Arts & Sciences Electronic Theses and Dissertations   Arts & Sciences

Spring 5-15-2015

Excessive Complement Activation Due to Genetic Haploinsufficiency of Regulators in Multiple Human Diseases

Michael Triebwasser
Washington University in St. Louis

Follow this and additional works at: https://openscholarship.wustl.edu/art_sci_etds

Part of the Biology Commons

WASHINGTON UNIVERSITY IN ST. LOUIS

Division of Biology and Biomedical Sciences

Human and Statistical Genetics

Dissertation Examination Committee:

John P. Atkinson, Chair
Alison Goate
Timothy Ley
Robi Mitra
D. Michael Nelson
John Rice

Excessive Complement Activation Due to Genetic Haploinsufficiency
of Regulators in Multiple Human Diseases

By

Michael Paul Triebwasser

A dissertation presented to the
Graduate School of Arts & Sciences
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

May 2015

St. Louis, Missouri

Table of Contents

List of Figures

List of Figures (continued)

Chapter 5

List of Tables

List of Tables (continued)

Chapter 5

List of Abbreviations

1kG	1000 Genomes Project
Ab	Antibody
aCGH	array comparative genomic hybridization
aHUS	Atypical hemolytic uremic syndrome
AMD	Age-related macular degeneration
ANA	Anaphylatoxin domain
AP	Alternative pathway
APLA	Anti-phospholipid antibody
C3-T	C3-Target
C3GN	C3 glomerulonephritis
C3NeF	C3 nephritic factor
C4BP	C4 binding protein
CA	Cofactor activity
CCP	Complement control protein
CD11b	Complement receptor 3
CD11c	Compement receptor 4
CP	Classical pathway
CR1	Complement receptor 1
CR2	Complement receptor 2
CR3	Complement receptor 3
CR4	Compement receptor 4
Crry	complement receptor 1-related gene/protein-y
CUB	C1r/C1s, Uegf, Bmp1 domain
CV	Coefficient of variance
CVF	Cobra Venom Factor
d	days
D+HUS	Diarrheal hemolytic uremic syndrome
DAA	Decay accelerating activity
DAB	3,3'-Diaminobenzidine
DAF	Decay accelerating protein
DDD	Dense deposit disease
DNA	deoxyribonucleic acid
DPBS	Dulbecco's Phosphate Buffered Saline
dpc	Days post conception
EDTA	Ethylenediaminetetraacetic acid
ESRD	End-stage renal disease
FACS	Fluoresence activated cell sorting
FB	Factor B

FBS	Fetal bovine serum
FCN1	Ficolin-1
FH	Factor H
FI	Factor I
GAG	Glycosaminoglycan
GATK	Genome analysis toolkit
GWAS	Genome-wide association study
h	hours
HELLP	Hemolysis, elevated liver function tests, low platelets
HRP	Horseradish peroxidase
iC3b	Inactivated C3b (by factor I)
kb	kilobase
LP	Lectin pathway
mAb	monoclonal Antibody
MAC	Membrane attack Complex
MAF	Minor allele frequency
MASP-2	MBL associated serine protease 2
Mb	megabase
MBL	Mannose binding lectin
MCP	Membrane Cofactor Protein
MG	Macroglobulin domain
min	minutes
MPGN	membranoproliferative glomerularnephritis
NHLBI ESP	National Heart Lung and Blood Institute Exome Sequening Project
OR	Odds ratio
P	Properdin
PCR	Polymerase chain reaction
PE	Preeclampsia
PlGf	Placental growth factor
qPCR	quantitative PCR
RBC	Red blood cell
RVCD	Rare variant common disease hypothesis
sFlt-1	soluble fms-like tyrosine kinase 1 (VEGF Receptor)
SNP	single nucleotide polymorphism
SPR	Surface plasmon resonance
TBS	Tris buffered saline
TED	Thioester domain
TF	Tissue factor
TMA	Thrombotic microangiopathy

VEGFA	Vascular endothelial growth factor A
VWF	von Willebrand Factor
WES	Whole exome sequencing
WGS	Whole genome sequencing
WT	Wild type

Acknowledgements

I owe a great deal to my mentor, John Atkinson, for all of his support. He took me on as a student while I was working on the effects of patient mutations in complement regulators. He supported my interests in genetics as they went far afield of his previous experience and we learned a lot together. I never doubted that he believed in me when I was not sure I believed in myself. John has also been an excellent role model. I am not sure I have ever met someone so generous with their time. He sets a great example of what it means to be a good person.

I would like to thank everyone in the Atkinson group as well. Lorraine Schwartz and Madonna Bogacki made seemingly endless administrative tasks tolerable. Kathy Liszewski always had an open office door: always ready to write another grant or listen about a failed experiment. I would like to thank Paula Bertram too for taking the plunge into genomics and always making time to help. Richard Hauhart also provided invaluable assistance, and usually with only a moments notice. His attitude is inspiring.

Parul Kothari, Elizabeth Schramm, and I grew to be good friends over the years and I appreciate all of the time they spent listening about topics they have very little interest in. I probably should have done more of the listening. I would also like to thank Elizabeth Schramm for the collaborative spirit during the last six months of my PhD; it was a blessing.

My thesis committee offered continual guidance and support. I would especially like to thank Tim Ley for helping me stay focused on the path and not getting lost in the forest. I would

Dedicated to the memory of my mother.

There was never a more selfless and gentle person.

None of this would have been possible without her.

She is loved and missed daily.

ABSTRACT OF THE DISSERTATION
Excessive Complement Activation Due to Genetic Haploinsufficiency
of Regulators in Multiple Human Diseases
by
Michael Triebwasser
Doctor of Philosophy in Biology and Biomedical Sciences
Human and Statistical Genetics
Washington University in St. Louis, 2015
Professor John P. Atkinson, Chairperson

The complement system is an ancient and powerful form of innate immunity. The

alternative pathway (AP), a positive feedback loop, is at the core of the complement system.

Activating components and regulators of the AP are genetically implicated in atypical hemolytic

uremic syndrome (aHUS) and age-related macular degeneration (AMD). aHUS features kidney

failure, and often affects young children, but may occur throughout life and can be precipitated

by pregnancy. aHUS associated variants are extremely rare and are considered highly penetrant.

At the opposite end of the spectrum, AMD affects the retina leading to loss of central vision with

a late age of onset. Risk variants in AMD are common in the population and have smaller effect

sizes.

I endeavored to understand the role of rare variants of large effect, similar to those

causative in aHUS, in common diseases involving the kidney, specifically preeclampsia and

lupus nephritis. I also examined the role of such variants in AMD. Because thousands of people

must be studied to assess the impact of rare variation, I developed novel approaches allowing

these experiments to be done with 10- to 100-fold reductions in both cost and labor.

aHUS-like variants are present in preeclampsia, a syndrome that affects pregnant women

and shares multiple pathologic findings with aHUS. These variants are present in ~1% of

preeclamptic individuals. Additionally, aHUS-like rare variants are found in severe AMD

patients. A diversity of variants in factor H and factor I are enriched in AMD cases. Subsets of these alleles have high penetrance in families and defective function.

To study the effect of unregulated AP activation *in vivo* I studied a mouse deficient for the ubiquitous membrane regulator of complement Crry. Embryos that lack Crry are not viable due to attack by the maternal AP early in development. Damage in this model is unrelated to traditional forms of complement-mediated inflammation such as neutrophil activation or anaphylatoxin signaling.

The findings of this body of work are an important step forward in understanding the risk individuals have of the common diseases PE and AMD and has implications for how these patients could be treated.

Chapter 1

Introduction

Complement System

The complement system is an ancient component of innate immunity. In humans, its most appreciated role is in defense against infection. It opsonizes targets for phagocytosis and can lyse them as well. Activation of the pathway is initiated through three arms: the classical (CP), lectin pathway (LP) and the alternative (AP) (Figure 1-1). Each pathway is highly regulated and controlled, as excessive activation will lead to disease (1). C3 activation and the terminal pathway (C5-C9) are common to each cascade.

Activation

The classical pathway is antibody-dependent and activation is highly coordinated. It begins when C1q in the C1 complex (C1q, C1r, C1s) is engaged by IgM or IgG subclasses 1 and 3 bound to a target. Upon Fc receptor binding by C1q, the pro-enzyme C1r autoactivates and cleaves C1s (also a proenzyme) to form active C1ś. C1ś then cleaves C4 to generate C4a and C4b. This exposes a reactive thioester in C4b, allowing C4b to covalently attach to a target. When C4b binds C2, C1s cleaves the proenzyme C2, leading to the active enzymatic fragment C2a. The complex of C4bC2a forms the CP C3 convertase.
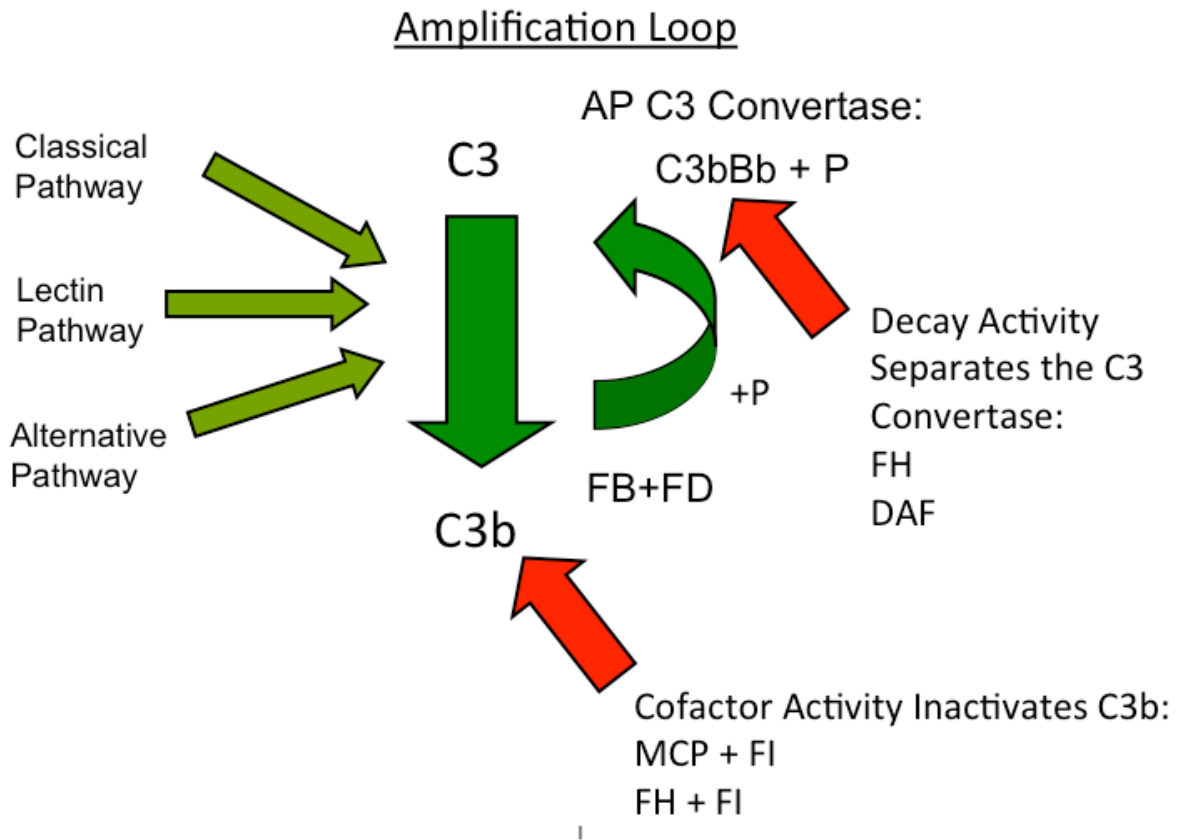
Figure 1-1. Alternative Pathway (AP) Amplification Loop.

The AP, CP, and LP converge on the cleavage of C3 to C3b (Figure 1-2). C3b engages in the

amplification loop by factor B and forming a new AP C3 convertase. Properdin can bind the AP

C3 convertase and stabilize it. The AP C3 convertase cleaves additional C3 to C3b, thus feeding

forward. Negative regulation is critical and occurs at the level of the C3b and the C3 convertase.

Cofactor activity cleaves and inactivates C3b, preventing it from binding FB. Decay acceleration

accelerates the natural decay of C3bBb to its two components C3b and Bb. C3b can still bind FB

again and form a new convertase.

This complex cleaves C3 to C3b, generating the anaphylatoxin C3a and C3b. C3b's thioester, like that in C4b, is also exposed upon cleavage, which leads to covalent binding of the majority of C3b to the target and some to the C4bC2a complex.

The combination of C3bC4bC2a is the classical pathway C5 convertase, which cleaves C5 to C5b and C5a, a potent anaphylatoxin. C5b binds C6 and then C7; C5b-7 deposits on cell membranes. C8 then binds the trimeric complex followed by multiple copies of C9 to form a pore. The C5b-9 complex is known as the membrane attack complex (MAC). MAC deposition leads to lysis or, in the case of sublytic attack, membrane perturbation (2-4).

The LP of complement is initiated when mannose binding lectin (MBL) or a ficolin protein (both analogous to C1q) bind a target and activate mannose activated serine proteases (MASP) -2 (analogous to C1s) (5). MASP-2 then cleaves and activates C4 and C2, and at this point the CP and LP share the same C3 and C5 convertases. Instead of binding antibody like C1q, MBL recognizes sugar moieties on pathogens rather than specific antigens.

In contrast to the CP and LP, where activation is targeted through adaptive and innate immunity, the AP is not activated specifically. The AP also proceeds through an ordered stepwise activation, but it is unique from the CP and LP in that it consists of a positive feedback loop. AP activation relies on the continuous "tickover" of C3. The thioester of C3 is inherently unstable and approximately 1-2% of C3 turns over every hour (6). Much of this will bind water, forming C3(H_2O). But should the thioester bond hydrolyze and bind a sugar, protein or phospholipid on the surface of a cell or piece of debris, it will covalently attach (sometimes called C3-Target, C3-T). Upon hydrolysis of its thioester, C3 undergoes a massive conformational change, which allows factor B to bind (Figure 1-3). When B is bound, factor D

cleaves the proenzyme B to the active enzyme Bb. C3-TBb is the AP C3 convertase and cleaves C3 to C3b and C3a.

C3(H$_2$O), C3-T and C3b all share the same conformation with the thioester exposed and are able to bind FB and form a new C3 convertase (the positive feedback loop). In exposing C3b's thioester, it can attach either to a surface (target) or the C3bBb convertase to form the AP C5 convertase. Additionally, properdin (P) binds the C3 and C5 convertases dramatically stabilizing them. C3bBb has a half-life of 2 to 3 minutes, whereas properdin increases this ten-fold (7). The AP is a positive feedback loop capable of exponentially increasing complement activation on a target.
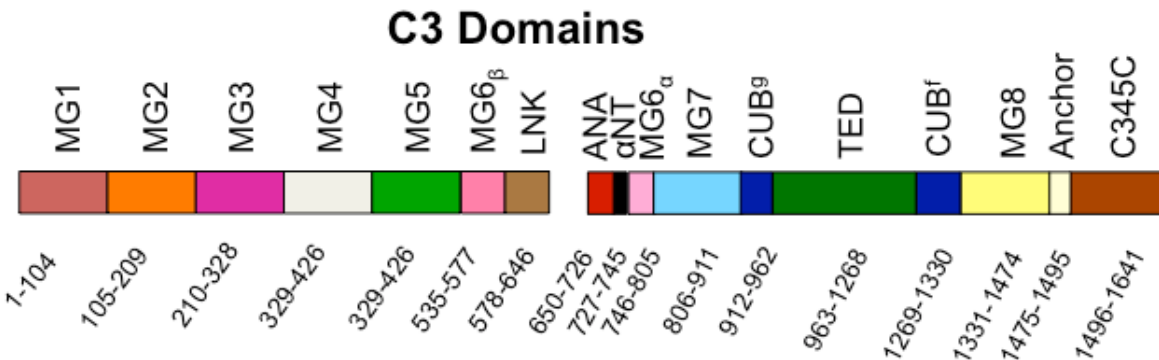
C3 Domains

Figure 1-2. Linear diagram of the C3 protein.
The β chain is comprised of MG 1-6 (left side) and the α chain is comprised of the ANA, αNT, part of MG6, MG7, CUB, TED, MG8 and C345C domains. The TED domain is where the cysteine of the thioester bond is contained. In the presence of a cofactor, Factor I cleaves in the CUB domain converting C3b to iC3b. Cleavage of C3 to C3b removes the ANA domain. The β and α chain are joined by a disulfide bond.
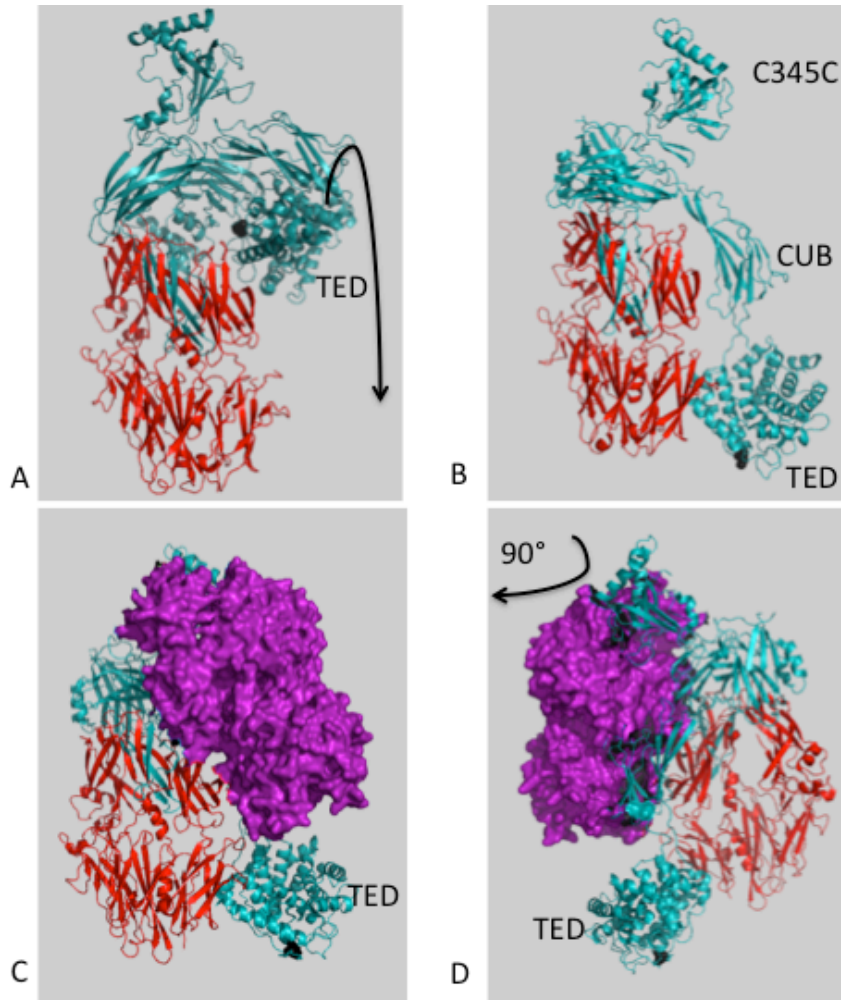
Figure 1-3. C3 Changes Conformation With the Conversion to C3b.
A. C3 protein with the β chain (red) and α chain (teal) with the thioester in black. PDB ID 2A73.
B. The TED domain (C3d fragment site of the thioester) of C3b radically changes position upon cleavage of C3 to C3a and C3b. C3-T, C3(H_2O), and C3b all share the same conformation. PDB ID 2I07. C. Factor B (purple) bound to C3b. Factor B makes contacts to C3b at the CUB domain, C345C domain and some of the MG (macroglobulin domains). PDB ID 2XWJ.

Regulation of Complement Activation

Complement activation is subject to negative regulation particularly at the level of the C3

and C5 convertases, both through accelerating the natural decay of the enzyme complexes (decay

acceleration activity, DAA) and through proteolytic cleavage of C4b and C3b by the serum

protease factor I to prevent formation of additional C3 convertases (Figure 1-4) (8). Host cells

6

express negative regulators on their surface to perform both roles. Decay acceleration activity

(DAA) (define) is provided by the GPI-anchored protein decay accelerating factor (DAF, CD55)

(Figure 1-4). On the other hand, cleavage by factor I in concert with an absolutely required

cofactor protein (Figure 1-4B). On the cell surface, the most widely distributed protein with

cofactor activity (CA) is membrane cofactor protein (MCP, CD46).
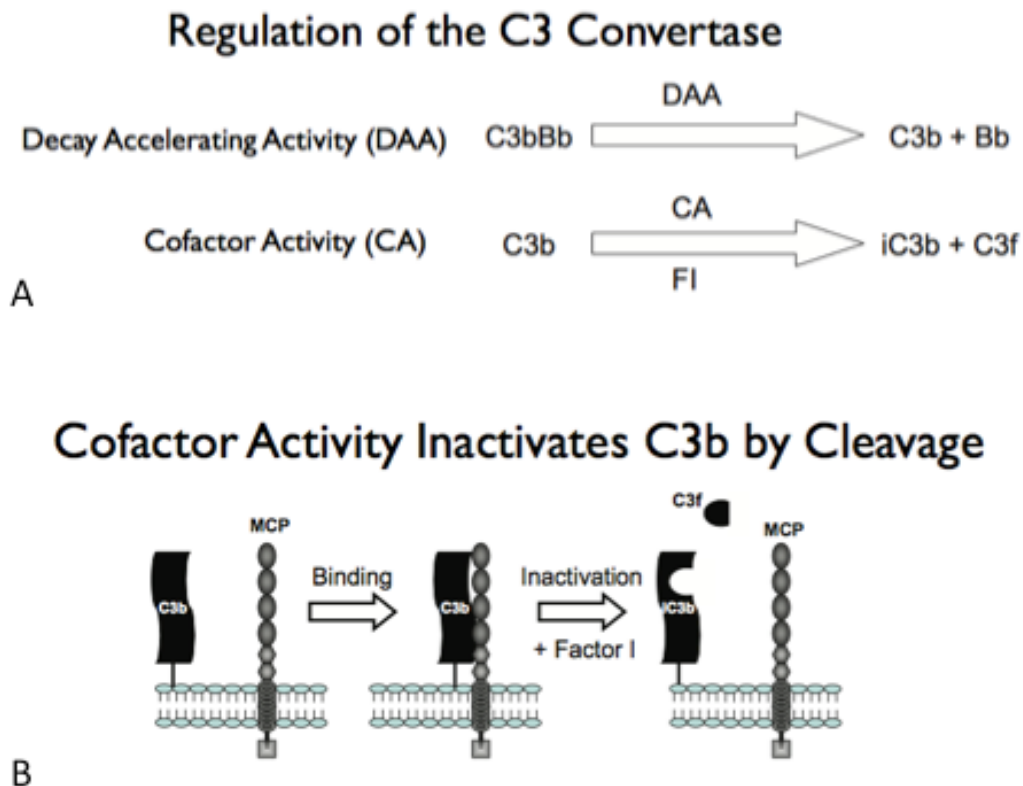
Figure 1-4. Negative Regulation of the AP Occurs at the Level of the C3 Convertase
A. Decay acceleration separates the C3bBb into C3b and Bb. Bb cannot rebind to C3b. A new
FB must rebind C3b and factor D to reform the C3 convertase. Cofactor activity allows the
protease FI to cleave C3b to iC3b and a small fragment C3f. iC3b cannot rebind factor B and
thus cannot form a new C3 convertase. C3f is not covalently linked to iC3b. B. A diagram
illustrating how MCP (or FH) participates in cofactor mediated cleavage on the cell surface.

A more limited variety of cells also express complement receptor 1 (CR1, CD35), which contains sites for both decay and cofactor activity. Lastly, on the cell membrane is CD59, a GPI-anchored protein that inhibits C8 and C9 pore formation.

In blood, factor H (FH) carries both DAA and CA for the AP C3 convertase. FH functions in the fluid phase by limiting the ability of $C3(H_2O)$ to form the C3 convertase. It is also able to deposit on surfaces with exposed glycosaminoglycans (GAGs) as well at previous sites of complement activation (C3b/C3d) to limit AP activity on surfaces and debris (9) (Figure 1-5).

Factor H, DAF, MCP and CR1 are all comprised largely or entirely of complement control protein repeats domains (CCP, sushi domains). CCPs are approximately 60 amino acids and have a conserved disulfide structure: the first cysteine bonds to the third, and the second cysteine to the forth (10). MCP, DAF and FH all contain their regulatory activity in the first 4 CCPs (MCP and DAF each only have 4 CCPs). CR1 has three separate regulatory sites, but their organization is similar.

Factor H has 20 CCPs. Besides the complement regulatory activity of CCP 1-4 (Figure 1-6), CCP7 and 8 comprise a heparin (anionic) binding domain, and CCPs 19-20 have both heparin binding and C3b and C3d binding activity (Figure 1-5). In the case of CR1 and factor H, the CCPs between known functional sites may be conserved in evolution to allow optimal spacing.
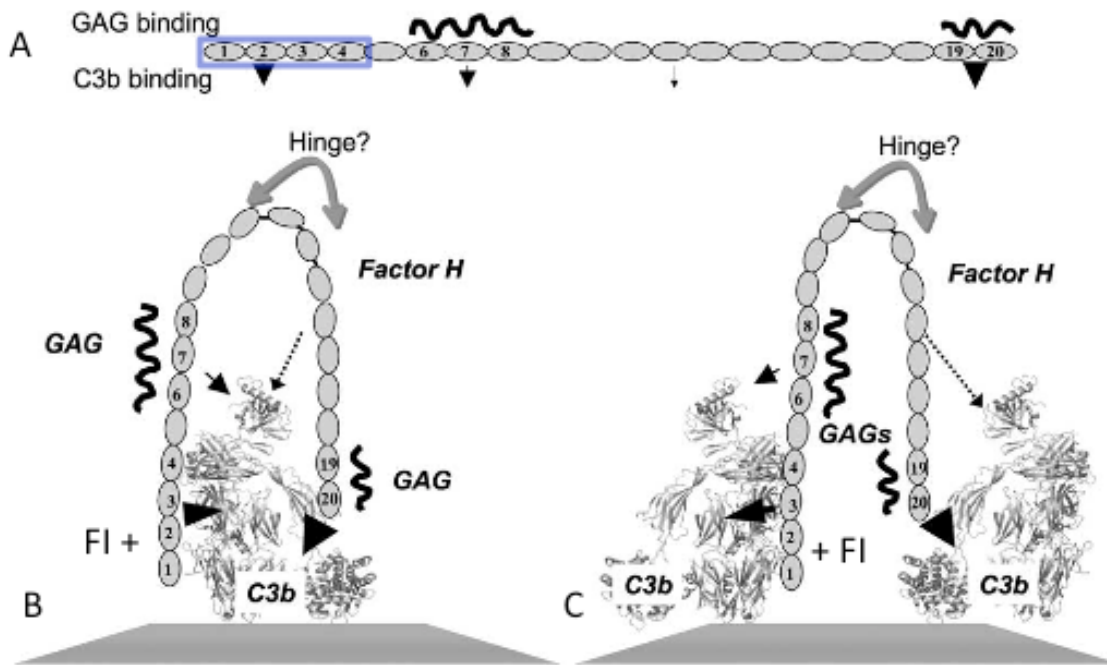
Figure 1-5. FH's Domains Allows it to Bind GAGs and C3b/C3d on Surfaces
A. FH's structure contains regulatory activity in CCP1-4 (blue box). GAG binding domains (squiggle above FH) present in CCP6-8 and 19-20. C3b binding (arrow below FH). Shows sites with evidence for C3b binding. Arrow size is proportional to the binding affinity. B. FH may bind and regulate the same C3b molecule through CCP19-20 binding of the TED domain, while the CCP1-4 domain can perform DAA and CA. C. FH may also bind to a surface with CCP19-20 by binding either the C3d scar (TED domain) that will permanently tag a site of complement activation or though GAGs present on the cell or extracellular matrix. FI would interact with the FH:C3b complex to perform CA. Figure adapted from Schmidt, CQ, et al. J. Immunology, 2008, 181: 2610-2619
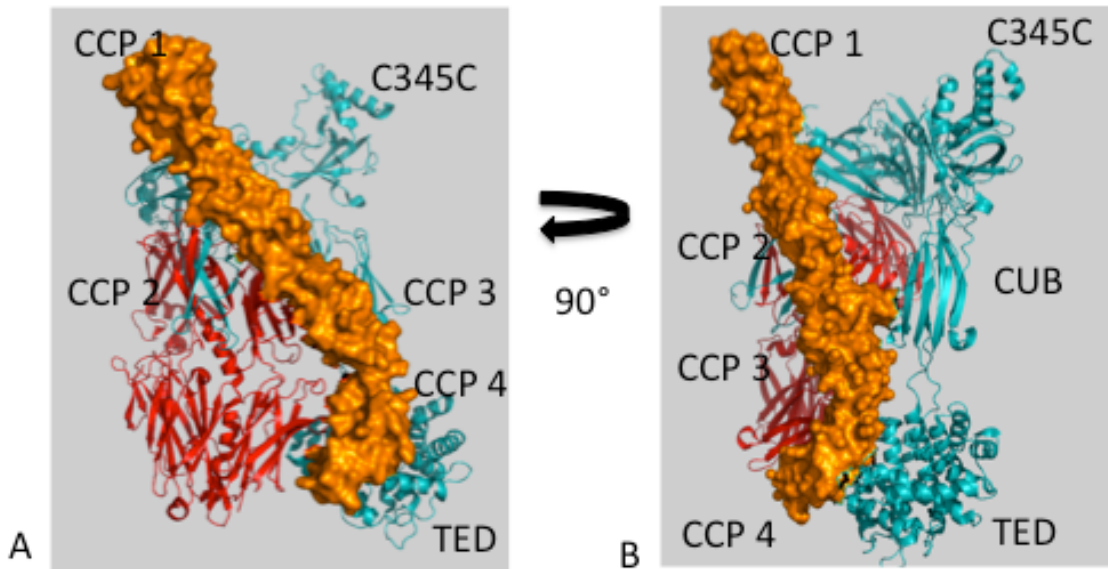
Figure 1-6. FH Binds in a Groove over the CUB and TED Domain of C3b
A. Factor H CCP1 to 4 (orange) binds to both the β chain (red) and α chain (teal) of C3b. CCP2-3 make the majority of contacts with C3b, including the CUB domain and TED domain. B0. Rotated 90°. PDB ID 2WII.

C3b is cleaved to iC3b by FI in concert with MCP, FH, or CR1. CR1, complement receptor 2 (CD21), and complement receptor 3 (CR3/integrin M/CD11b) all have a high affinity for iC3b (11). CR1 and FI can cleave iC3b further to C3dg, which remains covalently bound to the target, and can be bound by CR2, and complement receptor 4 (CR4/integrin X/CD11b) (12, 13). These degradation products of C3b are an "immunologic scar" and lead to adherence and signaling in those cell types expressing the requisite receptors. For example, B cells express CR2, and without CR2 signaling mice have dramatically reduced adaptive immune responses (14). CR3's recognition of C3b is one of the main mechanisms by which phagocytic cells traffic into areas of inflammation and by which they phagocytose opsonized targets (15).

Rare Variation in Human Disease

Rare variation has previously been assumed to be the purview of devastating Mendelian disorders. Common diseases were postulated to be the result of nearly neutral alleles with small effect sizes that have risen to high frequencies in the population though a mixture of migration, selection, and, largely, genetic drift (Common Disease Common Variant hypothesis) (16). After increasingly large genome wide association studies (GWAS), not all of the heritability of common disease is fully explained by these common SNPs, many of which have odds ratios on the order of 1.1 - 1.5 (17).

Heritability estimates may be overstated as many of these estimates were based on twin studies, not on large populations (18). Additionally, interactions between the common SNPs may yield non-additive effects; however, detecting these is problematic because of the burden imposed by testing so many hypothesizes. Ever larger GWAS studies will likely uncover alleles with increasingly small effect size, further capturing the spectrum of genetic risk. It is possible that genetic interactions with the environment, and in the case of common autoimmune diseases, pathogens such as viruses, may explain some of the risk. Another possible source of genetic risk is a diversity of rare variants with relatively large effects on disease risk (19).

New mutations enter the population at every cell division. The empirically determined mutation rate is ~1x10-8 per meiosis (20). The implication of this is that every person has ~60 de novo point mutations at birth, along with structural variations (copy number changes), the frequency of which is harder to estimate. If a population is of a fixed size, selection will remove even mildly deleterious alleles while neutral ones will be fixed or lost due to genetic drift (i.e., random chance) (16). The human population has been rapidly expanding for 5-10,000 years

(roughly coinciding with the advent of agriculture), which obfuscates the ability of natural selection to exert strong pressure on deleterious alleles and remove them from the population (21, 22). Rare mutations are more likely to have arisen recently, whereas common SNPs are likely to have arisen a long time ago (thousands versus hundreds of thousands of years).

Rare variants, by their low frequency, are likely to be detrimental (23). We propose here to look for such mutations that have a relatively large effect size compared to the majority of those identified recently by GWAS. This paradigm, the Rare Variant Common Disease (RVCD) hypothesis, is proving to be of significant interest for at least several forms of common diseases including cardiovascular disease, AMD and Alzheimer's.

Estimates of the study sizes required to detect rare variants tell us that large sample sizes will be required. Early estimates were that 660 individuals (half cases, half controls) would be required to detect an association between a gene that carried rare variants in 1% of controls and for which rare variants increase the risk of disease eight-fold (24). This is an optimistic assumption as many genes have more than 1% of individuals carrying a rare variant (<0.5% MAF). The fraction of people carrying a rare variant will generally increase with gene size. More importantly, if some variants have an effect this large those variants would likely be observed segregating with disease in pedigrees in Mendelian or near-Mendelian modes of inheritance (25). Based on recent studies, rare variation in a gene is likely to only increase the rate of mutations in cases 2 to 3-fold above that seen in the control group (26, 27). This will necessitate studies comprised of thousands to tens of thousands of individuals.

Atypical Hemolytic Uremic Syndrome

Atypical hemolytic uremic syndrome (aHUS) is a thrombotic microangiopathy (TMA) characterized by glomerular endotheliosis and formation of clots throughout the microvasculature, with a preference for the glomerulus of the kidney (28, 29). In addition to renal failure, cases exhibit refractory hypertension, platelet consumption and a hemolytic anemia.

Over the past decade, aHUS has been shown to be primarily caused by loss of function mutations in the negative regulators of complement activation. This is in contrast to the more common shiga-toxin associated HUS, or diarrheal HUS, where exposure to a Shiga toxin expressing bacteria, often E. coli, precipitates the TMA. In D-HUS the Shiga toxin itself damages the endothelial cells lining the glomerulus (30, 31).

Initially, a small number of pedigrees showed linkage to FH and MCP (32-34). Subsequent sequencing of FI, C3 and FB revealed rare highly penetrant mutations in both familial and sporadic aHUS (35-37). Mutations in the negative regulators abrogate binding or cofactor activity in relation to C3b, and mutations in C3 and FB lead to escape from interactions with negative regulators or increased stability of the AP convertase. Thus, the common mechanism in aHUS is loss of proper regulation of the AP leading to excessive activation. Such mutations are seen in 40-60% of aHUS patients (38). Autoantibodies to FH that either deplete or impair function are seen in another ~10% of aHUS patients (39). Surprisingly, no mutations in DAF have led to loss of function (40). CA appears to be the most crucial mode of AP regulation. It fundamentally stops the AP, whereas DAA only slows it down as the C3b remains, ready to generate a new C3 convertase.

The degree of haploinsufficiency that is sufficient to cause disease is striking: individuals heterozygous for a mutation may have one functional allele of MCP and one allele with 50% activity, leading to 75% of the regulatory activity of a sibling that does not carry the risk allele; however, 75% activity is not enough, and they are at risk of developing aHUS (41).

Mutations that cause aHUS act in a dominant manner with incomplete (~50% penetrance) (42, 43). Often children are reported to have had a viral infection immediately prior to an episode and pregnancy is also a trigger (44, 45). If a mutation carrier does not encounter an exposure while young, they may graduate into lower risk adulthood. A "stressor" may be required to trigger an aHUS episode, similar to the Shiga toxin that precipitates D+HUS.

Until recently, therapy for aHUS involved plasma exchange, with the thought that one could replace the defective factor H or factor I of the host with that of a healthy donor. Renal transplants are often required due to kidney damage, and only in individuals with MCP mutations is this curative because the defect, haploinsufficiency of MCP on glomerular endothelial cells, is corrected. Based in part on the finding that mice deficient for C5 are rescued from an aHUS like disease (46), there is now a mAb to C5 that prevents cleavage, which is effective in most aHUS cases. This proved that C5 cleavage triggered by the AP is critical to developing microthrombi (47).

Alternative Pathway Activation is Pathologic to the Kidney

Excessive AP activation is central to a variety of diseases of the kidney. This is illustrated by the central role of dysregulated complement activation in not only aHUS but also membranoproliferative glomerular nephritis type II (MPGN II, also known as Dense Deposit Disease, DDD), C3 glomerulonephropathies, and mesangioproliferative glomerulonephritis (48).

DDD is characterized by progressive loss of kidney function that primarily affects children between the age of 5 and 15 and generally leads to end stage renal disease in 10 years (49). It is characterized by subendothelial deposits, which contain C3b (50).

More than 80% of patients have an autoantibody that stabilizes the AP C3 convertase, termed C3 nephritic factor (C3NeF) (49). In ~10% of DDD patients, there are mutations, often recessive, in factor H (51). The end result of C3NeF and the mutations are the same, constant activation of the AP and turnover of C3. Other kidney diseases also feature C3NeF and varying rates of AP gene mutations.

In contrast to aHUS, FH mutations in MPGN II are largely in the N-terminal regulatory domain in contrast to the C-terminal localization domain. Such a mutation would lead to excessive amounts of C3b being produced in the fluid phase, like a C3NeF, and being deposited in the kidney. The majority of aHUS mutations in FH are in the C-terminal domain that localizes FH to surfaces, leading to only excessive local activation (Figure 1-5).

It may be the amount of this C3b that is depositing in the kidney in MPGN II, not necessarily the local activation occurring on exposed basement membranes or damaged cells as in aHUS. Some mutations are shared between aHUS and DDD and showcase the variability with which AP mutations can present.

The kidney may be extremely sensitive to complement activation because of the fenestrated endothelium of the glomerulus. This structure allows filtration, but it also puts the blood in direct contact with the basement membrane, which, unlike cells, does not contain intrinsic negative regulators of complement such as MCP and DAF.

Preeclampsia is a Common TMA

Preeclampsia (PE), a pregnancy specific syndrome, is the leading cause of maternal and fetal

mortality and morbidity globally (52). Diagnosis is based on the new onset of hypertension and

proteinuria in the latter half of gestation (53). Glomerular endothelial cells swell occluding the

vessel lumen, and clots are seen in the kidney and placenta (54). PE occurs in 3 to 8% of first

pregnancies and is responsible for ~15% of preterm births, largely because delivery is the only

cure (55). A unifying histopathological finding is shallow invasion of the uterine spiral arteries

by the placenta and failure to remodel them into high flow, low resistance vessels.

Human biomarker studies and other work have led to the model that the placenta, in

reaction to this abnormal vascular development, elaborates anti-angiogenic and nephrotoxic

molecules that cause maternal decompensation. (56-59). Specifically, multiple studies have

found low levels of placental growth factor (PlGF) and high levels of the soluble VEGF (and

PlGF) receptor 1 (Flt1) are correlated with the development of PE. This soluble receptor opposes

the VEGF-dependent glomerular endothelial cells, damaging them. The kidney pathology of PE

mimics an adverse reaction to anti-VEGF antibodies used in oncology (60).

Fragments generated by AP activation are also elevated early in pregnancy in those who

will later develop PE (61) suggesting that even from a very early point in pregnancy there is

more AP activation at the feto-maternal interface (Figure 1-7). In this model, complement

activation occurs and exacerbates the initial effects of sFlt-1 (62).

There is an incidence of aHUS associated with pregnancy, likely due to the vascular stress

imposed by the fetus on the maternal vasculature and specifically the kidney. Individuals with

pregnancy-associated aHUS often have rare mutations in complement regulatory proteins (63).

Four out of eleven patients with hemolysis, elevated liver enzymes, and low platelets (HELLP), a serious complication of PE, have been reported to have complement regulatory mutations (64). We observed rare mutations in PE patients that result in uncontrolled complement activation AP (65); one had been previously described in two aHUS patients that presented in the first year of life (66).
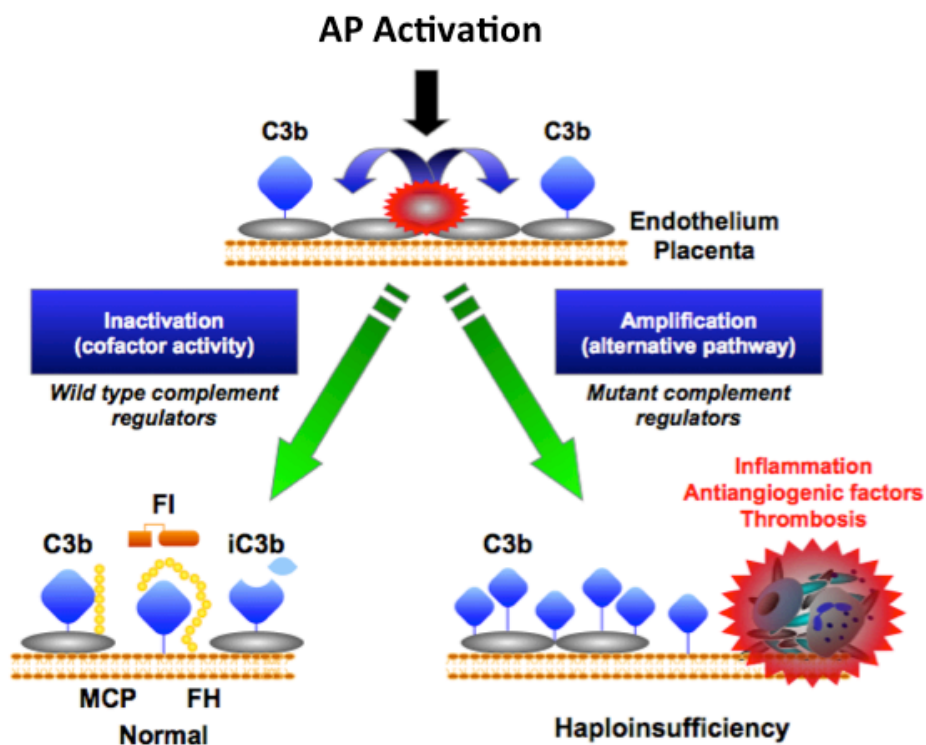
Figure 1-7. Complement Activation is a Shared Feature of aHUS and PE.
The AP activates on damaged or stressed endothelium and/or placenta cells. In individuals with normal complement regulators AP activation is controlled. In those with mutant regulators, AP activation amplifies leading to inflammation, anti-angiogenic factors and thrombosis. In our model these effects are involved in the syndrome of preeclampsia. Figure adapted from Salmon, JE, Heuser, C, Triebwasser, MP, et al. PLoS Medicine, 2011, 8(3): e1001013.

AP Genes Contribute Substantially to AMD

Age-related macular degeneration (AMD) AMD is a common disease of the retina and features loss of central vision. This area, called the macula, is responsible for the high resolution we have in the center of our field of view. The characteristic finding is the formation of proteinaceous deposits, known as drusen, between the retina and the supporting tissue behind. Damage to the retina occurs either because blood vessels from behind the retina proliferate and damage the retina (wet form) or because of loss of retinal cells in areas of deposition (geographic atrophy, dry form).

