGGAATTATGCAGTGCTGCCATAACCA
TGAGTGATAAACACTGCAGGGCCAACCTACTTCTGCAACAACGATCGGAGGACCGAA
GGGAGCTAACCGCTTTTTTGCAACAACCATGGGAGGATCATGTA
ACTCGCCTTGATC
GTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGAGCGTGACACCA
CGATGCCTGTAGCAATGGCAACAACGGTGGCAAAACTATTAACCTGGGACTC
ACCTACTCTAGCTTCCCGGCAACAATTAATAGACTGGAACAGGCGGATACA
GTTGCAGGACCACCTTTCTGCGCTCGCCCGCCCTTCGGGCTTGGCTTATTGCTGA
TAAATCTGGAGCCCGGTGAGCGTGGGCTCTCAGGTATCATATTGCAGACTGGG
CCAGATGGTAAGCCCTCCGATCGTAGTATCTACACGACGGGGGACCGAG
CAACTATGGGTAGACAAATAGACAGATCGCGACGATAGGTGCCTCACTGAT
TAAGCATCTTGGTAACTGTCAAGCAAAGTTTACTCATATATATACTTTAGATT
TAAAACTTCATTTTTAAAATTTAAAAAGATCTAGGATGAGGATCCTTTTTTT
CTCATGACCGAAATCTCCCTTTAAGCTGAGGGCTTTCTGCCACTGAGCGTCA
CGTAGAAAAGATCAAAGGATCCTTCGCTTGAGATCCTTTTTTTCTGCGCGTAATCT
GCTGCTTGCAAAAAAAAACCAACCGCTACCAAGCGGTGGGTTTGGTGGCGGGA
TCAAGAGCTACCAAACCTCTTTTTCCGAAGGTAACCTGGGCTCAGCAGACGCAG
ATACCAAAATACCTGTTTTCTAGTGTAACCGTGAGTTAGGCCACACTACCTAAAG
CTCTGTAGCGCCGCTACATACCTGGCTGCTGTAATCTCCTGTATTACAGTGCGT
GACAACGAGAGCTGCACATGCAGGATCTACAGGAAGGACAGGAACCATCCT
ATCTGCGTCAACCAGTGAGTGATGGATGAGTGCAGAAGAGAACCCTTCAAGACCA
GAGAGGGCAGAGAGATCACAGGACCCTGGCAGTCCCATCCTAAGAGAATGC
TCAGACACAAGGCCATGATCCAGTGCCAGCTCGCTTTTGGCTTCGCCGG
CATCTATGATAAGGACGAAGCGAGAGATGCAGATGAGTGCAGAAGAGAACCCTTCAAGACCA
GCTGAAAGGCAGCCCAGAAAGGGACATTACACCCGTGAACGAGACGAGACCATG
CAGGAGATCAATAACCCTGCTCATTGCCTGACTCCCACTTGAGAGAACGAGACGAGACCATG
TCCTCCCTCTCTCGAGCAAAATCTTCAGGAGGGACATCAGGGCCACGGCTCCGA
ACTCACACAAGCCGAAGCTGTGAAGGCCCTCGGATTCCTCAAACAGGAAGGCC
GCTGACAGAAGGTCGCCGCCGGTTTAAACGAATTACTGCGCGCTCGTATTAC
AACGTCGTGACTGGGAAAACCCTGGGCTTAATCCGACTCTGGCAGC
ACATCCCCCTTTCGCCAGCTGGGGAATAGCGAAGAGGCCACCGATCGC
CCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGCCTGATGCGGTATTTC
TCTCCTACGCTCTGTGCCTTTACACCCGCATATGGTGATGCGGTATTTC
CAATCTGCTCAGTGCGCATAGTTAAGCACCAGCCCCCGACACCCCGCAACCACCG
CGCTGACGCAGCCCTGACGCTGGGCTTTTACACCCGCATCCCTGCTTACAGACCA
GCTGACCTGCAGCTCCGGAGGACTGACATGTCAGAGGGTTTTCCACCGTCATCACC
GAAACGCGCGA
Plasmids used in construction of pSLIK1

>pFN24K::Crimson/P2A

AKA pFN24K::Product N.

ATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTGGTTAG
TGAACCCTGATCATGACTAGAAGCTTAATACGACTCATAGATAGGTACGATTT
AGGTGACACTATAAGAATAAGGAGAACCATCATCATGTGGCGAGAATTCTGACTGGGC
TTTCATCCGACCCCCATTATGTGGGAAGTCTACGACTGAGCTGAG
CGATATGTGGTCGGCGATGGCACCCTCCTGTGCTTTCTCGAAATCCCGAG
CCTCCTCTGACATGTGTCGATCGCAGATCTAGAAGGAGGTCGGCCACCCCATGGA
TAGCACTGAGAACGTCATCAAGCCCTTCATGCGCTTCAAGGTGACATCGACATGGAG
GGCTCCGTGAACGGCCACCGAGGAGGAGGGCGTGGGCGAGGGCAAG
CTGATGCAATGCGGCGGCTGCATACGCTTGATCCGGCTACCTGCCCATTCGAC
CACCAAGCGAACAATCGCATCGAGCGGACGCACGCAGGATGGAAGGCGGAC
CTTGTGATCAGGATGATCTGGACAGGAAGAGCTCAGGCTCAGGACATCGCAGACAT
AACCTGGGCTACCTGGAAGCTTGCGGAGCTGGGGTGCGGACCGGCTGTCGAGT
GACTCATGGCCAGCTGCTCCAGCATATCATGGATGGAAAATGGCCGCTTTT
CTGGATTCACTGACGCTGGAAGCCGCTGGAAGAGCCATCGAGGACATCGCAGACAT
AGCGTTGGCTTACCGGCTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGAC
CGCTTTCCTGCTGTTTACGGAATTGCAGCAGTTTCCAGCAATACGAGGT
CTATCGCCTTCTTGACGAGTTCTCTCTGAGCAGGGACTCTGGGGTTCGAAATGAC
CGACCAAGCGACGCCCAACCGGTATCAGCTCAGCTCAGCTCAGCTCAGCTCAGCT
TATCCACAAGATACGGGATAACGCAAGGAAAGACATGAGCAGAAAAAGGCC
AGCAAAAAGCCAGGAACCGTAACCAAAATGCACGCTCAAGTGAGGTTGAGG
CGAAACCCGACAGGACTATAAAGATACCGCAGGGCTTTCCCCTGGAAGCTCCC
TCGTGCGCTCTCCTGACCAGCCGCTTACCCGATACCTGTCCCGGCTTTC
TCCCTTCGCCAGGTGCTGTTTCTCATAGCTCAGCTGTAAGGTATCTCAGT
TCGCTGTAAGTTGCTGGCTCCCCGCCGCTTACCACGCTCAGCTGTAAGGTATCTCAGT
TCGCTGTAAGTTGCTGGCTCCCCGCCGCTTACCACGCTCAGCTGTAAGGTATCTCAGT
GCCCCGACCTGCGCCTTATCCCCGTAACTATCGTCTTGTAGCTCCAACCGGTAAG
GACACGACTCTATCGCCACTGGGACGCAGCAGCCACTGGTAACAGGATTAGCAGA
GAGGTATGAGGCGGCTGCTACAGAGTCTTGGAGTGGGTGGGCTCTAATCTAGGC
TACACTAGAAGGACAGTATTTGGTATCTGGCTCTGCTGCTGAAGGCTACACCTT
CGAAAAAGAGTTGGTAGCTTTGATCCGGCAACAAAACCACCAGCCTGGTACG
GGTGGTTTTTTGTTTGCAAGCAGCAGATTACCGCAGAAGAAAAGGATTTC
AAGAAGATCCTTTTGCATCTTTTCTCACGGGCTGCTAGCTGGAACGAAAA
CTCACGTTAAGGGAATTGGTCTGACGATTATCAGACAGGAAACGACCTGAGACCTGAATGCTTTCTTAGTCCGGAAATACAGGAAACGACGCTGGGTGGATGCCCTTCGCTGG
GATGGTAAACCAGAAAATGTCGACGCTTCGATGGATAACGTGAGGTAATGGTAGG
TGGGCTGTACCCTCTGTTGACATTAGGTATTACATTACGTTAATAATTTATCAGGCGGATTTTCGCGCAGTTTTTCGGGTGTTGTGCTGGCCATTTTACCTGTCTGCTGCC
GATGACGCGCTGAACGCGGTTTCTGCTGGTACATTAAGGGATTATTGTA
AATCCACTTTACTGTCTGCCCCTCGTAGCCATCGAGATAAAACCGGAGACTCCGG
CCACGATTCGTCCGCGGTAGAGGATCGAGATCTCT

>pFN22K-Crimson/P2A-ICP8
>pFN22K-Crimson/P2A-ICP8

CCAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCCGTTGACG
CAAATGGGCGGTAGGCCTGTGACGCGTTGGAGGCTATATAAGCAGAGCTGGTT
TAGTGAACCGTCAATAGCTAGAAGCTTAATACGACTCACTAGTAGGGTACGA
TTAGGGTGACACTATAGAATAAGGAATTCATCATGCGCAGAAATCGGTACT
GGCTTTCCATTGCACCCCCCATTATGTTGGAGACTCTGCGGAGGCGCATGACAA
CGTCGATGTTGGGACGCGATGCCACCCCTGTGCTGTTCCCATGACGGTAACC
CGACCTCCTCCCTCCTAGTTGGGCAACATCATCCCGCATGCTGGACCCGACC
CGCTGCAATTGCTCCAGACCTGATCGGTATGGGCAAATCAGACAAACCAGACC
TGGGTATTTTTCTTGGACACCACGCCTCGCCTCATGGATGCTCTTTCATCCAAGCC
CTGGGTCTGGAAAGGACGTCGTCTGGTCATTTCACGACTGGGGCTGGCCTTCT
GTTTCCACTGGGCAAGCAGCAATCCAGAGCAGGCTAAAGCTATTGCATTTAT
GGAGTTTCCATCGCCCTATCCCGACCTGGGACGAATGGCCAGTTCCTCATCA
GAGACCTTCCAGGCGCTTCCGCACCACCGAGTCCGGCGCAAGCCTGATCATCG
ATCAGAACGTTTTTATTCGAGGGTACGCTGCCGGATGAGAGTTGCTCGCCCGCGCTG
ACTGAAGTCGAGATGGACCATTTCCACGAGCCGTCTTCTGAATCCTCTTGACC
GCGAGCCACTGTGGGCCCTTTCCCAACAGCAGCTGAATCCGCGCGGTGAAGCCAGC
GAACATCGTCGGCTGGTGCAAGAAATACATGGAAGTGGCAGCAGGCTCCAG
GTCCGAAGCAGCAGGGGCAACCCCGGCGGTTCATTGATCCCACCGCGCAAGCC
GGGGTCTGAATCTGCTGC
AAGAAGACAACCCGGACCTG
ATCGGCAGCGAG
ATCGCGCGCTGGCTGTCGAC
GCTCGAGATTTCCGGCGAGC
CAACCACTGAGG
ATCTGTACTTTCAGAGCGATA
ACGCGATCGCAGATCTAGGA
GGTCCGCCACC
ATGGATAGCAGTGAACGGTCATCAAGCCCTTCATGCGTTTCAAGGTGCA
TGGAGGGCTCGTGAACGGCTGGAGACTGATCGAGGCGGTGCGAGGG
GCAAGCCCTACGGAGGGACACCCAGACCCGGACATGCAAGTGACAAAGGG
GCCCTGGCCCTTCGCGATCTCCTGGAGACATCTCCTGCTCCCCCAGTTCTTCTATCGGCTCC
AAGGCGTACATCAAGCACCAGCCGACATCCCGGACTACCTCAAGGACTGCT
TCCCGAGGGGTTCAAGTGGGAGCGGCGTGATGAACTTCGAGGACGGCGCGGT
GGTGACCGTGACCCAGGACTCCTCCTCGAGGACGCACCCCTCATCTACAC
GTGAAAGTTCTCGGCGTGAATTCTCCCCTCCGACGGGCCCCCTGTAATGCAGAAGA
AGACTCTGGGCTGGGAGCGCTCCACTGAGCGCAACTACCCCGCGACAGGCGT
GCTGAAAGGGCGAAGACACATGGCGCTGAAAGCTGAAAGGGCGGCGGCGACT
CCTGTGACTGTTCAAGTCTCATACATTGCGCAAGAGCCCCTGGAAGCTGCCC
GGCTACCACACGTGAGACTAAGCTCGACATCACCTCCACACAGCGGAGGACT
ACACCGTGGTGGAGCAGTCGAGCGCGCCGAGCCCGCGACCACTGTTCCA
GGATCCGCCAGCCACGTAATTCTCTCTGTTAAAGCAAGCAGGAGACGTGGA
GAAAAACCCCGGTCCTCCATCTGCTGCTGATCGGCCATGGAGACAAAGCCCCAAGA
CGGCAACCACCATCAGGTCCTCCCCCGGCGGCTGGGATACGTGTACGCTCG
CGCTGTCCCGTCCGAAGGCTCGAGCTTCTGGCGTTACTGTCGGCGCGCGACG
GGCGATGCCCGACGTCGCCCGTGCGGCTGCCCTGGTCTGCTGGCCTGACCCTGGA
GCGGCTTTTGAGGCACACGTAGCCTGTGGTCTGCTGGGCTGCTGACGCTG
CGGGGGTGATCCGCGGTGTCCCTGAAAACTGACGCATCCGACACGCTCG
GTGTACGTCTTTTCACGGCGGCACCGCAGCTGGGACCCAGCACCAGGCCCA
ACCTGACGCGACTTCGAGAACGCGACACGACGGGAGGCGCTGTTG
GAGCGCCTCGGCTGAGCCCGAGCGCCGCTGATCTGATCGTCTGTAACCG
AGGGCTTCAAGGAGGCGGGTGGCATTCAACAAACACTTTTCTGACCTGGGGAGG
CTCGGACAGGTAACCATAAGGCAGGTCGCTTGGAGCTACCCCATACCCGCTGTAT
CCGGTGAGCGCTGCATGGCCGAATTTATAGCCGGGTCATCGCCCGAGGCGGTCA
CCTGAACCACCGATCGATCGGGAAGAGAATTACCTACCCGCTTCCGTTTTTFA
ACCAGCCCCCTCACCAGCCTCTGTTCGAGGGCGTGCTGGAACCGGCCGCCGCT
GGCAGCTGCGATGCCGAAACGATGGACGCGCTGGCCCGCGCGGCCGCCCACCTG
CGTTTAACGAAAACAGGAGGCGCCGCCGCTCCCTCCCCGCCGACATTACGTCTCAGC
CGGCCTTGAAACGACGCGCTGCAGTCTATCTGTGCATGGCCGCTTTCGAC
GAGCCGCCCAACGACATCTCGCCGCTGCGCGCTGTGCAGGAGGCCAGGACACGG
CCGCGGCGCGCAAGCCTCGGGGCGCTACCTGCGCGCGCCGCCGCGGACT
CGTGGGAGGCGCATGAGTATATAGCACAACACTCGGCGCTTCATCTCGACAGGTG
GACGACGCGGCTCGGCGGACCCAAAGGACCACAGCAACACCTCCTCCTTTTACC
GCTTCTCTCCTGCTGCCGCGGACCCACGTCGTCGGGGGCTACCGCGCGCGCGCCG
CGAGGGGACACGTGTTGCGCGCGGCTTACGGGCTGCGCCACCCCGCCTCCGTC
GGCGGAACCCAGGAATTTTGCCGCGGAGCAGCTTGGCCCAGTCGTGTGTTGGTTT
CCCGCGCCTGCTGCAAGATGCTTATTTACTGAGCGCTGCGACGGCGCGG
CGTGATCGTGGGCACAGAGTAGGCGCTTTCGATACGCTGCGCGACTCC
AACCAGACCGACGCTGCCCTGGCAAACCTGTCACCTTCGACACCGCCACCGCT
GCCTACACACGACGCTCAGCTCCGCGCGAGCGGCGCTCCGAATCCGACCG
CGCGCGCGCGGAGAGCATCGGGGTGTCCGGGACCAGCATGAACACGCTACAGC
GACTGCGACGTGCTGGGAAAATACGCGCCGCTTCTCGGCCCTGAAGCAGCGCG
ACGGATCCGAGACCGCCCGGACCATCATGCGAGAGCTACCCGCAGCGCG
CCGAGCGGCTCATGGCCGAACTCGAGACCCCTGCAGTCAGTCTGGAGACGGAT
CCCCACGGCCATGGGGCGGGTGGAGACCATCATCACCACCCCGAGAGGCGCCC
CATACG GGTTGAACACGTCAGGCGAGGTGTGACGGACCCGAGGTGGAGACAG
CTGATGCACAACCTGTTGGAGGGAGGAATCTTCAAGTTTCGCGACGCTCTGG
GCGAGGCCAACCACGCAATGCTCCTGACGCTGGAGACCCGTACGCGTCGAGGG
ATGCCCCCTTGGTTCAGCTTTCGCGGCGGATCCAAACCCTCGGCGTGTATCGG
ACCTGGCCCCGAGCCAGTGCAGCGGGTCATCAGGCGGTCAGTCGAGGG
GCGCAACTTTCGCAATACTTCCAAACCGGCGTCTGCAGCGCGCCTGC
AGGGGCGGCTATCTGCGCCCTAGCCTAACGCGGCGCAGCGGCACCACCACC
CGAGACGAGCTTTCCGAGGGCGACGTTGGCCTCCCGCGTGACCCCTGGGCTTT
CACAAGAGCTGCGCTGGAGATAAACTTTAATTAACCTGGGCCCCCAACAAGCTG
GTTCACGCACGGCATTCTCCCGAGTGCTGGGCTTTCTGTGCTGAAAGCA
ACCCGCAGATGGTCTGGACCGCCTCCAACGCAACAGCTTCTCCCGC
TCCTGTCCCGCCAGGACATCGAGGATCGCTCATTAAAAATGTCTCCCTG
GACTACGCGCGATATAAATCTTAATACCTGGCCCGCCCAACACGCTGAGCG
TGCGATGTACTCATAGGCAAAACAGATTTCTGCGGACTGCGATCATCGAC
ATACTTCATCAACACCTCAACCGCCATCATCGCGGGTTCCCGCCCGTTCCCTG
GCGTGCAGGCAGCAGCGCCGCGCTGCTGGCCAGGGGGGGGGGCGGCTGGAGG
CCGGGGCCCGCGCGCTGATGACGCCTGGACGCATCCGGGCGCGCTGGGAC
GTCCATGGTTCGCGGAGCTGAACCTGTGCTGCGGCGCCCTCATGGCGGCGGCGCCCC
ATGGTCGTTGTTGGTTTGAGCATCAGCAAATACTACGGCAGGCGGCGCAACCG
ACCCTGGTGTACCAGGGCGGAACCTGGGCCAGCCTGATGGGCCGGCCAGCGGCAAGC
CTCCCGCTCTCTTATTTTGACCGCAACCGCAAGTTCTCGTCTGCGGCTGTCCCC
GGGCGGGGTGTTGTGCGCCGGCCTCAGAACCCTCGGGCGCGAGCGACCGACGAAAAG
CTCGCTGTGCGAGCAGCTCCGGGGCATTATCTCCGAGGGCGGGCGGGCGGCGTC
GCCAGTACGCTGTTCTGCGGAGCCCTGAGAAAGCTGGGGCCGCCCGCAACCCAGC
AGCTGCAGATCGAGGACTGGCTGGCGGTTCAGGAGGAGCGAGTAGCTAAAGCGA
GGAGATGATGGAGCTGACGCGCCGGCGCTGCGCCTGGAGCGGGCGGGCGAGTG
GTCGACGGACGCGCGTGCCCTGAGGGGTGAGGACGAGCTTGAACTCAGAAGCGA
CCACCTCGGCAACCGCGGGGAGGTGTTAACTTTGAGGGATTTTGGCTGCGAG
GACGACACCGCGCGCCGCTCCTGGCGCGGGCCGGGGCCGGCCGCGCCGCTTCTT
CCGCGGCAAACCGGGCAGGATGACCCCGTTGGGAGGGGAGGGGCCGCCGCC
CGACAAAAAGGGAGACCTGAGCGTGGGATATGCTGTTATAAACGAAATTCCGGGC
TCGTTACCGGGGATCCTCTAGACTGACCTGCAGGCAGCTGCAAGCTGATCCTCG
GCTGCTAAACACAGCCGGAAAGGGCTGAGGTGGCTGCTGCTGCTGCTGACCCGCTGAGC
AATAACTAGCATAAACCCTTGGGCGGGCGGCGGCTTGAGCAGCAGCATGAAGAT
ACATTGATGATGGTTGGACAAACCAACTAGAATGCGAGTTGAGTGGCTGCTGCTGCA
TATTTGGGATATGGGATGCTATTGCTTTATTGTAACCTATATAAGCTGCAA
TAAAACAGTTAAAACAAACTGATTCATTTATGTATTTTGCTCCAGGTTCCAGGGG
AGATGTGGGAGGTTTTTTAAAGCCAGTAAACCTCTCAAAATGTGGTAAAAAT
CGAATTCTAATGGGATCTCTTTTGCGGTTCCTTCCCTTGTCCAGATAGCCCA
GTAGCTGACATTCACTCCGGGGTCAAGCACCAGTTCTGCGGACTGCTTTCTACG
TGTTCCGCTTCTTTAGCAGGCTTTTCTGCGGCTGAGTGTGCCTTGCAGGCAAGCTGAG
CTTCAAAAGAATTGCCAGCTGGGGGCAGCCCTCTGGTGTAAGGTGTTGGAGAAGCGGTCC
CAAAAGTAAACTGGAATTGCTTTTCTTGGCCGGCAAAGGATGCTGATGGCAGGCAAGGGA
TCAAGATCTGATAAGAGACAGGATGACGGTCTTTCGCACTGCTTGAACAAAG
ATGGATTCACGCAGATTTCTCCGCGGCGGCTTTGGGTGGAGAGGTATTTCCGCTACT
TGACTGGGCACACAAACAGAAATGGCTGCTGCTGGCGGCTGTCGCTGCTTCCGGCTG
TCAGCGCAGGGGCAGCCGGTCTTTTTGTTCAAGACCGACCTGTCCGGTGCCT
GAATGAAGTCAGGAGGACAGGACAGCCGCGCTACTGGCTGCTGGCGCCACAGCGGG
CGTTCCCTTGGCCAGCCTGCTCCTGCTCCTGCTGCTGCTGCTTGCCCCCT
CTGCTATTGGGCGAAGTGGCCGGGCGGAGGATCTCTCTGCTCATCTCACCTTGGCCC
TGCCGAGAAATGATCCATCATGGCGCTGATGGCAATGCGGCGGCTGACCATCGCT
GATCCGGCTACCTGCCCATTCGGACCAACAGGCAAAACTCGACATCGAGCGAG
CACGCACCTCGGATGGAAGCCGGTCTTGTCGATCAGGATGATCTGGGCGAAGA
GATCAAGGGGTCTCAGCCAGCAGCGAACTGTTTCCGCCAGGCAGCTCAAGGGCGCTATG
CCGGATGTTGAGATCTCGTCGGACTCATGGCGATGCTGGCGCTGGGCCAATAT
CATGGTGAAAATGCGCCGCTTTTTCTGGATTTCCGACTGTGGCCGGCTGGGTG
TGGGCGGACCCTAATACGAGCTGATTGCTACCCGTGATGATTGCTGGAAGA
GCTTGGCGGGAATGGGCTGACCCGCTTTCTCCTGCTTACGGTGATCGCCGCTC
CCGATTCCAGCCGATCGCCTCTCCTACGACGATTCTTTTCGAGCG
GGACTCTGCGGGGCTGAAATGGCAAGGACAAGCGGACGCCAACCAGGTATACGCT
CAGCTAAAAGGCGGTAATACGGTTATCCACAGAATCGGAGGGAACAGCGAGGA
AAGAACATGTAAGCAGAAAGGCCAGCAGAAAGGCCAGGAACCGTAAAAAGGCCC
>pENTR/TREPitt::Crimson/P2A-ICP8

CTTCCCTGCGTTATCCCCCTGATTCTGTGGATAACCCTATTACCCTAGCATGG
ATCTCGGGGACGTCTAACTACTAAAGCGAGAGTAGGGAACCTGCAAGCATCAA
ATAAAAACGAAGAGCTCAGTCGGAAAGACTGGGCTTCCCATTTTATCTGTTT
GTGGAGTGAACCGCTCCTCCTAGTGGACAAAATCCGCCGGAGACGGAATT
CTTACAAAATCTCCTTGTTAGTTTACTTAAGCTGGCCGGCCCACAATTTG
ATTTTAACGTGATAGGTGGGCAACAAAATTGATAAGCAATG
CTTTTTATAATGCAACACTTTTGACCAAAAAAGCAGGCTGCGCCGAACCAA
TTCAGTATCTGCGAGCTCTCAGTGCAACAAATTTGAAGAT

> pENTR/TREPitt::Crimson/P2A-ICP8

CTTCCCTGCGTTATCCCCCTGATTCTGTGGATAACCCTATTACCCTAGCATGG
ATCTCGGGGACGTCTAACTACTAAAGCGAGAGTAGGGAACCTGCAAGCATCAA
ATAAAAACGAAGAGCTCAGTCGGAAAGACTGGGCTTCCCATTTTATCTGTTT
GTGGAGTGAACCGCTCCTCCTAGTGGACAAAATCCGCCGGAGACGGAATT
CTTACAAAATCTCCTTGTTAGTTTACTTAAGCTGGCCGGCCCACAATTTG
ATTTTAACGTGATAGGTGGGCAACAAAATTGATAAGCAATG
CTTTTTATAATGCAACACTTTTGACCAAAAAAGCAGGCTGCGCCGAACCAA
TTCAGTATCTGCGAGCTCTCAGTGCAACAAATTTGAAGAT
ACGTGTACGCTCGCGCTCGTGTCGCTCCGAAGGCATCGAGCTTTCTGGCCTACTG
TCGGCGCGCAGCGGGCAGATGCCGACGTGCCCGTGGC GCCCCTTGGTGCTGGTTGCTCG
TGACCGTGAGAGCCGGCTTTGAGCCGACCTGGAGCCACCATCGCAC
TACAGCTCGTCCGCTGTACTCTTTTCAGGCGGCGGACCTGGAGCCACCATCGCAC
CCCAGGCCCAACCTGAAGCGGACTCTGCGAGCGGGCACGCGGGCCATTTTGG
CTTTTCGGACTACACCCCCGGCCCGGCAGCGGACCTCAAAACAGAGACGACGGGG
GAGGCGCTGTGTAAGCGCCTCGGCCCTGGAGCCCCGACGCGGCCTCGGCTATAC
TGTCGTTACCGAGGCTTTCAAGGAGGCGGTGATCATCAAAACACACTTTTCT
GACCTGGAGGCTCGGACAAAGTGTAACCATAGCCGGGGCGGAGGTGACGGC
CATACCGGTATCCCGTGCAGCTGTTCATGCGGGATTTTAGCCGGGTATCG
CCGAGCCGTTCAACGCAGGCCACCGATCGATCGGGAGAATTGACCTAACC
GCTTCGGTTTTTAACCGCCCCCTCAACCGCCTCCTGTGTCGGGACTCGGCTTGG
GACCCGCGCGCTGGCAGCTCGATGGCCGAAACGTGGAGCGGGCGTGGCGC
GGCCGCCCACCTGGGGTGGAGCCAAACACAGAGGGCGGCCGGCCTCCCGGC
GACATTACGTTGCGCGCCTTCGAGGCCAGAACGGGGTTAGAAGACCCCGCCGGGTG
GGCGCGACGGCGCGGCAGGCCCGGCGGCGGCGGCGGAGGTCGACTTCTGGGGC
CCTCCGTCATGGCAGACGGCGCCCTGGCCCTCAGTCTATCGTGTCGATG
GCCGTCTTCGACGAGCCGCCACCTGGGCTTGCGTCGCTGTGAGCTGCGTGG
GCCAGACACGGCCCGCGCCCGGCGCCAACCGCGTCCGGGCGGCGCTTCGGGC
GCCAGGACAGGCACGCGGCCCACCGGAGGCGACATCTCCGCTGGGCGCTGTGCGAGG
GCCAGGACACGGCCCGCGCCCGCCCGCAACCGCGTCCGGGCGGCGCTTCGGGC
GCCAGGACAGGCTGCTGGGGGCATTGATTTTAGACCAAACCTGCGCCTCAA
TCTCACCAGGGTGACGGCGCCGTCGGGCGGAGCCAACGAGACCACAGCAG
ACCCCTCTTTTACCGCTTTCTTCGCTGCGCCGGGACCCACGTTGCGGCGCCACC
CACAGGTGGACCCGCAGGGACACGTGGTGGCCGCCGGTTTCAGGGGTCGGCCCA
CGCGCCCCCTCGTCGCCGGAAACCAGGAATTTGCGGCCGAGACCTTGCGCCAT
GCTGTGTTGGATTCTCCCCCGCGCCTGCTGATGAACTGCTGGTGGCTTCGAGCG
CGTCGCGAGCGGCGACGGGATTCGAGGATGGACGCTGGCTTTTCGATA
CGTCGCGAGCGGCGACGGGATTCGAGGATGGACGCTGGCTTTTCGATA
ACGCGCCACGCCTGCTCAGACACAGACGCTGCTGGAACCTGGCAGCTACAGC
CCAAGTTCCAGCCAGCCGCCAGGGGAGCCATCGGCGTCTTCGGGACCATGAA
CAGCAATGTACAGCGACTGCGACGTGGGAAACTACGCCGCCTTCTCAGGCC
CTGAAAGCGCGCGGAGGGACCCGAGGACCGCCGCGGACCATTGCAGGAGACG
TACCCCGCGGCGGACCCGAGGCGGTCTGCAGAACCCTGCGAGTGACGTACG
TGGACCAGGCAGTTCCTACCCACCCATGGGGCGATGAGGACCATCATCACCAC
CCGCGGAGCCCTTCATACGGGTGAACACGTCAGGCGAGTGCTGGACCGCG
GAGGTGGAGCGAGCCTGATGCGCAACCCTGGTGTGAGGGGAGGAACTTTCAAGTTTC
GCGACGCGTCTGGCGAGGGCGGCAACCACCGCATGCGGCTACGCTGGACCCGTA
CGCGTGCAGGCGGCATGCGCCAAGCTGCGGCGGATCCACCTCGC
CCGTTATACGGACCTGGCCCTGAGCCAGTGGCGCCAGGGGTTGCAATGCAAT
GTCGGTGCAGGGCGCAACTTCTCGCAATACCAATCTCCACGGGTGCTGCGCCGG
CGCTGTGATGGACATGTATTTAAACACGGGGTTTCTGCGGCCAAAACGTGACGG
TGGCCGCTCTCGGAGGGCGGCTATCTGCGCCCCAGCCTAACGGCCGGCA
GACGGCCCGCCGCCAGCCGACGTTCGCCAGGGCGGACGTGTTCCCGGCCGTGACCTG
GGGTTTCCAAAGGAGCTGCGCTCAAGAGCAGGGCGGCTGTGGTTCCCGGCCGGCA
GCAGCAACGCCTCAGGCGCCAGCAAGGCCGCGGCTGCGCCAGCCTGCAAGGCG
CCTACCAGAAGCCCGACACAGCGCTGAGCATCCTCCTGGGACCGGTGGCCTT
TAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCCTTCTTAGTAGGACG
TAGTTAGGCCACCACACTTCAAGAAACTCTGTAGCACCAGCCTACACATTACGCTCT
GCTAATCCTGTTACCAGTGTGGCGTGGCTGCCAGTGGGCGATAGTGTTGTTACC
GGTGGGACTCAAGACGATAGTTACCGGATAAAGGCAGCGCGTTCTGGCTGAAC
GGGGGGTTTCGTGCACACAGCCCAGCTTTGGAGCGAACGACCTACACCCGAAC
AGATACCTACAGCGTGAGCTATGAGAAAGCCACCGCTTTCCCGAAAGGGAGA
AAGGCGGACAGGTATCCGGTAAGCGGCAAGGCCTACACAGCCGCGCACAG
AGGGAGCTTCCAGGGGAAACGCCTGTATCTTTATAGTCCTGTGGGTTTCG
CCACCTCTGACTTGAGCGTCGATTTTGATGCTCGTCAGGGGGCCGGAGCC
TATGGAAGACCCAGCAACGCGGCCTTTTTCGTTTCTTGCCTTGTGCTG
CCTTTTGCTCACATGT
>pSLIK/TREPitt::Crimson/P2A-ICP8

AKA 1.pSLIK-Zeo/TREPitt::Crimson/P2A-ICP8 or pSLIK1. Sequence confirmed from 2607-7842 bp.

