2018-11-15

An Integrated Bioinformatic and Statistical Analysis of NGS Data Relevant to Diabetic Retinopathy

David Wayne Sant
University of Miami, davidwsant@gmail.com

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

Recommended Citation
https://scholarlyrepository.miami.edu/oa_dissertations/2200

This Open access 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

AN INTEGRATED BIOINFORMATIC AND STATISTICAL ANALYSIS OF NGS DATA RELEVANT TO DIABETIC RETINOPATHY

By

David Wayne Sant

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 2018
UNIVERSITY OF MIAMI

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

AN INTEGRATED BIOINFORMATIC AND STATISTICAL ANALYSIS OF NGS
DATA RELEVANT TO DIABETIC RETINOPATHY

David Wayne Sant

Approved:

Gaofeng Wang, Ph.D.
Associate Professor of Human Genetics

Rong Wen, M.D., Ph.D.
Professor of Ophthalmology

Eden R. Martin, Ph.D.
Professor of Human Genetics and
Public Health Sciences

Zane Zeier, Ph.D.
Assistant Professor of
Psychiatry and Behavioral
Sciences

Juan Young, Ph.D.
Associate Research Professor of
Human Genetics

Guillermo Prado, Ph.D.
Dean of the Graduate School
Diabetic Retinopathy (DR) is a complication of both Type 1 diabetes mellitus (T1DM) and Type 2 diabetes mellitus (T2DM) and is the leading cause of vision loss in working age adults around the world. DR is estimated to be present in 93 million individuals and most of the pathological events in DR can be attributed to the breakdown of the blood-retinal barrier (BRB).

Hyperglycemia in diabetes causes a local ascorbate (vitamin C) deficiency in the eye. Importantly, the role of ascorbate in epigenetic regulation of the genome has been largely ignored until recently. Our lab was the first to publish the role of ascorbate as a cofactor for ten eleven translocation (TET) methylcytosine dioxygenases, the enzymes that actively demethylate cytosines from methylcytosine (5mC) to hydroxymethylcytosine (5hmC), and further to intermediates that are replaced by unmodified cytosine. In the absence of ascorbate, TET enzymes cannot work at full capacity and there is a global reduction of 5hmC in the genome. Both 5mC and 5hmC play a role in
transcriptional regulation, and a global reduction of 5hmC has been associated with many diseases.

The known ascorbate deficiency in the eyes of diabetics integrated with the new knowledge of the role of ascorbate in epigenetics has led to the hypothesis that hyperglycemia-induced ascorbate deficiency in the eyes of diabetics contributes to epigenetic dysregulations in the retina that promote the development of diabetic retinopathy. Further, restoring the ascorbate levels in the eyes of diabetic patients may prevent further progress of DR symptomology.

The BRB has two sides, the inner BRB and the outer BRB. Both sides break down in the diabetic eye, but the role of the outer BRB in DR is less understood. The outer BRB is composed primarily of retinal pigment epithelial (RPE) cells that transport nutrients to the photoreceptors from the blood and remove waste from the photoreceptors, as well as produce many hormones that promote cell survival in the retina. Using cultured RPE cells as a model, I have investigated the influence of ascorbate restoration on the hydroxymethylome in cells of the outer BRB. In particular, a global increase in 5hmC was observed and corresponded with a very large portion (nearly half) of the regions of 5hmC enrichment shifted in strength after the addition of physiological levels of ascorbate to the culture media (Chapter 2). In addition to the changes in the hydroxymethylome of RPE cells, the transcriptome was altered by the addition of physiological levels of ascorbate to the culture media (Chapter 3). Cellular signaling pathways associated with ascorbate induced transcriptional changes were found to be related to DR or eye development.
One gene, vascular endothelial growth factor-A (VEGFA), is well known for its role in both eye development and diabetic eye complications and was found to decrease dramatically after the addition of ascorbate to the culture media. This gene is known to be overexpressed in DR and is thought to be responsible for the breakdown of the BRB. Pharmacologically reducing VEGF (the protein encoded by VEGFA) in the eye leads to improvement of vision, indicating that perhaps using ascorbate to suppress expression of VEGFA may prevent worsening of DR symptoms. In Chapter 4 I discuss the validation of this important finding. The decrease in VEGF was confirmed at both the RNA and protein levels, and further confirmed in primary cells from three separate donors. Finally, additional ascorbate in the drinking water of gulonolactone oxidase deficient mice (Gulo<sup>−/−</sup>) was shown to decrease the levels of VEGF in the eyes at both the RNA and protein level, indicating that dietary vitamin C can affect the production of VEGF in the eye and may be therapeutic in preventing onset or progression of diabetic retinopathy.
To my wonderful wife, Brooke:

Thank you for supporting me through all of the years I have spent chasing down my dreams.

I wanted to write you a beautiful love song, but this is the only kind of writing I know how to do.
Acknowledgments

First I would like to thank my mentor, Dr. Gaofeng Wang for his help and guidance with the design and execution of my thesis project. I would like to thank him for his enthusiasm toward completing projects and for pushing me to finish my projects. I would also like to thank Dr. Wang for allowing me to branch out and explore bioinformatics instead of just letting an analyst do the work in exchange for a position on the publications.

I would like to thank the members of the lab for their help and support. I would like to thank Dr. Vladimir Camarena for caring for the mice for all of those months and for helping me finish up the final pieces of many aspects of my project. I would like to thank Dr. Camarena for training me on lab techniques including cell culture and western blotting. I would like to thank our previous lab member Christopher B. Gustafson, who is currently attending medical school at Lake Eerie College of Medicine, for training me in lab techniques including qPCR and dot blots. I would like to thank Dr. Sushmita Mustafi for her help with aspects of my projects and for helping me keep things organized. I would like to thank Tyler Huff and Zachary Wilkes for helping teach me how to perform immunofluorescent staining. I would also like to thank Mr. Huff and Dr. Camarena for their insightful conversations that spurred deep scientific thinking on a daily basis. I wish Mr. Huff the best of luck with completing his studies.

I would like to thank our collaborators, Dr. Rong Wen and Dr. Yiwen Li. I would like to thank Dr. Wen for spending many hours with me discussing how to design the experiments for my thesis and helping me attain the resources needed. I would also like to thank Dr. Wen for teaching me how to obtain primary human fetal RPE cells from the National Eye Institute, which were vital for my thesis work. I would like to thank Dr. Li for performing all of the mouse retinal dissections for my projects.
I would like to thank Derek Van Booven for taking so much time to train me to use the Pegasus system for Linux work and for showing me examples of how to analyze many types of data. I would like to thank Dr. Eden Martin for meeting with me individually multiple times to discuss the statistical aspects of my projects. I would like to thank Dr. Bill Scott, the previous Human Genetics and Genomics graduate program director for meeting with me individually many times to discuss aspects of my thesis in addition to giving me career advice. I would like to thank Dr. Susan Blanton, the current graduate program director for helping me determine what was needed to complete my work. I would also like to thank Dori McLean for helping me organize all of my meetings and getting things scheduled.

Finally, I would like to thank my wife, Brooke and our four children: Aaron, Matthew, Kathryn and Karly. They have been my emotional support team and have helped me to have the strength to continue working on my thesis, even when it seemed the work would never be finished.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of Figures</td>
<td>viii</td>
</tr>
<tr>
<td>List of Tables</td>
<td>x</td>
</tr>
<tr>
<td>List of Abbreviations</td>
<td>xi</td>
</tr>
</tbody>
</table>

## Chapter

1. **Introduction** ........................................................................................................ 1
   - What is Diabetes Mellitus? .................................................................................. 1
   - Incidence of Diabetes Mellitus ......................................................................... 3
   - Anatomy of the Retina .......................................................................................... 3
   - The Inner Blood Retinal Barrier ....................................................................... 4
   - The Outer Blood Retinal Barrier ....................................................................... 7
   - Changes in the Retina in Diabetes .................................................................... 8
   - Screening and Prevention of Diabetic Retinopathy ....................................... 12
   - Treatment of Diabetic Retinopathy .................................................................. 15
   - The Genetic Aspect of Diabetic Retinopathy .................................................. 18
   - The Role of VEGF in the Eye ............................................................................ 21
   - Environmental Influence on Diabetic Retinopathy ....................................... 23
   - Vitamin C Deficiency in the Diabetic Eye ....................................................... 28
   - Vitamin C in Epigenetics ................................................................................... 30
   - Hydroxymethylcytosine Association with Disease .......................................... 32
   - Hypothesis .......................................................................................................... 34

2. **Ascorbate Induces Generation of 5hmC and Shifts the Hydroxymethylome in Retinal Cells** ......................................................... 36
   - Background ......................................................................................................... 36
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.1</td>
<td>The Anatomy of the Retina</td>
<td>5</td>
</tr>
<tr>
<td>Figure 1.2</td>
<td>The Cascade of Events Leading to the Breakdown of the Blood Retinal Barrier in Diabetic Retinopathy</td>
<td>13</td>
</tr>
<tr>
<td>Figure 2.1</td>
<td>Induction of 5hmC in ARPE-19 Cells by Ascorbate Treatment</td>
<td>45</td>
</tr>
<tr>
<td>Figure 2.2</td>
<td>Correlation Plots for Spike-In Methylated and Hydroxymethylated Controls Used in TAB-seq</td>
<td>48</td>
</tr>
<tr>
<td>Figure 2.3</td>
<td>hMeDIP-qPCR of Diagenode Spike-In Controls Was Used to Validate the Antibody Used for hMeDIP-seq</td>
<td>50</td>
</tr>
<tr>
<td>Figure 2.4</td>
<td>Peak Models and Cross-Correlation Plots from MACS2 for Each Sample</td>
<td>54</td>
</tr>
<tr>
<td>Figure 2.5</td>
<td>Irreproducible Discovery Rate (IDR) Graphs and Fingerprint Plot for hMeDIP-seq on ARPE-19 Samples</td>
<td>59</td>
</tr>
<tr>
<td>Figure 2.6</td>
<td>hMeDIP-seq Results of ARPE-19 Cells Treated with either 0 μM or 50 μM Ascorbate</td>
<td>61</td>
</tr>
<tr>
<td>Figure 3.1</td>
<td>Quality Control for RNA-seq on ARPE-19 Cells Treated with either 0 μM or 50 μM Ascorbate</td>
<td>67</td>
</tr>
<tr>
<td>Figure 3.2</td>
<td>The Transcriptome of ARPE-19 Cells Shifts after Treatment with Ascorbate</td>
<td>71</td>
</tr>
<tr>
<td>Figure 3.3</td>
<td>Hydroxymethylcytosine Levels Differ between Transcripts of Different Expression Levels</td>
<td>73</td>
</tr>
<tr>
<td>Figure 4.1</td>
<td>RNA Decreases and 5hmC Increases in the Gene Body of VEGFA after Ascorbate Treatment</td>
<td>87</td>
</tr>
<tr>
<td>Figure 4.2</td>
<td>qRT-PCR Verification of the Suppression of VEGFA by Ascorbate</td>
<td>89</td>
</tr>
<tr>
<td>Figure 4.3</td>
<td>The Suppression of VEGF by Ascorbate was Confirmed at the Protein Level by ELISA Testing</td>
<td>90</td>
</tr>
</tbody>
</table>
Figure 4.4  HIF-1α and Antioxidant effects are Not Responsible for the Suppression of VEGF by Ascorbate in Culture Conditions .......... 92

Figure 4.5  Ascorbate Reduces the Expression of VEGFA in Primary hfRPE Cells ........................................................................................................... 94

Figure 4.6  Decrease in the Levels of VEGF in the RPE/Choroid and Vitreous Humor of Ascorbate–Deficient Mice Supplemented with Additional Ascorbate .............................................................................. 95
List of Tables

Table 2.1  Alignment Rates for hMeDIP and Input Control Samples .......... 52
Table 2.2  Number of Peaks below the Irreproducible Discovery Rate (IDR) Threshold ................................................................. 57

Table 3.1  Alignment Rates for RNA-seq Samples ................................ 66
Table 3.2  Pathway Analysis of Differential Genes that Contain a DhMR from the RNA-seq Data from ARPE-19 Cells using the Gene Ontology Database ................................................................. 75
Table 3.3  Genes in the Eye Morphogenesis Pathway ............................... 76

Table 4.1  Primers used for qPCR Validation in Cell Based Studies .......... 82
Table 4.2  Primers used for qPCR Validation in Mouse RPE/Choroids ........ 84
Table 4.3  hMeDIP Peaks within the VEGFA Genomic Region............... 86
List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$^1$H-NMR</td>
<td>Hydrogen Nuclear Magnetic Resonance Spectroscopy</td>
</tr>
<tr>
<td>2HG</td>
<td>2-Hydroxyglutarate</td>
</tr>
<tr>
<td>2OG</td>
<td>2-Oxoglutarate</td>
</tr>
<tr>
<td>5caC</td>
<td>5-Carboxylcytosine</td>
</tr>
<tr>
<td>5fC</td>
<td>5-Formylcytosine</td>
</tr>
<tr>
<td>5hmC</td>
<td>5-Hydroxymethylcytosine</td>
</tr>
<tr>
<td>5mC</td>
<td>5-Methylcytosine</td>
</tr>
<tr>
<td>AA</td>
<td>Ascorbic Acid</td>
</tr>
<tr>
<td>ADA</td>
<td>American Diabetes Association</td>
</tr>
<tr>
<td>AMD</td>
<td>Age-Related Macular Degeneration</td>
</tr>
<tr>
<td>AML</td>
<td>Acute Myeloid Leukemia</td>
</tr>
<tr>
<td>ASD</td>
<td>Autism Spectrum Disorder</td>
</tr>
<tr>
<td>ATCC</td>
<td>American Type Culture Collection</td>
</tr>
<tr>
<td>BAM</td>
<td>Binary Alignment Map</td>
</tr>
<tr>
<td>BRB</td>
<td>Blood Retinal Barrier</td>
</tr>
<tr>
<td>ChIP-seq</td>
<td>Chromatin Immunoprecipitation Sequencing</td>
</tr>
<tr>
<td>CoCl$_2$</td>
<td>Cobalt Chloride</td>
</tr>
<tr>
<td>DAPI</td>
<td>Diamidino-2-Phenylindole</td>
</tr>
<tr>
<td>DCCT</td>
<td>Diabetes Control and Complications Trial</td>
</tr>
<tr>
<td>DHAA</td>
<td>Dehydroascorbic Acid</td>
</tr>
<tr>
<td>DhMR</td>
<td>Differentially Hydroxymethylated Region</td>
</tr>
<tr>
<td>DME</td>
<td>Diabetic Macular Edema</td>
</tr>
<tr>
<td>DMR</td>
<td>Differentially Methylated Region</td>
</tr>
<tr>
<td>DNMT</td>
<td>DNA Methyltransferase</td>
</tr>
<tr>
<td>DNMT1</td>
<td>DNA Methyltransferase-1</td>
</tr>
<tr>
<td>DNMT3A</td>
<td>DNA Methyltransferase-3A</td>
</tr>
<tr>
<td>DNMT3B</td>
<td>DNA Methyltransferase-3B</td>
</tr>
<tr>
<td>DR</td>
<td>Diabetic Retinopathy</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
</tr>
<tr>
<td>ECM</td>
<td>Extracellular Matrix</td>
</tr>
<tr>
<td>EPO</td>
<td>Erythropoietin</td>
</tr>
<tr>
<td>FACS</td>
<td>Fluorescence Activated Cell Sorting</td>
</tr>
<tr>
<td>FDA</td>
<td>Food and Drug Administration</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>FDR</td>
<td>False Discovery Rate</td>
</tr>
<tr>
<td>FPKM</td>
<td>Fragments per Kilobase per Million Mapped Reads</td>
</tr>
<tr>
<td>GCG</td>
<td>Glucagon</td>
</tr>
<tr>
<td>GLUT1</td>
<td>Glucose Transporter 1</td>
</tr>
<tr>
<td>GLUT3</td>
<td>Glucose Transporter 3</td>
</tr>
<tr>
<td>GSH</td>
<td>Glutathione</td>
</tr>
<tr>
<td>Gulo</td>
<td>Gulonolactone Oxidase</td>
</tr>
<tr>
<td>GWAS</td>
<td>Genome-Wide Association Study</td>
</tr>
<tr>
<td>hfRPE</td>
<td>Human Fetal Retinal Pigment Epithelial Cells</td>
</tr>
<tr>
<td>HIF-1α</td>
<td>Hypoxia-Inducible Factor-1 Alpha</td>
</tr>
<tr>
<td>hMeDIP-qPCR</td>
<td>hydroxymethylated DNA immunoprecipitation quantitative polymerase chain reaction</td>
</tr>
<tr>
<td>hMeDIP-Seq</td>
<td>hydroxymethylated DNA immunoprecipitation sequencing</td>
</tr>
<tr>
<td>HPLC</td>
<td>High-Performance Liquid Chromatography</td>
</tr>
<tr>
<td>IDH1</td>
<td>Isocitrate Dehydrogenase-1</td>
</tr>
<tr>
<td>IDH2</td>
<td>Isocitrate Dehydrogenase-2</td>
</tr>
<tr>
<td>IDR</td>
<td>Irreproducible Discovery Rate</td>
</tr>
<tr>
<td>ITGA2</td>
<td>Integrin Alpha 2</td>
</tr>
<tr>
<td>LDH</td>
<td>Lactate Dehydrogenase</td>
</tr>
<tr>
<td>MHC</td>
<td>Major Histocompatibility Complex</td>
</tr>
<tr>
<td>MLL</td>
<td>Mixed Lineage Leukemia</td>
</tr>
<tr>
<td>MMP9</td>
<td>Matrix Metalloproteinase-9</td>
</tr>
<tr>
<td>NEI</td>
<td>National Eye Institute</td>
</tr>
<tr>
<td>oxBS-seq</td>
<td>Oxidative Bisulfite Sequencing</td>
</tr>
<tr>
<td>P4H</td>
<td>Prolyl 4-Hydroxylase</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate-Buffered Saline</td>
</tr>
<tr>
<td>PDR</td>
<td>Proliferative Diabetic Retinopathy</td>
</tr>
<tr>
<td>PEDF</td>
<td>Pigment Epithelium Derived Factor</td>
</tr>
<tr>
<td>POLG1</td>
<td>Mitochondrial DNA Polymerase Gamma-1</td>
</tr>
<tr>
<td>PRC2</td>
<td>Polycomb Repressive Complex-2</td>
</tr>
<tr>
<td>PVDF</td>
<td>Polyvinylidene Fluoride</td>
</tr>
<tr>
<td>QC</td>
<td>Quality Control</td>
</tr>
<tr>
<td>RCPM</td>
<td>Read Counts per Million Mapped Reads</td>
</tr>
<tr>
<td>RNA-seq</td>
<td>RNA-Sequencing, Whole Transcriptome Sequencing</td>
</tr>
<tr>
<td>ROS</td>
<td>Reactive Oxygen Species</td>
</tr>
<tr>
<td>Term</td>
<td>Description</td>
</tr>
<tr>
<td>------------</td>
<td>-------------------------------------------------------</td>
</tr>
<tr>
<td>RPCs</td>
<td>Retinal Peripapillary Capillaries</td>
</tr>
<tr>
<td>RPE</td>
<td>Retinal Pigment Epithelium</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis</td>
</tr>
<tr>
<td>SELP</td>
<td>Selectin-P</td>
</tr>
<tr>
<td>SNP</td>
<td>Single Nucleotide Polymorphism</td>
</tr>
<tr>
<td>T1DM</td>
<td>Type 1 Diabetes Mellitus</td>
</tr>
<tr>
<td>T2DM</td>
<td>Type 2 Diabetes Mellitus</td>
</tr>
<tr>
<td>TAB-Seq</td>
<td>Tet-Assisted Bisulfite Sequencing</td>
</tr>
<tr>
<td>TCF7L2</td>
<td>Transcription Factor 7 Like 2</td>
</tr>
<tr>
<td>TER</td>
<td>Transepithelial Resistance</td>
</tr>
<tr>
<td>TES</td>
<td>Transcription End Site</td>
</tr>
<tr>
<td>TET</td>
<td>Ten-Eleven Translocation Methylcytosine Dioxygenase</td>
</tr>
<tr>
<td>TET1</td>
<td>Ten-Eleven Translocation Methylcytosine Dioxygenase-1</td>
</tr>
<tr>
<td>TET2</td>
<td>Ten-Eleven Translocation Methylcytosine Dioxygenase-2</td>
</tr>
<tr>
<td>TET3</td>
<td>Ten-Eleven Translocation Methylcytosine Dioxygenase-3</td>
</tr>
<tr>
<td>TSS</td>
<td>Transcription Start Site</td>
</tr>
<tr>
<td>UCSC</td>
<td>University of California Santa Cruz</td>
</tr>
<tr>
<td>UV</td>
<td>Ultraviolet</td>
</tr>
<tr>
<td>VEGF</td>
<td>Vascular Endothelial Growth Factor</td>
</tr>
<tr>
<td>VEGFA</td>
<td>Vascular Endothelial Growth Factor-A</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

What is Diabetes Mellitus?

Diabetes Mellitus is a common metabolic disorder characterized by chronic hyperglycemia. Chronic hyperglycemia causes disturbances to the metabolism of fat, protein and carbohydrates\(^1\). The classification of diabetes type is dependent upon the cause of hyperglycemia. Type 1 diabetes mellitus (T1DM) generally occurs in children and is caused by destruction of the pancreatic islet beta cells, resulting in insulin deficiency that must be corrected by insulin injections for survival\(^1\). Initial symptoms of T1DM include weight loss, excessive thirst, constant need for urination (polyuria), fatigue and blurring of vision. T1DM is usually diagnosed in children when they present with a fasting plasma glucose above 7 mM (126 mg/dL), but it is common for children with T1DM to present with very high plasma glucose levels (>11.1 mM or 200 mg/dL), fruity smelling breath from ketones and ketonuria, indicating a likely complete loss of insulin production. Insulin injections need to be initiated immediately and will be required for survival\(^1\).

The remaining cases of diabetes are type 2 diabetes mellitus (T2DM), which generally presents in adults. T2DM is caused by an insufficient production of insulin, a resistance to insulin or both\(^1\). Despite hyperglycemia and insulin imbalances, the majority of T2DM patients do not require insulin for survival, but some use insulin injections to maintain glycemic control. By the time diagnosis is made, most patients present with both inadequate production of insulin and resistance to insulin. T2DM is commonly diagnosed by an HbA\(_{1c}\) test result, where a result between 5.7% and 6.4% is classified as prediabetic and greater than 6.5% is indicative of diabetes\(^2\). Due to testing limits, two separate tests
are required for diagnosis of diabetes. Other indications of T2DM include a fasting plasma glucose level $\geq 7$ mM, a random blood plasma glucose level $\geq 11.1$ mM or a plasma glucose level $\geq 11.1$ mM two hours following ingestion of 75 grams of glucose in water.

Although T2DM may not necessarily be caused directly by obesity, the majority of patients with T2DM are obese and increased body fat has been shown to be related to insulin resistance. Many of the patients with T2DM that are not obese by traditional weight criteria have an increased proportion of abdominal body fat. The additional adipose tissue of obese individuals releases increased amounts of non-esterified fatty acids and chemokines that both induce insulin resistance and impair the function of pancreatic $\beta$-cells, which secrete insulin. Most obese individuals are not hyperglycemic because their $\beta$-cells maintain a high enough level of insulin secretion to match their needs, but they are still at risk of developing a reduced ability to secrete insulin after years of overstimulation of the $\beta$-cells.