After efforts to find the cause of AMD through linkage, the best candidate region was chosen for an association study. Using 44 SNPs in this region spanning 2.2 Mb, the factor H gene, CFH, was found to harbor a missense variant Y402H, which increases the risk of AMD (odds ratio=2.5) (67). This common variant is present in 30-40% of controls (and 60-65% of cases); it was estimated to explain ~40% of the disease burden in AMD (68).

Functional studies of this part of FH and of the H402 allele have shown that it affects the ability of the protein to bind to glycosaminoglycans and oxidized lipids like those that might be present basement membranes or damaged cells (69, 70). The risk variant may have experienced positive selection because it reduces binding of FH to some S. pyogenes proteins, which would lead to increased complement activation on them and decreased virulence in vivo (71).

Subsequent genome-wide association studies in AMD have identified common variants that increase the risk of AMD in CFI, CFB, and C3 (72-74). Common SNPs account for ~50% of the risk in AMD (75). AP activation is strongly implicated by genetics in the pathophysiology of AMD. The idea of a person's complotype, their genotype at all of the AMD risk SNPs in the AP,

has arisen. The greater the burden of AMD risk SNPs in a complotype, the higher the AP activity, even when each variant separately has only modest effects on function (76)

In addition to these common variants that have a small effect size, a single rare variant in CFH has been reported in aHUS and also associated with AMD (25). The R1210C variant in FH leads to a diminished heparin binding activity similar to the common Y402H allele (70, 77, 78). The inability to bind correctly leads to a local deficiency in FH activity, even though serum levels are normal.

AMD has no robust therapies currently. Anti-VEGF mAbs are able to stop, and in some cases reverse, the disease seen in wet AMD. Therapeutics that modulate complement are a promising therapy, and an anti-factor D antibody therapy recently showed benefit in a phase II trial of AMD patients with geographic atrophy. The largest predictor of response was the common factor I risk allele (Roche Press Release, August 27th, 2013).

Complement Regulator Deficient Mouse Models of Disease

Mice deficient for FH develop glomerular nephritis similar to DDD (79). Due to unregulated fluid phase consumption, these animals have profoundly low C3 and C5 levels; Neutrophils are present around the glomerulus (80). Further the development of renal disease is C5, but not C5b-9 dependent (80). A mouse model of aHUS utilizes a FH allele with CCP 16-20 deleted. These mice have higher C3 levels in the blood than do FH$^{-/-}$ animals (81), and the development of kidney disease is also C5 dependent (46).

The ubiquitously expressed membrane bound regulator with cofactor activity is Crry in the mouse. It is analogous to MCP in humans. Deficiency leads to embryonic lethality (82), and loss is dependent on the AP (C3 and FB), but not the CP or C5b-C9 (83, 84). Maternal AP

attacks the developing embryo, and if the mother is FB or C3 deficient, it is possible to generate

Crry$^{-/-}$ mice (85, 86). Crry$^{-/-}$ mice have low levels of C3, but higher than FH$^{-/-}$ animals, and do not

develop renal disease (84, 85). C5 deficiency had a minor effect on the rate of embryonic loss.

This model is unique in that is not entirely C5 dependent and the mechanism linking AP

activation to loss is not understood.

Specific Aims

This thesis has two focuses. The first is to investigate a mouse model of AP activation in

which the main regulator that possesses CA on the surface of mouse cells, Crry, has been deleted.

One effect of this is unrestricted activation by the maternal AP on Crry$^{-/-}$ embryos and fetal loss

(82). The goal was to further understand the components of the AP involved, what part of the AP

is mediating loss, and the impact of AP activation on development.

The second focus was to assess the effect of rare variants in AP genes on a variety of

human diseases, including AMD, PE, lupus nephritis and diarrheal HUS. To address this second

component, significant technical innovations had to be made to allow large scale sequencing of

the complement system, and related pathways like coagulation, in large enough patient

populations to yield sufficient statistical power.

References

1   Java, A., J. Atkinson, and J. Salmon. 2013. Defective complement inhibitory function predisposes to renal disease. Annu. Rev. Med. 64:307-324.

2   Triantafilou, K., T. R. Hughes, M. Triantafilou, and B. P. Morgan. 2013. The complement membrane attack complex triggers intracellular Ca2+ fluxes leading to NLRP3 inflammasome activation. J. Cell Sci. 126:2903-2913.

3   Lueck, K., S. Wasmuth, J. Williams, T. R. Hughes, B. P. Morgan, A. Lommatzsch, J. Greenwood, S. E. Moss, and D. Pauleikhoff. 2011. Sub-lytic C5b-9 induces functional changes in retinal pigment epithelial cells consistent with age-related macular degeneration. Eye (Lond) 25:1074-1082.

4   Tedesco, F., M. Pausa, E. Nardon, M. Introna, A. Mantovani, and A. Dobrina. 1997. The cytolytically inactive terminal complement complex activates endothelial cells to express adhesion molecules and tissue factor procoagulant activity. J. Exp. Med. 185:1619-1627.

5   Fujita, T. 2002. Evolution of the lectin-complement pathway and its role in innate immunity. Nature Reviews. Immunology 2:346-353.

6   Charlesworth, J. A., D. G. Williams, E. Sherington, P. J. Lachmann, and D. K. Peters. 1974. Metabolic studies of the third component of complement and the glycine-rich beta glycoprotein in patients with hypocomplementemia. J. Clin. Invest. 53:1578-1587.

7   Hourcade, D. E. 2006. The role of properdin in the assembly of the alternative pathway C3 convertases of complement. J. Biol. Chem. 281:2128-2132.

8   Lublin, D. M., and J. P. Atkinson. 1989. Decay-accelerating factor and membrane cofactor protein. Curr. Top. Microbiol. Immunol. 153:123-145.

9   Schmidt, C. Q., A. P. Herbert, D. Kavanagh, C. Gandy, C. J. Fenton, B. S. Blaum, M. Lyon, D. Uhrin, and P. N. Barlow. 2008. A new map of glycosaminoglycan and C3b binding sites on factor H. J. Immunol. 181:2610-2619.

10   Ault, B. H., B. Z. Schmidt, N. L. Fowler, C. E. Kashtan, A. E. Ahmed, B. A. Vogt, and H. R. Colten. 1997. Human factor H deficiency. Mutations in framework cysteine residues and block in H protein secretion and intracellular catabolism. J. Biol. Chem. 272:25168-25175.

11   Arnaout, M. A., M. Pierce, N. Dana, and L. K. Clayton. 1987. Complement receptor type 3 (CR3): structure and function. Meth.Enzymol. 150:602-615.

12   Szakonyi, G., J. M. Guthridge, D. Li, K. Young, V. M. Holers, and X. S. Chen. 2001. Structure of complement receptor 2 in complex with its C3d ligand. Science 292:1725-1728.

13   Law, A. 2000. CR4. In The Complement FactsBook. B. J. Morley, and M. J. Walport, eds. Academic Press, San Diego. 198-203.

14   Molina, H., V. M. Holers, B. Li, Y.-F. Fang, S. Mariathasan, J. Goellner, J. Strauss-Schoenberger, R. W. Karr, and D. D. Chaplin. 1996. Markedly impaired humoral immune response in mice deficient in complement receptors 1 and 2. Proc. Natl. Acad. Sci. USA 93:3357-3361.

15   Fiane, A. E., T. E. Mollnes, V. Videm, T. Hovig, K. Hogasen, O. J. Mellbye, L. Spruce, W. T. Moore, A. Sahu, and J. D. Lambris. 1999. Compstatin, a peptide inhibitor of C3, prolongs survival of ex vivo-perfused porcine xenografts. Xenotransplantation.

16   Reich, D. E., and E. S. Lander. 2001. On the allelic spectrum of human disease. Trends Genet. 17:502-510.

17   Manolio, T. A. Genomewide association studies and assessment of the risk of disease. N. Engl. J. Med. 363:166-176.

18   Zuk, O., E. Hechter, S. R. Sunyaev, and E. S. Lander. 2012. The mystery of missing heritability: Genetic interactions create phantom heritability. Proc. Natl. Acad. Sci. USA 109:1193-1198.

19   Pritchard, J. K. 2001. Are rare variants responsible for susceptibility to complex diseases? Am. J. Hum. Genet. 69:124-137.

20   Roach, J. C., G. Glusman, A. F. Smit, C. D. Huff, R. Hubley, P. T. Shannon, L. Rowen, K. P. Pant, N. Goodman, M. Bamshad, J. Shendure, R. Drmanac, L. B. Jorde, L. Hood, and D. J. Galas. 2010. Analysis of Genetic Inheritance in a Family Quartet by Whole-Genome Sequencing. Science 328:636-639.

21   Fu, W., T. D. O'Connor, G. Jun, H. M. Kang, G. Abecasis, S. M. Leal, S. Gabriel, M. J. Rieder, D. Altshuler, J. Shendure, D. A. Nickerson, M. J. Bamshad, and J. M. Akey. 2013. Analysis of 6,515 exomes reveals the recent origin of most human protein-coding variants. Nature 493:216-220.

22   Tennessen, J. A., A. W. Bigham, T. D. O'Connor, W. Fu, E. E. Kenny, S. Gravel, S. McGee, R. Do, X. Liu, G. Jun, H. M. Kang, D. Jordan, S. M. Leal, S. Gabriel, M. J. Rieder, G. Abecasis, D. Altshuler, D. A. Nickerson, E. Boerwinkle, S. Sunyaev, C. D. Bustamante, M. J. Bamshad, and J. M. Akey. 2012. Evolution and functional impact of rare coding variation from deep sequencing of human exomes. Science 337:64-69.

23   Kryukov, G. V., L. A. Pennacchio, and S. R. Sunyaev. 2007. Most rare missense alleles are deleterious in humans: implications for complex disease and association studies. Am. J. Hum. Genet. 80:727-739.

24   Altshuler, D., M. J. Daly, and E. S. Lander. 2008. Genetic mapping in human disease. Science 322:881-888.

25   Raychaudhuri, S., O. Iartchouk, K. Chin, P. L. Tan, A. K. Tai, S. Ripke, S. Gowrisankar, S. Vemuri, K. Montgomery, Y. Yu, R. Reynolds, D. J. Zack, B. Campochiaro, P. Campochiaro, N. Katsanis, M. J. Daly, and J. M. Seddon. 2011. A rare penetrant mutation in CFH confers high risk of age-related macular degeneration. Nat. Genet. 43:1232-1236.

26   Seddon, J. M., Y. Yu, E. C. Miller, R. Reynolds, P. L. Tan, S. Gowrisankar, J. I. Goldstein, M. Triebwasser, H. E. Anderson, J. Zerbib, D. Kavanagh, E. Souied, N. Katsanis, M. J. Daly, J. P. Atkinson, and S. Raychaudhuri. 2013. Rare variants in CFI, C3 and C9 are associated with high risk of advanced age-related macular degeneration. Nat. Genet. 45:1366-1370.

27   Cruchaga, C., C. M. Karch, S. C. Jin, B. A. Benitez, Y. Cai, R. Guerreiro, O. Harari, J. Norton, J. Budde, S. Bertelsen, A. T. Jeng, B. Cooper, T. Skorupa, D. Carrell, D. Levitch, S. Hsu, J. Choi, M. Ryten, J. Hardy, M. Ryten, D. Trabzuni, M. E. Weale, A. Ramasamy, C. Smith, C. Sassi, J. Bras, J. R. Gibbs, D. G. Hernandez, M. K. Lupton, J. Powell, P. Forabosco, P. G. Ridge, C. D. Corcoran, J. T. Tschanz, M. C. Norton, R. G. Munger, C. Schmutz, M. Leary, F. Y. Demirci, M. N. Bamne, X. Wang, O. L. Lopez, M. Ganguli, C. Medway, J. Turton, J. Lord, A. Braae, I. Barber, K. Brown, P. Passmore, D. Craig, J. Johnston, B. McGuinness, S. Todd, R. Heun, H. Kolsch, P. G. Kehoe, N. M. Hooper, E. R. Vardy, D. M. Mann, S. Pickering-Brown, K. Brown, N. Kalsheker, J. Lowe, K. Morgan, A. David Smith, G. Wilcock, D. Warden, C. Holmes, P. Pastor, O. Lorenzo-Betancor, Z. Brkanac, E. Scott, E. Topol, K. Morgan, E. Rogaeva, A. B. Singleton, J. Hardy, M. I. Kamboh, P. St George-Hyslop, N. Cairns, J. C. Morris, J. S. Kauwe, and A. M. Goate. 2013. Rare coding variants in the phospholipase D3 gene confer risk for Alzheimer's disease. Nature.

28   Kavanagh, D., A. Richards, and J. P. Atkinson. 2008. Complement regulatory genes and hemolytic uremic syndromes. Annu. Rev. Med. 59:293-309.

29   Taylor, C. M., C. Chua, J. J. Howie, and R. Risdon. 2004. Clinico-pathological findings in diarrhoea-negative haemolytic uraemic syndrome. Pediatr. Nephrol. 19:419-425.

30   Aldick, T., M. Bielaszewska, W. Zhang, J. Brockmeyer, H. Schmidt, A. W. Friedrich, K. S. Kim, M. A. Schmidt, and H. Karch. 2007. Hemolysin from Shiga toxin-negative Escherichia coli O26 strains injures microvascular endothelium. Microbes Infect 9:282-290.

31   Morigi, M., M. Galbusera, S. Gastoldi, M. Locatelli, S. Buelli, A. Pezzotta, C. Pagani, M. Noris, M. Gobbi, M. Stravalaci, D. Rottoli, F. Tedesco, G. Remuzzi, and C. Zoja. 2011. Alternative pathway activation of complement by shiga toxin promotes exuberant c3a formation that triggers microvascular thrombosis. J. Immunol. 187:172-180.

32   Warwicker, P., T. H. J. Goodship, R. L. Donne, Y. Pirson, A. Nicholls, R. M. Ward, P. Turnpenny, and J. A. Goodship. 1998. Genetic studies into inherited and sporadic hemolytic uremic syndrome. Kidney Int. 53:836-844.

33   Richards, A., M. R. Buddles, R. L. Donne, B. S. Kaplan, E. Kirk, M. C. Venning, C. L. Tielemans, J. A. Goodship, and T. H. Goodship. 2001. Factor H mutations in hemolytic uremic syndrome cluster in exons 18-20, a domain important for host cell recognition. Am. J. Hum. Genet. 68:485-490.

34   Richards, A., E. J. Kemp, M. K. Liszewski, J. A. Goodship, A. K. Lampe, R. Decorte, M. H. Muslumanolu, S. Kavukcu, G. Filler, Y. Pirson, L. S. Wen, J. P. Atkinson, and T. H. J. Goodship. 2003. Mutations in human complement regulator, membrane cofactor protein (CD46), predispose to development of familial hemolytic uremic syndrome. Proc. Natl. Acad. Sci. USA 100:12966-12971.

35   Kavanagh, D., E. J. Kemp, E. Mayland, R. J. Winney, J. Duffield, G. Warwick, A. Richards, R. Ward, J. A. Goodship, and T. H. J. Goodship. 2005. Mutations in complement factor I predispose to the development of atypical hemolytic uremic syndrome. J. Am. Soc. Nephrol. 16:2150-2155.

36   Goicoechea de Jorge, E., C. L. Harris, J. Esparza-Gordillo, L. Carreras, E. A. Arranz, C. A. Garrido, M. Lopez-Trascasa, P. Sanchez-Corral, B. P. Morgan, and S. Rodriguez de Cordoba. 2007. Gain-of-function mutations in complement factor B are associated with atypical hemolytic uremic syndrome. Proc. Natl. Acad. Sci. USA 104:240-245.

37   Fremeaux-Bacchi, V., E. C. Miller, M. K. Liszewski, L. Strain, J. Blouin, A. L. Brown, N. Moghal, B. S. Kaplan, R. A. Weiss, K. Lhotta, G. Kapur, T. Mattoo, H. Nivet, W. Wong, S. Gie, B. Hurault de Ligny, M. Fischbach, R. Gupta, R. Hauhart, V. Meunier, C. Loirat, M. A. Dragon-Durey, W. H. Fridman, B. J. Janssen, T. H. Goodship, and J. P. Atkinson. 2008. Mutations in complement C3 predispose to development of atypical hemolytic uremic syndrome. Blood 112:4948-4952.

38   Kavanagh, D., and T. H. Goodship. 2010. Atypical hemolytic uremic syndrome. Curr. Opin. Hematol. 17:432-438.

39   Dragon-Durey, M.-A., C. Loirat, S. Cloarec, M.-A. Macher, J. Blouin, H. Nivet, L. Weiss, W. H. Fridman, and V. Fremeaux-Bacchi. 2005. Anti-factor H autoantibodies associated with atypical hemolytic uremic syndrome. J. Am. Soc. Nephrol. 16:555-563.

40   Kavanagh, D., R. Burgess, D. Spitzer, A. Richards, M. L. Diaz-Torres, J. A. Goodship, D. E. Hourcade, J. P. Atkinson, and T. H. Goodship. 2007. The decay accelerating factor mutation I197V found in hemolytic uraemic syndrome does not impair complement regulation. Mol. Immunol. 44:3162-3167.

41   Richards, A., M. K. Liszewski, D. Kavanagh, C. J. Fang, E. A. Moulton, V. Fremeaux-Bacchi, G. Remuzzi, M. Noris, T. H. J. Goodship, and J. P. Atkinson. 2007. Implications of the initial mutations in membrane cofactor protein (MCP; CD46) leading to atypical hemolytic uremic syndrome. Mol. Immunol. 44:111-122.

42   Esparza-Gordillo, J., E. G. Jorge, C. A. Garrido, L. Carreras, M. Lopez-Trascasa, P. Sanchez-Corral, and S. R. de Cordoba. 2006. Insights into hemolytic uremic syndrome: segregation of three independent predisposition factors in a large, multiple affected pedigree. Mol. Immunol. 43:1769-1775.

43   Esparza-Gordillo, J., E. Goicoechea de Jorge, A. Buil, L. C. Berges, M. Lopez-Trascasa, P. Sanchez-Corral, and S. Rodriguez de Cordoba. 2005. Predisposition to atypical hemolytic uremic syndrome involves the concurrence of different susceptibility alleles in the regulators of complement activation gene cluster in 1q32. Hum. Mol. Genet. 14:703-712.

44   Noris, M., J. Caprioli, E. Bresin, C. Mossali, G. Pianetti, S. Gamba, E. Daina, C. Fenili, F. Castelletti, A. Sorosina, R. Piras, R. Donadelli, R. Maranta, I. van der Meer, E. M. Conway, P. F. Zipfel, T. H. Goodship, and G. Remuzzi. 2010. Relative role of genetic complement abnormalities in sporadic and familial aHUS and their impact on clinical phenotype. Clinical journal of the American Society of Nephrology : CJASN 5:1844-1859.

45   Fakhouri, F., L. Roumenina, F. Provot, M. Sallee, S. Caillard, L. Couzi, M. Essig, D. Ribes, M.-A. Dragon-Durey, F. Bridoux, E. Rondeau, and V. Fremeaux-Bacchi. 2010. Pregnancy-Associated Hemolytic Uremic Syndrome Revisited in the Era of Complement Gene Mutations. J. Am. Soc. Nephrol. 21:859-867.

46   de Jorge, E. G., P. Macor, D. Paixao-Cavalcante, K. L. Rose, F. Tedesco, H. T. Cook, M. Botto, and M. C. Pickering. 2011. The development of atypical hemolytic uremic syndrome depends on complement C5. J. Am. Soc. Nephrol. 22:137-145.

47   Legendre, C. M., C. Licht, P. Muus, L. A. Greenbaum, S. Babu, C. Bedrosian, C. Bingham, D. J. Cohen, Y. Delmas, K. Douglas, F. Eitner, T. Feldkamp, D. Fouque, R. R. Furman, O. Gaber, M. Herthelius, M. Hourmant, D. Karpman, Y. Lebranchu, C. Mariat, J. Menne, B. Moulin, J. Nurnberger, M. Ogawa, G. Remuzzi, T. Richard, R. Sberro-Soussan, B. Severino, N. S. Sheerin, A. Trivelli, L. B. Zimmerhackl, T. Goodship, and C. Loirat. 2013. Terminal complement inhibitor eculizumab in atypical hemolytic-uremic syndrome. N. Engl. J. Med. 368:2169-2181.

48   Bomback, A. S., and G. B. Appel. 2012. Pathogenesis of the C3 glomerulopathies and reclassification of MPGN. Nature reviews. Nephrology 8:634-642.

49   Appel, G. B., H. T. Cook, G. Hageman, J. C. Jennette, M. Kashgarian, M. Kirschfink, J. D. Lambris, L. Lanning, H. U. Lutz, S. Meri, N. R. Rose, D. J. Salant, S. Sethi, R. J. Smith, W. Smoyer, H. F. Tully, S. P. Tully, P. Walker, M. Welsh, R. Wurzner, and P. F. Zipfel. 2005. Membranoproliferative glomerulonephritis type II (dense deposit disease): an update. J. Am. Soc. Nephrol. 16:1392-1403.

50   Cortes, C., V. P. Ferreira, and M. K. Pangburn. 2011. Native properdin binds to Chlamydia pneumoniae and promotes complement activation. Infect. Immun. 79:724-731.

51   Dragon-Durey, M.-A., V. Fremeaux-Bacchi, C. Loirat, J. Blouin, P. Niaudet, G. Deschenes, P. Copp, W. H. Fridman, and L. Weiss. 2004. Heterozygous and homozygous factor H deficiencies associated with hemolytic uremic syndrome or membranoproliferative glomerulonephritis: Report and genetic analysis of 16 cases. J. Am. Soc. Nephrol. 15:787-795.

52   Dunn, P. M. 1995. Major ethical problems confronting perinatal care around the world. Int. J.Gynaecol. Obstet. 51:205-210.

53   2002. ACOG practice bulletin. Diagnosis and management of preeclampsia and eclampsia. Number 33, January 2002. Obstet. Gynecol. 99:159-167.

54   Stillman, I. E., and S. A. Karumanchi. 2007. The glomerular injury of preeclampsia. J. Am. Soc. Nephrol. 18:2281-2284.

55   Ilekis, J. V., U. M. Reddy, and J. M. Roberts. 2007. Preeclampsia--a pressing problem: an executive summary of a National Institute of Child Health and Human Development workshop. Reprod Sci 14:508-523.

56   Hutchinson, E. S., P. Brownbill, N. W. Jones, V. M. Abrahams, P. N. Baker, C. P. Sibley, and I. P. Crocker. 2009. Utero-placental haemodynamics in the pathogenesis of pre-eclampsia. Placenta 30:634-641.

57   Maynard, S., F. H. Epstein, and S. A. Karumanchi. 2008. Preeclampsia and angiogenic imbalance. Annu. Rev. Med. 59:61-78.

58   Karumanchi, S. A., S. E. Maynard, I. E. Stillman, F. H. Epstein, and V. P. Sukhatme. 2005. Preeclampsia: a renal perspective. Kidney Int. 67:2101-2113.

59   Venkatesha, S., M. Toporsian, C. Lam, J. Hanai, T. Mammoto, Y. M. Kim, Y. Bdolah, K. H. Lim, H. T. Yuan, T. A. Libermann, I. E. Stillman, D. Roberts, P. A. D'Amore, F. H. Epstein, F. W. Sellke, R. Romero, V. P. Sukhatme, M. Letarte, and S. A. Karumanchi. 2006. Soluble endoglin contributes to the pathogenesis of preeclampsia. Nat. Med. 12:642-649.

60   Ferrara, N., K. J. Hillan, H. P. Gerber, and W. Novotny. 2004. Discovery and development of bevacizumab, an anti-VEGF antibody for treating cancer. Nat. Rev. Drug Discov. 3:391-400.

61   Lynch, A. M., R. S. Gibbs, J. R. Murphy, T. Byers, M. C. Neville, P. C. Giclas, J. E. Salmon, T. M. Van Hecke, and V. M. Holers. 2008. Complement activation fragment Bb in early pregnancy and spontaneous preterm birth. Am. J. Obstet. Gynecol. 199:354 e351-358.

62   Lynch, A. M., and J. E. Salmon. 2010. Dysregulated complement activation as a common pathway of injury in preeclampsia and other pregnancy complications. Placenta 31:561-567.

63   Fakhouri, F., L. Roumenina, F. Provot, M. Sallee, S. Caillard, L. Couzi, M. Essig, D. Ribes, M. A. Dragon-Durey, F. Bridoux, E. Rondeau, and V. Fremeaux-Bacchi. 2010. Pregnancy-associated hemolytic uremic syndrome revisited in the era of complement gene mutations. J. Am. Soc. Nephrol. 21:859-867.

64   Fakhouri, F., M. Jablonski, J. Lepercq, J. Blouin, A. Benachi, M. Hourmant, Y. Pirson, A. Durrbach, J. P. Grunfeld, B. Knebelmann, and V. Fremeaux-Bacchi. 2008. Factor H, membrane cofactor protein, and factor I mutations in patients with hemolysis, elevated liver enzymes, and low platelet count syndrome. Blood 112:4542-4545.

65   Salmon, J. E., C. Heuser, M. Triebwasser, M. K. Liszewski, D. Kavanagh, L. Roumenina, D. W. Branch, T. Goodship, V. Fremeaux-Bacchi, and J. P. Atkinson. 2011. Mutations in complement regulatory proteins predispose to preeclampsia. PLoS Medicine 8:e1001013.

66   Bienaime, F., M. A. Dragon-Durey, C. H. Regnier, S. C. Nilsson, W. H. Kwan, J. Blouin, M. Jablonski, N. Renault, M. A. Rameix-Welti, C. Loirat, C. Sautes-Fridman, B. O. Villoutreix, A. M. Blom, and V. Fremeaux-Bacchi. Mutations in components of complement influence the outcome of Factor I-associated atypical hemolytic uremic syndrome. Kidney Int. 77:339-349.

67   Hageman, G. S., D. H. Anderson, L. V. Johnson, L. S. Hancox, A. J. Taiber, L. I. Hardisty, J. L. Hageman, H. A. Stockman, J. D. Borchardt, K. M. Gehrs, R. J. Smith, G. Silvestri, S. R. Russell, C. C. Klaver, I. Barbazetto, S. Chang, L. A. Yannuzzi, G. R. Barile, J. C. Merriam, R. T. Smith, A. K. Olsh, J. Bergeron, J. Zernant, J. E. Merriam, B. Gold, M. Dean, and R. Allikmets. 2005. A common haplotype in the complement regulatory gene factor H (HF1/CFH)

predisposes individuals to age-related macular degeneration. Proc. Natl. Acad. Sci. USA 102:7227-7232.

68   Haines, J. L., M. A. Hauser, S. Schmidt, W. K. Scott, L. M. Olson, P. Gallins, K. L. Spencer, S. Y. Kwan, M. Noureddine, J. R. Gilbert, N. Schnetz-Boutaud, A. Agarwal, E. A. Postel, and M. A. Pericak-Vance. 2005. Complement factor H variant increases the risk of age-related macular degeneration. Science 308:419-421.

69   Weismann, D., K. Hartvigsen, N. Lauer, K. L. Bennett, H. P. Scholl, P. Charbel Issa, M. Cano, H. Brandstatter, S. Tsimikas, C. Skerka, G. Superti-Furga, J. T. Handa, P. F. Zipfel, J. L. Witztum, and C. J. Binder. 2011. Complement factor H binds malondialdehyde epitopes and protects from oxidative stress. Nature 478:76-81.

70   Clark, S. J., L. A. Ridge, A. P. Herbert, S. Hakobyan, B. Mulloy, R. Lennon, R. Wurzner, B. P. Morgan, D. Uhrin, P. N. Bishop, and A. J. Day. 2013. Tissue-specific host recognition by complement factor H is mediated by differential activities of its glycosaminoglycan-binding regions. J. Immunol. 190:2049-2057.

71   Nilsson, O. R., J. Lannergard, B. P. Morgan, G. Lindahl, and M. C. Gustafsson. 2013. Affinity purification of human factor h on polypeptides derived from streptococcal m protein: enrichment of the y402 variant. PloS one 8:e81303.

72   Fagerness, J. A., J. B. Maller, B. M. Neale, R. C. Reynolds, M. J. Daly, and J. M. Seddon. 2009. Variation near complement factor I is associated with risk of advanced AMD. Eur. J. Hum. Genet. 17:100-104.

73   Maller, J. B., J. A. Fagerness, R. C. Reynolds, B. M. Neale, M. J. Daly, and J. M. Seddon. 2007. Variation in complement factor 3 is associated with risk of age-related macular degeneration. Nat. Genet. 39:1200-1201.

74   Gold, B., J. E. Merriam, J. Zernant, L. S. Hancox, A. J. Taiber, K. Gehrs, K. Cramer, J. Neel, J. Bergeron, G. R. Barile, R. T. Smith, A. G. C. S. Group, G. S. Hageman, M. Dean, and R. Allikmets. 2006. Variation in factor B (BF) and complement component 2 (C2) genes is associated with age-related macular degeneration. Nat. Genet. 38:458.

75   Maller, J., S. George, S. Purcell, J. Fagerness, D. Altshuler, M. J. Daly, and J. M. Seddon. 2006. Common variation in three genes, including a noncoding variant in CFH, strongly influences risk of age-related macular degeneration. Nat. Genet. 38:1055-1059.

76   Heurich, M., R. Martinez-Barricarte, N. J. Francis, D. L. Roberts, S. Rodriguez de Cordoba, B. P. Morgan, and C. L. Harris. 2011. Common polymorphisms in C3, factor B, and factor H collaborate to determine systemic complement activity and disease risk. Proc. Natl. Acad. Sci. USA 108:8761-8766.

77   Jozsi, M., S. Heinen, A. Hartmann, C. W. Ostrowicz, S. Halbich, H. Richter, A. Kunert, C. Licht, R. E. Saunders, S. J. Perkins, P. F. Zipfel, and C. Skerka. 2006. Factor H and atypical hemolytic uremic syndrome: mutations in the C-terminus cause structural changes and defective recognition functions. J. Am. Soc. Nephrol. 17:170-177.

78   Hyvarinen, S., K. Uchida, M. Varjosalo, R. Jokela, and T. S. Jokiranta. 2013. Recognition of malondialdehyde modified proteins by C-terminus of complement factor H is mediated via the polyanion binding site and impaired by mutations found in atypical hemolytic uremic syndrome. J. Biol. Chem.

79   Pickering, M. C., H. T. Cook, J. Warren, A. E. Bygrave, J. Moss, M. J. Walport, and M. Botto. 2002. Uncontrolled C3 activation causes membranoproliferative glomerulonephritis in mice deficient in complement factor H. Nat. Genet. 31:424-428.

80   Pickering, M. C., J. Warren, K. L. Rose, F. Carlucci, Y. Wang, M. J. Walport, H. T. Cook, and M. Botto. 2006. Prevention of C5 activation ameliorates spontaneous and experimental glomerulonephritis in factor H-deficient mice. Proc. Natl. Acad. Sci. USA 103:9649-9654.

81   Pickering, M. C., E. G. de Jorge, R. Martinez-Barricarte, S. Recalde, A. Garcia-Layana, K. L. Rose, J. Moss, M. J. Walport, H. T. Cook, S. R. de Cordoba, and M. Botto. 2007. Spontaneous hemolytic uremic syndrome triggered by complement factor H lacking surface recognition domains. J. Exp. Med. 204:1249-1256.

82   Xu, C., D. Mao, V. M. Holers, B. Palanca, A. M. Cheng, and H. Molina. 2000. A critical role for the murine complement regulator Crry in fetomaternal tolerance. Science 287:498-501.

83   Mao, D., X. Wu, C. Deppong, L. D. Friend, G. Dolecki, D. M. Nelson, and H. Molina. 2003. Negligible role of antibodies and C5 in pregnancy loss associated exclusively with C3-dependent mechanisms through complement alternative pathway. Immunity 19:813-822.

84   Ruseva, M. M., T. R. Hughes, R. M. Donev, B. Sivasankar, M. C. Pickering, X. Wu, C. L. Harris, and B. P. Morgan. 2009. Crry deficiency in complement sufficient mice: C3 consumption occurs without associated renal injury. Mol. Immunol. 46:803-811.

85   Wu, X., D. Spitzer, D. Mao, S. Peng, H. Molina, and J. Atkinson. 2006. Membrane protein Crry prevents excessive turnover of the alternative complement pathway in vivo. Mol. Immunol. 44:258.

86   Wu, X., and J. Atkinson. 2006. Maternal complement insufficiency leads to the generation of the Crry single knockout mice. Mol. Immunol. 44:258.

Chapter 2

Fetal Loss in the Crry$^{-/-}$ Mouse Is Not Dependent Upon Traditional Mediators of Complement Dependent Damage

Introduction

The alternative pathway (AP) of complement activation is central to both the ability to defend against a variety of pathogens but also to a number of human disease processes. In the kidney, AP activation is pathogenic as illustrated by atypical hemolytic uremic syndrome (aHUS), membranoproliferative glomerular nephritis type II/dense deposit disease (MPGN II/DDD), and C3 glomerulopathies (*1-3*). In these diseases, overlapping rare genetic mutations have been identified that lead to excessive AP activation. Likewise, in AMD complement regulation due to genetic risk factors is associated with retinal disease (*4*).

The AP is composed of four serum proteins C3, factor B (FB) , factor D (FD) and properdin (P). C3 continuously turnovers over to C3b, which can then bind FB. Bound to C3b, FB is cleaved by factor D to generate Bb and Ba. C3bBb is an active enzyme (the AP C3 convertase) and efficiently cleaves C3 to form additional C3b. Properdin binds the C3bBb complex and stabilizes it, increasing the half-life approximately five-fold. Nascent C3b can then form additional C3 convertases, a positive feedback or amplification loop. Factor I can cleave C3b, with a required cofactor protein, to inactivate it and prevent the formation of new AP convertases.

The AP also leads to the formation of membrane attack complex (MAC). A C5 convertase is formed when an additional C3b binds to the C3 convertase. The AP C5 convertase $(C3b)_2Bb$ cleaves C5 to C5b and C5a. C5a is a potent anaphylatoxin. C5b binds C6 and C7 and this complex can bind to cell membranes. Upon binding of C8 and multiple C9 molecules, a pore is formed (C5b-C9, MAC).

Regulation at the level of C3b is required to maintain homeostasis. If regulation is lacking, the positive feedback mechanism will cause activation to outstrip regulatory activity. At

29

least in aHUS, C5 cleavage is implicated by animal models and the response of most aHUS patients to anti-C5 mAb therapy (*5, 6*). The mechanism linking excessive AP activation and disease is not as clear in AMD and other diseases of the kidney.

To study how complement activation affects disease, we studied a complement dependent model of fetal loss in which the ubiquitous regulator of complement in the mouse, Crry, is deleted. Crry is a type I transmembrane protein with cofactor activity (CA). (*7*). Without Crry, the AP activates continuously on all surfaces. *Crry*$^{-/-}$ embryos in *Crry*$^{+/-}$ x *Crry*$^{+/-}$ matings are lost *in utero* because maternal AP activates on the developing ectoplacental cone.

Work with C3 and factor B deficient mice demonstrated that embryonic loss is dependent on the AP (*8*). The same investigation reported that demise was not dependent on other mechanisms of complement activation (*C4*$^{-/-}$) or B cells (μMT). Studies by another group showed that a *C6*$^{-/-}$ background did not rescue and that the MAC was not involved in loss (*9*). Both groups found a small effect of C5 deficiency. The *C5*$^{-/-}$ background led to ~5% of offspring (3 of 62 pups) being *Crry*$^{-/-}$ versus the expected 25% of offspring (*8*). Studies with an anti-mouse C5 monoclonal antibody (BB5.1) produced a single *Crry*$^{-/-}$ mouse in approximately 20 treated liters (*9*). These results argue against a central role for C5 in loss. Because C6 deficiency did not rescue, if C5 does have a role, it is likely through the anaphylatoxin C5a.

Crry$^{-/-}$ mice have low systemic blood levels of C3 and FB (10-25%) because of heightened, continuous turnover. *Crry*$^{-/-}$ female mice mated to *Crry*$^{+/-}$ males result in a full rescue of embryonic lethality (*10*). This is because the *Crry*$^{-/-}$ mother's AP levels are so reduced that they are insufficient to cause loss. Heterozygosity for C3 had no effect and heterozygosity for FB had a partial effect on rescue. *Crry*$^{+/-}$ mice have normal C3 and factor B levels at baseline.

There is another pertinent mouse model of fetal loss that is AP dependent (*11*). In this model, human anti-phospholipid antibodies (APLA) are injected into mice and bind in the placenta to activate complement (*11*). The proposed mechanism is Ab dependent AP triggering leading to C5 activation. The liberated C5a engages the C5aR and signaling leads to upregulation of tissue factor (TF) (*12, 13*). The TF:VIIa complex then forms, which cleaves protease activated receptor 2 (PAR-2). PAR-2 in turn activates neutrophils and leads to an oxidative burst (*14*). Neutrophil depletion at the time of APLA exposure rescues fetal loss. Neutrophils have been reported around $Crry^{-/-}$ embryos prior to loss (*8*). Based on these findings, *we hypothesized that complement activation of neutrophils mediates fetal loss in the $Crry^{-/-}$ model.*

Materials and Methods

Mouse Breeding and Genotyping

All mice were bred and maintained under pathogen-free conditions at Washington University School of Medicine in St. Louis, MO in accordance with institutional animal care guidelines. The *Crry* knockout mouse was originally generated by Hector Molina and colleagues (*7*) and maintained at Washington University. The $Crry^{-/-}$ allele was genotyped by PCR as described (*7*). The *C3aR* knockout mouse was generated by and obtained from Rick Wetsel (University of Texas, Houston). It was genotyped by PCR as previously described (*15*).

Timed Matings and Harvesting Embryos

Female mice were placed in the male's cage. Each subsequent d the female was checked for a vaginal plug. The d of its observance was considered 0.5 d post conception (dpc). Mice were checked daily for seven to ten d. If a new plug appeared, it was then considered 0.5 dpc. Mice were expected to deliver at 19.5 dpc. Pregnant mice were sacrificed by CO_2 asphyxiation in

accordance with institutional guidelines. The uterus was then removed and each implantation site

was separated surgically. The muscular uterus was removed under a dissecting microscope while

the implantation site was in cold PBS. All liters sacrificed at or before 13.5 dpc were weighed.

Intact implantation sites, consisting of the embryo within the intact chorion surrounded by

decidua, were weighed to confirm the dpc. Genotyping was performed on all liters. To

accomplish this, the embryo was removed, washed 7 times in cold PBS and then digested in

proteinase K overnight at 55°C. DNA was precipitated, resuspended in 10mM Tris, 0.1mM

EDTA and analyzed by PCR.

Transcardial Perfusion

Mice were anesthetized with ketamine/xylazine and transcardial perfusion with 50 ml 20

U/ml heparin (Sigma-Aldrich) in DPBS was performed to remove serum and red blood cells

(RBCs) from the vasculature.

Frozen Section Histology

Implantation sites were harvested as above and dehydrated in 20% sucrose o/n at 4°C.

They were then flash frozen in OCT with 2-methylbutane and cooled with dry ice. Cassettes

were stored at -80° C. Frozen sections (7 μm) were made on a Leica CryoStat. For Gr-1 staining,

frozen slides were fixed in pre-chilled acetone at RT. Endogenous peroxidase was quenched with

0.3% H_2O_2 in methanol. Blocking was performed in PBS, 1% BSA, 5% mouse serum, and 5%

goat serum. RB6-8C5 (BioXCell 3.5 mg/ml) was used at 1:500 in the blocking buffer. 1A8 was

used at (1:500, BioXCel). The secondary Ab was goat anti-rat light chain HRP (Jackson

Immunoresearch). Staining was visualized with DAB (Vector, IMPACT DAB KIT).

Formalin Fixed Parafin Embedded Histology

Embryos were harvested as above, excluding transcardial perfusion. They were fixed in 10% formalin overnight then embedded. Sections were rehydrated, antigen retrieval was performed, blocked, and stained. In the case of Crry staining, antigen retrieval was in 10 mM citric acid (anhydrous), 0.05% Tween-10, pH 6.0 in a pressure cooker for 3 min. Blocking and staining was done in 1% BSA, PBS, 10% donkey serum, 5% mouse serum. Rabbit anti-Crry (1:1000) was used in blocking buffer overnight. Donkey anti-rabbit HRP (Jackson Immunoresearch) was used at 1:200. Staining was visualized with DAB (Vector, IMPACT DAB KIT).

For TROMA-I staining, 7.5 dpc embryos were collected as described, without transcardial perfusion. Antigen retrieval was done with 10 mM Tris EDTA ph 9.0 for 3 minutes in a pressure cooker. Staining was accomplished with TROMA-I(1:50 dilution hybridoma supernatant) (Developmental Studies Hybridoma Bank, Univ. Of Iowa). goat anti-rat light chain HRP (Jackson Immunoresearch). Staining was visualized with DAB (Vector, IMPACT DAB KIT). Goat anti-rat light chain HRP (Jackson Immunoresearch) was used in blocking buffer. Staining was visualized with DAB (Vector, IMPACT DAB KIT).

FACS of d7.5 Embryos/Implantation Sites

Implantation sites were harvested as described above. Sites were kept in cold PBS on ice until all sites were obtained. Each site was then cut into 12 pieces and placed in RPMI 5% fetal bovine serum (FBS). These were digested in RPMI containing 5% FBS, 300 μg/ml collagenase F (Sigma-Aldrich), 200 μg/ml collagenase L (Sigma-Aldrich), 500 μg/ml dispase (Gibco), and 2 U/ml DNAse-1 (Roche) at 37°C for 30 to 45 min with a magnetic stir bar for agitation. Cells were passed over a 70 μm strainer (BD) to create a single cell suspension. Sites were washed in DPBS, 1% FBS, 25 mM EDTA. Cells were stained for FACS with anti-CD45 (30-F11, BD),

anti-CD11b (M1/70, BD), anti-Gr-1 (RB6-8C5, BD), and rabbit anti-Crry with a donkey anti-rabbit DyLight 488 (Jackson ImmunoResearch). Blocking for FACS was carried out employing DPBS, 1% FBS, 25 mM EDTA with 5% a donkey serum and 5% mouse serum. Tissues were examined employing a FACScan (BD) retrofitted with a Cytek Upgrade.

Cobra Venom Factor (CVF) Treatment

20 µg of CVF (Quidel, A600) was administered intraperitoneally (IP) with a 31G insulin syringe (Terumo).

Neutrophil Depletion with RB6-8C5 and 1A8

Neutrophils were depleted by IP injection of RB6-8C5, a mAb against Gr1 (Ly6G/C). RB6-8C5 is a rat IgG2b Ab. A 250 µg dose of RB6 depleted neutrophils in the periphery for 5 d and a 500 µg dose depleted for 6 d. 1A8 is a rat IgG2a Ab (BioXCell). A 500 µg dose of 1A8 depleted ~50% of neutrophils when peripheral blood was assayed at 72 h.

Sources of RB6-8C5

In the initial experiments, we used RB6-8C5 that was a gift from Emil Unanue (*16*). RB6-8C5 was also produced within the laboratory by growing the hybridoma cells. The mAb was purified from supernatants on a protein G column and then dialyzed against PBS. RB6 was also purchased from BioXCell.

C5aR Antagonist Treatment

AcF-[OP(D-cyclcohexylalanine)WR was produced and purified to 99% in the acetate salt form by Biomatik. It was initially dissolved in DMSO (30 µg/µl) and then further diluted in sterile water. C5aR antagonist was given IP at 5 mg/kg. Because the half-life of this C5aR antagonist in the blood is approximately 4 h (after an initial rapid clearance in the first h) it was administered either every 12 or 24 h. Daily dosing spanned 4.5-7.5 dpc with twice a d dosing 6.5 and 7.5 dpc.