>1.pSLIK-Zeo/TREPitt::Crimson/P2A-ICP8

GTCGACGGATCGGGAGATCTCCCGATCCCCCTATGGTGCACTCTCA GTACAAT
CTGCTCTGATGCGCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGT GGG
AGGTGCCTGACGATTGGCCGAGCAGAAATTTAAAGCTACAAAAAGGCAAGGCTT
GACCGACAATTTGCATGAAGAATCTGCTTAGGGTTAGG
CGTTTTGCGCTGCTTC
GCGATGTACGGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTAAT
TAGTAATCAATTACGGGGTTACATTAGTTCATAAGCCCATATATAGGAGGTTCGCC
TAGTACCGTAAATGGCCCCGCTGGGTAGCCGCCCCAACGACCCCACG
CCATTGACGTCAAATAATGACGTATGTTCCCATAGTAACGCAATAGGGACTTT
ATCAGTGATAGAGAAGGTACGTCTAGAACGTCTCCCTATCAGTGATAGAGAA
GTCGACACGTCTCGAGCTCCCTATACGTGATAGAGAGGATACGTCTAGAACG
TCCCTATCAGTGATAGAGAAGGTACCCCCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCG
CCTGGAGACGCCATCCACGCTGTTTTGACCTCCGTAAGAAGACACCGGAGCCG
ATCCAGCCTGGATCTAGGAGGTCGAGCTCCGCCACCATTGGATAGCAGCTGAACG
TCAAGCCCTTCATGCCTTCAAGGTTGACACATGGAGGGGCTCGTTGAGACGGC
CGAGTTGCAGATCGAGGGGCAGGGCGAGGAGGCAAGCCCTACGAGGCGACCCCA
GACCGCAAGCTGCAAGTGACCAAGGGCCGCCCCCCTGCCTCTCCGGCTGAGAC
ATCCTGTCCCCCAGTTCTTTTCTACGGCTCCACAGGCGTACATCAAGGACCCGCCG
CGACATCCCCGACTACTCAGTCAGTTCCCCGAGGGCTTTCAAGTGAGGAG
CGCGTGATGAACTTCGAGGACGGCGTGGTGAGCGTGACCCGACTCCCT
CCCTGCAGAGCGAGGCACCCTCTCATCTACACGTAAGTTATCGGCGTGAACTT
CCCCTCCGATCGGCCGCCGCCTGCTGAGGGCGGAGCCCTCC
ACTGAGCGCAACTACCCCGCGACCGGCGGCTGCTGAGGCGAGGACACCATG
GCGCTGAAAGGTAGGGCCGCGCCACTACCTGTGTGAAGTTAAAGTTCATCT
ACATTGCGAAGAGCCGCTTGGCCGTACACGCTTGGAGACCTACGAG
CGCGCCGAGGGCCGCCACCACCTGTCTCCAGGGATCCCGGAGCCACGAACTTCT
CTCTGTAAAAGCAAGCAGAGGAGAGCGTGAGAAGAGAAACCCCCGTCCTACATGTC
GCTGACACATCCTCCTCCCCAACAGGGGACTACCCCGTGGTAGGAGCACTACGAG
CGCGCCGAGGGGCGCCACCACCTGCTCTCCAGGGATCCCGGAGCCACGAACTTCT
CTCTGTTAAGCAAGCAGAGGAGAGCGTGAGAAGAGAAACCCCCGTCCTACATGTC
TGGAGACCCGGAAGCAGACAGCCAGGCAACCACCACCAGTAAGGTCCCC
CCCAGGCCCTTGAGATACGTGACGCTCGCGCGCTTGCCGATCCGGAAGGCACTG
AGCTTTGCGTACCTGTCCGCGCGAGCAGCCGCGATGCGCCGACGTCGGCCTGCGC

TCCAACGCAACCAGCTTCCGGCCCGGCTCTCTGTGCGGCGAGGAGGAGATCGAGAC
CATCGCGTTTATTTTCTGAGACTACCGGCGGATAATAACTTTATTACCTGAGAGAGAC
ACCTGGGCGGCACAACTGAGCGAGCGGAGGGCGGGCGGTGAGTGCAGGACGTCGCT
TCATCGCGGCTCCCAGGTCCCCCGAGCTGAGCAGGGGAGGCGGCGGCGGGTGTC
CGCGCAGGGCGGGGCGGCTGAGCCCTGAGGCGGCAGGGCGGCGGCGCTGAGGACG
CGTGACGCGATCCGGCGCTGAGGGAGTCTTGCAATGTTGGGTTGAGGCATCA
GAAATATCAGGCGCATGGCGCGACGGCGACGGCGACGGGTGACCTGAGCTGAACTG
GGCCAGCTGAGAGCAGACGCTGGCTGC
TTTCAGGCCGGAACTG
GGCCAGCCTGAGGGCAGGCGGTTTCTATCTTTGACCGCATCGGGAGCTGCTG
CGCGGATGCAGATCTGGAATTCCCATGAAAAGGGAGACCTGAC
GGTGGATATGCTGGTTAAGGATACCTGAGCAGGAACAGCAGCAGCAGCTTCG
GCGGGGCTGCGGCGGCGGCGGAGGAGGAGGAGGAGGAGGCAGGAGGAGGAGGAGG
CGGGGGCTGCGGCGGCGGCGGAGGAGGAGGAGGAGGAGGCAGGAGGAGGAGGAGG
CGGGGCTGCGGCGGCGGCGGAGGAGGAGGAGGAGGAGGCAGGAGGAGGAGGAGG
CGGGGGCTGCGGCGGCGGCGGAGGAGGAGGAGGAGGAGGCAGGAGGAGGAGGAGG
CTCCACCACCGTACCGTCTGCAGTCAATTCCACCACATTGCTAGACTGGACAA
GAGCAAAAGTCAATAAAGCGAGCTCTGGAATTACTCAATGGTGTGCTTGATCGAA
GGCCTGACGAACAGGAAAATCTCGCTAAAAAGCTGGGAGTTGAGCAGCAGCTACCC
TGATGGTGGACAGGCATGACTGCCATACCACTTCTGGGCCCTGGAAAGGCGAGTCATGG
CAAGACTTTGCTGGAACAAACGGGAAATGCAATACCGGTGCTGTGCTTCTCCTCTACA
TCGCGACGGGGCTAAATGAGCTCTCAGCCACCCGGCCCAACAGAGAAAACAGTAC
GAAAACCTGGAATATCAGCCTGGCTTTCTGTGTGTCAGCAAGAGCTTCTCCCTAG
AGAAGCCACTGTACGCTCTGTCGCCCGCTGGCCACTTTACACTGGGCTGCTG
ATTGGAGGAACAGGAGACATCAAGTAAGCAAAGAGGAAAAGAGAGACACCTAC
CACCAGTTCTATGGCCCTACCTCTGTAGAAGCAATGGGAGCTGTGCTGACCGGC
AGGGAGCCGAACCTGCTCTTCCTTTCGGGCTGGAACTAATCATATGTGGGCTG
GAGAAACAGCTAAAGTGCGAAAGCGGCGGGCCGACCGGCCTTGACGATTTTGACTT
GAGAAGCTTTAGACATGCTCCAGCGATGCCCCCTGAGCACCTTTGACCTTGATAT
GCTGCGCTGCTGAGCTCTTCGTACGATTTTGACCTTGACTTTGACCTTGATAT
TAAGTAAGGATCCGCACGCGAATCTAGAGGAATTCGGCCCTCTCCTCTCCCTCCCT
CCCCCTAACCTACTGGCGAAGCCGCTTGGGAATAAGGCGGCTTGTTGTTGGT
CTATATGTTATTTTCCACCATATTTGGCGTCTTTTGGAATGGTGAAGGCACCGGA
AACCTGGCCCTGTCTTTTGACGAGCATTCATAGGGTCTTTCCCTCCCTCCTGC
AAAGGAATGCAAGGTGTGCTGCTTGGGAATGGTGAAGGAGCAAAGACCTTCTGGAAG
CTTCTTGAGAACAAACAGCTCTGTAAGCGACCCTTTCGAGGAGCGAGCAGCCACCC
CCCAGCTGACGAGGTGCTCTGCTGCGCCCTGGCAAAAGGCCACGCGTGATAAGATACA
CCTGACAAAGGGCAGCACACACCCAGTGGCGCTTGGGATAGTTGCGG
AAAGAGTCAAATGGCTCTCCTCCTCAAGCGTAGTCAACAAGGGGCTGAAGGATGC
CCAGAAGGTACCCCATTGTATGGGAATCTGATCTGGGGCCTCGGTGCACATG
CTTTACATGTGTCTTAGTGAGGTTTCTTCTAAGCAGTATAAGCTTTACCGGTACGCGT
GATGTGTTGACAACATATCCTGGCATTGATATATCGGCATAGTATAAACTACGA
CAAGGTAAGGAACAGCTAGGCAAGTTGACCAGTGGCTCCGCGTCTGGGCTC
ACCGCGCGAGCCTGCGCGGAGCGCGTCTGAGTGCTTCTGGACCGACCGGCTCCTGGGT
TCTCCCCGGAACCTCCTGCTGGAGACAGCAGGTCGTCGCGGTCTGCTGGTCGAGCT
GACCCTGTTCATCAGCGCGGTCCAGACCAGTTGCTGGCGGACACACACCTCTG
GCCTGGGGTGTTGCGCGCCTGGAGACGCTCTCCGGTGTTGCTGTCGCGACGAGT
TCGTGTCCAGAACTTCCGGAGCCTCTCGGGCCGGCCCGGACCAGATGGAGT
CGAGCAGCCGTGGGCGCGAGTTCCCGCCTGCCGAGCCCGCGGCACTGC
GTCGACTTCGTTGCGCCAGGAGACCAGACTGACACATCTGTACACATCTACGTC
CCGCGACTCTAGATCATATACTCCGATAGGTCGTTTACTTTACATCTG
TTAAAAAAAACTCCTCCACACCTCCCCCTGAAACCTGAAACATAAAATGAATGCA
ATTGTTGGTTTATAGCCCTCCTCCAAATTCTGATATCAAAGCTTATCTGATAATCAAC
CTCTGGATTACAAAAATTGTAAGGATTGACTGTTATTCTTAACTATGTGTCT
CCTTTTGACTATGTGGAACACGTGCTTTTATGCCCTTTGTATCATGTATATTGCT
TCCGCTATGGGTTTCTATCTTTCTCTATTAATGTTCCTTTGTATATGCTATATTGCT
ATGGACCGATTTTGGGCCGTTTGTGACGCAACGTGGGCGTTGCTGACTGTTT
GCTGACGCACCCCACTGGTTGGGCATTGCCACCACCTGTCAGCTCCCTTC
CGGCACTTTCGCTTTCCCCCCTCTCTATTGCCACGGGAACTCATCAGCGGCT
GCCTTGCCCGCTGCTGGACAGGGGCTCAGTCTGGTGGGCACAGAAATCCCGT
GGTTTGGTTCGGGGGAAAATCATCGTCTCTTTCTGGCTGCTCTGCTTGGCTCCGA
CTGGATATTCTCGCGCGGACGCTCTTCTGTACGTCCCTTCCGCCCCTCAATCCA
CGGACCTTCCCTCCCCCGGCGCCTGCTGCTGGCTCTGGCCCTTCTCCGCCTCT
TCGCCTTCGCCTACGACGAGTGCGATCTCCTCTGCTAGAAGACAAAATCGAGCAATCCAAATGA
GGAGGAGGTTGATGTTCCAGTCACACCTCAGGTAACCTTTAAGGACAAATGACT
GAAGACGCTGATTTTAGCGCACTCTTTTCTCAACACAGATCTAGTACCAGTTGAGC
AAGAGAAGGTAGAAGAGCAATGAAGGAGAGAACCCGCTTGTTAACCC
CTGTAGATATCCACTGACCTTTGGATGGTGCTACAAGCTAGTACCAGTTGAGC
GGTTTGACAGCCGCCTAGCATTTCATCGACATGCGGCCAGAGCTGACTCGG
CTGTACTGCTCTCTGTGGAAGATGCTGCTCGTTGCAGTCTCTCGTCTGGCT
AACAGCAACCCACTGCTTAAGCGCTCAATAAACGTTGCCTTGAGTGCTTTCA
AGTAGGTGTGCGCCCTGCTGTTGTGTGACTCTGGTAACTAGAGATCCACACACAGCAC
CCTTTTAGTGACTGTGGAGAAAATCCTCTAGCAAGGCCGTTAACCCCGCTGTATC
AGCCTCAGCTGCGCTCTAGTTGCAAGGAGCCCTGTTAACCCCGCTGTAC
GCCCTTCTGGAGGGCTGGCCAACCTCCCCACTGCTCTCTTCTCAATAAAATAG
AGGAAATTGATCGACATGCTTGAGTAGTGCTCATCTATCTTCTGGGGGTGGG
GTGGGGCAGGACAGCAAGGGGAGGGTAGTGGAAGAAGACATAGCAGGACATGCT
GGGGATGCGTGGGGCTCTATGCGTCTCGAGGCGAAAGAAACAGCTGGGGCT
AGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAA
AAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGG
TTTTTTTGTTGGCAAGCAGCAGATTACGCGCAGAAAAAGGATCTCAAGAA
GATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAAACTCAG
TTAAGGGGATTITGTCATGAGATTATCAAAAGATATATGAGTAAAACCTTG
GTCTGACAGTTAACAATGCTTAACTCAGTGAGGCACCTATCTCAGCGATCTGTC
TATTTCTTCATCCTAAGTGGCTCAGCTCAGGCACTCTCGGTGAGATAACTACGATA
CGGAGGGGTTACCACATCTGGCCCAAGTGCATGATACCGCGAGACCCAC
GCTACCGGCTCCAGATTATCAGCAATAAAACGCAAGCCGAAGGCGGGA
GCGCAGAAGTGTTGGTCTCTGGAACTCTATCCGGCCTCCATCCAGTCTATTTAATTGTT
GCCGGAAGCTAGAGTAAGTGGTTCGGCCAAGTTAAATAGGTTGCGGAACGTTG
TGCCATTGCTACAGGCACTCGTGGTGTGTCACGCTCGTGGTTTGGTATGAGCTTCAT
TCAGCTCCGGTCCCACAGGATGCAAGGCGGATTTACATGATCCCTGGCTTGGTC
AAAAAGCGGTTAGCTCCTTCCGTCTCCAGATCAGTTGCTCTCGGCGAGATGTGC
CCGCAGTGTATACACTCATGGTTATGGCAGCAGCTGCAATAATCTCTTTACTGTC
ATGCCATCAGTAAGTGCTTTCTCTGTAGCTGACTGGTAGTAGTACTCAACCAAGTCATT
CTGAGAATAGTGATGCGGCCGACCCAGTTGGTCTGCTTGGCCCGCCGCTCAATACG
GATAATTACCAGCGCCACATAGCAGAAGTTAAAGTGCATCATCTTTGGAAAAAC
GGTCTTCGGGGCGAAAACCTCTCAAGGATCTCCGCTCTCGTGGAGATCCAGTTCG
ATGTAACCCACTCGTGCACCCAACCTGATCTCTTTACTTTACTTTCAACCAG
CGTTTCTGGGTGAGCAGAAGAACAGGAGGCATGGCATGGCCAAAAAGGGGAAT
AAGGGCGACACGGAAATGTTGAAATACCTCAGTCTCTCTCTTCTTTCAATATATT
GAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT
TAGAAAAATAAACAAATAAGGGTTCCGCACATTTCCCCGAAAAGTGCCAC
CTGAC

Plasmids used in construction of pSLIK2

>pUC19::Crimson/P2A-NLS/HA-HumBeta

GACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTACATGATAATA
ATGGTTTCTTAGACGGTCGTGCACCTTTTCGGGGAAATGTGCGCGGAACCC
CTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAA
AACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCA
ACATTTCCGTGTCGCCCTTTATTCCCTTCCGTCGATTTTGCCCTTCTGTTTTT
CTGCCAGTGGCGATAAGTCGTCTTACCAGGTGACTCAAGACGATAGTT
ACCGGATAAGGCAGCGCTCCGCTGACAGCAGGCGTAACGAGGAAAGGCGGAGACGATAGTT
CTGCCGCTCTACCTACCAGAACCTAGGACGACCTCAGGTGGCTCCAGGGCGAAGG
AGCCGACCCCTCGCTGCAGGAGCACGAGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGA
GGAAGCGGAAGAGCGCCCAATACGCAAACCGCCTCTCCCAGCCTTTGGCCG
ATTCATTAATGCGAGTGAGCACGAGGTGTTCGCTGATCTCACTGCTACGACG
AGCGCAACGCAATTAATGTGAGTCTGACTGAGTCTCAGACGACGACGACG
ACACTTTATGCTCCGGCTCGTATGTTGTGTGGAATTTGTGAGCGGATAAACAAT
TTCACACAGGAAACAGCTATGACCATGATTAACGCAAGGTACGCTCCGTAACCG
ATGTGCGATCGCAGATCTAGAGGTCCGCTGACCATTGAGTAGACTGAGAACG
TCATCAAGCCCCCTGTCCGCTTCAAGGTGACATGGAAGGCCTCCGCTGAAACGG
CCACGAGTTGCAGATCGAGGGCTGGGGCGAGGGGAAAGGCCTACGAGGGGCAC
CCAGACCCCAAGCTGCAGTGACCAAGGCGGCCCCCTGCCCTTCCGCTGG
GACATCCCTGTCCTCCCAAGTTCTTCTAGGCGTCCCAAGGCGTGACATCAAGCACC
CGCCGACATCCCCGACTACCTCAAGCAGTCTCCTCCCGAGGGGCTTCAAGTGG
GAGCGCGTGATGAACTTCGAGGACGCGCGGCTGGTGACCGTGACCAGTGCAGGACT
CCTCCTGCAGGACGAGCGGCCCTCATCTACACCAGTGTTTCATCGGCGTGAA
CTTCCCCCTCCGACGCCCCGTAATGCAGAAGAGAAGACTCTGGGGCTGGGAGCCC
TCCACTGAGCGCAACTACCCCGCCGACCGGCGTGCTGAAGGGCGAGAACACA
TGGCGCTGAAAGCTGAAAGGCGGGCAGCCCACACTACCTGTGAGTTCAATCCAT
CTACATGGCAAAGAGCCCTGTAAGCTGCAGGCTACCTACAGTGGACTAC
AAGCTCGACATCACCTCCACAACGAGGACTACACCCGTTGAGGAGCAGTACG
AGCGCGCCGAGGCCCAGACACCTGTTCCAGGATCCGAGGACACGAACTT
CTCTCTGCTAAAGCAAGCAGAGAGATGGTGAAGGAAAAACCCCGGTCTCTCAGT
GTGCCTCCCAAGAAGAGGAGGAGAGTGCAAGGACACCGCAAGCTATGAC
GTGCCCGATTACGCTTGGCGAGGAGGCCAGCTCGCGGCGAGGAGGATCC
GGAGGAGGCAGCATGTCACACAGTCTCAGCCACCTGGGCGGAAAACTTGCGG
AGAGAGTGGAAGGCAAGACTCCGTGGATCCCCAAGGAGCTGACACACTGAG
GCAAACCGCCTTCAGGGGAGATGCTTCCAGCCAGCCAAATTCATCGGCTC
ATCGTGCCCAACCAGTACGCGCCTGAACCCCTTTGGAACCAAGGAGATTTAGCCT
TCCCGCAAGCAGAATGGCATTGTCCCTGTGGTCGGCGTGGACGGCTGGAG
CAGAATCAT
AGGACGAAGGTGGCATGAGATCTACGAGGAGAGGAGCAATCCCTATCT
GCCTACCCGAGTGGAAGGAGATGCTGCAAGAGAGAACCCTTCAGAGCAGAG
AGGACAGAGAGCATCCGACGAGGATCTACAGGAAGGACAAGGAAATCTCTATCT
GACACAAGGCCATGATCCGACGTGACGCTTCGCCTTTTGACTTCCGCGGAGT
CTATGATAAGGACGAAGCCGAGAGGAGATCGTGAGTAAATACCCCTATACCCGCT
GAAAGGCAGCCGAAAGGGACATTACACCCGTAAGCAGGACACCATGCG
GAGACACATATACCTGCTCTTTGACGGAAACCTGGGAGCAGTACGACCTCC
TCCCCCTCTGCAAGGCCAAATCTTCAGGAGGGACATCAGGGCCAGCTCCGAACT
CACACAAGCCGAAGCTGTGAAGGCCCTCGGATTCCTCAAACAGAAGGCCGCT
GAGCAGAAGGCTCAGCCGCTGCCTTAACAGAATTCACTGGGCGTCGTCTTAACAAC
GTCTGACTGGGAAAAACCTGGCGTTACCCAACTTAATCGCCTTTGCAAGCA
TCCCCCTTTCGGCAGCTGCGTAATAGCGAAGAGGCCGCAACCAGATCGCCT
TCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGCTGACGTTACATTTTCT
CCTACGCTCATCTGTGCCTGTATTTCAACCCGCTATATGGGCGACTCTCTCAGTACAA
TCTGCTCTGATGGCCGATAGCTAAGCCAGCCCGCAACCCCCGACACCCGACACCCGC
TGACGCGCCCTGACGGGGCTTTGTCTGCTCCCGCGCATCCCGCTTACAGACAAAGCT
GTGACCGGTCTCCCGGAGCTGCTGTGTCAGAGGTTCACGTCATCACCGAA
ACGCGCGA

>pENTR2B/TREPitt::Crimson/P2A-NLS/HA-HumBeta

[pENTR2B-TREPitt-HumBeta-P2A-Crimson]
4353 bp
>pENTR2B/TREPitt::Crimson/P2A-NLS/HA-HumBeta

CTTTCTGCGTTATCCCTGTGATTCTGTGGATAAACCCTATTACCGCTAGCATGG
ATCTCGGGGACGTCTAAACTACTAAGCAGAGAGTAGGGAACTGCCAGGCATCAA
ATAAAAACGAAAGGCTCAGTGGAAAGCTGGGCTTCTCCTGTTTTATCTGTTTAT
GTGGTGAAACGAGCTCTCTGAGTAGAGACAAATCCGCCGAGGAGCGGATTTGAAAC
GGTGGTGAAGCAACGCCCAGGAGGGTGGCGGCGAGGACGCCGCACCATAAACT
GCCAGGCATCAAACCTAAGGAGGAGCCATCTGACGGATGGCCCTTTTTGTGCT
TTATACAAACTTCTCCTGTATTTAGTATTTAAGCTCGGGCCCCAAATAATG
ATTTTATTGACTGATAGTGACCTGCTGGCAACAAAATGATAAGCAGATAT
CTTTTTATAATGCCAAACTTTGTAACAAAAACGAGGCTGGCGGAGCCGAACCA
TTAGatatcAAAGTGCCACCTGACGTCTCCCTTATCAGTGATAGAGAAGTGCACAC
GTCTCGAGCTCCCTATCAGTGATAGAGAAGGTACGTCTAGAACGTCTCCCTAT
CAGTGATAGAGAAGGTACGTCTAGAACGTCTCCCTATCAGTGATAGAGAAGG
TACGTCTAGAGACGTCTCCCTATCAGTGATAGAGAAGGTACGTCTAGAACGTCTC
TACGTCTAGAGACGTCTCCCTATCAGTGATAGAGAAGGTACGTCTAGAACGTCT
CAGGCCATCTGCTCTGCTGCTGCTGCTGTCTGCTGCTGCTGCTGCTGCTGCTG
GACAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAGACGCCATCC
ACGCTGTTTTGACCTCCATAGAAGACACCGGGACCGATCCAGCCT
GGAGGTCCGCCACCATGGATAGCACTGAGAACGTCATCAA
GGCGTGGGCGAGGGCAAGCC
CTTCTACGGCTCCAAGGGCGTACATCAAGCAACCCCGCCGACATCCCCGACTAC
CTCAAGCAGTCCTCCCGAGGGCTTCAAGTGGGAGCGCGTGATGAACCTTCG
AGGACGGCGCGTGTTGAGCTGACCAGTGACCCAGGACTCTCCTCCCTGCGAGGACGGCAC
CCTCATCTACCAGTGAGGCTCATCGGCCTGAACTTTCCCTGCCAAGCCTCCC
TAATGCAAGAAGAAGACTCTGGGCTGGGAGCCTCCCCTGACCTGAGCAGCAGTACCC
CCGCGACGGCGTGCTGAAGGCGGAGAACCACATGGGCGCTGAAAGCTGAAGGG
CGGCGCCACTACCTGTTGAGTCTCAAGTCCATCTACATGGGCAAGAAGCCC
GTGAAGCTGCCCCGCTACCTGAGCACTCAAGCTGAAAGCTGACATCACCTCCC
ACAACGGAGGACTACACCCGTGTTGAGCAGTACGAGTACGAGCGCGCCAAGGCCC
ACCACCTGTTCAGGATCCCGAGCCACGAACTTCTCTCTGTTCAGAAGCAAGC
AGGAGACGTGGAAGAATCCCCCGGTCTTCCACATGGGTGCTTCCCAAGAAGAA
GAGGAAGGTGCAGGGACCAGAAGTTTCTATGACGTGCCCGGATTACGCTGTC
GGAGGCGAGGCTCCGCGCGAGGAGGAGGATCCGGAGGAGGAGCAGCAGTGC
ACAGCTCTCGCCACCCTGGCCGCCAAAACCTGGGCCAGAGATGGAATGGCG
CCGTTGATCCCGAGCTGATCACACACTGAGGCAAACCCGGCTTCAAGGG
AGATGCTTTCCGACGCCCATTATCATCGCCCTCTCTCATCGTGGCCAAACCAGTACG
GCCTGAAACCCTTTGGAGCCAGGAGTTTACGCCTCTCCCGGACAAGCAGATGG
CATTGTCCTGTGTTGCTCGAGGGCTGGAGGACGAGCAGAATGGATTTAACCGGAGAC
CAGCAGTTTGGACGAGGTCTTCGAGCAGCAGACAGAGCTGACATGCA
GGATCTACAGGAAAGGAAGGACCCATCTATCTGCTCGTCACCCGAGTGGAATGGA
TGAGTGCGAGAAGAGAAACCCTTCAAGACGACGAGGCGAGAGAGACATCACAG
ACCCTGGCAGTCCCACTCTAAGGAGAATGCTCAGACACAAAGGCCATGATCCAG
TGTCAGGCTGCTTGTTGCTCGCCCGGACCTATGTGAATAAGGACAAGGCCG
AGAGGATCTGGGAAATACCCGGCTACACCGCTGAAAGGCGGCGCCAAAGG

582
ACATTACACCCGTGAACGAGACGAGACCATGCAGGAGATCAATACCCTGCTCAT
TGCCCTGGACAAAAACCTGGGACGATGACCTCCTCCCTCCCTCTGCAGCCAAATC
TTCAGGAGGGACATCAGGGCCAGTCCGAAACTACACAAGCCGAAGCTGTGA
AGGCCCTCAGATTTCTCTTCTCAAAACAGAAGGCGCTGAGCAGAAGGTCGCCGCCGT
TTAAACGAATTTCGCGGCCGCACCTCGAGATATTCTAGACCCAGCTTTTCTTGATACA
AAGTGGCATTATAAGAAAGCATTGCTTTATCAATTTGTGCAACGAAACAGGT
CACTATCAGTCAAATATAATACATTATTGTGCAAAACAGATACAAAGGGGTGTTATGAGC
CATATTCAACGGGAAAAGTCAGGCAGGCCGCAGTAAATTCCACACATGGATGCTG
TTAATATGGGTATAAATGGGCTTCGCGATAATGTGGGCCATACAGGTGAGCC
AATCTATCGCTTTGTATGGGAAGGGCGTCTGCGAGTTTGTTTCTGAAACATG
GCAAAGGGTACGCTGGCAATGATGGTACAGTGAGATGGTCAGACTAAACTG
GCTGACGGAAATTATGCTCTTCTTCGACCATCAAGCATTTATTCGTCGCACTACTG
ATGATGCATGGTTACTCACCACAGTTCGATCCCCGAAAAACAGCATTCCAGGT
ATTAGAAGAATATCCTGATTCCAGGTAAAAATTGTTGATGCCTGCGACGTG
TTCTCGCCGGGTTATCAGTTCTCTGCTCTGCAACTCAACGAAATACGGTTTGGTTGA
TGCGAGTGATTGGAGTGACGAGCGTAATGGCCTGCTGGAACGACGTCTGG
AAAGAAATGCATAAATTGTGCAAGTCGTCACTCATGG
TGATTTCTCACAATCTTATATTGTGATAGGGGAAAATTAAGTGGTTGTA
TTGATTTGGACGAGTCGGAATCGCAGACCGATACCAGGATCTTGCCATTC
ATGGAACCTCGTGGTGAAGGTTCTCCTCCATTACAGAAAACGGGCTTTTTTCAAA
AATATGGTATTGATAATCCTGATATGAAATTGCAGTTTCATTTGATGCTC
GATGAGTTTTTCTAACGAGAATTTGTTAATTGGTTGTAACATTATTCAGATTG
GGCCCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCT
TGAGATCCTTTTTTTCTGCAGCGTAAATCTGCTGCTTGCAAAACAAAAACACCC
GCTACCAGCGGTGGTTTGTTTTGGCGGATCAAGAGCTACCAACTCTTTTTTCCGA
AGGTAACCTGGCCTTCAGCAGAGCGCAGATACCAAAATACTGTCTTCTCTAGTGTA
GCCGTAAGTTAGGCCACGACAATTCCTGCTGAGCACCAGCCTAATACCTACCTC
GCTCTGCTAATCCTGGTACAGTCGTGCTGCACCAGTCGATAAGTCTGTGCT
TACCAGGTTGACTCAAGACGATAGTTACGGATAAGGCACGCAGTGCAGGC
TGAACCGGGGGTTGCTGACACAGCAGCCAGTTGGAGCGAAGCACCTACCCAG
AACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCCTCCCCCCAGG
GAGAAAGGCACGGACAGGTATCCGGTAAGCGGCGAGGTCGAGACAGGAGAGCG
CACGAGGGAGCTTCCAGGGGGAAAACGCCTGATTATCTTATGTCCTGTGG
TTTCGCCACCTCTGAGCAGTGGATTTTGATGCTGCTGTCAGGGGGCG
GAGCCTATGGAAAAACGCACGCAACGCAGGCTTTTTTACGGTCTTGCCCTTTT
GCTGGGCTTTTGCTCAGATGTT
>pSLIK/TREPitt::Crimson/P2A-NLS/HA-HumBeta

AKS pSLIK-Zeo-TREPitt-Crimson-P2A-HumBeta or pSLIK2. Sequence verified from 2633 to 4868 bp.

>2.pSLIK-Zeo/TREPitt::Crimson/P2A-HumBeta

GTCGACGGATCGGGAGATCTCCCGATCCCCCTATGGTGCACTCTCAGTAACATCTGCTCTGATGCCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTTCGCGATGTACGGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTATTACATACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGC
CCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTT
CCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTA
CATCAAGTGTATCATATGCAACTGACGCCCCCTATTGACGTCAATGACGGTA
AATGGCCCCGCCTGGCATTATGCCCAAGTACATGACTATTATGGGACTTCTCTACT
TGGCAGTACATCTACGTATTATGCTACGCTATTACCATGTTGATGCGGTTTTTG
GCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGT
CTCCACCCATTTGACGTCAATGGGAGTTTTGGCACCAAAATCAACGGG
ACTTTCCAAAAATGTCGTAACAAAACACTCCGCCCCATTGAGCAAATGGGCGGTAG
GCGTGTACGGGTGGGAGGTCTATATATAAGCAAGCGCCTTTTGCTCTGTAATGGGTCT
CTCTGGTACCCAGATCTGACTGGCAGCTCTCTGCTCTGCTAATAGGAACC
CACTGTAAAAGCCCTCAATAAAGCCTTTGGACGTCTTCAAGTAGTGTGTGCCC
CGTCTGTGTGTGACTCTGGTAACATAGAGATCCCTCAGACCCCTTTTAGTCAGT
GTGGAATAATCTCTAGCAGTGGCGCCGAACAGGACTTTGAAAGCGAAAGGG
AAACCAGAGGAGCTCTCTAGCAGACTCGCGCTTTGCTGAAGCAGCCGCACCG
CAAGAGGCGAGGGGCAGCGACTGGTGAGTGACGCCAAAAATTTTGACTAGCG
GAGGCTAGAGGAGAGATGGGATGCGAGGCTCGGCTAGTTAAGCGGGGGA
GAATTAGATCGCGATGCCAAAAATTCGTTAAGCCAGGGGAAGAAA
AATATAAAATAAAAACATAATAGTATGGGCAAGCGGAGGAGCAGTAAAGCGCATTCGC
AGTTAATCCTGGCCTGTAGAAACATCAGGAAGGCTTAGACTAGCAAAATACTGGGA
CAGCTACAACCATCCCTCAGACAGGATCAGAAGAACTTAGATCATTATATA
ATACAGTACGCAACCCCTCTATTGTGTCATCAAAGGATAGAGATAAAGACAC
CAAGGAAGCTTTAGACAAAGTAGAGGAAAGAGCAAAAACAAAAAGTAAGCCAC
CGCAGTACAGCGCGCTGACTCTACAGACCTGGAGAGGAGGATATGGAGG
ACAATTGAGAGAAGTATATAATAAATATTAAAAGTAGTAAAAATGGACATT
AGGAGTAGCACCCACCAAGCAGAGAGGTGGTGCAGAGAGAAAAAA
GAGCAGTGGAATAGGAGCCTTTGTTCTTCTTGGGTTCTTGGGAGCAGCAGGAAG
CACTATGGGCAGCGTGCAATGACGCTGACGGTACAGGCCAGGCAGAA
TCTGGTATAGTGCAGCAGCAGAAACATTTGCTGAGGCTATTGAGGCAGC
AGCATCTGTGGCAACTCAGTCTGGGCAATCAAGCGCTCCTCTGGGGATTTGGGGT
TGCTCTGGAAAAACTCATTGGCACCACTGTGTGCCTTGGGAATGCTAGTTGGAG
TAATAAATCTCTGGAAACAGATTTTGGGAATCACAACGCCTGGGATTTGGGGAC
AGAGAAATTAACAAACTACAAAGCCTTAATACACTCTCTTTAAATTGAAAGATCGC
AAAAACCAGCAAAGAAAGAATGAACAAGAAATTAGTGGAAATTGAGAAAAATGGG
CAAGTTTGTGGAATTGTTACATAAACAATTTGGCTGTGGTATATAAAATTA
TTCTATAATGATAGTAGGAGCTTGGGTAGTTTTAAAGATAGTTTTTGCTGTACT
TTCTATAGTGAAATAGAGTTTAGGCAGGGATATTCCATTACATTTCGTTCAGACCC
ACCTCCCAACCCCGAGGGGACCCGACAGGGCCGAGGAATAAGAAGAAGAAG
GTGGAGAGAGAGACAGAAGAGATCCATCGTGAACGGGCAGCCAGC
TGCCTGCGCCAATATTGTGCAGACAAATGGCAGTATTCATCCAAATTAAAAA
GAAAAAGGGGGGAGGTTGGGGGTAGCTAGTGCGAGGGGAAATAGTAGACATAA
TAGCAACAGCATAACAACTAAAAAAGAATACTACAAAAAAATCCTTCCTTCT
AAAAATTTTCGGGTTTATTACAGGGACAGAGATCCAGTTGGTTAATCTA
GACAGTTATCACAAGGTGGTGACAAAAAGCTGAAACGAGAACGCTAAAATG
ATATAAAATATCAATATATTAAATTTGATTTTCATAAAAACAGACTACATA
ATACTGTAAAAACACAAATATCCAGTCATATGGGCAGCGGAAACCAATTCAG
ATAATAGTGCCCACCGTCTCCCTATCATGATAGAAGATCGACACGT
AGCAGAAGGACATTTTAGGACGGAACCTTGCGACTCGGCACTGACTGTTTTT
CTTTCCAGAGAGCAGAGACAGGGAAGGAAAATGAGTCCCTTCGCGAGATTCT
GCAGGAGGGGACTCCTCCGCTGCGAGGCAGACGAGGATTTGGGTCGCGGCTCTG
TTTGTGGAGTCGCTGTGATCGTCACGTTGGGAATAGCGGGGCTGCTGGGCCTGG
GGGCTTCGCTGGGCCGCATGGCGTTTGGGAGCGCAGCAAAATGGCGGCTGTTCC
GGGGCTTTCGTGGCCGCCGGCCGCTCGGCTGGGAGCGGACAGCGTGTGGAGAGA
CCGCCAAGGGCTGATAGTCTGAGGAGCAGGACGAGGAGGCTCGAATAGGGGAG
TTGCGGAGCAGCGCAGCAAAATGGCGGCTGCTGGGCCTGGGCCGCATGGGCTA
GGGCAAGAACCAAGGTCTTGAGGGCTTCGCTAATGCGGGAAAGGCTCTATT
CGGGTGAGATGGGCTGGGGCACCAGCATCTGGGAGCAACTCGAGTGAAGTTTG
CTGACTGGAGAAGCTCGGCTCGGAGGCTCCTTGTTGGGGAGGGGCGGAGTTATG
GGCAGCTGCCGCTTCTGCTGGTATAGTCTGAGGATAGGGGAGGGAAGGATAAGT
GAGGCCGTCCATTCTGGGAGTGCTGCTGGGAGGGGTGGAGGAGGTAAGGT
AGCTCCGTTTTTGAACATGCGCTCGGAGGTTTGGCGAGGTGCTTTTTGTGAAGT
TTTTagGCACCTTTTGAAATGTAAATCTCATTTGGGTCGCAATATGTAAATTTC
TAGACTAGTAATTTGTCGACTAAATCTGCGGCTTTTTGCTGCTTTTTTTTGTA
TAGACTGGTACCGAGTCGAGTCCGCTTGACCTCCGCCTCTCGCTGCAGTCGAA
TTCAACCAGTGCTAGAGCTGACAGGAAGAGCAGAAATAGTCTAAAAACGGAGCTG
TACTCAATGCTGTGCTGATCGAAGGCCTGAGCAGAAGGAAAACCTCGCTCAAAA
GCTGGGAGTTGAGCAGCCTACCCCTGTACTGGCAGTGGAAGAAACAGCAGGCCC
CTGCTCGATGCCCTGCCAATCGAGATGCTGGACAGGCATCATACCCACTTCTG
CCCCCTGGAAGGCGAGTCATGGCAAGACTTTCTGGAAGAATACGCTCGCTTCTG
TGTCAGCAAGGCTTTTCCCTGGAAGAAGCGCAGACTGTACGCTGTGCTGCCGCGTGG
GCCACTTTACACTGGCTGCGTATTGGAGGAACAGGAGCATCAAGTAGCAA
AGAGGAAAAGAGAGACACCTACCACCAGATTCTATGCCCCACTTCTGAAGACAA
GCAAATTGAGCTGGTGACCCGCGAGGGAGCCGACCTGCTCCTTTTTCGGCCT
GGAACTAATCATATGTGGCCTGGAGAAACAGCTAAAGTGCGAAAGCGCGG
GCCGACCGACGCCCTTGACGAGATTGACTTAGACATGCTCCAGCCGATGCC
TTGACGACTTTGAGCTGTGATATGTGCTGCTGACGCTCTTTGACGATTTTGAC
CTTGACATGCTCCCCGGGTAACTAAGTAAGGATCCGCGGCCGCGACTAGAGGA
ATTCCGCCCTCTCCCTCCCCCCCCCTTAAGCTGGAAGCCGCGGGAGCGCTGG
AATAAGGCCCCGTGTGTGTGGTTGGTCTATAGTTATTTTCCAACCATATTTGGCCTT
TTGGCAATGGAAGGGCCGGAAACCTGGCCTGTCTTTTGACGAGCATTCCTCT
AGGGGTCTTTCCCTCCTCGCACAAGGAATGCAAGGTCTGTTGAATGTCGTA
AGGAACGATTCCTCTTGGAAGCTTCTTGGAATGTCGTAAGCGTACGAC
CCTTTGCAAGGCAAGCGAACCCCCCAACTGCGCACAGGGCTCGCTGCCGAA
AAGCCACGTGTATAAAGTACACCTGCAAAAGGCGCACAACCCCGAGTGCTCAG
TTGTAAGTTGGATAGTTGGAAGAGCTCAAATGGCTCTACCTCAAGCGTAGT
CAACAAGGGGCTGAGAGATGGCCAGAGGTTACCTGATGATTGGGAATCTGA
TCTGGGGCCTCGGTGACACATGCTTTTACATGTTTGGTAGCAGGTAAAAAACC
GTCTAGGCCCCCGAACCACGGGGACGTGGTTTTCTTTTGAAAAACACGATG
ATAAGCTTACCGGTACGCTGATGTGTTGACAATTAAATCATCGGCGATAGTATA
TCGGCATAGTTAATAACGACAAGGTAAGGGAACACAAACATGGCCAAGTTGAC
CAGTGCCGGTTCGGGCTCTCACCACGGCGACGGCAGTCGCGGAGACGAGCTG
TGGACCAGCCGGCTGGGTTCCTCCTCGAGACTTCGTGAGGAGACACCTGCTCCGCG
GTGATGCTCAGGGAGCCGAGCTGACCCTGTTATCATGACGCGGCTCAGGACCAGGT
GGTGCCGGACGACAAGGGGACCTGGGATTTCGACCCCGGACCCCGGGACGGGGAC
GCTGCCGGAGGCAGCTCTCACTGCAGCCGGAGACGAGAATTCGCGGTTGCTGGAGG
CAGGGACGAGTGGTCGGA
GGTGCCTGGCCCGAGCTGACATCGGTCGGGTCCACGAACTTCGCGGAGACGAGCCTCCG
CGCCATGACCGAGTCAAGCAGAGCCCTGGGGGACGGAGCTGCTGCCCTG
CGGCCATGACCGAGATCGGCGAGCAGCCGTGGGGGCGGGAGTTCGCCCTGC
GCGACCCGGCCGGCAACTGCGTGCACTCTCGTGCCCGGAGGAGTTCGCCCTGC
CGGCCATGACCGAGATCGGCGAGCAGCCGTGGGGGCGGGAGTTCGCCCTGC
CATCTGTACAAGTAAAGCGCCCGCGACTCTAGACATCAATAACGCCCATAACAC
ATTGTAGAGGTTTTTACTTTGCTTTTTTTAAAAACCTCCCCACACCTCCCCCTGAACC
TGAAACATAAAAATGAATGCAATTGTTGTTTAGTCCCTCCAATGGACGACGCTCCG
CAAGCTTATCGATAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTG
GTATTCTTAACATGTGGCTCCTTTTACGCTATGTGGATACGCTGCTTTTAAATGC
CTTTGTATCATGCTATTGCTTCCCGGTATGGCTTTTCTATTTCCTCTCCTTGTATA
AATCTCTGTTGCTGTCTTTTATGAGGAGTGTGCGCCGTGGTCAGCGGCAACGT
GGCGTGGTGTGCAGCTGTGTTGCTGACGCAACCCCGACCCGACTGTTGCGGGGAC
CACCCCTGTCACGGCTCTTTCCCAGGGACTTTCGCTTTTTCCCTTTCCCCCTATTGAC
GGCGGAACCTCATGCGCCGCTGGCTCTGGCGCGGTAGCGACGGGGCTCGGCTG
TGGGCACTGACATTTCCGTTGTGTGGTGCTGGGGAGAATCATCGTCCTTTTCTTGG
GCTGCTCGCCTGTGTTGCGCACCTGGATTCTCGCGGGAGACGCTCTCTTTGCTACG
TCCCTCGGCCCTCAATCCAGCGGACCTTCCTCCTCCCAGGGGCTGCTGCGGCT
CTGCGGCTCTTTCCGCCTCTTCGCCTCGCCCTCAGACGAGTGCGGATCTCCCT
TTGGGCGCCTCCCCGCATCACATCAGCTCACAGCTGACCTGAGATCGTGATCTCCCT
ACATGGAGCAATCAAGTAGCAATACACAGGCTACCAATGCTGATTGTGCC
TGCTGAGAAGCAACAAGGGAGGGAGGGAGGGAGGTTGCTGATCTTTGCCACTACCT
AAAGAAAGGGGGACTGGAAGGGCTAATTCACTCCCAACGAAGAACAGAGAT
AATCCTTGAGTTGAGATCCTACACACAACAAAGCTACCTCCCTGTATTGAGCAGAA
CTACACACCAGGGCCACCGATCAAGATATCCACTGACCTTTGGATGGGTGCTAC
AAGCTAGTACCAGTTGAGCAGGAAAGGGTAGAAGAAGGACATGAAGGAGAG
AACACCCGCTTTGATTACACCTGTGAGCCCTGATGGAATGGATGCCTAC
GAGAAGTAGATTTAGAGTTGAGGTTGACAGCCCGCTAGCATTTCACTACATGGC
CCGAGAGGTGCTCAGTGGACTGTACGTGGGCTCTCTGGTTGATAGCCAGATCTGA
GCCTGGGAGCTCTCTGCTTTACTAGGGAAACCACCGTATTTAAGCCTCAATAAA
GCTTGCCTTGAGTGTTTCAAGTAGTGTTGAGCGCTCTGTGTTGTGACTCTGCT
AATCCTGTGATCCCTCAGACCCTTTTAAGTCAGTGGAAATCTCTAGCAGGG
CCCTTTAAAACCCGCTGATCAGGCCTCAGACTGTGCGCTCTTCTAGTGTGCCAGCATC
TGTTGGTTCGCCCCCTCCCCGTGGCTCTTCCTTGACCTGGAAAGGTTGCCACTCCCA
CTGTCTTTCTCATAAATAAGGAAATTCGATCGCATTGTCTGAGTAGGTGT
CATTCTATTTCTGGGGGGTGGGCTGGGCGAGACAGCAAGGGGAGGGAGATGG
GAGAAGCATAGCAGGATGCGGATGCCCCGTGGGGCTCTATGGCTTGTAGGTGT
CGGAAAGGAACCAGCTGGGGCTCTAGGGGATATCCACGCGCCCTGTAGCGG
CGCATTAAGCGCGCGGGGTGTTGTTGTTACGCGAGCGTGGACCGCTACACTT
GCCAGCGCCCCATAGCGGCGCCCTCCTCTTCTGCTTCTCCCTCTCCCTTCCTCCGACG
GGACGCCGGCTGGATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTC
GCCCAACCCAACTTGTTTATGGCAGTTATAATGGTTACAAATAAAAGCAATAG
CATCACAATTTTCACAATAAAAGCATTTTTTTTCCTGCATTCTAGTTGGTGGTT
GTCCAAAATCTCATCAATGTATCTTTATCATGTCTGTATACCGTCGACCTCTAGCT
AGAGCTTGCGTAATCATGGTCTAGCAGCTGGTTTCTCTGTTGTGAAATTGTATACCCG
CTCACAATTCACACAAACATACGAGCCCGAAGCATATAAGTGAAGCTACGCTG
GTGCTTAATGAGTGAAGCTAATCATAATATTGCTTGCGCTCAGTGGCCGCT
TTCCAGTCGGGAAAACTGTGTGCTGCCAGCTGCTATTAATGAAATCGGCAACGC
CGGAGAGAGCGGTGTGCTATTGCGGCGCTCTCCTGGCTTCTCGCTGCTACTGAC
TCGCTGCGCTCGGTCTCCGGTCGAGCGGCGAGCGGTATCAGCTCACTCAAAGG
CGGTAATACGTTTATCCACAGAATACAGGGGATAACGCGAAGAAAGCATGT
GAGCAAAAAAGGCAGAAAAAGCCAGGAAACGTAAAAGGCCCGCTGGCTGG
CGTTTTTCATAGGCTCCGCCGCCCTGAGCACGACATCACAATATCGAGGCTCA
AGTCAGAGGTGCCGGAAACCAGAGAAGACTATAAAGTACCAGGCCTTTCCCC
CTGGAAGCTCCTCGGCTCGCTCTGTTTCGGACCCCTGCGTTACCCGGATAC
CTGTCGGGCTTCCTCCGGGAAGCGCTGTGCTTTCTCATAGGCTCAGCGCTG
TAGGTATTCAGTTCCGGGTAGGTGTCGTTCTGCTCTGCTCAAGCTGGGCTGTCGACG
AACCCCCCCGTTCAAGCAGCACTATCCGGTAGTTACACGACTATCGTCTTGAG
TCCAACCCCGTAAAGACAGACTTATCGCGGACTGCGCAGCAGCCACTTGGTACAC
GGATTAGCAAGACGCCGAGGATGATGACCAGGTCCAGGCTACAGAGTTGCTGA
GCCTAACTACGCTACACTAGAAGAAGACTATTTGTTAAGATCTGCCTCTGCTGA
AGCCAGTTACCTCGGAAAAGAGTTGGTAGCTTTGATACCGGCAAAACAAAC
CACCAGTGATAGCGGTTGTTTTTTGTTTGCAAGCAGCAGATTACCGGCAGAA
AAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAG
TGGAACGAAACACTCAGTTAAGGGATTITTTGTCATGAGATTATCAAAGAGAA
TCTTCACCTAGATCTTTTAAATTTAAATGAGTTTTAAATCAATCTAAAGTG
ATATATGAGTAACCTTGGT CGACAGTTAAGCAATGCTTAAACTCAGTGAGGCAC
CTATCTCAGCGATCTGCTCATTTTCCATCATGCTTGCTGACTCCCGTGCG
TGTTGATAACTACGATACGGGAGGGCTTTACCATCGTCCGCCCCAGTGCTGCAAT
GATACCGCGAGACCACGCCTACGCCGCTCCAGATTTATCAGCAATAAACAGA
CCAGCCGGAAGGGCCGAGGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCA
TCCAGTCTATTAAATTCTCTCTGCGGGAAGCTAGAGTAAGTAGTTTCGCCAGTTAAT
AGTTTGCGCAAGCGTTGTTGCCATTTGCTACAGGCGATCGTGGTGTCACGCTCGTC
GTTTGGTATGGCCTTCATTGACGCTCCTCCCTCCCAAGGATCAAGCGAGTTACAT
GATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGGTTCTCCTCGATCGTT
GTCAGAAGTAAAGTTGCGGAGCTGTATATCACTCATGTTATGGGCGACGACTGC
ATAATTTCTCTTACTGTACATGCAATCTTTCTGCTGGACTGAGGTTGAG
TACTCAACCAAGTGCTTCTGAGAATAGTGTATGCGGAGCCGAGTTGCTCCTTG
CCCGGCGTCAATACGGGATAATAATCGGCACCATAGCGAATTAAAAGTG
CTCATTGAAAACGTTTCCTGCGGCGAAAACCTCCTCAGGATCTTACCGCT
GTTTAGATCCAGTGGAGATGCTACCTAATCCTGACCCACTCTGACCACTTCAACAT
CTTTATCTTCCACCACAGTTTCTGGGTGAGCAAAAAACAGGAAAGGCACAATGC
CGAGGAAAAAGGGAATAAGGGCGACAGGGAATGTTGAATATACCTATCTACTCTTC
CTTTTTCAATATTATGAAGCATTATTCAGGGTTATTTGGCTCATGAGCGGATA
CATATTTGAAATGTATTGAGAAAAATAAATAAATAAGGGTTCCGCACAGTTT
CCCGAAAAGTGGCCACCTGAC
Plasmids used in construction of pSLIK3