Diabetic symptoms of patients with T2DM are usually less severe, and many individuals do not realize they are experiencing complications until years after the onset of hyperglycemia. For this reason, it is recommended that overweight individuals (body mass index, BMI $> 25$ kg/m$^2$) or with central obesity (waist to hip ratio $> 0.90$ for males and $> 0.85$ for females) be screened for diabetes. Additionally, it is suggested that all adults over age 45 be screened for diabetes. Although the majority of these individuals will not be diagnosed as having T2DM, an HbA$_{1c}$ test will likely determine that a large proportion of these individuals are prediabetic (HbA$_{1c}$ level between 5.7% and 6.4%). At this stage, controlling diet and exercise can decrease body fat and lower the risk of progressing to T2DM.
Incidence of Diabetes Mellitus

T1DM presents in approximately 1 in 500 people and accounts for only ~5-10% of all diabetics, while the remaining cases of diabetes are T2DM. It is estimated that there are 382 million people in the world and 29 million people in the United States living with diabetes. From the year 2001 to the year 2009 T1DM increased by 21% and T2DM increased by 30.5%, and occurrence is estimated to grow from 2.8% of the world population in 2000 to 4.4% of the total world population by the year 2030. By the year 2040 the number of diabetics throughout the world is estimated to reach 600 million. It is also estimated that approximately half of people with diabetes under age 50 are undiagnosed and about 23% of those over age 50 are undiagnosed. Additionally, it is estimated that 352 million adults between ages 20 and 79 could be classified as having prediabetes by the American Diabetes Association standards. It is estimated that 5-10% of prediabetic individuals progress to T2DM annually. Additionally, the prevalence of overweight and obese people in the United States is very high, and is estimated to plateau by the year 2030 at 28% of the population being classified as overweight (25 ≤ BMI < 30), 32% of the population being classified as obese (30 ≤ BMI < 40) and 9% of the population as extremely obese (40 < BMI).

Anatomy of the Retina

The retina is the organ responsible for converting visible light signals that enter the eye into the electrochemical signals that can be sent to the brain for interpretation. It is divided into eight separate layers, listed from anterior to posterior: nerve fiber layer, ganglion cell layer, inner plexiform layer, inner nuclear layer, outer plexiform layer, outer nuclear layer, outer segments of rods and cones, and finally the retinal pigment epithelium (RPE, Figure 1.1). The posterior side of the RPE is bordered by the Bruch’s membrane,
the external limiting membrane divides the outer nuclear layer from the outer segments of the rods and cones, and the internal limiting membrane borders the anterior of the nerve fiber layer. The outer segments of the photoreceptors (rods and cones), located in the outer segment, receive the light waves and convert them to electrochemical signals. The photoreceptor cell bodies in the outer nuclear layer convey this signal through the bipolar cells, amacrine cells and horizontal cells in the inner nuclear layer and then to the retinal ganglion cells in the ganglion cell layer\textsuperscript{14}. To support these cells of the neural circuit the retina also contains microglia cells and macroglia cells (Müller cells and astrocytes), as well as macrophages throughout the inner layers.

Interestingly, to reach the photoreceptors light must pass not only through the aqueous humor, the lens and the vitreous humor, but through all of the retinal cell layers except the RPE. To allow for the least obstruction of vision, the cell layers of the retina (aside from the RPE) are almost entirely transparent\textsuperscript{14}. However, the cells of the retina are very metabolically active and require a large amount of nourishment from the blood.

**The Inner Blood Retinal Barrier**

Like the brain, the retina has a specialized barrier to keep any unwanted flow of nutrients (including fluid) from entering, known as the blood-retinal barrier (BRB). The BRB is divided into two sides, the inner BRB and the outer BRB. The majority of the cells of the retina receive their nourishment from the inner BRB. The main components of the inner BRB are the capillary endothelial cells of the retina and pericytes, which cover the outside of the endothelial cells. The endothelial cells have an elaborate network of tight junctions to contain the blood\textsuperscript{15}. Interestingly, co-culture systems have been used to show that brain endothelial cells produce tight junctions because of the secretions of pericytes and reduce the permeability of sodium fluorescein\textsuperscript{16,17}. 
Figure 1.1. The Anatomy of the Retina. This diagram represents the multiple layers and cell types of the retina. The blood vessels of the inner blood retinal barrier are shown by red dashed lines, where the different types of dashes represent the different layers of blood vessels. The green arrows represent the direction of nutrients coming from the choriocapillaris through the outer blood retinal barrier, while wastes are carried the opposite direction through the RPE cells back to the choriocapillaris. Image modified from The Retina Reference (http://www.retinareference.com).
This same relationship is thought to exist within the retina as well\textsuperscript{14}. This network of tight junctions controls the flow so tightly that even sodium ions don’t pass through freely\textsuperscript{18}. The endothelial cells of the inner retina branch off from the central retinal artery into three segments of capillary networks: the radial peripapillary capillaries (RPCs), the inner capillaries, and the outer capillaries\textsuperscript{19}. The innermost branch (the RPCs) lies within inner part of the nerve fiber layer. The middle segment is the inner capillaries and runs through the retinal ganglion cells and underlies the RPCs. The final segment is the outer capillaries and runs from the inner plexiform layer through the inner nuclear layer to the outer plexiform layer (Figure 1.1).

The retina’s transparency is important for vision and the blood vessels that provide nourishment must cause minimal interference with the passage of light. In primates, including humans, there is a capillary free zone of the retina that is highly enriched with cones. This area, known as the macula, is required for high contrast central vision\textsuperscript{14}. The outer and inner capillary networks merge into a single network around the region that surrounds the macula, known as the parafoveal region\textsuperscript{19}. Despite the fact that the majority of the cell types of the retina receive their nourishment from the inner BRB, only 15%-35% of the total blood flow to the retina passes through the inner retina\textsuperscript{19}.

One important feature of the inner BRB is the regulation of blood flow. The endothelial cells are covered with a thick basement membrane and pericytes\textsuperscript{20}. Pericytes contract and relax to regulate the tone of the capillaries and control of blood flow, and they synthesize the majority of the structural and adhesive proteins of the extensive extracellular matrix of the retina\textsuperscript{21}. Approximately 95% of the endothelial cell layers in the retina are covered by pericytes, which is higher than that of any other tissue in the body including the brain, indicating a tight regulation of blood flow in the retina\textsuperscript{22,23}. Retinal blood
flow is autoregulated\textsuperscript{24}, and in absence of disease the retina has the capability to maintain a constant perfusion pressure even when blood pressure is greatly increased\textsuperscript{15}.

**The Outer Blood Retinal Barrier**

The main constituent of the outer BRB is the RPE, which is a highly polarized monolayer of cells that divides the Bruch’s membrane and choroid from the photoreceptors\textsuperscript{15}. The choroid has a very rich vasculature that covers almost the entirety of the back of the retina. The capillaries of the choroid (choriocapillaris) are very permeable to small molecules, which are instead blocked from the retina by the RPE\textsuperscript{18}. Interestingly, unlike the retinal endothelial cells, the RPE is able to form a tight barrier from the blood without the presence of pericytes\textsuperscript{14}. Although the outer BRB controls the passage of nutrients from the blood to only the outer nuclear layer and the outer segments of the photoreceptors, the choroidal blood vessels receive 65\%-85\% of the blood flow to the retina\textsuperscript{19}.

Interestingly, although the RPE creates a barrier from the blood, it also has many other functions including recycling all-\textit{trans} retinal to 11-\textit{cis} retinal as part of the visual cycle\textsuperscript{25}. The retina has very high oxygen consumption for the metabolism of the photoreceptors and high oxidative stress from direct exposure to light, leading to a breakdown of the photoreceptor outer segments with entire renewal every 7 to 10 days\textsuperscript{26,27}. The RPE is responsible for phagocytosis of the outer segments of photoreceptors as they degrade. The RPE contributes to the immune privilege of the eye and expresses essential major histocompatibility complex (MHC) molecules\textsuperscript{28}.

In addition, the RPE secretes many of the growth factors present in the retina, including vascular endothelial growth factor (VEGF) and pigment epithelium derived factor (PEDF)\textsuperscript{29}. While VEGF is a hormone that promotes neovascularization and increases
permeability of the BRB, PEDF inhibits both neovascularization and permeability, acting as an opposing factor to VEGF. VEGF is mainly secreted basally toward the choroid and is largely responsible for the rich vascular supply of the choroid. PEDF is mainly secreted apically and is partially responsible for keeping the choriocapillaris from growing through the Bruch’s membrane to the retina. The choroid has very rich vascularization and has the highest blood flow per gram of any tissue in the body, which is why it receives the majority of ocular blood flow. It is important to maintain rich vascularization in the choroid because all nutrients for the photoreceptors must come from the choroid through the RPE.

Nutrients pass through the RPE in the baso-apical direction including glucose and vitamins (Figure 1.1). Waste products of the metabolism in the retina including lactate, chloride, and water are transported through the RPE in the apico-basal direction back to the blood in the choroid. Although not traditionally thought of as a waste product, the extremely high metabolic rate of the photoreceptors generates a large amount of heat that must be removed by a large volume of blood flow. The retina has been known to be above 40°C in healthy individuals. Choroidal blood flow is not autoregulated like the retinal blood flow, and sympathetic stimulation leads to constriction of choroidal blood vessels, probably to prevent overperfusion during sympathetic stimuli.

Changes in the Retina in Diabetes

Although the most common causes of T1DM and T2DM are different and the management of symptoms is different, both lead to hyperglycemia and therefore share complications. These complications include nephropathy that can lead to complete renal failure, peripheral neuropathy that can lead to foot ulcers and Charcot joints, autonomic dysfunction including sexual dysfunction, dementia, cardiovascular complications and changes in the eye. While some diabetic complications are of great concern because
they can be lethal if not kept in check (including renal failure and heart complications), one of the most devastating complications is vision loss that occurs because of diabetes, which has been assessed to have a greater impact on quality of life than even heart disease and renal failure\textsuperscript{36}.

Diabetic Retinopathy (DR) is the disease of the diabetic retina that leads to vision loss. Owing to the frequency of diabetes, DR is the most common cause of vision loss in working-age adults worldwide\textsuperscript{9}, and frequency is expected to increase because of the growing population of diabetics\textsuperscript{10}. DR is estimated to be present to some degree in \textasciitilde{}35\% of individuals with diabetes around the world (93 million people) and vision threatening DR in estimated to be present in \textasciitilde{}10\% of individuals with diabetes (28 million people)\textsuperscript{37}.

There are four main stages of DR, ranging from mild nonproliferative DR to proliferative DR (PDR), graded on the number and severity of the symptoms of DR\textsuperscript{38,39}. Mild nonproliferative DR is graded as the presence of microaneurysms with no other symptoms. Moderate nonproliferative DR is when there are signs outside of microaneurysms (some internal hemorrhages or hard exudates), but not enough hemorrhages to be classified as severe. Severe nonproliferative DR is determined by the presence of more than 20 intraretinal hemorrhages in each of the 4 quadrants of the eye and definite venous bleeding in two or more quadrants, but no neovascularization. Proliferative DR is when neovascularization is present or vitreous hemorrhage\textsuperscript{39}. If left untreated, patients with mild forms of DR have a high likelihood of progressing to PDR and experiencing vision loss.

The first sign of DR is microaneurysms in the retinal capillaries, which are visible upon examination as small widened portions of the retinal capillaries (red spots in fundus photographs)\textsuperscript{29}. DR has been characterized as low grade inflammation in the eye\textsuperscript{29,40,41,42}, and leukocytes have been observed sticking to the retinal capillaries around the same
time as leakage of the inner BRB in rats\textsuperscript{43}. Leukocyte stiffness increases in diabetes, and it is hypothesized that the leukocytes occlude the capillaries in the retina and lead to increased pressure that causes these microaneurysms\textsuperscript{40}.

Microaneurysms are followed by either detachment of pericytes from the extracellular matrix (ECM) or the death of pericytes\textsuperscript{22,44,45}, which is known as “pericyte dropout”. Pericytes are smooth muscle cells that contract and relax to regulate the tone of the capillaries and control blood flow\textsuperscript{21}. Pericyte dropout leads to increased vascular permeability. Following pericyte dropout, endothelial cells die, which is thought to be caused by apoptosis\textsuperscript{46}, leaving acellular capillaries that are essentially empty, collagenous tubes\textsuperscript{40}. The heavily altered composition of the extracellular matrix in diabetes has been shown to contribute to the hyperpermeability of the BRB, and treatments using antisense oligonucleotides to decrease the levels of certain extracellular matrix proteins (fibronectin, laminin and collagen IV) have been shown to decrease death of both endothelial cells and pericytes\textsuperscript{47,48}. DR is traditionally known for pericyte dropout and endothelial cell death in the retinal endothelial cells. It is not well studied, but the choroid has been shown to have similar patterns of endothelial cell death and cell-free capillaries which correlate with the thickest areas of deposits along the Bruch’s membrane\textsuperscript{40}. Along with the endothelial cell death there is an increased leakage. Lipoproteins build up and make deposits known as hard exudates\textsuperscript{49}. Hemorrhages occur, one of the hallmarks of early DR\textsuperscript{50}. One more sign of early DR is cotton-wool spots, which are seen as yellowish white areas in the retina in fundus photographs. The exact cause of the cotton-wool spots is controversial, but it is thought to be a debris from interruptions in organelle transport in retinal ganglion cell exons caused by ischemia. Over the course of several months, cotton-wool spots tend to resolve themselves which sometimes leads to improvements in vision over time\textsuperscript{51}. 
The neurons of the retina have very high metabolic requirements, so the lack of viable blood vessels (ischemia) is followed by hypoxia. This hypoxia causes an increase in VEGF secretion from within the retina, which stimulates increased vascular permeability and the growth of new vessels, known as neovascularization. Furthermore, the inflammation in diabetes causes the recruitment of macrophages in the eye, and the presence of these macrophages has been shown to contribute to neovascularization. Neovascularization of the retina is a hallmark of PDR and can lead to the growth of blood vessels into the vitreous, resulting in vision loss. Additionally, the increase in contractile fibrous tissue with the new blood vessels can lead to tractional retinal detachment.

Diabetic Macular Edema (DME) is another component of DR and can occur at any stage of DR. DME is caused by breakdown of the BRB and refers to retinal thickening and accumulation of intraretinal fluid, primarily between the inner and outer plexiform layers around the center of the macula. Whereas PDR is generally more severe and more common among individuals with T1DM, DME is far more common in individuals with T2DM. DME is classified as mild, moderate or severe based on the location of the hard exudates and retinal thickening. DME can be classified as either diffuse or focal. Diffuse DME is caused by general leaking from retinal capillaries and is generally found in both eyes, whereas focal DME is found only in small sections of the eyes or a single eye and is usually accompanied by multiple microaneurysms and hard exudates.

In 2015, DME was estimated to be present in 15 million individuals in the United States (approximately half of the individuals with diabetes in the United States) and in over 150 million individuals throughout the world, making it far more prevalent than PDR. Like some of the other complications of DR, DME is caused by the breakdown and hyperpermeability of the BRB. This is caused in large part by the lack of viable vessels and related hypoxia in the diabetic retina. In healthy individuals there is always a large
amount of fluid that moves into the retina and a large amount of water produced as a result of the high ATP synthesis rates, but the RPE removes the fluid and excess proteins by active transport as quickly as they enter the retina, leading to no net buildup of fluid\textsuperscript{55}. However, when there is a breakdown of the BRB, the accompanying hyperpermeability causes water to accumulate within the retina because of hydrostatic pressure differences and effective osmotic pressure differences (Starling’s Law)\textsuperscript{56}. Additionally, the autoregulation becomes disturbed soon after the onset of diabetes, and a lack of efficient autoregulation causes a net buildup of fluid\textsuperscript{57}.

The buildup of fluid tends to occur predominantly at the center of the retina and is thought to be because of the unique anatomy of the macula. Specifically, the macula has extremely high cell count and an extremely high metabolic activity and the fovea around the macula is thicker than the macula allowing for fluid to gather in the center. The retina has no lymph system, so the removal of fluid is dependent on the vascular system. The macula is avascular which means that the blood vessels are not able to remove the fluid as efficiently\textsuperscript{58}. Due to changes in the influx and removal of proteins and fluid, it is common among patients with DME to have visual acuity fluctuate throughout the day or between days\textsuperscript{53}.

**Screening and Prevention of Diabetic Retinopathy**

Vision loss associated with DR can cause dramatic reduction in quality of life and can be a great burden on society. Even as late as the 1990’s, there were no routine screening programs to determine which patients were most at risk of developing DR and counsel them on preventing progression. One study in the United Kingdom found that more than half of the patients presenting with severe vision loss from DR had known they were diabetic, but had never received retinal examinations\textsuperscript{59}. 
The American Diabetes Association (ADA) has established guidelines for screening of patients. The recommendations were revised in 2017 to the following:

- Adults with type 1 diabetes should have an initial dilated and comprehensive eye examination by an ophthalmologist or optometrist within 5 years after the onset of diabetes.
- Patients with type 2 diabetes should have an initial dilated and comprehensive eye examination by an ophthalmologist or optometrist at the time of the diabetes diagnosis.
• If there is no evidence of retinopathy for one or more annual eye exams, then exams every 2 years may be considered. If any level of diabetic retinopathy is present, subsequent dilated retinal examinations for patients with type 1 or type 2 diabetes should be repeated at least annually by an ophthalmologist or optometrist. If retinopathy is progressing or sight-threatening, then examinations will be required more frequently.

• Women with preexisting type 1 or type 2 diabetes who are planning pregnancy or who have become pregnant should be counseled on the risk of development and/or progression of diabetic retinopathy.

• Eye examinations should occur before pregnancy or in the first trimester in patients with preexisting type 1 or type 2 diabetes, and then these patients should be monitored every trimester and for 1 year postpartum as indicated by the degree of retinopathy.

• While retinal photography may serve as a screening tool for retinopathy, it is not a substitute for a comprehensive eye exam, which should be performed at least initially and at intervals thereafter as recommended by an eye care professional.

The standards include screening at the diagnosis of diabetes in T2DM and five years after the onset of diabetes in T1DM. One of the strongest techniques for evaluating the condition of the retina is optical coherence tomography (OCT) angiography. Methods have been developed to automate the analysis of OCT images and provide robust analysis without the requirement of significant hands-on time from physicians.

The establishment of routine screening for DR can help identify those patients at the highest risk of developing vision loss so that stricter control of hyperglycemia can be implemented. A study in the United States known as the Diabetes Control and Complications Trial (DCCT) investigated intense control of hyperglycemia in patients with T1DM versus conventional treatment and found that the patients with intensive treatment were 27% less likely to develop any retinopathy and 76% less likely of progressing to clinically significant retinopathy than those in the conventional treatment group.
Additionally, it was found that therapy was most beneficial when initiated early after the onset of diabetes. In the United Kingdom a randomized study investigated the role of tight regulation of blood pressure by giving one group either and angiotensin converting enzyme inhibitor (Captopril) or a beta blocker (Atenolol) and no drug to the other group\textsuperscript{66}. After 9 years, the group that took medication showed significantly lower blood pressure, a 34\% reduction in risk of developing DR and a 47\% reduced risk of deterioration of visual acuity. These studies indicate that tight control of hyperglycemia and blood pressure reduce the risk of progression of visual complications in diabetes.

**Treatment of Diabetic Retinopathy**

Despite good glycemic control and low blood pressure, many years of living with hyperglycemia still leads to the development of DR. However, treatments exist that can help prevent retinopathy from progressing and even improve vision. One of the major treatments used is laser photocoagulation of blood vessels in the eye either at regions where leakage is present (focal photocoagulation), or at small sections across the entire retina (pan-retinal photocoagulation). In a study in which one eye was treated and the other served as a control, it was shown that laser photocoagulation reduced the risk of moderate vision loss by 50\% and increased the chance of improvement in visual acuity in DME\textsuperscript{67, 68}.

Although laser photocoagulation was clearly associated with better visual outcomes, a large proportion of patients appeared to have no benefit from treatment\textsuperscript{69}. The vitreous humor has been found to contribute to macular edema by storing chemoattractant factors and by increasing traction on the macula because of the increased glycation in hyperglycemia, so a technique for removing the peripheral vitreous humor has been developed. This procedure is called vitrectomy, and also includes the
removal of the inner limiting membrane throughout the macula\textsuperscript{70,71}. When retinal breaks are present, photocoagulation is administered in conjunction with vitrectomy. Studies have found that visual recovery in the vitrectomy group was far more likely than in the control group, and vitrectomy led to complete resolution of cystoid macular edema. Additionally, mean capillary blood flow velocity in the retina increased, leading to less hypoxia throughout the eye\textsuperscript{72,73}.

Aside from laser photocoagulation and vitrectomy, a couple of pharmaceutical approaches have been shown to improve visual acuity in patients with DR. The breakdown of the BRB, vascular leakage, hemorrhages and neovascularization have all been attributed to excess VEGF found in the eyes of diabetic patients, so most pharmaceutical approaches have been aimed at reducing the levels of VEGF in the eye\textsuperscript{74,75}. There are four main drugs that are used to decrease VEGF levels in the eye. Ranibizumab (trade name Lucentis) is a monoclonal antibody fragment that is injected into the eye (0.3 mg or 0.5 mg) to bind up the VEGF in the eye. It is produced by Genentech in San Francisco. This was the first drug made to decrease VEGF in the eye and the one best studied for effects in DR. Bevacizumab (trade name Avastin) is a similar monoclonal antibody fragment, also produced by Genentech. Whether any actual differences between the two drugs exist, Bevacizumab is less expensive because only Ranibizumab has been approved by the food and drug administration (FDA) for treating age-related macular degeneration (AMD). Aflibercept (trade name Eyela) is a soluble VEGF receptor produced by Regeneron Pharmaceuticals. Similar to the antibody approaches, this soluble VEGF receptor binds the VEGF in the eye and causes it to not bind to the receptors in the cells, therefore lowering effective VEGF levels in the eye. Pegaptanib (trade name Macugen) is a 28-base RNA aptamer that is also designed to bind to VEGF when injected into the eye.
Many clinical trials have evaluated the effectiveness of Ranibizumab in treating or preventing vision loss because of DR and DME, and they have overwhelmingly shown great benefits in restoring vision\textsuperscript{76-82}. In the RESOLVE study, patients receiving injections of Ranibizumab gained an average increase in visual acuity of 7.8 letters on a vision chart and were more than three times as likely to gain 10 or more letters on a vision chart\textsuperscript{79}. In the READ-2 study, patients were given Ranibizumab, photocoagulation or both. The greatest increase in visual acuity was seen in the Ranibizumab only group, but the group that received both Ranibizumab and laser photocoagulation required fewer injections\textsuperscript{80}. The RESTORE study found similar results to the RESOLVE study, that Ranibizumab either alone or in combination with photocoagulation was far more effective in restoring vision in patients with visual impairment\textsuperscript{76}. The RISE, READ-2, and DRCR.net studies reproduced the results and showed that Ranibizumab with photocoagulation was far superior to either sham injections or photocoagulation alone\textsuperscript{78-80}.

Bevacizumab has so far been in only a few clinical trials, but the results have been very similar to those in Ranibizumab. The studies found that injection of Bevacizumab in conjunction with laser photocoagulation was superior to laser treatment alone\textsuperscript{83-85}. Similarly, studies of Aflibercept showed an much greater increase in visual acuity in patients receiving injections rather than laser treatment\textsuperscript{86, 87}. Pegaptanib clinical trials have shown that patients receiving Pegaptanib injections showed significantly greater improvements in vision than those receiving sham injections and required fewer laser treatments\textsuperscript{88}.

Even though Ranibizumab has been used in far more studies than any of the other anti-VEGF treatments, the results overwhelmingly demonstrate the effectiveness of using anti-VEGF treatment either alone or with laser photocoagulation, and that this treatment is far superior to laser treatment alone. Little information is available regarding a
combination of anti-VEGF treatment in combination with vitrectomy, but the success of the anti-VEGF treatments has led many groups to conclude that anti-VEGF treatment should be the first line of defense in treating vision loss associated with DME rather than laser photocoagulation.

The Genetic Aspect of Diabetic Retinopathy.

Despite good blood glucose control, good blood pressure, and short duration of diabetes, some patients still develop DR, while some patients with long term hyperglycemia do not develop DR, indicating a possible genetic influence on the development of DR. Differences in rates of DR among diabetic patients from different ethnic backgrounds has also been found in American studies, also indicating the influence of genetics on DR. Studies have estimated the broad sense heritability of DR to be between 18% and 27%. A single study in Finnish patients with T1DM estimated broad sense heritability of proliferative DR to be as high as 52%, but the approaches to heritability estimation of this study have been called into question and the Finnish population is an isolated population that differs from other parts of the world, including Europe, with respect to diabetes and diabetic complications.