<u>Results</u>

Neutrophils Are Not Responsible for Crry$^{-/-}$ Embryo Loss

The absence of B cells, T cells, and macrophages in the decidua has been reported and argues against their involvement in fetal loss [#3117]. Genetic evidence has ruled out a role for humoral immunity, as *Crry$^{-/-}$* embryo rescue did not occur on the μMT background that lacks B cells. This was further reinforced by the lack of rescue in the *C4$^{-/-}$* background. This result rules out a role for the lectin or classical pathway of complement. T cell deficient mice did not rescue either (unpublished data). A previous report showed neutrophils were present around 7.5 dpc in *Crry$^{-/-}$* embryos (*8*).

To investigate if neutrophils are required for complement dependent fetal loss in the *Crry$^{-/-}$* model, we utilized the mAb RB6-8C5 that depletes Gr-1 (Ly-6C/G) positive cells , the majority of which are neutrophils. A 250 μg dose administered IP depletes neutrophils from the peripheral blood for 5 d. A 500 μg IP dose depletes for 6 d. Both doses are followed by a strong rebound neutrophilia (approximately doubling of pre-depletion levels), which we were unable to overcome with an additional IP dose 4 d after the first dose.

Embryo loss was reported to occur by 9.5 dpc so we depleted prior to that focusing on initiating therapy around 5.5 dpc. The embryo does not reach the uterus and implant until 4.5 dpc (*17*). 5.5 dpc ensured depletion prior to establishment of the maternal vascular connection to the embryo and up to the previously noted time of loss. Neutrophil depletion with anti-Gr1 did not rescue *Crry$^{-/-}$* embryos *in utero* when mothers were sacrificed between 10.5 and 13.5 dpc (Table 2-1). This was further born out when liters treated and allowed to deliver (Table 2-2). Multiple sources of anti-Gr1 were tried with the same effect including purified IgG from mouse ascities,

mAb produced by the hybridoma cell culture facility, and, finally, mAb purchased from

BioXCell. Depletion of neutrophils with an anti-Ly6G specific antibody (1A8 500 μg) also did

not have an effect.

Table 2-1. Neutrophil Depletion Does Not Prevent Embryonic Loss

Female x Male	Treatment / dpc	Full Size Embryos[A]	Resorbed Embryos[A]	[o]Embryos {Liters}	Rescue
$Crry^{+/-}$ x $Crry^{+/-}$	RB6 / 5.5	19 (70%, 75%)	8 (30%, 25%)	27 {3}	No
$Crry^{+/-}$ x $Crry^{+/-}$	RB6 / 6.5	13 (76%, 75%)	4 (24%, 25%)	17 {2}	No
$Crry^{+/-}$ x $Crry^{+/-}$	PBS / 6.5	17 (68%, 75%)	8 (32%, 25%)	25 {3}	No
$Crry^{+/-}$ x $Crry^{-/-}$	RB6 / 4.5	6 (43%, 50%)	8 (57%, 50%)	14 {2}	No
$Crry^{+/-}$ x $Crry^{-/-}$	RB6 / 4.5, 6.5	9 (60%, 50%)	6 (40%, 50%)	15 {2}	No
$Crry^{+/-}$ x $Crry^{-/-}$	1A8 / 5.5	13 (54%, 50%)	8 (46%, 50%)	21 {3}	No

RB6: RB6-8C5 (mAb anti-Gr1) 250 μg IP in $Crry^{+/-}$ x $Crry^{+/-}$ matings and 500 μg IP in $Crry^{+/-}$ x $Crry^{-/-}$ matings. 1A8 (mAb anti-Ly6G) 500 μg IP. Two dates listed for dpc indicates the dose was given twice. [A]First number in parenthesis is the experimentally obtained percent and the second number is the percent of embryos expected to be resorbed when all $Crry^{-/-}$ embryos are lost given the parental genotypes. [o]Bracketed number is number of liters. No treatment condition significantly differed from the expected rate of resorptions (p=0.05). Embryos were examined at 11.5 dpc.

Table 2-2. Neutrophil Depletion Does Not Lead to the Birth of $Crry^{-/-}$ Pups

Female x Male	Treatment / dpc	$Crry^{+/-}$ Pups[A]	$Crry^{-/-}$ Pups[A]	[o]Total Pups	Rescue
$Crry^{+/-}$ x $Crry^{-/-}$	RB6 / 3.5	6 (100%, 100%)	0 (0%, 0%)	6 {1}	No
$Crry^{+/-}$ x $Crry^{-/-}$	1A8 / 4.5	4 (100%, 100%)	0 (0%, 0%)	6 {1}	No
$Crry^{+/-}$ x $Crry^{-/-}$	1A8 / 5.5	4 (100%, 100%)	0 (0%, 0%)	6 {1}	No

RB6: RB6-8C5 (mAb anti-Gr1) 500 μg IP. 1A8 (mAb anti-Ly6G) 500 μg IP. [A]First number in parenthesis is the experimentally obtained proportion and the second number is the proportion of the liter with that genotype in seven untreated $Crry^{+/-}$ x $Crry^{-/-}$ matings. [o]Bracketed number is number of liters. No treatment condition significantly differed from the expected rate of resorptions (p=0.05)

Neutrophils are Present Around All d7.5 Embryos

Using both FACS and immunohistochemistry, we detected neutrophils around $Crry^{-/-}$ and $Crry^{+/-}$ 7.5 dpc embryos ($Crry^{+/-}$ x $Crry^{-/-}$ female x male mating) (Figure 2-1A, C). There was no correlation between genotype and the number of neutrophils present (Figure 2-1D). Additionally, there was no correlation between CD45^{+} cells (a marker of hematopoietic derived cells) (or CD11b+ Gr1-) with $Crry^{-/-}$ or $Crry^{+/-}$ genotypes at d7.5 (Figure 2-1B, F). We confirmed that the anti-Gr1 Ab was able to deplete neutrophils at the site of embryo implantation at 7.5 dpc, three d after administration at 4.5 dpc (Figure 2-2).
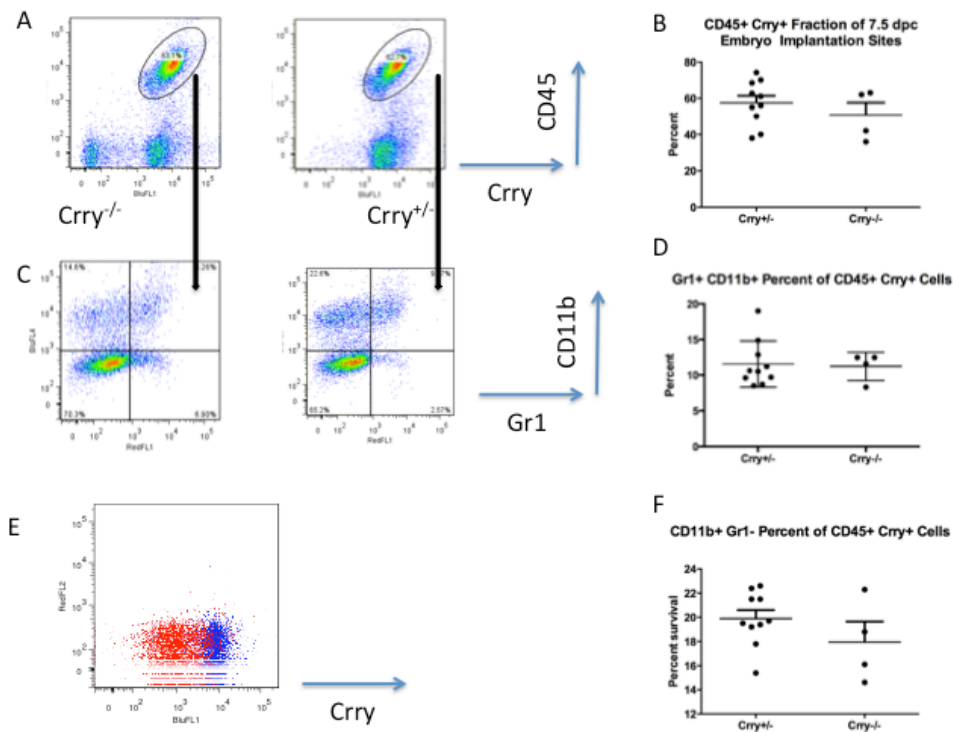
Figure 2-1. Neutrophils are Present Around All Embryos

A. *Crry*$^{-/-}$ embryos (left) have three cell populations when stained with anti-Crry and anti-CD45. Crry$^+$CD45$^+$ population (maternal hematopoietic derived cells), Crry+CD45- population (maternal decidua derived cells), and Crry- CD45- population (embryo and trophoblast derived cells). *Crry*$^{+/-}$ embryos (right) lack the Crry$^-$ CD45$^-$ population; the embryo and trophoblast derived cells cluster with maternal decidua derived cells. B. There is no trend towards different proportions of hematopoeitic derived (CD45$^+$) cells around *Crry*$^{+/-}$ and *Crry*$^{-/-}$ embryos. C. The Crry$^+$ CD45$^+$ population contains Gr1$^+$ CD11b$^+$ positive cells, neutrophils. D. There is no difference in the proportion of neutrophils dependent upon genotype. E. Staining control for anti-Crry. *Crry*$^{-/-}$ splenocytes in red and *Crry*$^{+/-}$ in blue. F. There is no difference in the proportion of CD11b$^+$ Gr1$^-$ cells dependent upon genotype (subset of Crry$^+$ CD45$^+$).
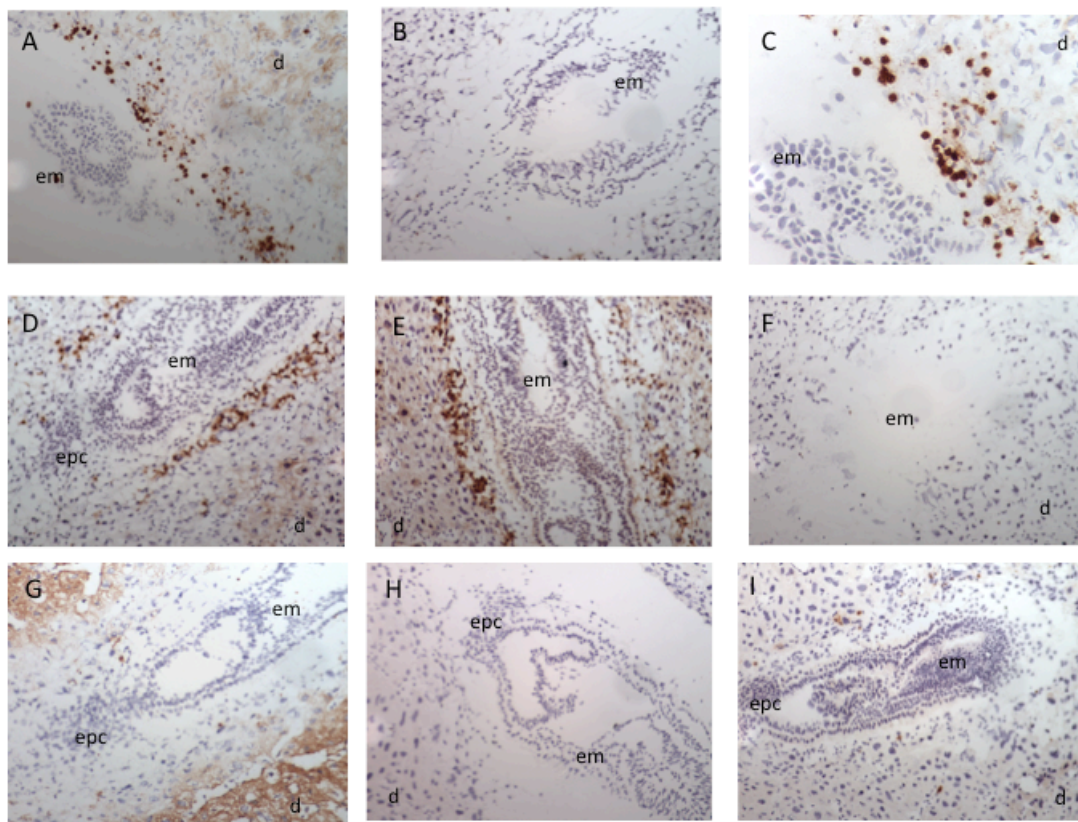
Figure 2-2. Neutrophils Are Present Around 7.5 dpc Embryos and RB6-8C5 Depletes Them From Around Embryos

A. Neutrophils are present around the embryo at 7.5 dpc (anti-Gr1 staining) (200x). B. This staining is specific for Gr-1 and is not present in the isotype control. C. At higher magnification (400x) the banded nuclear pattern characteristic of neutrophils can be seen inside of Gr-1+ cells. D,E. Identical staining patterns are observed in staining for Gr1 (RB6-8C5, D) and Ly6G (1A8, E). F. Ly6G staining is specific to 1A8 and not seen in the isotype control. G-I. Neutrophil depletion with 500 µg RB6-8C5 at 4.5 dpc leads to complete depletion in the tissue at 7.5 dpc of Gr1+ (G) and Ly6G+ cells (H and I). em embryo. epc ectoplacental cone; d, decidua.

Transient Depletion of the AP Rescues Crry$^{-/-}$ *Embryos*

A fully functional AP is required for embryo loss, so we transiently depleted C3 with cobra venom factor (CVF) (*18, 19*). This component of snake venom is a C3b analog that forms an AP convertase with Bb (CVFBb), but unlike the host's C3 convertase (C3bBb) it is <u>not</u> susceptible to decay acceleration activity by regulators such as factor H. CVF is however highly antigenic and a second dose, after 5 or 6 d, does not deplete because of the development of neutralizing Abs.

Treatment with CVF is sufficient to rescue *Crry$^{-/-}$* embryos from attack by the maternal AP. Western blot analysis of serum showed no detectable C3 at four d post CVF treatment with levels returning to 50% of WT at d 11 (7 d post 20 μg CVF IP). CVF treatment between 3.5 and 5.5 dpc fully rescued *Crry$^{-/-}$* embryos at 11.5 dpc. Treatment at 6.5 or 7.5 dpc gave a ~50% rescue (Table 2-3). However, CVF at 8.5 dpc had no effect, indicating that the events that result in loss are irreversible by this time point.

Early doses of CVF were also able to rescue the *Crry$^{-/-}$* pups at birth (Table 2-4). Pups born to CVF treated mothers survived to adulthood and *Crry$^{-/-}$* females were fertile.

Table 2-3. Transient Depletion of the AP with CVF Prior to 8.5 dpc Prevents Embryonic Loss

Female x Male	dpc of treatment	Full Size Embryos^A	Resorbed Embryos^A	^{o}Embryos {Liters}	Rescue
Crry$^{+/-}$ x *Crry*$^{+/-}$	3.5	26 (96%, 75%)	1 (**4%, 25%**)	27 {3}	YES**
Crry$^{+/-}$ x *Crry*$^{+/-}$	4.5	19 (95%, 75%)	1 (**5%, 25%**)	20 {2}	YES**
Crry$^{+/-}$ x *Crry*$^{+/-}$	5.5	10 (100%, 75%)	0 (**0%, 25%**)	10 {1}	YES
Crry$^{+/-}$ x *Crry*$^{+/-}$	6.5	13 (87%, 75%)	2 (**13%, 25%**)	15 {2}	PARTIAL*
Crry$^{+/-}$ x *Crry*$^{+/-}$	7.5	24 (89%, 75%)	3 (**11%, 25%**)	27 {3}	PARTIAL**
Crry$^{+/-}$ x *Crry*$^{+/-}$	8.5	17 (68%, 75%)	8 (32%, 25%)	25 {3}	NO

20 µg CVF. ^{A}First number in parenthesis is the experimentally obtained proportion and the second number is the proportion of resorbed embryos expected when all *Crry*$^{-/-}$ embryos are lost. ^{o}Bracketed number is number of liters. Partial rescue indicates the number of resorbed embryos was between a full rescue and no rescue. **p<0.01 and *p<0.05 compared with the proportion of full size embryos when not treated. 5.5 dpc had only one treated liter; a statistical test was not performed.

Table 2-4. Transient Depletion of the AP with CVF Leads to the Birth of *Crry*$^{-/-}$ Pups

Female x Male	dpc of treatment	*Crry*$^{+/-}$ *Pups*^{A}	*Crry*$^{-/-}$ *Pups*^{A}	^{o}Total Pups	Rescue
Crry$^{+/-}$ x *Crry*$^{-/-}$	3.5	10 (59%, 100%)	7 (**41%, 0%**)	17 {3}	YES***
Crry$^{+/-}$ x *Crry*$^{-/-}$	4.5	6 (37%, 100%)	10 (**63%, 0%**)	16 {2}	YES***
Crry$^{+/-}$ x *Crry*$^{-/-}$	5.5	12 (46%, 100%)	14 (**54%, 0%**)	26 {5}	YES***

20 µg CVF. ^{A}First number in parenthesis is the experimentally obtained proportion and the second number is the proportion of the liter with that genotype in seven untreated *Crry*$^{+/-}$ x *Crry*$^{-/-}$ matings. ^{o}Bracketed number is number of liters. ***p≤0.005 compared to the proportion of *Crry*$^{+/-}$ pups when not treated.

Properdin (P) Neutralizing Antibody Rescues

The function of P is critical for efficient AP activity as it stabilizes the C3 convertase by a factor of five. The *P*$^{-/-}$ knockout mouse rescues *Crry*$^{-/-}$ lethality, similar to *FB*$^{-/-}$ and *C3*$^{-/-}$ (*20*). We utilized a rabbit polyclonal Ab to mouse P (Dennis Hourcade) and it rescued *Crry*$^{-/-}$ viability in utero and at birth. The dose required for rescue was 1 mg at 6.5 dpc and 7.5 dpc (Table 2-5). A single dose at 7.5 dpc did effect the size of *Crry*$^{-/-}$ embryos at 11.5 dpc but did not rescue them, some were ~2/3 the size of *Crry*$^{+/-}$ litermates (Supplemental Figure 1). Western blots confirmed that this Ab is not depleting. A mAb to properdin (1 mg at 6.5 and 7.5 dpc) was also able to

rescue *Crry*[-/-] lethality with 32% of offspring being *Crry*[-/-] versus the expected 0% without

treatment (p=0.0005, 25 pups, 4 liters). Though there was an effect with the mAb to P at 1 mg it

appeared to rescue only 75% of the expected *Crry*[-/-] offspring, a higher dose was not explored.

Table 2-5. Blocking Properdin Prevents Embryonic Loss					
Female x Male	dpc(s) of treatment	Full Size Embryos[ᴬ]	Resorbed Embryos[ᴬ]	°Embryos {Liters}	Rescue
Crry[+/-] x *Crr*[-/-]	3.5 to 7.5	15 (88%, 50%)	2 (**12%, 50%**)	17 {2}	YES***
Crry[+/-] x *Crry*[-/-]	6.5 to 7.5	20 (83%, 50%)	4 (**17%, 50%**)	24 {3}	YES***
Crry[+/-] x *Crry*[-/-]	6.5 to 7.5[†]	12 (71%, 50%)	5 (29%, 50%)	17 {2}	NO[‡]
Crry[+/-] x *Crry*[-/-]	7.5	14 (48%, 50%)	15 (52%, 50%)	29 {3}	NO

1 mg rabbit anti-properdin. [†] 500 µg rabbit anti-properdin. [ᴬ]First number in parenthesis is the experimentally obtained proportion and the second number is the proportion of resorbed embryos expected when all *Crry*[-/-] embryos are lost. Brackets is number of liters. ***p<0.005 difference between the expected number of resorptions in an untreated liter. [‡]p=0.05 difference between resorptions observed and those expected.

C3aR[-/-] *Does Not Rescue Crry*[-/-] *Lethality*

      C3aR[-/-] mice were mated with *Crry*[--/-] mice to generate *Crry*[+/-]*C3aR*[-/-] mice. Absence of

C3aR did not rescue *Crry*[-/-] embryos (Table 2-6). This confirms our finding that the small

molecule C3aR antagonist SB 290157 was also unable to rescue *Crry*[-/-] embryos (data not

shown). CVF was able to rescue *Crry*[-/-]*C3aR*[-/-] embryos, similar to its effect in the *C3aR*[+/+]

background. In the *Crry*[-/-]*C3aR*[-/-] mother (a product of CVF treatment), the low levels due to lack

of Crry allowed *Crry*[-/-]*C3aR*[-/-] embryos to survive similar to *Crry*[+/-]*C3aR*[-/-] embryos. The

reciprocal mating did not rescue because the *Crry*[+/-]*C3aR*[-/-] mother had normal AP component

levels. C3aR transduces the signal of C3a, which is only generated as a result of complement

activation and is relatively unstable. C5a is implicated because the *C6*[-/-] background does not

rescue, ruling out the role of the MAC complex. The *C5*[-/-] background does lead to a small

change in phenotype, with ~5% of pups being *Crry*[-/-] versus the expected 25% (*8*). This small

effect may be mediated by the C5a fragment. Our hypothesis had been that C5a:C5aR or

C3a:C3aR signaling, or a mix of the two, could drive an inflammatory response around the

embryo (*21, 22*).

Table 2-6. The *C3aR*$^{-/-}$ Background Does Not Rescue *Crry*$^{-/-}$ Pups					
Female x Male	*Crry*$^{+/+}$ *C3aR*$^{-/-}$ Pups[A]	*Crry*$^{+/-}$ *C3aR*$^{-/-}$ Pups[A]	*Crry*$^{-/-}$ *C3aR*$^{-/-}$ Pups[A]	[o] Total Pups	Rescue
Crry$^{+/-}$ *C3aR*$^{-/-}$ x *Crry*$^{+/-}$ *C3aR*$^{-/-}$	16 (27%, 33%)	43 (73%, 67%)	0 (0%, 0%)	59 {11}	NO

[A]First number in parenthesis is the experimentally obtained proportion and the second number is the proportion of resorbed embryos expected when all *Crry*$^{-/-}$ embryos are lost. [o] Brackets is number of liters. There is no difference between what was observed with *Crry*$^{+/-}$*C3aR*$^{-/-}$ x *Crry*$^{+/-}$ *C3aR*$^{-/-}$ and eleven *Crry*$^{+/-}$ x *Crry*$^{+/-}$ liters at birth (p-value>0.05)

The C5aR Does Not Compensate in C3aR Deficient Animals

To examine the role of anaphlyatoxin signaling in mediating $Crry^{-/-}$ loss, we administered a small peptide C5aR antagonist to $Crry^{+/-}C3aR^{-/-}$ matings (*23*). Knowing that complement activation during this 6.5 to 7.5 dpc window causes the loss of $Crry^{-/-}$ embryos, a dose of 5 mg/kg was used either every 12 h on 6.5 dpc and d7.5 dpc or every 24 h from 4.5 to 7.5 dpc. No treatment regiment was able to rescue $Crry^{-/-}$ embryos or pups (Table 2-7). DMSO vehicle treatments (2%) did not have an adverse effect on viability of the $Crry^{+/+}C3aR^{-/-}$ or $Crry^{+/-}C3aR^{-/-}$ pups.

Table 2-7. C5aR Antagonist in the C3aR Background Does Not Rescue $Crry^{-/-}$ Lethality					
A. Effect on Delivered Liters					
Female x Male	dpc(s) of Treatment	$Crry^{+/+}$ $C3aR^{-/-}$ Pups[A]	$Crry^{+/-}$ $C3aR^{-/-}$ Pups[A]	$Crry^{-/-}$ $C3aR^{-/-}$ Pups[A]	[o] Total Pups
$Crry^{+/-}$ $C3aR^{-/-}$ x $Crry^{+/-}$ $C3aR^{-/-}$	6.5 to 7.5 bid	4 (33%, 25%)	8 (67%, 75%)	0 (0%, 0%)	12 {2}
$Crry^{+/-}$ $C3aR^{-/-}$ x $Crry^{+/-}$ $C3aR^{-/-}$	5.5 to 7.5 qd	2 (22%, 25%)	7 (78%, 75%)	0 (0%, 0%)	9 {2}
$Crry^{+/-}$ $C3aR^{-/-}$ x $Crry^{-/-}$ $C3aR^{-/-}$	3.5 to 7.5 qd	0 (0%, 0%)	5 (100%, 100%)	0 (0%, 0%)	5 {1}
B. Effect on Embryos					
Female x Male	dpc(s) of Treatment	Full Size Embryos[A]	Resorbed Embryos[A]		
$Crry^{+/-}$ $C3aR^{-/-}$ x $Crry^{-/-}$ $C3aR^{-/-}$	4.5 to 7.5 qd	1 (25%, 50%)	3 (75%, 50%)		4 {1}
$Crry^{+/-}$ $C3aR^{-/-}$ x $Crry^{-/-}$ $C3aR^{-/-}$	5.5 to 7.5 qd	5 (63%, 50%)	3 (37%, 50%)		8 {1}

5mg/kg C5aR antagonist IP. bid: every twelve hours. qd: every 24 h. No $Crry^{-/-}C3aR^{-/-}$ full size embryos or pups were recovered. No condition differed from the expectations of no rescue (p=0.05). [A]First number in parenthesis is the experimentally obtained proportion and the second number is the proportion of resorbed embryos expected when all $Crry^{-/-}$ embryos are lost. [o] Brackets is number of liters.

Crry$^{-/-}$ Embryos are Lost Starting at 8.5 dpc

We observed that loss actually begins at 8.5 dpc where there is a difference in the size of the allantoic vessels, a lack of proliferation of the labyrinthine trophoblasts, and a smaller embryo in *Crry$^{-/-}$* embryos (Figure 2-3). This had previously been noted at 9.5 dpc (Figure 2-4), but this process actually begins with fusion of the allantois to the ectoplacental cone at 8.5 dpc. The embryo enters the uterus at 4.5 dpc, but it is not until 6.5 dpc that the embryo comes into direct contact with maternal blood (*17*). It is clear that by 7.5 dpc the chorion is bathed in maternal blood, especially the ectoplacental cone where the placenta would normally form (Figure 2-5). From 6.5 dpc onwards, maternal complement will be in constant contact with the developing trophoblast.
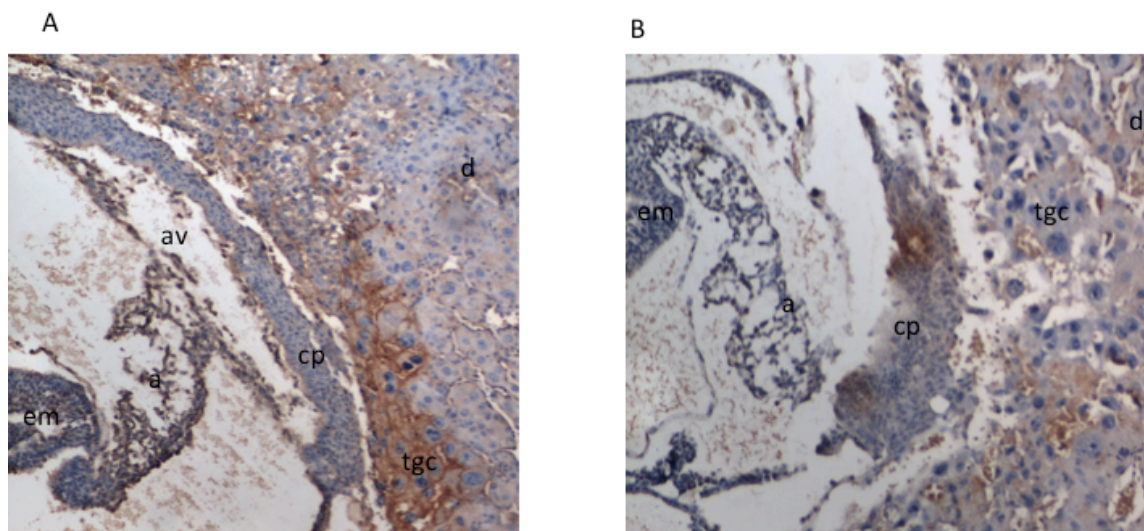
Figure 2-3. *Crry$^{-/-}$* Embryos Die at 8.5 dpc due to a Failure of the Allantoic Vessels to Attach to the Chorionic Plate
A. The allantois of *Crry$^{+/-}$* embryos attaches to the ectoplacental cone at 8.5 dpc and expands the chorionic plate. *Crry$^{+/-}$* trophoblast giant cells stain positive (brown) for Crry. B. The allantois of *Crry$^{-/-}$* embryos does not attach properly to the ectoplacental cone at 8.5dpc and the allantoic vessels and trophoblast fail to develop normally. *Crry$^{-/-}$* trophoblast giant cells do not stain positive (brown) for Crry. Areas of Crry positive cells in the decidua that are not of fetal origin are maternal blood cells within maternal vessels. FFPE sections: em, embryo; a, allantois; av, allantoic vessels; cp, chorionic plate; tgc, trophoblast giant cell; d, decidua.
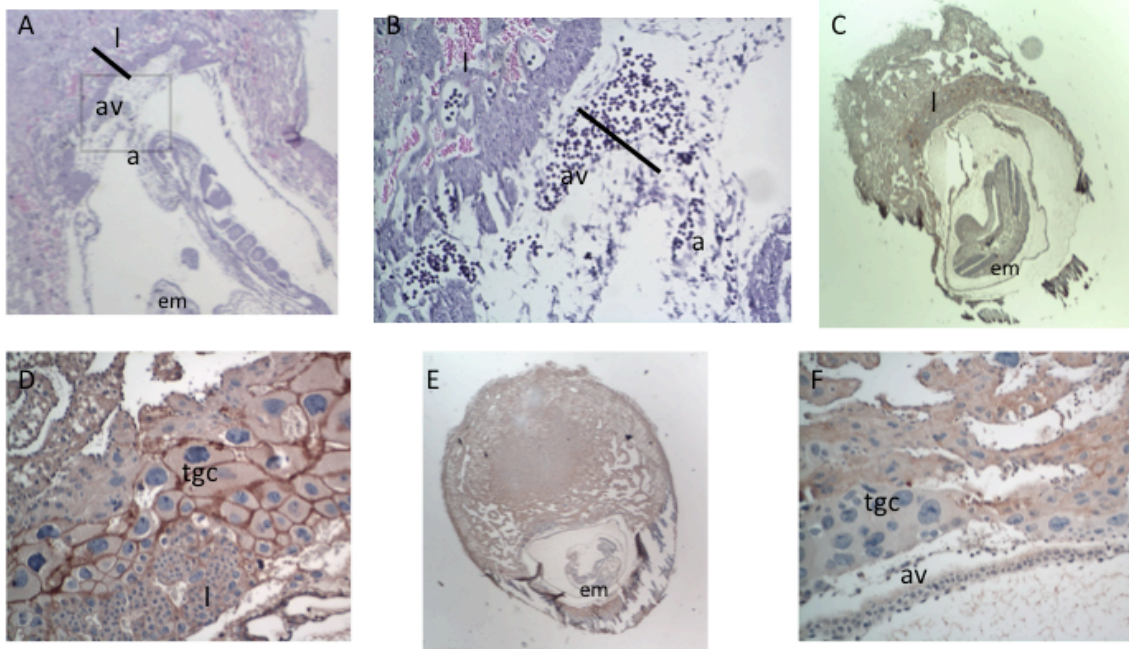
Figure 2-4. 9.5 dpc embryos fail to develop allantoic vessels and labyrinth

A. *Crry*$^{+/+}$ embryo is fully developed at 9.5 dpc. Its allantois has fused with the mesoderm overlaying the ectoplacental cone (HE, 20x) Black bar delineates the labyrinth. B. The allantoic vessels (black bar) have expanded and nucleated fetal red blood cells (RBCs) are visible passing into the labyrinth where they come into close proximity with maternal RBCs (enucleated) (200X of box in A). C. Crry positive embryo at 9.5 dpc (20x). D. Trophoblast giant cells strongly express Crry (brown, rabbit anti-Crry) and separate the maternal decidua from the labyrinth. E. *Crry*$^{-/-}$ embryo at 9.5 dpc has failed to develop. F. The *Crry*$^{-/-}$ trophoblast giant cells do not stain for Crry, the labyrinth has not developed and allantoic vessels did not expand at 9.5 dpc. av allantoic vessels. l labyrinth. em embryo. tgc trophoblast giant cell.
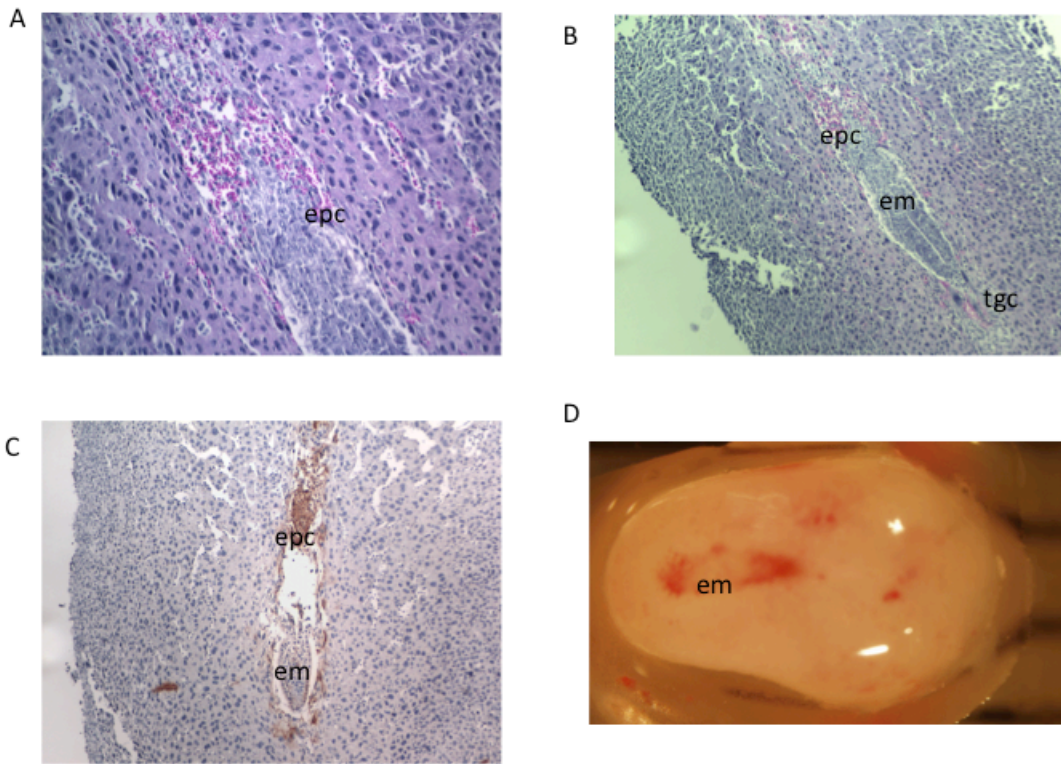
Figure 2-5. Maternal blood contacts the trophoblast at 7.5 dpc
A. Maternal RBCs are in contact with the ectoplacenta cone (200x). B.RBCs are visible at both ends of the amnionic sac. C. The cells in the epc are trophoblast cells and stain positive (brown) for cytokeratin 8 (TROMA-I). D. Gross dissection of 7.5 dpc implantation site shows blood around both ends of the implantation site. Space between black bars (right side) is 1 mm. All panels were from a 7.5 dpc liter, sacrificed without perfusion to retain the maternal blood. Histology is from FFPE. A. and B. are HE stains. em embryo. epc ectoplacental cone. tgc. trophoblast giant cell.

Discussion

The AP is a constant sentinel activating continuously on surfaces and in the fluid phase. Cells carry regulators on their surface and in the plasma to limit this activation. Crry is expressed on the surface of mouse cells and carries cofactor activity (*24*); in plasma, factor H (FH) performs this role. Crry is able to permanently stop AP activation on cells and protect them from complement activation (*25*). In this mouse model of excessive AP activation, this protection is absolutely necessary as *Crry*$^{-/-}$ embryos are lost at 8.5 dpc of development. *Crry*$^{-/-}$ mice can be rescued by modulating the levels of the AP in the mother (because AP attack on the placenta cannot occur). *Crry*$^{-/-}$ mice have low AP levels in the plasma; 10-25% the amount of *Crry*$^{+/+}$ and *Crry*$^{+/-}$ mice due to unrestricted complement activation on cellular surfaces throughout the animal (*10*). Similarly, *FH*$^{-/-}$ animals have even lower levels of AP components in the plasma due to unrestricted fluid phase turnover (*26*). The *FH*$^{-/-}$ mouse subsequently develops renal failure similar to patients with MPGN II due to a C3 nephritic factor antibody, which stabilizes the AP convertase and turns over the AP (*27*).

C3 deposition occurs on the placenta at 7.5 dpc because there is <u>no</u> regulator to prevent activation (*8*). The GPI-anchored regulator decay-accelerating factor is not expressed on the placenta until after ~10.5 dpc (*8*). Other regulators of complement such as CD59 and FH may be present but cannot compensate for Crry. CD59 is an inhibitor of the terminal pathway of complement and its expression is very high on human trophoblast cells and begins as early as six weeks (*28*). It is not known at which d CD59 is expressed on the mouse placenta, but it is likely present from an early time point. Maternal derived factor H is insufficient to limit the diffuse AP

activation seen in $Crry^{-/-}$ mice, indicating that its ability to localize to the sites of AP activation is limited in this model.

$Crry^{-/-}$ trophoblast cells appear remarkably healthy despite the AP activation that occurs on their surface (figure 2-3 and 2-4). Embryos are lost not due to excessive activation and lysis upon contact with maternal serum, but because fusion of the allantois to the chorion, a developmental process, fails to occur. This is demonstrated by the 8.5 and 9.5 dpc chorionic plate failing to develop in $Crry^{-/-}$ placentas. At 9.5 dpc it is clear that the labyrinthine component of the placenta has failed to mature because fetal vessels originating in the allantois do not infiltrate into the placenta (figure 2-4). The embryo needs to make the connection to the maternal vasculature at 8.5 dpc (figure 2-3). In contrast, the trophoblast giant cells that surround the implantation site are still present from 7.5-to 9.5 dpc. Intact embyronic and trophoblast tissue from $Crry^{-/-}$ implantation sites can be recovered at 9.5 to 12.5 dpc reliably and genotyped, even though the embryo is lost at 8.5 dpc.

Only transient depletion of the AP, either with CVF or an anti-properdin antibody, is required to rescue. There is a critical window at 6.5 to 7.5 dpc where AP activation leads to embryo loss. In the case of CVF treatment, treatment during this window produces approximately 50% rescue, whereas treatment prior to 6.5 dpc yields 100% rescue. Interestingly, if CVF is given at 3.5 dpc, prior to implantation of the fertilized embryo in the uterus, maternal C3 levels will have recovered to ~50% by 7.5 dpc (4 d after CVF treatment) and $Crry^{-/-}$ embryos survive. This replicates the finding in $Crry^{-/-}$ mothers that full blockade of the AP is not required as haploinsufficiency of AP components will rescue. The polyclonal anti-Properdin Ab that was able to fully rescue if given at 6.5 and 7.5 dpc produced a somewhat intermediate phenotype if only given at 7.5 dpc. Some of the $Crry^{-/-}$ embryos were developmentally delayed and smaller in

comparison to their litermates at 11.5 dpc, while other $Crry^{-/-}$ embryos in the same liter were fully resorbed similar to a rabbit IgG control treated liter.

Though immune cells (CD45$^+$) are present around 7.5 dpc embryos, we saw no trend for an increase of CD45$^+$ cells around $Crry^{-/-}$ embryos and specifically, there was no increase in neutrophils around $Crry^{-/-}$ embryos. Additionally, there was no correlation between $Crry$ genotype and the number of CD11b positive cells at 7.5 dpc. Correspondingly, depletion of neutrophils failed to rescue loss. Neutrophils have been found to mediate the link between complement activation and embryonic loss in a model of antiphospholipid syndrome (APLS) mediated fetal loss. In this model of APLS, the classical pathway is activated by antibodies that bind to the placenta, and AP activation leads to C5 cleavage, neutrophil recruitment and reactive oxygen species (*11, 12*).

Based on past experiments we know that the terminal pathway does not play a role in $Crry^{-/-}$ loss, and that if C5a is involved, it is not the dominant mediator of AP activation (*8, 9*). We ruled out the role of C5a and C3a by blocking both receptors (C3aR knockout and C5aR antagonist treatment). We were concerned that, because the $C5^{-/-}$ mouse had a limited effect on $Crry^{-/-}$ lethality, there might be some overlap in their ability to drive inflammation. C5a is critical in a number of mouse models of disease, including colitis, (*29*), Alzheimer's disease (*30*), athlerosclerosis (*31*), and lupus nephritis (*22*).

We did not directly rule out the role of C3b's interactions with CR3 (CD11b) and CR4 (CD11c) on immune cells; thus, it is possible that immune cells recruited to $Crry^{-/-}$ embryos could be responsible for loss and mediate the developmental defect observed. No uterine decidua cells or embryo derived cells express CD11b. Increasing evidence implicates C3b in developmental processes. Neuronal cell pruning in development and in models of Alzheimer's

disease implicate C3b deposition on unused synapses and subsequent engagement by microglial cells carrying CD11b (*32*). Previous efforts to identify macrophages (F480+ cells) around embryos at 7.5 dpc have failed (*8*). Efforts to examine CD11c/CR4 around the 7.5 dpc embryo have been limited, but the formation of its ligand, iC3b, would require maternal FH mediated cleavage of C3b.

It is also possible that C3b deposition itself could affect the placental cells directly. C3b deposition on erythrocyte membranes has been shown to change the membrane properties by clustering a variety of cytoskeletal proteins together on the surface (*33*). The current evidence for embryo loss implicates an unknown mechanism that operates through C3b activation and leads to a failure of development of the embryonic vessels in the trophoblast from the very beginning, (8.5 dpc). This AP dependent mechanism is not mediated by any of the classical mediators of complement, including the MAC, anaphylatoxins or neutrophils. This model may be related to human diseases that do not involve these effects, but instead feature C3b deposition, specifically C3 glomerulonephritis, DDD and AMD, potentially Alzheimer's disease, and the physiologic process of synapse remodeling.

References

1. D. Kavanagh, A. Richards, J. P. Atkinson, Complement regulatory genes and hemolytic uremic syndromes. *Annu. Rev. Med.* **59**, 293 (2008).
2. A. Hawfield, S. S. Iskandar, R. J. Smith, Alternative Pathway Dysfunction in Kidney Disease: A Case Report and Review of Dense Deposit Disease and C3 Glomerulopathy. *Am. J. Kidney Dis.*, (Feb 4, 2013).
3. R. J. Smith, C. L. Harris, M. C. Pickering, Dense deposit disease. *Mol. Immunol.* **48**, 1604 (Aug, 2011).
4. J. Ambati, J. P. Atkinson, B. D. Gelfand, Immunology of age-related macular degeneration. *Nat. Rev. Immunol.* **13**, 438 (Jun, 2013).

5.	C. M. Legendre *et al.*, Terminal complement inhibitor eculizumab in atypical hemolytic-uremic syndrome. *N. Engl. J. Med.* **368**, 2169 (Jun 6, 2013).

6.	E. G. de Jorge *et al.*, The development of atypical hemolytic uremic syndrome depends on complement C5. *J. Am. Soc. Nephrol.* **22**, 137 (Jan, 2011).

7.	C. Xu *et al.*, A critical role for the murine complement regulator Crry in fetomaternal tolerance. *Science* **287**, 498 (2000, 2000).

8.	D. Mao *et al.*, Negligible role of antibodies and C5 in pregnancy loss associated exclusively with C3-dependent mechanisms through complement alternative pathway. *Immunity* **19**, 813 (2003).

9.	M. M. Ruseva *et al.*, Crry deficiency in complement sufficient mice: C3 consumption occurs without associated renal injury. *Mol. Immunol.* **46**, 803 (Feb, 2009).

10.	X. Wu *et al.*, Membrane protein Crry maintains homeostasis of the complement system. *J. Immunol.* **181**, 2732 (Aug 15, 2008).