>pFN24K::Crimson/P2A-Control

AKA pFN24K::Product N-Control.

>pFN24K::Crimson/P2A-Control

ATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTGGTTTAG
TGAACCGTCAGATCACTAGAAGCTTTAATACGACTCAGACTATAGGGTACGATT
AGGTGACACTATAAGGAGAATTTCATCATGGCCAGAAATCGGTACTGGC
TTTCCATTCCGACCCCATATTATGGAGGTCTTGGGCGAGCGCATGCACTACGGT
CGATGTTGGTGAGGAGTGGCACCCTCTGTGGCTGCTTCTGACGTCATGGAACG
CCTCCCTCTACCAGATCAGGAGGTTCCGACCACATGGATAGCAGTGAGACG
TCATCAAGGCCCTTTCATGCGTCTCAAGGGTGCACTGGAGGGCTCCGTTGAAACGG
CATCTCACCTTGCTCCTGCGAGAAAGTATCCATCATGGGTGATGCAATGCGG
CGGCTGCATACGCTTGTATCCGGCTACCTGGGCCATTTGGGCTACCACCAAGCGAAAC
ATCGCATCGAGCGACGCACCTCGGATGGAAGCCGGTCTTGTCGATCAGGATGCTGGAGGATCTCGTCTGAGACTCATGGCTGTGGACCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGAC
TGATATTTGTGAAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTT
TACGTTATCGCCGCTCCGATTTCCAGCGACACTCGGATGCCTCCTTCTATCGCCTTCTTGA
CGAGTTCTTTCTGAGCGGGAATCTCTGGGCTTTCGAAATGACCGACGCGACGC
CCAAACCGGTATCAGCTCACTCAAAGGCGGTAAATACGGTTATCCACAGAATCA
GGGGATAACGCAAGGAAAGACATGTGAGCAAAAGGCCAGG
AACCGTAGAAAAGGCCCGTCTTGCGGCTTTTCCATAGGCTCCGCCGCCCCCCTGAA
CGAGCATCACAAAATTCAGACGCTCAAGTCAAGGTTGGCGAAACCCGACAGG
ACTATAAGATACCCAGGCTTCTCTTCCCCCCTTGGAAAGCTCCCTCGTGCTCTCGT
TTCCGACCCCTGCGGCTTTACCCGATACCTGTCCGCCTTTTCTCCTTCCGGAAGGC
GTGGCGGCTTTTCTCATAGCTCAGCCTGTGAGGTATCTCAGTTCGTTGGTAGTGCTG
TTCGCTCCAAGCTGGGCTGTGTCGACGAAACCCCCCGGTCCGACGCGACCGCGTGC
GCCTTATCCGGTAAACTATCGTCTTTGAGTCCAACCCGGATAGACAGAGACTTATC
GCCACTGCGACGCACCCTTTGTAACAGGGATAGCAGACGGAGGGTATCTGAGGC
GGTGCTACAGAGTTCTTGGAAGTGGTGTCGCTACTACGGCTACACTAGAAGGA
CAGTATTTGGTATGCTCGCTCTGTAAGGCCAGTTACTCTCGGAAAAAGAGGT
GGTAGCTCTTGGATCCGGAACAACAAACCCACCGCTGGTACGGGTGTTTTTTTGT
TTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATTTTCAAGAAGATCCTTTG
ATCTTTTCTACGGGTCTGAGCAGTCGTGGAACCGAAAACCATTCTACGTCTCTTATAGGCC
TTTTGTCATGAGATTTATCAAAAGGATTTTACCTACCTAGATCCTTTTATAGTG
GGAAATACAGGAACGCACGCTGGATGCGGCTTCGTGGATGGGTGAAACCATT
GAATATGGCAGCTTCAGTGAAATAGTGATGTTGGAATATGGGTGCTAACCCTC
TGGTTGATAGGTATTTTCGTTAAAATTTTATCACGGCGGATTTGCGGAGT
TTTTCGCGGTGTTTTGTTGCCATTCTACCTGTCTGCTGCCGTGATCGCGCTGAA
CGCGTTTTAGCGGTCGTCAATAATTATGGGATTATGGTAATCCACTTACTGTC
TGCCCTCGTAGCCATCGAGATATAAACCCGCACTACCGCCCACGATCGTCCG
GCGTAGAGGAGATCGAGATCT

>pENTR2B/TREPitt::Crimson/P2A-Control
>pENTR2B/TREPitt::Crimson/P2A-Control

CTTTTCTGCTTTATCCCCGTATAATCGTGGATAAACGCTATTACCGCTAGCATGG
ATCTCGGGGGAGCTCTAACTAAGCGAGGAGTGGGAAGTGCCAGGCAATCAA
ATAAAAACGAAAGGCTCAGTCGGAAGACTTGGGCTTTTCCGTTTTATCTGTTGTTT
GTCTCGTGAACGCTCTCTTAGGACAAATCCGCCGGGAGCAGGAGATTTGAAC
GTCTCGTGAAGCAACCGCAGGGGCTGCGGCGAGGACGCCGCCCATAAACT
GCCAGGCATCAAATAAAGCAGAAAGCCCATCTGACGGATGGCCTTTTGCGT
TTCTACAAACTCTTCCTGTAGTTAGTTACTTAAGCTCGGGCCTGAAATAATG
ATTTTATTTGGACTGTAGTGACCTGTTGTTGCAACAAATTGATAAGCAATG
CTTTTTATAATGCCAACKTTTGTACAAGAACGACGGCTGCGCCGGGAAACCAAA
TTCAgatatcAAGTGCCACCTGACGCTCTCCCTATCAGTGTAGAGAAGTCGACAC
GTCTCGAGCTCCCTATCAGTGATAGAGAAGGTACGTCTAGAACGTCTCCCTATCAGTGATAGAGAAGTCGACACGTCTCGAGCTCCCTATCAGTGATAGAGAAGGTACCCCCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAGACGCCATCC
ACGCTGTTTTGACCTCCATA
GAAGACACCGGGACCGATCC
AGCCTGAGCTA
GGAGGTCCGCCACCATGGATGCACTGAGAACGTCATCAA
GCCTTCATGCG
CTTCAAGGTGCACATGGAGGGCTCCGTGAACGGCCACACGAGTCAGTCCGTGAGATCGAG
GGCGTGGGCGAGGGCAAGCCCTACGAGGGCACCCAGACCGCCAAGCTGCAA
GTGACCAAGGGCGGCCTGGCCCTTCCGCTGGAGATCATCTGTGCCTCCCCAGTT
CTTCTACGGGCTCAGACATCAAGGCACCCCGGCGACATCCTCCCCGACTAC
AATCAGAATAACGGTTTGTGATGCGAGTGATTGATGACGAGCG
TAATGGCTGGCCCTGGTGAACAAAGTCTGGAAAAGAAATGCAATAACCTTTTGCCA
TTTCACCCGATTTCAGTCGTCAGCTCATGTTGATTCTTCACCTTGATAACCTTATT
TTTGACGAAGGAATATTAATAGGGTTGATTGATGGTGACGAGCTCGGAATCG
CAGACCAGATCCAGATTCTTGGCATTCTATGGAAACTGCTCGGATAGTTTTTCT
CCTTCAATACAGAAACGGCTTTTTTCAAAAAATATGGTATTGATAATCTCGATAT
GAATAAAATTGCAGTTTTTCATTGTAGCTGACGATTGGTTTTCTAATCGAAATGG
TTAATTGGTGTGTAACATTATTAGATGGGCCCCGTTCCACTGAGCCGTCAGAC
CCGTTAGAAAAAGATCAAAGGATCTTCTTTGAGATCTCTTTTTTTCTGCGCCTGAAT
CTCGTCCTGCAAAACAAAAAACCACCGCTACCAGCGGTTGTTTTGGTGCCG
GATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCACAGGCG
AGATACAAATACTGTTCTTCTTAGTGATGAGCCGTAGTTAGGCCACCACACTTCAAG
AAGCTCGTACGCACCGCCTACATACCTGCTCTGCTAATCTGGTTACAGGTCG
TGCTGCCAGTGCGATAAGTCGTTCTATCCGGGTAGGACTCAAGACGATAG
TTACCCGATAAGGCAGCAGCGCTCGGCTGACCTGAGCTGAGGATACCCGGAAATG
CCAGCTGGAGCGAACGACCTACACCCGAACGTGACATACCTACAGCGTGAGCT
ATGAGAAAAGCACCACGCCCTTCCGAAGGGAGAAAGCGGACAGGAGTGATATCCGCT
AAGCGGCGAGCGTGACGCAAAACGGAGAAGCGCGAGCCAGGAGGCCCGAGGCTCCAGCGGAA
CCGCTGGTAATCTTTTATAGTCTGCTGCTGCTTTGCCAACCTCTGACTTGAGCTG
GATTTTTGTGATGCTGTCGAGGCGGCGAGGCTATGGAAGATCACCCAGGCA
CGCGGCGCTTTTACGGTTTCTGGCTTTTTCGTGCTACATGTT
>pSLIK/TREPitt::Crimson/P2A

Sequence verified from 2770 to 3963 bp.

>3.pSLIK-Zeo/TREPitt::Crimson/P2A-Control

GTGCACGGATCGGGAGATCTCCCGATCCCCCTATGGTGCACTCTCAGTACAAT
CTGCTCTGATGCGCCGATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTTGG
AGGTCGCTGAGTACTAGTGCAGCAGAGCAAATTTAAGCTACAACAAGGCAAGGCTT
GACCGACAATGCAAGAATCTGCTTAGGTTAGGGCAGGCTTGGGCTGGCTTCC
GCGATGTACGGGCGATATACGCGTTGACATTGATTTGACTAGTTATTAA
TAGTAATCAATACGCGGCTAGTTCTAATAGCCATATATGGAGTTCCGCGT
TACATCTACTACGTTCCATAGTACGCTGCTGCCCAACTGAGGACTTT
CCATTGACGTCAATGGGGTGAATTATTTACGCTAATAGGCAACTGACCCACTTTGCGAGTA
GAAGCAGTGGAATAGGAGCTTTTGTTCCCTTGGGTTCTTGAGGAGCAGCAGGAAGGCAGTGGAATAGGAGCTTTTGTTCCCTTGGGTTCTTGAGGAGCAGCAGGAAG
CACTATGGGCACCTAATGACGCTGACGGGCTACAGCCACCTGCTCCAGGCAAGAT
TCTGGTATAGTGCAACGCACGAAACAAATTTGCTGAGGGCTATTGAGGCGCAAC
TAATAAATCTCTGGAACAGATTTGGAATACACACGACCTGGAGATGGAGTGGGAC
AGAGAAATTAAACACAAAGCTGTTAAATACACTCCTTAAATTGGAAGAATCGC
AAAACCAGCAAGAAAGAAATGAACACTTGAATTTGAATTAGATATAATGGG
CAAGTTTTGTGGAATTGTTAAACATAACAAAATTTGCGTGTGGATATATAAAAATTA
TTTATAATGATAGTAGAGGCTTGGAGTGGTTTTAAAGAATAGGTTTTTGCTGACT
TTTATAAGTAGAGTTAGGCGAGGATATTCCACCATTATCGTTTCAGACCC
ACCTCCCAACCCCCGAGGGGACCCGACAGGGCCCAAGGAATAGAAGAAGAG
GTGGAGAGAGAGACAGACAGATCCATTGCATTAGTGAAACGGATCGGCAC
TGCGTGCGCCCATTTGCAGACAAATGGCAGTTACATCCAATTTTAAAA
GAAAGGGGGGATTGGGGGATCAGTGACGGGAAAGAATAGTAGACATAATTAGCAACAGACATACAAACTAAAGAATTACAAAAACAAATTACAAAAATTC
AAAATTTTGCGGTTTTATTACAGGGCAGCAGATCCAGTTGGTGTTAATTCTA
GACAGTTATAAACAAGTTTTGTAACAAAAAGCTGAACGAGAAACGTAAAAATG
ATATAAATATCAATATATTAAATTAGATTGGTGATAAAAACAGACTACATA
ATACTGTAACACACAACTATCCAGTCACTATGGGCACCCGAAACCAATTCAG
atcAAGTGCCACCTGACGTCTCCTCCTATCAGTGATAGAAGATCGACACGTCTC
GAGCTCCCTATCACTGATAGAGAAGGTACGTCTAGAAGCTCAGGCTCCCTATCAGT
GATAGAGAAGTCGACACGTCTCGAGCTCCCTATCAGTGATAGAGAAGGTACG
TCTAGAACGTCTCCCTATCAGTGATAGAGAAGTCGACACGTCTCGAGCTCCCT
ATCAGTGATAGAGAAGGTACG
TCTAGAACGTCTCCCTATCAGTGATAGAGAAG
TCGACACGTCTCGAGCTCCCTATCAGTGATAGAGAAGGTACG
TCTAGAACGTCTCCCTATCAGTGATAGAGAAGGTACG
AGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAGACGCCATCCACGCT
GTTTTGACCTCCATAGAAGACACCGGGGACCAGATCCAGCTGGATCTAGGAGG
TCCGCCACCATGGATAGCCTAGAAACGTCATCAAGCCCTTCATGCCTCCTCA
AGGTGCACTGGAGGGGCTCCGTGAACCGGCGACAGGCTACGAGGAGGT
GGGCAGGGCAAGCCCTACGGAGGGCACCCAGACCCGGAAGCTGCAAGTGAC
CAAGGGCGGCCCCCTGCCCTTGCGGATCTCCTGGATGCTCCCCCATTTCTCTTCT
ACGGCTCCAAAGGCGTACATCAAGCAACCACCAGGCGCATCCCCGACTACCTCAA
GCAGTGCTCTCCCCGAGGGCTTTCAAGTGGGAGCGCCTGGATGAACTTTCGAGGAC
GGCGGCGTGTGACCCGTCGACCAGGAGACTCTCTCCTCAGACGGACGACCCCTCA
TCTACCACGTGAAGTTTCACTCGGCCTGAACCTTCTCCCCTCCGACCGCCCCGTATAG
CAGAAGAAGACTCTGGGTGGAGGATGCCCTCCACTGAGCGCAACTACCCCCCCG
ACGGCGTGTGCAAGGCGAAGACAGACGAGCTAAGGAGGGCGGC
GCCACCTACCTGTGTGAAGTTTCAAGTCCCATCTACATGGCAAGGCCGCTGAA
GCTGCCCCGCTACCACTACGTGGAACCTACAGCTGCAGATCACTACCTTCCCCACAC
GAGGACTACACCCGTGTTGGAGCAGTACGAGCGCCGCGAGGCCCAGGACCACAC
CTGTTCAGGGATCCCGAGCCACAGAACTTCTCTTCTGTAAAAGCAAGCAGGAG
ACGTGAAAAGAAAAACCCGGGTCTCCATCGAATATTTGACATTTATATGGTCTCCTTGGCTC
CATTTGGGTCAATATGTAATTTTCAGTGTTAGACTAGTAAATTGTCCGCTAAA
TTCTGCGCCGTTTTTGGGCTTTTTTGTTAGACGAAGCTTGCTTACCGAGCTCGGAT
CTCCACCCCGTACCGGTCTCGCAGTGAATTCCACCATGCTAGACTGACTAGAA
GAGCAAAGTCATAAAGCGAGCTCTGGAAATTACTCAATGTTGCTGGTATCGAA
GGCCTGACGACAGAAACCGGACCCCGATCTGCTGCTGCTGTGGCTGCAGAC
TGACTGTCAGGCTCATACCACTTACTTCGTCGCCCTGCGAGCGGAGCTATGG
CAAGACTTTCTGCGGAACAAACGCGCCAGATCGATACCCGGCGCCAGCGAGAC
GAAAACCTGGAAATGACGCTCGGTTCCTGCTGTACGAGCAAGGCTTCTCCCTGG
AGAACGCACTGTACGCTCTGCTGGCCCGGTGCGCGCTGCGAGGCTGCGT
ATGGAGGAACAGGAGCATCAAGTAGCAAGAAAGAGAACAGTGACCTAC
CACCAGATACGTGCATCTCGGCTGAAACTAATCATGATGAGCTGGCTC
AGGGAGCCGAACCTGCCCTTCTTTCGGCCGTGAACTAATCATGATGAGCTGGCTC
GAAACAGCTAAAGTGCGAAAGCCGGCCGACCGACGCCCTTGACGAG
TTTGACTTAGACATGCTCCCAGCCGATGCCCTTGGACACTTGGCTTGATAT
GCTGCTGCTGACGCTCTTGAGATTTTGACCTTGACGACATGCTCCCAGGTAAC
TAAGTAAGGATCCTCGGCGCCGCACTAGAGGAATTCCCGCCCCCTCCCTCCTCCCCC
CCCCCTAACGTATTACTGGCCGAAGCCCGCTTGGAATAAGGCCGGGTGTGTGTGTG
CTATATTTTATTTCCACCATATTGCGCCTTTTGGCAATGTGAGGGCCCGGGA
AACCTGCGCCCTGCTCTTGTGACGACATTCTAGGGGTCTTCCCCTCTCGCC
AAAGGAATTGCAAGGTCTGTGTAATGTCTGTAAGGAAGACAGCTCTTGAGAAG
CTTCTGGAAGCAAAAAACGCTCTGCTAGCGACCCTTTTCAGGCGAGCCGGAAACCC
CCCACCTGGGCAACAGGTCCTCTCGGGCCAAAAGCCACGTGTTAAGATACA
CTGCAAAGGCCACCAACCCCATGTCACCAGTTTGAGTTGGATAGTTTGG
AAAGAGTCAAATGGCTCTCCTCAAGCGTAGTCAACAAGGGGCTGAAGGATGC
CCAGAAGGTACCCCATTGTATGGGATCTGATCTGGGGCCTCGGTGCACATG
CTTTACATGTGTTTAGTGCGAGGTATTTTGGAAAAAAAAACGTCTAGGCCCCCGAACCAC
GGGGACGTGTTTTTCTTTTGAACACAGATGATAAGCTAACCGGTACGCGTG
GATGTTGTAACATTAATCTATCGGACATGTATATCTGCGATAGTATAATACGA
CAAGGTAGGAACTAAAAACGAGGGCAGTTGACCAGTGCGTTCCGGTGCTC
ACCGCGCGCGAGCTCGCCGGAGGCGGTGAGTTCTGGACCGACCGGCTCGGGT
TCTCCCGGGACTTCTGTTAGGAGCACTTCTCCGCGCTGAGTCCGGGACTGCGTG
GACCCTGTTCATACGCGCGTCCAGGACCAGGTGGTGCAGGCAACACACCTG
GCCTGGGCTGTTGGGTGCGCCGGCTGAGCTGTACGCGAGTGTTGGTGCGAGG
TCGTGTCCACAACTTCCGGGACGCCTCCGGGCCGGCCATGACCGAGATCGG
CGAGCGAGCGCTGGGGGCGGAGTTCCGCTGCGCATCAGGACCCCGCGCACTGC
GTGCACTTTCGTCGGCGAGGACAGCTGCAGACATCTGTACAAAGTAAAGCGG
CCGCGACTCTAGATCATAAATCGCCATACACATTTGTAGAGGTTTTACTTGC
TTTAAAAACCTCCCACACCTCCTCCCTGAAACCATGAAAAATGAACTGA
ATTGTTGGTTTATGTCCTTTATCAGCATAACAAGCTTTATCGATAATCAAC
CTCTGGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTGCT
CCTTTTACGCTATGTGGGATACGCTGTCTTTAATGCTTTGTATCATGCTATTGCT
TCCCGTAGGACGCTTTTCTCCTCTCAATTTCTTATTGCTAAAATCTTGTGCTGTCTCTT
ATGAGGAGTGGCCTTTTGTCAGCAACGCTGCGGTGCTGACTGTTGTT
GCTGACGCAAACCCACACTGGTTGGGCGCATGTGCAACATCTGGTCTCCTTTT
GTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCT
GGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCT
CTTTTCGCTTTCTTCTTCTTCTCAGCCACGTTCCGGGCTTTCCGGGCTAA
GCTCTAAATCGGGGGCTCCCCCTTTAGGGTCCGATTTAGTGCTTTTACGGCAGCT
CGACCCCCAAAAACTTGATTAGGATGATGTCACGTAAGTGAAAGCCATCGCCC
TGATAGACGGTTTTTCGCTTTACGGGAGGATCAGTCTTTAACATG
ACTCTTGTTCCTCAAACTGGAACAAACTCTAACTTCTCTGCTCTTTAATTG
ATTATAAGGGATTTTGCCGATTTCCGGCCTATTGCTTTAAAAATGAGCTGATT
TTACAAAAATTTAAGCGAGTTAATATCTGTGGAATGGTGGTAGTACGGGAGGTG
GGAAAGTGCTCCCAGGCTCCCCACGAGAGGAAGATATGCAAAGCATGCATCTCA
ATTAGTCAGCAACCAGGTGGAAGAGTTCCTCAGGCTCCCAGCAGGCAAGAG
TATGCAAGATGCACTCTCAATTAGTCAGCAACCAGTCTCTGGGCTACACTC
CGCCCATCCCGCCCCCTAACCTCCGGCAGTTCCGGCCCCATTCTCCGGCCCAGGC
TGACTAATTAAAAATTTATGCAAGAGGCCAGGCGGCGACTCTGCTCTTGAGCT
ATTCCAGAAGGTAGTGGAGGAGGTTTTTTTGAGGCTAGGCTTTTGCAAAAAG
CTCCCAGGAGGCTTGTATATCTCTTTCGGATCTGATCAGCACGTGTTGACAAT
TAATCATCGCCATAGTATATCGGCTAGTATAATACGAAGATTGAGGAACCT
AAACCATTGGAAGCTTGACCAGTGGCCGTTCCGGTGCTCACCGGCCG CGCACGT
CGCCGGAGCGGTCCGGATCTGGCAAGCCGGCTCCGGGCTCCGGGACTCTTC
GTAAGGGGACAGCTTCCGGGTGTGGGTCCGGGACGACGTGACCCTGTTCATCA
CGCCGGTCCAGGACAGGTGGTGCGGGGACACACACCTGGCCTGGGTGGGT
GCGCGGCTGGACGAGCTGTACGCCGAGTGCTGGAGGTCGTGCTCCACGAAC
TTCCGGAGCCTCGGCGGCCAGATGACGAGATCGGCGAGCAGCCGTGGGC
GGCGGGAGTTGCCTCGGCACTGGCCGACCCGGCCGCAACTGCGTGCACCCTTCGTCG
CGAGGAGCAGGACTGACACTGTCAGGAGATTCTCGAATTCACCCAGCCGCTTC
TATGAAAGGTTGGGCTTCGGAATCGTTTTTCGCGGAGCAGCCGGCTGGAGATCCT
CCAGGCAGGGGATCTCATGCTGGAGATTCTTCGCCCCACCCCAACTGGTTATTG
CAGCTTATAATGGTACAAAATAGAATAGCATCAAAATTTTCACAAATAA
AGCATTTTTTTCACTGCTATTCTAGTTGTGGTTGTCTACAAACTCATCAATGTATC
TTATCATGTCTGTATACCGTCGACCTCTAGCTAGCTAGCTTAGGCGTAATCATGGT
CATAGCGTCTTCTGTGTAATGGTTATACCCGCTCAAAACTTCCACACAAATA
CGAGCCCCAGCATTTAGTTAGCTGGGAGCTGAATGAGGTGCAAGCTAAAC
TCACATTAAATTTCGTTGCGCTCTCGCTCCCGCTTTCCAGTGCGGGAAACCTGTCG
TGCCAGCTTCATTGAAATCAGCAGAGCGCCGAGAGGGCCGTTGATGCTCAAGCTGTCG
TTGGGCGCTCTTCCGCTTCTCGCTCACCTGACTGACTCGCTGCTCGCTGCTCGGGCTG
CTCGGGCGAGCGGTATCAGCTCACTCAAAGGCCTGTAATACGGTTATCCAGT
AATCAGGGGATAACGCGAAGGAAGAACATGTAAGCAGGAAAAAGGCGAGGCAAGG
CCAGGAAACGTAAAAAGGCCGCGTGCTGGCGTTTTCTCATTAGGCCTCGGCCC
CCCTGACGACGCTCAAAAAATCGACGCTCAAGTCAGAGGTGGCAAGCCGGCC
GACAGGACTATAAGATACCAGCAGTTTCGTCTAAGTCCAGAGGTGGCAGAAAACC
TCTCCTGTACTCCGACCTCGCTCTACCGGATACCTGTGCCGCTTTCTCCTTCG
GGAAACGGTGGCGCTTTTCATAGCTAGCTAGCTAGGTATCTCAGTTCGGGTGA
GGTCCTCGCTCAAGCTGGCTGTGCACGAACCCCTCGCTTCAGCCCATG
CGCTCGCCCTATCCGGGTAACATCGTCTTGAAGTCCACACCCCGGTAAGACACG
Plasmids used in construction of pSLIK4