Many genome-wide association studies (GWAS) and candidate gene studies have investigated the role of genetic variation in DR and individually reported polymorphisms to be significantly associated with DR. Most of these associations were found by candidate gene studies and do not reach the stringent significance threshold of $5 \times 10^{-8}$ that is generally considered to be genome-wide significant, but found results that were significant with a threshold of approximately $5 \times 10^{-3}$. Most identified variants have yet to be confirmed by other reports, and other variants have conflicting results with a positive association in one study, but a negative association in another study. The reason behind
the conflicting results is likely due to the relatively small sample sizes in most studies (usually below 1,000 affected individuals) and the heterogeneity of the phenotype studied (T1DM vs T2DM vs all types of diabetes, any DR vs PDR only vs DME only)\(^\text{97}\). Many of these studies have reported a result as significant, despite not reaching significance after Bonferroni correction but state that it is still significant because of the function of the associated gene, giving more likely candidacy to the gene. Despite most papers using lower significance thresholds, there was one variant that reached genome-wide significance in a single study\(^\text{98}\). This variant was rs4838605, which lies in the gene \textit{ARHGAP22}, which codes for a Rho GTPase activating protein on chromosome 10. This association was found in a Taiwanese GWAS study to be associated with DR in patients with T2DM in comparison to patients with T2DM without DR. No studies have investigated the possible functional role of \textit{ARHGAP22} in DR.

Although the results have been largely conflicting or marginally insignificant, a few genes have been replicated by more than one study. A very thorough review of candidate gene and GWAS studies by Broadgate \textit{et al.} was used as a reference for studies performed up until 2017, and a PubMed search was used to find publications containing the terms “genetics” and “diabetic retinopathy” in August, 2018 to gain a thorough overview of the genetic studies relating to the genetics of DR\(^\text{97}\). To select the most significant genes, a significance threshold of \(5\times10^{-3}\) was used to filter for only the most significant genes. This is not as stringent as the threshold typically used for GWAS studies \((5\times10^{-8})\) but almost no variants reached this threshold and many of the studies were candidate gene studies involving fewer markers than GWAS studies. Only five genes were found to be significant by this threshold in more than one study, even if no association was found for the same marker in other studies. These five genes are \textit{EPO}, \textit{ITGA2}, \textit{SELP}, \textit{TCF7L2} and \textit{VEGFA}. 
One gene with variants found by multiple studies was the erythropoietin gene, **EPO**, which was found in three separate studies: one in Caucasian Americans and two in Han Chinese in patients with DR as opposed to T2DM without DR\textsuperscript{99-101}. The protein coded by this gene, also known as EPO, inhibits apoptosis and promotes proliferation in erythroid precursor cells. It was found that injection of EPO can help prevent neuronal injury following reduced oxygen tension (hypoxia) and even provided neuroprotection after traumatic brain injury and spinal cord injury\textsuperscript{102}. Administration of EPO was also found to improve vision and resolve hard exudates associated with DME in a small study when it was given as a treatment for anemia and renal failure, indicating a likely role of the **EPO** gene in DR\textsuperscript{103}.

A second gene found by three studies is integrin alpha 2, **ITGA2**. This gene was found in two meta-analyses and in one Japanese study comparing patients with DR to patients with T2DM without DR\textsuperscript{104-106}. This gene has been found to be profibrotic, increase collagen expression, increase reactive oxygen species (**ROS**), be anti-angiogenic, and affect platelet receptor density\textsuperscript{107-110}. In diabetes, platelets aggregate more easily in the presence of agents such as collagens, and this change in platelets has been suggested to potentially play a role in DR\textsuperscript{111, 112}.

**Selectin P (SELP)** was found to be significant in two separate candidate gene studies, one in Iran and on in the United States using mixed ethnicities comparing patients with DR to those with T2DM but no DR\textsuperscript{113, 114}. Selectin P has been found to be necessary for leukocyte recruitment by allowing for leukocyte adhesion to endothelial cells\textsuperscript{115}. Mouse models carrying a mutation in the **Selp** gene (**Selp\textsuperscript{tm1Bay}**) show deficits in leukocyte rolling and can be used as a model for leukocyte deficiency syndrome\textsuperscript{116}. Selectin P is expressed in the veins in the choroid of the eyes of healthy individuals, and this expression is elevated...
Leukocytes are observed sticking in the retina in diabetes, and this is hypothesized to contribute to microaneurysms and hemorrhage\textsuperscript{40, 43}.

In one candidate gene study investigating associations in patient with DR versus people without diabetes and in a multi-ethnicity meta-analysis, the gene transcription factor 7 like 2 (\textit{TCF7L2}) was found to be significant\textsuperscript{118, 119}. This gene was investigated in candidate gene studies specifically for its role in diabetes. \textit{TCF7L2} (the protein coded by \textit{TCF7L2}) is a transcription factor that is a key component in the Wnt signaling pathway. \textit{TCF7L2} has been found to regulate the proglucagon gene (\textit{GCG}), which helps regulate blood glucose homeostasis. Variants in this gene have also been found to be associated with insulin resistance and with diabetic coronary heart disease\textsuperscript{118, 120}.

**The Role of VEGF in the Eye**

By far, the best studied gene in the genetics of DR is vascular endothelial growth factor-A (\textit{VEGFA}). The rationale behind the study of this gene is that VEGF, the protein coded by \textit{VEGFA}, is a protein that stimulates the growth of blood vessels, leading to vascularization in areas where it is expressed. The neovascularization of proliferative DR is thought to be caused largely by overexpression of VEGF in the eye. Additionally, the most effective treatments for DR, particularly DME, are anti-VEGF therapies\textsuperscript{89}. Variants in \textit{VEGFA} have been found associated with DR in many studies of varying ethnicities and study designs. Twelve of these studies have reported associations more significant than the cutoff of $5 \times 10^{-3}$ that I have used as my threshold\textsuperscript{105, 121-131}.

VEGF is expressed in many cell types throughout the body and is known to be expressed during hypoxia in response to the upregulation of the transcription factor hypoxia-inducible factor 1-alpha (HIF-1\textalpha)\textsuperscript{132}. VEGF is very important for normal physiological processes. In mammalian development, VEGF is necessary for proper blood
vessel formation. In fact, in mice even if a single copy of Vegfa is knocked out, the mice die embryonically due to impaired blood vessel formation\textsuperscript{133,134}. In the retina, VEGF has been found to be necessary for the formation of choriocapillaris and for the development of the eye, even though the retina is one of the most well oxygenated tissues in the body. Conditional knockout of Vegfa in the RPE cells led to mice with no choroid, very small eyes (microphthalmia) and no visual function\textsuperscript{135}. Although heterozygous mice were not entirely blind, they still had fewer RPE cells, and those that were present showed abnormal pigmentation granules. The knockout of HIF-1α caused no changes in eye development or function, indicating that hypoxia is not the only factor to regulate the production of VEGF in the RPE. Even into adulthood, VEGF has been shown to be constitutively expressed in the RPE\textsuperscript{136}. Reduced expression of VEGF in adult RPE has been shown in both mice and cell culture to lead to increased apoptosis of the RPE and severe choroidal remodeling including atrophy of the choriocapillaris\textsuperscript{137,138}.

Despite the necessity of VEGF in the eye, particularly in the RPE, it must be regulated. When the endothelial cells and pericytes die in response to diabetic complications, it creates ischemia in the retina. This is followed by hypoxia that will induce the activity of HIF-1α. Although VEGF is constitutively expressed, increased activity of HIF-1α can lead to an overexpression of VEGF\textsuperscript{139}. This idea largely came to light when a study found that VEGF was at significantly higher levels in patients with DR and other ocular diseases than in healthy controls\textsuperscript{74}. An increase in VEGFA, induces hyperpermeability of the BRB\textsuperscript{140}. This is thought to be the cause of the increased leaking of the BRB that leads to an increase in lipoproteins that cause hard exudates, excess proteins that contribute to cotton-wool spots and accumulation of fluid associated with macular edema. In addition to causing hyperpermeability, VEGF has been found to cause neovascularization. Mice and rats kept in high oxygen chambers at birth and then removed
to atmospheric air receive a hypoxic shock that both increases VEGF and leads to retinopathy similar to the retinopathy of prematurity seen in humans\textsuperscript{141,142,143}. Overexpression of human VEGF in the photoreceptors of mouse eyes in a model known as the Kimba model showed the effect of increased VEGF in the eye. The mice showed retinal vascular leakage, venous bleeding, capillary dropout, neovascularization and even focal retinal detachment\textsuperscript{144}. A small number of the mice even exhibited edema. Essentially, an excess of VEGF in the retina can create all of the symptoms of diabetic retinopathy. Furthermore, injection of soluble VEGF neutralizing proteins was found to significantly reduce the retinal neovascularization in these mouse models, demonstrating that the problems are directly related to VEGF\textsuperscript{75}.

**Environmental Influence on Diabetic Retinopathy**

While genetics may play a role in the development of DR, estimates of broad sense heritability, with a single exception from a study in Finnish patients with T1DM, are below 30\%\textsuperscript{95}. The remaining effects are likely caused by the environment. The prevalence of DR increased with duration of diabetes (21.1\% in individuals with diabetes for less than 10 years, but 76.3\% in individuals with diabetes for >20 years) and was more common in individuals with T1DM than T2DM (77.3\% for T1DM, 25.2\% for T2DM)\textsuperscript{37}. Similarly, DME increases with the duration of diabetes (3\% among individuals with diabetes for less than 10 years, but 20\% in individuals with diabetes for >20 years). This increase with the duration of the disease is likely indicative of the environmental link because of increased time for environmental contributors to affect the body.

Studies of the major risk factors of developing vision loss in diabetes have found that multiple markers of management of hyperglycemia are good predictors for whether or not patients will develop DR. The top predictor for progression of DR, aside from
duration of diabetes, is HbA1C. Only 18% of patients with low HbA1C (<7.0%) had DR but it was present in 51.2% of individuals with high HbA1C (>9.0%)\textsuperscript{37}. Similarly, the DCCT study found that in patients with T1DM the strongest predictor of retinopathy progression was HbA1C, and a 10% lower HbA1C (for example, 7.2% instead of 8%) was associated with a 43% lower risk of progression\textsuperscript{145}. Additionally, studies showed that intensive treatment (keeping the HbA1C below ~7%) in patients with T2DM reduced the incidence of DR by 76% and progression of DR by 54%\textsuperscript{146}.

A second environmental factor that is associated with DR is blood pressure. In the United States in 2012, 39.6% of patients with T2DM and high blood pressure (>140/90) had DR, but only 30.8% of patients with lower blood pressure (<140/90) were found to have DR\textsuperscript{37}. This corroborated an earlier longitudinal study in the United Kingdom that found that patients with T2DM that had tightly controlled blood pressure had lower incidences of microaneurysms, hard exudates, hemorrhages, cotton-wool spots and less deterioration on the vision chart\textsuperscript{147}.

A third environmental factor that may be associated with DR is control of lipids in the blood. Studies have been inconsistent, but some have found associations between serum cholesterol and DME. One study found that LDL cholesterol was not associated with DR or non-clinically significant DME, but was associated with clinically significant DME\textsuperscript{148}. Another study found that high serum total cholesterol, LDL cholesterol and HDL cholesterol were related to DME\textsuperscript{149}. A third study investigated the effect of giving fenofibrate, a long-term lipid-lowering drug, to patients with T2DM. The risk of needing laser treatment for DR was significantly lower in the fenofibrate than it was in the control group, although they did not have better vision on average\textsuperscript{150}. A meta-analysis study in 2018 found that lipid-lowering agents show a protective effect on DR progression, but that
they had no effect on vision loss\textsuperscript{151}. More studies will be needed to determine if lipid-lowering can reduce risk of DR.

From these studies, we see that the biggest predictors of whether or not a person with diabetes (either T1DM or T2DM) will develop DR are HbA\textsubscript{1C}, blood pressure and control of lipids. Additionally, the risk of developing vision problems increases greatly with the duration of diabetes. Together, this information highlights the importance of managing the environmental factors of diabetes, particularly the level of hyperglycemia. The benefit of disease symptoms that are predominantly associated with environmental factors is that, unlike genetics, the environment can be modified to prevent the progression or onset of disease symptoms.

One interesting phenomenon found by the Diabetes Control and Complications Trial in the United States was the effects that intensive control had on patients ten years after the study was completed. The patients that had been in the “intensive control” group had significantly lower HbA\textsubscript{1C} levels during the study but returned to the same level as the “control” group within a short time. After ten years, even though the HbA\textsubscript{1C} levels of both groups had been the same for the past six years, the “intensive control” group had significantly lower incidence of eye complications than the control group\textsuperscript{152}. In an experimental animal model of diabetes, dogs with T1DM for 5 years that received intensive metabolic care within two months of the onset of diabetes showed significantly fewer eye complications than those without intensive control. However, the introduction of intense metabolic care at 2.5 years did not make a difference in comparison to the group with poor metabolic control for all 5 years, indicating that the metabolic memory was already set\textsuperscript{153}. A similar result was observed in diabetic rats\textsuperscript{154}. This observation of long lasting effects of hyperglycemia has been termed “metabolic memory”. These long lasting effects are thought by many to be the result of persistence of epigenetic modifications\textsuperscript{155, 156}.\textsuperscript{156}
Epigenetic modifications are changes to DNA or proteins that form histones that affect gene expression without making changes to the genetic code. The most studied epigenetic modification of DNA is 5mC, which accounts for approximately 1% of all DNA bases, but affects ~70%-80% of all CpG dinucleotides. The addition of the 5mC mark and maintenance of the 5mC mark are controlled in the cell by three different enzymes, termed DNA methyltransferases (DNMTs)\textsuperscript{157}. Two of the enzymes, DNMT3A and DNMT3B are responsible for establishing 5mC patterns in cells and can add 5mC \textit{de novo}, while DNMT1 is predominantly responsible for maintenance of 5mC during cell replication\textsuperscript{158}. The distribution of 5mC throughout the genome is not uniform. This is not unusual considering that the density of CpG dinucleotides is not uniform, but it is surprising that areas of high CpG density, known as CpG Islands, have lower levels of 5mC than other regions of the genome. These areas tend to be clustered around promoter regions of genes, and methylation at these sites can sometimes reduce the expression of the associated genes\textsuperscript{159}. Additionally, 5mC has been found to be required for X-chromosome inactivation in females\textsuperscript{160}. 5mC is enriched in heterochromatin, and the presence of 5mC inhibits the binding of many transcription factors\textsuperscript{161}.

A few groups have investigated a potential link between DNA methylation and diabetic retinopathy. In whole mouse retinas, long term hyperglycemia increased the expression of DNMT1\textsuperscript{162}. Similarly, in cultured RPE cells (ARPE-19), long term hyperglycemia was found to increase expression of DNMT1 and increase DNMT activity. Another group found that hyperglycemia increased DNMT1 expression in a second type of retinal cells, retinal endothelial cells\textsuperscript{163}.

The increase in DNMT activity could lead to a global increase in 5mC in diabetes. In lymphocytes the global level of 5mC was found to be higher in patients with DR than in diabetic patients without retinopathy. Furthermore, the global 5mC levels were also
significantly higher in patients with proliferative DR than in patients with non-proliferative 
DR\textsuperscript{164}. A second group investigated DNA methylation levels at 485,577 specific CpG 
dinucleotides throughout the genome and found 349 CpG's to be differentially methylated 
between the DNA from blood cells of patients with T1DM with or without retinopathy\textsuperscript{165}. 
Surprisingly, in contrast to the previous study that found a global increase in DNA 
methylation, the majority of the differentially methylated CpG's (79\%) were lower in the 
group with DR. Expression at the genes corresponding to these methylation changes was 
not investigated, but methylation levels at these CpG's could possibly be markers for DR.

One group investigated DNA methylation in the promoters of specific genes. Mitochondrial DNA polymerase gamma 1 (\textit{POLG1}) codes for the protein that forms that 
catalytic subunit of mitochondrial DNA polymerase. \textit{POLG1} is downregulated in diabetes 
and this is thought to contribute to the buildup of DNA damage in response to oxidative 
stress. It was found that the promoter of \textit{POLG1} is hypermethylated in retinal endothelial 
cells after three months of hyperglycemia, and remains hypermethylated after another 
three months of intensive metabolic control in rats\textsuperscript{166}. Matrix metalloproteinase-9 (\textit{MMP9}) 
damages mitochondria when activated, accelerating apoptosis in retinal endothelial cells. 
In high glucose conditions, the promoter region of \textit{MMP9} was hypomethylated and 
expression was increased in cultured retinal endothelial cells\textsuperscript{167}.

Collectively these studies show that DNA methylation is altered in hyperglycemia, 
and these alterations may persist after alleviation of hyperglycemia, contributing to 
metabolic memory of diabetes. These findings also emphasize the importance of early 
detection and management of diabetes.
Vitamin C Deficiency in the Diabetic Eye

One of the ways the environment can influence disease processes is through an impaired distribution of metabolites, such as vitamins, minerals and even sugars such as the overabundance of glucose in diabetes. One very important vitamin to humans is vitamin C (L-ascorbic acid, AA), which is an essential micronutrient and a very important water-soluble antioxidant in human blood\textsuperscript{168}. Most mammals produce endogenous ascorbate in the liver, but humans, guinea pigs, bats and primates have a dysfunctional gulonolactone oxidase enzyme necessary for the conversion of glucose to ascorbate\textsuperscript{169,170}. Vitamin C exists predominantly as ascorbate anion under physiological pH\textsuperscript{171}. Most cells in the body absorb ascorbate through sodium-dependent vitamin C transporters (SVCTs), but ascorbate enters the retina almost exclusively in its oxidized form, dehydroascorbic acid (DHAA), through glucose transporters (GLUT1 and GLUT3) by active transport\textsuperscript{171,172}. Although ascorbate is transported across the retina in its oxidized form, it is concentrated in the retina in the reduced form, indicating that it is converted back to ascorbate anion inside the retina\textsuperscript{172}.

Even though DHAA is the predominant form of ascorbate absorbed by the retina, glucose has been found to act as a competitive inhibitor to absorption of reduced ascorbate in cultured RPE cells, indicating that it may be locally oxidized before transport into the cells\textsuperscript{173}. Glucose has also been shown to act as a competitive inhibitor of DHAA in rat retinas ex vivo, and high concentrations of glucose (30 mM) in rat retinas have been shown to inhibit uptake of DHAA by 86.4\%\textsuperscript{172}. Chronic hyperglycemia in diabetes is thought to inhibit uptake of DHAA in the retina of diabetic patients. The retina cannot be accessed until after death limiting the studies using human retinas, but studies have found that alterations in the nutrient levels in the vitreous humor are indicative of alterations in the retina and vitreous samples can be collected upon vitrectomy\textsuperscript{174}. By high-performance
liquid chromatography (HPLC), one group in Japan found that ascorbate levels in the vitreous humor of patients with proliferative diabetic retinopathy were only approximately one third the level of patients without retinopathy\textsuperscript{175}. Using Hydrogen nuclear magnetic resonance spectroscopy (\textsuperscript{1}H-NMR) to study metabolites in the vitreous humor, another group in Spain found that the vitreous humor of patients with PDR contained less than one fourth as much ascorbate as the vitreous humor of patients with macular hole (386 $\mu$M vs 1,677 $\mu$M)\textsuperscript{176}. This was confirmed in a second set of samples using a ferric reducing power assay. In an experimental model, diabetes was induced in rats using streptozotocin. After the rats had hyperglycemia for only three weeks, blood-to-retina transport of ascorbate was reduced by 65.5\%, giving further support to the notion that hyperglycemia inhibits absorption of ascorbate in the retina\textsuperscript{177}.

The concentration of ascorbate is not consistent across tissues. In human plasma the concentration of ascorbate is generally $\sim$50 $\mu$M, but can reach up to 100 $\mu$M with increased vitamin C in the diet\textsuperscript{168}. DHAA, on the other hand, is generally only $\sim$5 $\mu$M in the plasma. In the heart and skeletal muscle, ascorbate concentration is 200-400 $\mu$M. In the kidney ascorbate is around 300-500 $\mu$M. In the liver and the lung ascorbate is around 1,000 $\mu$M, and in the brain the concentration can be between 2,000 $\mu$M and 10,000 $\mu$M\textsuperscript{178}. Concentration of ascorbate in the retina is not as high as it is in the brain, but is generally found to be between 1,000 $\mu$M and 1,600 $\mu$M\textsuperscript{172,175,179-181}. The relatively high concentration of ascorbate in the retina highlights its importance and is consistent with the active transport used to collect ascorbate from the blood. This importance is indicative that the local ascorbate deficiency caused by hyperglycemia may play a role in the pathogenesis of diabetic eye complications.
Vitamin C in Epigenetics

The high concentration of ascorbate found in the retina of healthy individuals shows its importance, and this importance has been thought to be related to oxidative stress. The retina is the only organ in the body where light is directly focused on a group of cells, and the light produces very high amounts of oxidative stress in the form of free radicals\textsuperscript{182}. By measuring the amount of lactate dehydrogenase (LDH) released from injured cells, one group was able to study the influence of ascorbate in protecting the retina. It was found that ultraviolet light (UV) significantly raised LDH levels in the rat retina. This increase was greatly diminished when ascorbate was added to the solution at 1 mM\textsuperscript{183}. Surprisingly, glutathione (GSH), a metabolite with stronger antioxidancy than ascorbate, did not protect the retinas from damage. This indicates that perhaps the importance of ascorbate in the retina is not solely for antioxidancy, but may be related to the other functions of ascorbate.

Ascorbate is well known for its function as a general antioxidant and is an essential cofactor for many enzymes\textsuperscript{171}. Deficiency of ascorbate is known to cause scurvy due to its role as a cofactor for collagen hydroxylase prolyl 4-hydroxylase (P4H), which is independent of its role as an antioxidant. Ascorbate deficiency in scurvy is known to cause abnormal bleeding, particularly in the mouth and nose as a result of deficient collagen hydroxylase action\textsuperscript{184}. The levels of ascorbate in DR are likely not so low that the collagen hydroxylases are insufficient and thus leading to the breakdown and bleeding, but it is interesting to note that both are related to bleeding and associated with a decrease in available vitamin C. Additionally, this emphasizes the importance of ascorbate as a cofactor of enzymes.

Importantly, one role of ascorbate that was ignored until recently is its role in epigenetics. The Wang lab made the discovery that ascorbate is a necessary cofactor for
the ten-eleven translocation (TET) methylcytosine dioxygenases. Both collagen P4H and TET dioxygenases are part of the iron and 2-oxoglutarate (2OG) superfamily that require reduced iron (Fe\textsuperscript{2+}) for full activity. In the absence of ascorbate, both TET enzymes and P4H can initially proceed with full catalytic activity, but Fe\textsuperscript{2+} is quickly oxidized to Fe\textsuperscript{3+} which cannot fulfill the requirement of the enzymes. In essence, ascorbate acts as a cofactor for both enzymes because it serves to reduce Fe\textsuperscript{3+} to Fe\textsuperscript{2+} and replenishes the environment to allow for the function of the enzymes.

TET dioxygenases have only relatively recently been discovered to catalyze the oxidation of 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC). Both are covalent modifications of cytosine which occur almost exclusively at CpG dinucleotides. TET proteins can further oxidize 5hmC to 5-carboxylcytosine (5caC) and 5-formylcytosine (5fC), which are further excised by thymine-DNA glycosylase to produce an abasic position that will further be replaced by unmodified cytosine.

The most prevalent and best studied epigenetic mark of DNA is 5mC, as discussed previously. It is found almost exclusively at CpG dinucleotides, enriched in heterochromatin and generally associated with repression of gene expression. During embryonic development, DNA methylation dynamics change very rapidly. 5mC levels drop to only a fraction of their level in somatic tissues and 5hmC levels are very high, largely mediated by the TET enzymes. These changes in methylation pattern are thought to affect cell pluripotency and differentiation.