11.	V. M. Holers *et al.*, Complement C3 activation is required for antiphospholipid antibody-induced fetal loss. *J. Exp. Med.* **195**, 211 (2002).

12.	G. Girardi *et al.*, Complement C5a receptors and neutrophils mediate fetal injury in the antiphospholipid syndrome. *J. Clin. Invest.* **112**, 1644 (Dec, 2003).

13.	P. Redecha *et al.*, Tissue factor: a link between C5a and neutrophil activation in antiphospholipid antibody induced fetal injury. *Blood* **110**, 2423 (Oct 1, 2007).

14.	P. Redecha, C. W. Franzke, W. Ruf, N. Mackman, G. Girardi, Neutrophil activation by the tissue factor/Factor VIIa/PAR2 axis mediates fetal death in a mouse model of antiphospholipid syndrome. *J. Clin. Invest.* **118**, 3453 (Oct, 2008).

15.	J. Kildsgaard *et al.*, Cutting edge: targeted disruption of the C3a receptor gene demonstrates a novel protective anti-inflammatory role for C3a in endotoxin-shock. *J. Immunol.* **165**, 5406 (Nov 15, 2000).

16.	H. W. Rogers, E. R. Unanue, Neutrophils are involved in acute, nonspecific resistance to Listeria monocytogenes in mice. *Infect. Immun.* **61**, 5090 (Dec, 1993).

17.	K. Theiler, *The House Mouse.* (Springer Verlag, 1989).

18.	C. G. Cochrane, H. J. Muller-Eberhard, B. S. Aikin, Depletion of plasma complement in vivo by a protein of cobra venom: its effect on various immunologic reactions. *J. Immunol.* **105**, 55 (Jul, 1970).

19.	U. K. Laemmli, Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680 (1970, 1970).

20.	Y. Kimura, L. Zhou, T. Miwa, W. C. Song, Genetic and therapeutic targeting of properdin in mice prevents complement-mediated tissue injury. *J. Clin. Invest.* **120**, 3545 (Oct 1, 2010).

21.	L. Bao, I. Osawe, M. Haas, R. J. Quigg, Signaling through up-regulated C3a receptor is key to the development of experimental lupus nephritis. *J. Immunol.* **175**, 1947 (Aug 1, 2005).

22.	L. Bao *et al.*, C5a promotes development of experimental lupus nephritis which can be blocked with a specific receptor antagonist. *Eur. J. Immunol.* **35**, 2496 (Aug, 2005).

23.	D. R. March *et al.*, Potent cyclic antagonists of the complement C5a receptor on human polymorphonuclear leukocytes. Relationships between structures and activity. *Mol. Pharmacol.* **65**, 868 (Apr, 2004).

24.   Y. U. Kim *et al.*, Mouse complement regulatory protein Crry/p65 uses the specific mechanisms of both human decay accelerating factor and membrane cofactor protein. *J. Exp. Med.* **181**, 151 (1995, 1995).
25.   T. Miwa, L. Zhou, B. Hilliard, H. Molina, W. C. Song, Crry, but not CD59 and DAF, is indispensable for murine erythrocyte protection in vivo from spontaneous complement attack. *Blood* **99**, 3707 (2002).
26.   M. C. Pickering *et al.*, Uncontrolled C3 activation causes membranoproliferative glomerulonephritis in mice deficient in complement factor H. *Nat. Genet.* **31**, 424 (2002).
27.   H. Ohi, T. Yasugi, Occurrence of C3 nephritic factor and C4 nephritic factor in membranoproliferative glomerulonephritis (MPGN). *Clin. Exp. Immunol.* **95**, 316 (Feb, 1994).
28.   C. H. Holmes *et al.*, Preferential expression of the complement regulatory protein decay accelerating factor at the fetomaternal interface during human pregnancy. *J.Immunol.* **144**, 3099 (1990, 1990).
29.   U. Jain, T. M. Woodruff, A. W. Stadnyk, The C5a receptor antagonist PMX205 ameliorates experimentally induced colitis associated with increased IL-4 and IL-10. *Br. J. Pharmacol.* **168**, 488 (Jan, 2013).
30.   M. I. Fonseca *et al.*, Treatment with a C5aR antagonist decreases pathology and enhances behavioral performance in murine models of Alzheimer's disease. *J. Immunol.* **183**, 1375 (Jul 15, 2009).
31.   H. D. Manthey *et al.*, Complement C5a inhibition reduces atherosclerosis in ApoE-/- mice. *FASEB J.* **25**, 2447 (Jul, 2011).
32.   A. H. Stephan, B. A. Barres, B. Stevens, The complement system: an unexpected role in synaptic pruning during development and disease. *Annu. Rev. Neurosci.* **35**, 369 (2012).
33.   P. Karnchanaphanurach *et al.*, C3b deposition on human erythrocytes induces the formation of a membrane skeleton-linked protein complex. *J. Clin. Invest.* **119**, 788 (Apr, 2009).

Supplemental Figure 2-1. Administration of anti-P at 7.5 dpc Partially Delays Embryonic Loss

A. When compared to anti-properdin at 6.5 and 7.5 dpc, a single dose at 7.5 dpc results in some *Crry*[-/-] embryos weighing between what full size embryos weigh (all are *Crry*[+/-], green) and what resorptions weigh. All *Crry*[-/-] embryos in the liter treated with a single anti-properdin dose at 6.5 dpc were resorbed. Black symbols indicate embryos where no embryonic material could be recovered.

Chapter 3

Improvements to Targeted Sequencing and Methods for Illumina Sample Preparation

Introduction

The expectations of the rare variant common disease hypothesis are that a diversity of rare, relatively recent mutations contribute to common diseases (1, 2). These mutations must be directly sequenced because of their diversity and rarity. This is in contrast to common variants, which are observed efficiently by genotyping. Genotyping platforms have grown significantly and millions of SNPs can be interrogated simultaneously, but the majority of variants in a person's genome are rare and will be invisible to genotyping methods. Ever larger genotyping platforms and those focused on rare coding alleles (Illumina "Exome" Chip) are available but their performance in terms of sensitivity and specificity on very rare variation has yet to be fully studied.

It is now possible to sequence the entire genome of a human for $1,000; however, this cost is still substantial when we consider the sample sizes required. Efforts to select portions of the genome using hybridization-based methods have been incredibly successful (1) but introduce significant additional labor and expense to the process. Recently, the cost of sequencing even the coding 1-2% of the genome (the exome, whole exome sequencing, WES) was approximately 10 to 15% of whole genome sequencing (WGS). As the cost of sequencing itself has fallen, and thus the cost of WGS, WES is now only 25-50% of WGS. As sequencing technologies evolve, WES and WGS will converge in terms of obtaining the data. The analytic and storage costs will remain higher for WGS for longer.

We undertook to significantly decrease the cost and labor required to partition specific areas of the genome, such as all known genes associated with a disease based on GWAS results and/or the coding parts of genes within a pathway. Such an approach would leverage our existing

knowledge about the genetics of disease and what we know about the underlying biology in order to increase our power relative to WES.

Initial efforts to use pooled sequencing of amplicon libraries were daunting because of the amount of follow-up genotyping required (3). When we initially utilized this approach on the Illumina Genome Analyzer IIx, the machine error rate was >1% and even covering each allele 30-fold yielded an overwhelming number of potential variants. The lower error rate and substantially higher output of the HiSeq instrument likely improves the reliability of pooled sequencing. Additionally, because it is based on PCR, the number of genes that could be targeted was limited and did not scale well from five to 50.

Our goal was to adapt a method that allowed for unambiquous variant calling at the level of the individual. This would allow for highly confident genotype calling from the beginning and permit immediate comparison to a number of public datasets that were underway, such as the 1000 Genomes Project (4). Hybridization based strategies that utilized short, custom oligos were an excellent candidate because of their rapid scalability and adaptability. Initially this approach utilized modified array comparative genomic hybridization (aCGH) methods (5) and later transitioned to biotinylated oligos in solution (6, 7). Array based methods were low throughput due to high labor commitments and they were technically challenging. Also, both strategies were very expensive on a per individual basis.

Illumina sequencing allows for the incorporation of a molecular barcode into the Illumina adapter and therefore multiple samples to be sequenced simultaneously on the same flow cell. Utilizing indexing adapters dramatically affects the efficiency of the capture (8-10). A larger proportion of the DNA recovered is not from the targeted regions. This is because the barcode in the adapter allows intermolecular base pairing of the adapters. The targeted fragments of DNA

57

anneal specifically to the oligos, but they bring non-specific fragments with them through these adapter interactions.

We solved this problem by overcoming the challenges introduced by multiplexing samples prior to capture. This allowed us to both decrease the cost per individual of the capture step and also reduce the number of captures being done, resulting in a savings of labor. We validated this method in a number of commercial platforms and ultimately used the Nimblegen platform on over 2,000 individuals.

The other roadblock to scaling targeted capture is library preparation. I simultaneously improved a method of generating Illumina sequencing libraries for capture and sequencing. I focused on optimizing the enzymatic steps required and removed numerous steps yielding a protocol that requires a single cleanup and takes only a few hours to process 96 samples.

Materials and Methods

Illumina library preparation

A detailed step-by-step protocol for making 96 Illumina libraries is in Supplement 3-1.

Illumina Adapter Oligos

Illumina Multiplex Adapter 1:

5′-/5Phos/<u>GAT CGG AAG AGC</u> ACA CGT C*T-3′

Illumina Multiplex Adapter 2:

5′-ACA CTC TTT CCC TAC ACG AC<u>G CTC TTC CGA TC</u>*T-3′

*=phosphorothioate bond. <u>Underlined</u> sequence is shared and the basis of annealing. These were HPLC purified. They must be annealed before use. See supplement 3-1 for protocol for annealing adapters.

58

Indexing and Dual Indexing

We designed 108 indexes, 96 for routine use and an additional twelve. They are 7 bp in length and separated from each other and the 96 7 bp indexes the Washington University Genome Technology Access Center has by 3 errors (11) (Supplementary Table 1). We also designed six base pair indexes that are separated by four errors; these were incorporated into primers that index the other, non-indexed, Illumina adapter arm. This "dual index" is read by priming off the Illumina adapter flow cell following read one and subsequent uracil cleavage.

Indexing Samples with PCR

A detailed protocol is included in Supplement 3-1. Multiplex Primer 1 was purchased HPLC purified or PAGE purified.

Multiplex Primer 1: 5′ AAT GAT ACG GCG ACC ACC GAG ATC TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC *T-3′

The indexing primers used are in Supplement 3-2. They were ordered PAGE purified.

Introducing the Dual Index with PCR

This used the same conditions as as in supplement 3-1, except the annealing temperature was 57°C. Dual index oligos are in Supplement 3-2.

MyGenostics GenCap Capture

We followed the manufacturer's protocol. No custom blocking oligos were used.

Nimblegen SeqCap Arrays and SeqCap EZ Capture

We followed the manufacturer protocol expect the blocking oligos were custom and reflect the indexing strategy we used. 1000 pmol of each oligo was used for a single capture reaction (1 μg of a DNA pool input).

These blocking oligos are:

Block 1: 5′-AAT GAT ACG GCG ACC AACC GAG ATC TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC T -3SpC -3 ′

Block 2 dI: 5′- CAA GCA GAA GAC GGC ATA CGA GAT <u>III III I</u>GT GAC TGG AGT TCA GAC GTG TGC TCT TCC GAT CT -3SpC-3′

Inosines are underlined. Some type of terminator group is required at the 3′ of Block 2 dI to avoid priming off of it in the post-capture PCR and thus erasing the barcode. Both oligos were HPLC purified.

Post-Capture PCR

  We followed the Nimblegen SeqCap protocol, except for these modifications: "post-capture oligo 1" and "post-capture oligo 2" or post-capture oligo 2 and a dual indexing oligo (Supplement 3-2) were used to either preserve the initial index and/or add a second index. Oligos were at a final concentration of 1 μM. Annealing was at 57°C for 30 seconds. Each capture was split into four 50 μl PCR reactions. No betaine was included in the post-capture PCR; it will cause DNA to bind to the streptavidin beads and inhibit the PCR substantially. Post-capture PCR was carried out for 10 cycles if the capture was to pooled prior to sequencing or 12 cycles if it was not. PCR reactions were cleaned up twice with 1.4-fold volume of Serapure beads (12), twice to remove PCR primers.

Post Capture Oligos (desalted, not purified):

Post-Cap 1: 5′-AAT GAT ACG GCG ACC ACC GAG ATC-3′

Post Cap 2: 5′-CAA GCA GAA GAC GGC ATA CGA GAT-3′

Evaluation of Uniformity of Interval Capture

The Genome Analysis Toolkit (GATK) Depth of Coverage function was used to determine the average coverage of each targeted region (i.e. exon) (13).

SybrGold DNA Quantification

SybrGold (Life Technologies) was diluted 1:2000 in 10mM Tris pH 8. 19 μl was mixed with 1 μl of sample in duplicate in a 384-well plate (Greiner #781076). Fluoresence was read on a BioTek Synergy with excitation at 495 nm and emission at 537 nm. Values were compared with standard curve with 11 values between 75 ng/ul and 1 ng/ul.

Read Alignment and Data Processing

See Chapter 4.

Results

Novel Illumina Library Preparation Method Without Cleanup

Core labs and third party contractors often require 1 μg of genomic DNA for preparing Illumina libraries. This is a significant amount of DNA for many repositories and can be an impediment to conducting sequencing experiments. Using <500 ng, and often only 200 ng (~1000 samples), we constructed nearly 3,000 Illumina libraries using a modified protocol that removes all cleanup steps until after ligation. This protocol also keeps the DNA in the same tube minimizing losses due to transfer.

This is accomplished by using temperature to modulate the activity of each enzyme. T4 DNA polynucleotide kinase and T4 DNA polymerase both are active at 25°C and heat inactivatable. End repair of physically sheared DNA is carried out with these enzymes in 10x

ligation buffer with dNTPs. Subsequently, Taq polymerase is added instead of the Klenow exonuclease deficient (exo-) polymerase fragment because it is active in the temperature range that inactivates the end-repair enzymes. Like Klenow exo-, Taq has a ~1,000 fold preference for adenine over guanine addition to a non-overhanding 3′ end (14). Finally, ligation of the annealed Illumina adapters occurs at temperatures at which Taq is not functional (25°C). This allows the entire protocol to be done quickly with no cleanups (Figure 3-1). This method was highly reproducible in terms of the size of library recovered and the success rate. Of 690 libraries made in one set, only one failed.

In a second group of 950 individuals using 200 ng of DNA, 24 failed in a first attempt. When these 24 individuals were re-processed, including shearing, 22 of 24 were successful.
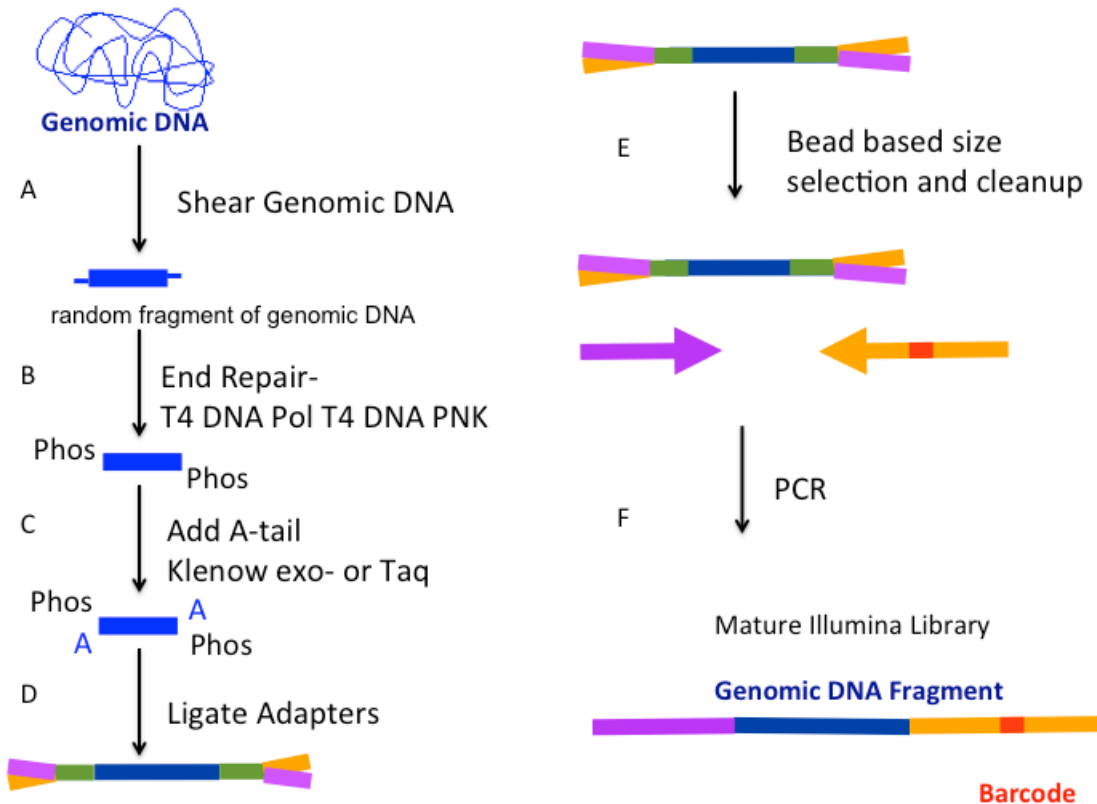
Figure 3-1. Illumina Library Preparation
A. DNA is randomly fragmented. B. Fragmented DNA has overhanging edges which is filled in with T4 DNA polymerase. T4 polynucleotide kinase phosphorylates the 5′ hydroxyl. C. An adenosine is added to the 3′ position at each end of the DNA fragment with either Klenow (exo-) fragment or Taq polymerase. D. Illumina adapters with an overhanging "T" are ligated onto the DNA fragment, utilizing TA cloning. These adapters are asymmetric with the common part shown in green and non-common part in purple and orange. E. Bead based size selection removes adapter-dimers and fragments below the desired size. F. PCR is done with primers specific for each of the non-complimentary Illumina adapter arms. One primer also contains a unique set of nucleotides, the barcode (red).

Covaris Fragmentation Is Optimal Compared with Plastic PCR Plates

DNA must be randomly fragmented prior to making an Illumina sequencing library. We required scalable methods that were reliable and wasted very little DNA. 96 well PCR plates can be used with the Covaris system in place of Covaris microTubes to reduce the cost of library preparation (15). However, in practice this protocol is extremely labor intensive as it is done in a very large volume relative to the plate well and requires subsequently concentrating the DNA by binding it to beads. Using such large volumes and additional manipulations also makes cross contamination more likely. Finally, the shearing that is produced is not nearly as tightly focused as that produced using Covaris microTubes on the Covaris E-series instrument (Figure 3-2). We utilized microTubes because they were extremely reliable, did not waste DNA in the fragmentation process, and the small volumes achievable simplified the workflow.
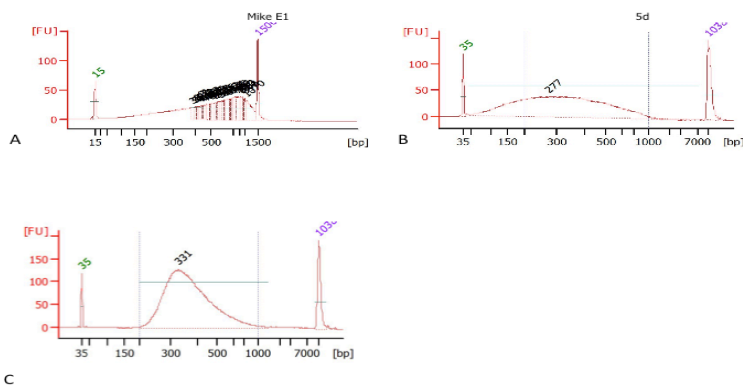
Figure 3-2. Covaris Shearing Is Narrowly Focused Compared to Plastic Plate Based Shearing
A. Initial attempts to do shearing based on the protocol in Fisher et. al. (15). B. Best results were obtained in a 96-well PCR plate (Applied Biosystems MicroAmp PCR plate). C. An Illumina library (post-PCR) using DNA sheared in a Covaris microTube.

Cleanup with "SeraPure" Beads versus AMPure XP Beads

We utilized a previously described "home-made" magnetic bead purification slurry based on the same properties utilized by AMPure beads (12) (16). This method was able to perform size selection as well. It removed any Illumina adapter dimers created during ligation prior to PCR (Figure 3-3). We saw no decrease in yield for ~1,200 libraries compared to ~1,500 libraries that utilized AMPure XP beads.
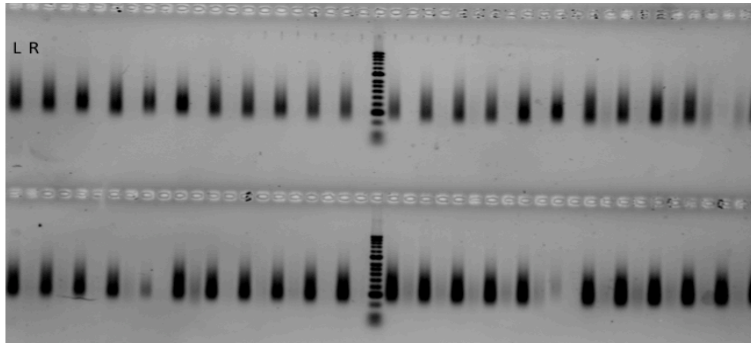
Figure 3-3. Post Indexing PCR Versus Pre-PCR Library
Illumina libraries, after 10 cycles of PCR with an Illumina indexing primer followed by cleanup with 1.2x magnetic beads (L). An equivalent amount of sheared library, after adapter ligation and bead clean-up (R). They are loaded in pairs and run on a 2% agarose gel to assess successful library construction

Nimblegen Sequence Capture Can be Efficiently Multiplexed with Degenerate Blocking Oligos

The original Nimblegen SeqCap protocol utilized the Illumina PCR primers for blocking the adapter arms during hybridization at ~100-fold molar excess compared to the Illumina library being hybridized. The incorporation of indexes into these barcodes means that the blocking oligos will only pair properly with one end. To correct this, we utilized blocking oligos with the an additional 7 bp where the index resides. Further, to make a generalized, oligo we utilized deoxyinosine, which pairs promiscuously with adenine, cytosine, or thymine (Figure 3-4).

Figure 3-4. Capturing Indexed Samples Requires Blocking the Index.
A. Our preferred strategy utilized a longer blocking oligo for the indexed adapter arm. The nucleotides encoding the barcode were represented as either a set of inosines, which will base pair promiscuously, or degenerate bases. B. An alternative strategy is to use two short oligos for each part of the barcoded adapter. This leaves the barcode itself unpaired and was not our preferred method. The shorter oligos may require a locked nucleic acid to make the base pairing stable during hybridization conditions.

Nimblegen Sequence Capture Outperforms MyGenostics GenCap

In terms of the percentage of reads on target and the diversity of the Illumina libraries on a per-individual level, SeqCap EZ captures outperformed Nimblegen Sequence Capture Arrays and the MyGenostics GenCap (Table 3-1). The MyGenostics capture had widely varying capture efficiencies. The SeqCap EZ was technical simple, robust and performed well with multiplexes up to 200 individuals.

Table 3-1. Nimblegen SeqCap Outperforms Other Capture Methods in Multiplex Capture

Capture	Ind./Capture	Avg % On Target	% Duplicates
SeqCap Array	24 {2}	67	24
SeqCap Array	48 {1}	70	48
MyGenostics	12 {3}	38 ± 5	26 ± 3
MyGenostics	24 {4}	37 ± 5	41 ± 9
MyGenostics	50 {3}	31 ± 6	42 ± 11
SeqCap EZ	100 {8}	62.5 ± 3	8.7 ± 6
SeqCap EZ	200 {1}	67.9 ± 2	10.4 ± 2

Ind./Capture: Number of individuals uniquely indexed and captured together (ie "plex"). Number in brackets is number of captures at that "plex". SeqCap EZ samples summarized here were sequenced to a median of ~70x average coverage for a 1 Mb target. SeqCap Array: Nimblegen 2.1M SeqCap Arrays. MyGenostics: GenCap capture. Both the 2.1M SeqCap arrays and GenCap targeted ~580 kb. SeqCap EZ is a solution-phase capture utilizing long biotinylated DNA oligos.

High Levels of Multiplexing Do Not Affect Capture Efficiency

We systematically increased the number of individuals captured at one time from 24 to 200 and saw very little effect on the specificity of the capture reaction and only a small effect on the diversity of post-capture libraries (percent duplicate reads). The difference between 100 and 200 individuals was 1.7% of reads being duplicates at 70-fold average coverage for a 1 Mb target region (Table 3-1).

Concordance to Array Based Genotypes

91 individuals that were sequenced also had array based genotyping data at 277 SNPs overlapping our targeted regions. At the individual level, we observed on average 99.8 ± 0.4% concordance (± standard deviation, maximum 1.0, minimum 96.4%) for the 277 SNPs between sequencing and arrays at a median depth of 70-fold coverage. Similar statistics were obtained for the average concordance for each SNP, 99.8% (maximum 100, minimum 96.7%).

Empirical Rebalancing of Captures Improves Uniformity

Nimblegen capture probes are designed to be isothermal, which should mean that they all perform similarly in terms binding DNA during hybridization. In practice they do not perform equally well. The CV (standard deviation/average) for target interval average coverage in one initial design was 0.53 when we considered the coverage of all targets across 690 individuals captured (Figure 3-5). We empirically rebalanced the number of times a probe was replicated in the capture design for a second set of individuals captured for the same regions. We did this by assessing how many fold below the median an interval was and increasing the number of probes accordingly, up to a 12 fold increase. This improved the uniformity of the capture design. 950 individuals captured for the same regions had a CV of 0.40.

68
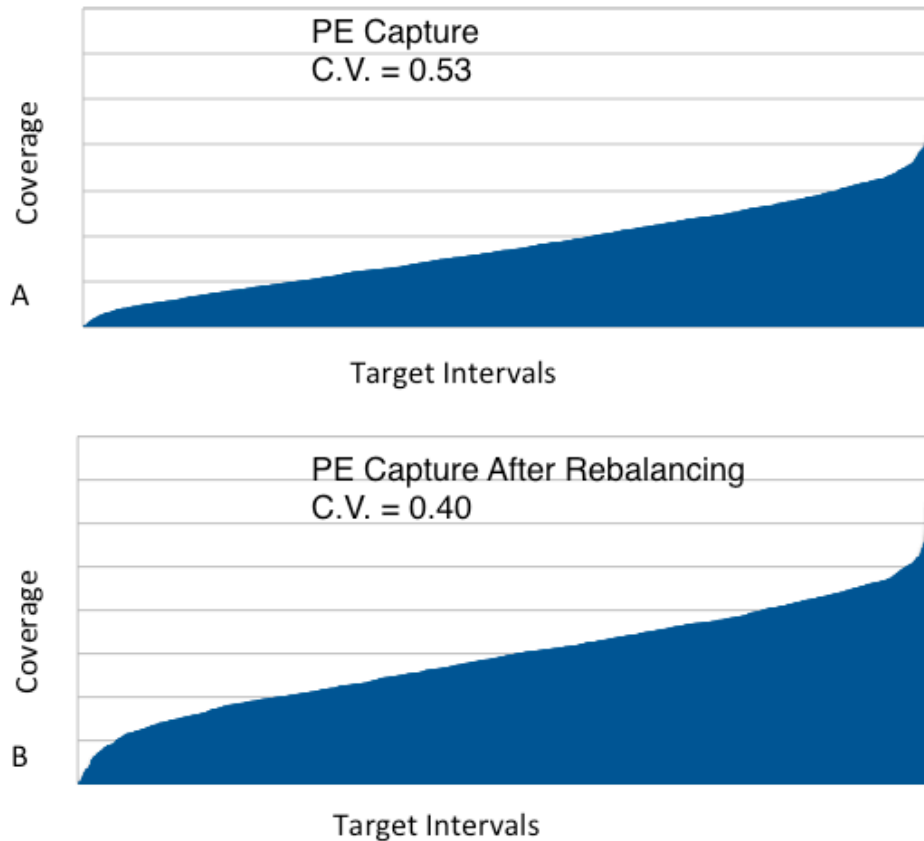
Figure 3-5. Empirical Rebalancing of Nimblegen SeqCap EZ Probe Density Improves Capture Uniformity
A. Variability in the coverage of intervals for the initial sequence capture design for our preeclampsia sequencing (737 kb) yielded a coefficient of variance of 0.53 across all 700 individuals. B. Increasing the probe density for regions that were captured below the median improved the uniformity across all captured regions. The coefficient of variance was 0.40 across a different 950 individuals.

Read Based Pooling is Superior to Pooling of Equivalent Masses

10 pools of 95 individuals were made using an equal volume of each sample. These were then assigned a unique dual index using PCR. All ten pools were sequenced on a MiSeq using a 2x30 bp run with two index reads (7 bp index and 6 bp dual index). The relative concentration of each individual to the pools was assessed.

69

Pooling using relative read-depth based concentration estimates outperformed mass based pooling. Read-depth based pooling yielded pools exhibiting CV of 0.04-0.0.13, in contrast to mass based pooling that exhibited CVs of 0.25 to 0.3 (Figure 3-6). Mass based pools could be corrected based on their relative read-based concentrations to some extent (Table 3-2).

Over multiple captures, the actual CV of the reads assigned to each individual after the capture enrichment step was greater than 0.2 (not all captures shown). In the ten pools we evaluated for relative read-based capture versus mass based capture the CV of pools went up for samples that had been pooled based on relative read-depth. Capture had more of an effect on these pools than it did on mass based pools, but it in the final post capture sequencing, read-based pools were still better because they started with such a uniform composition (Table 3-2).

Table 3-2. Read Depth Based Pooling Outperforms Mass Based Pooling

VIP Group	CV of Pool	CV After Correction	Capture CV
1	4.1	3.6	12
2	5.7	5.6	11.6
3	9.4	-	16.8
4	30.5[†]	20.4[‡]	14.3
5	8.4	-	14.5
6	6.5	5.6	17.4
7	13.1	-	17.5
8	10.9	-	12.7
9	6.6	-	9.1
10	26.8[†]	19.0[‡]	11.2

CV: coefficient of Variance (average/standard deviation). [†]These pools were initially constructed using mass based metrics. [‡]After sequencing the pool, corrections were calculated based on read depth, and this was sequenced again. Capture CV is based on the sequencing done after capture.
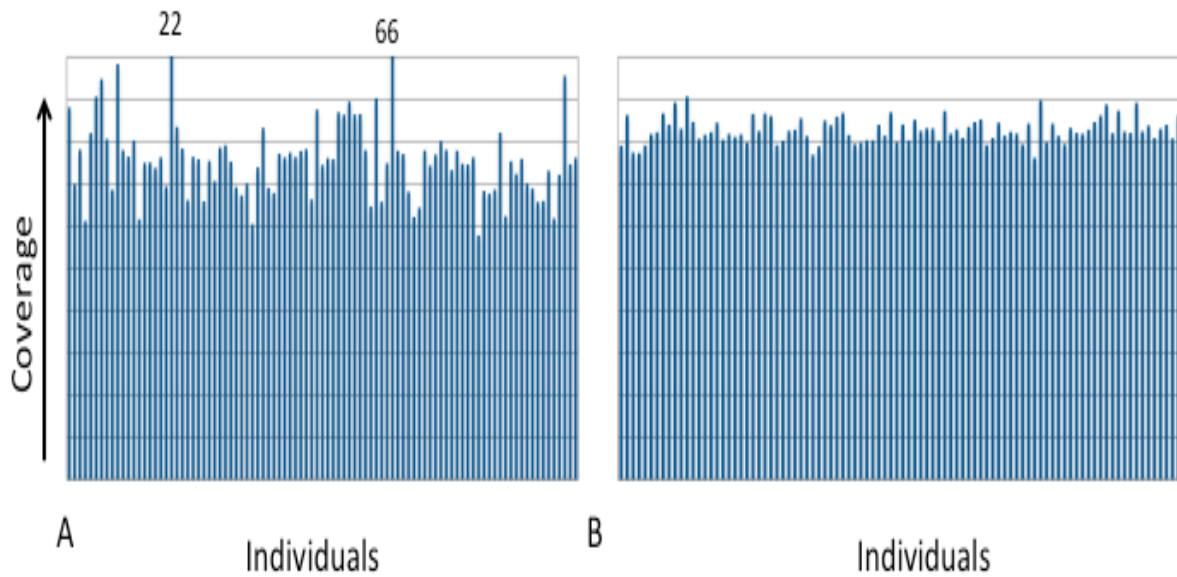
Figure 3-6. MiSeq Read-Depth Based Normalization is Superior to Mass Based Pooling.
A. Pool of 96 samples based on combining equal masses of each sample. Masses were derived
from a SybrGold based assay utilizing a standard curve. All samples were quantified in duplicate
in a 384-well plate. Samples 22 and 66 are off the scale (actually double the height of the
presented y-axis). The coefficient of variance of the pool in panel A is 30%. B. Pool of 96
individuals combined based on their relative read depths from a pool of equal volumes. The
coefficient of variance of this pool is 3.6%. Using the MiSeq platform and dual indexing,
hundreds to thousands of samples can be normalized at the same time.

Discussion

Partitioning the genome into regions of interest, such as subsets of genes will continue to

be of interest until WGS becomes so affordable it is widely available in society. Until then WES

and partitioning even smaller subsets will remain an attractive way to identify rare variants.

Partitioning the genome introduces substantial labor and costs that do not scale linearly with the

subset of the genome targeted.

This work helps reduce the cost in terms of labor and money in two ways: first by

streamlining library prep, and second by reducing the number of captures that must be done.

Others recognized the need to simplify these steps too. Independently, Qiagen began marketing a "one-tube" library preparation kit. Nimblegen now supports multiplex capture, though our method was developed independently. A variety of other multiplex capture strategies have either been published or developed since we undertook to do this as well (17) (12) (Andrey Shaw, unpublished). While our pools of 100 to 200 individuals worked well for the purpose of detecting germline mutations, we did not evaluate if smaller pools would be suitable for applications that require very high depths of coverage (>200x).

Normalization of samples is incredibly important to maximize the number of individuals that can be sequenced simultaneously. For small numbers of samples, qPCR works well. But for hundreds or thousands of samples, read based pooling is superior to mass based pooling and requires significantly less time and labor while being more accurate.

Multiplex capture is another tool that can be used to identify rare variation in genes already associated with a disease. It is more scalable than MIP based methods (18) and allows immediate assignment of individual genotypes compared to pool based methods (3).

References

1   Tennessen, J. A., A. W. Bigham, T. D. O'Connor, W. Fu, E. E. Kenny, S. Gravel, S. McGee, R. Do, X. Liu, G. Jun, H. M. Kang, D. Jordan, S. M. Leal, S. Gabriel, M. J. Rieder, G. Abecasis, D. Altshuler, D. A. Nickerson, E. Boerwinkle, S. Sunyaev, C. D. Bustamante, M. J. Bamshad, and J. M. Akey. 2012. Evolution and functional impact of rare coding variation from deep sequencing of human exomes. *Science* 337:64-69.

2   Fu, W., T. D. O'Connor, G. Jun, H. M. Kang, G. Abecasis, S. M. Leal, S. Gabriel, M. J. Rieder, D. Altshuler, J. Shendure, D. A. Nickerson, M. J. Bamshad, and J. M. Akey. 2013. Analysis of 6,515 exomes reveals the recent origin of most human protein-coding variants. *Nature* 493:216-220.

3   Vallania, F. L., T. E. Druley, E. Ramos, J. Wang, I. Borecki, M. Province, and R. D. Mitra. 2010. High-throughput discovery of rare insertions and deletions in large cohorts. *Genome Res.* 20:1711-1718.

4   Pennisi, E. Genomics. 1000 Genomes Project gives new map of genetic diversity. *Science* 330:574-575.

5   Hodges, E., M. Rooks, Z. Xuan, A. Bhattacharjee, D. Benjamin Gordon, L. Brizuela, W. Richard McCombie, and G. J. Hannon. 2009. Hybrid selection of discrete genomic intervals on custom-designed microarrays for massively parallel sequencing. *Nat Protoc* 4:960-974.

6   Hodges, E., Z. Xuan, V. Balija, M. Kramer, M. N. Molla, S. W. Smith, C. M. Middle, M. J. Rodesch, T. J. Albert, G. J. Hannon, and W. R. McCombie. 2007. Genome-wide in situ exon capture for selective resequencing. *Nat. Genet.* 39:1522-1527.

7   Gnirke, A., A. Melnikov, J. Maguire, P. Rogov, E. M. LeProust, W. Brockman, T. Fennell, G. Giannoukos, S. Fisher, C. Russ, S. Gabriel, D. B. Jaffe, E. S. Lander, and C. Nusbaum. 2009. Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. *Nat. Biotechnol.* 27:182-189.

8   Teer, J. K., L. L. Bonnycastle, P. S. Chines, N. F. Hansen, N. Aoyama, A. J. Swift, H. O. Abaan, T. J. Albert, E. H. Margulies, E. D. Green, F. S. Collins, J. C. Mullikin, and L. G. Biesecker. 2010. Systematic comparison of three genomic enrichment methods for massively parallel DNA sequencing. *Genome Res.* 20:1420-1431.

9   Nijman, I. J., M. Mokry, R. van Boxtel, P. Toonen, E. de Bruijn, and E. Cuppen. 2010. Mutation discovery by targeted genomic enrichment of multiplexed barcoded samples. *Nat. Methods* 7:913-915.

10   Shearer, A. E., M. S. Hildebrand, H. Ravi, S. Joshi, A. C. Guiffre, B. Novak, S. Happe, E. M. LeProust, and R. J. Smith. 2012. Pre-capture multiplexing improves efficiency and cost-effectiveness of targeted genomic enrichment. *BMC Genomics* 13:618.

11   Meyer, M., and M. Kircher. 2010. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harbor protocols* 2010:pdb prot5448.

12   Rohland, N., and D. Reich. 2012. Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. *Genome Res.* 22:939-946.

13   McKenna, A., M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytsky, K. Garimella, D. Altshuler, S. Gabriel, M. Daly, and M. A. DePristo. 2010. The Genome

Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20:1297-1303.

14   Clark, J. M. 1988. Novel non-templated nucleotide addition reactions catalyzed by procaryotic and eucaryotic DNA polymerases. *Nucleic Acids Res* 16:9677-9686.

15   Fisher, S., A. Barry, J. Abreu, B. Minie, J. Nolan, T. M. Delorey, G. Young, T. J. Fennell, A. Allen, L. Ambrogio, A. M. Berlin, B. Blumenstiel, K. Cibulskis, D. Friedrich, R. Johnson, F. Juhn, B. Reilly, R. Shammas, J. Stalker, S. M. Sykes, J. Thompson, J. Walsh, A. Zimmer, Z. Zwirko, S. Gabriel, R. Nicol, and C. Nusbaum. 2011. A scalable, fully automated process for construction of sequence-ready human exome targeted capture libraries. *Genome biology* 12:R1.

16   DeAngelis, M. M., D. G. Wang, and T. L. Hawkins. 1995. Solid-phase reversible immobilization for the isolation of PCR products. *Nucleic Acids Res* 23:4742-4743.

17   Ramos, E., B. T. Levinson, S. Chasnoff, A. Hughes, A. L. Young, K. Thornton, A. Li, F. L. Vallania, M. Province, and T. E. Druley. 2012. Population-based rare variant detection via pooled exome or custom hybridization capture with or without individual indexing. *BMC Genomics* 13:683.

18   O'Roak, B. J., L. Vives, W. Fu, J. D. Egertson, I. B. Stanaway, I. G. Phelps, G. Carvill, A. Kumar, C. Lee, K. Ankenman, J. Munson, J. B. Hiatt, E. H. Turner, R. Levy, D. R. O'Day, N. Krumm, B. P. Coe, B. K. Martin, E. Borenstein, D. A. Nickerson, H. C. Mefford, D. Doherty, J. M. Akey, R. Bernier, E. E. Eichler, and J. Shendure. 2012. Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders. *Science* 338:1619-1622.

<u>Supplement 3-1</u>
<u>Illumina Multiplex Library Protocol</u>

<u>Shearing:</u>
NOTE: shearing takes place in 55ul. 6ul of that is 10x Ligase Buffer. The rest is DNA and H_2O.
NOTE: Pulse spin is a table top swinging bucket centrifuge spinning up to ~1200.

This protocol works very well for 100ng-1ug. We shoot for 500ng, but there are obviously quantitation errors.

Quantities in excess of 1ug may need adjustment of one or more steps, specifically the amount of adapter.

Component	Volume for 1 sample
10x Ligase Buffer	6ul
H20	55-6=49-Vol DNA
Total	55-Vol DNA

1) In a prep plate, place the H_2O and 10x Ligase Buffer mix in each well of a 96 well PCR plate. "Shearing Prep Plate".
2) Check orientation of gDNA plate and "Shearing Prep Plate". <u>A1 in upper left.</u> Add DNA to H_2O and 10x Ligase Buffer.
3) Pulse spin the Shearing Prep Plate.
4) Check orientation of Shearing Prep Plate and Covaris 96 well plate (PN: 520078).
5) Transfer samples from Shearing Prep Plate to Covaris 96 well plate.
6) Cover with aluminum seal provided. Cover with ELISA Plate film.
7) Spin Plate.
8) Take to GTAC.
   a. Make an appointment in advance. email Matt Minton: mminton@pathology.wustl.edu or Toni Sinnwell tsinnwell@path.wustl.edu to reserve.)
9) Use the following for 55 μl volume and a 250 bp size.

Duty Cycle	**10%**
Intensity	5
Cycles/Second	200
Time (sec)	85
SHORTER SHEARING TIME WILL RESULT IN A LARGER FRAGMENT DISTRIBUTION.	50seconds = 300bp 85seconds = ~250bp 120 seconds = 200bp

10) Check that the number of steps/samples on the Sonolab software is the same as the number of samples you will do.
11) Save the protocol with a unique name. You must save before starting.
12) Click Check Alignment (after you have finalized the protocol). This will place the carrier arm at the same height and position it will start the protocol at.

13) Check water bath level. Add Plate. Water should be up to top of skirt but not over the aluminum foil cover. Water can be added from MQ Carboy on bench.
14) Use "Change Bath" to raise transducer to adjust water level if too much water.
15) Shear. Check to make sure the AFA fiber is shaking in the first well.
18) Can leave plate overnight at 4°C after shearing is done.

THAWING/PREP STEP:
Thaw 25mM each dNTPs (Bioline).
Thaw Enzymatics 10x Taq Buffer.

End Repair:
1) Pulse spin the Covaris plate.
2) Transfer samples out of Covaris plate into a new plate, the "Construction Prep Plate". Using a multichannel with range ~5-200.
3) Pulse spin the Covaris plate.
4) Use a P10 multichannel to transfer any residual sheared DNA (uses 10ul skinny tips to get right to the bottom).
5) Spin the "Construction Prep Plate."
6) Seal "Construction Prep Plate".
7) Make End Repair Mix

Component	1x	100x
Sheared DNA in ~1x Ligase Buffer	50	NOT IN MIX. ADD LATER
dNTPs (25mM each)	1	100ul
Enzymatics T4 DNA Pol (3k U / ml)	2.25	225ul
Enzymatics T4 PNK (10k U / ml)	2.25	225
10x Ligase Buffer	0.545	54.5
CLEAN H20	3.955	395.5
10ul of mix		

8) Use a P10 Multichannel to put 10ul End Repiar Mix in each well of a semi-skirted plate, "Library Making Plate".
9) Transfer 50ul of the Sheared DNA from "Construction Prep Plate" to "Library Making Plate". Pipette each row/column up and down 2x.
10) Cover with sealer film.
11) Place in Thermo cycler. 40 min at 20°C and then 25 min at 75°C to heat inactivate. Use Heated Lid.