>pFN22K::ICP8-P2A/Crimson

CCAAAATCAACGGGACTTTCCAAAATGTGTAATAAACCCCGCCCCGGTACG
CAAATGGGCCTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTGGTT
TAGTGAACCGTCAGATCAGATAGTTAATACGACTCAGTTAGGGTACGA
TTAGGTGACACTATAGAATAAGGAAGAATTGCAATGACGAAATCGGTACT
ACCAGACCGACGTGCCCTGCAACCTGTGCACCTTCGACACGCGCCACGCCTG
CGTACACACGACGCTCATGCGCTTCGCGGACCATCCAAAGTTCGCCAGC
GCCGCCCGCGGAGCCATCGGCGTTCCGGACCATGAACAGCATGTACAGCG
CGAGCAGCTCATGCGCAACTCGAGACCCCTGCACTACGTGGACCAGCAGCGG
CCCACTGCGCCATGGGGCGGCTGGA
GACCATCATCACCAACCGCGAGGCCCTGC
ATACGGTGGTAACACGCTAGGCAGGTCTAGTGGAACCGCGAGGTGGACGC
TGATGCGAACCCTGGTGAGGGAGGAACCTCAAGTTTCGCGACGGTGAGG
CGAGGCAACACAGCAGCCATGTCCCTGAGCGCTGGACCCGTACGCGTGCGG
CCACGGGGTGTTCGCCGGGCAGTCGGTCAGGGG
CGCAACTTTCGCAATCAATTCACCCGCTTGCTGCGGCGCGCGCTGGAC
TGTTTAACACGGGTTTCTGCTCGGCGCAGTGCCACCTCGCGCGATGGAGG
GAGAGCCGGCTATCTCGCGCAGCTAAGCGCGTGAGGCGCTGGGAC
CGAGACCAACCAGCCATGTCCCTGAGCGCTGGACCCGTACGCGTGCGG
CCACGGGGTGTTCGCCGGGCAGTCGGTCAGGGG
CGCAACTTTCGCAATCAATTCACCCGCTTGCTGCGGCGCGCGCTGGAC
TGTTTAACACGGGTTTCTGCTCGGCGCAGTGCCACCTCGCGCGATGGAGG
GAGAGCCGGCTATCTCGCGCAGCTAAGCGCGTGAGGCGCTGGGAC
CGAGACCAACCAGCCATGTCCCTGAGCGCTGGACCCGTACGCGTGCGG
CCACGGGGTGTTCGCCGGGCAGTCGGTCAGGGG
CGCAACTTTCGCAATCAATTCACCCGCTTGCTGCGGCGCGCGCTGGAC
TGTTTAACACGGGTTTCTGCTCGGCGCAGTGCCACCTCGCGCGATGGAGG
GAGAGCCGGCTATCTCGCGCAGCTAAGCGCGTGAGGCGCTGGGAC
CGAGACCAACCAGCCATGTCCCTGAGCGCTGGACCCGTACGCGTGCGG
CCACGGGGTGTTCGCCGGGCAGTCGGTCAGGGG
CGCAACTTTCGCAATCAATTCACCCGCTTGCTGCGGCGCGCGCTGGAC
TGTTTAACACGGGTTTCTGCTCGGCGCAGTGCCACCTCGCGCGATGGAGG
GAGAGCCGGCTATCTCGCGCAGCTAAGCGCGTGAGGCGCTGGGAC
CGAGACCAACCAGCCATGTCCCTGAGCGCTGGACCCGTACGCGTGCGG
CCACGGGGTGTTCGCCGGGCAGTCGGTCAGGGG
CGCAACTTTCGCAATCAATTCACCCGCTTGCTGCGGCGCGCGCTGGAC
TGTTTAACACGGGTTTCTGCTCGGCGCAGTGCCACCTCGCGCGATGGAGG
GAGAGCCGGCTATCTCGCGCAGCTAAGCGCGTGAGGCGCTGGGAC
CGAGACCAACCAGCCATGTCCCTGAGCGCTGGACCCGTACGCGTGCGG
CCACGGGGTGTTCGCCGGGCAGTCGGTCAGGGG
CGCAACTTTCGCAATCAATTCACCCGCTTGCTGCGGCGCGCGCTGGAC
TGTTTAACACGGGTTTCTGCTCGGCGCAGTGCCACCTCGCGCGATGGAGG
GAGAGCCGGCTATCTCGCGCAGCTAAGCGCGTGAGGCGCTGGGAC
CGAGACCAACCAGCCATGTCCCTGAGCGCTGGACCCGTACGCGTGCGG
CCACGGGGTGTTCGCCGGGCAGTCGGTCAGGGG
CGCAACTTTCGCAATCAATTCACCCGCTTGCTGCGGCGCGCGCTGGAC
TGTTTAACACGGGTTTCTGCTCGGCGCAGTGCCACCTCGCGCGATGGAGG
GAGAGCCGGCTATCTCGCGCAGCTAAGCGCGTGAGGCGCTGGGAC
CGAGACCAACCAGCCATGTCCCTGAGCGCTGGACCCGTACGCGTGCGG
CCACGGGGTGTTCGCCGGGCAGTCGGTCAGGGG
CGCAACTTTCGCAATCAATTCACCCGCTTGCTGCGGCGCGCGCTGGAC
TGTTTAACACGGGTTTCTGCTCGGCGCAGTGCCACCTCGCGCGATGGAGG
GAGAGCCGGCTATCTCGCGCAGCTAAGCGCGTGAGGCGCTGGGAC
CGAGACCAACCAGCCATGTCCCTGAGCGCTGGACCCGTACGCGTGCGG
CCACGGGGTGTTCGCCGGGCAGTCGGTCAGGGG
CGCAACTTTCGCAATCAATTCACCCGCTTGCTGCGGCGCGCGCTGGAC
TGTTTAACACGGGTTTCTGCTCGGCGCAGTGCCACCTCGCGCGATGGAGG
GAGAGCCGGCTATCTCGCGCAGCTAAGCGCGTGAGGCGCTGGGAC
CGAGACCAACCAGCCATGTCCCTGAGCGCTGGACCCGTACGCGTGCGG
CCACGGGGTGTTCGCCGGGCAGTCGGTCAGGGG
CGCAACTTTCGCAATCAATTCACCCGCTTGCTGCGGCGCGCGCTGGAC
TGTTTAACACGGGTTTCTGCTCGGCGCAGTGCCACCTCGCGCGATGGAGG
GAGAGCCGGCTATCTCGCGCAGCTAAGCGCGTGAGGCGCTGGGAC
CGAGACCAACCAGCCATGTCCCTGAGCGCTGGACCCGTACGCGTGCGG
CCACGGGGTGTTCGCCGGGCAGTCGGTCAGGGG
CGCAACTTTCGCAATCAATTCACCCGCTTGCTGCGGCGCGCGCTGGAC
TGTTTAACACGGGTTTCTGCTCGGCGCAGTGCCACCTCGCGCGATGGAGG
GAGAGCCGGCTATCTCGCGCAGCTAAGCGCGTGAGGCGCTGGGAC
CGAGACCAACCAGCCATGTCCCTGAGCGCTGGACCCGTACGCGTGCGG
CCACGGGGTGTTCGCCGGGCAGTCGGTCAGGGG
CGCAACTTTCGCAATCAATTCACCCGCTTGCTGCGGCGCGCGCTGGAC
TGTTTAACACGGGTTTCTGCTCGGCGCAGTGCCACCTCGCGCGATGGAGG
GAGAGCCGGCTATCTCGCGCAGCTAAGCGCGTGAGGCGCTGGGAC
CGAGACCAACCAGCCATGTCCCTGAGCGCTGGACCCGTACGCGTGCGG
CCACGGGGTGTTCGCCGGGCAGTCGGTCAGGGG
CGCAACTTTCGCAATCAATTCACCCGCTTGCTGCGGCGCGCGCTGGAC
TGTTTAACACGGGTTTCTGCTCGGCGCAGTGCCACCTCGCGCGATGGAGG
GAGAGCCGGCTATCTCGCGCAGCTAAGCGCGTGAGGCGCTGGGAC
CGAGACCAACCAGCCATGTCCCTGAGCGCTGGACCCGTACGCGTGCGG
CCACGGGGTGTTCGCCGGGCAGTCGGTCAGGGG
CGCAACTTTCGCAATCAATTCACCCGCTTGCTGCGGCGCGCGCTGGAC
TGTTTAACACGGGTTTCTGCTCGGCGCAGTGCCACCTCGCGCGATGGAGG
GAGAGCCGGCTATCTCGCGCAGCTAAGCGCGTGAGGCGCTGGGAC
CGAGACCAACCAGCCATGTCCCTGAGCGCTGGACCCGTACGCGTGCGG
CCACGGGGTGTTCGCCGGGCAGTCGGTCAGGGG
CGCAACTTTCGCAATCAATTCACCCGCTTGCTGCGGCGCGCGCTGGAC
TGTTTAACACGGGTTTCTGCTCGGCGCAGTGCCACCTCGCGCGATGGAGG
GAGAGCCGGCTATCTCGCGCAGCTAAGCGCGTGAGGCGCTGGGAC
CGAGACCAACCAGCCATGTCCCTGAGCGCTGGACCCGTACGCGTGCGG
CCACGGGGTGTTCGCCGGGCAGTCGGTCAGGGG
CGCAACTTTCGCAATCAATTCACCCGCTTGCTGCGGCGCGCGCTGGAC
TGTTTAACACGGGTTTCTGCTCGGCGCAGTGCCACCTCGCGCGATGGAGG
GAGAGCCGGCTATCTCGCGCAGCTAAGCGCGTGAGGCGCTGGGAC
CGAGACCAACCAGCCATGTCCCTGAGCGCTGGACCCGTACGCGTGCGG
CCACGGGGTGTTCGCCGGGCAGTCGGTCAGGGG
CGCAACTTTCGCAATCAATTCACCCGCTTGCTGCGGCGCGCGCTGGAC
TGTTTAACACGGGTTTCTGCTCGGCGCAGTGCCACCTCGCGCGATGGAGG
GAGAGCCGGCTATCTCGCGCAGCTAAGCGCGTGAGGCGCTGGGAC
CGAGACCAACCAGCCATGTCCCTGAGCGCTGGACCCGTACGCGTGCGG
CCACGGGGTGTTCGCCGGGCAGTCGGTCAGGGG
CGCAACTTTCGCAATCAATTCACCCGCTTGCTGCGGCGCGCGCTGGAC
TGTTTAACACGGGTTTCTGCTCGGCGCAGTGCCACCTCGCGCGATGGAGG
GAGAGCCGGCTATCTCGCGCAGCTAAGCGCGTGAGGCGCTGGGAC
CGAGACCAACCAGCCATGTCCCTGAGCGCTGGACCCGTACGCGTGCGG
CCACGGGGTGTTCGCCGGGCAGTCGGTCAGGGG
CGCAACTTTCGCAATCAATTCACCCGCTTGCTGCGGCGCGCGCTGGAC
TGTTTAACACGGGTTTCTGCTCGGCGCAGTGCCACCTCGCGCGATGGAGG
GAGAGCCGGCTATCTCGCGCAGCTAAGCGCGTGAGGCGCTGGGAC
CGAGACCAACCAGCCATGTCCCTGAGCGCTGGACCCGTACGCGTGCGG
CCACGGGGTGTTCGCCGGGCAGTCGGTCAGGGG
CGCAACTTTCGCAATCAATTCACCCGCTTGCTGCGGCGCGCGCTGGAC
CCAAGGCGTACATCAAGCACCACCCCAGACATCCCCGAGCTACCTCAAGCAGTC
CTTCCCCGAGGGCTTTCAAGTGGAGCAGCGTGATGAACTTCGAGGACGCGGCGGC
GTGGTGACCGTGACCAGGAATCCCTCTCCCTGAGGACGGACCCCTCTCATCTACC
ACGTGAAGTCTCATCGGCTGAATCTCCCCTTCGCCAGCGCCCGTAAATGCAGAA
GAAGACTCTTCGGCTGGAGCCTCCACTGAGGCTACACTACCCCGGCGCCGACGMC
GTGCTGAAGGGCCAGAACCACATCGCGCTGAAGCTGAAAGGCCGGCGACGCAC
TACCTGTGTGAGTTCAAGTCCATCTACATGGCAAAGAAGCCCGTGAAAGCTGC
CCGCTACACACTACGGACTACAGCTCGACATCACCTACCTACCACAACGAGGA
CTACACCGTGGTGGAGCAGTACGAGCGCGCAGGCGGAGGCCACCCACCACCTGTC
CAGTAGAAACGATTCGGGCTCGGTACCCGGGATCTCTTAGAGTCGACCTG
CAGGCGATGGAAGCTGATCCCAAAGCTGGGGAGCTCCACTGAGCGCAACTAC
GGCTGCTGAAGGGCGAGAACGGCTGAAAGCTGAAAGCTGC
GTGCTGCTGCAAACGCTGACGATCATTGAGTTTGGACAAACCACAACTACG
TCTAACCATTATAAGCTGCAAATAGTTAACAACAAACAAAATGGTAGATGTTTGGACAAAACACACAACACTAGA
ATGCAGTGAAAAAATGCTTAATTTGTAATTTGTAATTTTGATTGCTATTGTTTATT
TGTAACCATTATAAGCTGCAAATAGTTAACAACAAACAAAATGGTAGATGTTTGGACAAAACACACAACACTAGA
TTTATGGTTTCAGGTTTCAGGGAGATGGGGAACGGTTTGGTTTTTTTAAAGCAAGTAAAA
CCCTACAACATGTGGTAAATTCGAAATCTACTATATGGATCCTTTCTCTCTTGCTGCTG
TTCCCTTGCTCCAGATAGCCAGTAGCTGACATTCTACCAGGGGTACAGACCGGT
TTCTCGGAGCTGGCTTTCTCCGTCTTCTCGGCTCTCCAGGGGCAGGCCCCTG
GAGTGCTTGGGCGAGCTGGCTTCAAAGAATGGCAGCTGGGAGGCCGTGCTGCTG
TTGGTAGGTTGGAGCTGAGGCTTTCTCTTTCTTTCTTTACGACGCCCTTCGCGCC
TGGTAGGTTGGAGCTGAGGCTTTCTCTTTCTTTCTTTACGACGCCCTTCGCGCC
AGGATCTGATGAGCGCGAGGGATTAAGATCTGATCAAGAGACAGGATGACGG
TCGTTTCGCAATGCTTGAACAAAGATGGATTGGCAGCAGTTTCTCCGGCCCCTTG
CGCTGCCTTTATCCGGTAACATC>TCTTGTGAGTCCAAACCUGGTAAAGACACG
ACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAAGCAGAGCGAGGT
AGAAGGACAGTATTTGGTATCTGCGCTCTCTGGTAACAGGATTAGCAGAGCGAGGT
AAAGAGTTGGTAGACTTTTGATCCGGCAACAAACACCACCCGCTGGGTAGAGTGG
TTTTTTTAGTTTGCAAGCGACAGATTACGCAGCAGAAAAAGGATTTTCAAAGAA
GATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGAACAGAACACCACCGT
TTAAGGGATTIGGTGTATGAGATTATCAAAAAAGGATTCACCTAGATCCTTTT
TATAGTCCGAAAATACAGGAAACGCACGCAGCTGGATGGGCTGCTGGGATGGGT
GAAACCATGAAAAATGCGAGCTTCAGTGAATGAGTGGGTAATGTGGCCT
GTACCCTCTGGTGCCATAGGGTATCTATACGGTTAAAATTTATCAGGCGCGATT
GCGGCAGTTTTTCGCGGGTTTGGTTGCCCATTTTTCGCTGCTGGCCTGATC
GCGCTGAAACGCCTTTTAGCGGTGCATCAATTAAGGGGATTATGGTAAATCCA
CTTACTGTCTGCCCCGTAGCCAGATGAAAACCAGCATCTACTCCCCGACG
ATGCGTCCGCGTAGAGGATCGAGATCT
CTTTCCCTGGTTATCCCTGATTCTGTGGGATAACCCGTATTACCAGCTAGCATGG
ATCTCGGGGACGTTCTTAACTAACTTGAGTCAGGAAACTGAGAGCTAACTAGCAAA
ATAAAAACGGAAGGGCTCAAGGTCAAGACTTTGCTCTTGTATTCTGGTGTGTTGTTG
GTCGTTGAAACGCTCTCTCCTCAGGTAGGACAAATCCCGCCGGAGACGGGATTGTGAAC
GTTGTGAAGCAGCCGCGAGGGTGCGGCGCGGACGACGCAGGCCAACATTAAACT
GCCAGGCAATCAACTAAGCAAGGCAAGGCGATCTGACGGGACTTGGTCTGGTGCT
TTCTCAAAAATCTCCTCCTGAGTGGTTAGTTACCCAAGCTCGGCCCACAAATAATG
ATTTTATTTTACTGATATCAGCCTGTTGCAACAAATTGATAGCAGAAC
CTTTTTTATAATAGCAACTTGTACAAAAAAGCAGGCTGCGCCGGGACCAAAC
TTCAGATATCAAGTGCCACCTGACGTCTCCCTATCAAGTAGAGAAGTGCA
CACGTCTCGAGCTCCCTATC
CTATCAGTGA TAGAGAAGTGCAACGCCTCTCGAGCTCCCTATACGTAAGAG
AAGGTACGTCTAGAAACGTCTCCTCATATCAGTAGAAGAGTGCAGACGCAGTCTC
GAGCTCCCTATCATGATAGAGAAGGTACGTCTAGAAACGTCTCCCTATACGT
GATAGAGAAGTGCAACGCCTCTCGAGCTCCCTATACGTAAGAGTGCAGACGC
CCCTATATAAGCAAGAGCTCGTTTAGTGAACCGTCAGATCGCTGCGCTGGAACGCC
ATCCACGTGTGTTTGACCTCCAATAAGAAGACCCGAGCGAGCATCCAGCTAG
TCTCGCCATGGGAGACAAGCCGCCCACCCCCACCAGTTACAGCAGCTGCTGG
GGGCCCTTGGGATACGTGTCGCTCGCCTGCTGCTGCGCTCGAAGGACGCTG
TTCTGGGCTTTACTGTCGGGCACTGCCCGAGGCGATGCCGGAGGCGCGCC
CCTGTTGCGTGAGGCCTGAGAGAGCGGCTTTGAGGCAACAGTGAGCCGTG
GTCGTGGGTCTCAGCACGCAGGGCTCGGGGGGTACCGCGGTGCTGCTGCTG
TAGACGCCATGCACTACAGCTGCTCCGTTGACTGCTTTTACCGGGCGCGCCGC
CTGGACCCAGCAACCGAGGCCCCAAACCTGACGCACCTGCGACGGGCAG
GCCGCCATTGGGCTTTCGGACTACACCCCCCGCGCGACCTCAAACAC
GAGACGACGGGGGAGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGC
GCCCTCCTGTATCTGGTCTACCGAGGGCTTCACGGGAGGCGGTCTGCTGATCAA
CAACACCTTTCTGCACCTGGAGGGCTCGGACAAAGTAACCATAGGCGCAGGC
GAGGTGCACCCGATACCGGTCGTATTGAGGCTTGTGCAGCTTGTCCACGT
CGCCGTCGTGGGACCCGCCGC
GTTAAGGCACTGCGATGCCGAAAACCA
CGAGGGCGCCGC
GCGGGTCGTGGGACCCGCCGC
CCTGATCGCCGAGCGCCCTCAACCGCAACCCAGGATCGAGCGGGAGAAT
TTTACCTACCCCGCTCTGGGTTTTATACCACCGCCCCTCAACCAGGCTCGTGGACGC
GCGGTCTGGGACAACCGGCGGGCTGGCAGCTGCGATGCCGAAACGCTGGGACGC
TGGCCCGCGCGCGCCCGCCCACCTCGGCCTTGGACGAAAACACGAGGCGCCGC

624
625

CCTCCCCGCGACATTACGTTTCACGGCCTTCGAAGCCAGCCAGGGTAAGACC
CCGCGGGGTGGGCGCGACGGGCGGCAAGGGCCCGGCGGCGGCGGCTGGG
CAGCGCCTGGCCTCCGTCATGGCCGAGACGCCCGCTGGCCCTGCGAGTCTA
TCGTTGTGATGGCCGCTTTCGAGGACGCGCCGACCGACATCTCGCGTGCC
GCTGTGCGAGGGCCAGGACACGGCGCGGCGGCGGCGGCGGCGGCGTCCGAG
GTACCTGGCGCGCGCGCGGACTCTGTTGGGGGCATGTTATTAGCACCAACG
TCGGCCCTCCATCTCAACGAGGTTGAGACGACGGCGTCCGCCGAGGCCAAGG
ACCACAGCAACACCTCCTTTTACGCTTTCTCTCTGTCGCCCCGAGCCCACGTG
GCGGCACACCCACCAGTGACCGACGAGGACAGCTGTGTCGCCGGGCTGGAG
GGTCGGGCCCACCGCGCCCCCTCGTGCCGCGGAGAACCAGAATTTGCGCCGCGAGC
ACCTGGCCATGCTGTGTGTTGTTTCCCCGCGCCTGCTGGCCAAGATGCTGTTT
TACCTGGGAGCGCTGCACGGCGGCGTGATCGTGCGCCGGCCAGAGATGGAC
GTGTTTCGATACTGTCGCGGACTCCAACCAGACCACGCTGCCTCTGTCAACCTGT
GCACCTTCGACACGCACCACCGCTGCTACACAGACGCTCATGCGCCCTCCG
GGCGCGCCATCCCAAGTTCCGACGCGCCCGCCCGCGGAGCCATCGCGGCTTTC
GGGACCATTGAACAGCATGTACAGCGACTGCGACGTGCTGGGAAACTACGCC
GCCTTCTCGGCCCTGAAGCCGCGGACGGATCCGAGACCAGCGCGGACCACATCA
TGCGGAGACGTACCGCGCGGACCGAGCCGCGCTGACCGGGGAACTCGAGA
CCCTGCGAGTCGTGGACCGGACGCGGCTGCCCCACGCCCAGGGGGGCTGGGAGAC
CATCATCAACACCGCGAGGCCCCCTGCACTACCGGTGGGTGAAACACGTCAGGCA
GTGCTGGGACCGAGGTGGAGACGCTGTGATCGCAACCTGTGGGAGGGGAGG
AACTTTCAAGTTTCGCACGGTCTGGGCGAGGCCACACCAGCCCACGGTCTCCTG
CGCTGGACCCGCTACGCGTCCGCGGCGGCAATGCCCCCTGCTTTCACTGCGGCG
GCGATCCAACCTCGCCGTGTATCAGGACCTGGCCCTGAGCCAGTGCCTACCAGG
GTTTCCGCGGCGACGTGCTGCCAGGAGGCGGCAACTTTTCGCAATCAATTCCAAC
CGGTGCTGCGGCGGCGGTGATGGACATGTGTTAAACACCGGTTTTCTGCGGCA
CACAAACGTGACGCTCGCCTCGGAGAGGGGGCGCTATCTGCGCCGCCACG
CTAACGGCCAGACGCGCCAGGGGCTCCTGGAGAGAGGCAGTGTGGA
GTTCGCAGGGCGAGCGCAACGCAGTCCGAGGCGCCAAGGCAGCGGAGGTCG
CAGCCTCCAGAGCGCTACCAGAAGCGCACAAGCAGTGACATCGAGACC
GGACCGCTGGGCTTTCTGCTGAGCAGCTTCCAGCGGGCACCATTCTCCCCAAACG
GCAAGCCCCCGGGGTTCCAAACCAGCGCGAACCCCGACTGCTTCTGAGCGGCCCT
CCAACGCAACAGCTTTCCGCCCGGCTCCTGTCGCGCAGGACATCGAGACC
ATCGCGTTCTTAATAAGTTTTTCTGACTACGGCGCGATAAACACATTTATTAA
CCTGGCCCCAAACAACGTGAGCGAGCTGGCGATGATCACTACAGCGAAAACAG
ATTCTGCGGTACTGCGATCACTCGACATCATCTCAACACCCCCTACCGGCACT
CATCGCGGTTCCCAGCGCTCCCCCACAGCGCTGCGCACGGCGGCGCGGCTGTC
GCGCAGGGCCGGGCGCGGCTGGAAGCCGGCCGGCCGGCCGGCTGATGGCAGCC
GTGGAGCGCATCCGGCGCTGGACGCTCATGTTCCAGCTGCAACCTGCGG
TGGCGCCGTATGCGGGCCGCGCCGCCCATGGTCGCTGTGGGTTGAGCATCA
CAAATACGTCGCGATGCGGCGGCAACGACCCTGTTTCAGGCCGGGAACTTG
GCCAGCCTGATGCGGGCCAAAAACGCCTGCCCCGCTCCTTTATTTTGAAGGCA
CCCGCAAGTGGCCTCCCTGCTCGAGGGCCGGGGTTTGTGTTGCGCGGCTCG
AACCTCGCCGGCGAGCGACGAAAGCTCGCCTGTCGAGCAGCAGCTCCGGG
ATATATCCGAGGGCGGCGGCGGCGGCGGCGTCGAGGTGTGCTTCGAGGCACCG

626
TGAAAAGCCTGGGGCCCCGCACCCACGCAGCTGCAGATCGAGGACTGGGTGGGC
GCTCGGTGGAGACGAGGTACCTGAAGCGAGAGATGATGGAGCTGACCACGCCTG
GCCCTGGAGCGCGGCAACCGGCGAGTGGTCACGGACGGACCGGCCCCCTGGAGGT
GCGCAGACGGCGAGGCCCCCTAGTCAGCACAACTCGGCAACCGCCGGAGGGAGGT
TTTAACCTTTGGGGATTTTGGCTGCGAGGACGACAACGCGAGGGCCGGTCACCGG
GCCGGGGGGCCCCGGAGCGCGGCAACCGGCCCCTGGAGCGCGGCAACGGACCTG
GGATAGCTGGTGTCTCGGAGGACCGACGAATCTGCGACTGAGACGTCTCAT
CAAGCCCTTCATCGCCTTCAAGGTCACATGGAGGGCTCCCGTGAACCGGCCAC
GAGTTGGAGATCAGGGGCCTGGGCGAGGGCAAGCTACGAGGGCCACCCAG
ACCGCCAGCTGCAAGTGACCAAGGGCGGCCCTGCCCTTCGCTGGGACA
TCCTGGCCCCCATTTCTTCTACGCGCTCCAAGGCGTAGATCAAGACCCCGGCC
GACATCCCCGACTACCTCAAGCAGTCTCTTCCCCGAGGGCTTTCAAGTGGGAGC
CGTGATGAACCTTCAGGACGGCAGGGCGCTGGTGACCGTGACCCAGGACTTC
CCTCAGGAGCGACCCCTCATCTACACAGTGAAGTTCATCGGCGTGAACTTC
CCCTCCGACGGCCCCGTATGCAGAAGAAGACTCTGGGCTGGGAGCCCTCCA
CTGAGCGCACTACCCCCGCGACCGGCGTGCTGAAGGCGAGGAACCACATGG
CGCTGAAGCTGAAGGGCGCGGCCACTACCTGTGTGAGTTCAAGTCCATCTCA
CATGGCCAAGAAGCCCCGTAAGCTGCCGGCTACCACTACGCTGGAACGTCAAG
CTCGACATCACCTCCCAACAGAGGACTACACCGTGGTGAGGAGCAGTACGAGC
GCGCCAGGGCCCCGCAACCGTGCCTAGCGAAGGATTTGCATTATAAAGAAAGC
TCGAGATATCTAGACCCAGCTTTCTTGTACAAGAAGTTGGCATTATAAAGAAAGC
CGCAGATACCAAATACTGTTCCTCTAGTGTAAGCCGTAAGTTAGGCCCAPCCTTC
AAGAACTCTGTAGCACCCGCCTACATACCTCGCTCTGCTAATCCTGTACACAGT
GGCTGCTGCAGTGGCGATAAAGTCGTTGCTTTACCAGGTTTGGAAGACTCAAGACGA
TAGTTACGGGATAAGCGCAGCGGTGCTGGCTGAACGCGGTCAGGTTTACGTCAGE
AGCCCAGCTTGGAGCGAACCACCGTACACCGAACTGAGATACCTACAGCGTGA
GCTATGAGAAAGCGCCACCTCTCCCGAAAGGAGAAAGCGGACACAGGTATCC
GGTAAGCGGCAGGGTCGGAACAGGAGACGCACGAGGGAGCTTCCAGGGGG
AAACGCCTGGTATCTTTATAGTCCTGTCGGTTTCGCCACCTCTACACCCAGGGG
GTCGATTTTGTGATGCTCGTCAGGGGGCGAGGCCCCTATGGAAAAACGCCAG
CAACGCAGCTTTTTTACGGTTCTGGCCTTTTGCTGGCCTTTTGCACATGTT
>pSLIK /TREPitt::ICP8-P2A/Crimson

AKA 4-pSLIK-Zeo/TREPitt::ICP8-P2A/Crimson or pSLIK4. Sequence verified from 2410 to 7547 bp.

>4-pSLIK-Zeo/TREPitt::ICP8-P2A/Crimson

GTCGACGGATCAGGGAGATCTCCCGATCCCCCTATGGTGCACCATCTCAGTACAAT
CTGCTCTGATGCGCGCATAGTTAAGCCAGTATCTGCTCCCCTGCTTTTGCTGTGG
AGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTT
GACCGACAATTGCATGACCCAGCTATTAGGGTACCCGTGTTGGTCTGCTGCTTCT
GCGATGTACGCGCCAGATATACGCGTGACTTAGTTATTGCTAGTTATTTAC
TAGTAAATCAATTACGGGCTATTAGTATCATGTCAGCGCTGCTAGTTTTGCT
TATAAATTACGGGTAAATGGCCGCCTGGCTGGACCGCCACGACCCCCGGCC
CCATTGACGTAATAATGACGTATGTGCTTTCTAAGATGTAACAGACTAGTTTT
CCGCACCCAGCAGCTGCAGATCGAGGACTGGCTGGCGCTCCTGGAGGAGACGA
GTACCTAGCGAGGAGATGATGAGCTGACCGCGCTGCCCTGGAGACGCAGGG
CAACGGGAGTGGTGACCGACGAGCAGGCCTGGAGGTGACCTGAAGCTGAGG
GGCCCTAGTCAGCAACTCGGCAACGCAGGAGGGAGGTGTTTAACTTTGGGGGAT
TTTGACTGCGAGGACGAACACGCGACCGCTTCGGCGGCCCGGGGGCCCGGGG
GACCAGGATTTGCGGGCCGCAAAACGAGGGCGTACGACCCACTCCAGGGATGAGC
GGAGGGGCCCGACAAAAAGGGAGACCTGACGTTGGATATGCTGGTTTcc
GGATCCGAGCCACGAACTTCTCTCTCTGTAAAGCAACAGGAGAGCCTGAGGGAG
AAAAACCGGTACTATGGATAGCAGACTGAAGACGTCAAAAGCCTTCCTCATCGG
CTTCAAGGTGACATGGAGGGGCTCCGGTAACCGCCACAGGATTCGAGATCGAG
GGCGTGGGGCCAGGCAAAGCCCTACAGAGGGCCACCCAGACCCGCGCAAGCTGCA
GTGACCAAGGGGCGGCCCTGCCCTGGAGACATCCCTGTTCGCCCCCTAGTT
CTTCTACGGCTCCAAGGCGTACATCAAGCACCACGCGGACATCCCCGACTAC
CTCAAGGAGCTCTCTCCCGAGGCTCTCAAGTGGGAGCGCTGATGAACTTCCG
AGGACCGCGCGGCTGTTGAGGACCGCTGACCCAGGACTCCCTCCCTGAGGAGGCAC
CCTCATCTACCACGTAAGTTCACTCCGCGCTGAATTCCCTCCCCGACCCGGCCCG
TAAATGCAGAAGAAGCTCTGGGCTGGGAGGCCCTCCACTGAGCGCAACTACCC
CCGGAGGCGCGTGGCTGAAGGGGCGAAAACATGGCGCTGAAGCTGAGGG
CGGCAGGCCACTACCTGCGATGTTCAAGTCCCATCACATGCCAAGAAGCCC
GTGAAGCTGGGCGCTACCACTAGCTGGGAAGCTCAAGCTDCAGACATCACCCTCC
ACAACGAGGACTACACCGTGGTGGAGCGATCAGAGCGCGCGCCGAGGCCGCC
ACCACCTGTCTCAGTAAAACGAAATTCGCGCGCAGCTCAGATTATAGTACTGGTTT
AGTGACTGGATATGTGTTGTTTTACAGTATTATGTAGTCTGTTTTTTATGCAAA
CGAGTGTTTTTGTAAGTTTTTACCCCTTGGAAATGTAATCATTTGGG
TCAATATGTAATTTTCAGTGTTAGACTAGTAAATTGTCCGCTAAATTCTGGCC
GGTTTTGGCGTTTTTGTTAGACGAAGCTTGGGATACCGAGCTCGGATCTCCACCC
CGTACCGGTCTCTGCACTGAATTCAACCTAGTCTAGACTGGACAAGAGCAGAAAG
TCATAAACCGGAGCTCTGGAATTTACTCAATGGTGTTCTGATATCGAAGGCTTGAC
GACAGGAAGAACAACGCGGCGGCTGCTCGATGCCCTGCCAATCGAGATGCTGG
ACAGGACATCATACCCCATTCTGCCCCCTGGGAAGGGCGAGTCATGGCAAGACTT
TCTGCAGAACAACGCGAAGTCAATCGCTGTGCTCTCCTCTCGACACTCAGCGACG
GGGCTTAAATGCTCTCAGCCTTCTTGTTCAAGCAGCTTCTCCCTGGGAAGAACGCA
CTGTACAGCTCTGCGCCCTGGGCCACTTTACACTGGGCTGCTGTATTGGAGGA
ACAGGAGCATCAAGTAGCAAAAAAGGGAAGGAGACAGACACCTACCACCGATTCT
TATGGCCCCCCACTTCTGAGACAAGCAATTTGAGCTGTTCGACCGGCGAGGAGGCC
GAACCTGCCCTCTTTTCCGGCCTGGAACCTATCATATGTGGCCTGGAGAAAC
GCTAAAGTGCAGAAGCCGCGGGCCGCCACGGGCCCTTGGACGATTGTGACTTA
GACATGCTCCCCAGCGATGCCCCTTGACGACTTTGACCTTGATATGCCTGCCTGC
TGAGCCTCTTTGAACTTTTGAACCTTGAATGCTCCCCGGAATAACTAAGTAAGG
ATCCCGCGGCGCAGGACTAGGAATTCCGCCCTCTCCCTCTCCCCCCCCCTAAACG
TTACTGGCGGAAGGCCGTGGGAATAAGGCGGCTGTTGTGGTTGTCTATATGTTA
TTTTCCACCATATTGCGCTTTTGGCAATGGGACTGAGGGCGGCAAGCTTGCC
TGCTTTCTTGACGAGCATTCAGGGGTCTTTCCCCTCTCGCCAAAGGAATGC
AAGGTCTTGTTAATGTGTCGTGAAGGAAAGCGACTTCTCTCGGAAGCTTCTTAAAGG
ACAAACAACGTCTGTAGCGACCCTTTGCAGCCAGCGGAAACCCACCTGGC
GACAGGTGCCTCTGCGCCAAAAAGGCCACGTGTATAAGGATAACCTGCAAAAGG
CGGCCAACCCCAGTGCCACGGGTGAGTGGTGGAAAGTGAAGAATGGCTCTG
TTTAGTCGAGGTTAAACAGTCGTTAGCTAGCTAAGGCAGGATGCCCAGAAGGT
TTTCCTTTGAAAAACAGGATGATAAGCTTACCGGTACGCGTGATGTGGTGACA
TTAAATCATCGGCATAGTATATCCGACATAGTATAATACGACAAGGGTGAAGA
CTAAACCATGGCAGAGTTGACCAGTGGGCTTCCGCTGCTCACCAGCGCGCAGC
GTCGCGGACGCGTCGAGTTCTGGACGCCGACCACGGCTCGGTTCCTCCGCGGACT
TCGGGAGGCAGGACTTCCGCGGCTGTCGGACCCGAGCTGACCCCTGGTACAT
CAGCGCGGTCCAGGACAGGTGGTGCCGAGTAACACCCTGGCTGGGTATGGG
GTGGCCGGCCCGGGCTGGAGGTTCTGGACCCGAGCGGTGCCCGGTTCTCCGGGACT
TCGGGAGGACTGACACATCTGTACAAAGTAAAGGGCGCGCGACTCTACTG
GATCATAATCAGCCTACCACTACATTGGTAGAGGTTTTACTTGCTTTAAAAAACC
TCCCAACACCTCCCTGAACCTGAACATTTAATGAATGCAATTGGTTGTTGTT
TAGTCCCTCCCAAATTCGATATCAAGCCTATTGCTACCTCGGATATTACA
AAATTTGTGAAGAGATTGACTGGTTTTACTTGCTTCCTCTTACGCTAT
GTGGATAACGTGCTTTAATGCGCTTTGTATCATGCTATTGCTTCGCCTTACGCTT
TCATTTTCTCCTCGTATAAAATCCTCGTGTCTCCTTTATGAGGAGTTGT
GGCCCGTTGTCAGCGACGGCGTGGTGTTGCACCTGGTATGGTGGACGCAAC
CCCCACTGGTTGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCG
CTTTCCCCCTCCTATTGCCACAGGCAGGAACCTCATCGGCCGCTGCTTGCCCGC
TGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGTGTTGTCGG
GGAAATCATCTGCTCCTTCTCCTGGCCTGCCTGTGTTGGGCACTTGGGGATTTCTG
CGCGGGACGTCTCTTCTGCTACGTCCCTTCGGCCCTCAATCCAGCGGACCTTCC
TTCCCCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCC
CTCAGACGAGTCGGATCCTCCCCTTTTGGGCCGGCCTCCCGCATCGATACCGTCA
CCTCGATCGAGACCTAGAAAAACATGGAGCAATCACAAGTAGCAATACAGC
AGCTACCAATGCTGATTGTGCTGGCTAGAAGCAACAGAGGAGGAGGAGGT
GGGTTCCTCCAGTCACACTCTAGACTCTTTAAGAAATGACTCTAACAGACG
GCTGTAGATCTCTGAGCCACTTTTTAAAAAGAAAAGGGGGGACTGGAAGGGCTAA
TTCACTATCCTCTATGCTAGATATCCTTCTGTGACTCTGACCAACTGACACACA
AGGCTACTCTCCCTGACTTTGAGGAGACCTACACACAGGGCCAGGGATAGATAT
CCACTGACCTTTTGGATGTTGCTACAAAGCAGTAGTTGAGCAAGAGAAGG
TAGAAGAAGCCAATGAAAGGAGAGAACACCACCTGGCTTACACGCTGAGCCT
GCATGGGATGGATGACCAGGAGAGAAGATATTAGAGTGGAGGTGTTGACAG
CCGCTAGCATTTCATCAGATGGCCCGAGAGCTGCACTCCGGGACTGTACTGGG
TCTCTCTGGTATGACCAGATCTGAGCCTGGAGCTCTCTGTGACCTAGGGAA
CCCAGCTGTTAAAGCCTCAGAGCCTTCTGAGTCCTTCAAGATGTGTTGGTG
CCCGAAGCTGTTGCTGACTCTGTAAGTAGAGATCTCAGACCCCCTTTTAGCTA
GTGTGAAATACTCTAGCAGGCGGGCTTAAACCGCTGATCGACCTCGACT
GTGCTTCTAGTGGCCAGCCTACCTGTGTTCCGGCTCCCTCCCCGTTTCTCCTTGC
ACCCTGGAAGGAGGCTCGCCTCCACTGTCCTTTCTCAATAAAATGAGGAAATTGC
TTATCCGGTAACTATCGTCTTGAGTCCAAACCGTGTAAGACACGACTTTATCGCC
ACTGGCAGCAGGCACTGGTAAACAGAGATTAGCAGAGCGAGGTATGTAGGCCGT
GCTACAGAGTTTCTTGAGTGGTGGCCTAACTACCGGCTACACTAGAAGCTAAGTCTTGGT
AGCTCTTGATCCGGCAAACAAACCACCAGCTGGTAGCGGTGGTTTTTTGTTTG
CAAGCAGCAGATTACGCGCAGAAAAAGAGATCTCAAGAAGATCTTTTGAAGAGTATTTT
GGTCATGAGATTATCAAAAAAGGATCTTACCTAGATCCTTTTAAATTAAATAAAAAAT
GAAGTTTTAAATCAATCTAAAGTAGATATATGAGTTAAACTTTGTCTGCAGAGTTAC
CAATGCTTAACGTAGGACCACCTATCTCAGCGATCTGTCTATTTCTCTCATC
CATAGTTGGCTGACTCCCGTCCGTGTAGATAACTACGATACGAGGGAGGTCTTA
CCATCTGGCCCAGTGCTGCAATGATACCGAGACCCACGCTCACCGGTCC
CAGATTTATCGCAATAACACCAGGCGGAAAGGCGAGCGCAAGATG
GTCTGCAAACCTTATTCGCTCCATCCATCCAGTCTAATTATGTTGCGGGGAAAGCT
AGAGTAAGTGATTTGCCAGTAAAATAGTTTGCGCAACCTGTGTGCTGGACGTGTAT
AGGCATCGTGTTGCTACGCTCGTGTGGTTGATGCTTACCAGCTCAGGTTT
CCCAACCGATCAAAGGCCGAGTACGATCTCCCCCATGTTGCAAAAAAGCGGT
TAGCTCCTCGGTCTCCGATCGTTGTCAGAGTAAGTTGCCGCGCAGTGTGAT
CACTCATGTTATGCGACACTGCAATAATTCTCTACTGTCTACATGCCATCCGTA
AGATGCTTTTCTGGACTGTTGAGTACTCAACCAAGTCTATTCTTGAGAATTAGTG
TATGCGGGCGACCAGGTTGCTCCTTGGCCGCGCTCAATACGGGATAATACCGCG
CCACATAGCAGAACTTTAAAGTGCTCATTGGAAGACGGTTCTCTCGGGGC
GAAAAACTTCTAAGGATTTACGCTGTTGAGATCCAGTGCTGATGTAACCCACT
CGTGCAACCCAACTGATCTTCAGCATCTTTTACTTTTACCCAGCGTTTCTGGGTG
AGCAAAAAACAGGAAAGGCDAATGCCCCAAAAGGGAAATAGGGGCACAC
GGAAATGTGAATACACTCAGCTTTCTCTCTCTCTCTCTCTCTCTCTGATTTATGAGAAGCATTATTC
AGGGTTATTGTCTCATGAGGGATACATATTTTGATGTATTTTAGAAAATAA
ACAAATAGGGGTTCCGCACACATTTCCCGAAAAAGTGCACACCTGAC

Plasmids used in construction of pSLIK5

>pUC19::NLS/HA-HumBeta-P2A/Crimson

AKA pUC19ΔKpnI-Kan::NLS/HA-HumBeta-P2A/Crimson.