Although 5hmC is an intermediate product in the process of active demethylation, it is relatively stable and present at roughly 1-10% of the level of 5mC in most cells. Whereas 5mC levels are stable across tissues, levels of 5hmC change greatly. Levels are highest in tissues of the nervous system with 5hmC levels at 0.5%-0.7% of cytosines in the cerebral cortex, brain stem, spinal cord and cerebellum. Levels were medium
levels at 0.2% of all cytosines) in many soft tissues such as kidney, bladder, heart, lung and muscle. Levels were very low (5hmC levels at 0.05% of all cytosines) in the liver, spleen, pituitary gland and testes. Both 5mC and 5hmC are known to be modifications of DNA that act as an interface between a dynamic environment and the genome, helping to allow for cellular responses to changes in the environment and development. In contrast to 5mC, 5hmC is enriched in euchromatin and is generally enriched in gene exons. Presence of 5hmC is associated with enhanced gene expression\textsuperscript{193}. The presence of 5hmC in the gene body has been consistently shown to be correlated with gene expression\textsuperscript{191}. Although generally known for increasing gene expression, presence of 5hmC in the promoter region has been shown to interfere with transcription factor binding and can silence gene expression, similar to 5mC\textsuperscript{161}. Some studies have found 5hmC in the promoter region near the transcription start site (TSS) to be associated with gene repression\textsuperscript{191}. Additionally, knocking out TET1 leads to the increase in gene expression of many genes, particularly Polycomb-targeted developmental regulators and 5hmC is thought to be necessary for PRC2 (Polycomb Repressive Complex-2, a component of the Polycomb complex) to bind to the DNA, which alters chromatin structure and represses transcription\textsuperscript{194}.

**Hydroxymethylcytosine Association with Disease**

The importance of ascorbate in the methylation/demethylation equilibrium helps explain its role in many diseases that cannot be explained by the insufficiency of collagen hydroxylases or by a decrease in antioxidant levels. Leukemias are known to have hypermethylation and a global loss of 5hmC. In fact, the TET1 enzyme was given the name of ten-eleven translocation-1 because of the translocation that was found in acute myeloid leukemia (AML) and mixed lineage leukemia (MLL)\textsuperscript{195}. Leukemia is frequently
associated with somatic mutations in the TET enzymes (TET1-TET3), in addition to mutations in isocitrate dehydrogenase 1 (IDH1) and IDH2. The mutations in IDH1/2 produce 2-hydroxyglutarate (2HG), which competitively inhibits 2-oxoglutarate (2OG) from binding with the TET enzymes and thus inhibits their function. Additionally, increased ascorbate levels through intravenous injection have been shown to decrease malignancy of leukemia and enhance sensitivity to chemotherapy196.

Most if not all types of solid tumors have also been associated with hypermethylation and global loss of 5hmC197,198. Epidemiological studies have shown evidence that increased consumption of ascorbate plays a role in prevention of many cancers168,199,171. A global loss of 5hmC in breast cancer, particularly in triple negative breast cancer, is a poor prognostic marker200. Increasing vitamin C in the culture media of breast cancer cells to higher physiological levels has been shown to increase apoptosis and decrease the invasion of breast cancer201. Melanoma, similarly, has been associated with a global loss of 5hmC and the decrease in 5hmC was much greater in metastatic melanoma cases than benign cases202. Addition of ascorbate in culture media of melanoma cells has also been shown to decrease the invasiveness of melanoma cells203. Furthermore, ascorbate in vitro and in vivo has been shown to sensitize melanoma to certain chemotherapy treatments204.

Aside from cancer, 5hmC has been implicated in other diseases. The levels of 5hmC are highest in the brain and nervous tissue, so it is natural that multiple studies have investigated the role of 5hmC in neurodegenerative diseases and psychological conditions. A global decrease in 5hmC was found in the entorhinal cortex, hippocampus and cerebellum in the brains of Alzheimer’s patients in comparison to healthy controls205, 206. Interestingly, in another study a global increase in 5hmC was found in the middle frontal gyrus and middle temporal gyrus in Alzheimer’s patients207. In mouse models of
Alzheimer’s it has also been demonstrated that there is a reduced level of vitamin C in the brain, which likely contributes to these shifts in the hydroxymethylome\textsuperscript{208}. A reduction in 5hmC has also been associated with major depression. Although measured in leukocytes and not in nervous tissue, it was found that elderly patients with major depression had significantly lower global levels of 5hmC than age-matched controls\textsuperscript{209}. A shift in 5hmC caused by environmental stressors has also been hypothesized to contribute to the development of both schizophrenia and autism spectrum disorders (ASD)\textsuperscript{210}.

A shift in hydroxymethylation was also observed in trinucleotide repeat disorders. Trinucleotide repeat disorders are genetic diseases and caused by an expansion of a three-base code in a given location of the genome. Although they are inherited and related to the expansion of the repeat, the actual pathological mechanisms are not well understood. These diseases include Huntington’s disease, Fragile X syndrome and spinocerebellar ataxia\textsuperscript{211}. It has been hypothesized that, although they do not cause the disease, epigenetic mechanisms that occur because of the repeat expansion may be related to the pathogenesis of the disease. In a mouse model of Huntington’s disease, a global loss of 5hmC was found in GABAergic neurons\textsuperscript{212}. It has also been found that in Huntington’s the flux of vitamin C within the brain is impaired, which is likely responsible, at least partially, for the global loss of 5hmC\textsuperscript{208}. Similarly, in a mouse model of Fragile-X found a global decrease in 5hmC in the cerebellum, but an increase in 5hmC at cerebellum specific enhancers and an increase 5hmC in the repetitive sequences\textsuperscript{213}.

Hypothesis

I hypothesize that hyperglycemia-induced ascorbate deficiency in the eyes of diabetics contributes to the epigenetic dysregulations in the retina that promote the development of diabetic retinopathy.
To investigate this hypothesis, I used ARPE-19 cells as a cellular model of the RPE, the main component of the outer BRB, to test the influence of ascorbate deficiency in the outer BRB. Specifically, I treated cells with culture media supplemented either with physiological levels of ascorbate (50 μM) or without ascorbate and extracted the DNA to investigate the global levels and local changes in 5hmC throughout the genome in RPE cells. The results of this study are discussed in Chapter 2. Next, I used whole transcriptome sequencing to investigate the changes in the transcriptome and integrated these changes with the changes found in Chapter 2 to find genes that are the most influenced by the ascorbate deficiency. These results are discussed in Chapter 3. Finally, I used multiple cell lines, primary human cells and mouse models to confirm reduction in the expression and secretion of VEGF in RPE cells associated with treatment with physiological levels of ascorbate (Chapter 4).
Chapter 2

Ascorbate Induces Generation of 5hmC and Shifts the Hydroxymethylome in Retinal Cells

Some of the data presented in this chapter were previously published in *Investigative Ophthalmology and Visual Sciences* (IOVS, Volume 59, Issue 8)\textsuperscript{214}. I performed all of the work presented in this chapter with a few exceptions. The cell culture and DNA extraction for the dot blots, along with the actual southern blot procedure were performed by Christopher Gustafson. The TET-assisted treatment and bisulfite treatment, followed by PCR amplification were performed by Vladimir Camarena, MD/PhD and Sushmita Mustafi, PhD. Sequencing of the bisulfite treated DNA and TET-assisted bisulfite treated DNA was performed by the Sequencing Core of the John P. Hussman Institute for Human Genomics at the University of Miami. Sample preparation and sequencing of hydroxymethylated DNA immunoprecipitation sequencing (hMeDIP-seq) was performed by the Epigenomics Core at the University of Michigan.

Background

*Perspectives*

Chronic hyperglycemia of diabetes has been shown to lead to vitamin C (ascorbate) deficiency in the eye\textsuperscript{176}. Ascorbate acts as cofactor for ten-eleven translocation (TET) methylcytosine dioxygenases, which actively demethylate DNA, rendering methylation/demethylation dynamics disrupted during ascorbate deficiency\textsuperscript{185}. This disruption has been shown to contribute to the development of many diseases, and I hypothesize that this contributes to the development of diabetic retinopathy (DR)\textsuperscript{171}. As
a first step to testing this hypothesis, I determined that I would investigate the hydroxymethylome of retinal cells after being restored to normal physiological levels of ascorbate. I used dot-blots and immunofluorescent staining to determine if there was a global increase in 5hmC, as is expected with the restoration of ascorbate. Additionally, I investigated the methods for mapping 5hmC in the genome and chose one of the methods (hMeDIP-seq) to determine the genomic location of the changes in 5hmC.

The Choice to Study the Outer BRB

Both the outer and the inner BRB change greatly during the development of DR. The inner BRB has been studied extensively in DR, as it is generally thought to contribute the greatest amount to the leakage that leads to most of the retinal pathological complications seen in DR. Additionally, it is the location of neovascularization in proliferative DR\textsuperscript{52}. The outer BRB, on the other hand, has not been well studied but has been shown to change just as greatly as the inner BRB during the development of DR. The outer BRB receives the majority of the blood flow to the retina and is responsible for the nourishment of the photoreceptors\textsuperscript{19}. The RPE cells of the outer BRB are responsible for the removal of much of the water buildup during macular edema and responsible for the phagocytosis of the outer segments of the photoreceptors to prevent buildup of extracellular debris, which may contribute to hard exudates\textsuperscript{30}. Additionally, the RPE cells secrete many hormones into the eye that influence the cells of the inner BRB and may contribute to pathogenesis in DR\textsuperscript{29}. For this reason I chose to study the RPE cells of the outer BRB and determine how they change when ascorbate is restored, and potentially uncover more of the role of the RPE in DR.
The Choice of a Cellular Model of the Outer BRB

With the decision to study the RPE came the task of determining an appropriate model. Fresh, human RPE can be obtained only post-mortem and are generally very poor quality. Additionally, this does not allow for experimentation with differing levels of ascorbate. Rodent models of DR are not available. Additionally, even if mice or other rodents are treated with different levels of ascorbate, the collection of the RPE is difficult and unreliable. Laser capture microdissection does not separate the RPE from the Bruch’s membrane and often the cells of the choroid remain attached. Additionally, cell sorting such as fluorescence activated cell sorting (FACS) is not feasible because the high content of extracellular matrix in the retina renders it impossible to break into a single-cell suspension required for sorting.

Cell culture allows for the isolation of only a single cell type and the manipulation of that cell type. The most commonly used cell line to represent the RPE is ARPE-19 cells\textsuperscript{215}. These cells exhibit most of the characteristics of native RPE, they are very homogenous, grow quickly and can be passaged indefinitely. For this reason I chose to use ARPE-19 cells. Except where otherwise noted, ARPE-19 cells were grown either in the absence of ascorbate to represent ascorbate deficiency in diabetes or with 50 $\mu$M ascorbate to represent physiological levels in healthy individuals. Although the level of ascorbate varies throughout the body and is greater in the eye than in the blood, the level of 50 $\mu$M ascorbate was chosen to represent physiological levels because ascorbate enters the RPE cells through active transport from the blood where levels are generally around 50 $\mu$M\textsuperscript{168}. 
Methods

Cell Culture

Human retinal pigment epithelial ARPE-19 cells were purchased from American Type Culture Collection (ATCC, Manassas, VA). ARPE-19 cells were maintained in DMEM-F12 medium (Life Technologies, Carlsbad, CA), which does not contain any ascorbate in its formulation. After seeding approximately $3 \times 10^5$ cells in 6-well plates or $10^6$ cells in 10 cm plates for 24 hr, cells were treated with sodium ascorbate or L-ascorbic acid (Sigma-Aldrich, St. Louis, MO) at different concentrations for varying durations. Each treatment group consisted of at least three wells or three plates for every experiment. In all experiments cell culture media was changed daily with or without ascorbate to maintain a constant concentration of ascorbate.

5hmC Dot-Blot Assay

Genomic DNA was extracted from cultured ARPE-19 cells using QIAamp DNA mini kits (Qiagen, Hilden, Germany) according to the manufacturer’s instructions. A Qubit Fluorometer (Life Technologies, Carlsbad, CA) was used to quantify the concentration of DNA. The dot-blot was performed following a previously published protocol using two spots per sample, one with 0.4 μg and the other with 0.8 μg to allow for visualization of contrast between samples with both low and high levels of 5hmC. Briefly, DNA samples were diluted with 2 N NaOH and 10 mm Tris-Cl, pH 8.5, and then loaded on Hybond N+ nylon membrane (GE Healthcare, Chicago, IL) using a 96-well dot-blot apparatus (Bio-Rad Laboratories, Hercules, CA). After being baked in 80 °C for 30 min and blocked by 5% nonfat milk for 1 hour at room temperature, the membrane was incubated in a 1:10,000 dilution of polyclonal anti-5-hmC antibody (#39769, Active Motif, Carlsbad, CA) at 4 °C overnight. 5hmC was visualized by using a Pierce ECL Western blotting substrate kit.
40

(Thermo Scientific, Waltham, MA). To ensure equal loading, the membrane was stained with methylene blue post-immunoblotting. The densities of the dots on the membrane were captured by ImageJ. Statistical significance of differences in 5hmC content between treatments and control was assessed by one-way ANOVA with a significance threshold of \( \alpha = 0.05 \).

5-hmC Immunofluorescent Staining

Immunofluorescent staining of 5hmC was performed as previously described\(^{216}\). Glass 12-mm coverslips were first acid washed in 1M HCl at 56°C for 4 hours with occasional agitation, followed by a thorough rinsing in water and wash in 100% ethanol. Cells were seeded in 24-well plates with acid-washed coverslips at 10,000 cells per well. After one day, cells were conditioned in media either with (50 \( \mu \)M) or without ascorbate for 5 days. After treatments, cells were washed once with phosphate-buffered saline (PBS, Life Technologies, Carlsbad, CA) and fixed with 4% paraformaldehyde for 20 minutes. Cells were then washed again with PBS and then incubated with 2N HCl at 37°C for 20 min and neutralized with 100 \( \mu \)M Tris-HCl for 10 min to denature the DNA, which is required to allow for the binding of the anti-5hmC antibody. After washing three times with PBS and blocking with 0.4% Triton X-100 with 10% FBS in PBS for 1 hour, cells were incubated with a 1:500 dilution of anti-5hmC antibody (#39769, Active Motif) at 4°C overnight. In the morning the coverslips containing cells were washed with PBS and incubated with a 1:10,000 dilution of Alexa Fluor 488-conjugated donkey anti-rabbit IgG. Cells were then counterstained with diamidino-2-phenylindole (DAPI).

Cells fluorescence images were acquired into a 512 \( \times \) 512 frame size by averaging 16 times at a bit depth of 8 using a Zeiss LSM 710 confocal microscope (Zeiss, Oberkochen, Germany). To ensure an unbiased selection of cells, DAPI staining was used
for focus before acquisition. Fluorescence intensity was quantified using Fiji (ImageJ)\textsuperscript{217}. Average intensity values were measured from every cell within the image field from a minimum of five $20 \times$ images per condition (Around 350 cells per condition). The intensity values from individual cells were plotted using the ggplot2 package in R and statistically analyzed by student’s t-test with a significance level of $\alpha = 0.05$.

**Bisulfite Sequencing and Tet-Assisted Bisulfite Sequencing**

ARPE-19 cells were treated with or without sodium ascorbate (50 uM) for 10 days. Genomic DNA was extracted using the QIAamp DNA mini kit (Qiagen). Spike-in controls were added to the DNA (#S001 and #S002, WiseGene, Chicago, IL) and the DNA was sonicated. Part of the sample was used for bisulfite conversion using the EZ DNA methylation Gold kit (Zymo Research, Irvine, CA) according to the manufacturer’s protocol. The remainder of the DNA was then glycosylated and reduced via the Tet1 enzyme using the 5hmC tet-assisted bisulfite sequencing (TAB-Seq) Kit, following the manufacturer’s protocols (#K001, WiseGene). DNA was then purified with Micro Bio-Spin 30 columns (Bio-Rad Laboratories) and QIAquick PCR purification Kit (Qiagen). TET-oxidized DNA was further directed to bisulfite conversion using the EZ DNA methylation Gold kit (Zymo Research) according to the manufacturer’s protocol. PCR was performed using the primers provided with the spike-in kit and submitted for sequencing at the John P. Hussman Institute of Human Genomics. Amplicons were sequenced on an Illumina HiSeq 2000 (125 bp paired-end reads; Illumina, San Diego, CA).

Bisulfite sequencing and Tet-assisted bisulfite sequencing reads were aligned to a custom fasta file containing the sequences of the spike-in controls using BSMAP, a program that converts the genome to one containing a Y in the place of each cytosine, representing either a cytosine or thymine\textsuperscript{218}. BSMAP was used to calculate the percentage methylation and hydroxymethylation for each base, and methylKit was used for statistical
analysis. Individual cytosines with an adjusted $P$ value below 0.05, a minimum of 10% difference in methylation between treatments and a minimum of a 1.5-fold change between treatments were considered “differential”.

**hMeDIP qPCR**

hMeDIP-qPCR was performed following previously published protocols. Genomic DNA was sonicated to approximately 150 base pair (bp) fragments using the Covaris S220 (Covaris, Woburn MA). Spike-in controls (5C, 5mC, 5hmC) were added to 2 μg of sonicated genomic DNA (#C02040011, Diagenode, Searing, Belgium). Ten percent was removed to save as unprecipitated input control (input). DNA was denatured at 99°C for 10 minutes and then incubated overnight with anti-5hmC antibody (#39791, Active Motif, Carlsbad, CA). The following day, antibodies and bound DNA was precipitated using Protein G magnetic beads (#161-4021, Bio-Rad, Hercules, CA). Beads were then resuspended in proteinase K buffer for 3 hours at 55°C. Input samples were incubated with proteinase K buffer alongside immunoprecipitated samples. All samples were then purified using Ampure XP beads (#A63880, Beckman Coulter, Brea, CA).

qPCR was performed on the QuantStudio 12K Flex (Applied Biosystems, Foster, CA) using PowerUp SYBR Green Master Mix (Applied Biosystems) with 10 μl reactions. Input crossing thresholds (CT) were adjusted by subtracting 3.32 cycles (to adjust for using only 10% of the sample as input), according to the methods published by Taiwo et al. for calculating percent recovery. The percent recovery for each amplicon was calculated using $2^{(\text{adjusted input CT} - \text{immunoprecipitated CT})} \times 100$. Specificity was calculated using $1 - (\% \text{ recovery } 5\text{C or } 5\text{mC spike-in} / \% \text{ recovery } 5\text{hmC spike-in})$ and was calculated to be 99.6% when using 5C as the standard and 99.1% when using 5mC as the standard. Percent recovery for 5C spike-in and 5mC spike-in amplicons were below
0.5% and below 1.1% respectively. Differences between treatments were assessed by the student’s t-test with a significance of $\alpha=0.05$.

**hMeDIP-seq**

ARPE-19 cells were cultured in confluent monolayer and treated with or without sodium ascorbate (50 $\mu$M) for 10 days. Genomic DNA was extracted from ARPE-19 cells using QIAamp DNA mini kits from Qiagen (Hilden, Germany) according to the manufacturer’s instructions. A Qubit Fluorometer from Life Technologies (Carlsbad, CA, USA) was used to quantify the concentration of DNA. A Bioanalyzer 2000 was used to measure the quality of DNA. DNA was submitted for hydroxymethylated DNA immunoprecipitation sequencing (hMeDIP-seq) to the Epigenomics Core at the University of Michigan using previously published protocols. Briefly, DNA was sonicated to approximately 100 bp, ligated with Illumina adaptors, and immunoprecipitated using anti-5hmC antibody (#39791, Active Motif, Carlsbad, CA). DNA of both immunoprecipitated and un-immunoprecipitated (input control) samples were subsequently sequenced on a HiSeq4000 sequencing system (50 bp single-end reads, 3 samples per lane; Illumina, San Diego, CA). Analysis of hMeDIP-seq was carried out using a previously reported pipeline. Briefly, after quality control, reads were aligned using the burrows-wheeler aligner and peaks were called using MACS2 then filtered using the irreproducible discovery rate (IDR) pipeline. Reads within peak regions were quantified using HTSeq-count. Differential enrichment was calculated using edgeR, and only peaks below an adjusted $P$-value (false discovery rate, FDR) of 0.05 with a minimum of a 2X fold change were considered differential to minimize false positives. Heatmaps of read density for peaks in each sample and fingerprints to represent the enrichment in each sample were generated using version 2 of DeepTools.
Results

5hmC increases in ARPE-19 cells after the addition of ascorbate

To test the effects of ascorbate on the outer BRB I first treated ARPE-19 cells for three days with media containing varying concentrations of ascorbate. Without any added ascorbate, 5hmC was barely detectable in the extracted DNA by dot blot. Addition of ascorbate at different concentrations (10, 100, 500 μM) for 3 days induced a ∼2.5 fold increase in the global content of 5hmC, as shown by the dot blot and semi-quantitation of the dot blot (Figure 2.1 A & B). The average concentration of ascorbate in the plasma of healthy humans is ∼50 μM and can reach up to ∼150 μM. Ascorbate at a high pharmacological concentration (500 μM) did not exert additional benefits in 5hmC generation. Furthermore, ascorbate increased 5hmC in a time-dependent fashion. The 5hmC content increased to approximately 1.5 fold after 1 day, ∼2.5 fold after 3 days, and ∼3 fold after 5 days as shown by the dot blot and semi-quantitation of the dot blot (Figure 2.1 C & D).

To confirm the increase in 5hmC after the addition of ascorbate to the culture medium by another method, immunofluorescent staining was used. ARPE-19 cells were treated for five days with media supplemented with (50 μM) or without ascorbate. Nuclear staining was performed using DAPI, and the blue channel was used to focus on the cells so as to eliminate the potential bias of picking the most brightly fluorescent cells. The cells treated with 50 μM ascorbate exhibited nearly 2-fold the fluorescence intensity as the cells deficient in ascorbate (Figure 2.1 E & F). Together, these results indicate that ascorbate profoundly increased the global content of 5hmC in ARPE-19 cells.
Figure 2.1. Induction of 5hmC in ARPE-19 Cells by Ascorbate Treatment. (a-b). Dot-blot and semi-quantitative analysis of the dot-blot show that ascorbate (0, 10, 100 \( \mu \text{M} \)) dose-dependently induced 5hmC in ARPE-19 cells. Ascorbate at 500 \( \mu \text{M} \) did not cause further increase in 5hmC, possibly due to its toxicity to the cells at this concentration. Methylene blue staining was used as a loading control. (c-d) Dot-blot and semi-quantitative analysis of the dot-blot show that ascorbate (50 \( \mu \text{M} \)) time-dependently (0, 1, 3, and 5 days) induced 5hmC in ARPE-19 cells. (e-f) Immunofluorescent staining and quantification confirm the ascorbate induced increase in 5hmC. Bold points represent outliers. (*** \( P < 0.001 \))
Determining a Method for Investigating the Genomic Locations of 5hmC Change

The global increase in 5hmC is not expected to be evenly distributed, but rather thought to be indicative of local changes in many small regions of the genome. Upon literature searches, I found three separate methods to profile for 5hmC throughout the genome. The first method is known as oxidative bisulfite sequencing (oxBS-seq)\textsuperscript{229}. This method uses a combination of two different types of sequencing and uses a subtraction technique to determine 5hmC levels. The first round of sequencing is traditional bisulfite sequencing. Bisulfite sequencing involves a treatment of genomic DNA with bisulfite, which converts cytosines to uracils, except where they are protected by possessing either a 5mC or 5hmC mark. Upon sequencing, unmodified cytosines (or 5fC or 5caC) are read as thymine and only the 5mC or 5hmC are read as cytosine. This method effectively allows for the determination of the levels of 5mC + 5hmC at single-base resolution throughout the genome. The second round of sequencing involves an oxidation of 5hmC to 5fC, followed by bisulfite treatment and sequencing. This second round of sequencing allows for the determination of levels of 5mC throughout the genome. A subtraction technique is applied to determine the levels of 5hmC throughout the genome separate from the 5mC. This technique sounds ideal as it allows for single-base resolution of both 5mC and 5hmC levels throughout the genome. However, it has been shown that this method is not effective for determining levels of 5hmC in cells that have relatively low levels of 5hmC and is generally regarded as giving inaccurate measurements of 5hmC\textsuperscript{230}. Due to the published inaccuracy of this method, I did not attempt this method.