A-tail:
We use Taq here. It is possible to use Klenow Exo (-) here instead in the same way. Replace Taq Buffer with T4 Ligase Buffer and adjust volume, 3ul of Klenow (Exo -) and temperature accordingly (20 min 37°C, 20 min 75°C).

We can add Taq or Klenow exo (-) here because they have a ~1000-fold preference for A over T addition to a blunt end, and very low affinity for G or C transfer to blunt ends.

It may be possible to add the dATP and Taq at the beginning and length the heat-inactivation step of the End Repair so it adds the A-tail too.

A-Tail Reaction

Component	1x	104x
60ul End-Repair Rxn	60	NOT IN MIX!
dATP (10mM)	0.65	67.6
CLEAN H$_2$O	1.95	202.8
10X Enzymatics Taq Buffer	0.5	52
Enzymatics Taq	1.9	197.6
5ul of mix!		

12) Pulse spin.
13) Add 5ul of A-tail mix with P10 multichannel.
14) Set P200 multichannel to 60ul. Use to mix each sample 3x.
15) Incubate at 70°C for 25 min.

THAWING/PREP STEP
Make FRESH 80% EtOH with MQ H$_2$O and clean EtOH.
Get 25uM Multiplex Adapter out of -20°C
Get 2X Rapid Ligase Buffer out of -20°C.
Thaw 5xHF Buffer.

Adapter Annealing:
Resuspend adapters in 1x HF Buffer (From Phusion Kit) to 100 μM.
Combine them 1:1.
98°C 5 min and then cool at 0.1°C/second down to 15°C.
Dilute 1:1 with 10mM Tris, 0.1mM EDTA. Final concentration is 25 μM.

Adapter Ligation:
Prepare the following 104x reaction ON ICE.

Reagent	1x	104x
A-tailed Rxn	65ul	NOT IN MIX!!!!
2x Rapid Ligase Buffer	18.5ul	1924ul (962x2)
25uM Multiplex Adapter Annealed	1.5ul	156
Enzymatics High Conc. Ligase	1ul	104-ADD LAST!
CLEAN H$_2$O	1ul	104
22ul to each sample		

Keep Ligation mix on ice.
Transfer 22 μl of mix to Library Making Plate. With Plate on ice.
Mix 2x with multichannel set to 80 μl.
Incubate 25°C for 20 min.
CAN STOP HERE AND FREEZE!

THAWING/PREP STEP
Thaw Multiplex Primer 1 and Indexing Primers.
Gather 5M Betaine, 5X HF Buffer and dNTPs (25mM each).

The ligation already has 3.2% PEG from the ligation buffer. We must take this into account
when adding the cleanup beads or any Illumina adapter dimers will be carried forward into PCR.

Beads retain DNA fragments >150bp at ~8% PEG. Below 6.7% PEG almost no DNA was bound
to beads.

Clean with 0.65x SERAPURE / AMPURE XP BEADS (Rohland et. al.) (12):
   1) Transfer 56.5ul of beads (0.65x) to each well.
   2) Mix 10x.
   3) Cover. Allow DNA to bind beads for 10 minutes.
   4) Place on magnetic plate for 10 minutes. The beads are much more stable after 10 minutes
      than after 5 minutes.
   5) Remove Bead/PEG Supernatant from wells (set pipette to 150ul) with multichannel. Use
      a smooth action to prevent disturbing beads.
   6) Wash with 200ul FRESH 80% EtOH.
   7) Incubate 30sec with 80% EtOH. Remove EtOH. Repeat 30 sec EtOH wash.
   8) Let wells air dry 3minutes.
   9) Aspirate any EtOH that has gathered in the base of the well.
   10) Dry 2-5 more minutes. Do not dry longer than 10 minutes.
   11) Elute in 40ul Buffer HF/H20 mix.

Component	1x	110x
5x Buffer HF	10.7	1177 (2x588.6)
H20	29.3	3223 (4x805.75)

   12) Place plate on magnet and transfer the whole volume to a new plate. Cover.
   13) Gently spin new plate. Separate on magnet.
   14) Second separation is a good idea to prevent ANY bead carryover into PCR. Betaine acts
   like PEG and if there are sufficient beads the bead+Betaine combo will substantially inhibit
   the PCR reaction (qPCR estimates of 1000-fold inihibition).

Multiplexing PCR: This will select for fragments with an adapter on each end and will ensure that the finished library is predominately molecules with a different adapter on each end.

Get a plate with 1 µl of Indexing Primer in each well or alternatively place the primer in after the PCR master mix and Cleaned Ligation Product.

Reagent	1x	104x
Cleaned Ligation Product	36 µl	NOT IN MIX
5x HF Buffer	1.37 µl	142.48
5M Betaine	11ul	1144
dNTPs, 25mM each mix	.44	45.76
Multiplexing Primer 1 @ 25mM	1	98.8
Indexing Primer @25mM	1 already in plate	NOT IN MIX
Phusion polymerase	1	104
CLEAN H$_2$O	3.19	352.56
55 ul rxn	18 to each (Indexing primer is separate)	

1) Add 18 µl to each well.
2) Transfer the 36ul from the 2nd separation of the post-adapter ligated, cleaned sample to the PCR rxn. Mix 5x.
3) Save remained for gel analysis
4) Cover. Pulse spin.

Indexing PCR: (10 cycles). If yield is a concern either because very low input or will capture one individual at time, do 12 – 14 cycles.
1) 98°C 30 sec
2) 98°C 10 sec
3) 62°C 30 sec
4) 72°C 45 sec
5) go to 2, 9 times.
6) 72°C 5 min
7) 8°C forever.

We want to limit the number of cycles to prevent artifacts from being introduced and to limit the skew in fragment representation. Skewing results from the molecules that are "lucky" enough to be amplified in the first round; they will have a small chance of becoming a disproptionately large percentage of the mixture.

Can freeze this reaction or run on gel the same day.
Run 2ul of the pre-indexing pcr (cleaned ligation product) with 2.75ul of uncleaned PCR product on a 2% agarose gel.

Goal: See a fragment of ~the correct size in all lanes (200-500bp). PCR-Matured Adapter Dimer will be ~143bp. This must be removed with a 1 x bead cleanup prior to capture.

Purify with 1.2x SERAPURE/ Ampure XP Beads:
Check volume of remaining PCR reacations.
Clean with 1.2x beads as above.

Elute with 75ul of Clean H_2O.
Let sit for 5 minutes, place on magnet for 5 minutes,.
Transfer supernatant to a new plate labeled with SAMPLES, DATE, INDEXES USED, and
CLEANUP VOLUME and ELUTION VOLUME/TE or H_2

Supplement 3-2   Multiplexing Indexes

Index	Name	Primer Sequence
AACCAGG	MP_T_1	CAAGCAGAAGACGGCATACGAGATcctggttGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
AACCGTC	MP_T_2	CAAGCAGAAGACGGCATACGAGATgacggttGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
AACGCGA	MP_T_3	CAAGCAGAAGACGGCATACGAGATtcgcgttGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
AAGATGA	MP_T_4	CAAGCAGAAGACGGCATACGAGATtcatcttGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
AAGTATA	MP_T_5	CAAGCAGAAGACGGCATACGAGATtatacttGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
AATAGAA	MP_T_6	CAAGCAGAAGACGGCATACGAGATttctattGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
ACCGGCC	MP_T_7	CAAGCAGAAGACGGCATACGAGATggccggtGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
ACGAAGT	MP_T_8	CAAGCAGAAGACGGCATACGAGATacttcgtGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
ACGCATC	MP_T_9	CAAGCAGAAGACGGCATACGAGATgatgcgtGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
ACTACTA	MP_T_10	CAAGCAGAAGACGGCATACGAGATtagtagtGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
ACTCAAG	MP_T_11	CAAGCAGAAGACGGCATACGAGATcttgagtGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
AGAACGG	MP_T_12	CAAGCAGAAGACGGCATACGAGATccgttctGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
AGAGCAA	MP_T_13	CAAGCAGAAGACGGCATACGAGATttgctctGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
AGCTCTT	MP_T_14	CAAGCAGAAGACGGCATACGAGATaagagctGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
AGGCCGC	MP_T_15	CAAGCAGAAGACGGCATACGAGATgcggcctGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
AGTAGGT	MP_T_16	CAAGCAGAAGACGGCATACGAGATacctactGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
AGTCTAC	MP_T_17	CAAGCAGAAGACGGCATACGAGATgtagactGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
ATAAGCG	MP_T_18	CAAGCAGAAGACGGCATACGAGATcgcttatGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
ATACGAT	MP_T_19	CAAGCAGAAGACGGCATACGAGATatcgtatGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
ATCATCA	MP_T_20	CAAGCAGAAGACGGCATACGAGATtgatgatGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
ATGAGGC	MP_T_21	CAAGCAGAAGACGGCATACGAGATgcctcatGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
ATTAAGG	MP_T_22	CAAGCAGAAGACGGCATACGAGATccttaatGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
ATTCCAA	MP_T_23	CAAGCAGAAGACGGCATACGAGATttggaatGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
ATTGATT	MP_T_24	CAAGCAGAAGACGGCATACGAGATaatcaatGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
CAAGGAC	MP_T_25	CAAGCAGAAGACGGCATACGAGATgtccttgGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
CAATTCT	MP_T_26	CAAGCAGAAGACGGCATACGAGATagaattgGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
CAGAAGG	MP_T_27	CAAGCAGAAGACGGCATACGAGATccttctgGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
CAGCCTA	MP_T_28	CAAGCAGAAGACGGCATACGAGATtaggctgGTGACTGGAGTTCAGACGTGTGCTCTTCCGA

Barcode	ID	Sequence
CATGGGTA	MP_T_29	CAAGCAGAAGACGGCATACGAGATtaccatgGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
CCAAGAG	MP_T_30	CAAGCAGAAGACGGCATACGAGATctcttggGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
CCGCTCG	MP_T_31	CAAGCAGAAGACGGCATACGAGATcgagcggGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
CGACCAG	MP_T_32	CAAGCAGAAGACGGCATACGAGATctggtcgGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
CGATCTA	MP_T_33	CAAGCAGAAGACGGCATACGAGATtagatcgGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
CGATGCG	MP_T_34	CAAGCAGAAGACGGCATACGAGATcgcatcgGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
CGCGAGG	MP_T_35	CAAGCAGAAGACGGCATACGAGATcctcgcgGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
CGCGGAT	MP_T_36	CAAGCAGAAGACGGCATACGAGATatccgcgGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
CGGAGAA	MP_T_37	CAAGCAGAAGACGGCATACGAGATttctccgGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
CGGTATT	MP_T_38	CAAGCAGAAGACGGCATACGAGATaataccgGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
CTAGACT	MP_T_39	CAAGCAGAAGACGGCATACGAGATagtctagGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
CTATGAA	MP_T_40	CAAGCAGAAGACGGCATACGAGATttcatagGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
CTATTGC	MP_T_41	CAAGCAGAAGACGGCATACGAGATgcaatagGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
CTCGGTC	MP_T_42	CAAGCAGAAGACGGCATACGAGATgaccgagGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
CTCTAGT	MP_T_43	CAAGCAGAAGACGGCATACGAGATactagagGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
CTGACGT	MP_T_44	CAAGCAGAAGACGGCATACGAGATacgtcagGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
CTGATAC	MP_T_45	CAAGCAGAAGACGGCATACGAGATgtatcagGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
CTGCGCA	MP_T_46	CAAGCAGAAGACGGCATACGAGATtgcgcagGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
CTTCCTG	MP_T_47	CAAGCAGAAGACGGCATACGAGATcaggaagGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
CTTGAAC	MP_T_48	CAAGCAGAAGACGGCATACGAGATgttcaagGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
GAACCAA	MP_T_49	CAAGCAGAAGACGGCATACGAGATttggttcGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
GAATAAT	MP_T_50	CAAGCAGAAGACGGCATACGAGATattattcGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
GAATCGG	MP_T_51	CAAGCAGAAGACGGCATACGAGATccgattcGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
GACGTAC	MP_T_52	CAAGCAGAAGACGGCATACGAGATgtacgtcGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
GAGATAT	MP_T_53	CAAGCAGAAGACGGCATACGAGATatatctcGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
GAGCTGC	MP_T_54	CAAGCAGAAGACGGCATACGAGATgcagctcGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
GCAGTAA	MP_T_55	CAAGCAGAAGACGGCATACGAGATttactgcGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
GCATGAC	MP_T_56	CAAGCAGAAGACGGCATACGAGATgtcatgcGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
GCCGAAG	MP_T_57	CAAGCAGAAGACGGCATACGAGATcttcggcGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
GCCGTCT	MP_T_58	CAAGCAGAAGACGGCATACGAGATagacggcGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
GCGACCG	MP_T_59	CAAGCAGAAGACGGCATACGAGATcggtcgcGTGACTGGAGTTCAGACGTGTGCTCTTCCGA

GCGGTTC	MP_T_60	CAAGCAGAAGACGGCATACGAGATgaaccgcGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
GCGTTAG	MP_T_61	CAAGCAGAAGACGGCATACGAGATctaacgcGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
GCTCCGT	MP_T_62	CAAGCAGAAGACGGCATACGAGATacggagcGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
GCTCTCC	MP_T_63	CAAGCAGAAGACGGCATACGAGATggagagcGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
GCTGCTG	MP_T_64	CAAGCAGAAGACGGCATACGAGATcagcagcGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
GGAATAG	MP_T_65	CAAGCAGAAGACGGCATACGAGATctattccGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
GGATGGA	MP_T_66	CAAGCAGAAGACGGCATACGAGATtccatccGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
GGCCGGT	MP_T_67	CAAGCAGAAGACGGCATACGAGATaccggccGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
GGTCAGG	MP_T_68	CAAGCAGAAGACGGCATACGAGATcctgaccGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
GGTTCCG	MP_T_69	CAAGCAGAAGACGGCATACGAGATcggaaccGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
GTAGCAG	MP_T_70	CAAGCAGAAGACGGCATACGAGATctgctacGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
GTCGGCA	MP_T_71	CAAGCAGAAGACGGCATACGAGATtgccgacGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
GTCTATC	MP_T_72	CAAGCAGAAGACGGCATACGAGATgatagacGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
GTTAGAC	MP_T_73	CAAGCAGAAGACGGCATACGAGATgtctaacGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
TACGACC	MP_T_74	CAAGCAGAAGACGGCATACGAGATggtcgtaGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
TAGAATT	MP_T_75	CAAGCAGAAGACGGCATACGAGATaattctaGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
TAGCCGG	MP_T_76	CAAGCAGAAGACGGCATACGAGATccggctaGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
TAGTAGC	MP_T_77	CAAGCAGAAGACGGCATACGAGATgctactaGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
TAGTCCT	MP_T_78	CAAGCAGAAGACGGCATACGAGATaggactaGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
TATATAC	MP_T_79	CAAGCAGAAGACGGCATACGAGATgtatataGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
TATCTCT	MP_T_80	CAAGCAGAAGACGGCATACGAGATagagataGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
TATGCTT	MP_T_81	CAAGCAGAAGACGGCATACGAGATaagcataGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
TCATATT	MP_T_82	CAAGCAGAAGACGGCATACGAGATaatatgaGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
TCCTCGT	MP_T_83	CAAGCAGAAGACGGCATACGAGATacgaggaGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
TCCTTCA	MP_T_84	CAAGCAGAAGACGGCATACGAGATtgaaggaGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
TCGGATG	MP_T_85	CAAGCAGAAGACGGCATACGAGATcatccgaGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
TCTAGCC	MP_T_86	CAAGCAGAAGACGGCATACGAGATggctagaGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
TCTCTGG	MP_T_87	CAAGCAGAAGACGGCATACGAGATccagagaGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
TCTGAGT	MP_T_88	CAAGCAGAAGACGGCATACGAGATactcagaGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
TGAATCT	MP_T_89	CAAGCAGAAGACGGCATACGAGATagattcaGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
TGAGCGC	MP_T_90	CAAGCAGAAGACGGCATACGAGATgcgctcaGTGACTGGAGTTCAGACGTGTGCTCTTCCGA

Index	Name	Primer sequence
TGCCGCC	MP_T_91	CAAGCAGAAGACGGCATACGAGATggcggcaGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
TGCGTTC	MP_T_92	CAAGCAGAAGACGGCATACGAGATgaacgcaGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
TTACTTA	MP_T_93	CAAGCAGAAGACGGCATACGAGATtaagtaaGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
TTATACC	MP_T_94	CAAGCAGAAGACGGCATACGAGATggtataaGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
TTCCGTT	MP_T_95	CAAGCAGAAGACGGCATACGAGATaacggaaGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
TTGATCG	MP_T_96	CAAGCAGAAGACGGCATACGAGATcgatcaaGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
AAGGCTC	MP_T_97	CAAGCAGAAGACGGCATACGAGATgagccttGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
ACGTTGC	MP_T_98	CAAGCAGAAGACGGCATACGAGATgcaacgtGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
ACTCGCA	MP_T_99	CAAGCAGAAGACGGCATACGAGATtgcgagtGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
ATAGTAC	MP_T_100	CAAGCAGAAGACGGCATACGAGATgtactatGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
ATGCAGA	MP_T_101	CAAGCAGAAGACGGCATACGAGATtctgcatGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
CATACCT	MP_T_102	CAAGCAGAAGACGGCATACGAGATaggtatgGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
CCTCGAC	MP_T_103	CAAGCAGAAGACGGCATACGAGATgtcgaggGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
GACCATT	MP_T_104	CAAGCAGAAGACGGCATACGAGATaatggtcGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
GCCATTG	MP_T_105	CAAGCAGAAGACGGCATACGAGATcaatggcGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
GGAACTT	MP_T_106	CAAGCAGAAGACGGCATACGAGATaagttccGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
TGAGACG	MP_T_107	CAAGCAGAAGACGGCATACGAGATcgtctcaGTGACTGGAGTTCAGACGTGTGCTCTTCCGA
TTCAGAG	MP_T_108	CAAGCAGAAGACGGCATACGAGATctctgaaGTGACTGGAGTTCAGACGTGTGCTCTTCCGA

Index is the 7bp sequence read by the Illumina chemistry. Primer sequences contain a unique 7bp sequence flanked by a common sequence corresponding to the Ilumina adapter sequences. The unique primer sequence is the reverse complement of the Index.

Supplement 3-2 (continued)

Dual Indexing Primers

AATCCT	MP_T_Dual_Index_1	AAT GAT ACG GCG ACC ACC GAG A TCT ACAC ACAC TC TTT CCC TAC ACG ACG CTC T	{AATCCT}
CCGCGA	MP_T_Dual_Index_2	AAT GAT ACG GCG ACC ACC GAG A TCT ACAC ACAC TC TTT CCC TAC ACG ACG CTC T	{CCGCGA}
CTTGCC	MP_T_Dual_Index_3	AAT GAT ACG GCG ACC ACC GAG A TCT ACAC ACAC TC TTT CCC TAC ACGACG CTC T	{CTTGCC}
GAAGTC	MP_T_Dual_Index_4	AAT GAT ACG GCG ACC ACC GAG A TCT ACAC ACAC TC TTT CCC TAC ACG ACG CTC T	{GAAGTC}
GGCAAG	MP_T_Dual_Index_5	AAT GAT ACG GCG ACC ACC GAG A TCT ACAC ACAC TC TTT CCC TAC ACG ACG CTC T	{GGCAAG}
TCTTTC	Un-Indexed Adapter	AAT GAT ACG GCG ACC ACC GAG A TCT ACAC TC TTT CCC TAC ACG ACG CTC	{NA}

Un-indexed adapter has no index {NA}, instead the normal adapter arm will be read as the index (underlined).

Chapter 4

Rare Variants In The Complement Pathway in Two Kidney Diseases

Introduction

PE is a disease of pregnancy characterized by hypertension and proteinuria. It has clinical and pathological similarities to TMAs such as aHUS, specifically glomerular endotheliosis and clots in the microvasculature (1, 2). TMAs may be precipitated by pregnancy, likely due to the vascular stress imposed by the fetus on the maternal vasculature, and the kidney. Pregnancy-associated aHUS is more commonly observed after delivery, but some individuals experience renal failure and hypertension as well as hemolytic anemia during pregnancy (4 of 21 individuals that had pregnancy-associated HUS) (3). Of 100 female patients with a diagnosis of aHUS, 21 had their episode in connection to a pregnancy. Of these, 18 carried a mutation in MCP, FH, or FI. Many of these mutations had been noted in aHUS (or since have been identified and evaluated in the context of AP dysregulation).

High sFlt-1 levels antagonize VEGF signaling in PE patients (4). This disproportionately affects fenestrated endothelial cells, which are located in the kidney, liver, and cerebral plexus (5-7). These three sites are also sites affected by PE and its more serious related conditions HELLP (hemolysis, elevated liver enzymes, and low platelets) and eclampsia. Four of eleven consecutive HELLP patients admitted to a nephrology service were reported to have complement regulatory mutations (8). In a second study, three of 33 HELLP patients had a mutation in a complement regulatory gene (9). In preeclamptic patients elevated Bb fragment levels, indicative of AP activation, predict the development of PE (10).

We sequenced *CFH, CFI,* and *MCP* in a cohort of women with lupus and/or anti-phospholipid syndrome (PROMISSE study) and in a non-autoimmune PE cohort and identified individuals carrying a rare mutation. There is evidence for a genetic component to the risk of PE. Men and women born to a mother with PE have a two- to three-fold increase in the risk that they

will be part of a PE pregnancy (11, 12). Also, women with a sister affected by PE are at an increased risk of developing the syndrome (sib-sib OR of ~8) (13).

Many candidate genes have been linked to PE, but meta-analyses have concluded that moderate effects may exist for a handful of genes (14, 15). The first GWAS in PE (~1000 individuals) found a result with a p-value of 10^{-7}, which did not replicate in two independent studies (16). Variants that increase the risk of PE will face negative selective pressure; thus, our focus on rare variants.

To expand our study of the effect of rare variants in the complement pathway to another renal disease, we sequenced of severe lupus nephritis patients, specifically with end-stage renal disease (ESRD). The autoantibodies that occur in lupus deposit in the kidney as part of immune complexes, leading to CP and AP activation (17). If a regulator of either the CP or AP is deficient, this would lead to excessive complement activation.

Because of the known complement activation occurring in PE and lupus nephritis, and the pleiotropic effects of AP mutations on other kidney diseases, *we hypothesized that rare variants in the complement pathway would be present in these diseases*.

Methods

Illumina Library Preperation, Targeted Capture, and Sequencing

Illumina libraires were prepared from genomic DNA as described in Chapter 3. Targeted Capture was done with Nimblegen SeqCap EZ Arrays (385k) for the initial PROMISSE PE (64) and SLE/APLS non-PE controls (34) in groups of 12 to 24. Additional PROMISSE controls, the FINNPEC study, and the VIP study were captured with Nimblgen SeqCap EZ in groups of 12 to 200. The lupus nephritis cases, and controls were captured with Nimblegen SeqCap Arrays (2.1M) or MyGenostics GenCap in groups of 6 to 48. Captures were done following the manufacturer's instructions. Alterations to the Nimblegen hybridization are described in Chapter 3. Pre-capture pooling for all groups except the VIP was done based on equal mass. The VIP pools were constructed based on relative read depths of equivolume pools followed by MiSeq sequencing. Sequencing was done on Illumina HiSeq at the Washington University Genome Technology Access Center using 2x101 bp, 2x135 bp and 2x150 bp reads.

Read Processing, Alignment, and Post-Alignment Processing

Reads from 2x135 bp and 2 150bp runs had any adapter sequences removed using fastq_scrubber.py (Eli Roberson). Reads were aligned with bwa aln (18). Duplicates were marked with Picard. Genome Analysis Toolkit (GATK) Best Practices were followed for indel realignment at the individual level (19-21). Base quality scores were recalibrated at the lane level.

Sample Coverage

We required all samples to be covered ≥ 10x at the targeted exons. The FINNPEC samples were sequenced to a median of 97% of bases with >10x coverage and with 94% of bases >20x coverage (median average coverage of the target of >100x). The VIP samples had a median average coverage of the target of >70x.

Variant Calling

Variant calling was done using the Unified Genotyper of the GATK on all individuals from a study simultaneously. Variants were filtered for quality in the FINNPEC, VIP and ESRD studies with Variant Quality Score Recalibration (4 Gaussians for SNPs and INDELS, 99% Truth Tranche). PROMISSE PE variants were filtered based on hard quality filtering metrics described in the GATK Best Practices for small studies.

Variant Annotation

Annotation was with SnpEff using GRCh37.70; variants that had a SNPEFF_IMPACT of High or Moderate were kept (22). SeattleSeq was also used for annotation.

Variant Analysis

PLINK/SEQ was used to manipulate genotype data. R was used for statistical tests.

Target Intervals

Target intervals were selected based on candidate genes. The complement pathway was central to all captures. The initial preeclampsia capture (used on the first 98 PROMISSE individuals) was approximately 500kb and targeted 118 genes. This evolved as areas like coagulation found additional support. The final capture used for the VIP samples was 1 Mb and

was designed to be used on patients with PE or alternatively individuals with thrombosis or hemostatic disorders. A list of genes is in Supplement 4-1.

For lupus nephritis we sequenced the entire complement pathway because of the role of the classical and alternative pathways in lupus and kidney disease respectively. We also sequenced additional genes identified as potentially involved in kidney disease and inflammation (largely selected by our collaborators) and all known genetic risk factors for lupus (common variants from GWAS as well as Mendelian causes of lupus). This was 171 genes. We also included all ENCODE sites derived from ChIP-Seq experiments, that were within 500 bp of a SNP with an $r^2 \geq 0.8$ with a SNP associated with SLE risk by GWAS (HuGE Navigator, February 2012). The total target was 578 kb.

Expression of FH Mutants and Functional Characterization

FH 1-4 mutations (L3V, R127H, R166Q) were expressed in 293T cells under serum-free conditions, and purified as described (23). Functional studies and biacore were performed as described (23). FH 18-20 mutations, C1077S and N1176K, were expressed in 293T cells under serum free conditions and purified as described with the exception that they eluted from a GE HisTrap column at 285 mM NaCl (FH 1-4 at 95 mM) (23). Heparin binding was assayed using a GE HiTrap Heparin column.

Patient Populations

PROMISSE

Predictors of pRegnancy Outcome: bioMarkers In antiphospholipid antibody Syndrome and Systemic lupus Erythematosus (PROMISSE), a prospective multi-center observational study to identify predictors of pregnancy outcome initiated in 2003. Patients with SLE (defined as ≥4

ACR criteria) and/or APL Ab (defined as at east one of following documented twice between 6 wk and 5 y apart: ACL [IgG or IgM ≥40 units], lupus anticoagulant, or anti-b2GPI [IgG or IgM ≥40 units]) and disease controls were recruited by 12 wk gestation. At screening, patients with renal disease (proteinuria .1,000 mg/24 h, RBC casts, or serum creatinine .1.2 mg/dl), taking prednisone .20 mg/d,diabetes mellitus (Type I and Type II antedating pregnancy),hypertension (blood pressure $140/90 mmHg), or multiple gestations were excluded. Preeclampsia was defined as new onset of elevated systolic blood pressure (≥140 mmHg) and/or elevated diastolic blood pressure (≥90 mmHg) after 20 wk gestation on two occasions at least 4 h apart and proteinuria of 300 mg or greater in a 24 h urine specimen or $1+ on dipstick at least 4 h apart in the absence of pyelonephritis or hematuria. All controls sequenced here were self-reported Caucasian, non-Hispanic. 12 of 64 PE cases were NOT Caucasian non-Hispanic, the other 52 were.

FINNPEC

The Finnish patient samples used in this study originate from the Finnish Genetics of Preeclampsia Consortium (FINNPEC) study cohort and the Southern Finland preeclampsia study cohort. FINNPEC is an ongoing multi-centre study where DNA samples and data have been collected prospectively at all university hospitals in Finland (i.e. Helsinki, Turku, Tampere, Kuopio and Oulu) from 2008. For each woman with preeclampsia, the next available woman giving birth at the same hospital, with no preeclampsia, is invited as a control. After initial review of hospital records by a research nurse, each diagnosis is confirmed by a study physician based on criteria described below.

Finnish women who suffered a preeclamptic pregnancy and had no medical history of chronic hypertension, type 1 diabetes, or renal disease were eligible for the study as cases.

Diagnostic criteria used for the FINNPEC study cohort were SBP≥140 mmHg and/or DBP≥90 mmHg on at least two occasions with new onset proteinuria (≥0.3 g/24 hrs, or ≥0.3 g/L, or in the absence of concurrent quantitative measurement, at least a '2+' or more, or two '1+' proteinuria dipstick readings) after 20 weeks gestation in a previously normotensive woman.

VIP

679 controls (women who did not develop pre-eclampsia) and 283 cases of pre-eclampsia, occurring either during the VIP pregnancy or a previous pregnancy. All patients were self-declared Caucasian. All women recruited to the VIP study were at risk of pre-eclampsia, due to one or more of the following factors: obesity, multiple pregnancy, pre-eclampsia in a previous pregnancy, essential hypertension, diabetes, renal disease or APS. Controls had a risk factor but did not develop preeclampsia.

Criteria for preeclampsia was SBP≥140 mmHg and/or DBP≥90 mmHg on at least two occasions with new onset proteinuria (≥0.3 g/24 hrs, or ≥0.3 g/L, or in the absence of concurrent quantitative measurement, at least a '2+' or more, or two '1+' proteinuria dipstick readings) after 20 weeks gestation in a previously normotensive woman.

Lupus Nephritis

Participants were consenting adults (>18 years old) who self-report their ethnicity as African-American or European-American and undergoing peritoneal dialysis, hemodialysis, or had received a prior kidney transplant. Inclusion criteria for the diagnosis of SLE utilize those of American College of Rheumatology. SLE-ESRD will have been diagnosed in those with ESRD who have had kidney biopsy evidence of prior Stage 3, 4, 5 or 6 lupus nephritis, using World Health Organization criteria; or, in the absence of a renal biopsy, a history of SLE for > 2 years prior to developing ESRD and with historical evidence of proteinuria exceeding 500 mg/24

hours (or equivalent: > 100 mg/dl proteinuria on dipstick or spot urine protein:creatinine ratio > 0.5 gram/gram) and/or presence of cellular casts (red cell, hemoglobin, tubular cell or mixed) prior to the initiation of dialysis or kidney transplantation. Participants with SLE-ESRD that lack renal biopsy information will have confirmed to the study coordinator that they had onset of SLE at least two years prior to developing ESRD; they will meet entry criteria based on urinary findings and they will confirm that they lack risk factors for other forms of kidney disease (e.g., diabetes mellitus, urologic disease, first degree relatives with autosomal dominant polycystic kidney disease or Alport's syndrome, surgical nephrectomy). Individuals who are unwilling or unable to provide written informed consent are excluded

Results

PROMISSE PE

We reported that 7 of 40 cases of preeclampsia from the PROMISSE study had a mutation in MCP, FI or FH ((24), Supplemental Manuscript 3). Importantly, we showed that the MCP mutation, K66N, had a reduced binding to and cofactor activity for C4b. Its regulatory activity for C3b was not affected. This PE case also had SLE and likely had complement-fixing immune complexes depositing in the kidney and possibly in the placenta. In a replication cohort of severe PE from Utah (without autoimmunity), we identified a variant in CFI that had previously been observed in aHUS and characterized as having a secretion defect (24, 25). Both cohorts also had enrichment for the MCP A304V allele that has been observed to have an impact on the ability of MCP to protect the cell from complement deposition in situ (26).

Targeted Sequencing of the PROMISSE Cohort

To extend these studies, we sequenced all of the genes in the complement system in 64 PE cases from the PROMISSE study, as well as all those previously reported to be associated with PE either by candidate gene studies or linkage. We also included genes in the clotting pathway. A number of these have been associated with PE in candidate gene studies, and coagulation is abnormal in PE, as increased clots have been reported in the placenta and kidney vasculature.

We sequenced 64 PE cases from the PROMISSE study (mostly Caucasian, non-Hispanic) and 34 controls with SLE but without PE (all Caucasian non-Hispanic). All SNPs within 200 bp of the targeted exons were included. Only variants that were genotyped in all individuals were included to minimize any possibility that a genotyping call rate was responsible for a difference between cases and controls. We excluded any variant that did not pass hard quality filtering. A gene-by-gene test of rare variant burden (MAF<5%) yielded 3 genes that met the Bonferroni corrected statistical cutoff of 0.0004 (118 genes tested) using the gene-based variable threshold test (Table 4-5). These genes were VEGFA, FCN1 and CR1 (Table 4-1).

Table 4-1. Genes Meeting Bonferroni Corrected Significance Threshold (VT test)			
Gene	Number of Unique Variants	p-value	# of Alleles Cases/Cntls
VEGFA	31	7.5x10-5	196/74
FCN1	37	1.9x10-4	495/171
CR1	60	4.3x10-4	526/236

64 PE cases and 34 controls from the PROMISSE study.

VEGFA: Multiple studies have identified low VEGFA signaling in PE due to elevated levels of the sFlt-1. Uncommon and rare variants (MAF <5% in this group) drive this signal, with 45 alleles present in cases and 4 in controls. Only one is missense (Ser77Asn) and it was seen in only one of ~6500 individuals reported in the ESP GVS. The others are all intronic which complicates the interpretation.

FCN1: Ficolin-1 (FCN1) initiates the lectin pathway of complement by recognizing sugars present on pathogens and altered self. There are 23 variants (4 missense) at <5%, of which 43 alleles are present in cases and three in controls. The 4 missense variants are only observed in cases.

CR1: Complement Receptor 1 (CR1) is a highly polymorphic gene that was originally cloned in the Atkinson laboratory. It can regulate both the CP and the AP (27). The locus is complicated because there is structural variation in this gene, with a 9 exon unit comprising a long homologous repeat (LHR) domain (28). Most individuals carry four LHRs, but some carry three and others five or six (Figure 4-1). Homology between the repeated LHR units is extremely high and no read aligned to this repeated unit has an unambiguous match (Map Quality (MAPQ)=0). The association is being driven by variants found in unique regions, specifically LHR D. It is unlikely that read mapping issues are responsible for a false positive.

Variation in CR1 LHRs Due to Insertion of B-like Duplication

Of B/B-Like LHRs:
CR1*1 -1 copies-~90%
CR1*2 -2 copies-~10%
CR1*3- 0 copies-~1%
*Genomic Reference Uses CR1*2

Non-Unique Portion of CR1 (MAPQ=0)
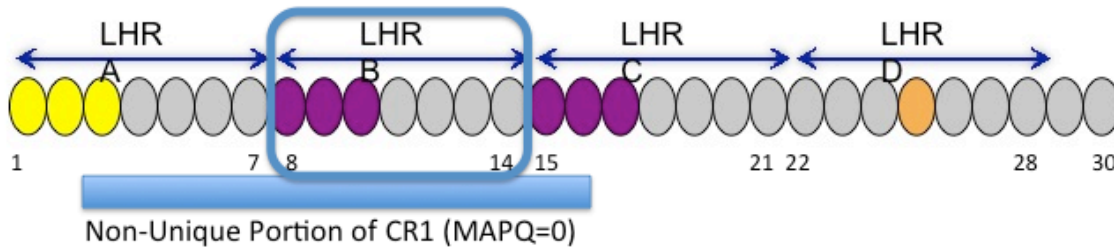
Figure 4-1. CR1 Is Composed of Homologous Repeats
Each repeat of CR1 (LHR) is composed of 7 CCP domains (9 exons). Decay activity resides in LHR A (yellow). Cofactor activity is present in LHR B and LHR C (purple). C1q binds in LHR D. The area of the gene duplicated in ~10% of the population is boxed in blue. The blue bar delineates regions were 2x100 bp reads are mapped ambiguously. LHR B is a hybrid LHR A and LHR C caused by a gene conversion event.

Population Stratification is a Confounding Factor in the PROMISSE Cohort

After sequencing additional Caucasian, non-Hispanic SLE only controls and healthy controls (no PE in either control group), it was concluded that the strong association in *VEGFA, FCN1*, and *CR1* was due to population stratification. Non-Caucasian haplotypes present in these three genes could predict >80% of the non-Caucasian individuals. Because intronic variants were included in the analysis and each haplotype might carry one to three variants in the captured regions, only 10 or 20 non-Caucasian cases (15-30% of the total) were able to contribute substantially to the excess of variants seen in cases. Because only a fraction of the cases were

non-Caucasian (and none of the controls) these variants appeared disease specific. This highlights how stratification can lead to very significant false positives.

Targeted Sequencing in a Finnish PE Population

We sequenced a more homogeneous Finnish population from the FINNPEC study. This included 500 cases, two-thirds of which had severe PE, and 190 controls (Table 4-2). Even with the expected homogeneity within Finland, these samples were matched between cases and controls across 8 geographic codes (Table 4-3). The Finnish population experienced a bottleneck as a result of migration and thus disease causing variants may be population specific and at a higher frequency than in other, related, European populations (29).

Table 4-2.FINNPEC Sample Phenotypes

Phenotype	Number of Individuals
Severe PE	362
Non-Severe PE	138
Non-PE Controls	190

Table 4-3: Geographic Distribution of Finnish Samples

Region	Cases (%)	Controls (%)
Southern Finland	22.0	27.4
Southern Finland	15.2	18.4
Southern Finland	7.2	6.8
Eastern Finland	8.4	12.6
Northern Finland	13.6	15.3
Central Finland	8.6	11.6
Western Finland	9.4	7.9
Other	15.6	0

Gene Burden Test in Finnish Samples

Association testing was done in plinkseq. We conducted the test with a MAF cutoff of

5% and 1% as the population was relatively small and predominantly cases. Some results were

unstable as the MAF changed, but generally the top ten results were consistent (Table 4-4).

Though a number of interacting genes were towards the top of both lists (C1R, C1S, and CD93;

INHA, TGFB3 and ACVR1), we initially have focused on *CFH* as it was consistent with our

initial hypothesis. C5 was also consistent with our hypothesis but the connection between renal

disease and FH was felt to be the strongest.

Table 4-4. FINNPEC Gene-Based Burden Test at MAF ≤ 1% and ≤ 5%					
MAF ≤ 1%			MAF ≤ 5%		
Rank	Gene	p-value	Rank	Gene	p-value
1	CD93	0.017	1	C1R	0.004
2	ADAM28	0.027	2	C4B	0.011
3	**CFH**	0.035	3	HEY2	0.018
4	C5	0.045	4	ADAM28	0.019
5	C1R	0.046	5	**CFH**	0.034
6	HEY2	0.049	6	CD93	0.041
7	INHA	0.051	7	FCN2	0.045
8	F8	0.096	8	C5	0.062
9	C1S	0.10	9	F8	0.070
10	TGFB3	0.10	10	ACVR1	0.078

499 cases and 190 controls. Burden testing conducted on 160 genes. Bonferroni corrected p-value is 0.0003.

AP Gene Mutations in FINNPEC

Factor H: We found rare factor H mutations in 7 of the 499 FINNPEC cases and zero of the 190

controls (Table 4-5). 6 of 7 individuals had severe PE and/or HELLP, but we would expect 4 or 5

individuals (of 7) to have severe PE given the fraction of cases with severe PE (Table 4-6). We

excluded variants that are known to be polymorphisms due to their frequency (≥ 1% MAF in the

NHLBI ESP), that the aHUS literature considers a polymorphism, or that was balanced in a large

case:control AMD sequencing cohort (30, Supplemental Manuscript 1). These rare mutations

cluster in CCP1-4 and CCP18-20, known sites of FH's regulatory activity localization specificity

(31-33).

Table 4-5. CFH Variants (MAF<1% in the NHLBI ESP) in FINNPEC				
AA	Cases	Controls (1kG)	Effect	NHLBI ESP Observations (Disease Notes)
L3V	3	0	Unknown	2 No association
R127H	1	0	Low secretion	Novel (aHUS, DDD, AMD-Case)
R166Q	1	0	Unknown	Novel (AMD-Case)
C1077S	1	0	Loss of conserved cysteine	Novel
N1176K	1	0	CCP20	Novel

499 cases. 190 controls. 1kG: Number of observations in the 93 Finnish individuals from the 1,000 Genomes Project. AMD: The variant was seen in AMD cases in the case:control sequencing experiment discussed. CCP20 is a hotspot for aHUS associated mutations. NHLBI ESP is the number of alleles observed in the 6500 individual release.

Table 4-6. Clinical History of Rare CFH Carriers	
Mutation	
L3V	HELLP @ 28 wks.
L3V	Severe PE @ 32 wks. History of PE
L3V	Severe PE @ 31 wks. History of PE
R127H	Severe PE @ 34 wks
R166Q	Severe PE @ 30 wks
C1077S	HELLP @ 33 wks
N1176K	PE @ 37 wks

$2/3^{rd}$ of all cases were severe PE. HELLP is more serious than severe PE.

<u>L3V</u>: Was secreted normally. Preliminary investigations into L3V's cofactor activity showed no effect on function (data not shown). It is in the signal peptide, and as the protein appeared to be the correct size, the only defect likely is if carriers are haploinsufficient.

<u>R127H:</u> Had previously been described to have a secretion defect and accumulate intracellularly (34) and we confirmed this (Figure 4-2).

<u>R166Q:</u> It had a profound defect in cofactor activity and a ten-fold lower affinity for C3b in SPR (Figure 4-3). R166Q had a unique cleavage pattern of C3b compared to wild type 1-4 (and other mutants in CCP 1-4 (23)). R166Q was delayed in cleaving the α_{43} fragment to the α_{40} fragment. While both fragments may be seen, the α_{40} fragment is the predominant fragment of WT, especially when using full length FH and as time elapses.
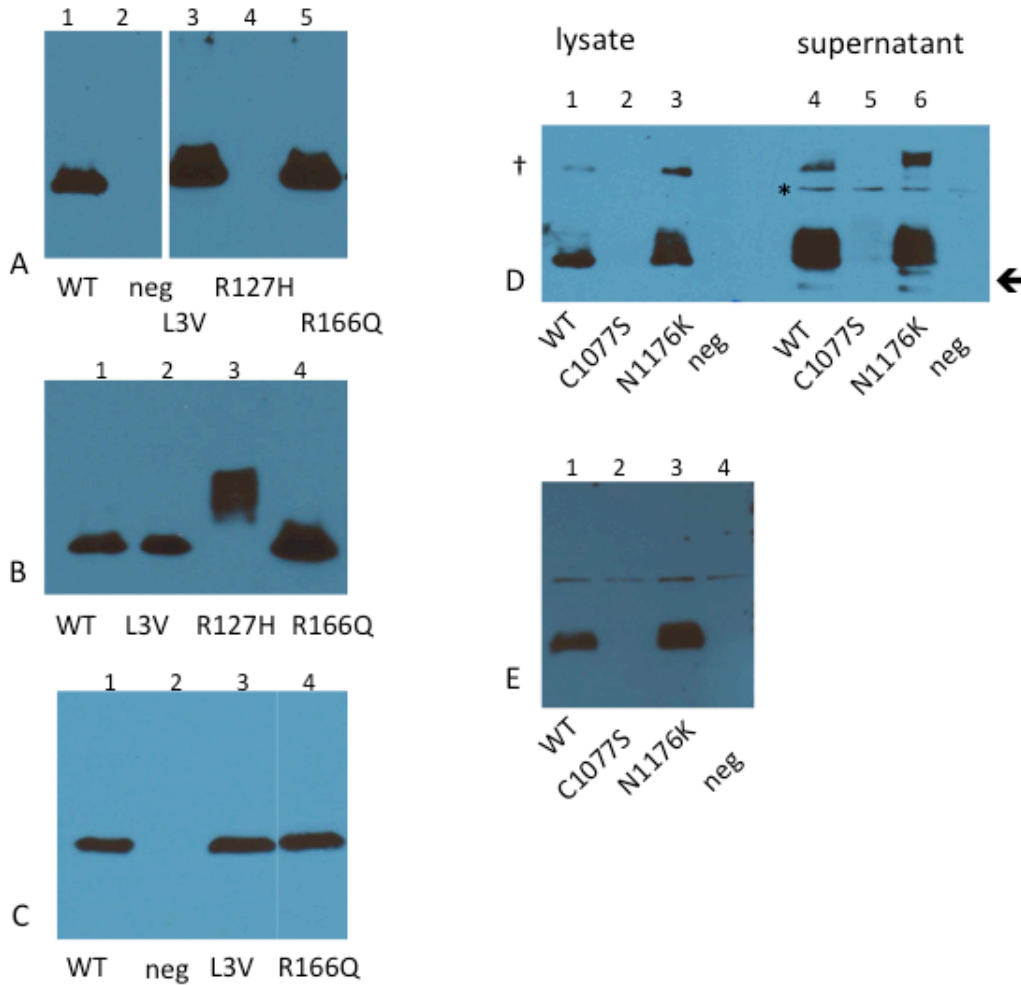
Figure 4-2. R127H and C1077S Are Not Secreted

A. Non-reducing WB of CHO supernatants. R127H is not secreted (lane 4). B. Non-reducing WB of CHO cell lysates. R127H is not processed correctly (lane 3). C. Non-reducing WB of equivalent masses of each protein. D. Non-reducing WB of FH 18-20 fragments. C1077S does not accumulate intracellularly and is not secreted (lanes 2 and 5). The N-glycosylation site in CCP 18 is likely used as a small fraction of protein always ran at a lower molecular weight (arrow). † It is possible some of the protein is in a dimeric form. *There is a non-specific band visible present in the negatve control transfection supernatant. E. Reducing WB of 18-20 fragments removes the dimer. neg is negative control transfection.