> pUC19::NLS/HA-HumBeta-P2A/Crimson

GACGAAAGGGCCTCGTGATACGCCTATTTTATAGGTTAATGTGTCATGATAATA
ATGGTTTTCTTAGACGTGACTTTGCTGGAATCTGCTGCGGAGAACCC
TCAAGAGCTACCAACTCTTTTTCCGAGGTAACTGGCTTCAGCAGAGCGCAG
ATACCAAAATACTGTTCTTCTAGTGTAGCCCTAGTACCCAGCACCAGCTTCAAAGAA
CTCTGTAGCACCCCTACATACCTCGCTCTGCTCTGCTATACTCCGTGAGCTAGTT
ACGGGATAAGGCAGCGGTACGCTCGGCTGAGGGGGGTTCCTGCACACACCCGCC
CAGCTTGGAGCGAAACAGCCTACACCCGACTGAGATACCGTACGGTACGCTA
TGAGAAAGCCGCAACGCTTCCGAAAGGAGAACAGCCGACAGGTATCCCAGGGAAAC
AGCGGCAGGGTCGGAACAGGAGACGCGACAGGGAGCTCCAGGGGGAAAC
GCCTGGGTATCTTTATAGTCCTGCGGGTTTCGACCTCTGACTTGAGCGTGAG
ATTGGTTGATGCTCGTCAGGGGGTCGAGCCTATGGAAAAACGCCCAGCAAC
GCGGCTTTTTACGGTTCTGGGCTTTGGCTCTACATGTCTTTCTT
CCTGCGTTATCCCCTGATTCTGAGTAACCGTATTAACGCCTTTTGAGTGAGGC
TGATACCGCTCAGCAGCGCAGCGCAACGCGAGCGCAGCGAGTCAGTGAGCGA
GGAAGCGGAAGAGCGCGCCAATACGCAAAACCAGCTCTCCCAGCGTTGAGCCG
ATTGACCAATGCGAGCTGAGCAGAGGTTCTCCGACTGGAAGCGGAAACGCG
AGCGCAACGCAATTAATGTGAGTTAGCTACTCATATTAGGCCACCCACGCTTT
ACACTTTATGCTCTCCGGCTCTGTATGTTGTGGAATTGTGAGCGGATACAAAT
TTCAACAGGGAAACAGCTATGACCATGATTACGCAAAGCTTTACGTACCACC
ATGGTGCCCTCCAAGAAGAAGAGGATCGAGGACCCCAGAATATTACGTCCTAT
GACGTGCCCCGATTACGCTGGCGGAGGCGGAGGCTCCGCGGAGGAGGAGGAGGA
TCCGGAGAGGCAGCATGCTCCACAGCTCTCCGCAACCTGCGCGCGGAAAACCTGG
CCAGAGAGTCGGGATGGAGTCCCGAACGAGCTGACTCACCACAACCT
GAGGCAAACCGCCTTCAAGGGAGATGCTTCCGACGCCCCAAATTCATCGCCCTC
CTCATCGTGCAACCAGTACGGCCTGAACCCTTGGACCAAGGGAGGATTTACG
CCTCCCCGACAAGCAGAATGGCATTGTCGCTTGGTGTCGCGTGAGACGGCTG
GAGCAGAATCATTAACGAGAACCAGCAGATTGACGGCATGGACTTCGAGCAG
GACAACGAGAGCTGACATGCAGGATCTACAGGAAGAGCAGGAACCACCATCCT
ATCTGCGTCACCAGTGGAGATGGAGATGAGTGACAGAAGAGAACCCTTCAAGACCA
GAGAGGCAGAGAGATCACAAGGACATGATCTACAGGAGAGGAGCAAGAGAGATG
TCAGACACAAAGGCCATTGAGAGGATCTAGGAGAAATACCGCCTACACC
GCTGAAAGGCAAGCCCGAAAGGGACATTACACCGGTGAAGCAGGACACGACCATG
CAGGAGATCAATACCCTCTGCTATGCCCCTGGAACAAAAACCTGGGACGATGACC
TCCTCCCCCTCTCGGACGACCAATCTTCAGGAGGACATCAGGGCCCCAGCTCCGA
ACTCACACAAGCCGAAAGCTGTGAAGGCGGCTTCCGATTTCTTCAACACAAGAGGCC
GCTGACGAGAAGGCCTCCGGCCTCTCCGGATCCGGAGCCAACCGAACCTTCTC
TGTTAAAGCAAGCAGGAGACGTGGAAGAAACCCCGTCTATGGATAGCA
CTGAGAAGCTCATCAAGCCGACTCTTCGCTGCAGCTTCAGGCAAGTGGAGGGCTC
CGTGAACGGCAGTTCCGAGATCGAGGGCGTGGGCGAGGGCAAGCCC
CGAGGGCAACCAGACCGCAGCTGGCAAGTGAACAGGCGGCGGCGGCGGCGGCGG
CAGGCAGCAACCCCGACAATCCCGAGCATTACCTCAAGCTGACATGGACGAGGCTG
CTTCCGCTGGGACATCTCCTGTCCCCCCAGTCTTCTCATACGGCCTCAAGGCGTACA
TCAAGCACCACCGACATCCCGAGACTTCACCATTCCAGGCTCAAGGCGGTACA
CTTCAAGTGGGAGCGCTGTAGAACCCTTCGAGGCAACGGCCGCGGTGGTGAGCGTG
ACCCAGGACTCCTCCCTGCAAGACCAGACCCCTCATCATTACCTACCCGTAAGGACTGCTGGG
CTGGGAGCCCTCCACTGAGCCAAACTACCCCCGGCAGGGCGGTGCTGAAGGGCG
GAGAACCACATGGCGCTGAAGCTGAAGGCGCGAGCCACTACCTGTTGTGAGT
TCAAGTCCATCTCATGGAAGCGAAGCCGTTGAAGCTGACCGGGCTTCTACCTA
CGTGGACTACAAGCTGAGACTACAACCCGAACAGGACTACACCGTGGTG
GAGCAGTACGACGCGGGAGGCCCAGCACCACCTGTTTCAATGAAACGA
ATTCACTGGCCGTCGTATTGACGACTGGAAAAACCCCTGGCCGTTACC
CAAATTAATGCCTTCGACAGCACTCCCCCTTTTCGCCAGCTGGCGTAATAGCGA
AGAGGCCGCACCCGATCGCCTTTCCAAACAGTTGCCGACGCTGAGATTGCGAA
TGGCGCCTGATGCGGTATTTTCTCTACTCGACATCTGTCGTGGTATTTCACACG
CATATGGTGACTCTCAGTACAATCTGCTCGATGCGCAGTATGTAAGCGCAG
CCCGACACCCCGCCAACACCGGCTGACGCGCCCTGACGGGCTTGCTCCC
GGCATCCGCTTACAGAAGCTGTGACCCGTCTCCGGAGAGCTGACCTGTCAG
AGGTTCATCCGCATCCACGGAAACCGCGCA
>pENTR2B/TREPitt::NLS/HA-HumBeta-P2A/Crimson

CTTTCCCTGCGTTATCCCTGATTCTGGATAAACCATATTACCGCTAGCATGG
ATCTCGGGGACGTCTAACTACTAAGCGAGAGTAGGGAACTGCCAGGCATCAA
ATAAAACGAAAGGCTCAGTCGGAAGACTGGGCCTTTCGTTTTATCTGTTGTTT
GTCGTTGAACGCTCTCTCTGAGTAGGACAAATCCGCCGGGACGGGATTTGAAC
GTTGTGAAGCAACGGCCGGAGGGGTCGCTCCTGATGGGATTTTCTTTTTGCGT
TTCTACAAAACCTCCCTCTGATTTAGTTACTTAAAGCTCGGCCCAGCCCAT
ATTTTATTTTGACTGATAGTGACCTGTTCGTTGCAAACCACATTGATAGCAATG
CTTTTTTAATAATGCCAACACTTTGTACAAAAAGCAGGCTGGCGCCGGAAC
TTCAgatatcAAGTGCCACCTGACGTCTCCCTATCAGTGAAGAAGAGTCGACAC
GGCGGACAGGTATCCGGTAAGCGGCAGGGTGCGGAACAGGAGAGCGCACGAG
GGAGCTTCAGGGGAAACGCCTGTATCTTTATAGTCTCTGTGCCTGTGTTTCGCC
ACCTCTGACTTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCTTA
TGGAAAAACGCCAGCAACGCAGCGGACTTTTTTACGGTTCTGGCCTTTTTGCTGGCC
TTTGCTCACATGTT

">pSLIK/TREPitt::NLS/HA-HumBeta-P2A/Crimson

AKA 5-pSLIK-Zeo/TREPitt::NLS/HA-HumBeta-P2A/Crimson or pSLIK5. Sequence verified from 2738 to 4863 bp.

>5-pSLIK-Zeo/TREPitt::NLS/HA-HumBeta-P2A/Crimson

GTCGACGGATCGGGAGATCTCCCGATCCCCCTATGGTGCACTCTCAGTACAAT
CTGCTCTGATGCCGACTTTTTACGGTTCTGGCCTTTTTGCTGGCC
TTTGCTCACATGTT
AGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTT
GACCGACAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTC
GCGATGTACGGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATGA
TAGTAATCTAATTACGGGGTCTGTTGGATAGTTTACGGTTAAACTGCCACTTGCCAGTA
TACATAAUCTACGGTAAATGGCCCGCCTGGCTGACCGGTTGACCGGTTTTTG
CCATTGACGTCAAATAAGCACGTATGTAGTTCCCATAGTAACGCCCATAGGGA
CCATTGACGTCAAATGGGTGAGAGATATTTACGGTAAACTGCCCACTTGCCAGTA
CATCAGGTTGAGATCATGCAATGACGTACTGCAACCTATTGAGTACGCTTATG
GAGTACATCATAATGGGCGTGGATAGCGGTATTGACTCAGGAGGGATTTCAA
CTCACCACCATGGACGTAAATGGGAGTTTGGGATTTGTCAGACGCAAAATGA
ACGGCGGTAG
GGGTACGCTGCTGTTGGGACTTAGGGGTCTATATAAGCAGCGCGTTTTGC
CAGTAATGGGCGTGGATAGCGGTACTGCTTGGTCCCTAACTAGGG
AAATATAAATTAAAACATATAGTATGGGCAAGCAGGGAGCTAGAACGATTCG
CAGTGTACGGGTGAGGTCTATATAAGCAGCGCGTTTTGCTGCTGACTGGA
CTCAGATCAGGAGCATCTGGAGGAGCTTGGATATTCAGATCAGATCAGA
CTGGAAAATCTCTAGCAGTGGGCAGCAGGAGCTGCTGACTGGA
AAACCAGAGGGAGCTCTGGGAGGAGAGACGGCTGCTGACTGGA
AAAGCAGGGAGCTCTGGGAGGAGAGACGGCTGCTGACTGGA
CAAGGAGGGCGGGGCGGCGACTGGTGAGTACGCACCCAAAATTTGACTAGC
GAGGCTAGAAAGGAGAGATGGGTGCGAGAGCGTCAATGATATAGCGGGGG
GAGATTAGATCGGATGGGAAAAATTTCGGTTAAGGCGGAGCTAGAAGCGATTGC
AATATAAATTTAATCAATTAGTATGGGCAAGGAGCGGAGCTAGAAGCGATTGC
AGTTAATCCTGGCCTGTAGAAACATCAGAAGGCTGTAGACAAATCTGGGA
CAGCTACAACCATCCCTCAGACAGGATCGAAGAAACTTAGATCATTATATA
ATACAGTAGCAACCCTCTATTTGTGTGACATCAAAGGATAGAGATATAGACAC
CAAGGAAGCTTTAGACAAGATAGAGGGAAGACAAAACAAAGTAAGACCAC
CGCACAGCAAGCGGCAGTGATCTTGACACCTGGAGGAGGAGATATGAGGA
ACAATTGGAGAATGTAAATTATAATATAAATAGTAGTAAAAATTGAACCAT
AGGAGTAGCACCACCAGGAAGAGAAGAGTGTGTCAGAGAGAAAGAAAA
GAGCAGTGGAATAGAGGCTTTTGTCTTGGGTCTTGAGGAGCAGCAGAAG
CACTATGGGCGCAGCGTCAATGACGCTGACGGGTACAGGCCAGACAATTATTG
TCTGGTATAGTGCAAGCAGCAGAACAATTTGCTGAGGCTATTGAGGCAGCCAA
AGCATCTGTGTGCAACTCAGTCTGGGGGATCAGATGGCTCCAGGCAAGAAT
CCTGGGCTGTGGAAAGATACCTAAAGGATACACAGCTTGGGGATTTGAGGT
TGCTCTGGAAAACCTCATTGGACACACTGCTGTGGCTTGGGAATGCTAGTGAG
TAATAAAATCTCTGGAACAGATTTGGAAATCAGAGCTGAGGAGTGAGGAC
AGGAAATTAAACACAGCTTAAATACACTCTTATGAAAGATCGGC
AAAACCAGCAAGAAAAAGATGAACAGAATTATTGGAAATAGGAAAATGGG
CAAGTTTTGTGGGAAATTGTTTTACATAACAAAATGGCTGTGGGTATAATAAATTAA
TTCTAGATATAGTAGGAGCTTGGTAGGTTTAAAGAATTAGTTTTTGTCTGTACT
TTCTATAGTAAATAGAGTTAGGCAGGGATATTGCCACATTATCGTTTCAAGCAC
ACCTCCCAACCACGGAGGGACCCGCAGGCGCGAGAATAGAAGAGAGAG
GTGGAGAGAGAGACAGAGACAGATCCATTGCATTAGTAGAAGCAGATCGGAC
TGCGTGCGGCAATTCTGCAGAACAAATGGCAGATTTGCTAGTACATCCACAAATT
GAAAAAGGGGGGATTGGGGGTACAGTGCAAGGGGAAGAAATAGTAGACATAA
GACCAGAGAGGGCAGAGAGATCACAGGACCCTGGCAGTCCCATCCTAAGAG
AATGCTCAGACACAAGGCCATGATCAGTGTGCCAGGCTCGCTTTTGGCTTCG
CCGGCATCTATGATAAGGACGAAGGGAGGATCTGGGAAAAATACCGCCTA
CACCGGTAAAGCGGCAAGGACATTACACCCGTGAAGGACGCAGGACT
CATGCAGGAGATCAATACCCTGCTCATTGCCCTGGACAAAACCTGGGAGAT
GACCTCCTCCCCCTCTGCAGGACATCAGGGGACCTGCCAGCT
CCGAACTCACACAAAGGCCAAGCTGTAAGGGCCCTCGGATTCCTCAACACAGAA
GGCCGCTGAGCAGAAGGTCGCCGCCGTTCCGATCCGGAGGCCACGAACCTTC
TCTCTGTAAAGCAAGCAGGAGACGTGGAAGAAACCCCGGTCTCATATGATA
GCACTGAGAACGTCATCAAGCCCTTCATGCCGTTCAAGGTGACATGGAGGG
CTCCGTTGAAACGGCAGCTTCGAGATCGAGGGCGTGCGAGGGCAAGCC
CTACGAGGGCACCAGACCCGAAGCTGCAAGGTGACCAAGGGCGGGCCGCCTT
GCCCTTCGCCTGGGACATCCTCAGTCCCCCAGTTTTCTACGGCTCCAAGGCGT
ACATCAAGCACCCCGGCGACATCCCCGACTACCTCAAGCAGTCTCCCCGA
GGGCTTCAAGTGGGAGCGCTGATGAACCTTCGAGGACGGCGCGGTGGAACC
GTGACCCAGGACTCCTCCCTGCAAGGAGCGGACCCCTCATCTTACCATGTAAGT
TCATCGGCGTGAACCTTCCTGCCAGGCGGCGCAGGGCCGCTGGTGACC
GGGCTGGAGACCTGAGGCTGACGCTGGAAGCTGGAAGGGCGGCGGCGGCTGGTGACC
GGCGAGAACCACATGGCGCTGAGCTGGAAGGGCGGCGGCGGCTGGTGACC
GGATTTCAAGTCCATCTAGGGCAGGAGACGGCACCCTCATCTTACCATGTAAG
ACTACGTTGGACCTACAAGCTCGACATCACCTCAGGACAACAGGAGCAGTACACCGT
GGTGGAGCAGTACGAGCGCGCCAGGGCCCGCCACACCTGTTCCAGTAGA
CGAATTCCGGCAGCGACTCGAGATATCTAGATAGTGACTGGGATATGTTGTGT
TTACAGTATTATGTAGTCTGTTTTTTATGCAAATCTAATTTAATATATTGATA
TTATATCATTTTACGTTTCTGTCAGTTTTCTTGTAACAGGTGTTGAATAAC
ATCTCGAGCCATAAAAGATGGTTAATTAAACCCACCAAGATCTGCGCTCCGCG
CCGTTTTGGCAGCTTCCCGCCGCGGCGGCGCCCCCCTCCCACGCGAGCGCTGCC
ACGTCAGACGAAGGCGACAGCGAGCGTCCTGATCTCCTCCGCCGCCGGCACGCTCA
GGACAGCGGCCCGCTGCTCATAAGACTCGGCTTTGAGAACCCAGATACAGCA
GAAGGAACATTATTAGACGGGACTTGGGTGACTAGGGCTACCTGGTTTTCTTTTC
CAGAGAGCGAAACAGGCGAGGAAAAATGATGCCCTTCTCGGCGATTTCTGCGGA
GGATCTCCGTGGGGCGGTGAACGCGCTATGATTATATAAGGACGCGCCGGGT
GTGCCACAGCTAGTTCCGTGCAGCCCGGGATTTGGGTGCAGGTCTGGTTTGTG
GGATCGCTGTATCGTCACTTGGTGAGTAGCGGCTGCTGGGCTGGCCGGGG
CTTCGTGGCCCGCCCCGCTCGGCTGGGGACGGAAAGCGTGTGTGAGAGACCGC
CAAGGGCTGTAGTCTGGTGGTCCGAGCAAGGTGGCCCTGAACTGGGGTTGG
GGGGAGCGCAAGAAATGCGCGCTGTTCCCCAGTCTTTGAAATGGAACGCGCTT
GTGAGGGCAGCGCTGTGAGTGGCAGGCTGCTGGGGCTGGCCGAG
CTTTCGTCGCGCAGCCGGGCTCGGTCGGAGTTGGTAGGCCCCTGAACGCTGGTGG
GGGAGCGCAAGAAATGCGCGCTGTTCCCCAGTCTTTGAAATGGAACGCGCTT
GTGCGAGCTGCCGCGGCTGTGAGTGGCAGGCTGCTGGGGCTGGCCGAG
CAAGGGCTGTAGTCTGGTGGTCCGAGCAAGGTGGCCCTGAACTGGGGTTGG
GGGGAGCGCAAGAAATGCGCGCTGTTCCCCAGTCTTTGAAATGGAACGCGCTT
GTGAGGGCAGCGCTGTGAGTGGCAGGCTGCTGGGGCTGGCCGAG
CTTTCGTCGCGCAGCCGGGCTCGGTCGGAGTTGGTAGGCCCCTGAACGCTGGTGG
GGGAGCGCAAGAAATGCGCGCTGTTCCCCAGTCTTTGAAATGGAACGCGCTT
GTGAGGGCAGCGCTGTGAGTGGCAGGCTGCTGGGGCTGGCCGAG
CAAGGAACCAAGGTCTTTGAGGCCCTTCCGCTAATGCGGAAAGACTCTTTATTCGG
GTGAGATGGCCTGGGACCACTTGGGGACCCCTGACGTGAAAGTGTGTGCTACTG
ACTGGAGACCTCGGTTTGTCTGCTGTTGGCGGGGGCGCAATATGGCGCTGTC
CGTTGGGCACGTGCACCCGCTACCTTGGGAGCGCGCGCCCTCGCTCGTGTGCTG
ACGTCACCCGGTCTGTGGCCTTAATATGCAGGGTGGGGCCACAGCCTGCGGGT
GTGTCGCGTAGCTTTTTCTCGTCAGCGACGGAGGCGGCTTGAGGCTAGGGTA
GGCTCTCCTGAAATCGAGGCGCCGGCACCTCTGGGTAAAGGGCCAGAGGTTAGT
AGGGCGTCAGTTTTCTTTTGTCGGTTTTATGTACCTATCTCTTAAATGAGCTGAA
AGGAAGCAGTTCCTCTGGAAGCTTCTTGAAGACAAACAACGTCTGTAGCGAC
CTTTTCGACGGCAGCGGAACCCCCCCACCTGGGCGACAGGTGCCTCTGCG
GCCAA
AAGCCACGTGTATAAGATACACCTGCAAAGGCGGCACAACCCCAGTGCCACG
TTGTGAGTTGAGTAGTTGGAAGAAGGATCACAATGGGTCTCCTCTAGCGTAGT
CAAAACAGGGGGTCGAGGATGCAGCAGACGACCATTGTATGGGAATCTGA
TCTGGGCGCTCAGGTCACATGCTTTACATGTGTGTAGTGGAGGTTAAAAAAC
GTCTAGGGCCCGGGAACCACCGGCGACGTGGTTTTCCTTTGAA
AAACACGATG
ATAAGCCTACCAGTACGCGCTAGTGTGTGTAGCAGAATATCATCGGGCATAGTATA
TCGGGCAATAGTATAATACGACAAGTTGAGGAACTAAACCATGGCCAAAGTGC
CACGTGCCGTTCGGGTCTCACCAGCGCAGCTCGCCGGAGCAGGTGGACTTC
TGGACCCAGCGCTCGGTTCTCCCGGGACTTCGTGGAGGAGCAGTCTCGCCG
GTGTGGGTCGGGGACGCAGTGACCTGTTCACATCGCGCGTCCAGGGACAGGT
GGTGCCGGAACACAACCCTGCGCCCTGAGGTTGTTGGGTGGCGCGCGGCTGGGAGT
TACGCGAGTGGTCGAGGTGCGGTAGTCACACGGAGGCCTCGGGGC
CGGCCATGACCGAGATCGGCGAGCAGCCGTGGGGGCGGGAGTTCGCCCTGC
GCGACCCGGCGCAGCAACTGCCTGTCGACCTTCTGGCGGAGGAGCAGGACTGACA
CATCTGTACAAAGATAAGCGGCGGAGCAGTCTCTAGATACAATACGCGCATACCAC
ATTGTAAGAGGTTTTACTTGCTTTAAAACCTCCACACCTCCCCCTGAACC
TGAAAACATAAAATGAAATGCAATTTGTGGTTGTTTGTAGTCCTCTCCTATATGATAT
CAAGCTTATCGATAATCAACCTCTGGGATTACAANATTTTGGAAGATGGACTG
GTATTTCTAATGTTGCTCTTTTACGTATGTGGATACGCTGCTTTAATGTC
CTTTGTATCATGCTATTGCTCTCCCGGTATGGCTTTTATTTCTCCTCTCTGTTAA
AATCCCTGGTTGCTGTCTTTATGAGGAGTTGTGGGCGGGGTTTCAGGGCAACGT
GGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCCACTGGTTGGGGCATTGC
CACCACCTGTGCTAGCTCTTTTCCGGGACTTTGTGCTTCCCCCTCCCTATCTCCAC
GGCGGAATCATCGCCGGGCTTGCTGCTGTTCGCTGGACAGGGGCTGGCTG
TTGGGCACTGACAAT
TCCGTGGTGTTGTCGGGGAAATCATCGTCCTTTCCTGTGCTCTACG
TCCCTTGCGCCCCCTCAATCCAGGGGACCCTCTCTTCCTCCGCGCGCCTGCTGCGGCT
CTGCGGCCCTCTCCGCTCTCGCCTCCCTCGCCACTACAGAGTGGATCTCCCT
TTGGGCGCCCTCCCCGCAATCGGTACCTCGACTCGATCGAGACTCGATCGAGATCTAGAAAA
ACATGGAGCAATCACAAGTGAATACAGCAGCTACCAATGCTGATTGTGCTG
TGCTAGAAGCAAGAGGAGGAGGAGGTGGGTTTTCCAGGCACTCCACTTTTGA
TACCTTTAAGACCAATGAATCACAAGGCAAGCTGAGATCTTAGCCACTTTTTTA
AAAGAAAAAGGGGACTGGAAGGGCTAATTCACTCCCAACGAAGACAAGAT
AAGCTAGTACCAGTTGAGCAGAAGAGGAGTTTAGTGATGCTAC
AACACCCGCTTGTACACCCTGTGTAGCCTGCATGGGATGGATGACCCGAGA
GAGAAGTTAAGTGGAGGTTAGTGAGCAGCCGCTACTTGCTTCTGGTTAGACCGAGATCTG
CGAGAGCTGCATCCGGGACTGTACTGGGCCTTCTCTTGGACTCTGGTAA
GCCTGGAGAGCTCTCCTGGCTAAGAGGGAACCGACTGCTAAAGGCATCATAAAA
GCTTGCCCTTGGACTTCAAGTAGTGTGCGGTCCGCTGCTTGGTGTGACTCTGGT
AATCGAGATCCCTACCCCTTAPGTATCAGGTGGAAAATCTCTAGCAGGGCGCTTTAAACCCGCTGATCCGACTGTGCTTTGACACCGCATC
TGTTGTTTGGCCCTCCCGGTGCTTCTTTGACCCCGCTGAGGTGAGTGCCCACCTCCAA
CTGTCCTTTCTAAATAAAAATGAGGAAATTCAGCATTGCTTGAGTAGGTTGCT
CATCCTATTTCTGFGGGGTTGTTGAGGGCAGAGACAGGAAGGGGAGGATTG
GAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGG
CGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCGAACGCCCTGTAGCGG
CGCATTAAGCGCGGCGGGGTGTTGTTGTTACGCGCAGCGTGACCCTACCTT
GCCAGCGCCCTAGCGCCCGCTCTTTTCTGCTTTTCTCTTTTATGTTCTCCG
TTCGCCGGCTTTTTCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCG
ATTTAGTGCTTTTCAGCACCCTCGACCCCCAAAAAATTTGATTAGGGGTAGGTT
CACGTAAGTGCGCCTACACGCTTCTAGAGACGGTTTTTCCGCTTTTACGTTGGAG
TCCACGTTTTTAAATAGTGGACTCTTTGTTCAAAAACTGGAACAACTCCCAACCC
TATCTCGGTTCTATTTTTTTGATTATAAGGGGATTTTTGCCGATTTCGCGCTATTG
GTTAAAAATGAGCTGATTTAAATTTGTTAACGGAATTAATTTCTGTGGA
ATGTTGCTCGTATTAGGGGTGGAAGTCCCCAGCTCCCATCTCAATTGTACGGCAA
GCTCCCCAGCAGGCAGAAGTATGCAATGCGATGCATCTCAATTAGTCAGCAA
CCATAGTCCCGCCCTAATCTCGCCCACTCCCGGCCCCTTAATCTCGCCAGTTCC
GCCCATTTCTCCGGCCCATGGCTGACTAATTTTTTATTTATGACGAGGCCGA
GGCCGCCTCTCCTGCTGCTAGTCTCCAGAAAGTATGAGGAGGAGGTTTTTTGGAG
GCCTAGGCTTTTGCAAAAAGCTCCCGGAGCTTTGATATATCATTCTCCGGATCT
GATCAGCAGCTGTTGCAATTATATCATCGGCATAATGATATATCGGCATAGTATA
ATACGACAAGGTAGGAGAACTAAAAACATGGCAAGTGAGCAGCGTTCCCG
GTGCTACCAGCGACGCGAGCGTCGCCGCGGAGCGGCCGAGCTGCGATTCTGGAGC
TCCGAGTTTCTCCCGGGACTTCTGGAGAGGACGACTTACCGCCGCTGTGAGTGCTCCGGA
ACCCCCCGTTCAGCGGACCGCTGCGCCTTATCGGTAACCTATCGTCTTGAG
TCCAACCGGTAAGAAGCACGACTTATCGGCACCTGGCAGCGCAGCCACTGGAACA
GGATAGCAGAGCGAGGTATGTATAGGGCGTGCTACAGAGTTTCTTGGAAGTGGTG
GCCTAACTACGGCTACACTAGAAGACGAGATTACGGCGCAGAA
AGCCAGTTACCTTCGGAAAAAGAGTTTGAGCAGCTCTTGATCAGCAGACGTTATAC
CACCGCTGGTAGCGGTGGTTTTTTGGTTTGCAAGCAGCAGATTACGCGCCAGAA
AAAAGGATCTCAAGAAGATCCTTTTGATCTTTTCCTACGGGCGTCTGACGCTCA
TGGAACGAAAACCTACGTTAAGGGATTTTGGGTCATGAGATTATCAAAAAAGGA
TCTTCACCTAGATCCTTTTTAAATAAAAATGAAGTTTTAAATCAATCTAAAAGT
ATATATGAGTAAACTTGGTCTGACAGTTACATGTTATACGAGGACAC
CTATCTCAGCGGATCTGTCTATTTTTCATCCTCCATAGTTGCCCTGACTTCCCCGTCG
TGTAGATAACTACGATACGGGAGGGCTTACCATCTGCGCCAGTTTTGCTGCAAT
GATACCCGGAGACCCACGGCTCACCAGGGCTCCAGATTTTATCAGAAATACAGAG
CCAGCCGGAAGGGCCGAGCCGAGAAAGTGTTGCGCTCAACTTTATCCGCCTCCA
TCCAGTCTATTAATTGTTGCCGGGAAAGCTAGAGTAAGTAGTTCGCCAGTTAAT
AGTTTCGCAACGGTTGGCATTGCTACAGGCACTGTTGGCTACGGCTCAGC
GTTTTGGTAGTTGCTTCATCCAGCTCCGGGTGTTCCCAACGATCAAGGCGAGTTACAT
GATCCCCCATGTTGCTCAAAAAAGCGGGTTAGCGCTCCTTCGCTCCTCCAGTACGGT
GTCAGAAGTAAAGTTTGCCGGCAGTTTATCTACATGTTATTAGCGCAAGACTGGC
ATAATTCTCTCTACTGTCATGCACTCCCCGTAAGATTTTTCTGTGACTGAGG
TACTCAACCAAGTCACTTGTAGAATAGTGATATCGGCAGCCAGTTGCTCTTG
CCCGGCCTCAATACGGGATAATACCGGCACGACTGCAAGAACTTTAAAAAGTG
CTCATCTGGAAAAACGTTCTTCGCGGCGAAAAACTCTCAAGGATCTTACCAGCT
GTTGAGATCCAGTTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTTCACCAGCGTTTCTGGGTGACCAAAAAACAGGAAGGCAAAAATGC
CGCAAAAAAGGGAATAAGGGCGACAGGAAATGTTGAATATCTCATACTCTCCTCTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATA
CATATTTGAATGTATTTAGAAAAATAAAACAAAATAGGGGTTCCGCACACATTTCCCGAAAAGTGCCACACCTGAC

Plasmids used in construction of pSLIK6

>pFN24K::P2A/Crimson-control

AKA pFN24K::Product C-Control.

>pFN24K::P2A/Crimson-control

ATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTGGTT
TGAACCCGTCAGATCAGTGAAGCTTAATACGACTCACTATAGGGTACGATTT
AGGTGACACTATAGAATAAGGAGAATTCCATCATGGCAGAAATCGGGTACCTGGC
TTTCATTCGACCCCCATTATGTGGAAGTCTGGGGGCAGCGCATGCACTACGT
CGATGTTGGTCCCGCGAGATGGCACCCTCTGTGTCTTCTGTACGGAACCCGA
CCTCTCTTACCAGATCTAGAGGAGTCTGCGACCACATGGCCCGATCCGGAGCCA
CGAACTTCTTCTGTTCAGGAAAGCAAGCAGGAGAGCTGGGAAGAAAACCCCGGTCC
TATGGATAGCACTGGAACGTCATCAAGCCCTTCTATCGCCTTCAAGGGTGAC
ATGGAGGCTCCGTGAAAAGGCGCCAGCAGTGTCCAGATCGAGGGCGTGGGCGAG
GGCAAGCCCCAGCAAGGGCAACCCAGACCACGGCAAAGCTGAAGTGAACAGGAGC
GGCCCCTGCCCCTCGCCTGAGAGATCTCCTTGCTCCCTCCAGTTCTTCTACGCTC
CAAGGCGTACATCAAGCACCACCGGCACATCCCCGACTACCTCAAGCAGTCC
TTTCGGGAGGGCTTTCAAGTGGGAGGCGTGTGATGAACCTTCGAGGACGGCGGCG
TGGTGACCCTGACCCAGGAGCTCCTGCTGCGAGGACGGCAACCCTCTCATCTACCA
CGTAAGTTCACTCCGGCGTGAACCTCCCCCTCCCGACAGGGCCCGCTATGCAAGAAG
AAGACTCTGGGCTGGGAGCCCTCCTGCAGCAGCAACTACCCCGCAGCGACGGC
TGCTGAAGGGCGGAGAACACATGGCGCTGAAGCTGAAGGCGACGGCGGACT
ACCTGTGTGAGTTCAAGTCCATCTACATGCGCAAGAAGCCCTGTAAGCTGCC
CGGCTACCACTACGTTGAGCTACAAACGTGCAGCATCACCTCCCCACACGAGGAC
TACACCGTGTTGGAGCAGTACGAGCGCGCAGGCGGAGGGCCCGCCACACCTGTCC
AGTAGGAATTCCGCGGCCGAGTGTTCGCCGTTCTTTAGCAGCCTTTCGCGCCTT
GAGTGCTTCGGCGACGGTGAGCTTCCAAAAAGAATGGCCAGCTGGGGCGGCCCTC
TGGTAAGGTTGGGAAAGCCTGCAAAGTAAMCTTGGATGGCTTTCTTGCGCCA
AGGATCTGATGGCGCAGGGGATCAAGATCTGATCAAGAGACAGGATGACGG
TCGTTTCGCAATGCTTTGAACAGATGGATGGCAGCAGGTCTCCGGGCAGGTG
CGCTGCCTTTATCCGGTAACGTATCGTCTTTGAGTCAACACCACGTGAAAGACACG
ACTTATCGCCACTGGCACGAGCAGCCACTGGTTAACAGGATTAGCAGAGCGAGGTAA
TGTAAGGCGGTGCTACAGAGTTTCTTTGAAGTGGTGGCCTAACTACGGCTACA
AGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTTCGGAA
AAAGAGTTGGTAGCTCTTTGATCCCCGGCAAACACACACCACGCTGTCAGGGTGGA
TTTTTTTTTGGCAAAGCAGAGATTACGCGCAAGAAAAAAAGGATTTCAAGAA
GATCCTTTTATCTTTTTCTACGGGGTCAGCTACGGTGAACGAAACGAAAACCTACG
TTAAGGGATTITGGTCTAGAGATTATTATCAAAAAAGGATCTTCACCTAGATCCTTT
TATAGTCGGAAATACAGGAACGCACGCTGGAACGATGCGCCCTTCGCTGGGATGGT
GAAAACATGAAAATGGCAGCTTCAGTGGATTAAGTGGGGATTAATGTGGCCCT
GTACCCCTCTGGTTGCATAGGTTATCTATACGGTTAAAATTTATCAGGCGCGATT
GCGGCAGTTTTTCGGGTGGTTGTGGCCATTTTACCTGTCTGCTGGCCGGTGATC
GCGCTGAACGCGTTTTAGCGGTGCGTACAATTAAGGGATTATGGTAAATCCA
CTTACTGTCTGGCTTGGAGGATACGAGGATATAAACGCGAGTACTCGCGCCACG
ATGCGTCGGCGTGATAGGGGATCGGATCT
> pENTR2B/TREPitt::P2A/Crimson-control

CTTTCCCTGCGTTATCCCTGATTCTGTGGGATAACCGTATTACCGCTAGCATGG
ATCTCGGGGACGTCTAAACTACTAAAGCAGAGTGGGAACCTGGCGGATTTGAAC
GTTGTGAAGCAACGGCCCGGAGGTTCGCTGCAAAGACGCTGCTCCCATGCTGCT
GCCAGGCATCAACACTAAAGCAGAGGCCCATCCTGACGGTGAGGCCTTCTTCTGC
TTACAAACTCTCTGATTAGTTACTTAAGCTCGGCGCCCAACATAATG
ATTTTATTTTGGACTGATAGTGACCTGTTCTGCAACGAAAATGGATAAGCAATG
CTTTTTTATAATGCCAACTTTGGTACAAAAAGCAGGCTGGCGCGCCGGAACCAA
TCAGGATATCAAGTGCCACCCTGACGTTCCCTTATCAGTGATAGAGAAGTCGA
AGCATTGCTTATCAATTGTTGCAACGAACAGGTCACTATCAGTCAAAAATAA
AATCATTATTTGCACATTCCAGCTGCAGCTCTGGCCCGTGTCTAAAAATCTCTGA
TGTTACATTGCACAAGATAAAAAATATATCATCATGAAACAATAAAAACTGTCTG
CTTACATAACAGTAATACAAGGGGTTGTTATGAGCCATATTCAACGGGGAAAC
GTCGAGGCGCGATTTAAATCCACATGGAGTCTGATTTTATATGGGTATAAAA
TGCGCTCGAGATAATGTCCGGCAATCAGGTTGCAAACTCAGGACAGGAAATTATG
CAATGTATTTACGAGTGAAGAGTGTCAGACTAAACTGGCTGACGGAATTATG
CCTCTTCCGACCATCAGCATTATATCCGTACTCCTGAGATGATGCGTTTACT
CACCAGTGCAGATCCCAGGAAAAACAGCATCCAGGTATTTGAGAATATCCT
GATTCAGGTAAGGATATTTGCTAATGCGCAGTGTTCCTGGCAGGCCTGCA
TTGATTCCCTGTGTTGTAATTTGCTCTTTTAACAGCGATCGCTGATTTTCGTTCG
TCAGGCAGAAATCACGAATGAATAACGGTTGTTGAGTTGGAGTGGATTTTTGAT
GACGAGCGTAATGGCGCTGGGAGAGTCTGGGAAAGGTAAGATGCATAAAC
TTTGGCCAATTCTCACCGGATTGAGTCGTCATCTCAGTGTTTCTCAGTAT
ACCTTATTGGCACTTAGGGGGAATAATAGGGTTGATATTGAGTGGGGGAGTAC
GGAATGCAGCAGACCAGATCCAGGACTTTGCATCCTATGGAACCTGCCCTGGT
AGTTTTCTCCTTCTATTACAGAAACGCTTTTTTTCAAAAAATAGGTATTTGATAAT
CCTGATATGAATAAATATGCTGTTTCTGAGATGATGTTTTTCTAATTC
AGAATTGTTAAATTGTTGTAACATTATTCGAGATGGGCCCCGTCCACTGAG
CGTCAGACCCCGTGAAGAAAAGATCAAAGGATCCTTTCTGAGATCCTTTTTTTCTG
CGCGTAATCTGCTGTGCTTCGAAAAAAAGAAACACCGCTACAAGCGGTGTTT
GTTTGCCCGGATCAGAGGCTACCAACTCTTTTTCCCAGGAAGTAACGTGCTTCAGG

671
AGAGCGCAGATACCAAATACTGTTCCTTCTAGTGTAGCCGTAGTTAGGC
ACTTCAAGAAACTCTGTAGCACCGCCTACATACCTCCTGTCTGCTAAATCTCTGT
CCAGTGGCTGCTGCGCAAGTGCTGGCTGCTTTACCGGGGTTGGACTCAA
GACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACCGGGGGGTTTCGTGC
CACACAGCCCAAGCTTGGAGCGAAGCGACCTACACCGAAACTGAGATACCTACAG
CGTGAAGCTATGAGAAAGCGCAGACGGCTTTCCGAAGGAGAAAGGCGGACAGG
TATCCGGTAAGCGGCAGGGTCGGGAACAGGAGAGCGACGAGGGAGCTTCCA
GGGGAAACGCCTGGTATCTTTATAGTCTCTGGTTTCGACCACCTCTGACT
TGAGCGTCGGTTTTTGTGATGCTCGTCAGGGGGGCGGAAGGCTATGGAATAAC
GCCAGCAACGCAGCTTTTTACGGTCTCGGGGCTTTTGCTGTCAC
ATGTT
>pSLIK/TREPitt::P2A/Crimson

AKA 6-pSLIK-Zeo/TREPitt::P2A/Crimson-Control or pSLIK6. Sequence verified from 2758 to 3966 bp.