A second method was available for 5hmC sequencing at single-base resolution. This method is known as Tet-assisted bisulfite sequencing (TAB-seq)\textsuperscript{231}. This method uses β-glucosyltransferase to glucosylate 5hmC specifically, and then uses a recombinant mouse Tet1 enzyme to oxidize 5mC to 5fC. After bisulfite treatment, 5hmC bases are protected but bases that were originally 5mC have been converted to 5fC and are now
converted to uracil. Upon sequencing, only 5hmC bases remain as cytosines and all others are read as thymine. This sequencing method is attractive because it allowed for sequencing the genome only one time instead of twice as was necessary for oxBS-seq. Additionally, because 5hmC is read through direct measurement this method has been found to be more sensitive and more accurate than oxBS-seq\textsuperscript{230}.

To test the accuracy of this method, I performed bisulfite sequencing and TAB-seq on spike-in standards from WiseGene, the company that produces the TAB-seq kits. These standards contain plasmid DNA that contains either 100% methylated cytosines, or highly hydroxymethylated cytosines (reported to be around 90%). These amplicons were spiked into genomic DNA of ARPE-19 samples either treated with (50 \(\mu\)M) or without ascorbate and processed for both bisulfite sequencing and TAB-seq. Primers provided with the kit were used to amplify the treated DNA, and sequencing was performed.

The 5hmC standard should read out as having no 5mC, but \(~90\%\) 5hmC at each individual cytosine. Analysis of the levels of 5mC (obtained from TAB percentage cytosine subtracted from bisulfite percentage cytosine at each location) for 5hmC spike-in standard showed very low 5mC levels, all below 4\% and no cytosines were differential between the two samples, as expected (Figure 2.2A). Additionally, the 5hmC readings of the 5hmC spike-in control were all high (with the exception of a single base) and very consistent between the two samples (Figure 2.2B). This indicates that if 5hmC is present, it will accurately be read as 5hmC by TAB-seq.

The 5mC standard should read as having 100\% 5mC, but 0\% 5hmC at all cytosines. Analysis of the levels of 5mC in this spike-in were surprisingly low, with each point (with the exception of a single base that was probably unmethylated) being between 50\% and 90\% 5mC. Additionally, one point came back as differential between the two samples, indicating that the Tet1 treatments were not equal between samples (Figure 2.2C). Analysis of the 5hmC levels showed that, although all points should have read as
0% 5hmC, many read back between 15% and 50% 5hmC (Figure 2.2D). Additionally, many points were differential between conditions, indicating that although the two samples were processed together there were differences in the Tet1 reduction between the two samples. This data indicates that where 5mC is present, it will be read as 5hmC often greater than 15% of the time, and that the Tet-assisted oxidation of 5mC may not be equal between samples. This indicates that the error rate of 5mC being read as 5hmC is higher than the absolute value of 5hmC at the majority of CpG sites throughout the genome, and this method is not accurate enough to differentiate between 5mC and 5hmC.
A third method has been published for investigating 5hmC levels genome-wide. Hydroxymethylated DNA immunoprecipitation sequencing (hMeDIP-seq) is a method that uses an antibody against 5hmC to preferentially enrich DNA that has higher 5hmC content. This method is similar to that of chromatin immunoprecipitation sequencing (ChIP-seq) and does not provide single-base resolution, but is accurate for determining relative levels of 5hmC throughout the genome\(^\text{221}\). To determine the specificity of the antibody used for hMeDIP-seq, I performed hMeDIP-qPCR on spike-in controls from Diagenode. These spike-in controls are reported to have 100% of cytosines as 5mC, 5hmC or 5C and come with primers designed for each individual standard. hMeDIP-qPCR employs the full immunoprecipitation followed by qPCR to determine the percent recovery of the individual amplicons in comparison to a portion of the sample that was saved from the immunoprecipitation process. The recovery of the 5hmC amplicon was 97%, while the recovery of the 5C and 5mC amplicons were only 0.3% and 0.91%, respectively (Figure 2.3). The specificity of the antibody to 5hmC was calculated to be 99.6% when calculated using the 5C spike-in and 99.1% when using the 5mC spike-in control. This experiment demonstrates that the antibody is highly specific and will have a strong preference to
hydroxymethylated DNA as opposed to methylated or unmodified DNA strands. Although this method does not provide single-base resolution, it will provide relative enrichment levels and can be used to accurately look for differences in 5hmC throughout the genome.

Figure 2.3. hMeDIP-qPCR of Diagenode Spike-In Controls Was Used to Validate the Antibody Used for hMeDIP-seq. To validate the specificity of the antibody used for hMeDIP-seq, hMeDIP-qPCR was used on oligos generated using only cytosines with a specific 5’ modification (5C, 5mC or 5hmC). The percent recoveries of the 5C, 5mC and 5hmC spike-in controls were 0.33%, 0.91% and 96.9%, respectively.

Determining a Method for the Analysis of hMeDIP-seq Data

Performing hMeDIP-seq was accompanied with the challenge of analyzing the data. Upon a thorough literature search, no programs were designed specifically for the analysis of hMeDIP-seq data. There was, however, a program designed for the analysis of MeDIP-seq data, a sister method to hMeDIP-seq that uses an antibody against 5mC instead of 5hmC. This program is called MEDIPS and processes data from fastq through
MeDIPS uses a sliding window approach to investigate regions throughout the genome. Differential enrichment is calculated using edgeR and a mixture between the \( P \) values obtained and fold changes between sample types are used to determine differentially methylated regions (DMRs).

The differences in distribution of 5mC and 5hmC throughout the genome, however, should not be forgotten. The majority of CpGs throughout the mammalian genome are methylated, and analysis types often focus mainly on regions that have low levels of methylation. For this reason, MEDIPS investigates all regions throughout the genome under the assumption that looking for highly enriched regions is not necessary. The final output gives DMRs, but does not give regions that are enriched but not differential.

In contrast to 5mC, 5hmC is relatively rare throughout the genome and present at only approximately 1-10% the level of 5mC. Additionally, the clusters of higher levels of 5hmC are the ones thought to change and contribute to differences between cell types as well as contribute to disease. Applying MEDIPS to hMeDIP-seq data is therefore not appropriate because it operates under the assumption that the majority of regions throughout the genome should be highly enriched, which is not true for hMeDIP-seq.

Although no gold standard program has been developed for determining highly enriched regions in hMeDIP-seq, there is a well-established method for determining highly enriched regions in ChIP-seq. The Irreproducible Discovery Rate (IDR) pipeline was developed by the ENCODE project to determine the enriched regions that are consistent between ChIP-seq samples. This method can be combined with normalization of reads and differential enrichment analysis using edgeR, similar to what is used by MEDIPS. The benefit of the IDR pipeline is that it gives both the differentially hydroxymethylated regions (DhMRs) and enriched regions that do not change between treatment types. Additionally, the IDR pipeline contains many quality control (QC) guidelines that ensure the quality of the data used.
Addition of Ascorbate to the Culture Media of ARPE-19 Cells Causes a Dramatic Shift in the Hydroxymethylome

ARPE-19 cells were cultured in media either supplemented with (50 μM) or without ascorbate and the DNA was extracted and processed for hMeDIP-seq. After alignment of raw reads with BWA, multimapped reads and PCR duplicate reads were removed, in accordance with the standards published by the ENCODE consortium\(^{233}\). All samples, both genomic input and hMeDIP had a minimum of 55 million reads, which is above the minimum of 40 million reads and close to the 60 million reads recommended for analysis of this type of data\(^{220}\). Alignment rates at each step are shown in table 2.1.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Input Reads</th>
<th>Alignment Rate</th>
<th>Percent Duplicates</th>
<th>Usable Reads</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 μM Ascorbate 1 Input</td>
<td>173,638,300</td>
<td>96.55%</td>
<td>62.16%</td>
<td>55,337,031</td>
</tr>
<tr>
<td>0 μM Ascorbate 1 hMeDIP</td>
<td>154,952,950</td>
<td>93.63%</td>
<td>40.56%</td>
<td>72,273,855</td>
</tr>
<tr>
<td>0 μM Ascorbate 2 Input</td>
<td>151,501,443</td>
<td>97.79%</td>
<td>25.86%</td>
<td>95,838,557</td>
</tr>
<tr>
<td>0 μM Ascorbate 2 hMeDIP</td>
<td>141,372,024</td>
<td>96.70%</td>
<td>23.19%</td>
<td>86,477,191</td>
</tr>
<tr>
<td>0 μM Ascorbate 3 Input</td>
<td>108,806,520</td>
<td>97.30%</td>
<td>33.67%</td>
<td>61,261,221</td>
</tr>
<tr>
<td>0 μM Ascorbate 3 hMeDIP</td>
<td>309,765,108</td>
<td>93.57%</td>
<td>67.87%</td>
<td>78,258,599</td>
</tr>
<tr>
<td>50 μM Ascorbate 1 Input</td>
<td>116,473,102</td>
<td>97.85%</td>
<td>22.33%</td>
<td>77,271,453</td>
</tr>
<tr>
<td>50 μM Ascorbate 1 hMeDIP</td>
<td>129,116,794</td>
<td>96.79%</td>
<td>23.86%</td>
<td>79,383,156</td>
</tr>
<tr>
<td>50 μM Ascorbate 2 Input</td>
<td>219,938,732</td>
<td>97.43%</td>
<td>61.98%</td>
<td>71,158,211</td>
</tr>
<tr>
<td>50 μM Ascorbate 2 hMeDIP</td>
<td>108,494,772</td>
<td>95.00%</td>
<td>34.12%</td>
<td>58,628,438</td>
</tr>
<tr>
<td>50 μM Ascorbate 3 Input</td>
<td>103,311,170</td>
<td>97.76%</td>
<td>23.57%</td>
<td>67,427,002</td>
</tr>
<tr>
<td>50 μM Ascorbate 3 hMeDIP</td>
<td>134,857,868</td>
<td>96.64%</td>
<td>42.07%</td>
<td>67,402,206</td>
</tr>
</tbody>
</table>
Table 2.1. Alignment Rates for hMeDIP and Input Control Samples. The percent alignment and percent duplicates for each sample are shown. All samples have greater than the 40 million reads generally suggested for hMeDIP-seq.

The next step of the IDR pipeline is peak calling using MACS2\textsuperscript{224}. Immunoprecipitation based sequencing experiments lead to reads beginning at a peak location and moving away from the center, meaning that reads that align to the forward sense tend to be shifted downstream and reads that align to the reverse sense tend to be shifted upstream. This is referred to as a strand-shift\textsuperscript{233}. MACS2, first calculates the strand shift and generates two figures. One of these figures is referred to as a “Peak Model” and shows the density of reads on the forward and reverse strand in relation to the center of peaks. When clear peaks are present (indicating enrichment in peaks), two hills separated by clear valley are present. The second figure is called a cross-correlation plot, and shows the strand shift relative to the overlap of reads to the center of the peak. When enrichment is present, this should look like a sine wave. Two of the samples treated with media containing no ascorbate produced figures that matched the description of the ideal peaks, but sample 0 \( \mu \text{M} \) #2 produced abnormal plots for both figures (Figure 2.4 A and B). Similarly, of the samples treated with 50 \( \mu \text{M} \) of ascorbate, two of the samples produced the expected figures while 50 \( \mu \text{M} \) #1 produced abnormal plots for both figures (Figure 2.4 A and B). For comparison, I ran MACS2 using the genomic input of sample 0 \( \mu \text{M} \) #1 to see what the graphs look like with a sample that has no enrichment. The peak model and cross-correlation plots from the genomic input sample looked similar to the figures generated for 0 \( \mu \text{M} \) #2 and 50 \( \mu \text{M} \) #1 (Figure 2.4 C and D). This indicates that there may have been little or no enrichment in these two samples.
Figure 2.4. Peak Models and Cross-Correlation Plots from MACS2 for Each Sample. The peak calling software MACS2 generates plots showing the distance from reads aligned to the forward sense and the reverse sense to the center of peaks and the correlation between the forward and minus tags. The distance between greatest correlation is considered the strand shift, also noted by MACS2 as alternative strand lag (alt lags). Peak models (a) and cross-correlation plots (b) for each sample are shown. Peak model (c) and cross-correlation plot (d) from a genomic input file with no immunoprecipitation match the shapes of the peaks in 0 \( \mu \text{M} \) 2 and 50 \( \mu \text{M} \) 1.

The next step of the IDR pipeline involves comparing each sample to the other samples within the treatment type and splitting each sample into subsets that are compared. First, the reads from each sample are shuffled and split into two equal size files. These are referred to as “pseudoreplicates”. Finally, reads from all samples are merged to make a pool, which is also split into two pseudoreplicate samples. Peaks are called on these pseudoreplicates just as they are called on the parent samples. After all samples (including pseudoreplicates) have had peak calling performed, the peaks are ranked from strongest peak to weakest peak by \( P \) value. These ranked peak files will then be compared between samples and between pseudoreplicates. Specifically, for a group of three samples as I have here, 7 comparisons are made. The first three are the individual sample comparisons: Sample 1 vs Sample 2, Sample 1 vs Sample 3, and Sample 2 vs Sample 3. The next three comparisons are the pseudoreplicate comparisons: Sample 1 Pseudoreplicate 1 vs Sample 1 Pseudoreplicate 2, Sample 2 Pseudoreplicate 1 vs Sample 2 Pseudoreplicate 2, and Sample 3 Pseudoreplicate 1 vs Sample 3 Pseudoreplicate 2. The final comparison is one of the pooled sample pseudoreplicates, Pooled Sample Pseudoreplicate 1 vs Pooled Sample Pseudoreplicate 2.

Comparisons are made using the files from one given condition (i.e. 0 \( \mu \text{M} \) peak files are compared only to other 0 \( \mu \text{M} \) peak files and not to 50 \( \mu \text{M} \) peak files). In these comparisons, the peaks are evaluated for similarity in strength using the scripts provided with the IDR pipeline. Each peak is given an IDR value, which is analogous to an adjusted
$P$ value or false discovery rate (FDR). The threshold used for individual sample comparisons and pseudoreplicate comparisons is 0.01, while the threshold for pooled sample pseudoreplicate comparisons is 0.001. The count of the peaks below a given IDR threshold should be within two-fold of all other comparisons. For instance, if 10,000 peaks are below 0.01 for the comparison Sample 1 vs Sample 2, then the number of peaks with an IDR below 0.01 for the comparison Sample 1 vs Sample 3 should be between 5,000 and 20,000. If the number of peaks below threshold is not within a two-fold range it is indicative that immunoprecipitation strength is not similar between all samples.

For the 0 $\mu$M samples, the Pooled Pseudoreplicate comparisons yielded 89,907 peaks below 0.001. The comparison of Sample 1 vs Sample 3 yielded 81,883 peaks below an IDR of 0.01. The pseudoreplicate comparisons of Sample 1 and Sample 3 were 59,755 and 59,882. All of these comparisons are within the two-fold threshold. However, the pseudoreplicate comparisons of Sample 2 yielded only 2,636 peaks below 0.01 and the individual comparisons of Sample 1 vs Sample 2 and Sample 2 vs Sample 3 yielded 5,966 and 5,387 peaks, respectively. All of the comparisons involving sample 2 gave very low numbers of peaks passing threshold, indicating that sample 0 $\mu$M 2 failed the immunoprecipitation and should be removed from the analysis. Samples 0 $\mu$M 1 and 0 $\mu$M 3 were then pooled and split into pseudoreplicates. The comparison between the new pseudoreplicate gave 92,593 peaks below the threshold of 0.001, which is within 2-fold of the other comparisons (Table 2.2).

For the 50 $\mu$M samples, the Pooled Pseudoreplicate comparisons yielded 115,147 peaks below 0.001. The comparison of Sample 2 vs Sample 3 yielded 72,084 peaks below an IDR of 0.01. The pseudoreplicate comparisons of Sample 2 and Sample 3 were 61,508 and 120,213. All of these comparisons are within the two-fold threshold. However, the pseudoreplicate comparisons of Sample 1 yielded only 7,591 peaks below 0.01 and the
individual comparisons of Sample 1 vs Sample 2 and Sample 1 vs Sample 3 yielded 29,982 and 8,199 peaks, respectively. All of these comparisons involving sample 1 gave very low numbers of peaks passing threshold, indicating that sample 50 \( \mu \text{M} \) 1 failed the immunoprecipitation and should be removed from the analysis, just like sample 0 \( \mu \text{M} \) 2. Samples 50 \( \mu \text{M} \) 2 and 50 \( \mu \text{M} \) 3 were then pooled and split into pseudoreplicates. The comparison between the new pseudoreplicate gave 118,898 peaks below the threshold of 0.001, which is within 2-fold of the other comparisons between samples that passed threshold (Table 2.2).

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Peaks below IDR threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0 ( \mu \text{M} )</td>
</tr>
<tr>
<td>Sample 1 VS Sample 2</td>
<td>5,966</td>
</tr>
<tr>
<td>Sample 1 VS Sample 3</td>
<td>81,883</td>
</tr>
<tr>
<td>Sample 2 VS Sample 3</td>
<td>5,387</td>
</tr>
<tr>
<td>Sample 1 Pseudoreplicates</td>
<td>59,755</td>
</tr>
<tr>
<td>Sample 2 Pseudoreplicates</td>
<td>2,636</td>
</tr>
<tr>
<td>Sample 3 Pseudoreplicates</td>
<td>59,882</td>
</tr>
<tr>
<td>All samples pooled pseudoreplicates</td>
<td>89,907</td>
</tr>
<tr>
<td>Samples passing QC pooled pseudoreplicates</td>
<td>92,593</td>
</tr>
</tbody>
</table>

Table 2.2. Number of Peaks below the Irreproducible Discovery Rate (IDR) Threshold. The threshold for pseudoreplicate comparison and for cross sample comparison is recommended threshold of 0.01. The threshold used for pooled sample pseudoreplicates is the recommended threshold of 0.001. The number of peaks passing threshold from all comparisons should be within a 2-fold change of the individual sample comparisons. Note that samples 0 \( \mu \text{M} \) sample 2 and 50 \( \mu \text{M} \) sample 1 fail the pseudoreplicate comparison and fail both cross-sample comparisons. Additionally, the pooled sample pseudoreplicates have greater numbers of peaks below the IDR threshold when the failed samples are removed, despite lower read counts.
The IDR program generates graphs that show the number of peaks below a given IDR threshold, where the IDR threshold is on the Y-axis and the number of peaks below a given threshold for the comparison is on the X-axis. It is clear in the comparisons that one individual comparison and two pseudoreplicate comparisons clearly have far more significant peaks at any given threshold than the comparisons involving the samples that failed immunoprecipitation (Figure 2.5 A-D).

As a final measure to confirm a difference in immunoprecipitation of the samples, I generated a plot known as a fingerprint plot using DeepTools2227. A fingerprint plot is an useful method to represent the strength of enrichment in any given sample. The genome is divided into equal size bins, and the reads within each bin are counted. The bins are then ranked from highest number of reads to lowest number of reads, and the number of reads within the bins is charted. The X-axis is the rank of the bin, and the Y-axis is the fraction of the total reads within that bin and all lower ranking bins. In theory, a sample with no immunoprecipitation should yield a perfectly diagonal line from the coordinates 0,0 to 1,1. A sample with enrichment will generate a curve below and to the right of this line. The stronger the enrichment within certain peak regions, the farther the line will be from the center of the graph. As a control, I generated the fingerprint plot using all of the genomic input samples as well as the hMeDIP-seq samples. The genomic input samples clearly cluster together and make a line very close to that expected for samples with no immunoprecipitation. The samples that passed all QC thresholds all cluster below and to the right of the genomic input samples, indicating enrichment. The two samples that failed QC cluster together and are closer to the genomic input samples than to the other immunoprecipitated samples, further validating the QC of the IDR pipeline that indicates that these samples failed IP (Figure 2.5 E).
After all peak calling was performed, peaks from both treatment types were merged using bedtools (https://bedtools.readthedocs.io/en/latest/). After merging peaks between treatment types, 132,956 regions of the genome were found enriched for 5hmC. Reads within peaks were counted using HTSeq\textsuperscript{225}. Read counts were normalized and differential enrichment was calculated using edgeR\textsuperscript{226}. Peaks were considered differential if they were given an FDR below 0.05 by edgeR and showed at least a 2-fold increase or decrease.

Using this method, 56,487 peaks were found to have significantly higher enrichment of 5hmC and only 7,591 peaks were found to have significantly lower enrichment of 5hmC in samples treated with ascorbate. The remaining 68,878 regions were found to have similar enrichment across groups. These results confirmed that ascorbate treatment indeed increases 5hmC in the genome of ARPE-19 cells, consistent with the results of the dot-blot assays. However, it is interesting to note that over half of the peaks have similar enrichment across samples and a small portion of peaks actually decrease in strength after treatment with 50 \( \mu \)M ascorbate. A ChIP-seq style heatmap was generated using DeepTools2\textsuperscript{227}. Each column represents a separate sample and each row represents a separate peak. Darker colors represent a greater read density (in read counts per million mapped reads, RCPM) (Figure 2.6A). Profiles were also generated,
where each line represents the average RCPM at a given location for that given sample (Figure 2.6B).

Figure 2.6. hMeDIP-seq Results of ARPE-19 Cells Treated with either 0 µM or 50 µM Ascorbate. (a) Heatmap of the 5hmC enriched regions investigated for differential analysis. Only regions below an FDR below 0.05 by edgeR and with a minimum of a 2X (or maximum of 0.5X) fold change were considered differential. Darker color represents greater coverage in read counts per million (RCPM). (b) Overall coverage profile plots show the average coverage in RCPM across all peaks. Blue lines represent samples treated with 0 µM ascorbate and green lines represent samples treated with 50 µM ascorbate.

Conclusions

ARPE-19 cells were used to represent the RPE of the outer BRB. These cells can be passaged and give a fairly accurate representation of native RPE. Cells were cultured
without ascorbate to represent the diabetic condition, or with physiological levels of ascorbate (50 μM) to represent repletion of ascorbate. Addition of ascorbate to the culture media increased global 5hmC in ARPE-19 cells, consistent with the known role of ascorbate as a cofactor for the TET enzymes. This increase in 5hmC was not further amplified by using pharmacological doses of ascorbate. This indicates that pharmacological intravenous injection of vitamin C will not likely increase 5hmC levels beyond what is reached in healthy individuals.

Three methods were investigated for determining the locations of 5hmC change throughout the genome. Two of the methods provide single-base resolution readouts of 5hmC, but these methods were both shown to be unreliable and inaccurate. While hMeDIP-seq does not give a single-base resolution readout, it was found to be highly specific to finding changes in 5hmC without interference from 5mC (like the other methods) and gives relative representation of 5hmC levels throughout the genome. The IDR pipeline developed by the ENCODE consortium was used for peak calling to find regions highly enriched in 5hmC.

Throughout the genome, 132,956 regions were found enriched for 5hmC. A very large portion of these regions showed greater enrichment in the samples treated with 50 μM ascorbate than in samples treated without ascorbate. This corroborates the dot blot and immunofluorescent staining data that indicates a global increase in 5hmC. However, a small proportion of peaks were also found to significantly decrease in the samples treated with 50 μM ascorbate. This is consistent with previous findings that knockout of TET enzymes actually increases 5hmC in some regions of the genome. This is likely because the TET enzymes are required to further oxidize 5hmC to 5fC and 5caC, and at some locations the oxidation is stopped at 5hmC without functional TET enzymes. Additionally, using the IDR pipeline it was determined that approximately half of the
regions enriched in 5hmC do not change with the addition of ascorbate to the culture medium, despite the global increase. This indicates that only some genomic regions will be affected by the repletion of ascorbate in the diabetic eye, which is likely to be important in regulation of transcription. The genes whose transcription is dependent upon 5hmC may become dysregulated in the ascorbate deficiency that occurs in hyperglycemia, and these genes may be contributing to the development and progression of DR.
Chapter 3

The Transcriptome of Retinal Cells Shifts after Treatment with Ascorbate

Some of the data presented in this chapter were previously published in *Investigative Ophthalmology and Visual Sciences (IOVS, Volume 59, Issue 8)*\(^2\)\(^\text{14}\). I performed all of the work presented in this chapter with two exceptions. The cell culture and RNA extraction for sequencing were performed by Christopher Gustafson. Whole transcriptome sequencing was performed by the Sequencing Core of the John P. Hussman Institute for Human Genomics at the University of Miami.