Figure 4-3. R166Q Has a C3b Binding and Cofactor Defects
A. Fluid phase cofactor assay of WT and R166Q at 20 and 30 min. R166Q cleaves the α′ chain less quickly. $α_1$ and $α_{40}$ accumulate more slowly. Additionally, R166Q consistently has a problem cleaving $α_{43}$ to $α_{40}$. B. Quantification of the fluid phase cofactor activity based on the amount of α′ chain remaining compared to the amount at time 0. There is a significant difference at all three time points. *** $p < 5 \times 10^{-4}$. **$p < 0.005$. *$p < 0.05$. C. Surface plasmon resonance traces with FH 1-4 fragments coupled to the chip. R166Q has a 10-fold lower affinity for C3b (64 vs 6.8 μM). WT had 300 RU and R166Q 600 RU coupled to the chip. K_D calculated using the BiaEval 1:1 binding model.

Y402H: The common allele Y402H has been reported to be enriched in cases of DDD (35). We

do not find it enriched in the PE cases (42%) compared to the controls (44%).

C1077S: The loss of the second conserved cysteine in CCP18 led to a complete secretion defect with no protein present intracellularly (Figure 4-2 D).

N1176K: It is produced similarly to wild type FH18-20 (Figure 4-2 D and E). It is on the face of CCP20 that faces the C3d fragment in the crystal structure. This is opposite face of CCP20 compared to R1210C, which affects heparin binding (Figure 4-4). Thus, it is expected to affect C3d/C3b binding, not heparin binding. Consistent with this interpretation, our experiments with heparin binding have shown it binds heparin comparable to wild type (not shown). Experiments to measure C3d binding have been limited by the quantity of C3d required to detect binding at 150 mM NaCl. These experiments are ongoing.
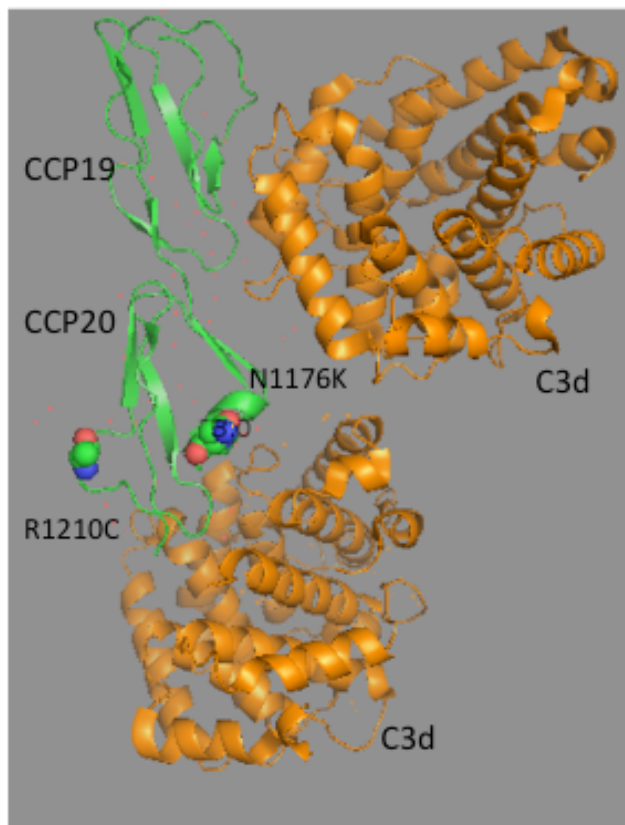

Figure 4-4. N1176K is Adjacent to C3d
Factor H 19-20 is in green. C3d fragments are in orange. N1176K is on the opposite face relative to R1210C. This places N1176K adjacent to the C3d fragment in the crystal structure (PDB ID 2XQW)

<u>Factor I</u>: *CFI* does not have a similar burden of variants associated with PE (6 alleles in cases and 2 in controls). If the T107A variant is removed (2 cases and 1 control), comparing singletons in each group, there are 4 variants in case and 1 in controls (Table 4-7). The G261D variant is considered a polymorphism in the aHUS literature (after substantial debate), and it was nearly perfectly balanced in the AMD case:control cohort described (30). Q462H is in a PE case; it was observed in a non-AMD control, who had a serum level near the bottom of the normal range (30).

Table 4-7. CFI Variants (MAF<1% in the NHLBI ESP) in FINNPEC				
AA	Cases	Controls (1kG)	Effect	NHLBI ESP Observations (Disease Notes)
T107A	2	1	Unknown	2
G119R	1	0	Secretion Defect	11 (aHUS and AMD-Case)
A258T	0	1	Splice Site	4 (AMD-Case)
G261D	1	0	No functional or quantitative defect.	22 (aHUS, AMD-Both)
G314V	1	0	Splice Site	Novel
Q462H	1	0	Pred. Secretion defect	2 (AMD-Control)

499 cases. 190 controls. 1kG: Number of observations in the 93 Finnish individuals from the 1,000 Genomes Project. AMD: The variant was seen in AMD cases or controls from the case:control sequencing experiment discussed. NHLBI ESP is the number of alleles observed in the 6500 individual release.

Factor B: *CFB* has an uncommon polymorphism (1.8% in the ESP) that is enriched in cases 2.2-fold, G252S. There is also a case that carries a variant previously reported in aHUS, I242L (36). These variants are of unknown significance; they sit in the loop containing the scissile bond involved in activation of factor B to Bb (37). Mutations here do affect Bb activation by factor D and other proteases (Hourcade, unpublished data).

MCP: The K66N variant, observed in a PROMISSE PE case, was present in 2 controls and 2 cases. The A304V variant, previously observed to be enriched in PROMISSE PE cases (25), had a minor allele frequency of 4.0% in cases and 4.4% in controls in the FINNPEC population. Two cases carried a rare variant, T353I, in one of MCP's alternatively spliced tail isoforms.

C3: C3 did not have any variants that are likely to affect AP activity.

Disease Related Clotting Mutations are Enriched in Preeclamptic Cases

PE likely represents a state of systemic inflammation and has been considered (by some) to be a thrombotic state. There is an enrichment of rare, OMIM-annotated variants in the clotting pathway in PE cases vs controls. We are using the OMIM database to derive meta-information about these variants, much as we used the aHUS literature and structure-function analyses in our interpretation of AP variants. We observed an enrichment of known disease associated variants in FV (*F5*), protein S (*PROS1*) and Anti-thrombin III (*SERPINC1*), all of which would contribute to an excess of fibrin rich clots being formed (Table 4-8). We also observe an enrichment of known disease associated variants in von Willebran Factor (*VWF*) and FX (*F10*) (Table 4-9).

Table 4-8. OMIM Annotated Variants in the Prothrombotic Pathway Are Enriched in Cases						
Gene	Variant	Cases	Controls (1kG)	P-value	ESP	Phenotype
FV	Q534R	29	5 (2)	0.118	278	Protein C cannot inactivate.
Protein S	N365K	16	2 (3)	0.181	2	Cofactor to Protein C
Anti-thrombin III	P73L	7	1 (0)	0.457	5	Interferes with heparin binding.

499 cases, 190 controls. Numbers in parenthesis are alleles present in 93 Finnish individuals from 1000 Genomes. ESP: Number of alleles observed in the NHLBI ESP 6500 individual release

Table 4-9. OMIM Annotated Variants Leading to Bleeding Are Enriched in Cases						
Gene	Variant	Cases	Controls (1kG)	P-value	ESP	Phentoype
VWF	P1266L	11	0 (0)	0.041	1	VWD type IIb.
F10	E142K	19	2 (0)	0.082	27	Mild functional defect.

499 Cases. 190 Controls. Numbers in parenthesis are alleles present in 93 Finnish individuals from 1000 Genomes. ESP: Number of alleles observed in the NHLBI ESP 6500 individuals. VWD: von Willebrand Disease. P1266L VWF has an abnormal response to ristocetin induced aggregation, but normal levels in the blood.

FV Leiden: While results have varied between studies, two recent meta-analyses (14, 15) of the

literature concluded that FV Leiden has a real association (OR=1.7). We replicate this finding in

the FINNPEC population (OR=2.2, 2.9% vs 1.3% MAF in cases vs controls) (Table 4-8). FV

Leiden predisposes to clotting because it is resistant to inactivation by Protein C, which requires

Protein S as a cofactor. This increases the amount of active Va, that in turn enhances the

formation of the active Xa:Va enzyme leading to clot formation (Figure 4-5).
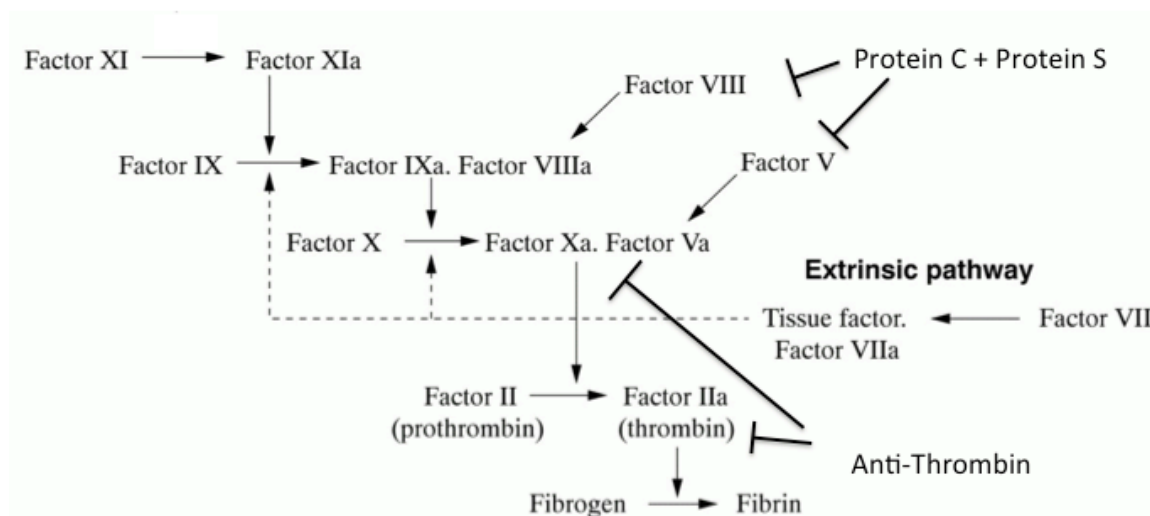
Figure 4-5. The Coagulation Cascade
Important points of negative regulation are highlighted. FINNPEC cases are enriched for
mutations in the negative regulators anti-thrombin and protein S. Adapted from Bowen D J Mol
Path 2002;55:127-144.

Protein S: We find an excess of alleles in cases with the OMIM annotation for Protein S

deficiency in cases as well, specifically N365K (Table 4-8). No missense changes were observed

in the protease Protein C, Protein S's binding partner.

Anti-thrombin III (AT III): We identified an excess of alleles in the AT III gene in cases, which

controls both thrombin (F2) and Xa. The "Clichy" mutation (P73L) is in 6 cases and 1 control

(Table 4-8). It interferes with heparin binding, leading to a secondary deficit in F2 and F10 regulation. One individual is homozygous and had early onset of preeclampsia (31 weeks).

von Willebrand Factor (VWF): This factor is at the center of a TMA, thrombotic thrombocytopenia purpura (TTP), as well as the bleeding disorder von Willebrand disease (VWD). It is a chaperone for FVIII in the blood and also mediates platelet adhesion to surfaces and clots. VWF responds in dynamic ways to increased shear stress, which would be present in the spiral artery that has not been appropriately remodeled in PE. We find a number of OMIM annotated variants in *VWF*. One in particular (P1266L) is seen in 11 cases and in no controls, none of the 1000 Genomes subjects of any ethnicity, and only a single time in the ESP (Table 4-9). P1266L affects platelet aggregation in response to ristocetin (38-40), while it has normal size multimers and levels in the blood. Because of this, it is unclear how it relates to a pro-bleeding phenotype.

There is a small allele bias issue with P1266L, with the non-reference base averaging 38% of reads at this site versus the expected 50% (vs. 0.07% in non-carriers). However, we are confident that our 2x101bp alignments are correct because there are 18 nucleotide differences in the 600 bp surrounding this variant between *VWF* and its pseudogene on chromosome 22. This variant is present on at least four haplotypes based on our sequencing, three of which have been reported, and is thought to have arisen through gene conversion.

There are two other OMIM annotated variants in *VWF*. They occur in cases and controls. A fourth variant is OMIM annotated, but is an uncommon polymorphism (~2% MAF) and is not related to disease.

Factor X: *F10* has an OMIM annotated variant that is over represented in cases (Table 4-9). This variant has a mild, but significant, decrease in function and is often observed in combination

with more severe variants (41). A significant decrease in F10 levels (<15% normal) must be

present in order for patients to present clinically with bleeding secondary to a F10 deficiency.

We observe a novel stop mutation in F10 in a separate case as well.

Interpretation of Coagulation Variants in Light of OMIM

Given our sample size, each of these clotting variants alone (and the many rare variants in

these genes) do not account for a large enough risk of disease at the population level to reach

significance in isolation. Considered as a risk class, OMIM recognized coagulopathy variants, in

aggregate are associated with PE (Table 4-10). Inclusion of the un-phenotyped 93 Finnish 1,000

Genomes participants does not change how we interpret these data.

This metric may underestimate the risk of disease attributable to coagulation pathway

mutations as there are many rare variants observed exclusively in cases that may be associated

with disease (though some may be a neutral, rare variation).

Table 4-10. Coagulation Pathway OMIM-Annotated Variants Are Increased in PE Cases			
	# of PE Cases	# Controls	Controls w/ 93 Finnish Individuals (1000 Genomes)
Individuals Carrying A Coagulation Pathway Variant	85	13	6
Individuals Not Carrying A Coagulation Pathway	414	177	87
FINNPEC SAMPLES: p=1.6x10^{-4} (one-tailed hypergeometric) p=3.8x10^{-4} (two-tailed exact)			
FINNPEC SAMPLES w/ 1000 Genomes: p=1.1x10^{-5} (one-tailed hypergeometric) p=2.7x10^{-5} (two-tailed exact)			

499 cases. 190 controls. 93 Finnish individuals from the 1,000 Genomes project.

Replication in the Vitamins in Preeclampsia (VIP) Study

We are replicating these findings in an unrelated cohort of 350 PE cases and 650 controls from the VIP study. To leverage the increased output of the HiSeq 2000 as longer read lengths have become available, we utilized 2x150 and 2x135 bp reads initially. This was problematic because some of our reads contained adapter sequence, as the insert sequence was smaller than 150 bp or 135 bp respectively (though far fewer had adapter in the 135 bp read length sequencing). This led to mapping problems for some reads. The bwa mem algorithm was very aggressive in aligning reads containing long Illumina adapter sequences to the genome. This resulted in spurious variant calls being made based on the mismatches. Inspection at the read level quickly revealed that something was wrong with the alignment.

The read were reanalyzed starting with removing the adapters with the use of fastq_scrubber.py (Eli Roberson). This tool was able to preserve read order to take advantage of read pairing when doing paired end alignment. Many other tools fail to do this correctly. Reads were then aligned with bwa aln, which is what was used with the FINNPEC samples. The mem algorithm was evaluated as overly aggressive when the adapters were present.

We are currently processing this data and evaluating it. However, we have some preliminary data. The variant calls were 99.8% concordant with array based genotyping results using 277 SNPs present in both platforms and 91 cases shared between the groups.

aHUS-associated FH Variants Are Present in the VIP PE Group

There is one aHUS associated variant in the cases and none in the controls. Q400K is in the anionic binding site in CCP7 neighboring Y402H. The aHUS patient that carried Q400K had a fatal episode at 15 days after birth (34). We have already defined the CA defect in R166Q.

There are two other variants in the case group in CCP6, G366Q and H373Y (Table 4-11). The controls have only one variant, in CCP8, Y481F. There were no variants in CCP19-20.

Y402H was distributed evenly between cases and controls (38% in cases and 40.6% in controls. MCP A304V was distributed evenly between cases and controls (10 vs. 25 alleles, 2.4x as many controls as cases).

Table 4-11. CFH Variants in Functional Domains in VIP				
AA	Cases	Controls	CCP	NHLBI ESP Observations (Disease Notes)
R166Q	1	0	3	Novel (AMD)
G366E	1	0	6	Novel
H373Y	1	0	6	Novel
Q400K	1	0	7	1 (aHUS and AMD)
Y481P	0	1	8	Novel

282 Cases. 678 controls. AMD: The variant was seen in AMD cases in the case:control sequencing experiment discussed. NHLBI ESP is the number of alleles observed in the 6500 individual release.

Thrombotic Risk Loci are Not Enriched in the VIP Cohort

FV Leiden was actually twice as prevalent in the control group (0.9% MAF in cases vs. 1.7% in controls (Table 4-16). The only homozygous individual was a control as well.

The other risk loci that were notably increased in the FINNPEC study were either not present or showed no trend towards enrichment in cases (Table 4-12. Only the E142K variant in F10 was enriched in cases. We did not observe the P1266L variant in any VIP individual.

Table 4-12. OMIM Annotated Clotting Variants Enriched in FINNPEC Presence in VIP					
Gene	AA	VIP Cases	VIP Controls	FINNPEC Observations (Case/Control)	NHLBI ESP
F5	Q534R	5	27	29/5	278
PROS1	N365K	0	0	16/2	2
AT-III	P73L	0	2	7/1	5
VWF	P1266L	0	0	11/0	1
F10	E142K	3	0	19/2	27

282 Cases. 678 controls. AMD: The variant was seen in AMD cases in the case:control sequencing experiment discussed. NHLBI ESP is the number of alleles observed in the 6500 individual release.

FINNPEC Extension with Additional Cases and Controls

I am sequencing C3, FI, FB, FH, and MCP in an additional 190 PE cases using the Fluidigm Access Array. This assay was developed by me for use by Alexion Pharmaceuticals as the basis for a CLIA approved aHUS genetic test. Also, we are genotyping 24 variants in 190 additional (non-severe PE) cases and 475 additional Finnish controls with a focus on the clotting pathway variants enriched in cases. Given the discrepancy between the size of the Finnish case group and control group (2.6 fold), we would like to get a better estimate of these allele frequencies in the Finnish population as well as validate these variants in the sequenced samples.

Lupus Nephritis

We sequenced 200 cases of end stage renal disease in the setting of lupus (ESRD-SLE), 192 lupus cases without renal disease (NN-SLE), and 190 healthy controls (Healthy). All were self-declared Caucasian, non-Hispanic. The ESRD-SLE group was from a number of centers around the country, while the NN-SLE and Healthy controls were largely from the University of Alabama-Birmingham. Our power to detect an association given the number of genes we were testing was estimated to be >70% if any gene was responsible for 7% of the population attributable risk of severe lupus nephritis (a 15-fold increase of variants over an anticipated aggregate rate of rare mutations in controls of <1 per 100 individuals (0.05% aggregate MAF).

We identified a single variant that we are confident will lead to haploinsufficiency in MCP (Y189D) (42). There were no other variants in AP genes where we were confident in their effect on function, or where there was any enrichment in cases versus controls for renal disease (ESRD-SLE vs NN-SLE). This data was transferred to our collaborators at Univ. Alabama Birmingham (Bob Kimberly) for further analysis of the effect of rare variants in genes known to increase the risk of lupus.

Discussion

AP Dysregulation in PE

Based on our results in sequencing genes of the complement system we conclude that dysregulation of the AP for genetic reasons is not a significant contributor to the prevalence of PE. There is a consistent trend in both the FINNPEC and VIP populations for enrichment of variants in the functional domains of FH in PE cases. However, this is on the order of 1% of cases. The R127H and C1077S variants have secretion defects as expected. While specifically

114

the R166Q mutation seen in both groups has a significant decrease in CA, and the Q400K is in close proximity to the AMD risk allele Y402H. It is also possible Q400K has a secretion defect (34). The other *CFH* variants are largely of unknown functional significance. The role of CCP6 (site of 2 rare variants in the VIP study) is debated, however there is a burden of variants in this domain in AMD cases verus controls (Chapter 5).

A counter argument for the role of AP regulatory mutations in PE, is that rare *CFI* variants known to cause haploinsufficiency or loss of function are present in cases and controls of the FINNPEC and VIP studies. Additionally, previous reports that A304V is a risk variant may be incorrect as it is not enriched in the FINNPEC or VIP studies. Its role in aHUS is also questionable due to its high frequency in the population.

We rule out rare variants in the AP or its regulators from contributing significantly to the risk of PE. Mouse models of preeclampsia are far from perfect, but pharmacologic blockade of the AP has been reported to reverse the blood pressure and renal dysfunction observed in these mice during pregnancy. AP activation may exacerbate the effects of low VEGFA levels on glomerular endothelial cells or other specialized vascular beds. AP blockade may be an effective therapy in a subset of PE patients.

Coagulation Pathway Variants in PE

The enrichment we see in the FINNPEC cohort for variants known to affect the clotting pathway is interesting considering the unique hemostatic environment of the placenta. Thus it is not surprising that a diversity of variants that alter hemostasis may increase the risk of preeclampsia. Meta-analyses have confirmed that both the F5 Leiden mutation and the prothrombin 3 ′UTR mutation (both prothrombotic) increase the risk of PE (14, 15). We replicate

115

the F5 Leiden finding in the FINNPEC group (the prothrombin mutation was not covered well in sequencing).

This is the first report of rare variants in multiple genes of the clotting pathway increasing the risk of PE. The same report that noted pregnancy associated aHUS predominantly occurs after delivery did note that episodes of ADAMTS13 deficient TTP (a TMA that features CNS involvement) predominantly occur during pregnancy (15 of 21 cases between 20 weeks GA and 37 weeks) (3). This is consistent with out data in the FINNPEC study that dysregulation of hemostasis may be exacerbated by pregnancy and increase the risk of PE.

While the VIP study did not replicate these findings, we are examining if study inclusion requirements and or study design can explain why this finding is not replicated and why the F5 Leiden allele is observed at ~50% the frequency in VIP cases versus controls. We are also genotyping all of the alleles discussed here in clotting genes in additional FINNPEC cases (190) and controls (475).

Variants in the AP are Not Increased in Lupus Nephritis Cases

We can conclude that genetic deficiencies of AP regulation are not a common factor that exacerbates the renal disease observed in individuals with lupus. We examined the most extreme phenotype and found only one individual with a variant where we are confident they will be haploinsufficient for AP regulatory activity. No such mutations in *CFH, CFI,* or *C3* were identified. This is counter to the idea that rare mutations in MCP and FH may lead to an early age of onset of lupus nephritis (43).

This may be due to a strict versus generous definition of which variants are loss of function in terms of if they are likely/proven to affect function and their MAF. Early screening of aHUS controls was limited to a few hundred people. Large collections of sequencing data are available, importantly the NHLBI ESP dataset, yielding much better estimates of the MAF of rare variants (44). Even more so than in PE, the case for complement dysregulation in the kidney leading to renal disease in a subset of patients is strong. The lack of genetic variants that increase risk does not preclude the idea that these individuals would benefit from complement therapeutics such as the anti-C5 antibody now used in aHUS (45).

References

1   Stillman, I. E., and S. A. Karumanchi. 2007. The glomerular injury of preeclampsia. *J. Am. Soc. Nephrol.* 18:2281-2284.

2   Taylor, C. M., C. Chua, J. J. Howie, and R. Risdon. 2004. Clinico-pathological findings in diarrhoea-negative haemolytic uraemic syndrome. *Pediatr. Nephrol.* 19:419-425.

3   Fakhouri, F., L. Roumenina, F. Provot, M. Sallee, S. Caillard, L. Couzi, M. Essig, D. Ribes, M. A. Dragon-Durey, F. Bridoux, E. Rondeau, and V. Fremeaux-Bacchi. 2010. Pregnancy-associated hemolytic uremic syndrome revisited in the era of complement gene mutations. *J. Am. Soc. Nephrol.* 21:859-867.

4   Levine, R. J., S. E. Maynard, C. Qian, K. H. Lim, L. J. England, K. F. Yu, E. F. Schisterman, R. Thadhani, B. P. Sachs, F. H. Epstein, B. M. Sibai, V. P. Sukhatme, and S. A. Karumanchi. 2004. Circulating angiogenic factors and the risk of preeclampsia. *N. Engl. J. Med.* 350:672-683.

5   Eremina, V., M. Sood, J. Haigh, A. Nagy, G. Lajoie, N. Ferrara, H. P. Gerber, Y. Kikkawa, J. H. Miner, and S. E. Quaggin. 2003. Glomerular-specific alterations of VEGF-A expression lead to distinct congenital and acquired renal diseases. *J. Clin. Invest.* 111:707-716.

6   Sison, K., V. Eremina, H. Baelde, W. Min, M. Hirashima, I. G. Fantus, and S. E. Quaggin. 2010. Glomerular structure and function require paracrine, not autocrine, VEGF-VEGFR-2 signaling. *J. Am. Soc. Nephrol.* 21:1691-1701.

7   Rana, S., C. E. Powe, S. Salahuddin, S. Verlohren, F. H. Perschel, R. J. Levine, K. H. Lim, J. B. Wenger, R. Thadhani, and S. A. Karumanchi. 2012. Angiogenic Factors and the Risk of Adverse Outcomes in Women with Suspected Preeclampsia. *Circulation*.

8   Fakhouri, F., M. Jablonski, J. Lepercq, J. Blouin, A. Benachi, M. Hourmant, Y. Pirson, A. Durrbach, J. P. Grunfeld, B. Knebelmann, and V. Fremeaux-Bacchi. 2008. Factor H, membrane cofactor protein, and factor I mutations in patients with hemolysis, elevated liver enzymes, and low platelet count syndrome. *Blood* 112:4542-4545.

9   Crovetto, F., N. Borsa, B. Acaia, C. Nishimura, K. Frees, R. J. Smith, F. Peyvandi, R. Palla, M. Cugno, S. Tedeschi, P. Castorina, E. Somigliana, G. Ardissino, and L. Fedele. 2012. The genetics of the alternative pathway of complement in the pathogenesis of HELLP syndrome. *J Matern Fetal Neonatal Med* 25:2322-2325.

10   Lynch, A. M., R. S. Gibbs, J. R. Murphy, T. Byers, M. C. Neville, P. C. Giclas, J. E. Salmon, T. M. Van Hecke, and V. M. Holers. 2008. Complement activation fragment Bb in early pregnancy and spontaneous preterm birth. *Am. J. Obstet. Gynecol.* 199:354 e351-358.

11   Esplin, M. S., M. B. Fausett, A. Fraser, R. Kerber, G. Mineau, J. Carrillo, and M. W. Varner. 2001. Paternal and maternal components of the predisposition to preeclampsia. *N. Engl. J. Med.* 344:867-872.

12   Lie, R. T., S. Rasmussen, H. Brunborg, H. K. Gjessing, E. Lie-Nielsen, and L. M. Irgens. 1998. Fetal and maternal contributions to risk of pre-eclampsia: population based study. *BMJ* 316:1343-1347.

13 Plunkett, J., I. Borecki, T. Morgan, D. Stamilio, and L. J. Muglia. 2008. Population-based estimate of sibling risk for preterm birth, preterm premature rupture of membranes, placental abruption and pre-eclampsia. *BMC Genet.* 9:44.

14 Buurma, A. J., R. J. Turner, J. H. Driessen, A. L. Mooyaart, J. W. Schoones, J. A. Bruijn, K. W. Bloemenkamp, O. M. Dekkers, and H. J. Baelde. 2013. Genetic variants in pre-eclampsia: a meta-analysis. *Hum. Reprod. Update* 19:289-303.

15 Staines-Urias, E., M. C. Paez, P. Doyle, F. Dudbridge, N. C. Serrano, J. P. Ioannidis, B. J. Keating, A. D. Hingorani, and J. P. Casas. 2012. Genetic association studies in pre-eclampsia: systematic meta-analyses and field synopsis. *Int. J. Epidemiol.* 41:1764-1775.

16 Johnson, M. P., S. P. Brennecke, C. E. East, H. H. Goring, J. W. Kent, Jr., T. D. Dyer, J. M. Said, L. T. Roten, A. C. Iversen, L. J. Abraham, S. Heinonen, E. Kajantie, J. Kere, K. Kivinen, A. Pouta, H. Laivuori, R. Austgulen, J. Blangero, and E. K. Moses. 2012. Genome-wide association scan identifies a risk locus for preeclampsia on 2q14, near the inhibin, beta B gene. *PloS one* 7:e33666.

17 Lech, M., and H. J. Anders. 2013. The pathogenesis of lupus nephritis. *J. Am. Soc. Nephrol.* 24:1357-1366.

18 Li, H., and R. Durbin. 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 26:589-595.

19 McKenna, A., M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytsky, K. Garimella, D. Altshuler, S. Gabriel, M. Daly, and M. A. DePristo. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20:1297-1303.

20 DePristo, M. A., E. Banks, R. Poplin, K. V. Garimella, J. R. Maguire, C. Hartl, A. A. Philippakis, G. del Angel, M. A. Rivas, M. Hanna, A. McKenna, T. J. Fennell, A. M. Kernytsky, A. Y. Sivachenko, K. Cibulskis, S. B. Gabriel, D. Altshuler, and M. J. Daly. 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* 43:491-498.

21 an der Auwere, G. A., M. Carneiro, C. Hartl, R. Poplin, G. del Angel, A. Levy-Moonshine, T. Jordan, K. Shakir, D. Roazen, J. Thibault, E. Banks, K. Garimella, D. Altshuler, S. Gabriel, and M. DePristo. 2013. From FastQ Data to High-Confidence Variant Calls: The Genome Analysis Toolkit Best Practices Pipeline. *Current Protocols in Bioinformatics* 43:11.10.11-11.10.33.

22 Cingolani, P., A. Platts, L. Wang le, M. Coon, T. Nguyen, L. Wang, S. J. Land, X. Lu, and D. M. Ruden. 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly* 6:80-92.

23 Yu, Y., M. P. Triebwasser, E. Wong, E. C. Schramm, B. Thomas, R. Reynolds, E. Mardis, J. P. Atkinson, M. Daly, S. Raychaudhuri, D. Kavanagh, and J. M. Seddon. 2014. Whole exome sequencing of families with macular degernation identifies rare functional mutations in CFH. *Am. J. Hum. Genet.* Submitted.

24 Salmon, J. E., C. Heuser, M. Triebwasser, M. K. Liszewski, D. Kavanagh, L. Roumenina, D. W. Branch, T. Goodship, V. Fremeaux-Bacchi, and J. P. Atkinson. 2011. Mutations

in complement regulatory proteins predispose to preeclampsia: a genetic analysis of the PROMISSE cohort. *PLoS Med* 8:e1001013.

25   Salmon, J. E., C. Heuser, M. Triebwasser, M. K. Liszewski, D. Kavanagh, L. Roumenina, D. W. Branch, T. Goodship, V. Fremeaux-Bacchi, and J. P. Atkinson. 2011. Mutations in complement regulatory proteins predispose to preeclampsia. *PLoS Medicine* 8:e1001013.

26   Fang, C. J., V. Fremeaux-Bacchi, M. K. Liszewski, G. Pianetti, M. Noris, T. H. Goodship, and J. P. Atkinson. 2008. Membrane cofactor protein mutations in atypical hemolytic uremic syndrome (aHUS), fatal Stx-HUS, C3 glomerulonephritis, and the HELLP syndrome. *Blood* 111:624-632.

27   Krych, M., L. Clemenza, D. Howdeshell, R. Hauhart, D. Hourcade, and J. P. Atkinson. 1994. Analysis of the functional domains of complement receptor type 1 (C3b/C4b receptor; CD35) by substitution mutagenesis. *J. Biol. Chem.* 269:13273-13278.

28   Hourcade, D., D. R. Miesner, C. Bee, W. Zeldes, and J. P. Atkinson. 1990. Duplication and divergence of the aminoterminal coding region of the complement receptor 1 (CR1) gene. *J. Biol. Chem.* 265:974-980.

29   Kittles, R. A., M. Perola, L. Peltonen, A. W. Bergen, R. A. Aragon, M. Virkkunen, M. Linnoila, D. Goldman, and J. C. Long. 1998. Dual origins of Finns revealed by Y chromosome haplotype variation. *Am. J. Hum. Genet.* 62:1171-1179.

30   Seddon, J. M., Y. Yu, E. C. Miller, R. Reynolds, P. L. Tan, S. Gowrisankar, J. I. Goldstein, M. Triebwasser, H. E. Anderson, J. Zerbib, D. Kavanagh, E. Souied, N. Katsanis, M. J. Daly, J. P. Atkinson, and S. Raychaudhuri. 2013. Rare variants in CFI, C3 and C9 are associated with high risk of advanced age-related macular degeneration. *Nat. Genet.* 45:1366-1370.

31   Kuhn, S., and P. F. Zipfel. 1996. Mapping of the domains required for decay acceleration activity of the human factor H-like protein 1 and factor H. *Eur. J. Immunol.* 26:2383-2387.

32   Gordon, D. L., R. M. Kaufman, T. K. Blackmore, J. Kwong, and D. M. Lublin. 1995. Identification of complement regulatory domains in human Factor H. *J. Immunol.* 155:348-356.

33   Sharma, A. K., and M. K. Pangburn. 1996. Identification of three physically and functionally distinct binding sites for C3b in human complement factor H by deletion mutagenesis. *Proc. Natl. Acad. Sci. USA* 93:10996-11001.

34   Dragon-Durey, M.-A., V. Fremeaux-Bacchi, C. Loirat, J. Blouin, P. Niaudet, G. Deschenes, P. Copp, W. H. Fridman, and L. Weiss. 2004. Heterozygous and homozygous factor H deficiencies associated with hemolytic uremic syndrome or membranoproliferative glomerulonephritis: Report and genetic analysis of 16 cases. *J. Am. Soc. Nephrol.* 15:787-795.

35   Abrera-Abeleda, M. A., C. Nishimura, K. Frees, M. Jones, T. Maga, L. M. Katz, Y. Zhang, and R. J. Smith. 2011. Allelic variants of complement genes associated with dense deposit disease. *J. Am. Soc. Nephrol.* 22:1551-1559.

36   Maga, T. K., C. J. Nishimura, A. E. Weaver, K. L. Frees, and R. J. Smith. 2010. Mutations in alternative pathway complement proteins in American patients with atypical hemolytic uremic syndrome. *Hum. Mutat.* 31:E1445-1460.

37   Milder, F. J., L. Gomes, A. Schouten, B. J. Janssen, E. G. Huizinga, R. A. Romijn, W. Hemrika, A. Roos, M. R. Daha, and P. Gros. 2007. Factor B structure provides insights into activation of the central protease of the complement system. *Nat. Struct. Mol. Biol.* 14:224-228.

38   Baronciani, L., A. B. Federici, G. Castaman, M. Punzo, and P. M. Mannucci. 2008. Prevalence of type 2b 'Malmo/New York' von Willebrand disease in Italy: the role of von Willebrand factor gene conversion. *J Thromb Haemost* 6:887-890.

39   Holmberg, L., J. A. Dent, R. Schneppenheim, U. Budde, J. Ware, and Z. M. Ruggeri. 1993. von Willebrand factor mutation enhancing interaction with platelets in patients with normal multimeric structure. *J. Clin. Invest.* 91:2169-2177.

40   Weiss, H. J., and Sussman, II. 1986. A new von Willebrand variant (type I, New York): increased ristocetin-induced platelet aggregation and plasma von Willebrand factor containing the full range of multimers. *Blood* 68:149-156.

41   Forberg, E., I. Huhmann, E. Jimenez-Boj, and H. H. Watzke. 2000. The impact of Glu102Lys on the factor X function in a patient with a doubly homozygous factor X deficiency (Gla14Lys and Glu102Lys). *Thromb. Haemost.* 83:234-238.

42   Fremeaux-Bacchi, V., E. A. Moulton, D. Kavanagh, M.-A. Dragon-Durey, J. Blouin, A. Caudy, N. Arzouk, R. Cleper, M. Francois, G. Guest, J. Pourrat, R. Seligman, W. H. Fridman, C. Loirat, and J. P. Atkinson. 2006. Genetic and functional analyses of membrane cofactor protein (CD46) mutations in atypical hemolytic uremic syndrome. *J. Am. Soc. Nephrol.* 17:2017-2025.

43   Jonsen, A., S. C. Nilsson, E. Ahlqvist, E. Svenungsson, I. Gunnarsson, K. G. Eriksson, A. Bengtsson, A. Zickert, M. L. Eloranta, L. Truedsson, L. Ronnblom, G. Nordmark, G. Sturfelt, and A. M. Blom. 2011. Mutations in genes encoding complement inhibitors CD46 and CFH affect the age at nephritis onset in patients with systemic lupus erythematosus. *Arthritis Res Ther* 13:R206.

44   Tennessen, J. A., A. W. Bigham, T. D. O'Connor, W. Fu, E. E. Kenny, S. Gravel, S. McGee, R. Do, X. Liu, G. Jun, H. M. Kang, D. Jordan, S. M. Leal, S. Gabriel, M. J. Rieder, G. Abecasis, D. Altshuler, D. A. Nickerson, E. Boerwinkle, S. Sunyaev, C. D. Bustamante, M. J. Bamshad, and J. M. Akey. 2012. Evolution and functional impact of rare coding variation from deep sequencing of human exomes. *Science* 337:64-69.

45   Legendre, C. M., C. Licht, P. Muus, L. A. Greenbaum, S. Babu, C. Bedrosian, C. Bingham, D. J. Cohen, Y. Delmas, K. Douglas, F. Eitner, T. Feldkamp, D. Fouque, R. R. Furman, O. Gaber, M. Herthelius, M. Hourmant, D. Karpman, Y. Lebranchu, C. Mariat, J. Menne, B. Moulin, J. Nurnberger, M. Ogawa, G. Remuzzi, T. Richard, R. Sberro-Soussan, B. Severino, N. S. Sheerin, A. Trivelli, L. B. Zimmerhackl, T. Goodship, and C. Loirat. 2013. Terminal complement inhibitor eculizumab in atypical hemolytic-uremic syndrome. *N. Engl. J. Med.* 368:2169-2181.

Supplement 4-1. PE Candidate Genes Selected for Sequencing

AP/MAC/Regulators	Classical Pathway	Inhibin Axis
C5AR1	C1QA	INHBB
GPR77	C1QB	INHBA
C3AR1	C1QC	INHBC
C3	C1S	INHBE
PZP	C1R	INHA
CRP	SERPING1	ACVR2A
C5	C2	ACVR2B
C6	C4A	ACVRL1
C7	C4B	ACVR1B
C8A	C4BPA	ACVR1C
C8B	C4BPB	ACVR1
C8G	C1QBP	INHBB
C9	CD93	
MCP	CALR	**TGFB/ENG**
DAF		TGFB1
CFI	**Lectin Pathway**	TGFB2
CFH	MBL2	TGFB3
CFHR1	MASP1	TGFBR1
CFHR2	MASP2	TGFBR2
CFHR3	FCN1	TGFBRAP1
CFHR4	FCN2	ENG
CFHR5	FCN3	
CFP	SFTPA1	
CFD	SFTPA2	**Angiopoietin and Rec.**
CFB	SFTPD	ANGPT1
CD59		ANGPT2
CLU	**Angiogenesis**	ANGPT3
VTN	VEGFA	ANGPT4
THBD	VEGFB	TIE1
	VEGFC	TIE2
Complement Receptors	FLT1	
CR1	KDR	
CR1L	FLT4	
CR2	PGF	
ITGAM		
ITGAX		
ITGB2		
VSIG4		

Clotting Pathway	Platelet Aggregation	Candidate Genes
TF	PEAR1	STOX1
TFPI	ADRA2A	STOX2
TFPI2	PIK3CG	Cdkn1c
PROS1	JMJD1C	EHD3
PROC	MRVI1	EHD4
SERPINC1	SHH	NOTCH2
F2R		JAG1
F2RL1		HEY1
F2RL2	**Cell Surface Ligands**	HEY2
F2RL3	SELP	ADM
F7	SELPLG	ADM2
F10	CD36	ROCK1
F3		ROCK2
F9	**VWF LEVEL:**	DEF6
F8	ABO	SWAP70
F11	STXBP5	ADAM10
F5	STX2	ADAM17
PROCR	SVIL	ADAM8
PROZ		ADAM9
SERPINA1	**VWF Binding:**	ADAM12
SERPINA5	ITGA2B	ADAM15
SERPINE1	ITGB3	ADAM19
SERPINB2	GP1BA	ADAM28
SERPINA10	GP1BB	TNF
FGA	GP9	LCT
FGG	GP5	GPR98
FGB	GP6	IL10
PLAT		TREX1
PLAU		CORIN
PLG	**Assc with CAD**	NODAL
SERPINE2	LIPA	CTLA4
SERPIND1	PDGFD	ACE
F13B	ADAMTS7	LPL
	KIAA1462	F13A1
	LRRFIP1	ITGB1
	COMMD7	ANTXR1
		ITGA2
		ITGB1
		DGKE
		IP6K1
		MAGI1
		MME

Chapter 5

**Rare Variants in the Alternative Pathway Increase the Risk of
Age-Related Macular Degeneration**

<u>Introduction</u>

Genetic evidence supporting the role of AP activation in the eye has already been discussed in Chapter 1. In summary, one of the most potent common SNPs discovered by GWAS is Y402H in *CFH* (OR=2.5). *C3, CFB,* and *CFI*, all components of the AP have also been found to associated with AMD. A highly penetrant mutation in *CFH,* R1210C, was found in families with AMD (1). This mutation had previously been described in aHUS and has a known defect in GAG binding, similar to the effect of the Y402H allele (2, 3). We identified additional that rare mutations in the AP that increase

The studies summarized below are in different stages relative to being published. The first paper (Seddon, JM, et. al. 2013. *Nature Genetics* – Supplemental Manuscript #1) was a sequencing study utilizing a case and control design to find rare variants that affect the risk of AMD. I aided in study design and target selection as well as interpretation of the *CFI, C3* and *C9* variants. I also made significant editorial contributions to the manuscript. Several additional unpublished studies relevant to the above findings are described below as well.

A second manuscript detailing haploinsufficiency of FI in the blood of *CFI* rare variant carrier cases versus cases without a rare variant and versus controls with a rare variant is in preparation.