>6-pSLIK-Zeo/TREPitt::P2A/Crimson-Control

GTCGACGGATCGGGAGATCTCCCGATCCCCCTATGGTGCACCTCTCAATACAT
CTGCTCTGATGCCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTTGG
AGGTCGCTGAGTAGTGCGGAGCAGCAAAAATTTAAGCTACACAAAGGCAAGGCTT
GACCGACAATTGCAGGAAGATCTGCTTACTACGCGTTAGCTATGGGTAGGCTT
GCGATGTACGGGCAGATATACGCGTTGACAT
TGATTATTGACTAGTTATTAA
TAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGT
TACATAACTTACGGGATCATTAGTTCATAGCCTATATGGGAGTTCCGCGT
CCATTGACGTAAATAGCAGTATGTTCATAGCCTATATGGGACATTAGGACTTT
CCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTA
CATCAAGTGTATCATATGCGCAAGTACGCCCCCTATTGACGTCAATGACGGTA
AATGGCCCGCCTGGCATTATGCGCCAGTAGACATGACCTTATGGGACTTTCCTACT
TGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTTG
GAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACCGGGATTTCCAAGT
CTCCACCCCATGGACGTCAATGGGAGTTTTGGCACCAAAATCAACGGG
ACTTTCCAAAATGTGTAACAATCCTCCGCCCCATGGACGCAATGCGGGGTAG
GCGTGTACGGTTAGAGGGTCTATATATAAGCAAGCAGCCTTTTGACTTCAAGTAGTGTTGACC
CTCTGGTTAGACCAGATCTGACTGGGAGCTCTGGCTAATCTAGGGAACC
CACTGCTTAAGCCCTCAATAAAGCCTTTGGACTTCAAGTAGTGTTGACC
CGTCTGGTTGTGACTCTGGTAATCTAGGAGATCCCTCAGACCCCTTTTAGTCAGT
GTGGAATAATCTCTAGCAGTGCCCGAAGCCGAGCCGACTTTGAAGAGCGAAAGGG
AAACCAGAGGAGCTCTCGACGCCAGGACTCGGCTCAGAAAGCGCCGACCG
CAAGAGGCAGGGGCGGCAGGGACTCCGAGTCCTTGTGAAGCGGAGG
CACTGCTTAAGCCCTCAATAAAGCCTTTGGACTTCAAGTAGTGTTGACC
GAGGCTAGAAGGAGAGATGGGTAGGGTGAGTCAGTATTAAGCGGAGG
GAATTAGATCCGAGTGAGAAAATTCGGTTAGGCCACGAGGGGAAGAAAAA
AATATAAAATTTAAACATATAGTGATGGGCAAGCAGGAGCTAGAAGCAGATTCGC
AGTTAACTCCGGCTTGTAAGAAAATCATAGAAGGGCTCTGAGACAAATACTGAGGA
CAGCTACAACCATCCCTCGACAGGATCAGAAGAACTTAGATCATTTATATA
ATACAGTGCAACCCCTCTATTGGTGATCAAGGATAGAGATAAAAAGACAC
CAAGGAAGCTTTAGACAAGATAGAAGGAAGAGCAAAAAAGTAGAAGACCAC
CGCACAAGCGGCGCTGGATCTCTGACTGACGACTGAGGAGTATGAGGG
ACAATTGGAGAAGTAGTAATATATAAAATATAAGTAGTAAAAATTGAAACCATT
AGGAGTAGCACCCACCAAGGCAAGAGAAGGTGGTGCAGAGGAAG
GAGCAGTGGGAATAGGAGCTTTTGTTCTTTGCTGGGTCTGGATAGGAGG
CACTATGGGCGCAGCGTCAATGACGCTGACGGTACAGGCCAGACAATATTG
TCTGGTATAGTGACAGCAGCAGAACAATTTGCTGAGGGCTAATGAGGCGCAAC
AGCATTCTGGTCAAATCTCACAGTCTCGGGCCATCAAGCAGCTGACGGTACAGGCCAGCAAGAAAAT
CCTGGCTGTGGAAAGATACCTAAAGGATGAACAGCTCCTGGGGGATTTGGGGT
TGCTGTGGAAAGACTCATTGGCAACCACGTCTGTCCTGGGAATGCTAGTGTGGAG
TAATAAAATCTCTGGGAACAGATTTGGAATACACACGACCTGGATGGAGTGGGAC
AGGAAAATTAACACACAAAGCTAATACACTGCTTAATTCAAGGAAATCGC
AAAACCAGCAAGAAAAGAATGAAAAAGAATATTGGAATTAGGAAATTGAGTAAATTG
CAAGTTTGTGGAAATTTGTTAACATAACAATTTGGCTGTGGTATATAAAATTA
TTGATAATGATAGTGGAGGCTCTGCTAGGTCTTTAAGAATAGTTTTTGCTGTACT
TTATAGTGAATAGTAGTGGAAGGTATTTCAATTCCCATATTACGTTTTTACGACCC
ACCTCCAACCCCGAGGGGACCCGACAGGCCCAGAAGAAATAGAAAGAAAG
GTGGAGAGAGAGACAGAGAGACAGGAGATCCCTTTGGAAGGGAATAGGAGTGGGAC
GAAAAGGGGGATTGGGGGGTACAGTGCAGGGGAAAGAATAGTAGACAT
TAGCAACAGACATACAAAATGAGCGATATTGATTTATCCCAACATTTAAAA
GAAAAAGGCGGTGGAATCCCAATGGGCGGAAATAGAGTGGGACCAT
TAGCAACAGACATAAAACTAAAGAATTACAAACAAAATAAAAAATTC
AAAATTTTCCGTTTATTACAGGGACAGGACAGATCCAGTTTTGGTTAATTTCA
GACAGTTATCAACCAGTTGGTACAAAAAAGGTGAACGAGAAACGTGAAATG
ATATAAAATATCAATATTTAATATTTGTCATAAAAAACAGACTACATA
ATACTGTAAAAACACACCATTCCACGTCACTATGGGGCGCCCAAATATCTAG
ATATCAAGTGCCACCTGAGCAGTCTCCCTATCAAGTACAGGAAAGTGGGACACGT
CTCGAGCTCCCTATCATGATAGAGAAAGGTACGTCCTAGAAGGTTACGTCTACCCTATCAGTGA
GTGATAGAGAAGTGACACACGTCTCGAGCTCCCTATCATGATAGAGAAAGGTACGTCCTAGAAGGTTACGTCTACCCTATCAGTGA
CGTCTAAACAGTCTCCCTATCATGATAGAGAAAGGTACGTCCTAGAAGGTTACGTCTACCCTATCAGTGA
CCTATCAGTGATAGAGAAGGTACGTCCTAGAAGGTTACGTCTACCCTATCATGATAGAGAAAGGTACGTCCTAGAAGGTTACGTCTACCCTATCAGTGA
GAAGTCGACACGTCCTCGAGCTCCCTATCATGATAGAGAAAGGTACGTCCTAGAAGGTTACGTCTACCCTATCAGTGA
ATAAGCAGAGCTCTTGGTAGAAACCCGTCCAGCTAGTCCCTGGGAGACGCACACATCCAC
GCTGTCTTGACCTCCATAGAAGAAGACACCAGGGCCAGATTCCAGCTCGAGCTACTAGGGACG
AGGTCCGCCACCATGGCCGGCCTGGAGGCTCCGTAGG
GCCGCCAGGAGTTCGAGATGGGCTGGGCGAGGGCAAGCCCTACGAGGGGCAC
CCAGACCGCAAAGCTGCAAGCCCTTCATGCACAGGATCCGGAGCCACGAACTTCT
TCATCAAGCCCTTCTCATGCGCTTTAAGGTGACATGGAGGGCTCCCTGGAACGG
CCACAGGGTTCGAGATCGAGGGGCTGGGCGAGGGCAAGCCCTACGAGGGGCAC
CCAGAACGCAAAGCTGCAAGCCCTTCATGCACAGGATCCGGAGCCACGAACTTCT
GACATCTGTGCCCCCAGTTCTTCTACGGGTCTCAACGGGTACATCAAGGACAC
CGCCGACATCCCCGACTACCCTCAAGCAGCTCTTCCCCAGGAGGCTTTCAAGATTT
GAGCGCGTGATGAACTTCGAGGACGGGCGCCCTGGTGTGACCGTGACCCAGGACT
CCTCCCTGCAGAGCCACCCTCCTCATCTACGCTGAAGTTCACGTCCGCTGAA
CTTCCCTCCCGAGGCCCGCCGTAATGCGAAGAAAGAAGTCTCTGGGCTGGGAGGCC
TCCACTGAGCGCAACTACCCCCCGCAGCGCGCTCGAAGGGCGAGAACCACA
TGCGCGCTGAAGCTGAAGCCGGCCCTACCTACCTGTGAGTTCAGATCCTCCAT
CTACATGGCAAGAAAGCAGGCGTGAAGCTGAGCTCCGCTACACTACGTGGAGCTAC
AAGCTCGACATCCACCTCCACAAACGGAGGACTACACCGTGGGTGGGAGCACTACG
AGCGCGCCAGGCCCGCCACCACCTGTTCCAGTAGGAATTCCGCGGCGCACT
CGAGATATCTAGATAGTGACTGGATATGTTGTGTTTTACAGTATTATGTAGTC
TGTTTTTTATGCAAAATCTAATTTAATATATTGATATTATATCATTTCACGTT
TCTCGTTTCAGCTTTTTCTTGTAACAAAGTGGTTGATAAACATCTCGAGCCATAAAAGA
TGTTTAATTAACCCACCCAAGACTGGGCCTGCCGCGCCGGGTTTTGGCGCCTCC
CGCGGGCGCCCCCCTCCTCACGCGGCAGCGCTGCTCCACGTCAGACGAAGGCGC
AGCGAGCGTCTGTGATCCTTTCGGCCCAGAGCAGTCAGACGACGCGCCGGCTGCT
CATAAAGACTCGGCTTAGAAGAACCAGATTACAGAGAGAGGATTTTAGGACG
GGACTTTGGGTGACTCTAGGGCAGCTGGGTTCCTTTTCCTCCAGAGAGGGGACAGGC
GAGGAAAAGTGTACCTCCTTTCGGCCTGAGAGGGGATCTCCTCGGGGCGG
GTGAACGCGAGTATTATATAAGGACGCAGCGCCGGGTTGGCAGCAGCTAGGTCC
GTCGCAGCCGGGATTGGGTCGCGGGTTTCTTTTGTGATCGGTATCAGTC
ACTTGGTGAGTACCGGGCTGCTGGCTGGCTGGCCGGGGCTTTCGCTGGCCCGG
CCGCTCCTGTTGGGACGGAAGCTGTTGGAGAGACGCAAGCGCCAAGGCTGTAAGTCGG
GTCCGCGAGCAAGGGGTCGCCCTGAACCTGGGGGTGTCGATGCTGCTGGG
TGAGGCTTCCGCTAATGCAGGAAAGCTCCTTTATCCGGGTTGAGATGGGCTGGGG
CACCATCTGGGGACCTCTGAGTGAAGGTTGCTCTGACTGAGAACACTCGGTTT
GTCGTCTGTGGGGGCGCCAGGGGAGCTATGCCCAGCTGGGCTGGGGCGAAGTCGCC
GTAACCTTTGGGACGCGCGCCCTCGTCTCTGAGTGCTGCACGCTACCCGGTTCTGTTG
GCTTATAATAGCAGGGTGAGGGCCACCTGCGGATGGTGCGGCTAGGAGCTTTTCTT
CCGTCGCAAGGACAGGCTTCCGGCCTAAGGGTAGGCGTCTCTCGGAAACTCGACAG
GCGCCAGACCTCTGTTGGAGGGAGGATAAGTAGGGCGCTGACTTTTTTCGGGT
CGGTTTTATGTCATTATCTTCTTAAAGTAGCTGAAGCTGCTCCGGGTTTGAACATATG
GAAGCTTCTTGAAGACAAACAACGTCTGTAGCGACCCTTTGCAGGCAGCGGA
ACCCCCCACCCTGGCGACAGGTGCTCTCTCCTGCGGCCAAAAGCCACGTGTATAAGA
TACACCTGCAAAGGGCGCAACAACCCAGTGCTACGTGTTGTAGTTGGATAGTT
GTGGAAAGAGGTCAAATGCTCTCTCTCAACGCTAGTCAAACAAAAAGGGCTGAAGG
ATGCCGAGAAGGTACCCCCATTTGTATGGGAATCTGATCTGGGGCCTCGGTGCA
CATGCTTTTACATGTGTTCAGGGTTTAAAAAAACGCTAGGATCGCTTACCGGTACG
CTGATGTGGACTGCAATTAATACATCAGGCATAGTGATATATCGGCATAGTATAATAC
GACAAGGTGAAGAACTAAACCATGCGCAAAGTTGACCAGTGCCGTTCCGGGTGC
TCACCACGCACCGGTACCTCGCCGGGAGCGGTCTGAGTCTGACTGGAGCGCCGACGC
TTCTCACCCGGACTCTCGTGGAGGACGACTCTCGCCGGTTGCTGGTCGGGACGAC
GTGACCCCTGGTTCATCAGCGCGTCCAGGACCAGGTTGCTGGGCCGAGAAAACACCC
TGGCCTGTTGGTTGGTTGCACGGCGCTGGACGAGCTGTACGCGAGTGGTGCAGG
GGTCGCTGTCACGAACATCTCCGGGACGCCTCCGGGGCCGAGCATGACGAGTGC
GGCAGGCAGCCGTGGGAGGCAGGTCTGCGCTGACTTGGAGGGTTTCTTACT
TGCTTTAAAAAACCTCCCACACCTCCCCCTGAAACCTGGAAACATACATGAAT
GCAATTGTGTGTTTGTTAGTCCCTCCTCATTCTGATATCAAGCTTTATCGATAATC
AACCTCTGAGTACAAATTTGTAAGAGTTGACTGGTATTTCTTAACTATGTT
GCTCCTTTACGCTATGTGGATACGCTGGTTAATGCCTTTGTACATGCTATT
GCTTCCGTCATGGCTTTCAATTTTCTCCTCCTTTGTATAAATCTCCTGTTGCTGCT
CTTTATGAGGAGTTGTGCGCCGTTGTCAGGCAACGTTGGCGTGTTGTCAGCAGTGT
GTTTGCTGACGCAACCACCTGGGTGGGGCATGCCACCACCTCTGTCAGCTCC
TTTCGGGACTTTCCGCTTTTCCCTCCCTACTATTGACGCGGGAACCTCATTGGA
GCCTGCTTTGCCCCGCTTGACGAGGGGCTCGGCTGTTGGGCACTGAATT
CGTGGTTGTTGTCGGGGAATCACTGGCTCCTTTTCTGGCTGCTGCCTGTGTT
GCCACCTGGATTCTGCGGCAGGTCTCTTCGTACGTCCCCTCGGACCCCTCAA
TCCAGCGGACTCTCCTTTCCGGCCGCGCTCTGCGGCGTCTGCGGCTCTCCGC
GTCTTCGCTTTGGCCCTCAAGCGAGGTGACACCTTTGGGCCGCCTCCCCG
CATCGATACCGTGACCTCGATCGAGACCTAGAAAAACATGGAGCAATCACA
AGTAGCAATACAGCACTACATGCTGATTGTCGCTGCTAAGACAAGCACAAG
AGGAGGAGGAGGAGGGGTGTTCTCCAGTACACCTCGATCTTCTTTAGGAAAA
GACTTACAAGGCAGCTGATGATCCTTATGCGACTTTTTAAAAAGAAAAAGGGGGA
CTGGAAGGCTAATTCACATTCCACATCACCACACAAACGAGATATCCTGGA
TCTACCACACACAAAGGCTACTTTCCGTATTGGCCAGAACTACACACACAGGC
AGGGATCAGATATCCACTGCATTTTGGATGGTGCTACAAGCTAGTACCAGTT
GAGCAAGAGAAGGTAGAAGAGCAAGCATGAAGGAGAGAACACCCGCTTGTA
CACCCTGTGAGCTGACGCTTCTCTCTCTGCGATGGGGAGACTCTGTTCTTG
GGAGGTTAAGGCCTGCTTAGCCTGCTATCGAGCGTTGGGAGCTCTCTG
GCTAACTAGGGAGACCCAGCTGTTAAGCCTCAATCAAGAGGTTGCCCGTGT
CAAGTAGTGTGTGCGGGCTCTCTGTGACTCTGGAACCTAGATCCCTCGAG
ACCCTTTTGTCTAGTGTGGGAAATCTCTAGCGAGGGCCGTTTAACCCGCTGA
TCAGCGTCGCCATTCTAGTGGCCAGCCATCTGTGTGTGGTGCCCCTCCCCC
GTGCCTCTTGGACCTGGGATGGGGCACCTCCACCTCGCTTCCTTAATAAAA
TGAGGAAAATTGCATCGCATTGTCTGAGTGATAGGTGTCATTCTATTCTGGGGGGTG
GGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATG
CTGGGGATGCGGTGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCAGCCTG
CTGCTCTTTGCTTTTCTTCCCTTCTTTCTGCACAGTCCCGGCTTTTCCCCGTCT
AAGCTCTAAATCGGGGGCTCCCTTTTAGGTTGCCGATTAGTGCTTTCACGGCAC
CTCGACCCCCAARACTTGAGTATTAGGGGTATGTTGTCACGTAGTGAGCCATCGC
CCTGATAGACCGGTTTTTCGCCCTTGACGTTGGAGTCCACGTTCTTTAATAGT
GGACTCTTGTTCACAACTGGAACAAACTCAACCCCTATCTCGGTCTATTCTTT
TGATTTTATAAGGGATTTTTCGCGATTTCGCGCTATTGGTTAAAAATGAGCTGA
TTAACAAAAATATTAACGCGAATTAATTCTGTGGAATGTGTGTCAGTTAGGGT
GTGGAAAGTCACCCAGCTCCCCCCAGCGCAGAGATGCAACAAGCATGCATCT
CAATTAGTCAGCAACCAAGGTGTGGGAAAGTCCCCAGGCTCCCCAGCAGGCAGA
AGTATGCAAAAGCATTGCATCTCAATTAGTCAGCAACCAACTCAAGTCCCGCTAA
CTCCGCCCATCCCCGCCCTAACTCCGCCCATTTCCGCCCATTTTCGCCCATCT
GGCTGACTAATTTTTTTATTTATGCAGAGGCCGACGCCGCCACCTGCTGCTCTGA
GCTATTTCCAGAAGTGATGTGGAGGAGCCCTTTTTTTGGAGGCCTAGGCTTTGCAAA
AAGCTCCCCGGAGCTTGTATATCCATTCCATTTTCCGATCTGATCAGCAGCTCTGATGAC
AATTATCATCCGGCATAGTATATCCGCAATAGTATAATACGCAAGGTGAAGGA
ACTAAACCATGGCAAGTGGACCAGTGCCGTTCCGGTGCTCACCAGGAGGAAGA
CGTCGCGCGAGCGGTGTGGACCGGCGGGCGGTGCCTGCAGGGCTGCTGGCTCGGGAC
TTCGTGGAGGACGACTTCCCGGGGTTGTGGTCCCGGGACGACGTGACCCCTGTTCA
TCAGCGCGGTCCAGGACCAGGTGGTGCACCGGAACACCCCTGGCCTGGGTGTGGTG
GGTGCGCGCCTGGACGAGCTGTACGCCGAGTGGTCGGAGGTCGTGTCCACG
AACTTCCGGACGCCTCCGGGCCGGCCATGACCGAGATCGGCGAGCAGCCGT
GGGGGCGGGAGTTCGCCCTGCGCGACCCGGCCGGCA
ACTGCGTGCACTTCGT
GGCCGAGGAGCAGGACTGACACGTGCTACGAGATTTCGATTCCACCGCCGCC
TTCTATGAAAGGTTGGGCTCAGGACTTTCGGATTCTCCGGAGCCGGCGGGTGATGAT
CCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCCACCCCAACTTGTTTA
TTGCAGCTTATAATGGTTACAAAAAATAGCATCACAATTTTCACAAAA
TAAAGCATTTTTTTTCAGTGCATTCTAGTTGTGTTTTGTCCAAACTCATCAATGT
ATCTTATCATGTCTGTATACCCGTACCTCTAGCGAGCTTGCGTGAATCAT
GGTCATAAGCTGTTCCTCTGTGTGAATTGTGTATCCGCTACAATTCACCACACAAC
ATACGAGCGGAAGCATAAAGGTGAAGCTGGGGTGCTAATGAGTGAGCT
AACTCAGTTAATTGCCTGGCTCACTGCCCCTTTCCAGTCGGGAAAACTTG
TCGTGCCCAGCTGCATTAAATGACGCAACCGCCGGGAGAGGCCTTTTG
GATATTGGGCCTTCCGCTTCCTCGCTCACTGACTGCTCGGCTGCTGGTGTT
CGGCTGCAGCGGTATCAGCTCATTCAAAGGCCGTTGATAAATCGGTTATCCA
CAGAATCAGGGGATAACGCAGGAAGAACATGTGAGCGAAGAAGCGCAGCAA
AGGCCAGGAACCAGTAAAAGGCCGCGTTGGCTGGCGTTTTGCACATAGGCTCCG
CCCCCTGAGCACGACATCACAAAAATCGAGCCTCAAGTCAAGTGTCGGGAAC
CCGACAGGACTATAAAAGGATTACACCGCCGTTTCCCCCTGGAAGTCCACCTCGTGC
GCTCTCCTGTCCGCACCCCTGGCGCTACGATACCTGTCCGCTTTTCTCCCTT
CGGGAAGCCTGGCCGCTTTTCTCATAGCTACGCTGTTGAGTGATTCTCAGTTCGGTG
TAGGTCGTTCCGCTCAAGCTGGGCTGTGTCACGAACCCCGGTTTACGCCCG
Plasmids constructed for SynExo interactome studies

>pFC15K::β

CCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTGGTT
GCCGCAAGGATCTGATGGCGCAGGGGATCAAGATCTGATCAAGAGACAGG
ATGACGGTCTGTTTCGATCGTCTTGAACAAAGATGGATTCAGGAGGGTTCTCCG
GCCGCTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCACAACAGACAATCG
GCTGCTCTGTATGCGCGCCGTGTTCCGGCTGTCAGCGCAGGGGCGCCCGGTTCTT
TTTGTAAGACCAGGACTGTCCGGTGCTGCCCTGAATGAACTGCAAGCAGAGGAG
CGCGGCTATCGTGTCGTGGCCACAGACGGCGTTCCTCGGCAGCTGTGCTCGA
CGTTGTCTCAGGAACGGGAGGACTGCTGCTATTGGGCGAAGTGCCGGGG
CAGGATCTCCTCTGTATCCTACCTTGCTCTGCTCGTCTGCCAGAAAGATATCCATCAGGC
TGATGCAATGCGCGCAGCTGCATACGCTTGGATCCGGCTACCTGACC
ACCAAGCGAAACATCGCATCGAGCGAGCACGCACTCGGATGGAAGCCGGTC
TTGTCGATCAGGATGTCTGGACGAAGAGCATCAGGGCCTCGGCAGCCAGCGA
ACTGTTCGCCAGGCTCAAGGCCTATGGCCGATGGGTGAGGATCTCGTCGT
ACTCATGGCGATGCTCTGCTTGGCCGAATATCATGGTGGAATTGGCGCTTTTC
TGGATTCATCGACTGTGGCCTGGGTGTGGGTGGCGACCGCTATCAGGACATA
GCCTTTGGCTACCCGGTGATATGGCTGAAAGAGCTTGGCGCAGATGGGCTGAC
GCTCTCTGCTTTACGGTATCGCCGCCTGCCGTTACCGGATACCTGGCTTTCT
TATCGCCTTCTTGGACGAGTCTTCTTCTGAGGCGGACTCTGGGCTTTCGAATGACC
GACCAAGCGACGCCCAACCCGGTATCGCTACCTCAAAGGCCTAATACGGTT
ATCCACAGAATCAGGGGATAACGCAGGAAAGAACATATGGAGCAAAAAGGCCA
GCAAAAGGCCAGGACCCGTAAAAGGCCGCTTGCTGGCGTTTTTCCATAGG
CTGCCGCCCGGTACGAGCTACAAAGATCGACGCTCAAAGGCTGATGGTGGC
GAAACCGACAGGACTTAAAGATAGTACCAGGCCTTCTCCCGTAAAAGGCTCCTT
CGTGCGCTCTCTGCTTGGCCTGCCCTGCTTACCGGATACCTGTCCGCCTTTCT
CCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTACGCTGTAGGTATCTCAGTT
CGGTGTAGGTCGCTCTCGCTCCAAAGCTGGGCCTGTGTCGACGAACCCCGGTCTCA
GCCCGAGCGCTCGCCTTTATCCGGGTAGTACTATCTGTCTTGGAGTCCAAACCAGGTA
GACACGACTTATCACCAGCAGCAAGCCACTGGTAGATACAGGATAGCAGG
GAGGTATGGGTAGGGCTACAGAGTGTTCTTGAAGTGGTGTTGGGCTAATGATCTC
TACTAAGAAGGACAGTATTTTGTATCTGCGGTCCTCCTGCTAAGGCAGGTACCTTT
CGGAAAAAGATGTGATAGCTCTCTTGATCCGGAAAAACACCCACCGGTGTAGG
GGTGGTTTTTTTGTTCGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATTTTC
AAGAAGGATCTCTTTGATCTTTTTCTACGGGTTGCTGACGTACGTGAACGGAAAA
CTCACGTAAAGGGATTGTTGTCATGAGATTATCAAAAAAAGGATCTTCACCTAGA
TCCTTTATAGTCGCGAATACAGGAAGCAGCAGCTGGATGGGCCCTTCGCTGG
GATGGTAGAAACCATGAAAAATGCGAGCTTCAGTGGAATAGTTTGGGGAATGG
TGCCCTGTACCCTCTGGTTGCATAGGTTTACATACGTTAAATTTATCAGGC
GCGATTCGGGCAGTTTTCGCGGTTTGGTGTTCGGCATTTCATCTGTCTGCTGC
GTGATCGCGCTGAACGGCGATGTTTAGCGGTGCTGACAATTAAAGGAGATTAGGTA
AATCCACTTACTGTCTGCCCCTCGTAGCCATCGAGATGAACCGCAGTACTCCGG
CCACGATCGCTCCGGCGTAGAGGATCGAGATCT
>pFC15K::β(1-177)

CCAAATCTAAGGGACCTTTCCAAAATGTGTAATAACCCCCGCCCCGTGACG
CAAATGGGGTGTTAGGCATGTCGATAATAGGACAGAGCTGGTT
TAGTGACACACTATAGAATAGGAGTCTAGTTAAGCAGAGCTGGTT
TTAGTGACACACTATAGAATAAGGAGCGATCGCCATGAGTACTGCACTCGCA
ACGCTGGCTGGGAAGCTGGCTGAACGTGTCGGCATGGATTCTGTCGACCCAC
AGGAACTGATACCTCTTTGCCAGCAGCCAGACCCATTTAAAAGTGCATGCTCCAGCG
TGCGCAGTTCATCGATTACCTGACTCGTACGGCATGAGTACTGCACTCGCA
GGACGAAGAGAAATTTACGCTTTTCTGTGATAAAGGCAATGTCATCGTTCGCT
GGTGGGCGTTGATGCTGGCTCAGCAGCATCATCAAATGAAAGCAGCAGTTGAT
GGCATGGACCTTTGAGCGGCAACTGAAATCCTGTACATGCTACGAGCCGAT
AAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCC
TTGGGGCGGCGCTTCGAGCAGACATGATAAAGATACATTGATGAGTTTGGAC
AAACCACAACCTAGAATGCAGTGAAAAAATGCTTTTATTITGAAATTTTGTA
TGCTATTGCTTTTATTTGTAACCATTTAAAGCTGCAATAAAACAAGTTAACAAC
ACAATTGCATTCATTATATGTTTTCCAGGTCAGGGAGATGTTGGAAGGTTTTT
TTAAGCAAGTAAAAACCTCTCTAAATGTTGTTAAAAATCGAAATTCTAATGGATCC
TCTTTCGCGTTGCCTTTCCCTTGGTCCAGATAGCCCGACTGAGCAGGATTTCCATCC
GGGGTCAGCACCCGTTCAGGACTGGCTTTCTACGTGTTCCGTTCTCTTTAG
CAGCCCTTGCGCCCTGAGTGCTGGCGACAGCTGACTTCAAAAGAATTGCC
AGCTGGGGCGCCCTCTGGTAAAGGGGAAGCCTTGCAAAATGAAACCTGGATG
GCTTTCTTGGCAGGATCTGATGGCGCAGGGGATCAAGATCTGATCAAG
AGACACGAAGATGACGGCTTTCCATGCTTTGAACAAAGATGGAATTGCACGCAG
TTCTCCGGCCGCTTGGGAGGAGGTATTCCGCTATGACTGGGCAACAACAG
ACAATCGGCTGCTCTGATGCCGCGCTTCCGGCTGTCAGCGCAGGGCGGCC
CGGTTCCTTTGGTCAAGACCGAACCGTCTGCCTGGTGCTGCTTCCGGCTGTCAGCGCAG
GAGGCAGCGCGCTATCGTGGCTGGCCACGACGGGGGCGTTCTTGCGCAGCTG
TGCTCGACGTTGTCACTGAAGCGGAAGGGACTGGGCTGCTATGACTGGGCAAGT
GCCGGGGCAGGATCCTCTGTCATCTCACAATTGTGCTTGCTGCTCCGAGAAAAATCCA
TCATGGCTGATGCAATGCGGCGGCTGCATACGCTTGATCCGGCTACCTGCC
ATTCGACCACAAAGCGAAGATCGCAGCAGGAGCAGCAGCAGCAGCCATCGGAGTGA
AGCCGGTCTTGTCCAGTGGATCTGGACGAAAGCAGCATCAGGGGCCGCG
CCAGCCGAACTGTTGCGCCAGGCTCAAGGCGCGTATGCCGGATGGTGAGGATCG
TCGTCGTAATCATGGCGATGGCCTGCTTGCGCTTGGCGAAATATCGTGGGAAATGG
CCGCTTTTCTGGATTACATCGACTGTGGCAGCCGCTGGGTGTGGCGGACCGCTATCACGAGCATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGAGCGGGACTCTGGGGTTCGAAATGACCGACCAAGCGACGCCCAACCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAAAGGGTGACGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTCAGTCAGCTCCCTTCTCCGGAAAGCGTGCGGCCTTCTCTCATAGCTCAGCTGTAGGTAATCTCAGTTTCGGTGTAAGTGCTTGCCTCAAGCTGGGCTGTGGTCACGAACCCCGTTCAGCCCGACCGCTGCCCTTATCCGGTAACTATCGTCTTGAGTCAAACCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATTTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGTCATGAGATTATCAAAAAGGATCTACCTAGATCCTTTTATAGTCCGGAAATACAGGAACGCACGCTGGATGGCCCTTCGGCTGGGATGGTGAAACCATGAAAAATGGGGTAATGTGGCCTGTACCCTCTGGTTGCATAGGTATTCATACGGTTAAAAT
TTATCAGGCGCGATTGCGGCAGTTTTTCGGGTGGTTTGTTGCCATTTTTACCTG
TCTGCTGCCGTGATCGCGCTGAACGCGTTTTAGCGGTGCGTACAATTAAGGG
ATTATGGTAAATCCACTTACTGTCTGCCCTCGTACGCCATCGAGATAAACCGCA
GTACTCCGGGCCACGATGCGTCCGGCCGTAGAGGATCGAGATCT