Background

The vitamin C deficiency in the eyes of diabetics was modeled using ARPE-19 cells, and addition of physiological levels of ascorbate to the media was used to represent restoring ascorbate to healthy levels. Upon the addition of ascorbate, there was a global increase in 5hmC in the DNA of ARPE-19 cells as seen by both dot blot and immunofluorescent staining. This global increase is indicative of a shift in the dynamic process of methylation/demethylation, as was expected by a repletion of ascorbate. Using hMeDIP-seq, I found that there were many regions throughout the genome that changed in relative level of 5hmC. Dysregulations in the hydroxymethylome have been associated with many diseases and this has largely been attributed to the fact that changing 5hmC patterns can shift transcription patterns\(^{171,\ 191}\).

Given the dramatic shift in the hydroxymethylome of RPE cells after treatment with ascorbate, I hypothesized that there would also be a dramatic shift in the transcriptome of the cells after treatment with ascorbate. To test this hypothesis I used the same cellular
model as was used in Chapter 2 and performed whole transcriptome sequencing (RNA-seq) analysis on RNA extracted ARPE-19 cells treated with media containing either no ascorbate or 50 μM ascorbate.

**Methods**

**Cell Culture**

Human retinal pigment epithelial ARPE-19 cells were purchased from American Type Culture Collection (ATCC, Manassas, VA). ARPE-19 cells were maintained in DMEM-F12 medium (Life Technologies, Carlsbad, CA), which does not contain any ascorbate in its formulation. After seeding approximately $3 \times 10^5$ cells in 6-well plates or $10^6$ cells in 10 cm plates for 24 hours, cells were treated either with 50 μM sodium ascorbate (Sigma-Aldrich) or without added ascorbate to the media. Each treatment group consisted of three wells. Cell culture media was changed daily with or without ascorbate to maintain a constant concentration of ascorbate.

**Whole transcriptome sequencing**

ARPE-19 cells were cultured into a confluent monolayer and treated with or without sodium ascorbate (50 μM) for 7 days. Total RNA was extracted from the cells using the RNeasy Mini Kit (Qiagen). A Bioanalyzer 2000 was used to measure the quality of RNA. All samples’ RNA integrity numbers (RIN) were above 9. Whole transcriptome sequencing (RNA-seq) was carried out at the Sequencing Core of the John P. Hussman Institute of Human Genomics at the University of Miami using the Epicentre Ribo-Zero Human/Mouse/Rat kit (Epicentre, Madison, WI). Briefly, after ribosomal RNA (rRNA) was depleted, sequencing libraries were constructed following the standard Illumina protocols.
and were subsequently processed by a HiSeq2000 sequencing system (125 bp paired-end reads, 4 samples per lane; Illumina, San Diego, CA).

Data was analyzed using a previously published pipeline with a few minor additions. Briefly, after quality control, reads were trimmed using trim_galore to remove Illumina adapters and bases below a sequencing quality of 30. Reads were then aligned to the human transcriptome (GRCh38, Ensembl.org) using STAR. Uniquely aligned reads within features were quantified using HTSeq. All samples had well above the 10 million reads recommended for differential analysis (Table 3.1).

<table>
<thead>
<tr>
<th>Sample</th>
<th>Input Reads</th>
<th>Alignment Rate</th>
<th>Multi-Mapping Rate</th>
<th>Reads within annotated features</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 μM Ascorbate 1</td>
<td>40,010,852</td>
<td>98.61%</td>
<td>8.28%</td>
<td>31,871,167</td>
</tr>
<tr>
<td>0 μM Ascorbate 2</td>
<td>36,640,725</td>
<td>98.19%</td>
<td>9.94%</td>
<td>26,698,386</td>
</tr>
<tr>
<td>0 μM Ascorbate 3</td>
<td>52,104,935</td>
<td>96.28%</td>
<td>15.65%</td>
<td>22,517,195</td>
</tr>
<tr>
<td>50 μM Ascorbate 1</td>
<td>57,303,170</td>
<td>98.16%</td>
<td>11.35%</td>
<td>43,311,678</td>
</tr>
<tr>
<td>50 μM Ascorbate 2</td>
<td>61,975,463</td>
<td>98.37%</td>
<td>9.47%</td>
<td>46,311,630</td>
</tr>
<tr>
<td>50 μM Ascorbate 3</td>
<td>62,594,734</td>
<td>97.57%</td>
<td>11.61%</td>
<td>47,242,193</td>
</tr>
</tbody>
</table>

Table 3.1. Alignment Rates for RNA-seq Samples. The percent alignment, percent of reads that were multi-mapping and read counts within annotated features for each sample are shown. All samples have greater than the 10 million reads within annotated regions, which is the general suggested threshold for differential analysis.

Reads were normalized to fragments per kilobase per million mapped reads (FPKM) using CuffQuant and CuffNorm of the Tuxedo Suite package. Transcript FPKMs were charted on a density plot where the log10 of the FPKM was charted on the X-axis and the proportion of the reads at that expression level were charted on the Y-axis.
This method, introduced in the Tuxedo Suite package is used to determine a cutoff of transcripts that have enough sequencing coverage to reliably use for differential analysis. Each sample generates two hills separated by a valley. The hill on the left is the transcript with too low of coverage to reliably use while the hill on the right is the transcripts with high enough coverage to use for differential analysis. The base of the valley is used as the dividing threshold between the highly expressed and lowly expressed transcripts and effectively serves as a limit of detection. By this method, the threshold for my transcripts to be considered highly expressed was 1.58 FPKM, leaving 12,776 transcripts to be investigated (Figure 3.1 A). All samples were then clustered using hierarchal clustering and Euclidean distances with the hclust function in R. All samples clustered within the appropriate group, indicating lower variability between samples than between groups (Figure 3.1 B).

Figure 3.1. Quality Control for RNA-seq on ARPE-19 Cells Treated with either 0 μM or 50 μM Ascorbate. (a) Density plot showing the distribution of expression across transcripts. The dashed line shows the division between the “expressed” and “not expressed” transcripts. (b) Histogram showing the clustering of RNA-seq samples using all genes considered “expressed” above 1.58 FPKM.
Differential expression was calculated using three separate packages. Two of the packages, DESeq2 and edgeR, are both run in R using the raw read counts within transcripts\textsuperscript{226, 237}. CuffDiff calculates differential expression per transcript using binary alignment map (BAM) files\textsuperscript{236}. Genes below an FDR of 0.05 by both edgeR and DESeq2 were considered “differential” (See Chapter 3 Results section for reasoning of the exclusion of CuffDiff). Heatmaps were generated using Z-scores per transcript to show differential gene relative expression. Scatter plots named volcano plots were used to demonstrate differential genes. Each transcript is plotted with the log\textsubscript{2} fold change between conditions on the X-axis and the –log\textsubscript{10} of the P value obtained from edgeR on the Y-axis. Differential transcripts are plotted as red points and nondifferential transcripts are plotted as black points.

\textit{Investigating 5hmC levels across specific transcripts}

Transcripts above 0.1 FPKM were divided into quartiles of their expression levels per treatment. Reads below 0.1 FPKM were classified as “very low” expression and transcripts without a single read were considered to have “no expression”. Binary coverage files (BigWig files) were generated from the hMeDIP-seq files from Chapter 2 using the software from the University of California Santa Cruz (UCSC) Genome Browser software\textsuperscript{238}. BigWig files from hMeDIP-seq were used to plot the relative 5hmC coverage across the promoter, transcript start site (TSS) and gene body region of genes in the form of a profile plot using DeepTools\textsuperscript{227}. Solid lines represent the average and shaded regions around the solid lines represent the standard error of 5hmC coverage across all transcripts within the select group. The coverage over the gene body was condensed into smaller fragments to allow all gene body lengths to be represented as 10,000 bp next to the 3,000 bp upstream that was charted as the promoter. The X-axis represents the position relative to the TSS and transcript end site (TES) and the Y-axis represents the
coverage in read counts per million (RCPM). 0 μM and 50 μM samples were mapped separately due to differences in ranking of transcripts.

A second set of profile plots was generated using the hMeDIP-seq data, but this time the separation of transcripts and samples was different. The 0 μM and 50 μM samples were mapped together in each graph, but a separate graph was used to represent each group of transcripts. Transcripts were divided into groups based upon their change upon addition of 50 μM ascorbate to the media. Genes that did not change in expression were classified as “nondifferential transcripts”, while genes that increased in transcription were classified as “up-regulated transcripts” and genes that decreased in transcription were classified as “down-regulated”.

Pathway Analysis

DhMRs from Chapter 2 were annotated using Region_Analysis in the diffReps package\textsuperscript{239}. Differential genes were filtered for only differential genes that contained at least one DhMR. The list of differential genes containing DhMRs was then run through pathway analysis using EnrichR and Gene Ontology Biological Process was queried\textsuperscript{240-242}. Despite the large number of differential genes, only 18 pathways were present with Bonferroni corrected $P$ values below 0.05. A single pathway, eye morphogenesis, was directly related to the eye.

Results

Addition of Ascorbate to the Culture Media of ARPE-19 Cells Causes a Dramatic Shift in the Transcriptome

ARPE-19 cells were cultured in media either supplemented with (50 μM) or without ascorbate and the RNA was extracted and processed for RNA-seq. Raw reads were
aligned and quantified, and differential expression analysis was performed using edgeR, CuffDiff and DESeq2\textsuperscript{226, 236, 237}. Only genes below an FDR of 0.05 by all three algorithms were taken to be differential. This pipeline had previously been used by our lab for multiple studies\textsuperscript{216, 243, 244}. This yielded 1,702 differential transcripts between the two conditions. However, a large number of transcripts were called as differential by only two of the three programs, indicating that this pipeline may be too stringent. Upon further investigation, 1,484 transcripts were called as differential by both edgeR and DESeq2, but not by CuffDiff, indicating that CuffDiff may give many false negatives and should not be included in this analysis. A proportional venn diagram shows that nearly as many transcripts were called as differential by the two remaining programs as were called by all three (Figure 3.2 A). A literature search revealed that comparisons of these three programs have previously been performed, and CuffDiff was found to give a high number of false negatives. Specifically, one group analyzed a large number of samples from the HapMap project. Performing differential expression between male and female donors, the genes on the Y-chromosome were picked out as differentially expressed between the two groups by both edgeR and DESeq2. CuffDiff failed to call any of the genes on the Y-chromosome as differential, indicating that it is too stringent and gives a large number of false negatives\textsuperscript{246}.

After removing CuffDiff from the pipeline, 3,186 genes were called as differential between the two conditions, as found by an FDR below 0.05 by both edgeR and DESeq2 (Figure 3.2 B). Of these differential genes, 2,222 genes increase in expression and 964 decrease in expression when ARPE-19 cells are treated with 50 $\mu$M ascorbate (Figure 3.2 C and D). This is a monumental shift in the transcriptome, but is consistent with the findings that nearly half of the regions of 5hmC enrichment throughout the genome also change after the addition of ascorbate.
Figure 3.2. The Transcriptome of ARPE-19 Cells Shifts after Treatment with Ascorbate. (a) Proportional Venn diagram showing the number of genes called as differential by CuffDiff, DESeq2 and edgeR. (b) Venn diagram showing the number of genes called as differential by DESeq2 and edgeR. (c) Heatmap showing differentially expressed genes (FDR < 0.05 by both edgeR and DESeq2). Colors represent Z-scores where downregulated transcripts are represented as blue and upregulated transcripts are represented as red. (d) Volcano plot representing the distribution of transcript fold changes and P values. Red points represent differential genes and black points represent nondifferential genes.

5hmC is Higher in the Transcripts with Higher Expression

It has previously been reported that transcription rates are positively correlated with 5hmC level in the gene body and the promoter\cite{191}. To investigate our data, the genes with expression above 0.1 FPKM were divided into quartiles and the hMeDIP-seq
coverage for the given genes was calculated. Consistent with previous reports, the highest quartile genes averaged the highest enrichment in the hMeDIP-seq, and each quartile down had lower enrichment in both the promoter and gene body. The enrichment at the TSS, however, showed the inverse pattern where the genes with high transcription had the lowest 5hmC enrichment. This is visible in the samples receiving 0 μM ascorbate, but much more accentuated in the samples treated with 50 μM ascorbate (Figure 3.3 A). Also consistent with previous reports, the transcripts with no expression or very low expression exhibited very low levels of 5hmC from promoter through TES\textsuperscript{191}.

Investigation of transcripts that were highly expressed (above 1.58 FPKM) but do not change after addition of ascorbate, it was clear that the 5hmC levels are much higher in both the promoter and gene body of the samples treated with 50 μM ascorbate. This is not surprising considering the global increase in 5hmC after addition of ascorbate to the culture media. However, the enrichment levels at the TSS were similar across treatments. Down-regulated transcripts showed a very similar pattern of 5hmC change between the two conditions, where there was a large increase in 5hmC in the promoter and gene body but not at the TSS. The up-regulated transcripts, however, showed a much greater increase in 5hmC at the promoter, TSS and gene body than was seen in the other transcripts (Figure 3.3 B). Higher sample numbers will be required to investigate an association between the increase in 5hmC and an increase in expression through a repeated measures ANOVA, but it visually appears as though there is a greater increase in 5hmC in the up-regulated transcripts than in the nondifferential or down-regulated transcripts.
Pathway Analysis Reveals Dysregulation of Genes Related to Eye Morphogenesis

The shift in the transcriptome is dramatic. To determine the likely biological effects, I turned to pathway analysis. EnrichR is a reputable online pathway analysis software that performs a proportion test (Fisher’s Exact Test) on a list of “interesting” genes supplied by the user in comparison to an arbitrary list of the same number of genes\textsuperscript{240, 241}. To investigate only differential transcripts likely affected by changes in 5hmC,
only differential genes containing a DhMR were used for pathway analysis, leaving 2,208 genes. Of the 4,311 pathways included in the Gene Ontology Biological Process database, only 18 pathways had a Bonferroni corrected $P$ value below 0.05, indicating that the differential genes containing a DhMR tend to be clustered in these pathways (Table 3.2)\textsuperscript{242}.

Multiple pathways in this list are of interest to diabetic eye complications. Extracellular Matrix Organization is the top pathway and is interesting because there is a change in the extracellular matrix proteins in diabetes, and this is thought to contribute to the death of pericytes and endothelial cells in the inner retina\textsuperscript{47, 48}. Although these extracellular matrix changes are predominantly in the inner retina, it is possible that there are also changes in the extracellular matrix of the outer retina. Some of the other pathways are likely related to this pathway, including regulation of cell migration and regulation of smooth muscle cell migration. Another interesting pathway is the fifth ranked pathway, Regulation of Apoptotic Process. It has been shown that endothelial cells in the inner retina undergo apoptosis in diabetes and this is thought to contribute to the breakdown of the inner BRB\textsuperscript{46}. Multiple other pathways listed are related to this including positive regulation of apoptotic process and intrinsic apoptotic signaling pathway in response to DNA damage. Although these changes are known to occur in the inner BRB, this process too may happen in the outer BRB.

Perhaps the most interesting pathway that is significantly enriched in our genes in the sixth ranked pathway, eye morphogenesis. Development of the eye is heavily dependent upon the hormones released from the RPE cell layer\textsuperscript{133, 134}. Furthermore, these hormones have been shown to be necessary for the maintenance of visual function of the eye and are not only important for development\textsuperscript{136-138}. This pathway is particularly interesting because it is the only pathway listed that is known to be directly related to the RPE cell layer. The 20 genes of this pathway are listed in Table 3.3.
<table>
<thead>
<tr>
<th>Rank</th>
<th>Pathway</th>
<th>Accession</th>
<th>Genes in Pathway</th>
<th>Differential Genes</th>
<th>Bonferroni P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>extracellular matrix organization</td>
<td>GO:0030198</td>
<td>230</td>
<td>62</td>
<td>5.97×10⁻⁸</td>
</tr>
<tr>
<td>2</td>
<td>regulation of GTPase activity</td>
<td>GO:0043087</td>
<td>189</td>
<td>48</td>
<td>9.45E-05</td>
</tr>
<tr>
<td>3</td>
<td>regulation of cell migration</td>
<td>GO:0030334</td>
<td>317</td>
<td>67</td>
<td>0.000500</td>
</tr>
<tr>
<td>4</td>
<td>cytoskeleton organization</td>
<td>GO:0007010</td>
<td>127</td>
<td>35</td>
<td>0.000889</td>
</tr>
<tr>
<td>5</td>
<td>regulation of apoptotic process</td>
<td>GO:0042981</td>
<td>816</td>
<td>137</td>
<td>0.001284</td>
</tr>
<tr>
<td>6</td>
<td>eye morphogenesis</td>
<td>GO:0048592</td>
<td>20</td>
<td>12</td>
<td>0.001549</td>
</tr>
<tr>
<td>7</td>
<td>protein localization to membrane</td>
<td>GO:0072657</td>
<td>161</td>
<td>40</td>
<td>0.002546</td>
</tr>
<tr>
<td>8</td>
<td>phosphorylation</td>
<td>GO:0016310</td>
<td>387</td>
<td>75</td>
<td>0.003454</td>
</tr>
<tr>
<td>9</td>
<td>positive regulation of apoptotic process</td>
<td>GO:0043065</td>
<td>308</td>
<td>63</td>
<td>0.003919</td>
</tr>
<tr>
<td>10</td>
<td>regulation of smooth muscle cell migration</td>
<td>GO:0014910</td>
<td>19</td>
<td>11</td>
<td>0.004056</td>
</tr>
<tr>
<td>11</td>
<td>regulation of type I interferon production</td>
<td>GO:0032479</td>
<td>86</td>
<td>26</td>
<td>0.004675</td>
</tr>
<tr>
<td>12</td>
<td>regulation of alcohol biosynthetic process</td>
<td>GO:1902930</td>
<td>35</td>
<td>15</td>
<td>0.006816</td>
</tr>
<tr>
<td>13</td>
<td>regulation of small GTPase mediated signal transduction protein phosphorylation</td>
<td>GO:0051056</td>
<td>141</td>
<td>35</td>
<td>0.012863</td>
</tr>
<tr>
<td>14</td>
<td>negative regulation of smooth muscle cell migration</td>
<td>GO:0006468</td>
<td>471</td>
<td>85</td>
<td>0.013958</td>
</tr>
<tr>
<td>15</td>
<td>positive regulation of cell migration</td>
<td>GO:0014912</td>
<td>15</td>
<td>9</td>
<td>0.027767</td>
</tr>
<tr>
<td>16</td>
<td>protein localization to cell periphery</td>
<td>GO:0030335</td>
<td>222</td>
<td>47</td>
<td>0.035263</td>
</tr>
<tr>
<td>17</td>
<td>intrinsic apoptotic signaling pathway in response to DNA damage</td>
<td>GO:1990778</td>
<td>118</td>
<td>30</td>
<td>0.038357</td>
</tr>
<tr>
<td>18</td>
<td>protein localization to cell periphery</td>
<td>GO:0008630</td>
<td>49</td>
<td>17</td>
<td>0.044023</td>
</tr>
</tbody>
</table>

Table 3.2. Pathway Analysis of Differential Genes that Contain a DhMR from the RNA-seq Data from ARPE-19 Cells using the Gene Ontology Database. Pathway analysis was conducted using Enrichr and the list of all differential genes that contain DhMRs in ARPE-19 cells after the addition of 50 μM ascorbate to the culture media. All results below a Bonferroni corrected P value of 0.05 are listed.
<table>
<thead>
<tr>
<th>Gene Symbol</th>
<th>FPKM Control</th>
<th>FPKM Vitamin C</th>
<th>Fold Change</th>
<th>FDR DESeq2</th>
<th>FDR edgeR</th>
<th>Direction</th>
<th>DhMRs</th>
</tr>
</thead>
<tbody>
<tr>
<td>VEGFA</td>
<td>458.63</td>
<td>47.32</td>
<td>0.10</td>
<td>8.89×10^{-150}</td>
<td>2.90×10^{-71}</td>
<td>Down</td>
<td>2</td>
</tr>
<tr>
<td>COL5A1</td>
<td>14.14</td>
<td>8.46</td>
<td>0.60</td>
<td>6.23×10^{-8}</td>
<td>3.91×10^{-5}</td>
<td>Down</td>
<td>1</td>
</tr>
<tr>
<td>EFEMP1</td>
<td>825.75</td>
<td>571.26</td>
<td>0.69</td>
<td>1.12×10^{-6}</td>
<td>0.049</td>
<td>Down</td>
<td>7</td>
</tr>
<tr>
<td>NKD1</td>
<td>0.007</td>
<td>0.006</td>
<td>0.81</td>
<td>1.000</td>
<td>1.000</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>AGTPBP1</td>
<td>15.62</td>
<td>12.81</td>
<td>0.82</td>
<td>5.80×10^{-5}</td>
<td>0.015</td>
<td>Down</td>
<td>2</td>
</tr>
<tr>
<td>COL5A2</td>
<td>121.23</td>
<td>113.21</td>
<td>0.93</td>
<td>0.295</td>
<td>1.000</td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>NF1</td>
<td>15.51</td>
<td>17.43</td>
<td>1.12</td>
<td>0.581</td>
<td>0.971</td>
<td></td>
<td>7</td>
</tr>
<tr>
<td>FBN1</td>
<td>301.06</td>
<td>367.83</td>
<td>1.22</td>
<td>0.482</td>
<td>0.436</td>
<td></td>
<td>26</td>
</tr>
<tr>
<td>NIPBL</td>
<td>15.81</td>
<td>19.45</td>
<td>1.23</td>
<td>0.673</td>
<td>0.236</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>STAT3</td>
<td>13.65</td>
<td>17.04</td>
<td>1.25</td>
<td>0.111</td>
<td>0.010</td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>AXIN1</td>
<td>4.90</td>
<td>6.54</td>
<td>1.33</td>
<td>0.200</td>
<td>0.086</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>SHROOM2</td>
<td>0.040</td>
<td>0.055</td>
<td>1.37</td>
<td>0.540</td>
<td>0.538</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>RARG</td>
<td>4.75</td>
<td>6.86</td>
<td>1.44</td>
<td>0.047</td>
<td>0.020</td>
<td>Up</td>
<td>3</td>
</tr>
<tr>
<td>IFT122</td>
<td>5.67</td>
<td>8.40</td>
<td>1.48</td>
<td>0.037</td>
<td>0.008</td>
<td>Up</td>
<td>12</td>
</tr>
<tr>
<td>ALDH1A3</td>
<td>35.81</td>
<td>57.78</td>
<td>1.61</td>
<td>2.42×10^{-7}</td>
<td>1.12×10^{-7}</td>
<td>Up</td>
<td>3</td>
</tr>
<tr>
<td>C12orf57</td>
<td>3.25</td>
<td>6.83</td>
<td>2.10</td>
<td>3.90×10^{-5}</td>
<td>1.05×10^{-5}</td>
<td>Up</td>
<td>1</td>
</tr>
<tr>
<td>TENM3</td>
<td>11.40</td>
<td>29.04</td>
<td>2.55</td>
<td>5.69×10^{-15}</td>
<td>4.11×10^{-17}</td>
<td>Up</td>
<td>28</td>
</tr>
<tr>
<td>SKI</td>
<td>4.50</td>
<td>11.86</td>
<td>2.64</td>
<td>5.41×10^{-8}</td>
<td>6.52×10^{-11}</td>
<td>Up</td>
<td>13</td>
</tr>
<tr>
<td>MFAP2</td>
<td>3.55</td>
<td>9.40</td>
<td>2.65</td>
<td>2.62×10^{-4}</td>
<td>7.89×10^{-5}</td>
<td>Up</td>
<td>2</td>
</tr>
<tr>
<td>FBN2</td>
<td>6.85</td>
<td>18.74</td>
<td>2.74</td>
<td>1.82×10^{-20}</td>
<td>6.72×10^{-23}</td>
<td>Up</td>
<td>8</td>
</tr>
</tbody>
</table>

**Table 3.3. Genes in the Eye Morphogenesis Pathway.** The genes listed as being part of the Eye Morphogenesis pathway in Gene Ontology (GO:0048592) and their change in transcription occurring with the addition of 50 μM ascorbate are shown. Only genes below an FDR of 0.05 by both DESeq2 and enrichR were considered as differential (Up or Down).