The third study (Yi, Y, Triebwasser, MP, Schramm, EC, Wong, E, et. al. Submitted. *American Journal of Human Genetics* – Supplemental Manuscript #2), describes rare variants in *CFH* that are causative for AMD in families. Exome sequencing of pedigrees was used to identify these variants. I did the functional studies of the *CFH* variants in the mammalian system, which involved significant assay optimization given the intermediate loss of function defects

125

these variants confer. I also wrote a substantial portion of the manuscript contributed significantly editorially; I am a co-first author.

A fourth paper is in preparation. It is entirely my work and is based on further analysis of the case control sequencing data that was also initially described in the Seddon, et. al. 2013 *Nature Genetics* paper. In this study I identified a series of rare variants in factor H that are associated with AMD.

A Diversity of Rare Variants in *CFI* Are Associated with AMD (Seddon, JM, et. al. 2013 *Nature Genetics)*

Study Population, Design and Analysis

1,712 AMD cases and 781 phenotyped controls (2,493 total) were sequenced for 688 candidate genes. They were chosen by identifying pathways involved in complement, retinal pigment epithelial cell biology, photoreceptor biology, HDL metabolism, inflammation and oxidative damage, as well as genes implicated by previous AMD GWAS results and those associated with diseases related to AMD. Genes were tested for either an increased burden in cases or an increased burden in controls for variants with a MAF <1% in the 2,493 samples and that led to missense, nonsense, read-through or splice-site changes.

A Diversity of Rare Variants in CFI Increase the Risk of AMD

CFI ($p=1.6x10^{-8}$), exceeded a Bonferroni corrected threshold, $3.6x10^{-5}$ (0.05/(688 * 2)). 7.9% of cases (136 individuals) carried a CFI variant and 2.3% of controls (18 individuals) carried such a variant. There were 59 unique variants in total in CFI. The association with CFI was significant when adjusting for ancestry ($p=5.0x10^{-8}$, OR=3.6) or when adjusted for age and gender ($p=6.7x10^{-9}$, OR=3.7). Enrichment was also unrelated to the presence of a common SNP

near CFI when a stratified analysis was done based on the genotype at rs4698775 (p=1.7x10^{-8} for rare variants increasing risk of AMD). The effect of the common SNP was not obviated by the rare variants (OR=1.15 when controlling for rare variant status), evidence against a synthetic association.

This finding is further supported by the fact that many of these variants (12 of 59) have been observed in aHUS. There were four nonsense mutations and one splice donor mutation that were present in a total of seven cases and zero controls. Importantly, no missense, splice site, nonsense or read through mutation in CFI was present at greater than 1% in the population.

There is an increased burden of variants in the catalytic (serine protease) light chain relative to the heavy chain (OR=4.85 vs 2.63, respectively). The heavy chain is physically larger 55 vs 35 kDa, and, if variants were distributed randomly throughout the protein, the heavy chain would be expected to have a larger share. Taken as a class, variants predicted to be loss of function or damaging by Polyphen-2 collectively had an OR of 7.5, whereas those predicted to be possibly damaging had an OR of 2.4.

Rare CFI Variants in AMD Cases Lead to Low Blood Levels of FI

Subsequent analysis of serum levels of FI in 77 cases carrying a rare variant versus 48 cases without a rare variant or 42 controls with a rare variant (and 48 controls without a rare variant) has demonstrated that roughly half of the variants in CFI cases lead to a production defect and haploinsufficiency (Figure 5-1 and 5-2) (Kavanagh, et. al. Hum Mol Genetics. Accepted).
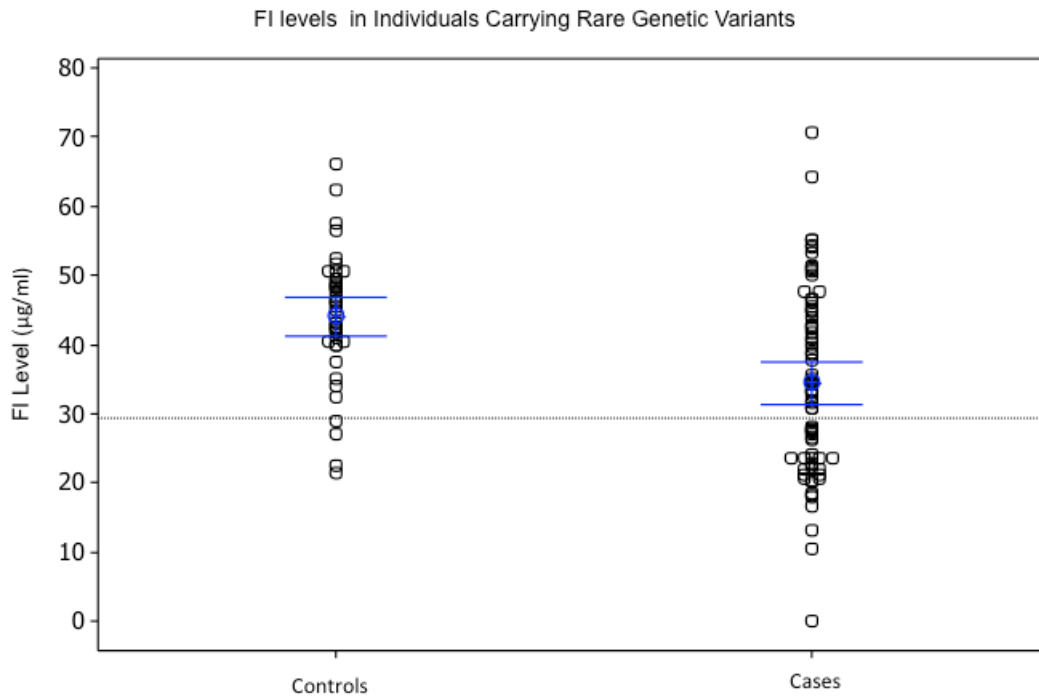
Figure 5-1. AMD Cases that Carry a Rare Variant in *CFI* Have Lower Serum Levels Than Controls With a Rare Variant

Controls (N=42) that carry a rare coding variant in FI have an average FI level of 44.1 μg/ml versus AMD Cases (N=77) which have an average FI level of 35.5 μg/ml (p<0.05). The FI level observed for each individual carrying a given FI variant. Grey circles are cases and open circles are controls. Grey line delineates the lower bound of the normal range, 29.7 μg/ml.
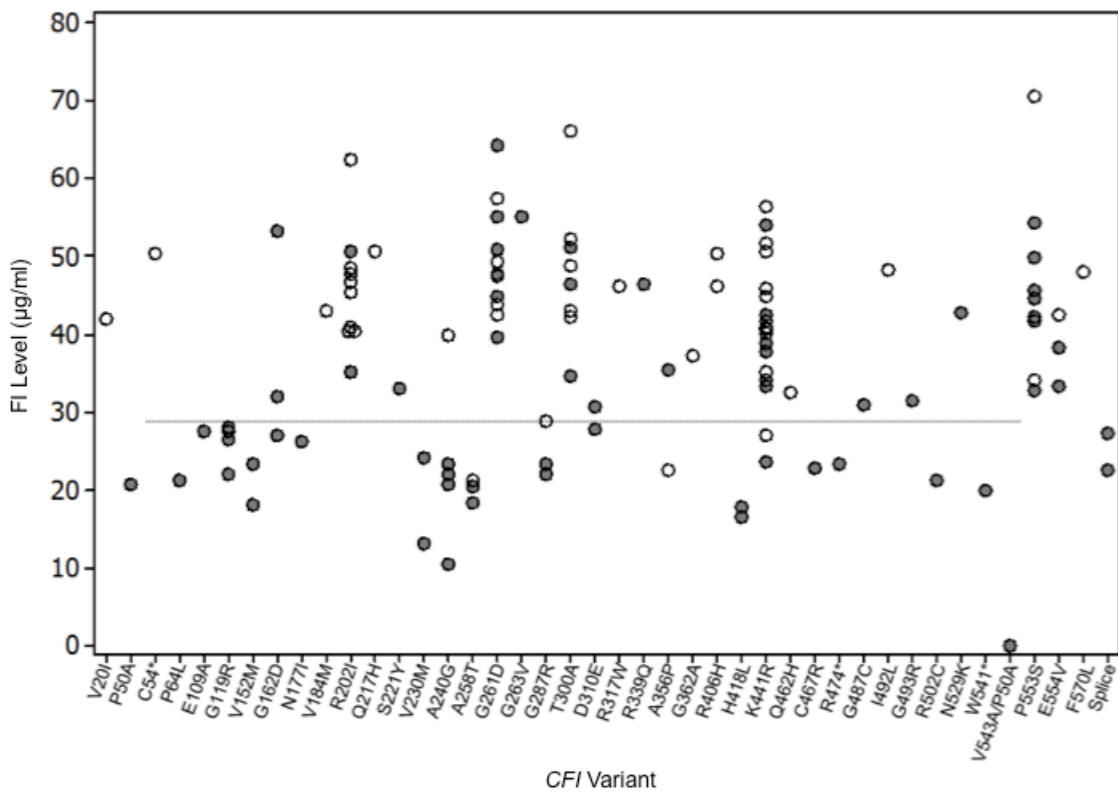
Figure 5-2. FI Haploinsufficiency is Common Among AMD Cases with a Rare Variant in *CFI*.

Grey circles are cases and open circles are controls. Roughly half of the cases with a rare variant have levels below the normal range (29.7 µg/ml) compared to 2 of 42 controls with a rare variant.

A Rare Variant in C3 and in C9 Increase the Risk of AMD

Two variants, K155Q in *C3* and P167S in *C9*, were found to be associated with AMD in this population and subsequent replication cohorts populations. K155Q of C3 had an odds ratio of 3.8 (MAF 1.3% in cases and 0.4% in controls). It results in lower binding to CFH and a significant defect in cleavage by CFI when FH serves as the cofactor, but not when MCP or CR1 is the cofactor. Two other groups reported the K155Q mutation in C3 at the same time we did (4, 5).

The C9 variant had an OR of 2.2 (MAF 1.8% in cases vs 0.9% in controls). Attempts to demonstrate an effect on the hemolytic activity of serum from P167S heterozygous individuals have been unsuccessful and P167S was indistinguishable from wild type when examined by Western blotting under reducing or non-reducing conditions (unpublished data, Triebwasser, MP and Schramm, E). There was no enrichment for missense, stop or splice site mutations in C9 (75 vs 37, 2-fold vs 2.2 fold expected).

Because C9 interacts with the regulator CD59, we thought that mutations in CD59 might be able to tell us more about how mutations in C9 increase risk. No individual in the 2,493 carried a coding mutation in CD59. In the 6500 individuals in the NHLBI ESP, 6 individuals carried a singleton or doubleton missense mutation in the mature protein sequence of CD59. From this summary, we conclude that CD59 function is critical and that coding changes in this gene are not well tolerated. The interaction of this C9 variant with CD59 requires further study.

Rare *CFH* Mutations are Causative in Families with a Low Burden of Common AMD Risk

Variants (Yi, Y, Triebwasser, MP, Schramm, E, Wong, E, et. al. Submitted. *AJHG)*

Family Selection and Variant Selection For Segregation Analysis

Ten families with a disproportionally low genetic risk given the burden of disease or with

early onset of disease had the exome of multiple affected members sequenced. An unaffected

family member was sequenced when one was available. Variants that were expected to alter the

protein product, had a MAF of <0.1% in public datasets, and satisfied an autosomal dominant

mechanism (present in all affecteds in that family and none of the controls) were considered.

Additionally, variants were required to be deleterious by SIFT or probably damaging or worse by

Polyphen-2. All such variants were genotyped in all available family members to confirm their

existence and segregation.

Two Families Have Rare or Novel CFH Variants That Segregate with Disease

Two of the ten families had a mutation in FH that segregated with disease. The R53C

variant in *CFH* was observed in eleven affected family members (LOD > 5). It had previously

been observed in aHUS and a different change at the position, R53H, had also been reported in

aHUS and DDD. The second family had a D90G variant in 5 affecteds (LOD=1.22). No other

family had a unique variant that satisfied these requirements, though two families had variants

with LOD scores >1, but in genes previously unassociated with AMD. Lacking a larger pedigree

for either family, additional evidence is required to implicate these genes.

Functional Analysis of R53C and D90G

We studied (myself, Elizabeth Schramm (St. Louis) and Edwin Wong (Newcastle, UK)) the CFH variants in two different groups using protein made in yeast *Pichia pastoris* (Newcastle) and in human 293T cells (St. Louis). Both groups performed functional analyses and they agreed. R53C has a severe defect in carrying out decay acceleration and a minor loss of cofactor activity. R53C was identified in 3 additional unrelated cases in the case:control study already discussed (6). The D90G mutation that segregated with disease in the second family had normal decay accelerating activity, but significantly lower cofactor activity. Generally, C3b binding was normal, though both groups found R53C to have a higher K_D. See Supplemental Manuscript #2 for the details on these functional assessments.

Rare Variants in the Functional Domains of *CFH* Are Associated with AMD (Triebwasser, MP, et. al. In Preparation).

The association observed in CFI was detectable because of its large effect size, prevalence in the population, and lack of neutral variants in the 0.5-1% MAF range in the control group. In fact, there were only two variants present at >0.1% (more than two observations in controls). Of these two, one was enriched 3.5-fold in cases and the other 1.4-fold. The ability to detect any signal is related to the "noise," and in the case of CFI there was very little noise.

I hypothesized that additional rare variants would be present in genes of the AP in AMD cases. Findings that support this include the penetrance of rare variants in *CFH* (1, 7), the presence of a rare variant in C3, and an excess of rare variants in *CFI* (6).

The Effect of CFH Variants is Dependent Upon the Definition of Rare

Using the same 2,493 individuals as described above (6), I examined whether there was an excess of rare missense, nonsense, splice-site or read-through mutations in FH in AMD cases. At a definition of rare at 5% MAF in the control group, there is no excess of rare variants in the cases (OR=1.1, p=0.3). With a definition of rare at 1% MAF in the control group, there is an excess in cases (134 vs 49, OR=1.25, p=0.097). This is the set of criteria used in the analysis described above that discovered the association with *CFI*.

This burden however becomes clearer as our definition of rare becomes increasingly stringent. At the 0.5% and 0.13% MAF points, the association is stronger with OR of 1.65 and 2.88, respectively (p=7×10^{-3} and p=2×10^{-5}) (Table 5-1) (MAF 0.13% in controls is 2 observations of an allele in the 781 controls). Attempts to utilize more sophisticated alternatives to the burden test that use variant frequency to weight variants (Variable-Threshold test) failed to identify stronger evidence for an association between FH and variants below 1% MAF.

Table 5-1. Burden of Rare Variants in *CFH* Increases as the Frequency of the Variants Decreases					
MAF in Controls	# Unique Variants	# Alleles in Cases	# Alleles in Controls	Odds Ratio	p-value
1%	75	134	49	1.25	0.1
0.5%	73	111	31	1.65	7×10^{-3}
0.13%*	70	93	15	2.88	2×10^{-5}

Minor allele frequency (MAF) in the control group (1,562 chromosomes). *MAF 0.13% corresponds to 2 or fewer alleles in the control group. P-value derived from Fisher's exact test using the number of alleles/chromosomes in each category.

Cases Exclusively Carry Loss of Function and aHUS-Associated Variants

Complement factor H consists of 20 CCP domains, each having two conserved disulfide bonds. Loss of a conserved cysteine typically leads to a failure of the protein to be secreted because it is not folded properly. There are 11 cases and zero controls that carry a variant leading to a nonsense change, loss of a conserved cysteine or in a canonical splice site (Table 5-2). Very rare, essentially private variants previously reported in aHUS also only occur in cases.

Table 5-2. Variants Likely to Impact *CFH* Serum Levels Are Present Exclusively in Cases

	Cases	Controls
Nonsense	4	0
Loss of Conserved Cysteine	4	0
Splice Site	3	0

Every CCP domain contains two conserved disulfide bonds. When one of the 4 conserved cysteines required to form these bonds is lost, the protein will not fold correctly.

Rare Variants in Functional Domains of CFH Confer Risk

The domain of factor H that encodes regulatory activity (CCP1-4) has an excess of rare variants (MAF < 0.5%, OR=4.0, p=0.0018) (Table 5-3). Variants in cases cluster at the interface between FH1-4 and C3b whereas those in controls are consistently found facing away from the C3b interface (Figure 5-3). There is a precedent through for variants in CCP1 that do not contact the C3b structure to affect decay acceleration (for example, R53C), as CFB binds up against the C345C domain (8, 9).

There is also an excess of rare variants in the anionic binding domains located at CCP 6-9 and in CCP 19-20 (Table 5-3). CCP7 is the site of the common AMD risk variant Y402H, which affects heparin binding (3). The variants found in CCP20 all cluster on the same face as the R1210C variant, which is known to be highly penetrant and to affect GAG binding (Figure 5-4) (1, 2). Interestingly, there are no variants in CCP19 or 20 that directly oppose C3d binding.
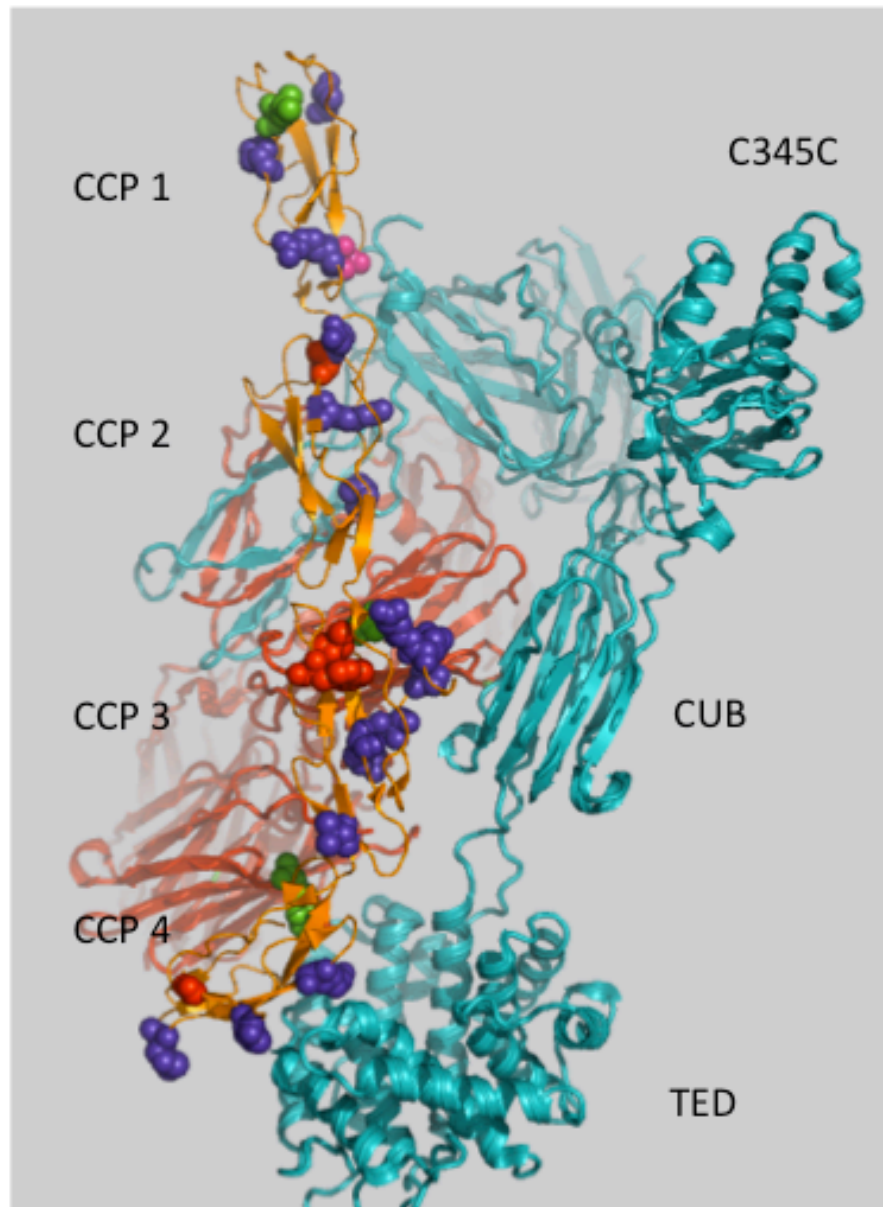
Figure 5-3. Rare *CFH* Variants in AMD Cases Within CCPs 1-4 Cluster at the C3b:FH Interface

Variants only in AMD cases are in purple. Variants that are nonsense or remove a conserved cysteine are in red (these are only seen in cases). Variants only in controls are in green. One variant was observed in cases and controls (magenta). The variants in CCP3 contact the CUB domain and variants in CCP4 contact the TED domain.
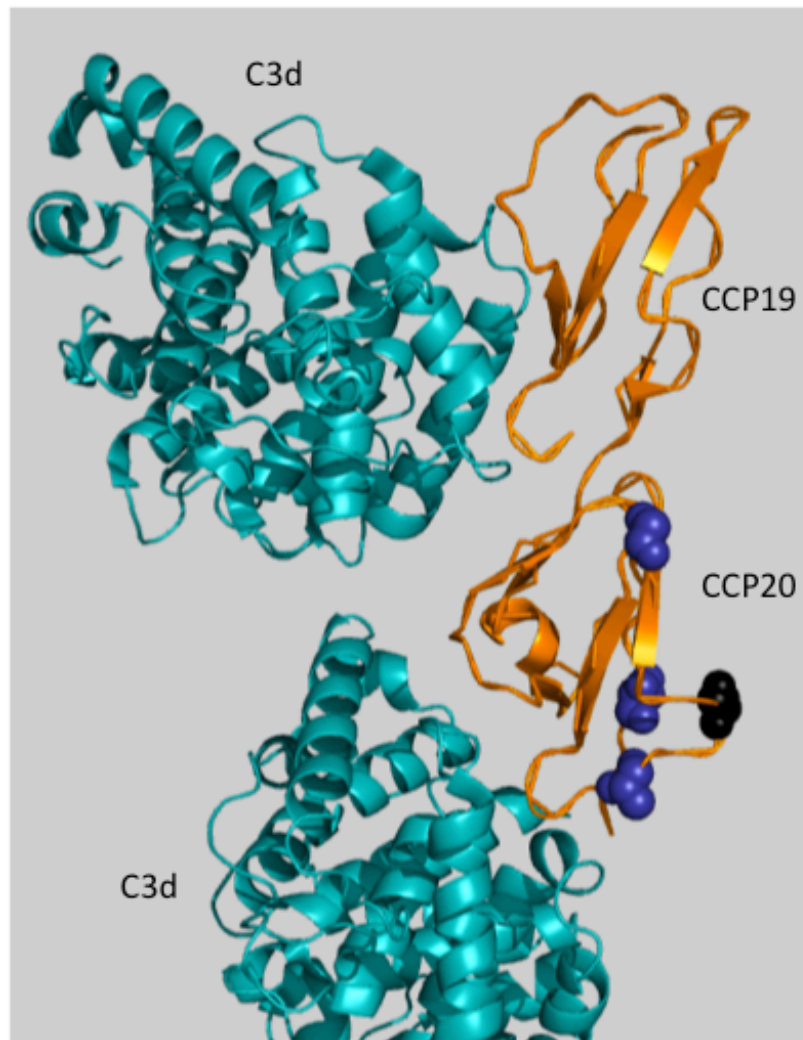
Figure 5-4. *CFH* **Rare Variants in CCP20 Cluster Near R1210C**
FH CCP 19-20 is in orange. C3d is in teal. R1210C variant is in black. The crystal structure contains two C3d fragments (PDB ID 2XQW). Mutations only in AMD cases are in purple. The CCP19 variant (Q1143, in 1 control) is not present in the crystal structure. The rare variants cluster on the same face as R1210C, which is known to affect heparin/GAG binding (2).

If we consider the variants (Table 5-3) in these three functional domains as a single class (rare variants that are in domains of functional importance of CFH), they are associated with AMD (MAF < 0.5% in controls, OR=5.1, p=$4.2x10^{-7}$). 4.4% of cases carry a rare variant versus 0.9% of controls. The previously reported R1210C variant does account for some of this association, but carrying a rare variant in these domains is still associated if we exclude R1210C (OR=4.6, p=$1.9x10^{-5}$).

Table 5-3. Individuals Carrying Mutations in Functional Domains of FH					
Domain	Function	Cases	Controls	Odds Ratio	p-value
CCP 1-4	DAA/CA	35 (2.0)	4 (0.5)	4.05	0.0018
CCP 6-9	Localization	15 (0.9)	1 (0.1)	NA	NA
CCP 19-20	Localization	24 (1.4)	2 (0.3)	NA	NA
CCP 19-20 w/o R1210C	Localization	7 (0.4)	1 (0.1)	NA	NA
All Functional Domains		76 (4.4)	7 (0.9)	5.1	$4.2x10^{-7}$
All Functional Domains w/o R1210C		69	6	4.6	$1.9x10^{-5}$

Number in parenthesis is % of individuals. If an individual carried two mutations mutation, they were assigned to be a carrier of the N-terminal mutation. All functional domains include missense mutations in CCP 1 to 4, 6 to 9, and 19 to 20, as well as nonsense and splice site mutations in other domains. NA: Test not done because few variants. R1210C is broken out as it is the single largest contributor to the effect seen in CCP19-20 and has previously been described. P-value derived from Fisher's Exact Test using the number of individuals in each category.

Finally, rare coding variants in CFH are associated with AMD independent of the two

common risk loci in CFH (rs1061170 (Y402H) and rs10737680 (intronic) (Table 5-4). Y402H

has an OR of 1.94, while rs10737680 has an OR of 1.85 in a conditional additive model. If all

rare coding variants in FH are considered, they have an OR of 1.65 (p=0.029) when both

common SNPs are accounted for. If only those rare coding SNPs in CCP 1-4 are considered, they

have an OR of 4.7 (p=0.0045) (Table 5-4). Given the small sample sizes, I also tested the

significance of the association of all rare coding variants in *CFH* or just those in CCP 1-4 using a

permutation based conditional logistic regression that accounts for Y402H and rs10737680 and

found the association to remain significant (p=0.024, all rare *CFH* variants, and p=$6x10^{-4}$ rare

variants in CCP 1-4).

Table 5-4. Rare Variants are Associated with AMD Independent of the Common SNPs in *CFH*

SNP	Consequence	OR	p-value
rs1061170	Y402H	1.94	$4.8x10^{-14}$
Rs10737680	Intronic	1.85	2.7x10
All Rare Variants		1.65	0.029
Rare Variants in CCP 1-4		4.7	0.0045

In a conditional regression, the effect of the common variants is not affected by the addition of

missense, nonsense and splice site mutations with MAF < 0.5%. Rare variants throughout the

gene are also significantly associated if permutation is used to estimate significance (p=0.024).

Rare variants in CCP 1-4 are associated independent of the common variants using either

conditional logistic regression or permutation based conditional logistic regression (p=$6x10^{-4}$).

<u>Discussion</u>

Rare variants in *CFI, C3* and *CFH* are present in severe AMD cases. A subset of alleles in *CFH* have very high penetrance, including R53C, D90G and the previously identified R1210C. There is also a rare variant in *CFB* that is present in 19 cases and 1 control. This variant has previously been reported in an aHUS case as well and we are pursuing further studies. These variants implicate cofactor activity as required for homeostasis of the AP, and AP activity in the degenerative disease AMD.

The mechanism linking excessive AP activity to the pathologic process in AMD is not known. But two different rare variants in C9 have been reported (6, 10). One is a nonsense mutation that is protective for wet AMD and the other is a missense mutation that increases risk of AMD. Sublytic C5b-9 deposition drives retinal pigment epithelial cells to produce vascular endothelial growth factor (VEGF) (11). This local production of VEGF would be central to the proliferation of vessels in wet AMD. If the null allele in C9 leads to decreased C5b-9, this would be consistent with a reduced production of VEGF and protection from wet AMD.

It is particularly interesting why these individuals present with AMD and not aHUS or DDD. Additional genetic and environmental factors must exist that affect the risk of developing kidney disease versus retinal disease. While unaffected relatives of aHUS patients carry the same mutation, they often have normal renal function or only mild impairment. Penetrance here may reflect both different levels of AP activity or susceptibility in carriers with the same mutation, and also the susceptibility of the kidney versus the eye.

The kidney is an organ with enormous functional reserve. Throughout evolution the kidney and liver were the front line in eliminating toxins from the body and thus would have been under constant assault. It is estimated that only when a person has <10% functional renal capacity will they start to be symptomatic. Thus even those individuals with AP risk variants may live a life free of kidney disease. For example of the 11 R53C carriers in the AMD family discussed here, none had abnormal measures of kidney function.

In contrast, AMD affects people in the last third of life and is a degenerative disease. It may be that both common variants and rare variants in the AP lead to excessive complement activation in the retina over decades, resulting in years of cumulative damage that destroys the excess capacity of the retina. The drusen deposits that accumulate between the retinal pigment epithelial cells and the vascular supply of the eye are known to activate complement and contain C3 fragments (12). Unlike the kidney, whose function most people do not monitor daily, eyes are used constantly. Whether the functional reserve in the eye is lower or we are more vigilant in monitoring it, those with retinal damage will usually be diagnosed.

The enrichment of *CFI* variants was easily detected at the gene level employing a widely utilized MAF cutoff of 1%. However, no cutoff was necessary because there are effectively no common coding variants in CFI (none present at >0.5% MAF in our 781 controls or the NHLBI ESP European-American population). This is in contrast to *CFH* where there is a great diversity of alleles across the frequency spectrum. Only when considering the rarest alleles do we see a strong association. Further subsetting on the domains of FH known to have functional importance in AP control only increases the significance of this association.

141

Our current work is focused on understanding the effects of these FH variants, both whether they lead to decreased protein in the blood and if they have functional consequences. We are developing a high throughput assay that utilizes patient serum as a screening tool. Additionally, we are trying to extend as many cases with *CFI* and *CFH* mutations into families as possible to estimate the penetrance of these mutations.

References

1   Raychaudhuri, S., O. Iartchouk, K. Chin, P. L. Tan, A. K. Tai, S. Ripke, S. Gowrisankar, S. Vemuri, K. Montgomery, Y. Yu, R. Reynolds, D. J. Zack, B. Campochiaro, P. Campochiaro, N. Katsanis, M. J. Daly, and J. M. Seddon. 2011. A rare penetrant mutation in CFH confers high risk of age-related macular degeneration. *Nat. Genet.* 43:1232-1236.

2   Jozsi, M., S. Heinen, A. Hartmann, C. W. Ostrowicz, S. Halbich, H. Richter, A. Kunert, C. Licht, R. E. Saunders, S. J. Perkins, P. F. Zipfel, and C. Skerka. 2006. Factor H and atypical hemolytic uremic syndrome: mutations in the C-terminus cause structural changes and defective recognition functions. *J. Am. Soc. Nephrol.* 17:170-177.

3   Clark, S. J., L. A. Ridge, A. P. Herbert, S. Hakobyan, B. Mulloy, R. Lennon, R. Wurzner, B. P. Morgan, D. Uhrin, P. N. Bishop, and A. J. Day. 2013. Tissue-specific host recognition by complement factor H is mediated by differential activities of its glycosaminoglycan-binding regions. *J. Immunol.* 190:2049-2057.

4   Zhan, X., D. E. Larson, C. Wang, D. C. Koboldt, Y. V. Sergeev, R. S. Fulton, L. L. Fulton, C. C. Fronick, K. E. Branham, J. Bragg-Gresham, G. Jun, Y. Hu, H. M. Kang, D. Liu, M. Othman, M. Brooks, R. Ratnapriya, A. Boleda, F. Grassmann, C. von Strachwitz, L. M. Olson, G. H. Buitendijk, A. Hofman, C. M. van Duijn, V. Cipriani, A. T. Moore, H. Shahid, Y. Jiang, Y. P. Conley, D. J. Morgan, I. K. Kim, M. P. Johnson, S. Cantsilieris, A. J. Richardson, R. H. Guymer, H. Luo, H. Ouyang, C. Licht, F. G. Pluthero, M. M. Zhang, K. Zhang, P. N. Baird, J. Blangero, M. L. Klein, L. A. Farrer, M. M. DeAngelis, D. E. Weeks, M. B. Gorin, J. R. Yates, C. C. Klaver, M. A. Pericak-Vance, J. L. Haines, B. H. Weber, R. K. Wilson, J. R. Heckenlively, E. Y. Chew, D. Stambolian, E. R. Mardis, A. Swaroop, and G. R. Abecasis. 2013. Identification of a rare coding variant in complement 3 associated with age-related macular degeneration. *Nat. Genet.* 45:1375-1379.

5   Helgason, H., P. Sulem, M. R. Duvvari, H. Luo, G. Thorleifsson, H. Stefansson, I. Jonsdottir, G. Masson, D. F. Gudbjartsson, G. B. Walters, O. T. Magnusson, A. Kong, T. Rafnar, L. A. Kiemeney, F. E. Schoenmaker-Koller, L. Zhao, C. J. Boon, Y. Song, S. Fauser, M. Pei, T. Ristau, S. Patel, S. Liakopoulos, J. P. van de Ven, C. B. Hoyng, H. Ferreyra, Y. Duan, P. S. Bernstein, A. Geirsdottir, G. Helgadottir, E. Stefansson, A. I. den Hollander, K. Zhang, F. Jonasson, H. Sigurdsson, U. Thorsteinsdottir, and K. Stefansson.

2013. A rare nonsynonymous sequence variant in C3 is associated with high risk of age-related macular degeneration. *Nat. Genet.* 45:1371-1374.

6   Seddon, J. M., Y. Yu, E. C. Miller, R. Reynolds, P. L. Tan, S. Gowrisankar, J. I. Goldstein, M. Triebwasser, H. E. Anderson, J. Zerbib, D. Kavanagh, E. Souied, N. Katsanis, M. J. Daly, J. P. Atkinson, and S. Raychaudhuri. 2013. Rare variants in CFI, C3 and C9 are associated with high risk of advanced age-related macular degeneration. *Nat. Genet.* 45:1366-1370.

7   Yu, Y., M. P. Triebwasser, E. Wong, E. C. Schramm, B. Thomas, R. Reynolds, E. Mardis, J. P. Atkinson, M. Daly, S. Raychaudhuri, D. Kavanagh, and J. M. Seddon. 2014. Whole exome sequencing of families with macular degernation identifies rare functional mutations in CFH. *Am. J. Hum. Genet.* Submitted.

8   Pechtl, I. C., D. Kavanagh, N. McIntosh, C. L. Harris, and P. N. Barlow. 2011. Disease-associated N-terminal complement factor H mutations perturb cofactor and decay-accelerating activities. *J. Biol. Chem.* 286:11082-11090.

9   Forneris, F., D. Ricklin, J. Wu, A. Tzekou, R. S. Wallace, J. D. Lambris, and P. Gros. 2010. Structures of C3b in complex with factors B and D give insight into complement convertase formation. *Science* 330:1816-1820.

10   Nishiguchi, K. M., T. R. Yasuma, D. Tomida, M. Nakamura, K. Ishikawa, M. Kikuchi, Y. Ohmi, T. Niwa, N. Hamajima, K. Furukawa, and H. Terasaki. 2011. C9-R95X Polymorphism in Patients with Neovascular Age-Related Macular Degeneration. *Invest. Ophthalmol. Vis. Sci.*

11   Lueck, K., S. Wasmuth, J. Williams, T. R. Hughes, B. P. Morgan, A. Lommatzsch, J. Greenwood, S. E. Moss, and D. Pauleikhoff. 2011. Sub-lytic C5b-9 induces functional changes in retinal pigment epithelial cells consistent with age-related macular degeneration. *Eye (Lond)* 25:1074-1082.

12   Johnson, L. V., W. P. Leitner, M. K. Staples, and D. H. Anderson. 2001. Complement activation and inflammatory processes in Drusen formation and age related macular degeneration. *Exp. Eye Res.* 73:887-896.

Chapter 6

Future Directions, Conclusions, and Final Remarks

The alternative pathway (AP) of complement is comprised of a powerful feedback loop that can spontaneously initiate as well as be engaged by adaptive and innate immunity. Regulation is critical to prevent tissue damage. A number of human diseases feature AP activation.

Rare mutations in *CFI, CFH, MCP, CFB* and *C3* are a known cause of the rare kidney disease atypical hemolytic uremic syndrome (aHUS) (1). This disease features microthrombi in the glomerular capillary network of the kidney, leading to renal failure (2). Antibodies in the AP and Ab that stabilize the C3 convertase (nephritic factors) are causative in several closely related renal diseases (DDD, C3GN, MPGN I) (3, 4). These diseases featuring renal failure showcase how the kidney is uniquely sensitive to AP activation.

Common variants in *C3, CFH, CFB,* and *CFI* increase the risk of age-related macular degeneration (AMD) (5). The Y402H risk allele is one of the most potent risk factors of any common disease, with an odds ratio of ~2.5 and a minor allele frequency of 30-40% (6). Together, these common risk factors in the AP describe a significant amount of disease risk (~50%) and strongly implicate excessive AP activation in this retinal disease (7).

The focus of this thesis was to understand the effects of unrestricted AP activation on tissue and to determine if individuals affected with common diseases featuring AP activation and kidney disease harbor rare variants in this pathway.

Summary and Conclusions for Crry$^{-/-}$ Embryonic Loss

Unrestricted AP activation leads to loss of *Crry$^{-/-}$* embryos before d8.5, earlier than previously reported (8). This is due to a failure in developmental such that the vessels of the

145

allantois (precursors to the umbilical vessels) do not proliferate and the labyrinth of the placenta is not developed at 9.5 dpc. We exhaustively ruled out the role of neutrophils as previous reports described them as present around *Crry*[-/-] embryos prior to loss. We did not observe evidence for an increased immune cell infiltrate (CD45[+] cells) around *Crry*[-/-] compared to *Crry*[+/-] embryos.

Embryos can be rescued by temporarily blocking the AP with anti-properdin Ab or by depleting C3 utilizing cobra venom factor. AP activation likely begins as soon as the embryo starts to contact maternal blood at 6.5 dpc, with increasing contact happening at each subsequent day. The process that fails to occur in *Crry*[-/-] embryos is complete by 8.5 dpc and embryos are not susceptible to loss after this point. This is interesting because DAF, another cell membrane regulator, is not present until after 10.5 dpc. This means AP activation proceeds unchecked by an intrinsic regulator for two days after the normal point of loss. The expression of CD59 is not relevant to loss, as C6 deficient mice do not rescue.

Many mouse models of disease, including lupus nephritis, anti-phospholipid antibody dependent fetal loss, and inflammatory arthritides are complement dependent, and specifically involve neutrophils and C5a:C5aR signaling (9-11). Even models of aHUS and MPGN may be due to C5a or C5 cleavage (12, 13). We showed that embryonic loss was not neutrophil dependent, and it was previously known that C5 is not the critical mediator of loss (14). We further showed that both C3a and C5a signaling are not involved in mediating fetal loss. This means that C3b or factor B products are likely to be mediating fetal loss (15-17). This is a unique model of AP dependent tissue damage.

Future Directions for Defining the Mechanism of Crry$^{-/-}$ Embryonic Loss

While we did not find evidence of an increased inflammatory infiltrate around *Crry$^{-/-}$* embryos, we did not exhaustively rule it out. A definitive look at what cell types are present around *Crry$^{-/-}$* and *Crry$^{+/-}$* embryos with a focus on CR2, CR3 (CD11b) and CR4 (CD11c) expressing cells would rule out a small population of cells that are recruited to the embryo by receptors for iC3b and/or C3d.

Because we now know the exact days preceding fetal loss, gene expression profiling at the level of embryonic structures such as the allantois, the chorion and the trophoblastic giant cells surrounding the ectoplacental cone at 6.5 and 7.5 dpc are an attractive way to assess changes that may be happening in *Crry$^{-/-}$* embryos as a result of exposure to the AP. Despite an effort to look at the RNA of grossly dissected 7.5 dpc implantation sites (chorion and embryo, with minimal decidua), we were never able to confidently distinguish the genotype of the embryos by real time PCR. Transcript from the knock-out allele persisted and we were comparing *Crry$^{+/-}$* to *Crry$^{-/-}$* in the setting of contaminating *Crry$^{+/-}$* decidua.

Summary and Conclusions on AP Variants in PE and Lupus Nephritis

Preeclampsia is a pregnancy specific condition whose only cure is delivery of the baby and removal of the placenta. It is thought that the placenta, in response to a high pressure, low flow maternal uterine spiral artery, releases vasculopathic molecules (18). These include sFlt-1, which antagonizes VEGFA and PLGF in the kidney, leading to hypertension and proteinuria. Clinical findings of a TMA are shared between aHUS and PE (2, 19). Individuals with complement regulatory gene mutations have been noted to have PE and HELLP (20-22).

Markers of AP activation are elevated early in pregnancy in individuals that will go on to develop PE (23).

We did identify rare variants in factor H that are haploinsufficient and likely lead to increased AP activity in two large independent groups. These alleles were in ~ 1% of cases. We did not find such alleles in MCP and those we observed in factor I were in both cases and controls. The common allele, Y402H, that is enriched in MPGN and AMD was not enriched in PE cases.

We did replicate two recent meta-analyses that confirm the association between the FV Leiden mutation and PE (24, 25). Further, our findings agree with these meta-analyses in concluding that other previously reported associations are not credible. We found a significant enrichment of known disease causing clotting pathway mutations in PE cases in a large Finnish cohort, but were unable to replicate the association in a separate UK cohort.

Complement activation is one way the autoantibodies present in lupus damage tissue. Our hypothesis was that, due to the kidney's unique sensitivity to complement activation, rare variants that increase complement activation would increase the risk of lupus nephritis. In 200 lupus nephritis cases we found only one rare variant that we predict would lead to this outcome, an aHUS associated mutation in MCP that has a secretion defect (26). We found no burden of mutations in C4 binding protein, a protein with function similar to factor H, but for the classical pathway. It is our conclusion that rare variants in the complement pathway that lead to excessive activation are not a significant contributor to lupus nephritis.

Future Directions for AP and Clotting Variants in PE

We await additional targeted sequencing on six AP genes in an additional 190 PE cases from the Finnish cohort. We have a particular interest in FH and FI. Also, we are awaiting additional genotyping from these 190 cases and 475 additional Finnish controls for the rare disease linked variants in the clotting pathway. Given there is a mix of prothrombotic and pro-bleeding variants, some may be shown to not be enriched in cases with the addition of more controls.

We are exploring the possibility that the VIP study does not replicate our findings in the clotting pathway due to study design. Our focus was on the complement pathway when we originally began the collaboration. Also, we will examine both groups for previously un-recognized relatedness. We did not explicitly receive families or pedigrees and generally the geographic origin of individuals with any given clotting pathway variant was diverse. This argues against a cryptic relatedness driving these trends. Given that we only have genotyping on 750kb to 1 Mb of coding regions, relatedness measures may be inaccurate. The VIP cohort sequencing data is undergoing an additional round of quality control and a gene based burden test to determine if there are genes with an excess of variants shared between the two groups.

Summary Conclusions from Rare Variants in AMD

We identified rare variants in the AP in severe AMD cases (5). These variants are in all the genes involved in cofactor-mediated control of the AP. They range from rare in the population (1%-0.1%) to extremely rare, essentially private. Study designs required to detect these variants differ: rare variants were successfully validated in case control cohorts, while strong statistical power for individual private variants can only be found in families.

We demonstrated that power to detect these variants is highly dependent upon the gene considered, something disheartening when considering genome or exome wide efforts. We utilized the functional domains of FH and the aHUS literature to show that rare mutations in *CFH* that are known to affect secretion, cofactor activity or localization to anionic surfaces are enriched in AMD cases independent of the common variants in *CFH*.

Future Directions for AMD

We would like to understand the effects of AMD associated *CFH* mutations. Serum from rare variant carries in *CFH* is being assayed for FH levels similar to the data presented in *CFI*. We hope to use serum from rare variant carriers to directly test FH function. Additionally, rare variant carriers in *CFH* and *CFI* are being extended into pedigrees where possible.