> pFC15K::Exo

CCAAAATCAACGGGACTTTTCCAAAATGTCGTAATAACCCCGCCCGTGAGACG
CAAATGGGCGGTAGGCGTGTACGGGAGGTCTATATAAGCAGAGCTGGTT
TAGTGAACCGTACAGTACGCTGAACGGCCATGGATTCTGTCGACCCAC
AGGAACTGATCACCACTCTTCGCCAGACGGCATTTAAGGTGATGCCAGCGA
TGCGCAGTTCATCGCATTACTGATCGTCTGGCAACCAGTACGGCTTAAATCCGTGGACGAAAGAAATTTACGCCTTTCCTGATAAGCAGAATGGCATCGTTCCGGTG
GGTGGGCCGTGGATGCTGGTGTCGCCACATCATCAATGAAACCAGTACGGCTTATCTGAAGCCATCAGATGATGGAATGCGCGGATTTACCGCA
AGGACCCTAAATCATCCGATCTGGTTACCAGATGGATGGAATGCGCGGATTACCGCA
GCATCTCCAAACGGATGTACGTGCTATAAAGCCATGATTCAGTGTGCCGTCTTG
GCCTTGCGATTTCGTGATCTCTATCTGACAAAGGATGAAGCCAGCTGGCAGATTGCA
ATCGGTACTGGCTTTCCATTCGACCCCCCATTATGTGGAAGTCTCCTGGGCAAGCG
CATGCACTACGTGATGTTGGTCCCGCGCAGGATGGCACCCTCTGTGCTGTTCCTGC
ACGGTAACCCCGACCTCCTCTACGTGTTGGCGCAACATCATCCTCCGCATGGTTC
CCGACCCATCGCTGCATTGCTCCAGACCTGATCGGTATGGAAGCCATCCGACA
AACCAGACCTGGAATTATTTCTTGACGACCACTCCGCTCTCATGGGATGCCCTTC
ATCTGAAGCCCTGGGTCTGGAAGAGGTCGTCCTGGTCATTCAGCTGGGGCT
CCGCTCTGCGCTTCCACTGGGCAAGCGCAATCCAGAGCGCGTCAAAGGTAT
TGCAATTTATGGAAGTTACCATCCTCCGACCTCCTGGGACGAATGGCGAATGGCCGAAT
TTGCCCGCGAGAGCTCTCCAGGCGCTCCGCAACCAGCGACGTCGGCGCGCAAGCT
GATCATCGATCAGAACGTTTTTATCGAGGGTACGCTGCCGATGGGTGTCGTCC
GCCCGCTGACTGACGATGGCACCATTACCGCGAGCGCGTCTCTGGCTCCATCC
TGGTGCCGACGACCTGCGCAGGGCGTCTGGCGAGAATACATGGGACTGGGTCCG
GAGCCACGCAACATCGTCGCGCAGGCTGCAAGAATACATGGAGCTGGCTGCACC
AGTCCCCCTGTCCGCGAAGCTGCTGTTCTCGGCGACCCCGAGGCGTCTGATCCCA
CCG GCC GAAGC GCCTCGCCCTGCGCCAAAGCCTGCTAACTGCAAGGCTGTGG
ACATCGGCCCGGGTCTGAAATCTGCTGCAAGAAGACAACC GGACCTGATCGG
CAGCGAGATCGCGCGCTGGCTGTCTACTCTGGAGATTTCCGGTTAATAGAATT
CTAGAGTCGACCTGCAGGCGCAGCATGCAAGGCTGACATTGATGAGTTTGGA
AAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAAGCAATAACTAGCATACCAACCA
TTGGGCCGCCCGCTTCGAGACAGACATGATAAAGATTTCCGGTTAATAGAATT
AAACCACAAACTGAATGCAGTGAAGCAACTTATTGAAATTGCTGAATTTTGA
TGCTATTGCTTTTATTAGCTTAAATAGCGCAATAAAGCGTCAATAAAAGTTAACCA
ACAATTGCATATTATTTATGTTTTCAGGGTCAAGGGAGATGGGAGGTTTTTT
TTAAGCAAAGTAAAAACCTCCTACAAATGTTGTGTAATATGCAGATTGCTGC
TCTTTGCCGTGGTTTTCCCTTGTGGTGATTGAGATGAGCTGACATTCACTCC
GGGTCAGCACCCTTTCTCCGGGACTGGCTTTTCTACGTGTTCCGGCTTTTAG
CAGCCCTGGCGCCCTGAGTGAGCTGACATTCACTCC
AGCTGGGGCCCGCCCTCTGGTTAGTTGGGAAGCGCAGGGATCGCTGGAAGGAT
GCTTTCTTGGGCAAGGATCTGGGCTGGCAGGGATCAAGATCTGCTGGGAT
AGACAGGATGACGGTCGTTTGGCATGCTTGAACAGATGGGATTGCACGCAGG
TTCTCCGGGCGCTGGGGATGGAGAGCTATTCCGGCTATGGACTGGGGCACAACAG
ACAATTGGCGCTCTGATGCAGCCGGCCTGTTCGGCGTGCAGCCAGGGCAGCC
CGTCTTTTTTGTCAGAGCAGCAGCTGCTGGGCTGGCAGGGAGACTGACGGGAG
GAGGCACAGCGCGTATCGTGCTGGCCACAGCAGCGCGTCTCAGCGGCTGAGT
TGCTCGACCTGTCGACTGAACGGGAAAGGGACTGGGCTGTAAAGTGGGCGAAGT
GCCGGGGCAGGATCTCCTGTCATCTCAGCTTGGCTCCGAGAAAGTATCCAG
TCATGGCTGAGCAATGCGGGCGCGCTGCAGGCTGCTGACCTGGCTACAAGGCTGCCCC
ATTCGACCACCAAGCGAACCACTCGACATCGAGCGGACGCACACTCGGATGGA
AGCCCGGTCTTGTGATCGATACGGGATGATCTGGACGAAGACGACATCGAGGCTCGG
CCAGCCGAAACTGTGGCAGCGCTTCTGCAAGCCTGGACGTGGGTGGAACGGCTATC
TCGTCGTAACCTGAGGCTGCTGCTTGGCCGAAATATCATGGTGGAAGATGG
CCGCTTTTCTGGAGTTATCGACTTGGTGGCGGCTTGTGGGCTGCGGACCCGTATC
AGGACATAGCGTGGCTACCGGTATATTGCTGAAGAGCTTGGGCAGCGGAATG
GGCTGACCCTTTCTCGTGCTTTACGGTATCGCCGCTCCGGATTGCACACGCGCA
TCG CCTTCTATCGCCTTCTGGACGAGTTCTTCTTCTGAGCGCGGACTCTGGGAGT
AAATGACCGACAAAGCGACGCCAAACCGGTATCAGCTCACTCAAAGGCGGTAA
ATACGTTATCCACAAGAATACGGGGATACGCAGGAAACAACTATGTGAGCA
AAAGGCCAGAAAAAGGCGAGGACGTAAGAAAGGCCGCAGGGCTGAGCGTTTT
TCCATAGGGCTCAGCCCTGACGATCATCCAAAATCGACGCTCAAAGGCTAG
GAGGGCCAGAAACCAGACCAGACTTATAAGATACCCAGGCGGTCTCCCTCCTGGA
AGCTCCTCTCGTGCCTCTCTGCTGCTCAACCCGGTATACCTGTC
CGCCTTTCTCCCTTCGGGAAGCGTGCCAGGCTTCTCTAGTCACTGACGCTCAGGCAT
ATCTCAGTTTCGTTAGGTCGTTTCGCTCAAGCTGCGTGTGCACGAAACC
CCCCTTCAGCCCCAGCCGCTGCCTTATCCGGTAACTATCGTCTTGGAGTCCAA
CCCCTAAAGCACGACTTATCGCCAATGGCAGCAGCCACTGGTAACAGGATT
AGCAGACGCGGTATGTTAGGCGGTGTACAGGAAGTGTTGGTGCGCA
ACTACGGCTACACTGAAAGACGATTATTTGTTATCTGCCGCTCTGCTGAAGCC
AGTTACCTTCCGAAAAAGGTTGTTAGCTTCTTGAACCGCAAAACCAACCCAC
GCTGGTAGGCTGCTTTTTTGTTTACGGAACAGCAGATACCGCGCAAGAAAA
AAGGATTTCAAGAAGACCTTTTGATTTTCTACGGGTTCTGACGCTAGTG
AACGAAAAACTCACGTAAAGGGATTTTGGTCATGAGATTATCAAAAAAGGATCT
TCACCTAGATCTTTTATAGTCCGAAATACAGGAACGCACGCTGGATGGCC
CTTCGCTGGGATGGTGAAACCATGAAAAATGGCAGCTTCAGTGGATTAAGTG
GGGGTAATGTTGGCCTGTACCCTCTCGTTGATAGGTATTCATACGGTTAAAAT
TTATCAGGCGCGGATTTGCGGCAATTTTTTCCGGTGTTTGTGCCATTCTTACCTG
TCTGCTGCCGTCGATCGCGCTGAAACCGGTTCCTTATGCGGTAGCAATTAAAGGG
ATTATGGTAATCCACTTACTGTCTGCCCTCGTAGCCATCGGATAAACCCGCA
GTACTCCGGCCACGATCGTCCCGGCGTAgAGGATCGAGATCT

>pFC15K::Exo(1-189)
>pFC15K::Exo(1-189)
CCAAATAACGGGACCTTCCAAAATGTGTAATAACCCCGCCCCCGTTGACG
CAAATGGCCGTTAGCCTGTACGGTGGGAGGCTATATATAAGCAGAGCCTGGTT
TTTAGGTGACACTATAAGGAGCGATCGCCATGACACCCGAACAGGGATGATGC
CTGCAGCGTACCGGGATCGATGTGAGAGCTGTCGAACAGGGGATGATGCGT
GGCACAAATTACGGCTCGGCGTCATCACCGCTTCAGAAGTTCACAACGTGAT
AGCAAAACCCCGCTCCGGAAAGAAGTGGCCTGACATGAAAATGTCCTACTTC
CACACCCTGCTTTGCTGAGGTGGTTCACCAGGTTGCTGCCGAAGTTAACAGCTAA
AGCAGTGGCCTGGGGAAACAGTACGAGAAGCGACGCAGAAGCAGAAGCTTGA
GTTCACTTCCGGCCTGAATGTTACTGAATCCCGATCATCTATCGCGACGAAA
GTATGCGTACCGCCTGCTCTCCCGATGGTTATGCAGTGACGGCAACGGCCTT
GAACTGAAATGCCCCGTTACCTCCGGGATTTTCATGAAGTTCAGCCTCGGTG
TTTGGACGCTTTGACGTTGATGTTTCTCTCGAGCCAACACTGAGGAT
CTGTACTTTCAGGCTATGGTACCACTTTGCACCCGAACAGCGATGGATT
GGTGACGCCGAAAAATGACCTGGTATATGGCACCCTATGACCCGCGGTATGAAGC
GTTAAGGCGCTGATTATGTCGATTGTTTCTCTCGAGCCAACCAGCTGAGGAT
CTGTACTTTTCAGAGCGATAACGATGGATCCGAAATCGGTACTGGCTTTCCATT
CGACCCCCATTATGTGGAAGTCTGGGCGGCGACGTGACTACGTCGATGT
GGTCGCCGCGATGGGCACCCCTGTGCTTCTGTCACGGAACCGACCTCCCTC
CTACGTGTGGCCACATCATCCCCGATGGTGACCGACCCCATCGCTGCATTG
CTCCAGACCTGATCGGTATGCGAGTCAATCCGACAACAGACCTGGGTTATTT
CTTCGACGCCCTGGGCTTCCATGAGTGCCTTCACGAAGACCTGGGTTGCTG
AAGAGGTCGTCCTGGTCAATTCCACGACTGGGGGCTCCGCCTCTGGGTTTCCACTGG
CGTTTCGCATGCTTGAAACAGATGGATTGCACGCAGGTTCTCCGGCGCTTGG
GTGGAGAGGCTATTCCGCATGACTGGGCACAACAGAATACGAGAGGAGCGAGCGCAGCTAT
CGTGGGCTGGCCACGGGCTGTCA
GCGCAGGGGCGCCCGTTTCTTGTCAAG
ACCGACCTGTCCGGATGTCATCGGATGTAAT
GCAGGGCTGCATACGC
TTGATCCGGCTACCTGCCCATTCGACCACCAAGCG
AAACATCGCATCGAGCGAGCACGCACTCGGATGGAAGCCGGTCTTGTCGATCG
AGGATGATCTGGACGAAGACAGCATCAGGGGTCGCCGAGGCCAAGCTTGTCCGC
CAGGCTCAAGGCCTATGCGCGGTGGTGAAGATCTCGTCGACTCATGGGC
GATGCCTGCTTGCCAAATATCATGGTGGAATGAGCCGGCTTTCTGGATTATCAT
CGACTGTGGCCGGCTGGTGGTGGCGGAGCACCCTATCAGGACATAGCGTTGGCT
ACCCGTGATATTGCTGAAGAGCTGGCGGCGAATGGGCTGACCGCTTCCTCG
TGCTTTACGGTATCGGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCC
CTTGACGAGTTCTTCTGAGCAGGGACTCTGGGTTCAATAGCCGAAAGCC
GACGCCCAACCGGTATCAGCTCAGTCAAGGGGTCGTACGGGTATACGGTTATCCACAAG
AATCAGGGGATAACGCAGGAAAGAACAGATGTGAGCAGAAAAAGGCGAGCAAAAGG
CCAGGAACCGTAAAGGGCCGCGTTGCTGGCGTTTTCTCATAGGCTCCGCC
CCCTGACGAGCAGCATCAAAAAATCGACGCTCAAGTCAGGGGCAGGAACCC
GACAGGACTATAAGATACCGAGGCGTTTCTCGGCCCTGGAAGCTCCCTCCTCG
TCTCCTGTACCTGAGACCTACCGGATAACCTGTCCGGCGCCTTCTTCTCCCTCC
GGAAGCGTGCGCCTTTCTCATAGCCTACAGCTGTAGGTATCTCAGGGGTGTA
GGTCGGTCGGGTCAAGGCTGGCTGTGACGACGCACACCCGTCAGCCGAGAC
CGCTCGCGCCTTATCCGGGTAAGCTATCTCTTGAGTCAAGTGAACACCCGGGTAAGAGAC
ACTTAGCCTACAGGCTTTGAAGGCAATGCTGGGCTAAACTACGGGGTACACT
AGAAGACAGATATTGCTGATCTGACGCTCTGGAGCTAAGCCACATGTTACCTTCCGGAA
AAAGAAGTCAGGCTTGCTACCCGCAAAACACCCAGCGCTTGTCATAGGGTG
TTTTTTGTGGCAACGCAGCAGATTACCACGCAGAAGAAAAAGGATTTTCAAGAA
GATCCTTTGATCTTCTTCTACGGGTCCTGACGCTCTAGTGGAACGAAACTCACG
TTAAGGAGATTTTGGTCATGAGATTATCAAAAAGGATCTCACCTAAGATCCTTT
TATAGTCGGAAATACAGGAACGACGCTGGATGGCCCTGCAGGGATGGGT
GAAACCAGGAAAAATGCAAGCTCCATGGAATTAGTGAGGGGTTAATAGTGCCCT
GTACCGCTGTGGATAGGTATTACGCTGGTAAACTATATACGGCCATTATCAGGCACGATT
GCCGCAGTTTTTCCGGGTGTTTTGGCCATTTTTACTTGCTGTGCTGCCGTATC
GCGCTGAACCGCGTTAGCGGTGCTACTAATTAAGGGATTATGGTAAAATCCA
CTTACTGCTGGCCTCGTATCGGAGATAAACCCAAGCTGAGAATGACT
ATGCGTCCGCGGTAGAGGATCGAGATCCT
>pFN22K::β

CCAAAATCAACGGGACTTTCCAAAAATGTGTAATAAACCCCGCCCCCGTTGACG
CAAATGGGCGGTAGGCTGTACGGTGGGAGGTCTATATAAGCAGAGCTGGTT
TAGTGAACGTCAGATCACTAGAAGCTT
AATACGACTCACTATAGGGTACGA

TTAGGGTACACTATAGAATAAGGAGAATTCATCATCAGCAGAAATCGGTACT
GGCTTTTCCATTGACCCCCCATTATGTGGAAGTCTGGGCGACGGCATGCACTA
CGTCGATGTTGGTCGCCGCGATGGGACCCTTCTGACTGTTTCCCTCGACGGTAAACC
CGACCTCCCTCTACTCGTGTTGGCGCAACATCATCACCCTGGCATGTTGCAACCGACC
CGCTGCAATTTGCTCCAGACCTGATCGGATATGGCAGAAAATCAGCAACAGACC
TGGGTTATTTTCTTTGACGACACGACTCGGCCTCTCCGCTCCCTTCACTGGAAGTGGCGGTACGA

CTGGGTCTGGAAGAGGTCGTCCTGGTCATTCACGACTGGGGCTCCGCTCTGG
TGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTT
GATCCGGCTACCTGCCCATTGACCAACCAAGCGAATATCGCATACGAGCGAG
CACGCACTCGGATGGAAGCCGCTTGTCGATCAGGATGATCTGGACGAAGA
CCGGATGGTGGAGGATCTCGTCTGACTCATGGCGATGCGCTGCTTTGCCGAAATAT
CATGGTGGAAAAATGCGCCCTTTCTGGATTTACATCGACTGTTGCGCCGCTGGGTG
TGCGGACCGCTATCAGGACATACGGCTTTGCTACCCGCTGATATTGCTGAGA
GCTTTGGCCGGAATGGGCTGACCCTCTTCTCGTGCTTTACGGTATCGCGGCTC
CCGATTTCGAGCGCATCGCCTTCTATCGCTTTCTGAGAGTTCTTCTGAGCG
GGACTCTGGGGTTCGAAATGACCGAGCCAAGCGACGGCCAACCGGTATAGCT
ACTTGGGTTCTCAGGAAATGACCGAGCCAAGCGACGGCCAACCGGTATACGCT
CAGCTCAAAGGGCGGTAAATACCGTTATCCACAGAATCAGGGGATAACGCAGGA
AAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCC
GCGTTGCTGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAA
TCGACGCTCAAGTCAGAGGCTCGAAGCCACGACTATAAAAGATACCA
GCCGTTTTTTTCGGAAGCTCCGCCCCCTGCGACGAGCATCACAAAAA
TTACCGGATACCTGTCCGCTTTTCCTCCCTTCGGGAACGGTGCCGCTTTTCTCAT
AGCTCACGCTGTTAGGTATCTCAGTTCGGTGTTAGTTCGGCTTCTCCGCTTACAGTGGG
CTGTCACGACGAAACCCCCGTTTCAGCCACGCCCTGCGCTTTTATCCGGTAACCT
ATCGTCTTGAGTCAACCCGGAAGACAGACGACTATTAGCAG CCTGCGACGAG
CACTGGCTAACCGATTAGCAAGACGGGTATGTTAGGCCTGCACGATGTTCTCTT
TTGAAATGGGTGGCCATACGACTTACAGAAGGACAGTATATTGCTATCT
GCCGCTCTGAGGCAAGGCCATGCTTTTCTCCGGAAAAAGAGTTGGTAGCTTGTGATCC
GGCAAACAACGACGGCCTGAGGCGTGGTTTTTTTGTGTTGCAAGCAGCGAG
TTACGCGCAGAAAAAAAGGATTTCAAGAAGATCCTTTGATCTTTTCTACGGG
GTCTGACGCTCAGTGGAAACGAAAAACTCACGTTAAGGGATTITTGGTCATGAGA
TTATCAAAAAAGGATCTTCACCTAGATCTTTTATAGTCCCGAAATACAGGAAC
GCACGCTGGATGGGCCCTTCGCTGGGATGGTAGGAAACCATGAAAAATGCGAGCT
TCAGTGGATTAAGGCGGCAATTGGGCCCTTGGTACCGATAGGTAT
TCATACGACGTTAATTTATACAGGCGCGATTGCGGCGATTTTTCGGGTGTTTG
TTGCCATTATTACCTGTCTGCTGCCTGATCGCGTGAACGCCTTTTACACGCT
GCGTACAATTAAGGGATTATGGTAAATCCACTTACTGCTGCCTCTGAGCCA
TCGAGATAAAACCCGAGTACTGCCGCAAATCAGTGTCGCCGCGTAGAGGATCGA
GATCT

>pFN22K::β(1-177)
>pFN22K::β(1-177)
CCAAAAATCAACGGGACCTTTCCAAATTGTCGTAATAACCCCGCCCGTTGACG
CAATGGGCCGTTAGGGCGTGAACGCTTTGAGAGTTTATATAAGCAGAGCTGGTT
TAGTGAACCGTCAGATCAGACTAGAAGCTTTAATACGACTCACTATAGGGTACGA
TTAGGTGACACTATAGAATAAGGAGAATTTCATCATGGCACAGAAAATCGGTACT
GGCTTTCATTCCAGAATGCCAATTATAGGGAAGCTTTGCTGGGCGAGCGCATGCC
CGTCGATGTGTTGTCGGCGGACCCACCATCATCCCAGCATGTGTTGCACCGACCCAT
CGCTGCAATTGCTCCAGACCTGATCGGTATGGGCAATCCGACAAACCAGACC
TGGGTATTTTCTTCCAGAACCACCTCCGCTTTCATGGATGCCTTCTCACGGAAGCC
CTGGGCTGGTTGGAAGGTTGCTTCAGCTTTCCCGCTTACGACTGGGGGCTCCGCTCTGG
GTTTCCTAGGGCAGCAGAATCCGAGAATCCGAGTACATTTAT
GGAGGGTCAGGTCAGATGGACATAGAGGCTTTGCTGAAACGGCTTCCAAACGAGCTGCCAATCGCCGGTGAGCCAGCGAACATCGTCGCGCTGGTCGAAGAACATGGACTGGCTGCACCAGTCCCCT
GTCCCGAAGCTGCTGTTCTGGGGCACCCCAGCGTTCTGATCCCACCGGCCGAAGCCGCTCGCCTGGCCAAAAGCCTGCCTAACTGCAAGGCTGTG
GACATCGGGCTGGCTGACGCTCGAGATTTCCGGCGAGCAACCACTGAGG
ATCTGTACTTTCAGAGCGATACGCGATCGCCATGAGTAC
TGCACTCGCAAC

AAGCCGCTCGGCTTGCCAAAAGGCTGCTAACTGCAAGGCTTGAGCACAG
CCCGGCTGTGAATCTGCTGGCAAGAAGACAAACCCGAGCCTGATCGGCAGCGAG
ATCGCGCGTGGCTGACGCTCGAGATTTCGCGAGAAGCGCAACCGATCGCAAC
GCTGGCTGGGAAGCTGGCTGATGGATTTCTGCGACCACAG
GAACTGATCCACACTCTTCGCCAGACGGCATTAAAGGTGATGCCAGCGATG
CGCAGTTCATCGCATATTACTGATCGTTGGCCAACCAGTACGGCCTTTAACGTCG
ACGAAAGAAATTACGCCTTTCCTGATAAAGCAGAAATGGCAGCTCGGATTTTACG
TGGGCCGGATGTTGGCTGGTGCCCAGCATCATAATGAAAAACCAGCAGTTTGATGG
CATGGAACTTTTAGGCAAGGAATACGATCTTGGATCCCGGATTTACCGAGTG
GACCGTAATCATCCGATCTGCTTGATCGCTACGAGCAGCTCCGCTCTGCGG
AACCATTTCAAACCTCGAAGCGAGAAATACCGGGGCGTGCAGTGC
ATCCAAACCGGATGTTACGTAATAAGCCATGATTGCTGGCCGTCTGGCC
TTCGGGATTTGCTGGTATCTATGACAAGGTGAAAGCCAGCGCCGTTTTAAAACGA
ATTCGGGCTCGGTACCCCGGGATTCCTCTAGAGTCGACCTGAGGCATGCAAG
CTGATCCCGCTGCTAAACAAAAGCCCCAAAAGGAAGCTGAGTTGGCTGGTGCTG
CGCTGAGCAATAACTAGCATAAACCCCTTTGGGGCGCCGCTCCGAGCAGACAT
GATAAGATACATTGATGAGTTTGGACAAAAACCAACTAGAATGCAAGGAAA
AAATGCCCCATTTTTGGTACAAAAGGTGATCTATTGCTTTATTTGTAACCATTATA
AGCTGCAATAAAACAAAGTTAACAAACCAATTGCATTCAATTATGTGTTACGGT
TCAGGGGGAGATGTGGGAGGTTTTTTTAAGCAAGATAAACCTCTACAAATGT
GGTAAAAATCGAATTTCAATGGATCTCCTCTTTGCGTTCGTGTTTCCCTTGCC
GATAGCCCGATGCATTACATCCCGGGGTACGACCCCTTTCTGCGGACTG
GCTTCTACGTTGTCGCCCTCTTCTTACGACCGCCGCTTGCGGAGTGCGTTTCG
CAGCGTGAAGCTCCAAGAATGCGGCGAGCTGAGGCGCCCTCTGCTGTAAGGTTGG
GAAGGCCCTGCCAAAAGATACGATGAGGTCTTTCTGCGCCAAGGATCTGATG
CGCAGGGGATCGATCACAGAGACAGGATACGGGTCGGTTTCGCGATGC
>pFN22K::Exo

CCAAAATCAACGGGACTTTCCAAAAATGTCGTAATAAACCCCGCCCCCGTTGACG
CAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTGGTT
TAGTGAACCTAGCTATAGAATACGACTCATGACGATTACGACTGATTACGA
TTTAGGTGACACTATAGAATAAGGAGAAATTATCATATGCGAACAAACTCGGTACT
GGCTTTCATTTCGACCCCTTATATGGAAGTGAGTTGCTGCTGCAGCGCGCATGAC
CGACCTCCTCATGCTGTGGCGCGAACATCGATCCGCAATTGATGCTGGCCTACCCCAT
CGCTGATGTTGTCGCGCGCATGGCAATCCGACAAACCAGACCCGTTGGTATTT
CTCGGTCACGGCACTGCTGGCCGGTTCGTTTCACTCAGGCACTTCAACGCGCTAA

> pFN22K::Exo

CCAAAATCAACGGGACTTTCCAAAAATGTCGTAATAAACCCCGCCCCCGTTGACG
CAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTGGTT
TAGTGAACCTAGCTATAGAATACGACTCATGACGATTACGACTGATTACGA
TTTAGGTGACACTATAGAATAAGGAGAAATTATCATATGCGAACAAACTCGGTACT
GGCTTTCATTTCGACCCCTTATATGGAAGTGAGTTGCTGCTGCAGCGCGCATGAC
CGACCTCCTCATGCTGTGGCGCGAACATCGATCCGCAATTGATGCTGGCCTACCCCAT
CGCTGATGTTGTCGCGCGCATGGCAATCCGACAAACCAGACCCGTTGGTATTT
CTCGGTCACGGCACTGCTGGCCGGTTCGTTTCACTCAGGCACTTCAACGCGCTAA
GGAGTTTCATCCGCCTCCTACCCGCACCTGGGACGAATGGCGCAGAATTTGCAC
GAGACCTGCTCCAGGCTCTCCGACACCAGCCGATCGGCCCGCAAGCTGATCATCG
ATCAGAAAGCTTTTATCGAGGCTACGCTGCCGATGGGTGTCGTCCGCCCGCTG
ACTGAAAGTCCAGATGGAGACCATTACCGCAGGCCGTTCTCTGAATTTGGTGA
GCGAGGCCACTGTGGCGGCTCTCCACAAGCGAGCTGCAATCGCCCGGTAGCGCAG
GAACATCGTCGCGCTGGTGCAAGGAATACATGGACTGGCTGCAACCAGTCCCCT
GTCCCGAAGCTGCTGTTCTGGGGCAACCAGCGGTCTTGATCCCACCGGCCG
AAGCCGCTCGCCCTGGGCAAAAGCTCCTGCTAATTGCAAGGCTGTTGGACATCGG
CCCGGGTCTGAATCTGCTGC
AAGAAGACAACCCGGACCTG
ATCGGCAGCGAAG
ATCGCGCGCTGGCTGTCGAC
GCTCGAGATTCCCAGCGAG
ATCTGTACTTTTCATAAGCGATAACCGCAGTCGCCATGAGTACTGCACTCGCAAC
GCTGCTGTGACGTGCTGCTGCGATGGCATGGCATGGATTCTGTCGACCCACAG
GAACGTGACACCACACTTCTTCGGCAGACGGCATTAAAAGGTGATGCGCAAGCGAT
CGCAGTTTCATCGCATTACTGATCGTGTGCGCAACGATCGCCCTTAATCCGTGG
ACGAAAGAAAATTACGCTTCTTCTTGTAAGCAAGAATGGCATCGTCCGCCTGG
TGGGCTTGGATGGCTGCTCCGCATCATAAGAAAACCAGCAGTTTGATGGA
CATGGACTTTTGAGCAAGGACATGATGCATGCGGATTTCACCGCAAAG
GACCGTAATCATCGGATCTGCTGGTTACCGAATGGATGGAATGGCAGCGCGC
AACCATTGAAACTCGCGAAGGCAGAATCGGCGGCGGTGGGCGCAGTGC
ATCCAAAACGGATGTTACGTCTAAAAACCGCATGATCTGCTGCCGCTGTCGGCC
TTCCGATTGTACGTCATTACAAAGCAGCATGATCGTGTGCCGCTGTCGGCC
ATCGGGGCTCGGTACCCGGGATCTTTCTCGGACCTGCGACGGCATGCAAG
CTGATCCGGCTGCTAAACAAAGCCGAAAGGAGCTGAGTTGGCTGCAAC
CGCTGAGCAATAACTAGCATAAACCCCTTGGGCGGCGGCTTCGAGCAGACAT
GATAAGATACATTGATGAGTTTGAGACAAACCACAACTAGAATGCAGTGAAAA
AAATGCTTTTATTTGTGAAAATTGTGATGCTATTGCTTTTATTTGTAACCATTATA
AGCTGCAATAAAACAAGTAAACAAACAAATTTGCATCATTCTTTTATTTGTTCAGGT
TCAGGGGGAGATGTTGGAGGTTTTTTTTAAGCAAGTTAAAACCTCTACAAATG
GGTAAAATCGAATTCTAATGGATCCTCTTTGCGCTTGCGTTTTCCCTTGTCCA
GATAGCCCATGAGCTGACACATTCCCGGGGTCAAGACCGTTCTCTGCGGACTG
GCTTTCTACGTGTTCCTCGTTCTCCTTTAGCAGCCCTTGTGGCTTGGCAGG
CAGCGTGAGCTTCAAAAGAATTGCAGCTGGGCGCCCCTCTGGAAGGTTG
GAAGCCCTGCAAAAAGACTGGATGCTTTTTTTTGCAGGGCAAGGATCGATGG
CGCAGGGGATCAAGATCTGATCAAGAGACAGGATGACGGTCGTTCTGCGATGC
TGAACAAGATGGGATTGCACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCT
ATTCGGCTATGACTGGGCACAACAGAAACTCGGTCTGTCTGTGATGGCCCGCTG
TTCCGGCTGTCAACGCAAGGGCGCCGGTTCGGGTTCTTGGGTGGAGAGGCT
CGTGCTGTGCAAGCAGACGAGCCAGACCGCCGGCTATCGTGCGACGTGGCC
ACGACGGGCGTCTCTTGCGCAGCTGTGCTCTGTCTGATGCTGTCGTGGCGCC
GGGACTGGCGTCTATTGGGCGAAATGCTCGGGGAGAGGATCTCTCGTGTAC
CCTTGTCCCTGCGCAGAAAATGATCCATCATGCTGATGCAATGCGGCAGGCT
CATACGCTTTGATCCGGCTACCTGCCCATTCGACCCAAACGCGAAACTCGCA
TCGAGCGAGCAACCTCGGATGGAAGCCCGGTCTTGTGCGATCAGAGTGATCT
GGCGAAGAGCATAGGAGGCTCAGCAGCGCAACTGTTGCGCAGGCTCAT
GGCCCGTATGCAGGATGAGTACCTCGTGACTCATGGCAGTGCTGCTG
TGTCCGAATATCATGGTGAAAAATGGCCCGCTTTTCTGAGATTCATCGACTGTGG
CCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGAT
ATTGCTGAAGAGCTTGGCGGCAATGGGCTGACCCTCCTGTCTTGTTTACGGG
TATCGGCCGCTCCCCGATTGCACGCATCGCCTTCCTCTCATCGCCTTCTCAGG
TATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGAGT
TCTTCTGAGCGGGACTCTGGGTTGCCAAATGACCGACCAAGCGACGCCCAAC
CGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGA
TAACGCAGGAAAACAGATGTAGACAAAGCCAGCAAAGGCCAGGAACCG
AAAAAGGCCGGCTTGTGGCGGCTTTTTCATAGGCTCCGCCCCCCTGACGAGC
ATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTAT
AAAAGATACCAGGGCTTTCCCCTGGAAGCTCCCTCCTGTGCCTCTCCTGTTCG
ACCCTGCCGGTTACCAGGATACCTGTCCGCTTTTCCTCCTTCCGGGAAGCGTG
GCTTTTCTCATAGCTCACTGCTGTAGGTATCTCAGTTCGCGGTAGGTGCTTCGCT
CCAAGCTGGGGCTGTGTACGGAACCCCCCGTTCAGCCCGACCGCTGCGCCTT
ATCCGGTAACATATCGTCTTTGAGTCCAACCCCGGTAAGCAGCACAGCTATCG
CTGGCAGCAGCCACTGGAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTG
CTACAGAGTTTCTTGAAGTGGTGCTGCTACACTCAGGCTACACTAGAAGGACAGT
ATTTGATATCTGCAGCTGTGCGGCAAGCTCCCTCAGGAAAAAGAGTTGGTA
GCTTTGATCCGGCAGCAAAAAACCACCAGGCTGGTAGCCTGTTGCTTTTTTTGGCT
AAGCAGCACAGATTACGCGCAGAAAAAGGATTCTCAAGAAGATCCTTTTGATCT
TTTCTACGGGGTCTGACGCTCAGTGGAAAACATTCACTACGTTAAGGGATT TT
GGTCATGAGATTATCAAAGGAGTCTCCTACCTAGATCCTTTTATAGTCCGGAA
ATACAGGAACGCACGCTGGATGGCCCTTCGCTGGGATGGTGAAACATGAAA
AATGGCAGCTTCAGTGGATTTATGTTGGGCTGTAATGTTGGCCTGACCTTCTGTT
GCATAGGTATTCATACGGTTAAAAATTTATCAGGCGGATTGGCGCAGTTTTTC
GGGTGGTTTGTGCCATTTTTACCTGTCTGCTGCCGTGATCGCGCTGAACGCGTTTTAGCGGTGCGTACAATTAAGGGATTATGGTAAATCCACTTAC
CTCGTAGCCCATCGAGATAAAACCGCAGTACTCCGGCCACGATGCATCGTGCCGGCGTAGAGGATCGAGATCT

>pFN22K::Exo(1-189)

CCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACG
CAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTGGTTAGTGAACCGTCAGATCACTAGAAGCTTAATACGACTCACTATAGGGTACGA
TTTAGGTGACACTATAGAATAAGGAGAATTCATCATGGCAGAAATCGGTACTGGCTTTCCATTCGACCCCCATTATGTGGAAGTCCTGGCGAGCGCATGCACTA

>pFN22K::Exo(1-189)

GGCTTTCCATTGGACCCCCATTATGTGGAAGTCCCCTGGGCAGCGCATGCACTA
TGCGTACCGCTCTCCCCGATGGTTATGCAGTGACGGCAACGGCCTTTGAA
CTGAAATGCCCCGTATTACCTCCCGGGATTTTCATGAAGTTCGCGCTCGGAGTTT
CGAGGCCATAAAGTCAGCTTACATGGCCACGGTGACGATACAGCATGTTGGGTG
ACGCGAAAAATGCGCTGGTACTTGGGCAAACTATGACCCCGCTAGTAGAAGCCTG
AAGGCCCTGCATTATGTCGTAATTGGTTAAAAACGAATTCGCGCTCGGTACCCCGG
GATCCTCTAGAGTCGACCTGCAGGATAGCTGATCCGGCTCGTAAACAAA
GCCCGGAAAGGAAGGTGATTGAGCTGCTGGACCCCGCTAGCAGCTACAAAAGAT
TGGACAAAAACACAACTAGAAGGAGAGTTGGGGAATGATGAGGAGGTGGGT
GGTTTTTTTAAGCAAGTAAAAACCTCTACCAAGGTAAGCTGAAAATCGAATTCTAAT
GGATCCTCTTTTGCCTTGCTTTCTTTTCTTTGTCCAGTACAGGCCATACGTGACA
TTGATCCGGGGTGTCAGCACCGTTTTCTGCAGGACTGGCTTTCTACGGTGTTCCCGT
CCTTACAGCAGCCTTGCAGGCCTGAGTGCTTGCGGCAGCGTGACAAAGCTAAG
AAATTGCCAGCTGGGGCGCCCTGCTCTGGAAGTTGGGGAAGCGTAAGGTTGAATG
CTGGAGTGCTTTCTTTGCCCGCAAGGATCTGATGGGCACGGGAGTCAAGATCT
GACAGAGAGACAGGGACGATGACGATGCTTTGCGCATGCTTGAAACAGATGGATTG
ACGCAGGGTTCTCCGCGGCTTTGGGTTGGAGAGCTATTCCGCTATGACTGCGGC
ACAACAGACAAACTGGCTGCTCTGATGGCAGGCGCTGTTCCGCTGACGCGGAG
GGGCGGCAAGTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACT
GCAGGACGCACGGCACGCAGCTATCGTGCTGGGCGCCAACAGGGCGGGCTTTGC
GCAGCTGTGACTGACGTTGACTGGAAGCAGGGACTGGCTGCTATTGG
GCGAAGTGCCCCGGGAGAGATCTCTGTATCTCACTCCTTGCTCGTCCGCAGAAGAA
AGTATCCATCATGGCTGATGCAATGCGCAGGCTGATACGCTTGGATATCAGCT
ACCTGCCCATTTGACAGCCAAACTGTTGCGCCAGGCTCAAGGCGGTATCCGGATAGGT
GAGGATCTCTGTCTGACTCATGGCGTGCTGGCTTGGCTTGGCAATATCATGGTG AGG
AAAATGGCGCGTTTCCTGGAATCCATCGACTGTGGCAGGCTGGCTGGCAGGCA
CCGCTATACAGGACATAGCGGTTGGGTACCCGGTATATTCGGTAAGAGGCTTGGC
GGCGAAATGGGCTGACCGCTTCTCTCTGGCTTTACCGTGATACCGCCGCTCCCGAT
GCAGCGCATCGCTCTCTACGCTTGACAGTGATTTCTGTGACGGGACTCT
GGGGTTCCGAAATGAGCCAGAAAAGCGACGCCAAACGGTATACAGCTCACACTCAA
AGGCGGTAATACGTTATACACAGAATCGAGGATAACGCAGGAAGGAACCA
TGTGAGCAAAAGGCCAGAAAGGCCAGGAAACCGTAAAAAGGCCGCGTTGCC
TGGCGTTTTCCTATAGGGCTCCGCCTTCTCCCTTGACGAGCATCACAATAATCA
GCTCAAGTGAGGGTGCGAAGAACCCGAACAGGACTATAAAAGGATACAGGCCT
CCCCCTGGGAAAGCTCCCTGCGCTCTCCTCTGGTCCGACCCCTGCGCCTTACCG
ATACCTGTCGCCCTTTCTCCCTCCGGAAGCGGTTGGCGCTTCTCTCATAGCTCAC
GCTGTAGGTATCTCAGTTCGGGTAGGTGGTTCGTCCCTCAAGCTGGGCTGTGTG
CACGAACCCCCTGTTCCGAGCCAGAAGCGCTGCGGCTTATCCGGTAACTATCGTCT
TGAGTCCAACCAGGTAAGACAAGAAGTATCCGCACTTGCAGCAGCAGCCACTTGG
AACAGGATTAGCAGACGAGGTATGTAAGCGGCTGCTACAGAGGCTTGAAGT
GGTGGCCTAACTACCGGCCTACACTGAGAAAGGACAGTATTTGCTATCTCGGCTCT
GCTGAAGCCAGTTACCTTCGGAAGAAGAGTTGGGTAGCTCTTGGATCCGGGCAA
CAAAACCACCGCTGGTAGCGGTGGTTTTTTTGTGCAAGCACAGATTACGC
CAGAAGAAAAAGGATTGCTTACGAAGGCACCCCTTTCTCTCCTACGCGGTT
GCTCAGTGGAACGAAACTCAGTTAAGGGATTTGGTCATGATTATCAA
AAAGGATTCTTACCTAGATCCTTTTTATATGTCGGAATACAGGAAAGCACGC
TGGATGGCCCTTTGGCTGGGTGGTAAACCACATTGGAATAGGCAGCTTCAGTG
GATTAAGTGGGGGTAATGTCGCCCTGTACCCTCTGGTTGCATAGCTATTCA
GGTTAAAAATTTATCGGCGCGATGCGGCGAGTTTTTCGGGTTGGTTGGGCA
TTTTTACCTGTCTGCAGCGATCGCGCTGAACCGCGTAAAAACGCGGAGCTAC
AATTAAGGGATTATGGATAACCTACTTACTGCTCCCTCGTAGCATCGAGA
TAAACCCGCACTCCGGCCACGATCGCCGCGTACGAGGATCGAGATCT
APPENDIX B. FLUORESCENT PROTEIN SEQUENCES

FP gene sequences

>gi|104717157 gfpmut3* (SCCI strain gift from Dr. Sze)

ATGCGTAAAGGAGAAGAACAAAAACTTTTCAGGAGTTGTCACCAATTCTGGTTGAATT
AGATGGTGTATATGTCACCAAAAATCTCTGTCAGGAGAGGTGAAGGT
GATGCAACATAAGAAAACCTTTACCTTTAATTATTTGCACTACTGGAAAAACT
ACCTGTCCATGGCCAACACTTGTCACTACTTTTCGTTATGGGTCTAATGCTT
TGCGAGATACCCAGATCATATGACAGCAGACTTTTTTCAAGAGAGAGCCATG
CCCAAGAGTTATTGTACAGGAAAGAACTATTTTTTCAAGAGATGACGGAACACT
ACAAGACACGCTGCAAGATGTTGATAACCTTGTTAATAGAAT
CGAGTTAAAAGGTATTTAAAAAGAGATGGAACATTCTTGGACACAAA
TTGGAATACAAACTATAACTCACAATAATGTATACATCATGGCAGACAAAAAAA
AGAATGGAATCAAAAGTTAATTTAGACACACACTTGAAGATGGAAG
CGTTCAACTAGCAGACCATTATCAACAAAAATACTCAAATTGCGATGCCCCT
GTCTTTTTACCAGACACAACTACCTGTCCACACAAATCTGCCGTTTTGAAAGA
TCCAACGAAAAAGAGAGACACATGCTTTCTTGGAGTTTTGAAACAGCTGCT
GGGATTACACATGGCATGGATAAATATAA