**Conclusions**

The large shift in the hydroxymethylome of ARPE-19 cells after treatment with 50 μM ascorbate was accompanied by a large shift in the transcriptome. Specifically, 3,186 of the 12,776 transcripts investigated were found to change after treatment with ascorbate.
Although this is a dramatic shift, it is consistent with the monumental shift in the hydroxymethylome found in Chapter 2 and the previously reported effects of 5hmC on transcription.

5hmC has previously been shown to be positively correlated with transcription, and higher 5hmC enrichment was found in the more highly expressed transcripts. Interestingly, higher expressed transcripts showed lower 5hmC enrichment at the TSS. It has previously been shown that 5hmC is enriched in euchromatin, and it is possible that the higher 5hmC is associated with more accessible DNA\textsuperscript{161, 193}. It has also been found that 5hmC inhibits the binding of some transcription factors and the lower levels of 5hmC at the TSS may be necessary for the binding of transcription factors and polymerases\textsuperscript{161}. Upon investigation of the 5hmC levels in differential transcripts with respect to nondifferential transcripts, it was found that downregulated and nondifferential transcripts both exhibit a similar increase in 5hmC with the addition of ascorbate, but no increase in 5hmC at the TSS. However, upregulated transcripts had a larger increase in 5hmC in the promoter, TSS and gene body upon the addition of ascorbate to the culture media.

Pathway analysis revealed that the differential genes that contain DhMRs in their genomic region tend to cluster within a handful of interesting pathways. Specifically, the genes clustered in pathways related to extracellular matrix organization and apoptosis, both of which have been associated with diabetic eye complications in the inner retina. Only one pathway, eye morphogenesis, was specifically known to be related to the outer BRB. Although this refers to development of the eye, the RPE cells have been shown to secrete hormones necessary for both the development of the eye and the maintenance of healthy retinal cells in the eye throughout life\textsuperscript{136-138}. Of the 20 genes in this pathway, 12 of them change in expression after the addition of ascorbate to the media, all of which contain at least one DhMR in the genomic region. This indicates that hyperglycemia induced ascorbate deficiency in the eyes of diabetics may contribute to the change in the level of
hormones related to the development of the eye, and an imbalance of these hormones may also be related to the development of diabetic eye complications.
Chapter 4

Functional Validation of the Ascorbate-Mediated Effects in Retinal Cells

The data presented in this chapter were previously published in *Investigative Ophthalmology and Visual Sciences* (IOVS, Volume 59, Issue 8)\textsuperscript{214}. I performed most of the work presented in this chapter with several exceptions. The RPE-J cell culture and qPCR validation of the reduction of Vegfa were performed by Vladimir Camarena, MD/PhD. The qPCR to confirm the isoforms of VEGFA was also performed by Vladimir Camarena, MD/PhD. The mice for these experiments were bred and given the correct water by Sushmita Mustafi, PhD and Vladimir Camarena, MD/PhD. The microdissection of mouse retinas was performed by Yiwen Li, MD. The ELISA on mouse vitreous humor was performed by Vladimir Camarena, MD/PhD.

Background

In the previous two chapters ARPE-19 cells were used to model the outer BRB, and media with no added ascorbate was used to represent the diabetic condition while media with 50 \(\mu\)M ascorbate was used to model the restoration of vitamin C to a healthy level. Upon the addition of ascorbate, there was a dramatic shift in the hydroxymethylome, as found by hMeDIP-seq. This change in the hydroxymethylome was accompanied by a large shift in the transcriptome, as determined by RNA-seq. Pathway analysis of the genes that contained DhMRs and changed in expression after addition of ascorbate revealed that the genes clustered within several pathways. Although multiple pathways could be related to diabetic eye complications in the inner BRB, only one significant pathway, eye
morphogenesis, has previously been demonstrated to be related to the outer BRB. The hormones released from the RPE, the main constituent of the outer BRB, are necessary for eye development and for maintenance of health in the retina. Dysregulations in the hormones secreted from the outer BRB have previously been shown to be related to the development of DR\textsuperscript{29}. In particular, the gene VEGFA is included in the eye morphogenesis pathway and it is well established that an overproduction of VEGF (the protein coded by VEGFA) contributes to DR. For this reason, VEGFA was the focus of functional validation.

To test the consistency of this change in VEGFA I measured transcription and secretion of VEGF in a second batch of ARPE-19 cells, another type of immortalized RPE cells from rat (RPE-J cells) and in primary RPE cells from three separate human donors. The final experiments were to test the changes in VEGFA in the retinas of mice after conditioning them for months with high or low ascorbate levels in the drinking water.

**Methods**

*Cell Culture*

Human retinal pigment epithelial ARPE-19 cells were purchased from American Type Culture Collection (ATCC, Manassas, VA). ARPE-19 cells were maintained in DMEM-F12 medium (Life Technologies, Carlsbad, CA), which does not contain any ascorbate in its formulation. After seeding approximately $3 \times 10^5$ cells in 6-well plates or $10^6$ cells in 10 cm plates for 24 hours, cells were treated sodium ascorbate (Sigma-Aldrich) at various concentrations or for varying durations as noted, or treated with glutathione (GSH) at 50 $\mu$M for 7 days. Each treatment group consisted of three wells. Cell culture media was changed daily with or without ascorbate to maintain a constant concentration of ascorbate.
Rat retinal pigment epithelial cells (RPE-J) were also purchased from the ATCC. Cells were seeded at approximately $3 \times 10^5$ cells in 6-well plates. Cells were cultured in DMEM medium (Thermo Scientific, Waltham, MA) for 24 hours before treatment either with (50 µM) or without sodium ascorbate (Sigma-Aldrich) for 7 days.

Primary human fetal retinal pigment epithelial cells (hfRPE) were a generous donation from the Dr. Sheldon Miller lab at the National Eye Institute (NEI). Cells were cultured according to the instructions given in the publication by the Miller lab describing these cells in culture\textsuperscript{246}. Glass 12-mm coverslips were first acid washed in 1M HCl at 56°C for 4 hours with occasional agitation, followed by a thorough rinsing in water and wash in 100% ethanol. Upon receipt, cells were seeded in 24-well plates with acid-washed coverslips at 10,000 cells per well. Cells were kept in MEM Medium (Cat#5650, Sigma-Aldrich) with added N1 medium supplement, taurine, hydrocortisone and triiodo-thyronin (Sigma-Aldrich) as well as GlutaMAX supplement (Thermo Scientific). Cells were left in culture for 6 weeks and checked for high levels of pigmentation to ensure native RPE characteristics were present before treatment with sodium ascorbate or glutathione as noted with each experiment. Each treatment group consisted of at least three wells for every experiment. In all experiments cell culture media was changed daily with or without ascorbate to maintain a constant concentration of ascorbate.

Quantitative Real-time RT-PCR

RNA was extracted from cultured cells using RNeasy kits (Qiagen). A nanodrop 8000 photospectrometer was used to measure the yield of RNA extraction (Thermo Scientific, Waltham, MA). The qScript Flex cDNA kit (Quanta Biosciences, Beverly, MA) was used for reverse transcription (RT) according to the manufacturer's instructions. Quantitative real-time RT-PCR (qRT-PCR) was performed in triplicate on an ABI 7900
(Life Technologies) using the PerfeCTa SYBR Green FastMix ROX (Quanta Biosciences, Beverly, MA) with 10 μl reactions. Primers were designed to span introns (Table 4.1). The transcript amplification results were analyzed with the ABI 7900 HT software (SDS), and all values were normalized to the levels of the ACTB using the $2^{-\Delta\Delta Ct}$ method. Statistical significance of differences in expression levels was assessed by Student’s t test or one-way ANOVA with a significance level of $\alpha=0.05$.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Species</th>
<th>Forward (5’→3’)</th>
<th>Reverse (5’→3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>VEGFA</td>
<td>Human</td>
<td>GCCTTGCTTGGCTGCTTAC</td>
<td>TGATTCTGCCCTCCTCTTCTG</td>
</tr>
<tr>
<td>VEGFA 165-a</td>
<td>Human</td>
<td>GAGCAAGACAAGAAAATCCC</td>
<td>CCTCGGCTTGTCACTCTG</td>
</tr>
<tr>
<td>VEGFA 165-b</td>
<td>Human</td>
<td>GAGCAAGACAAGAAAATCCC</td>
<td>GTGAGAGATCTGCAAGTACG</td>
</tr>
<tr>
<td>ACTB</td>
<td>Human</td>
<td>TCCCTGGAGAAGAGCTACG</td>
<td>GTAGTTTCGTGGATGCCACA</td>
</tr>
<tr>
<td>Vegfa</td>
<td>Rat</td>
<td>GAGCAACGTCACTATGCAGAT</td>
<td>CTTTTGCTCATTATCATCTGC</td>
</tr>
<tr>
<td>Actb</td>
<td>Rat</td>
<td>GCCACCAGCTCGCATGGATGACG</td>
<td>CACACCCRGGRGCCRAGGGCGG</td>
</tr>
</tbody>
</table>

**Table 4.1. Primers used for qPCR Validation in Cell Based Studies.** This table gives the sequence of all primers used for qPCR validations using RNA extracted from cultured cells. All primers are given in the 5’→3’ direction.

**ELISA Assays**

Cell culture media was saved from the final 24 hours of treatment of ARPE-19 or hRPE cells with sodium ascorbate or glutathione for varying durations and varying concentrations. VEGF levels in the cell culture conditioned medium were measured using Human VEGF DuoSet ELISA kits (R&D Systems, Minneapolis, MN) with antibodies mainly specific for VEGF-165 and VEGF-121, the two major isoforms of pro-angiogenic VEGF, according to the manufacturer’s instructions. Differences were assessed by one-way ANOVA with a significance level of $\alpha = 0.05$. 
**Immunoblot**

ARPE-19 cells were grown in DMEM:F12 medium (Thermo Scientific) with 10% FBS and pretreated for 7 days with or without ascorbate (50 μM). For the last 6 hours before collection, some wells of both ascorbate treated and untreated cells were treated with 100 μM of cobalt chloride (CoCl₂) to induce a hypoxia-like state as has been performed previously. Cell lysates were collected in RIPA buffer (Thermo Scientific) with protease inhibitor cocktail (Sigma-Aldrich), 1% SDS, and 0.5 mM dithiothreitol (DTT). The protein concentration was determined by Pierce BCA protein assay kit (Thermo Scientific). Cell lysates were resolved by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), transferred to Polyvinylidene Fluoride (PVDF) membranes (Bio-Rad Laboratories, Hercules, CA), and immunoblotted with anti-HIF-1α antibody and anti-β-actin antibody (#3716S and #4970S, Cell Signaling Technology, Danvers, MA). Proteins were visualized by using chemiluminescence using SuperSignal West Femto ECL (Thermo Scientific). Specific bands densities were quantified using ImageJ and analyzed by Student’s t test, at α = 0.05.

**Animal Studies**

All mouse experiments were performed in accordance with the guidelines for the care and use of laboratory animals published by the National Institutes of Health and in accordance with the ARVO statement for the use of animals in ophthalmic and vision research. The use of the animals and the research protocol has been approved by the IACUC (Institutional Animal Care and Use Committee) at the University of Miami. Gulonolactone oxidase knockout (Gulo⁻/⁻) mice were obtained from the Mutant Mouse Regional Resource Center (University of California, Davis, CA). Gulo⁻/⁻ mice were maintained at the animal facility in the University of Miami. Gulo⁻/⁻ mice (1 month old) were
maintained for 3 months with either high ascorbate (330 mg/L of L-ascorbic acid, n = 4) or low ascorbate (16.5 mg/L of L-ascorbic acid, n = 5) with 10 µM EDTA supplementation in drinking water. After the euthanasia of the mice, the entire mouse heads were fixed in 4% paraformaldehyde (Sigma-Aldrich) for 4 days, and then stored in phosphate buffered saline at 4°C until further processing. The eyes were then removed, and the RPE/choroid layer was dissected away from the remainder of the retina. Both RPE/choroid sections from a single mouse were pooled. Total RNA was extracted from the RPE/choroids using the RecoverAll Total Nucleic Acid Isolation kit (Thermo Scientific). The qScript Flex cDNA kit (Quanta Biosciences) was used for reverse transcription (RT) according to the manufacturer’s instructions. Quantitative real-time RT-PCR (qRT-PCR) was performed in triplicate on a QuantStudio 12K Flex (Thermo Scientific) using PowerUp SYBR Green Master Mix (Thermo Scientific) with 10 µl reactions. Primers were designed to span introns (Table 4.2). The transcript amplification results were analyzed with the QuantStudio software, and all values were normalized to the levels of the Actb using the $2^{-\Delta\Delta Ct}$ method.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Forward (5’→3’)</th>
<th>Reverse (5’→3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vegfa</td>
<td>GCCAGCACATAGGAGAGATGAGCTTCC</td>
<td>CTTTGTTGTCATTACATCTGCTGTGC</td>
</tr>
<tr>
<td>Actb</td>
<td>GTCGAGTCGCGTCCACC</td>
<td>GTCATCCATGGCGAAGTGGT</td>
</tr>
</tbody>
</table>

Table 4.2. Primers used for qPCR Validation in Mouse RPE/Choroids. This table gives the sequence of primers used for qPCR validations in RPE/choroids dissected from mice. All primers are given in the 5’→3’ direction.

In a separate experiment, male Gulo-/- mice (3 months old) were maintained for 4 months with or without ascorbate (330 mg/L of L-ascorbic acid) with 10 µM EDTA supplementation in drinking water (n = 3 per group). After the euthanasia of the mice, the vitreous humor from one eye was extracted and stored at -80°C until further processing.
VEGF protein levels were measured in the vitreous humor samples using the Quantikine ELISA Mouse VEGF (R&D Systems) according to manufacturer’s instructions. Briefly, 2 μl vitreous humor (from a single eye) was diluted in 8 μl RIPA buffer containing protease inhibitors (Millipore Sigma, Burlington MA). This 10 μl of this vitreous humor cocktail was added to 40 μl of sample diluent and loaded to the plate, along with an additional 50 μl of Assay Diluent, followed by a 2-hour incubation. Following a wash step, a conjugate solution was added to the wells for another 2 hours, followed by another wash step. A substrate was added to the wells for 30 minutes, and the reaction was halted with a stop solution. The absorbance was read at 450 nm with a correction wavelength of 540 nm. Statistical significance of differences in expression levels was assessed by Student’s t test, at α = 0.05.

Results

Choosing a Finding for Further Validation

Finding a dramatic shift in both the hydroxymethylome and the transcriptome after culture with physiological levels of ascorbate is indicative that the ascorbate deficiency in diabetes indeed affects the epigenetics of the RPE cells. However, not all changes in the RPE cells are likely to contribute to the changes involved in DR. Pathway analysis using EnrichR (Chapter 3) revealed a single pathway that is known to be related to both diabetic eye complications and the outer BRB: eye morphogenesis. Although eye morphogenesis specifically refers to the development of the eye, the hormones secreted from the RPE are necessary for both the development of the retina and for the maintenance of a healthy retina throughout life\textsuperscript{137, 138}. Either overproduction or underproduction of these hormones can lead to pathological conditions\textsuperscript{29}. 
Upon investigation of the genes listed in the eye morphogenesis pathway (Table 3.3), we find that 12 of the 20 genes are differentially expressed between the low ascorbate and physiological ascorbate conditions, and all 12 genes contain at least one DhMR in the genomic region. Of these 12 genes, the most extreme difference by both fold change and $P$ value is in the gene VEGFA. The expression of VEGFA reduces from 458.63 FPKM in the RNA from cells cultured without ascorbate to 47.32 FPKM in cells cultured with 50 μM ascorbate (Table 3.3). Upon investigation of the hMeDIP-seq data I found that three peaks of 5hmC enrichment were called in the VEGFA genomic region using the MACS2-IDR pipeline. Two peaks in intron 2 increased with 50 μM ascorbate treatment (FDR = 0.0177, 0.0041 respectively). A large peak was detected from intron 5 through the 3’UTR but was not called as differential (Figure 4.1, Table 4.3).

<table>
<thead>
<tr>
<th>Genomic Location</th>
<th>VEGFA Location</th>
<th>Size</th>
<th>Fold Change</th>
<th>Average FPKM* 0 μM Ascorbate</th>
<th>Average FPKM* 50 μM Ascorbate</th>
<th>FDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>chr6:43,775,008-43,776,044</td>
<td>Intron 2</td>
<td>1036</td>
<td>2.295</td>
<td>1.490</td>
<td>3.419</td>
<td>0.0177</td>
</tr>
<tr>
<td>chr6:43,776,137-43,777,624</td>
<td>Intron 2</td>
<td>1487</td>
<td>2.813</td>
<td>1.100</td>
<td>3.096</td>
<td>0.0041</td>
</tr>
<tr>
<td>chr6:43,779,807-43,783,537</td>
<td>Intron 5 - 3’UTR</td>
<td>3730</td>
<td>1.749</td>
<td>1.083</td>
<td>1.894</td>
<td>0.0746</td>
</tr>
</tbody>
</table>

Table 4.3. hMeDIP Peaks within the VEGFA Genomic Region. This table shows the peaks called within the VEGFA genomic region, along with the average coverage, fold change and false discovery rate (FDR) from edgeR. *FPKM = Fragments Per Kilobase per Million mapped reads.

VEGF is required for vascularization throughout the body and is necessary for survival\textsuperscript{134}. It is important in maintaining a healthy retina and constitutively expressed in the RPE cells\textsuperscript{137}. However, overproduction of VEGF has been implicated in DR.
Figure 4.1. RNA Decreases and 5hmC Increases in the Gene Body of VEGFA after Ascorbate Treatment. UCSC tracks show the coverage of the RNA-seq and hMeDIP-seq (5hmC) over the VEGFA genomic region. The two red stars indicate the 5hmC peaks that increase with 50 μM ascorbate treatment.

VEGF is required for vascularization throughout the body and is necessary for survival\(^{134}\). It is important in maintaining a healthy retina and constitutively expressed in the RPE cells\(^{137}\). However, overproduction of VEGF has been implicated in DR. VEGF levels in the eyes of patients with DR were significantly higher than in the eyes of patients with other eye complications\(^{74}\). Experimental increases in VEGF induce hyperpermeability of the BRB\(^{140}\). Mice generated to express human VEGF in the photoreceptors (known as the Kimba model) show many of the signs of DR including vascular leakage, venous bleeding, neovascularization and focal retinal detachment\(^{144}\). These findings have led to the development of four separate drug treatments that inactivate VEGF protein in the eye.
These treatments have been tested with varying levels of success relative to one another, but anti-VEGF treatments have been found to be more successful at treating vision loss from DME than laser photocoagulation. Many groups believe this should be the first line of defense of treatment for diabetic eye complications. Given the known role of VEGF in DR, the reduction of VEGFA in RPE cells upon treatment with ascorbate was the finding chosen for further validation.

**qRT-PCR Verification of the Reduction of VEGFA Expression by Treatment with Ascorbate**

To confirm the reduction of VEGFA after the treatment of RPE cells with 50 μM ascorbate, a second batch of ARPE-19 cells was thawed, plated and treated. By qRT-PCR, the level of VEGFA in cells treated with 50 μM was only 13% of the level in the cells treated with 0 μM ascorbate (Figure 4.2A).

To confirm the reduction in VEGFA in a second cell line, I chose RPE-J cells from rat. These cells are immortalized and passage easily, but give a fairly good representation of the native RPE. RPE-J cells were treated with (50 μM) or without ascorbate for 7 days. qRT-PCR showed that the RPE-J cells cultured with ascorbate had only 43% of the level of Vegfa as the cells without ascorbate, indicating that this is not specific to ARPE-19 cells (Figure 4.2B).

It is known that VEGFA can produce two families of transcripts – the ‘a-isoform’ with pro-angiogenic function and ‘b-isoform’ with anti-angiogenic capacity. The expression levels of the two isoforms in RNA from ARPE-19 cells were assessed by qRT-PCR using isoform specific primers. Results showed that the a-isoform was ~25 fold more abundant than the b-isoform in cells cultured without ascorbate. After treatment with
ascorbate, transcripts of both isoforms were significantly reduced and the a-isoform remained ~25 fold more abundant (Figure 4.2C).

Figure 4.2. qRT-PCR Verification of the Suppression of VEGFA by Ascorbate. (a) The inhibition of ascorbate on VEGFA transcription was validated in a second batch of ARPE-19 cells treated with 50 μM ascorbate for 7 days. (b) The transcription of Vegfa (homolog of VEGFA) in RPE-J cells was inhibited by treatment with 50 μM ascorbate for 7 days. (c) Pro-angiogenic VEGFA 165-a isoform is far more abundant than anti-angiogenic VEGFA 165-b in ARPE-19 cells. Both isoforms are significantly reduced after treatment with 50 μM ascorbate for 7 days. (** P < 0.01, *** P < 0.001)
Confirmation of the Reduction of VEGF at the Protein Level

To examine whether the reduced VEGF transcription leads to a decrease in VEGF protein secreted from the cells, we employed ELISA to test conditioned medium of ARPE-19 cells. ARPE-19 cells were treated with media containing 50 μM ascorbate for varying durations of time or treated with varying concentrations of ascorbate for 7 days. The media was changed daily, and the media tested was the media on the cells during the final 24 hours of treatment. ELISA assays indicated that ascorbate treatment reduced the secretion of VEGF in a time-dependent (Figure 4.3A) as well as dose-dependent manner (Figure 4.3B).

Figure 4.3. The Suppression of VEGF by Ascorbate was Confirmed at the Protein Level by ELISA Testing. (a) ELISA results shows that ascorbate time-dependently decreases the levels of VEGF protein in the conditioned culture medium from ARPE-19 cells treated with 50 μM ascorbate. (b) ELISA testing shows that ascorbate dose-dependently decreases the levels of VEGF in the conditioned culture medium after treatment for 7 days. (*** P < 0.001)

Suppression of VEGF is Not Correlated with HIF-1α or Antioxidant Effects

Hypoxia-inducible factor 1-alpha (HIF-1α) is an important transcription factor that upregulates the expression of VEGF under hypoxic conditions\textsuperscript{251}. Degradation of HIF-1α...
requires ascorbate because it is a cofactor for HIF-1α hydroxylases, enzymes that initiate the degradation of this transcription factor\textsuperscript{252}. Hence, the suppression of VEGFA by ascorbate could be mediated through an accelerated degradation of HIF-1α. To test this hypothesis, I measured HIF-1α levels in cultured ARPE-19 cells. HIF-1α was undetectable by Western blot in cultured ARPE-19 cells with or without ascorbate (Figure 4.4A). This result was expected, since the cell incubator environment contains 5% CO\textsubscript{2} and 95% atmospheric air, which contains nearly 21% oxygen. Addition of CoCl\textsubscript{2}, a compound that mimicks a hypoxic environment, induced a significant increase in HIF-1α protein which could be inhibited with ascorbate co-treatment (Figure 4.4A)\textsuperscript{247}. These results indicate that it is unlikely that the inhibition of VEGF by ascorbate is mediated by HIF-1α degradation since HIF-1α is not detectable in culture without hypoxia induction, yet ascorbate still reduces VEGF. This is consistent with previous studies that have shown that VEGFA expression can be mediated by factors other than hypoxia in the retina\textsuperscript{135}.