>Green (RIK424: gfpmut3* Y66, Green fluorescence)

ATGCGTAAAGGAGAAGAACAAAAACTTTTCACTGGAGTTGTCACCAATTCTGGTTGAATT
AGATGGTGTATATGTCACCAAAAATCTCTGTCAGGAGAGGTGAAGGT
GATGCAACATAAGAAAACCTTTACCTTTAATTATTTGCACTACTGGAAAAACT
ACCTGTCCATGGCCAACACTTGTCACTACTTTTCGTTACGGAATGCTTGAATGCT
TTGCGAGATACCCAGATCATATGAAAACAGCATGACTTTTTTCAAGAGTGCCAT
GCCCGAAGGGTTATGTCAGGGAAGAAGACTATATTTTTTTTCAAAGATGACGGGAAC
TACAAGACACGTGCTGAAGTCAAGTTTTGAAAGGTGATACCCCTTTGTTAATAGAA
TCGAGTTAAAAGGTATTTGATTTTAAAAGAGATGGGAAACATTCTTTGGACACAA
ATTGGAATACAACTATAACTCACAACATGTTACATCATGCGACACAAAACAA
AAGAATGGAATCAAAGGTAACTTCAAAATTAGACACAAACATTGGAAGATGGAA
GCGTTCAACTAGCAGACCATTATAACAAATACACTCCAAATGTGCGATGGCCC
TGTCCTTTTACCAGACAAACCATTACTTGCCACACAATCTGCCCTTTTCGAAAG
ATCCCAACGAAAAGAGAGACCATGGTCTCCTCTTGGAGTTTGTAACAGCTGC
TGGGATTACACATGGCAGCATGGGATGAAACTATACAAATAA

>Amarillo (RIK446: gfpmut3* Y203 (TTAC), Yellow fluorescence)
ATGCGTAAAGGAGAAGAACCTTTTTCTGGAATTTGTCCCCAATTCTTTGTTGATT
AGATGAGTATGTTAAATGGGCAAAAAATTTCCTCTGCAGTGGAAGGGGTGAAGGT
GATGCAACATACGGAAAACCTTACCTAAAAATTTATTTTGCACACTCTGGAAACACT
ACCTGTCTTGGCAGACACTTCGGTCACTACTTTTGGTTATGGTGTCAATGCTT
TGCGAGATACCCAGATCATAGAAAACGGCATGACTTTTTTCAAGAGTGCCATG
CCCGAAGGTTATGTACGAGGAAAAGAAGACTATATTTTTTTCAAAGATGACGGAACACT
ACAAGACACGCTGGAATTTGAAAGGTGATACCTTTGTGTTAATAGAAT
CGATTTAAAAGGTATTTTTAAAGAAGATGGGAAACATTCTTTGGACACAAA
TTGGAATACAACTATAACTCACAACATGTTACATCATGCGAGACAAACAAA
AGAATGGAATCAAAGGTTAACTTTCAAAATTAGACACAAACATTGGAAGATGGGAAG
CGTTCAACTAGCAGACCATTATCAACAAAAATACCTCAATTGGCGATGGCCCT
GTTCCTTTTACCAGACAAACATTACCTGCTTTACCAATCTGGCCCTTTTCGAAAGA
TCCCAACGAAAGAGAGACCACATGGCTCTTTCTTGAGTTTGTAAACAGCTGCT
GGGATTACACATGGCATGGATGAACTATACAAATAA

>Aqua (RIK411: gfpmut3* L_{64}T_{65} W_{66} (CtGAGCtGG), Cyan fluorescence)
ATGCGTAAAGGAGAAGAAGACTTTTCTCAGTGGAGTTGTCCCAATTCTTTTGTAATT
AGATGGGTGATGTTAATGGGCACAATAATTTTCTTCAGTGGAGAGGGTGAAGGT
GATGCAACATACGGAACACTTACCTTTAAATTTATATGGCACTACTGGAAAACT
ACCTGTTCCATGCGCCAACACTTGCACTACTCTGACCTGGGTTGCATTAATGCT
TTGCGGAGATACCCAGATCATATGAAACAGCATGACTTTTTTCAGAAGTGCCAT
GCCCGAAGGTTATGTACAGGAAAGAAGACTTATATTTTTAAAGATGAGCGGGAAC
TACAAGACAGTGGCTGAGGTCAAGTTTAAAGGTGACACTCTTCTTGACACAA
TGGGATACAAACTATAACTCACAACATGTACATCATGGCAGGAAACAA
AAGAATGGAATCAAAGTTAACTTCAAAATTAGACACAACATTGAAGATGGAA
GCGTTCAACTAGCAGACCATTATCAACAAATAACTCCAATGGCGATGCCC
TGTCCTTTTACAGACAACACCATTACCTGTCACACAATCTGCCCTTTGGAAG
ATCAGCAACGAAAAGAGAGACCACATGGCTCTTTCTTGAGTTTGTAAACAGCTGCT
GGGATTACACATGGCATGGATGAACTATACAAATAA

>Aqua (RIK414: gfpmut3* L_{64}T_{65} W_{66} (GACGtGG), Cyan fluorescence)
ATGCGTAAAGGAGAAGAAGACTTTTCTCAGTGGAGTTGTCCCAATTCTTTTGTAATT
AGATGGGTGATGTTAATGGGCACAATAATTTTCTTCAGTGGAGAGGGTGAAGGT
GATGCAACATACGGAACACTTACCTTTAAATTTATATGGCACTACTGGAAAACT
ACCTGTTCCATGCGCCAACACTTGCACTACTCTGACCTGGGTTGCATTAATGCT
TTGCGAGATACCCAGATCATATGAAACAGCATGACTTTTTCAAGAGTGCCAT
GCCCGAAGGTATGTACAGGAAGAAGAAGCTATATTTTTCAAAAGATGACGGAAC
TACAAGACAGTGCTGAAGTCAAGTGGTGAATCCCTTGTTAATAGAA
TCGAGTTAAAAGGTATTGATTTTAAAGAAGATGGAAACATTCTTTGGAACCAA
ATTGGAATACAATATAACTTACCAACTCCAAATGTACATCATGCGACACAAAC
AAGAATGGGAATCAGGAGTTAACTTCCAAAATTAGACACAAACATTTGAAGATGGAA
TCGAGTTAAAAGATATTAGATTTTAAAGAAGATGGAAACATTCTTTGGAACCAA
ATTGGAATACAACTATAACTCACACAATGTATACATCATGCGACAAACAA
AAGAATGGGAATCAAAGTTAACTTCAAAATTAGACACAACATTGAAGATGGAA
GCGTTCAACTAGCAGACCATTATCAAACAAAAATCTCCAAATTGCGATGGCCC
TGTCCTTTTACCAGACACACCATTACTCTCGACAAATCTGCCCTTTGGAACAG
ATCCCCAGAAAAGAGAGACCACATGGTCTCTTCTTGGAGTTTTGTAACAGCTGC
TGAGATTACACATGCGCATGGGAACATAATACCAAAAT
>Green (RIK425: \textit{gfpmut3*} \textit{L}_{64}^{T_{65}}{Y}_{66} (CtGACCtaC), Green fluorescence)
ATGCCTAAGGAGAAGAAGACTTTCTCAGTGGTTGGTCCCAATTTCTTTGGAATT
AGATGGGTGATGTAAATGCGACAAACACTTTCTGCATTGAGGGGTGAAGGT
GATGCAACATACGGAAACTTACATTTAATTTTATTTGGCAGCTACTGGGAACAACT
ACCTGTTCCATGGCCAAACTTGTCACTACTTCACTCTGACCTACGGTGTCTCAATGCT
TTGCGAGATACCCAGATCATATGAAACAGCATGACTTTTTCAAGAGTGCCAT
GCCCGAAGGTATGTACAGGAAGAAGAAGCTATATTTTTCAAAAGATGACGGAAC
TACAAGACAGTGCTGAAGTCAAGTGGTGAATCCCTTGTTAATAGAA
TCGAGTTAAAAGGTATTGATTTTAAAGAAGATGGAAACATTCTTTGGAACCAA
ATTGGAATACAATATAACTTACCAACTCCAAATGTACATCATGCGACACAAAC
AAGAATGGGAATCAGGAGTTAACTTCCAAAATTAGACACAAACATTTGAAGATGGAA
TCGAGTTCAACTAGCAGACCATTATCAAACAAAAATCTCCAAATTGCGATGGCCC
TGTCCTTTTACCAGACACACCATTACTCTCGACAAATCTGCCCTTTGGAACAG
ATCCCAACGAAAAGAGAGACCACTATGGTCCTTCTTTGAGTTGTAACAGCTGC
TGGGATTACACATGGCATAATGGAACACTTATACAAATAA

>Violeta (RIK419: gfpmut3* W66 (CC), Blue fluorescence)

>Mar (RIK413: gfpmut3* T65W66 (CCaCGT), Cyan fluorescence)
TTGCGAGATACCCAGATCATATGAAACAGCATGACTTTTTTCAAGAGGTGCCAT
GCCCGAAGGTATGTACAGGAAAGGAAACTATATATTTTCAAGAGATGACGGAAC
TACAAGACACGTGCTGAAGTCAAGGTTGGAAGGTGATACCCTTTGTTAATAGAA
TCGAGTTAAAAAGGTATTGATTTTTAAAGAGATGGAACACATTCTTTGGACACAA
ATTGGAATACAACACTATAACTCACAACAAATGTATACATCATGCGACAAAACAA
AAGAATGGAATCAAAGTTAACTTCAAAAATATGAGAGACATTTGAAGATGGA
GCGTTCAACTAGCAGACCATTATCAACAAATACTCCAAATTGGGCAGTGGCCC
TGTCCTTTTACCAGACACAAACATTACCTGCAACACAATCTGGCCTTTTCGAAAG
ATCCCAACGAAAAGAGAGACACATGGCTGCCTTTCTTGAAGTTTGGTAACAGCTGC
TGGGATTACACATGGCATGGATGAACTATACAAATTTTATGCACTACTGGAAAAC
ATGCGTAAAGGAGAAGAACCTTTTTTCACTGGAGTTTGTCCCAATTCTTTGGTAATT
AGATGCGTATGTTAATGGCAGAAATTTTTCTGTCAAGTGGAGAGGTGAAGGT
GAGTGCAACATACGGAAAACACCCTTTAAAATTATTTGCACCACACTGGAAAAC
ACCTGTTTCATGCCGTTGGACTTTTGCACACTTTTTCTGGTTATGGTTCAATGCTT
TGCGAGATACCCAGATCATATGAAACGATCTGGAAACAACTTTTTTCAAGAGTGCCAT
CCCGAAGGTTATGTACAGGAAAGGAAACTATATATTTTCAAGAGATGACGGAAC
ACAAGACACGTGCTGAAGTCAAGGTTGGAAGGTGATACCCTTTGTTAATAGAAAT
CGAGTTAAAAAGGTATTGATTTTTAAAGGGAAGATGGAACACATTCTTTGGACACAA
TTGGAATACAACACTATAACTCACAACAAATGTATACATCATGCGACAAAACAA
AAGAATGGAATCAAAGTTAACTTCAAAAATATGAGAGACATTTGAAGATGGA
CGTTCAACTAGCAGACCATTATCAACAAATACTCCAAATTGGGCAGTGGCCC
TGTCCTTTTACCAGACACCAATTTACCTGTCACACAATCTGGCCCTTTTCGAAAGA
TCCCAACGAAAAGAGAGAGCACCACATGGTCCTTCTTGAGTTTGTAAACAGCTGCT
GGGATTACACATGGCAGTGAACACTATACAAATAA

>Dark (RIK423: gfpmut3* Stop39 (G), Dark)
ATGCGTAAGAGAGAGAAGACTTTCTCAGTGAGTTGTCCTAATTCTTTGTAGATTT
AGATGGTGATGTTAATGGGCACAAAATTTCTGTCAGTGAGGAGGATGGTGAAGGT
GATGCAACATAGGGGAACCTTTACCTTTAATAATTTATTTGCACTACTGGAAAAC
TACCTGTTCCATGGCCAAACTTGGTCACTACTTTTTCGGTTATGTTGTTCAATGCT
TTGCGAGATACCCAGATCATATGAAACGAGCTTTCTCAAGAGTGCCAT
GCCCGAAGGGTTATGTACAGGAAAGAACT ATATATTTTTCAAAAGATGACGGGAAC
TACAAAGACACGTGCTGAAGTTGAGCTACCCTTTGTAAATAGAA
TCGAGTTAAAAGGTATTGATTTTTAAATTTATTTGAAAGATGGAAAACATTTCTTGGACACAA
ATTGGAATACAACTATAACTTCACAACATGTATACATCATGGCAGACAAACA
AAGAATGGAATCAAAGTAAAATTTAAGACACAACTGAAATGGGAAAC
GCGTTCACCTAGCAGACCATTATCAAACCAATACCTCCAATGGGCAGATGGCC
TGTCCCTTTTACCAACAGAAAACCATTACCTGTCACACAATCTGCCCTTTGAAAG
ATCCCAACGAAAAGAGAGAACCACATGGTCCTTCTTGAGTTTGTAAACAGCTG
TGGGATTACACATGGCAGTGAACACTATACAAATAA

>pDual-egfp
ATGGCTAGCATGACTGTTGGACAGCAAATGGGTCGCGGGGATCCACCGGTTCG
CCACCATGTTGAGCAAGGGCGAGAGTTGTTACCCGGGGGTGGTGCCCATCCT
GGTCGAGCTGGACCGCGACGTAAAACGGCCAAAGTTCAAGCAGTGCCGGCGA
GGGCGAGGGCGATGCCCACCTACGGCAAGCTGACCCTGGAATTGTTCTCGACAC
ACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGC
GCGTGCAGTGGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTT
CAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAG
GACGACGGCAACTACAAGACCCGCGCCGAGGAGGCTACCGACTTCAAGGAGGACCGACAAC
CTGGTGAACCCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACCGACAAC
ATCCTGGGGCACAAGGCTGGAGTCAACTACAACAGCCACAACCTGCTATATCA
TGGCCGACAAGCGAGAACCGCATCGAGACTTCAAGGAGGACCGACAAC
ACATCGAGGACGGGCAGCAGGCAGCTGCAGCTCGCCGACCACTACCAGCAGAACACCC
CATCGGGCAGGCCCCTGGCTGGCCTGCCGACAACCACTACCAGCAGAACACCC
TCCGCCCTGGGACAAAGCCCAACCCAGACGATCACATGGTCCTGCTGG
AGTTCGTGACCGGACGGCGATCGACTCCTCAGCAGCAGCAGCGAGCTCTGGCA
GGGACGGCGTGGCCGTGGACGCTGGAGGACCGCAGGAGCTGTACAGTCG
CGGACTCAGATCCGACCTGAGACCTTACAGAATACATGGGAGCAATGCAAGT
AGCAATAAAGACGAGCTAACAATG

>Celeste (pDual-egfp W67)

ATGGCTAGCATGACTGGTGAGCACAGAAATGGGTTCGGGGGATCCACCGGTCG
CCACCATGGTGAGCAAGGCGAGGAGCTGTGTCACCAGGGGGTGTTGCTCCCATCCT
GGTCAGCTGGAGCGCGCACTAAGGGCCAAAGGCTAGCTCGTCCGGGA
GGGCGAGGCGATGCGACCTACGGCGAAGCTGGGCCACCTGCGAGATGGTCCGACC
ACCAGGAAGGGTGCCTGGCCTGCCCACCCCTTCGTGAGCACCACCTGACCTGG
GCGTGCAGGTCTTCCGCAGCCGCTACCCGCACCACATGAAGCAGCAGACTCTTT
CAAGTCCGCCATCGCCGAAGGGCTACGTGCCAGGACCGACCACATGCTTCTTCTAAG
GACGACGGCAACTACAAGACCCGCGCCGAGGTGAAAGTTCGAGGGCGACACC
CTGGTGAACCCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACCGACAAC
ATCCTGGGGCACAAGCTGGAGTACAACGCTACAACAGCCACAACGTCTATATCA
TGCGCCGACAAGCAGAAGAACGGCATCAAGGTGAACCTCAAAGATCCGCCACA
ACATCGAGGACGGCAGCGTGAGCTCGCCGCCACACACTACACGAGAACAACCC
CATCGGCCGACGCCCGCCCGGATCACACTCTCGCCGATGGACAGCGTCGTAACGTC
CGACTCAGATTCGACCTCGAGACCTAGAAAAACATGGAGCAATCACAAGT
AGCAATACAGCAGCTAAACATG

>Dark (pDual-egfp Stop67)
ATGGCTAGCAGCATGACTGTTGGACAGCAATGGGTCGCGGGGATCCACCGGTCG
CCACCATGGGAGCGAAAGGCGAGGAGCTGTTTACCCGGGCTGGTGCCCATCCT
GGTCTGAGCTGGGACGGGCAGCGTGACATCGAAGCTGACCCTGAAAGATTCATCTGACC
TTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGGATCACATGGTCCTGCTGG

ATCCTGGGGCACAAGCTGGAGTACAACGCTACAACAGCCACAACGTCTATATCA
TGCGCCGACAAGCAGAAGAACGGCATCAAGGTGAACCTCAAAGATCCGCCACA
ACATCGAGGACGGCAGCGTGAGCTCGCCGCCACACACTACACGAGAACAACCC
CATCGGCCGACGCCCGCCCGGATCACACTCTCGCCGATGGACAGCGTCGTAACGTC
CGACTCAGATTCGACCTCGAGACCTAGAAAAACATGGAGCAATCACAAGT
AGCAATACAGCAGCTAAACATG

>Dark (pDual-egfp Stop67)
ATGGCTAGCAGCATGACTGTTGGACAGCAATGGGTCGCGGGGATCCACCGGTCG
CCACCATGGGAGCGAAAGGCGAGGAGCTGTTTACCCGGGCTGGTGCCCATCCT
GGTCTGAGCTGGGACGGGCAGCGTGACATCGAAGCTGACCCTGAAAGATTCATCTGACC
TTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGGATCACATGGTCCTGCTGG

ATCCTGGGGCACAAGCTGGAGTACAACGCTACAACAGCCACAACGTCTATATCA
TGCGCCGACAAGCAGAAGAACGGCATCAAGGTGAACCTCAAAGATCCGCCACA
ACATCGAGGACGGCAGCGTGAGCTCGCCGCCACACACTACACGAGAACAACCC
CATCGGCCGACGCCCGCCCGGATCACACTCTCGCCGATGGACAGCGTCGTAACGTC
CGACTCAGATTCGACCTCGAGACCTAGAAAAACATGGAGCAATCACAAGT
AGCAATACAGCAGCTAAACATG

>Dark (pDual-egfp Stop67)
ATGGCTAGCAGCATGACTGTTGGACAGCAATGGGTCGCGGGGATCCACCGGTCG
CCACCATGGGAGCGAAAGGCGAGGAGCTGTTTACCCGGGCTGGTGCCCATCCT
GGTCTGAGCTGGGACGGGCAGCGTGACATCGAAGCTGACCCTGAAAGATTCATCTGACC
TTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGGATCACATGGTCCTGCTGG
AGTTCGTGACCGCCGCGGGATCAGCTGTCACTCTCGGATCACTCTCGGATGCTGTACAAGTC
CGGACTCAGATCCGACCTCGAGACCTAGAAACATGGAGCAATCACAAGTC
AGCAATAACAGCAGCTAACAATG

> Azure (pDual-egfp H67)
ATGGCTAGCATGACTGTTGAGCAGCAATGGGCAGGCTCGGATCCACCGGCTCG
CCACCAGTGAGCAAGGCAAGGGAGGTCGTAACGGCCGACGGAGCTGTACAAGTC
CGGACTCAGATCCGACCTCGAGACCTAGAAAACATGGAGCAATCACAAGTC
AGCAATAACAGCAGCTAACAATG
>Mostaza (pDual-egfp Y203)

ATGGCTAGCATGACTGTGGACAGCAAGATGGTCGCGGGGATCCACCGGTCG
CCACCATGGTAGCAAGGCGAGAGCTGGTCACCACGGGGTTGCTGCCCATCCT
GCTCGAGCTGGAGCAGCAGAGGCTACGCAAGGCTGACCTGGATCTGGTCCAC
GGGCGAGGCGGATGCAACCCTACCGGAAGCTGACCCCTCTGGTACACCAGCTC
ACCAGGCAAGCTGACCCGCTCTGGGACAGCAAGGGCGAGGGAGCTGTGTCA
GGTACGACGGGTGCTCTGAGCGCTACCCACCACATGTACAGCAAGCTGACCC
GGTACGACGCTGACCCGCTCTGGGACAGCAAGGGCGAGGGAGCTGTGTCA
GGTACGACGGGTGCTCTGAGCGCTACCCACCACATGTACAGCAAGCTGACCC

**Fluorescent protein sequences**

> GFPmut3*

MRKGEELFTGVPVLVLDGDNVGHKFSVSSEGEGDATYGLTLKFICTTGKLP
VPWPTLVTFTGYGQCFARYPDHMVKQHDFFKSAMPEGYVQERTFFKDDGNYK
TRAEVKFEGLTDLVNRIELKGIHKEDGNIILGHKLEYYNISHVYIMADKQKNGI
KVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSAWLKDPNEKRD
HMVLLEFVTAAGITHGMDELYK

>Amarillo fluorescent protein (GFPmut3* Y203 variant)
MRKGEELFTGVVPIVLDGDVGKHFSVSGEGEGDATYGKLTLKFICTTGKLP
VPWPTLVTTFGVQCFARYPDHMKQHDFKKSAMPEGYVQERTIFFKDDGNYK
TRAEVKFEVGDTLVPNRIELKGIDFKEDGNILGHKLEYNYNSHNYIMADKQKNGI
KVNFKIRHNIEDGGSVLADHYQQNTPIGDGPVLLPDNHYLSTQSAWLKDPNEKRD
HMVLLEFVTAAGITHGMDELYK

>Aqua (GFPmut3* L64/T65/W66 variant)
MRKGEELFTGVVPIVLDGDVGKHFSVSGEGEGDATYGKLTLKFICTTGKLP
VPWPTLVTTFGWVQCFARYPDHMKQHDFKKSAMPEGYVQERTIFFKDDGNYK
TRAEVKFEVGDTLVPNRIELKGIDFKEDGNILGHKLEYNYNSHNYIMADKQKNGI
KVNFKIRHNIEDGGSVLADHYQQNTPIGDGPVLLPDNHYLSTQSAWLKDPNEKRD
HMVLLEFVTAAGITHGMDELYK

>Violeta (GFPmut3* W66 variant)
MRKGEELFTGVVPIVLDGDVGKHFSVSGEGEGDATYGKLTLKFICTTGKLP
VPWPTLVTTFGWVQCFARYPDHMKQHDFKKSAMPEGYVQERTIFFKDDGNYK
TRAEVKFEVGDTLVPNRIELKGIDFKEDGNILGHKLEYNYNSHNYIMADKQKNGI
KVNFKIRHNIEDGGSVLADHYQQNTPIGDGPVLLPDNHYLSTQSAWLKDPNEKRD
HMVLLEFVTAAGITHGMDELYK
>Mar (GFPmut3* T65/W66 variant)
MRKGEELFTGVVPILVELDGDVNGHKFSVSGEGEDATYGKLTLLKFICTTGKLP
VPWPTLVTTFTWGVQCFARYPDHMKQHDFFSAMPEGYVQERTIFFKDDGNYK
TRAEVKFEGDTLVRNIELKGIDFKEDGNILGHKLEYNYNSHNHYIMADKQKNGI
KVNFKIRHNIEDGSVQLADHYQQNTPIDGPVLLPDNHYLSQALTSDPNEKRDK
HMVLLEFVTAAGITHGMDLELYK

>Bronze (GFPmut3* T65/W66/Y203 variant)
MRKGEELFTGVVPILVELDGDVNGHKFSVSGEGEDATYGKLTLLKFICTTGKLP
VPWPTLVTTFTWGVQCFARYPDHMKQHDFFSAMPEGYVQERTIFFKDDGNYK
TRAEVKFEGDTLVRNIELKGIDFKEDGNILGHKLEYNYNSHNHYIMADKQKNGI
KVNFKIRHNIEDGSVQLADHYQQNTPIDGPVLLPDNHYLSyQSALSHPNEKRD
HMVLLEFVTAAGITHGMDLELYK

>Mostaza (eGFP Y203 variant)
MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEDATYGKLTLLKFICTTGKLP
VPWPTLVTTFTWGVQCFARYPDHMKQHDFFSAMPEGYVQERTIFFKDDGNYK
TRAEVKFEGDTLVRNIELKGIDFKEDGNILGHKLEYNYNSHNHYIMADKQKNGI
KVNFKIRHNIEDGSVQLADHYQQNTPIDGPVLLPDNHYLSyQSALSHPNEKRD
HMVLLEFVTAAGITLGMDLELYKSGRLSRQAASNSAVDGTAGPGSTGSR

>Celeste (eGFP W66 variant)
MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEDATYGKLTLLKFICTTGKLP
VPWPTLVTTFTWGVQCFARYPDHMKQHDFFSAMPEGYVQERTIFFKDDGNYK
TRAEVKFEGDTLVRNIELKGIDFKEDGNILGHKLEYNYNSHNHYIMADKQKNGI
KVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPEKRD
HMVLLEFVTAAIGITLGMDELYKSGLRSRAQASNSAVDGTPGSTGSR

> Azure (eGFP H66 variant)

MVSKGEELFTGVVPILVVELGDVNGHKFSVSGEGEDATYGKLTLKFICTTGKLP
VPWPTLVTTLTHGVQCFSRYPDHMKQHDFSAMPEGYVQERTIFKDDGNYK
TRAEVKFEGDTLVNRIELKGIKIDFKEGMLGHEKLEYNYNSHNVYIMADKQKNGI
KVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPEKRD
HMVLLEFVTAAIGITLGMDELYKSGLRSRAQASNSAVDGTPGSTGSR
APPENDIX C. VISUAL ABSTRACT

Recombineering reporter in bacteria

YELLOW, GREEN AND CYAN *E. coli*

Fluorescent proteins can be used as Recombineering Reporter

*Aqua, Amarillo* were created by Recombineering from GFPmut3* in *E. coli*. *Aqua, Amarillo* and GFPmut3* *E. coli* were grown independently to OD₆₀₀ of 0.4. Cells were collected from each culture and resuspended in minimal medium (OMBG) to a 10% of the initial volume. Cells were mixed and 1 µl of the mixed concentrated cells were added to a slide with mounting medium (P7481 Prolong Antifade, from Molecular Probes). The slide was covered and left to sit overnight before imaging. Tile scanning was taken with a Zeiss LSM710 confocal microscope. Maximum projection was performed with PerkinElmer/Improvision Volocity 64-bit software. Color merges were done in Adobe Photoshop CS3.

Picture taken by: Zeiss LSM710 demo technician.
What is Recombineering?

**Recombineering uses viral recombinases to engineer DNA inside cells**

A. Chromosome map of *E. coli*, indicating locations of the fluorescent reporter gene *gfpmut3*\(^*\), the viral source of Recombineering enzymes and of the origin of replication.

B. Recombineering enzymes are induced with heat just before mixing the cells with the mutagenic DNA.

C. Viral recombinases take the mutagenic DNA to replication forks and pair the mutagenic DNA to complementary ssDNA sequences. The pairing can tolerate changes if they are in the middle of the mutagenic DNA.

---

**Figure created by Dr. Rik Myers.**
Recombineering reporter in human stem cells

GREEN AND MOSTAZA BONE MARROW STEM CELLS

Stem cells expressing fluorescent protein genes to evaluate genome engineering in human cells

MIAMI cells were transduced with lentiviral particles of produced from pDual-Mostaza and pDual-eGFP. Cells were grown on coverslips, fixed and stained with DAPI. Images were taken in a Leica SP5 confocal microscope using a 63x oil-immersion 1.3 numerical aperture objective. Confocal acquisition parameters were determined at the beginning of the study and the same parameters (e.g., gains, slit aperture, laser intensity) were used for all the images. Confocal optical sections were 0.6 µm thick. Field selection was performed using the DAPI channel to identify MIAMI cells. Adobe Photoshop was used to merge the colors. In blue, DAPI staining of nuclei, in green, MIAMI cells transduced with pDual-eGFP and in Yellow, MIAMI cells transduced with pDual-Mostaza. The bar in white at the lower right corner corresponds to 50 µM.

Picture taken by Dr. Pedro Salas

What it takes to change fluorescent protein gene Mostaza to Green, or vice versa?

ssDNA is delivered to cells. The ssDNA is complementary to the target gene with the desired genetic change in the middle of its sequence. To change the Mostaza fluorescent protein gene inside the cells, the mutagenic ssDNA (oligo) differs from the Mostaza gene in the indicated green nucleotides.

<table>
<thead>
<tr>
<th>gacaaccactacctgTCcACcagtcgccctgag</th>
<th>Oligos T203 4s</th>
<th>Mostaza target</th>
</tr>
</thead>
<tbody>
<tr>
<td>ctgtttgtgatggaCTcGAgtcaggccgactc</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ctcaggccgactgACGTcGAcagtgtagtggttc</td>
<td>Oligos T203 4as</td>
<td>Mostaza target</td>
</tr>
<tr>
<td>gagtcgccctgaccATCGAgtccatccacaacag</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Recombineering reporter in human cancer cells

PROGERIN REPORTER IN HELA CELLS

Repairing Progeria by Recombineering

Cells were grown in flat bottom 96 well plates. Just before imaging, media was changed to HBSS. Images were taken using a Celigo adherent cell cytometer with live cells. Picture taken by Dr. TuAnh Dang.

What is the genetic condition in Progeria?

Recombineering could be used to repair a monogenic disease in a human cells like Progeria. Progeria leads to premature aging and death by the age of 13. The disease is caused by a de novo heterozygous synonymous substitution mutation (C1824T) of the lamin A/C (LMNA) gene. The mutation activates a cryptic donor splice site in exon 11 of LMNA that is used about 85% of the time to join exon 11 to exon 12 of the LMNA mRNA. A mutant lamin A protein (Progerin) with a 50 amino acid internal deletion is generated by this aberrant splicing. The deleted region includes a cleavage site for an endoprotease (Zmpste24) that normally removes the farnesylated C-terminus of lamin A. As a consequence, Progerin is maintained in a state of permanent carboxyfarnesylation. Progerin incorporates abnormally into the nuclear lamina, leading to a host of pleiotropic effects.
How efficient is ssDNA delivery to cells?

MUTAGENIC DNA (GREEN) DELIVERED TO STEM CELL NUCLEI

Lipofectamine 2000 delivers oligos to the nucleus of human cells for gene targeting experiments.

MIAMI cells grown on cover slips were transfected with 280 nM fluorescein-conjugated oligos using Lipofectamine 2000. Complexes and cells were incubated for five hours. Cells were then fixed and mounted with DAPI mounting medium. Fluorescein emits green fluorescence and DAPI stains the nucleus in blue. The image was obtained using a Zeiss confocal microscope with a 20 X objective. Nuclear localization of oligos was evaluated by examining colocalization of fluorescein with the blue fluorescence from DAPI (left) and by examining Z-sections of confocal planes (right). This image was processed using Adobe Photoshop. Picture taken by Dr. James Baker.
HERPES VIRUS RECOMBINASE ICP8 LOCALIZES TO SUBNUCLEAR COMPARTMENTS IN STEM CELLS

Viral recombinase is expressed in mammalian cells. Can ICP8 catalyze Recombineering?

Cells were transfected with pCMV-ICP8-GFP. 16 hours later cells were fixed with 4% PFA, stained with DAPI and mounted with Prolong Gold (Invitrogen). Images were taken with a Delta Vision OMXV4 fluorescent microscope and processed with their deconvolution software. On the top left is an image collected with the DAP filter (435 nm). On the top right is an image collected with the FITC filter (523 nm). On the bottom right is the merge of both images were the DAPI image was colored in blue and the FITC image was colored in green.

Picture taken by Marissa Hackett.
Neighbor cell from previous experiment. Are the green dots ICP8-GFP molecules or endogenous fluorescent molecules?

Marissa was trying really hard to find green fluorescence and increased the gate....

Cells were also incubated with red-fluorescent oligos, texas red labeled. In this picture the oligos are shown by the lamellipodia of MIAMI cells.
Marissa was trying really hard to find green fluorescence and increased the gate. ICP8-GFP fluorescence overexposed the CCD camera at gain levels necessary to see the green fluorescence of cells lacking ICP8-GFP.
HERPES VIRUS RECOMBINASE ICP8 LOCALIZES TO THE NUCLEUS OF HUMAN KIDNEY CELLS

Viral recombinase is expressed in easy to transfect and easy to grow 293T cells from inducible plasmids.

293T were co-transfected with pNEBR-R1 and pNEBR-X1-ICP8. 24 hours later inducer (RSL) was added to cells in the top panel. After 48 hours incubation, cells were processed for immunofluorescence. ICP8 protein was detected using an anti-ICP8 antibody (red) and nuclei were stained with DAPI. Cells were imaged using a Nikon fluorescent microscope.

Picture taken by Dr. Adriana Gomez.
HERPES VIRUS RECOMBINASE ICP8-GFP LOCALIZES TO THE NUCLEUS OF HUMAN KIDNEY CELLS

Viral recombinase is expressed in easy to transfect and easy to grow 293T cells from inducible plasmids.

293T were co-transfected with pNEBR-R1 and pNEBR-X1-ICP8-GFP. 24 hours later inducer (RSL) was added to cells in the top panel. After 48 hours incubation, cells were stained with DAPI and imaged using a Nikon fluorescent microscope.

Picture taken by Dr. Adriana Gomez.
Herpes virus recombinase ICP8 expression in HeLa cells
HERPES VIRUS RECOMBINASE ICP8-GFP LOCALIZES TO SUBNUCLEAR COMPARTMENTS IN HELA CERVICAL CANCER CELLS

Viral recombinase is expressed in progerin reporter cells. Can ICP8 promote the repair of the progeria mutation by Recombineering?

The cell sorter was used to isolate a derivative of the HeLa Progerin reporter following ICP8-mediated Recombineering that expresses DsRed but not GFP. This strain was subsequently transfected with pCMV-ICP8-GFP and live cells were visualized using the Zeiss LSM 780 laser scanning confocal microscope live in HBSS, 8 hours after transfection.
Herpes virus exonuclease UL12.5 expression in human HeLa cells

**HERPES VIRUS EXONUCLEASE UL12.5 (SYNEXO PARTNER OF ICP8) LOCALIZES TO MITOCHONDRIA IN HUMAN CELLS**

The viral SynExo exonuclease partner of ICP8 localizes to the nucleus and to the mitochondria in human cells. Can ICP8-UL12.5 catalyze Recombineering in mitochondria?

The UL12.5 gene from HHV1 strain 17 was cloned was fused to the HA epitope into RheoSitch mammalian inducible expression vectors from NEB and transiently transfected it into HeLa cells. Cells were grown on cover slips, induced for recombinase expression, and imaged by confocal microscopy using a fluorescent marker for mitochondria (mitotracker red) and a fluorescent antibody to the HA-tagged HHV1 recombinase (Anti-HA green).

Picture taken by Dr. Carlos Moraes.
Recombineering in 293T cells

Mostaza to Green Recombination in 293T

GENE TARGETING PRODUCES OLIGO DEPENDENT GENE CONVERSION

Genotypic Analysis pending

293T cells expressing the Mostaza recombination reporter were sequentially transfected with pCMV-ICP8 and with oligos that convert Mostaza to green. Oligo size is shown on the X axis. Oligo-as indicates that oligo sequence correspond to the antisense strand of egfp and oligo-s indicates that oligo sequence correspond to the sense strand of egfp. Recombinants were quantified using flow cytometry. The no oligo data for 293T or 293T-pCMV-ICP8 was subtracted to each data point. Data were analyzed and plotted using Prism 5 software.
Recombineering in HeLa cells

GENE TARGETING PRODUCES OLIGO DEPENDENT GENE CONVERSION

Genotypic Analysis pending

Cells were split and grown in 5% serum to 70-80% confluency for a total of 3 days. At about, 48 hours of growth, the incubator was found at 41.5 °C. Cells were collected and transfected simultaneously with 10 µg pCMV-ICP8 and 1 µg oligo using Nucleofector II and the program Q-013. Cells were grown until they reached 10^7 and were taken to be sorted for loss of green fluorescence.
Recombineering enzyme interactions with viral host proteins

The hypothesis of this work is that viral recombinases coordinate their activity with specific host proteins through protein-protein and functional interactions. The prediction was that viral recombination proteins coordinate with host DNA metabolic proteins to increase recombination efficiency. The results showed that viral proteins blocked host exonucleases from degrading the ssDNA recombination intermediates, viral-host protein interactions promoted increased activity of substrate processing and annealing, and provided evidence for recruitment of host proteins involved in the resolution of branched recombination intermediates, recruitment of host DNA replication machinery to assemble replisomes at recombination intermediates, and completion of recombination via disassembly of SynExo subunits that remain bound to recombinant products.

Interaction data from the λ SynExo pull downs, MS, genetic studies and from other protein interaction studies. Figure was created using additional information from STRING (http://string-db.org/) and IntAct (http://www.ebi.ac.uk/intact/).

Picture constructed by Dr. Rik Myers.
REFERENCES


doi:10.1038/mtna.2012.9

doi:10.1038/mtna.2012.51


Curtis, K. M. (2010). Rac1b regulates the neurotrophin-3 mediated neuronal commitment of bon. *Open Access Dissertations*