Alternatively, it is possible that ascorbate, a reducing agent, suppresses VEGF expression by attenuation of oxidative stress and associated reactive oxygen species (ROS), which can upregulate VEGF in RPE cells\textsuperscript{253}. To rule out this possibility, ARPE-19 cells were treated for 7 days with either ascorbate or glutathione, a potent antioxidant. No inhibition of VEGF expression was seen in cells treated with glutathione (Figure 4.4B). In contrast, VEGF expression was reduced to \sim 11\% of the control level in the presence of ascorbate. These findings rule out the possibilities that the suppression of VEGF expression by ascorbate is mediated through either the HIF-1α pathway or attenuation of ROS, leading to the conclusion that the decrease in VEGF is likely through the change in 5hmC levels in the gene body (Figure 4.1).
Figure 4.4. HIF-1α and Antioxidant effects are Not Responsible for the Suppression of VEGF by Ascorbate in Culture Conditions. (a) Western blot shows that there was no detectable HIF-1α in cultured ARPE-19 unless induced by CoCl₂ (100 μM) to mimic hypoxic conditions. After treatment with ascorbate (50 μM), the induced HIF-1α almost disappeared, but HIF-1α was not present in the experimental conditions of all other experiments. (b) qRT-PCR shows the suppression of VEGFA transcription by ascorbate, while a stronger reducing agent glutathione (50 μM) exerted no obvious change in VEGFA transcripts in ARPE-19 cells. (** P < 0.001)

Ascorbate Reduces VEGF Expression in Primary Human Fetal RPE cells.

Although immortalized RPE cells (ARPE-19 and RPE-J) are convenient to grow for many passages and maintain a uniform cell homogeneity, they are not entirely representative of native RPE246. In contrast, primary human fetal RPE cells (hfRPE) are
more difficult to grow and have higher cellular heterogeneity, but are an accurate representation of native RPE. They exhibit high levels of RPE markers, metabolize all-trans retinal, express high levels of pigmentation, polarize in culture, and develop high transepithelial resistance (TER)\textsuperscript{246}.

To test the effect of ascorbate in primary culture, RPE cells from three separate donors were obtained from the National Eye Institute (NEI). The initial culture media for hfRPE contains 250 $\mu$M L-ascorbic acid in its formulation. Following the first passage, cells were cultured with MEM media with the same formulation as the traditional MEM media used, except without the ascorbic acid. After six weeks of treatment with ascorbate deficient media, cells were treated with no additional compounds, 50 $\mu$M sodium ascorbate or 50 $\mu$M glutathione for 7 days. By qRT-PCR, it was confirmed that the transcription of VEGF decreased after ascorbate treatment to $\sim$50-70% of the pre-treatment levels in cells from all three donors (Figure 4.4 A-C). Additionally, ELISA on the conditioned culture medium revealed that levels of secreted VEGF from the cells of all three donors after ascorbate treatment were significantly reduced from the levels when ascorbate was absent (Figure 4.4 D-F). These experiments indicate that, similar to immortalized RPE cell lines, primary cells also downregulate VEGF in response to ascorbate treatment.

*Ascorbate Reduces VEGF in the Vitreous Humor of Mouse Eyes.*

To examine the effects of ascorbate on VEGF expression *in vivo*, we used *Gulo*\textsuperscript{-/-} mice. Unlike humans and other primates that cannot synthesize ascorbate due to loss-of-function mutations in the L-gulonolactone oxidase (*Gulo*) gene, rodents have fully functional L-gulonolactone oxidase. This is the enzyme that catalyzes the final step of ascorbate biosynthesis and thus they can synthesize ascorbate de novo in the liver.
Figure 4.5. Ascorbate Reduces the Expression of VEGFA in Primary hfRPE Cells. qRT-PCR shows a reduction of VEGFA transcripts after treatment with ascorbate (50 μM), but not after treatment with glutathione (50 μM) in donors 1-3 (a-c). ELISA testing shows a reduction in the amount of secreted VEGF in the cell culture conditioned medium after treatment with ascorbate (50 μM), but not after treatment with glutathione (50 μM) in donors 1-3 (d-f). (* P < 0.05, ** P < 0.01)

Genetic ablation of the Gulo gene in rodents creates mice that, like humans, require vitamin C through their diet to prevent scurvy\textsuperscript{248}. The long-term survival of Gulo\textsuperscript{-/-} mice relies on ascorbate supplements. When supplemented with 330 mg/L of ascorbate in the drinking water, the plasma ascorbate concentration of Gulo\textsuperscript{-/-} mice is maintained at ~50 μM, similar to the level of wild type mice\textsuperscript{248}.

Gulo\textsuperscript{-/-} mice (1-month-old) maintained with 330 mg/L of ascorbate in their drinking water, were divided into two random groups. The ascorbate supplementation was reduced
to 16.5 mg/L for one group, but maintained at 330 mg/L for the other group. Three months later, mice were sacrificed and the eyes were dissected. RNA was extracted from the RPE/choroid layers and the levels of Vegfa transcripts were measured by qRT-PCR. The mice with high ascorbate supplementation showed significantly reduced levels of Vegfa transcripts (0.75-fold) in the RPE/choroid layers in comparison to those mice given low levels of vitamin C ($P = 0.031$, Figure 4.6A).

In a second experiment, $Gulo^{-/}$ mice (3-months-old) maintained with 330 mg/L of ascorbate in their drinking water, were divided into two random groups. Ascorbate supplementation was removed from one group, while the ascorbate supplementation was continued for the other group of mice. Four months later, mice were sacrificed and VEGF protein level was measured in the vitreous humor by ELISA. The $Gulo^{-/}$ mice with ascorbate supplementation showed a significant decrease of VEGF (0.57-fold) accumulated in their vitreous humor in comparison to those without ascorbate supplementation ($P = 0.029$, Figure 4.6B). This in vivo data indicates that ascorbate inhibits VEGF expression in the eye, suggesting that our in vitro studies do mirror in vivo conditions.
**Figure 4.6. Decrease in the Levels of VEGF in the RPE/Choroid and Vitreous Humor of Ascorbate–Deficient Mice Supplemented with Additional Ascorbate.** (a) qRT-PCR shows a reduction in Vegfa transcripts in the RPE/choroid layers of Gulo−/− mice after being given sufficient ascorbate in the drinking water (330 mg/L) in comparison to mice kept with low ascorbate (16.5 mg/L) for three months. (b) The level of secreted VEGF protein, measured by ELISA, was decreased in the vitreous humor of Gulo−/− mice given sufficient ascorbate in the drinking water (330 mg/L) in comparison to mice given no ascorbate for four months. (* P < 0.05)

**Conclusions**

Of the pathways that had significant clustering of differential genes, only one of them was specifically related to the outer BRB. This pathway, eye morphogenesis, is specifically referring to the development of the eye but the same hormones that are important for eye development are important for the maintenance of a healthy retina throughout life. Of the genes in this pathway that change in transcription, the most dramatic change was a near 10-fold decrease in the expression of VEGFA. This gene codes for VEGF, a protein that is increased in the eyes of patients with DR. In experimental models it has been found that overexpression of VEGF leads to pathological conditions similar to those seen in the diabetic retina. Furthermore, drug treatment to inhibit VEGF from binding to its receptors in the eye has been shown to be very effective in preventing vision from worsening and has even been found to improve the vision of many individuals with DR. The dramatic reduction in VEGF after the addition of ascorbate to the culture media indicates that hyperglycemia-induced ascorbate deficiency in the eyes of diabetics may lead to the overproduction of VEGF in the RPE cells, which could contribute to the development of DR.

This decrease in VEGFA after the addition of ascorbate was verified in another batch of ARPE-19 cells and confirmed in RPE-J, another immortalized RPE cell line. This decrease was also seen at the protein level by ELISA test on the culture media used on the cells, indicating that the levels of secreted VEGF also decrease with ascorbate.
treatment. The decrease in VEGFA occurred when HIF-1α was not present, indicating that it is not related to hypoxia. Glutathione, an antioxidant with greater strength than ascorbate, was unable to reduce the expression of VEGFA, indicating that this effect was not mediated by the antioxidancy effects of ascorbate. However, there were two peaks of 5hmC enrichment in the second intron of the VEGFA gene that increased in strength after the addition of ascorbate to the culture media which may have mediated the changes in transcription.

To further confirm the findings of these cell lines, primary human fetal RPE cells were used. Cells from three separate donors showed a consistent decrease in VEGF at both the RNA and protein level upon treatment with 50 μM of ascorbate for 7 days, while glutathione had no effect. This confirms the ascorbate mediated decrease in VEGF in human cells that are a very good representation of native tissue.

To confirm that the decrease in VEGF in vivo, we experimented using Gulo−/− mice which, like humans, cannot produce vitamin C in the liver. In the first set of mice, some were given low ascorbate in the drinking water while others were given high levels of ascorbate in the drinking water to maintain ascorbate levels at ~50 μM. After three months, the mice that were given high levels of ascorbate in the drinking water had lower levels of Vegfa in the RPE/choroid layers, as determined by qRT-PCR. A second set of mice was conditioned for four months with either no ascorbate or high levels of ascorbate in the drinking water. The mice with high ascorbate in the drinking water had lower levels of VEGF in the vitreous humor as measured by ELISA testing.

Together these results indicate that VEGF expression and secretion in the eye is affected by the availability of ascorbate. These effects are likely mediated by the ability of the TET enzymes to hydroxymethylate the DNA in the genomic region of the VEGFA gene,
specifically at the second intron. This link between ascorbate and VEGF production was confirmed in multiple cell lines, primary human cells and in vivo using mice.
Chapter 5

Summary, Future Direction and Discussion

Summary of Findings

Diabetic retinopathy is a complication of both T1DM and T2DM that greatly reduces the quality of life\(^{36}\). Owing to the frequency of diabetes, DR is the most common cause of vision loss in working age adults around the world and is expected to become more common in future years\(^{9,37}\). It is evident that prevention strategies for DR are needed, and investigation of the unexplored aspects of the disease is needed.

The majority of pathological effects in DR are attributed to the breakdown of the BRB and the growth of new blood vessels. The breakdown of the BRB is largely attributed to the death of endothelial cells and capillaries in the retina. The lack of viable blood vessels in the retina is followed by hypoxia, which is accompanied by overexpression of VEGF that leads to the growth of new blood vessels to nourish the previously undernourished retina\(^{40}\). The most effective treatments for DR have been laser photocoagulation and anti-VEGF therapies\(^{67-69}\).

Ascorbate (vitamin C) is absorbed into the retina from the blood across the BRB in its oxidized form through glucose transporters, GLUT1 and GLUT3. High glucose found in hyperglycemia acts as a competitive inhibitor, thus leading to local ascorbate deficiency in the eyes of diabetics\(^{172}\). This ascorbate deficiency in the eyes of diabetic patients has been found in multiple studies and experimentally shown in diabetic animal models\(^{175-177}\).

Ascorbate serves as a cofactor for the TET enzymes, which actively demethylate cytosines from 5mC to 5hmC and further to other oxidized species of 5hmC that will be replaced by unmodified cytosine through base excision repair\(^{185}\). Without ascorbate, the
function of the TET enzymes is impaired and the methylation/demethylation dynamic in the cell is changed. This has been shown to lead to a decrease in global 5hmC and is implicated in multiple diseases\textsuperscript{195-212}. It is likely that the ascorbate deficiency in the eyes of diabetics leads to impaired function of the TET enzymes, which leads to epigenetic changes that contribute to the development of DR.

To test the role of ascorbate in DR, I chose to use ARPE-19 cells, a cell line that gives a fairly good representation of RPE cells, the main constituent of the outer BRB. These cells have been extensively used and can be grown for multiple passages. To represent the diabetic condition, cells were cultured in media with no ascorbate. To represent a restoration of ascorbate to healthy levels, most experiments featured 50 µM of ascorbate, the mean plasma concentration of healthy individuals\textsuperscript{168}. Upon treatment with media containing a physiological level of ascorbate, the global levels of 5hmC increased substantially in ARPE-19 cells, indicating a shift in demethylation dynamics. hMeDIP-seq was used to investigate relative 5hmC levels throughout the genome. When ascorbate was added to the media, there was a significant shift in approximately half (64,078) of the regions of 5hmC enrichment. The majority of differential peaks showed greater 5hmC enrichment in the ascorbate treated samples.

Whole transcriptome sequencing also showed a dramatic shift in the transcriptome of ARPE-19 cells after treatment with a physiological level of ascorbate. A total of 3,186 genes were found to differ in expression between groups, and 2,222 of the differential genes contained at least one differentially enriched 5hmC peak within their genomic region. Although acting as a cofactor for the TET enzymes is not the only function of ascorbate in the cell, this indicates that a large portion of the transcription changes might be related to the changes in 5hmC that occur as a result of the restoration of function to the TET enzymes.
Pathway analysis using the differential genes that contained DhMR’s revealed that these genes cluster within a handful of pathways. Some of the pathways may be implicated in diabetic eye complications, but only one pathway, eye morphogenesis, is related to both DR and the outer BRB. Although eye morphogenesis refers specifically to the development of the eye, studies have shown that the hormones that are responsible for the development of the eye are required to maintain visual function throughout adulthood^{136-138}.

Of the 20 genes listed in the eye morphogenesis pathway, 12 of them changed in expression after treatment with ascorbate. All 12 of these genes also contained at least one DhMR in the genomic region. The most dramatic change in this pathway was a near 10-fold decrease in VEGFA. This gene codes for VEGF, the hormone that is known to be upregulated in the eyes of diabetics and the target of drug treatments for DR. This could suggest a potential therapeutic role of ascorbate at maintaining the transcription of VEGFA at a healthy level.

This ascorbate mediated decrease in VEGF was confirmed at the RNA and protein level in a second batch of ARPE-19 cells and in a second immortalized line of RPE cells from rat. This decrease in VEGF is found in non-hypoxic conditions when HIF-1α is not present, indicating that this change is not dependent on HIF-1α. This is consistent with the previous findings that hypoxia is not the only regulator of VEGFA expression in the RPE^{135}. Glutathione, a strong antioxidant, also had no effect on the expression of VEGFA, indicating that the antioxidancy of ascorbate is not responsible for the decrease in VEGFA.

There were, however, two peaks of 5hmC enrichment in intron 2 of VEGFA that increase in strength after the addition of ascorbate. Generally, suppression of a gene is more frequently associated with an increase of 5hmC in the promoter rather than the gene body, but previous reports have shown that changes in the gene body of a gene can also be
correlated with a decrease in gene transcription\textsuperscript{254}. It is noteworthy that two single nucleotide polymorphisms (SNPs) in the second intron of VEGFA (rs735286 and rs2146323) were found to be associated with the severity of diabetic retinopathy, and these SNPs fall within the DhMRs found in this study\textsuperscript{255}. Additionally, another SNP within intron 2 (rs833069) is associated with the risk of AMD, and is located in close proximity to the first DhMR\textsuperscript{256}. AMD is a disease separate from DR, but both diseases are characterized by blindness caused by neovascularization in the eye related to an increase in VEGF in the eye. These associations indicate that this region within intron 2 may be important for regulation of VEGFA expression, and changes in hydroxymethylation in this regulatory region are likely to affect expression.

ARPE-19 cells give a fairly good representation of native RPE, but do not capture some aspects of native tissue. To confirm the ascorbate mediated decrease of VEGF, primary human fetal RPE cells were obtained from the Dr. Sheldon Miller lab at the National Eye Institute. The decrease in VEGF was found at both the RNA and protein level in RPE cells from three separate donors.

To validate the effect of ascorbate on VEGF expression \textit{in vivo}, a mouse model was used. Mice with mutations in the gulonolactone oxidase gene do not produce ascorbate in the liver and thus rely on dietary sources of vitamin C, similar to humans\textsuperscript{248}. These mice were treated with either deficient or high levels of ascorbate in the drinking water. After sacrificing the mice, it was found that the mice with high ascorbate in the drinking water had significantly lower expression of Vegfa in the cells of the RPE/choroid layer and significantly lower levels of VEGF protein in the vitreous humor. Together these results indicate that ascorbate contributes to the maintenance of VEGF expression and production at healthy levels in RPE cells, and this effect is likely mediated by hydroxymethylation.
Translational Implications

The expression of VEGFA in the RPE has previously been shown to be regulated by events other than hypoxia, but this study is the first time that its expression has been demonstrated to be affected by ascorbate. The findings of this study suggest that the link between ascorbate and VEGFA lies in hydroxymethylation in the second intron of the gene. VEGFA is overexpressed in DR and restoration of ascorbate to a healthy physiological reduces its expression dramatically in RPE cells. This finding supports my hypothesis that hyperglycemia-induced ascorbate deficiency in the eyes of diabetics contributes to the epigenetic dysregulations in the retina that promote the development of diabetic retinopathy.

In my studies with Gulo mice, an increase of vitamin C in the drinking water led to a decrease in VEGF at both the protein and RNA level within a few months. This indicates that oral supply of vitamin C can affect ascorbate levels in the eye, meaning that diet can affect the levels of VEGF in the eye. For individuals with diabetes, hyperglycemia will act as a competitive inhibitor to ascorbate absorption in the eye. However, higher levels of ascorbate in the diet will lead to higher levels of ascorbate in the blood, which can compensate for the lower absorption. Additionally, careful management of hyperglycemia will lessen the competition for transport and further enhance the benefit of dietary ascorbate. This increase in dietary vitamin C is likely to help prevent DR from occurring.

The benefit of a dietary means to attenuate the excess production of VEGF in the eye and prevent DR is apparent. Patients receiving anti-VEGF injections must pay multiple visits to a certified medical professional and receive injection of a drug through a needle into the eye. Patients receiving anti-VEGF treatment must receive and average of 12 injections. Vitamin C is a readily available nutrient found at high levels in many foods.
including citrus fruits, green leafy vegetables, peppers, potatoes and tomatoes. Additionally, where fresh fruits and vegetables will not supply a sufficient amount of vitamin C, inexpensive capsules of vitamin C are available.

When patients with diabetes begin to show signs of DR, a rigorous change in diet to include great amounts of vitamin C is likely not going to be as effective as anti-VEGF shots that immediately inactivate the majority of circulating VEGF protein in the eye and therefore should not replace anti-VEGF treatment. However, addition of a small amount of sodium ascorbate to the injection may be enough to raise ascorbate levels in the eyes and may help attenuate future excess production of VEGF.

**Future Directions**

This finding of a link between ascorbate and VEGF production in the eye is exciting, and has led to more questions with regard to ascorbate and its relationship to DR. Multiple experiments are needed to further validate these findings.

In this study I investigated the role of ascorbate on the hydroxymethylome of RPE cells to investigate the effect of the restoration of activity of the TET enzymes. I would be interested in performing bisulfite sequencing on the DNA of RPE cells treated with the same conditions to determine the location of changes in 5mC. Additionally, I would like to investigate the effect of repletion of ascorbate on cells of the inner BRB, particularly pericytes and retinal endothelial cells because they are the cell types thought to be most affected in the inner BRB during DR.

The validation part of my thesis focused on VEGFA because of its well-known role in breakdown of the BRB and neovascularization of DR. I would be interested in performing a more thorough investigation of the splice isoforms of VEGF and determining if there is any change in splicing events associated with the ascorbate induced change in
hydroxymethylation. Additionally, differential genes were found to cluster in pathways relating to the extracellular matrix organization and apoptosis (Chapter 3). These pathways were not pursued for validation because they have previously been implicated predominantly in the cells of the inner BRB. However, apoptosis of the RPE cells has also been observed in DR, and further experiments on apoptosis of RPE cells in high oxidative stress conditions (such as hyperglycemia) after the repletion of ascorbate could contribute further to an understanding of DR. Additionally, an increase of production of extracellular matrix proteins such as fibronectin, collagen and laminin in the RPE cells has previously been found in association with DR. I would also be interested in testing the effects of repletion of ascorbate on the production of these extracellular matrix proteins and determining if they contribute to the production of fibrotic membranes, such as have been found in DR.

In addition, my work focused on the changes that occur because of the change in concentration of ascorbate. However, it would be interesting to see the effects on the cell upon changing the concentration of glucose in the cell culture media. In particular I would like to perform a second set of experiments investigating the hydroxymethylome and transcriptome in RPE cells cultured with glucose at either 5.5 µM or 15 µM, which are concentrations of glucose in healthy and diabetic plasma, respectively.

Aside from cell-based work, I would like to perform more experiments to test the effects of ascorbate in mice. I have already shown from these studies that three months of high versus low ascorbate in the drinking water is sufficient to lead to changes in the levels of VEGF in the eyes of Gulo^+/- mice, but these mice did not show any eye disease as most eye complications take years to develop. In future experiments, I would like to treat the mice for 6 months with high or low ascorbate, followed by laser induction of choroidal neovascularization and one to two weeks of healing. I hypothesize that the mice
with low ascorbate will have higher production of VEGF in the eye as was found in my previous experiments, and that the areas of blood vessels will be significantly larger in the ascorbate deficient mice.

For another experiment, I would like to test the effect of ascorbate on Ins2\textsuperscript{Akita} mice. These mice, commonly referred to as Akita mice, produce insufficient levels of insulin. Without treatment, heterozygous mice have severe hyperglycemia but remain alive for nearly as long as their wildtype littermates. Additionally, these mice have been shown to exhibit the early signs of diabetic retinopathy and even progress to the presence of acellular capillaries in the eye at age 36 weeks\textsuperscript{260}. I would like to cross these mice with \textit{Gulo}\textsuperscript{-/-} mice to create diabetic mice that do not produce endogenous ascorbate and test them for the effects of ascorbate sufficiency versus deficiency. I would specifically look for the presence of apoptosis in the retina, increased vascular permeability, decreased number of cell bodies in the retinal ganglion cell layer and number of acellular capillaries to be different between groups. Furthermore, if the ascorbate deficient Akita mice developed signs of DR, I would like to investigate the effect of including ascorbate in anti-VEGF injections to treat the DR. I hypothesize that including ascorbate in the injections will allow the mice to remain clear of edema and neovascularization for a longer time after treatment.

These mice studies would ideally be followed up with longitudinal studies in humans. If another round of the Diabetes Control and Complications Trial is to be completed, I would suggest that they include one study where they give diabetic patients high doses of oral vitamin C and compare the frequency of diabetic eye complications between the high vitamin C group and the traditional group. This would be best if it could be used in conjunction with careful management of hyperglycemia, as the excess glucose in poor management will work as a competitive inhibitor against the absorption of
ascorbate. I hypothesize that the patients with careful management of hyperglycemia and high oral doses of vitamin C would have the lowest frequency of eye complications.

To overcome the competitive inhibition of glucose in patients with worse management of hyperglycemia, intravenous injection of ascorbate should be tested. Multiple studies have tested the safety and benefit of intravenous injections of up to 150 grams of ascorbic acid in patients. The results have shown that overall the injection of ascorbic acid is safe and well tolerated, even in sick patients such as those with cancer\(^{261, 262}\). Intravenous injection could raise the ascorbate levels in the blood much higher than could be achieved through the diet and overcome the competition of glucose to enter the eye. Given that this treatment has been well tolerated in clinical trials, this is perhaps less invasive than the eye injections of anti-VEGF drugs. When patients exhibit signs of early DR they could begin intravenous injections of ascorbate on perhaps a monthly basis to prevent further worsening of symptoms.

Finally, I would like to see a study investigating the effects of including sodium ascorbate in anti-VEGF injections. This would not introduce any more injections as it would be included with an already prescribed treatment, but the added ascorbate in the eye could help reduce the production of VEGF and possibly allow for anti-VEGF injections to be administered less frequently.

**Discussion**

In this thesis I present the work of investigating the impact of restoring ascorbate in the outer BRB in the diabetic eye. I have specifically investigated the hydroxymethylome and the transcriptome of RPE cells when cultured with and without ascorbate. A large shift was found in both the hydroxymethylome and the transcriptome. One particularly exciting finding is that restoring ascorbate to physiological levels significantly reduces the
expression and production of VEGF in the RPE. This was validated in multiple cell types and in mice. Importantly, the mice studies used dietary vitamin C and showed that even through the diet, ascorbate can affect the production of VEGF in the eye. This indicates that increasing dietary intake of vitamin C may be beneficial in preventing diabetic retinopathy.
References


119