There is a rare variant in *CFB* (I242L) that is highly enriched in cases as well, 19 cases and 1 control from our previously reported study. This variant was reported in aHUS but no one has investigated its effect on function (27). It resides in the loop factor D cleaves when activating FB to Bb. It could affect the efficiency by which factor D activates factor B. We are seeking to confirm its association in exome chip data (MAF 0.13% in NHLBI ESP EA population) and also identify families that carry it (~600 families available for evaluation).

Final Remarks

While we did not find AP variants in a substantial fraction of PE cases as our initial pilot study of 40 PROMISSE PE patients suggested, looking back at that in light of the targeted sequencing of large cohorts we performed demonstrates several important points. First, the interpretation of aHUS association was made based on the aHUS literature, which often used 100-200 controls from each center. This size control population is insufficient to truly estimate the frequency of variants in the 1% to 0.1% range. This is of great importance in all diseases as public datasets of disease-associated variants are being compiled. These must be constantly updated and curated as new variants are found and old variants are discredited. These datasets are currently of low quality with ambiguous annotations for the credibility of a deposition.

Second, population stratification in the context of rare variant analysis is a potent source of false positives. Rare variants show larger fluctuations in frequency between populations then do common variants and therefore small differences in case and control populations can lead to false positives.

It is clear from this work that the evolutionary constraint present on two different genes can vary widely and even within a gene it can vary significantly between domains. The association detected between rare variants in *CFI* and AMD was well powered because there is no missense variant present at >0.5% in *CFI* in the NHLBI ESP European American (EA) population. Similarly, a recent report that reported an association between *PLD3* and Alzheimer's was well powered on a population level because *PLD3* also has no missense variants present at >0.5% in the EA population (28). This means there was very little noise when considering the effect at the gene level, even if variants were clustered in certain areas of the protein.

Our efforts to use frequency sensitive variations of the general burden approach to detect an association between rare variants in *CFH* and AMD were not better than a simple burden test with a reasonable MAF cutoff. However, when we looked at those variants that were very rare in controls and in functional domains of FH, we saw a strong enrichment. This highlights the potential for increasingly sophisticated statistical tests. One approach might use curated lists of domains critical for protein function. Another approach would be to utilize a sliding window along a transcript with weights based on the frequency and diversity of mutations present in a large control population such as the NHLBI ESP. This would give more weight to areas such as the regulatory domain of FH, where there are generally few mutations per/100bp and those that are observed are often singletons or doubletons in the ESP.

There is significant pleiotropy present in variants that increase AP activation. Variants seen in aHUS, R53C and R1210C of *CFH*, can be nearly 100% penetrant for AMD (29). There are likely additional genetic factors responsible for increasing the risk of they have for kidney diseases. The difference in penetrance for the two diseases may also be due to the fact that visual problems are much easier to recognize than kidney disease.

I believe that in the degenerative disease AMD these variants will have high penetrance because they lead to a cumulative excess of AP activation over decades, and at a higher order of magnitude compared to common variants. This is of paramount importance for these families and has implications for their treatment. Such a scenario may play out in other common degenerative diseases as well.

It would be fortuitous if new risk variants in novel loci are already known to cause

disease, even if clinically unrelated. This is unlikely to be the case. Once a toehold is established

in the genetics of a disease, incorporating the review of experts in the structure/function

relationships of these genes will be essential to incorporating genomic information into

personalized medical care.

References

1   Kavanagh, D., and T. H. Goodship. 2011. Atypical hemolytic uremic syndrome, genetic basis, and clinical manifestations. *Hematology Am Soc Hematol Educ Program* 2011:15-20.
2   Taylor, C. M., C. Chua, J. J. Howie, and R. Risdon. 2004. Clinico-pathological findings in diarrhoea-negative haemolytic uraemic syndrome. *Pediatr. Nephrol.* 19:419-425.
3   Java, A., J. Atkinson, and J. Salmon. 2013. Defective complement inhibitory function predisposes to renal disease. *Annu. Rev. Med.* 64:307-324.
4   Fakhouri, F., V. Fremeaux-Bacchi, L. H. Noel, H. T. Cook, and M. C. Pickering. 2010. C3 glomerulopathy: a new classification. *Nature reviews. Nephrology* 6:494-499.
5   Seddon, J. M., Y. Yu, E. C. Miller, R. Reynolds, P. L. Tan, S. Gowrisankar, J. I. Goldstein, M. Triebwasser, H. E. Anderson, J. Zerbib, D. Kavanagh, E. Souied, N. Katsanis, M. J. Daly, J. P. Atkinson, and S. Raychaudhuri. 2013. Rare variants in CFI, C3 and C9 are associated with high risk of advanced age-related macular degeneration. *Nat. Genet.* 45:1366-1370.
6   Hageman, G. S., D. H. Anderson, L. V. Johnson, L. S. Hancox, A. J. Taiber, L. I. Hardisty, J. L. Hageman, H. A. Stockman, J. D. Borchardt, K. M. Gehrs, R. J. Smith, G. Silvestri, S. R. Russell, C. C. Klaver, I. Barbazetto, S. Chang, L. A. Yannuzzi, G. R. Barile, J. C. Merriam, R. T. Smith, A. K. Olsh, J. Bergeron, J. Zernant, J. E. Merriam, B. Gold, M. Dean, and R. Allikmets. 2005. A common haplotype in the complement regulatory gene factor H (HF1/CFH) predisposes individuals to age-related macular degeneration. *Proc. Natl. Acad. Sci. USA* 102:7227-7232.
7   Yu, Y., R. Reynolds, B. Rosner, M. J. Daly, and J. M. Seddon. 2012. Prospective assessment of genetic effects on progression to different stages of age-related macular degeneration using multistate Markov models. *Invest. Ophthalmol. Vis. Sci.* 53:1548-1556.

8   Mao, D., X. Wu, C. Deppong, L. D. Friend, G. Dolecki, D. M. Nelson, and H. Molina. 2003. Negligible role of antibodies and C5 in pregnancy loss associated exclusively with C3-dependent mechanisms through complement alternative pathway. *Immunity* 19:813-822.

9   Bao, L., I. Osawe, T. Puri, J. D. Lambris, M. Haas, and R. J. Quigg. 2005. C5a promotes development of experimental lupus nephritis which can be blocked with a specific receptor antagonist. *Eur. J. Immunol.* 35:2496-2506.

10   Redecha, P., R. Tilley, M. Tencati, J. E. Salmon, D. Kirchhofer, N. Mackman, and G. Girardi. 2007. Tissue factor: a link between C5a and neutrophil activation in antiphospholipid antibody induced fetal injury. *Blood* 110:2423-2431.

11   Tsuboi, N., T. Ernandez, X. Li, H. Nishi, X. Cullere, D. Mekala, M. Hazen, J. Kohl, D. M. Lee, and T. N. Mayadas. 2011. Regulation of human neutrophil Fcgamma receptor IIa by C5a receptor promotes inflammatory arthritis in mice. *Arthritis Rheum.* 63:467-478.

12   Pickering, M. C., J. Warren, K. L. Rose, F. Carlucci, Y. Wang, M. J. Walport, H. T. Cook, and M. Botto. 2006. Prevention of C5 activation ameliorates spontaneous and experimental glomerulonephritis in factor H-deficient mice. *Proc. Natl. Acad. Sci. USA* 103:9649-9654.

13   de Jorge, E. G., P. Macor, D. Paixao-Cavalcante, K. L. Rose, F. Tedesco, H. T. Cook, M. Botto, and M. C. Pickering. 2011. The development of atypical hemolytic uremic syndrome depends on complement C5. *J. Am. Soc. Nephrol.* 22:137-145.

14   Ruseva, M. M., T. R. Hughes, R. M. Donev, B. Sivasankar, M. C. Pickering, X. Wu, C. L. Harris, and B. P. Morgan. 2009. Crry deficiency in complement sufficient mice: C3 consumption occurs without associated renal injury. *Mol. Immunol.* 46:803-811.

15   Nicholson-Weller, A., and J. A. Halperin. 1993. Membrane signaling by complement C5b-9, the membrane attack complex. *Immunol. Res.* 12:244-257.

16   Fallman, M., R. Andersson, and T. Andersson. 1993. Signaling properties of CR3 (CD11b/CD18) and CR1 (CD35) in relation to phagocytosis of complement-opsonized particles. *J. Immunol.* 151:330-338.

17   Uwai, M., Y. Terui, Y. Mishima, H. Tomizuka, M. Ikeda, T. Itoh, M. Mori, M. Ueda, R. Inoue, M. Yamada, H. Hayasawa, T. Horiuchi, Y. Niho, M. Matsumoto, Y. Ishizaka, K. Ikeda, K. Ozawa, and K. Hatake. 2000. A new apoptotic pathway for the complement factor B-derived fragment Bb. *J. Cell. Physiol.* 185:280-292.

18   Levine, R. J., S. E. Maynard, C. Qian, K. H. Lim, L. J. England, K. F. Yu, E. F. Schisterman, R. Thadhani, B. P. Sachs, F. H. Epstein, B. M. Sibai, V. P. Sukhatme, and S. A. Karumanchi. 2004. Circulating angiogenic factors and the risk of preeclampsia. *N. Engl. J. Med.* 350:672-683.

19   Stillman, I. E., and S. A. Karumanchi. 2007. The glomerular injury of preeclampsia. *J. Am. Soc. Nephrol.* 18:2281-2284.

20   Fakhouri, F., L. Roumenina, F. Provot, M. Sallee, S. Caillard, L. Couzi, M. Essig, D. Ribes, M. A. Dragon-Durey, F. Bridoux, E. Rondeau, and V. Fremeaux-Bacchi. 2010. Pregnancy-associated hemolytic uremic syndrome revisited in the era of complement gene mutations. *J. Am. Soc. Nephrol.* 21:859-867.

21   Fakhouri, F., M. Jablonski, J. Lepercq, J. Blouin, A. Benachi, M. Hourmant, Y. Pirson, A. Durrbach, J. P. Grunfeld, B. Knebelmann, and V. Fremeaux-Bacchi. 2008. Factor H,

membrane cofactor protein, and factor I mutations in patients with hemolysis, elevated liver enzymes, and low platelet count syndrome. *Blood* 112:4542-4545.

22   Crovetto, F., N. Borsa, B. Acaia, C. Nishimura, K. Frees, R. J. Smith, F. Peyvandi, R. Palla, M. Cugno, S. Tedeschi, P. Castorina, E. Somigliana, G. Ardissino, and L. Fedele. 2012. The genetics of the alternative pathway of complement in the pathogenesis of HELLP syndrome. *J Matern Fetal Neonatal Med* 25:2322-2325.

23   Lynch, A. M., J. R. Murphy, T. Byers, R. S. Gibbs, M. C. Neville, P. C. Giclas, J. E. Salmon, and V. M. Holers. 2008. Alternative complement pathway activation fragment Bb in early pregnancy as a predictor of preeclampsia. *Am. J. Obstet. Gynecol.* 198:385 e381-389.

24   Staines-Urias, E., M. C. Paez, P. Doyle, F. Dudbridge, N. C. Serrano, J. P. Ioannidis, B. J. Keating, A. D. Hingorani, and J. P. Casas. 2012. Genetic association studies in pre-eclampsia: systematic meta-analyses and field synopsis. *Int. J. Epidemiol.* 41:1764-1775.

25   Buurma, A. J., R. J. Turner, J. H. Driessen, A. L. Mooyaart, J. W. Schoones, J. A. Bruijn, K. W. Bloemenkamp, O. M. Dekkers, and H. J. Baelde. 2013. Genetic variants in pre-eclampsia: a meta-analysis. *Hum. Reprod. Update* 19:289-303.

26   Fremeaux-Bacchi, V., E. A. Moulton, D. Kavanagh, M.-A. Dragon-Durey, J. Blouin, A. Caudy, N. Arzouk, R. Cleper, M. Francois, G. Guest, J. Pourrat, R. Seligman, W. H. Fridman, C. Loirat, and J. P. Atkinson. 2006. Genetic and functional analyses of membrane cofactor protein (CD46) mutations in atypical hemolytic uremic syndrome. *J. Am. Soc. Nephrol.* 17:2017-2025.

27   Maga, T. K., C. J. Nishimura, A. E. Weaver, K. L. Frees, and R. J. Smith. 2010. Mutations in alternative pathway complement proteins in American patients with atypical hemolytic uremic syndrome. *Hum. Mutat.* 31:E1445-1460.

28   Cruchaga, C., C. M. Karch, S. C. Jin, B. A. Benitez, Y. Cai, R. Guerreiro, O. Harari, J. Norton, J. Budde, S. Bertelsen, A. T. Jeng, B. Cooper, T. Skorupa, D. Carrell, D. Levitch, S. Hsu, J. Choi, M. Ryten, J. Hardy, M. Ryten, D. Trabzuni, M. E. Weale, A. Ramasamy, C. Smith, C. Sassi, J. Bras, J. R. Gibbs, D. G. Hernandez, M. K. Lupton, J. Powell, P. Forabosco, P. G. Ridge, C. D. Corcoran, J. T. Tschanz, M. C. Norton, R. G. Munger, C. Schmutz, M. Leary, F. Y. Demirci, M. N. Bamne, X. Wang, O. L. Lopez, M. Ganguli, C. Medway, J. Turton, J. Lord, A. Braae, I. Barber, K. Brown, P. Passmore, D. Craig, J. Johnston, B. McGuinness, S. Todd, R. Heun, H. Kolsch, P. G. Kehoe, N. M. Hooper, E. R. Vardy, D. M. Mann, S. Pickering-Brown, K. Brown, N. Kalsheker, J. Lowe, K. Morgan, A. David Smith, G. Wilcock, D. Warden, C. Holmes, P. Pastor, O. Lorenzo-Betancor, Z. Brkanac, E. Scott, E. Topol, K. Morgan, E. Rogaeva, A. B. Singleton, J. Hardy, M. I. Kamboh, P. St George-Hyslop, N. Cairns, J. C. Morris, J. S. Kauwe, and A. M. Goate. 2013. Rare coding variants in the phospholipase D3 gene confer risk for Alzheimer's disease. *Nature*.

29   Raychaudhuri, S., O. Iartchouk, K. Chin, P. L. Tan, A. K. Tai, S. Ripke, S. Gowrisankar, S. Vemuri, K. Montgomery, Y. Yu, R. Reynolds, D. J. Zack, B. Campochiaro, P. Campochiaro, N. Katsanis, M. J. Daly, and J. M. Seddon. 2011. A rare penetrant mutation in CFH confers high risk of age-related macular degeneration. *Nat. Genet.* 43:1232-1236.

Supplemental Manuscript #1

Seddon, et. al. 2013.

Nature Genetics

Rare variants in *CFI*, *C3* and *C9* are associated with high risk of advanced age-related macular degeneration

Johanna M Seddon[1-3], Yi Yu[1], Elizabeth C Miller[4], Robyn Reynolds[1], Perciliz L Tan[5-7], Sivakumar Gowrisankar[8], Jacqueline I Goldstein[9,10], Michael Triebwasser[4], Holly E Anderson[11], Jennyfer Zerbib[12], David Kavanagh[11], Eric Souied[12], Nicholas Katsanis[5-7], Mark J Daly[9,10], John P Atkinson[4] & Soumya Raychaudhuri[8,9,13-15]

To define the role of rare variants in advanced age-related macular degeneration (AMD) risk, we sequenced the exons of 681 genes within all reported AMD loci and related pathways in 2,493 cases and controls. We first tested each gene for increased or decreased burden of rare variants in cases compared to controls. We found that 7.8% of AMD cases compared to 2.3% of controls are carriers of rare missense *CFI* variants (odds ratio (OR) = 3.6; $P = 2 \times 10^{-8}$). There was a predominance of dysfunctional variants in cases compared to controls. We then tested individual variants for association with disease. We observed significant association with rare missense alleles in genes other than *CFI*. Genotyping in 5,115 independent samples confirmed associations with AMD of an allele in *C3* encoding p.Lys155Gln (replication $P = 3.5 \times 10^{-5}$, OR = 2.8; joint $P = 5.2 \times 10^{-9}$, OR = 3.8) and an allele in *C9* encoding p.Pro167Ser (replication $P = 2.4 \times 10^{-5}$, OR = 2.2; joint $P = 6.5 \times 10^{-7}$, OR = 2.2). Finally, we show that the allele of *C3* encoding Gln155 results in resistance to proteolytic inactivation by CFH and CFI. These results implicate loss of C3 protein regulation and excessive alternative complement activation in AMD pathogenesis, thus informing both the direction of effect and mechanistic underpinnings of this disorder.

AMD leads to irreversible vision loss in individuals over 55 years of age[1-3]. Most known genetic AMD variants are common, without established disease mechanisms[4]. So far, highly penetrant missense variants for AMD are only known in *CFH*[5]. We sought to identify additional rare missense risk variants with the goal of identifying specific disease mechanisms.

We applied targeted next-generation sequencing to coding regions of candidate genes within AMD-associated loci and pathways (Online Methods and **Supplementary Fig. 1**). Informed consent was obtained from each subject, and the institutional review board at each clinical site approved the study. To maximize power, we prioritized cases with severe disease, familial disease or early age of onset, as well as older controls without drusen. We sequenced 681 genes to >20× coverage for a median of 95.6% of the targeted region (**Supplementary Fig. 2**). After calling variants with the Genome Analysis Toolkit (GATK)[6] and applying strict quality control, we obtained genotype data for 1,676 cases, 745 controls and 36 siblings with discordant disease status ($n = 2,493$). Our data had <0.92% missing genotypes for each allele and <0.54% missing genotypes for each individual. We observed effect sizes comparable to published values for known AMD risk variants or proxies (**Supplementary Table 1**). Many variants discovered in our sequencing study were not genotyped or imputed in recent AMD genome-wide association studies (GWAS) (**Supplementary Fig. 3**).

We compared sequencing-based genotypes to genotypes from the HumanExome BeadChip array (exome chip)[7,8]. We observed 99.97% concordance for 2,426 missense SNPs that were captured by both sequencing and exome chip with allele frequencies (f) ≥0.001 (Online Methods). Allelic dosages were almost perfectly correlated ($r > 0.99$) for 96.5% of common variants ($f > 0.01$) and 93.0% of rare variants (**Supplementary Fig. 4**); only 0.2% and 3.6% of common and rare SNPs, respectively, had modestly correlated dosages ($r < 0.9$).

With the sequencing data, we tested whether any of 681 genes had a higher burden of rare variants in cases or in controls. We selected only rare variants ($f < 0.01$) within targeted genes that altered coding sequences (missense, nonsense, read-through or splice variants; $n = 15,762$ SNPs). We used a simple burden test to assess whether the proportion of case individuals who carried at least one rare variant was in excess of the proportion expected by chance (Online Methods and

[1]Ophthalmic Epidemiology and Genetics Service, New England Eye Center, Tufts Medical Center, Boston, Massachusetts, USA. [2]Department of Ophthalmology, Tufts University School of Medicine, Boston, Massachusetts, USA. [3]Sackler School of Graduate Biomedical Sciences, Tufts University, Boston, Massachusetts, USA. [4]Division of Rheumatology, Department of Medicine, Washington University School of Medicine, St. Louis, Missouri, USA. [5]Center for Human Disease Modeling, Duke University, Durham, North Carolina, USA. [6]Department of Cell Biology, Duke University, Durham, North Carolina, USA. [7]Department of Pediatrics, Duke University, Durham, North Carolina, USA. [8]Partners HealthCare Center for Personalized Genetic Medicine, Boston, Massachusetts, USA. [9]Program in Medical and Population Genetics, Broad Institute, Cambridge, Massachusetts, USA. [10]Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, Massachusetts, USA. [11]Institute of Genetic Medicine, Newcastle University, International Centre for Life, Newcastle-upon-Tyne, UK. [12]Department of Ophthalmology, Hôpital Intercommunal de Créteil, Hôpital Henri Mondor, Université Paris Est Créteil, Créteil, France. [13]Division of Genetics, Brigham and Women's Hospital, Boston, Massachusetts, USA. [14]Division of Rheumatology, Immunology and Allergy, Brigham and Women's Hospital, Boston, Massachusetts, USA. [15]Faculty of Medical and Human Sciences, University of Manchester, Manchester, UK. Correspondence should be addressed to J.M.S. (jseddon@tuftsmedicalcenter.org) or S.R. (soumya@broadinstitute.org).
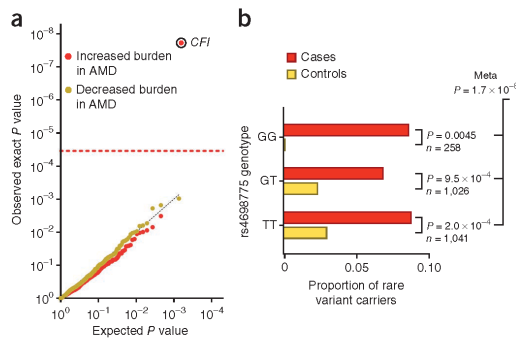
[157]

Figure 1 *CFI* burden of rare coding variants is greater in cases. (**a**) We tested 681 sequenced genes for increased burden of rare (*f* < 0.01) missense, nonsense, splice or read-through variants. Here we plot the observed *P* value for each gene as a function of the expected *P* value. We tested enrichment for carriers of rare variants in cases (red) and controls (yellow) separately. We indicate the threshold of significance after multiple-hypothesis testing (dashed red line, $P < 3.6 \times 10^{-5} = 0.05/(681 \text{ genes} \times 2 \text{ tests})$). We observed significant enrichment of rare variants in cases in only one gene, *CFI*. (**b**) We stratified individuals by their genotype at rs4698775, an associated common variant near *CFI*; the G minor allele has previously been reported to be associated with risk. We observed that rare variants are enriched on all genotypic backgrounds in cases compared to controls.

Supplementary Table 2a)[9]. We similarly tested increased burden in controls. To interpret statistical significance, we applied a Bonferroni-corrected significance threshold of $P < 3.6 \times 10^{-5}$ (0.05/(681 genes × 2 tests)). We saw a significantly increased burden in cases in only one gene, *CFI* (exact one-tailed $P = 1.6 \times 10^{-8}$; OR = 3.57; **Fig. 1a**); 7.8% of cases had a rare coding variant compared to 2.3% of controls. In contrast, using the C-alpha test[10], we found no evidence of any genes with alleles acting in opposite directions (**Supplementary Table 2a**).

The enrichment of rare coding variants in *CFI* is likely not due to confounding by population stratification, sequencing artifact, age or sex. First, if the enrichment was due to stratification, then other genes would also demonstrate enrichment for rare variants; however, burden tests for other genes did not exceed expectation due to chance (**Fig. 1a**). Second, stratification would also drive enrichment of synonymous variants; however, we observed no significant case enrichment for rare synonymous variants in *CFI* ($P = 0.89$) or in any of the 681 genes ($P > 0.001$; **Supplementary Tables 2a** and **3**). Third, to adjust for case-control stratification, we calculated principal components on the basis of ancestry-informative markers in a subset of 1,558 cases and 683 controls with exome-chip genome-wide genotyping data. We found that carriers of rare variants in *CFI* did not cluster along the top two principal components (**Supplementary Fig. 5**). Furthermore, adjusting for the top five principal components in a logistic regression framework caused no change in the case enrichment of rare *CFI* coding variants ($P = 5.0 \times 10^{-8}$, OR = 3.6 versus $P = 1.1 \times 10^{-7}$, $OR = 3.5$ after adjustment). To rule out effects from sequencing artifacts, we confirmed 137 of the 140 rare variant events assayed with Sanger sequencing (Online Methods and **Supplementary Fig. 6**). The observed enrichment remained significant after adjusting for age and sex ($P = 6.7 \times 10^{-9}$, OR = 3.7 versus $P = 1.1 \times 10^{-8}$, OR = 3.6 after adjustment). We note that, because extreme phenotypic samples were prioritized for sequencing, independent replication in prospectively ascertained samples will ultimately be necessary to accurately estimate relative enrichment of rare *CFI* variants within the general population.

The enrichment in cases of rare coding variants in *CFI* was independent of the nearby common risk allele rs4698775 (refs. 4,11). When we stratified our samples for rs4698775 genotypes, we observed consistent enrichment of rare coding variants in *CFI* across all three rs4698775 genotypes ($P = 1.7 \times 10^{-8}$; **Fig. 1b**). Moreover, the risk conferred by the common rs4698775 allele was independent of carrier status for rare variants ($P < 0.04$; allelic OR = 1.15), arguing against a 'synthetic association'.

We then used PolyPhen-2 (ref. 12) to classify the 59 separate *CFI* nucleotide variants into 4 categories on the basis of their potential predicted impact: (i) loss of function (nonsense or splice variants)[13], (ii) probably damaging, (iii) possibly damaging and (iv) benign (**Fig. 2a** and **Supplementary Table 4**). The distribution of variants predicted to result in loss of protein function (three nonsense variants and one splice variant) was the most skewed, and these variants were exclusively found in cases (seven cases compared to zero controls; **Fig. 2b**). The distribution of probably damaging variants was more skewed toward cases (41 cases compared to 3 controls) than possibly damaging variants (16 cases compared to 3 controls) and benign variants (70 cases compared to 12 controls). Indeed, probably damaging or loss-of-function variants conferred high risk of disease ($P = 1.3 \times 10^{-5}$; OR = 7.5). Further, some of the variants classified as benign are potentially pathogenic and are inaccurately classified by PolyPhen-2 (ref. 14).

CFI encodes complement factor I, which consists of a heavy chain and a light chain, with a catalytic serine protease domain that cleaves the C3 protein. There was a greater burden of rare variants that affected the catalytic light chain ($P = 2.2 \times 10^{-6}$; OR = 4.85) than affected the heavy chain ($P = 0.0012$; OR = 2.63). We observed no statistically significant evidence of association with AMD for any of the individual *CFI* variants ($P > 0.02$), possibly owing to their low frequency.

Notably, many of these rare variants in *CFI* have been found in individuals with atypical hemolytic uremic syndrome (aHUS), including ones encoding p.Pro50Ala[15], p.Gly119Arg[15], p.Ala240Gly[16], p.Gly261Asp[17], p.Arg317Trp[16], p.Ile340Thr[18], p.Tyr369Ser[16,19], p.Asp403Asn[15], p.Ile416Leu[19], p.Tyr459Ser[15], p.Arg474* (ref. 20) and p.Pro553Ser[15] changes. These variants likely cause aHUS by resulting in haploinsufficiency, thus mitigating the ability of CFI to cleave and thereby deactivate C3b. The p.Arg317Trp and p.Ile340Thr variants result in CFI functional deficiency[21], whereas p.Pro50Ala, p.Ile416Leu and p.Arg474* produce a quantitative deficiency in CFI protein[15]. Despite extensive study, no functional deficiency for the p.Gly261Asp variant has been identified[17].

We also tested rare variants individually for association with AMD (**Supplementary Fig. 1**). This study was not powered to detect association for variants observed fewer than five times; thus, we excluded extremely rare SNPs from this analysis. We tested 1,824 variants within the 681 sequenced genes that had an allele frequency of <1% in cases or controls (Online Methods and **Supplementary Table 2b**). To interpret statistical significance, we applied a Bonferroni-corrected significance threshold of $P < 1.4 \times 10^{-5}$ (0.05/(1,824 variants × 2 tests)).

We identified 5 risk and 15 protective variants that demonstrated nominal evidence of association in the discovery phase (exact one-tailed $P < 0.01$; **Supplementary Table 5**). Four of these variants were within or near *CFH*, including the previously reported *CFH* risk variant encoding p.Arg1210Cys ($P = 0.0012$)[5]. In addition, we observed association with the *CFH* variant encoding p.Asn1050Tyr, the *CFHR2* variant encoding p.Tyr264Cys and the *CFHR5* variant encoding p.Gly278Ser, which are protective. After phasing common variants within the *CFH* locus, we observed that the rare associated variants were in linkage disequilibrium (LD) with *CFH* haplotypes (**Supplementary Fig. 7**). We were uncertain whether these associations
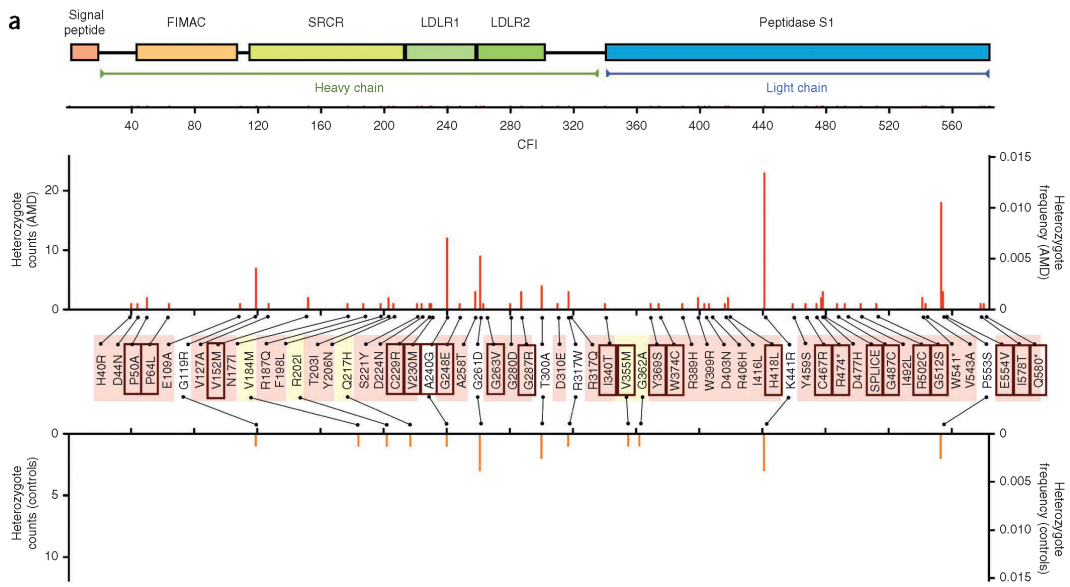
Figure 2 Burden of *CFI* rare variants. (**a**) Top, domain structure of the CFI protein, including the factor I membrane attack complex (FIMAC), the scavenger receptor cysteine-rich (SRCR) domain, the low-density lipoprotein (LDL) receptor (LDLR1 and LDLR2) domains and a peptidase S1 serine protease domain. Below, individual rare variants that we observed in our sequencing experiment in cases (red, upper) and controls (yellow, lower). In our data, we observed no common *CFI* coding variants with $f > 0.01$. We plot counts of variants along the left *y* axes and the proportions of heterozygote individuals along the right *y* axes. We list the actual coding changes conferred by the variants in the middle. Case-only variants are colored red, and control-only variants are colored yellow. We placed red boxes around variants that are predicted to be loss of function or probably damaging. (**b**) Plot of rare variants in *CFI* categorized with PolyPhen-2 as benign, possibly damaging, probably damaging and loss of function, with counts in controls (left) compared to counts in cases (right) for each category. Variants that are predicted to have greater likelihood of being damaging are more strongly skewed toward cases (one-tailed $P = 0.045$, logistic regression with permutation of only individuals with a rare variant).
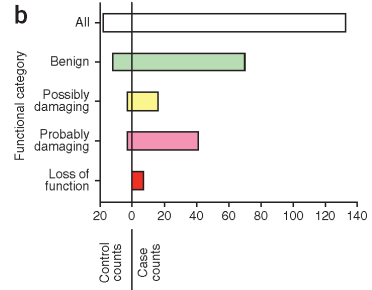
represented unique effects or were tagging *CFH* haplotypes with large effects, driven by the common risk allele at rs10737680 and by the common allele encoding p.Tyr402His[22–24].

We evaluated 11 of the 16 variants outside of the *CFH* locus in an independent set of 2,227 cases and 2,888 controls from Boston, Baltimore and France genotyped by exome chip or TaqMan assays

Table 1 *C3* p.Lys155Gln (rs147859257) association with AMD risk

| | Case-control | | | | | Discordant sibling pairs | | | p.Lys155Gln association | |
| | p.Lys155Gln heterozygotes | | Lys155 homozygotes | | Allelic OR | | | | One-tailed *P* value | |
Cohort	Unaffected	AMD	Unaffected	AMD		Concordant genotypes	p.Lys155Gln in affected sibling	p.Lys155Gln in unaffected sibling	No conditioning	Conditioning on rs2230199
Boston (sequencing, discovery)	1	40	744	1,636	18.0 (2.5–130.9)	36	0	0	4.8×10^{-6}	9.6×10^{-6}
Boston (replication)	19	20	2,014	726	2.9 (1.5–5.4)	13	0	0	9.3×10^{-4}	3.6×10^{-4}
France	4	18	676	935	3.2 (1.1–9.6)	–	–	–	0.018	0.011
Baltimore	3	17	163	467	2.0 (0.6–6.7)	–	–	–	0.24	–
All replication	26	55	2,853	2,128	2.8 (1.7–4.6)	13	0	0	3.5×10^{-5}	9.2×10^{-6}
Joint	27	95	3,597	3,764	3.8 (2.3–6.1)	49	0	0	5.2×10^{-9}	6.2×10^{-10}

Each row represents association statistics for the sample collections. The first row (Boston) represents our discovery data set and lists genotype counts in the case-control samples and in 49 discordant sibling pairs with exact one-tailed association *P* values. We list one-tailed *P* values with and without conditioning for the common variant rs2230199 in *C3* encoding p.Arg102Gly. Other rows list similar statistics for each of the replication cohorts. The last row represents the joint analysis of all samples. Because we initially tested a total of 1,824 variants for risk and for protection, the multiple-hypothesis-corrected significance threshold for joint analysis is $P < 1.4 \times 10^{-5}$.

[159]

Table 2 *C9* p.Pro167Ser (rs34882957) association with AMD risk

| | Case-control | | | | | Discordant sibling pairs | | | p.Pro167Ser association |
| | p.Pro167Ser heterozygotes | | p.Pro167Ser homozygotes | | Allelic OR | | | | One-tailed *P* value |
Cohort	Unaffected	AMD	Unaffected	AMD		Concordant genotypes	p.Pro167Ser in affected sibling	p.Pro167Ser in unaffected sibling	Unconditional
Boston (sequencing, discovery)	11	53	734	1,623	2.2 (1.1–4.1)	35	1	0	0.0068
Boston (replication)	39	29	1,994	717	2.0 (1.3–3.3)	12	1	0	0.0023
France	9	33	671	919	2.6 (1.3–5.6)	–	–	–	0.0045
Baltimore	4	21	127	438	1.7 (0.6–5.1)	–	–	–	0.22
All replication	51	79	2,828	2,128	2.2 (1.5–3.2)	12	0	0	2.4×10^{-5}
Joint	63	136	3,546	3,697	2.2 (1.6–3.0)	47	2	0	6.5×10^{-7}

As in **Table 1**, each row represents association statistics for the sample collections. The last row represents the joint analysis of all samples. The multiple-hypothesis-corrected significance threshold for joint analysis is $P < 1.4 \times 10^{-5}$.

(**Supplementary Table 6**). To increase power, we incorporated shared controls in this replication study (Online Methods); these controls may result in decreased allelic effect sizes because some may ultimately develop advanced AMD. We observed independent evidence of association for two variants in replication ($P < 0.0031 = 0.05/16$) and joint ($P < 1.4 \times 10^{-5}$) analysis: the rs147859257 (p.Lys155Gln) variant in *C3* (exact one-tailed $P = 4.8 \times 10^{-6}$ in the discovery phase) and the rs34882957 (p.Pro167Ser) variant in *C9* (exact one-tailed $P = 2.3 \times 10^{-3}$ in the discovery phase; **Tables 1** and **2** and **Supplementary Table 5**). Both variants had perfectly concordant sequencing-based and exome-chip genotypes ($r = 1$). To assess whether either association could be related to population stratification, we calculated principal components for a subset of samples ($n = 2,241$) as described above. Individuals who were carriers of either variant did not cluster along the first two principal components (**Supplementary Figs. 8** and **9**). Also, adjusting for the top five principal components caused no change in the effect mediated by the p.Lys155Gln variant ($P = 2.7 \times 10^{-5}$, OR = 16.1 versus $P = 2.0 \times 10^{-6}$, OR = 16.7 after adjustment) or in the effect mediated by the p.Pro167Ser variant ($P = 0.053$, OR = 2.0 versus $P = 0.053$, OR = 2.0 after adjustment). These associations retained significance after adjusting for age and sex ($P = 1.5 \times 10^{-5}$, OR = 9.5 versus $P = 1.5 \times 10^{-5}$, OR = 9.0 for p.Lys155Gln after adjustment; $P = 0.007$, OR = 2.3 versus $P = 0.01$, OR = 2.2 for p.Pro167Ser after adjustment).

The *C3* variant encoding p.Lys155Gln (position 133 excluding the signal peptide) demonstrated compelling evidence of association in replication analysis ($P = 3.5 \times 10^{-5}$; OR = 2.8) and was highly significantly associated with a large effect size in joint analysis with discovery samples ($P = 5.2 \times 10^{-9}$; OR = 3.8; **Table 1**). Whereas the estimated effects for p.Lys155Gln in the discovery and replication phases had overlapping confidence intervals, the effect in the discovery phase was larger, perhaps related to 'winner's curse' (ref. 25) or inclusion of shared controls in the replication phase. This risk was independent from that explained by a previously reported common risk variant, p.Arg102Gly (rs2230199)[26,27]. In a subset of samples with genotypes for p.Arg102Gly and p.Lys155Gln ($n = 6,935$, excluding Baltimore samples), we observed increased significance of the effect mediated by p.Lys155Gln when we stratified individuals on the basis of their genotypes for p.Arg102Gly (from $P = 5.4 \times 10^{-9}$ to 6.2×10^{-10}; **Table 1**). The risk variant encoding p.Lys155Gln was in phase with the protective allele rs2230199[G] encoding Arg102. The *C9* variant encoding p.Pro167Ser also demonstrated convincing association in replication analysis ($P = 2.4 \times 10^{-5}$; OR = 2.2) and in joint analysis ($P = 6.5 \times 10^{-7}$; OR = 2.2; **Table 2**). Whereas the

p.Pro167Ser variant is predicted with high confidence to be a probably damaging variant by PolyPhen-2 (ref. 12), there were no reported functional data for this variant. We note that one other independent nominal association has previously been reported in *C9* (ref. 28). Our study expands the repertoire of AMD-associated genes in the complement cascade, specifically implicating the membrane attack complex (C5-C9) that is formed downstream of C3 activation in the activation of the alternative complement pathway.

We note that p.Lys155Gln is on the surface of C3, close to the binding site for CFH (**Supplementary Fig. 10**)[29], which is a cofactor for CFI-mediated cleavage of C3. We therefore tested the hypothesis that the p.Lys155Gln substitution in C3 might confer resistance to its inactivation by CFI, using a previously described experimental strategy for evaluating mutations in aHUS[30]. Notably, we observed significantly reduced cleavage of C3 with Gln155 compared to C3 with Lys155 in fluid-phase cofactor assays (**Fig. 3a,b**). We also observed significantly reduced binding to CFH by the protein with Gln155 in surface plasmon resonance (~60%; $P < 0.03$; **Supplementary Fig. 11a**) and a trend toward reduced binding with ELISA (**Supplementary Fig. 11b**). In contrast, MCP cofactor activity was not reduced by the p.Lys155Gln substitution (**Supplementary Fig. 12a–c**).
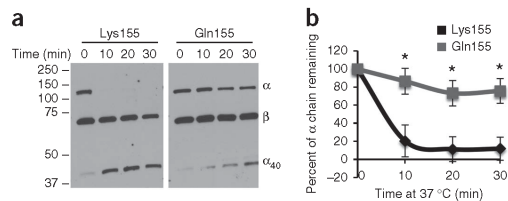

Figure 3 C3 Gln155 confers resistance to cofactor activity. (**a**) We incubated C3 proteins with Lys155 (wild type) or Gln155 (variant) at physiological salt concentration with 20 ng of CFI and 200 ng of CFH at 37 °C for 10, 20 and 30 min and then stopped the reaction by the addition of 3× reducing sample buffer. After electrophoresis and transfer to a nitrocellulose membrane, protein blots were performed for C3 proteins. This protein blot is representative of four similar experiments. We observe that the α chain is cleaved much less efficiency with Gln155 than with Lys155. In parallel, we see accumulation of the α_{40} cleavage fragment in proportion to α-chain cleavage. The α_1 cleavage fragment is not visualized because it migrates at the same molecular weight as the β chain. (**b**) We quantified the reduction in the amount of the α chain by densitometry of four independent experiments; here we present the mean ± s.e.m. at each time point. We assessed significance with an unpaired *t* test: *$P < 0.03$.

[160]

We conclude that the variant in *C3* encoding p.Lys155Gln primarily impairs C3b regulation by CFI with bound CFH, resulting in increased C3 convertase formation and feedback amplification of the alternative complement pathway (**Supplementary Fig. 13**). This pattern of reduced regulatory activity mirrors that of the highly penetrant p.Arg1210Cys substitution in CFH, which also disrupts CFH binding to C3b[31–33]. The burden of *CFI* mutations, particularly of the loss-of-function mutations affecting the catalytic domain of the protein, along with the large effect of the p.Lys155Gln and p.Arg1210Cys variants underscores the critical role of impaired alternative pathway regulation by dysfunction of the C3b-CFH-CFI trimolecular interaction. These observations parallel mounting evidence that common AMD risk alleles might also lead to increased alternative pathway activity[34–36]. These results suggest the potential of agents that prevent C3 activation or enhance CFI or CFH activity to reduce the risk or progression of disease.

Strikingly, our study did not identify rare variants outside the complement pathway, despite the large set of genes queried within other AMD-associated loci and genes involved in angiogenesis, lipid metabolism and extracellular matrix signaling[4,37–39]. One possibility to explore, as sequencing studies scale up in size, is that rare coding variants with even more modest effect sizes than those reported here may modulate AMD risk within those other genes as well.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

1. Lim, L.S., Mitchell, P., Seddon, J.M., Holz, F.G. & Wong, T.Y. Age-related macular degeneration. *Lancet* **379**, 1728–1738 (2012).
2. Seddon, J.M., Cote, J., Page, W.F., Aggen, S.H. & Neale, M.C. The US twin study of age-related macular degeneration: relative roles of genetic and environmental influences. *Arch. Ophthalmol.* **123**, 321–327 (2005).
3. Friedman, D.S. *et al.* Prevalence of age-related macular degeneration in the United States. *Arch. Ophthalmol.* **122**, 564–572 (2004).
4. Fritsche, L.G. *et al.* Seven new loci associated with age-related macular degeneration. *Nat. Genet.* **45**, 433–439 (2013).
5. Raychaudhuri, S. *et al.* A rare penetrant mutation in *CFH* confers high risk of age-related macular degeneration. *Nat. Genet.* **43**, 1232–1236 (2011).
6. McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
7. Goldstein, J.I. *et al.* zCall: a rare variant caller for array-based genotyping: genetics and population analysis. *Bioinformatics* **28**, 2543–2545 (2012).
8. Huyghe, J.R. *et al.* Exome array analysis identifies new loci and low-frequency variants influencing insulin processing and secretion. *Nat. Genet.* **45**, 197–201 (2013).
9. Li, B. & Leal, S.M. Methods for detecting associations with rare variants for common diseases: application to analysis of sequence data. *Am. J. Hum. Genet.* **83**, 311–321 (2008).
10. Neale, B.M. *et al.* Testing for an unusual distribution of rare variants. *PLoS Genet.* **7**, e1001322 (2011).
11. Fagerness, J.A. *et al.* Variation near complement factor I is associated with risk of advanced AMD. *Eur. J. Hum. Genet.* **17**, 100–104 (2009).
12. Adzhubei, I.A. *et al.* A method and server for predicting damaging missense mutations. *Nat. Methods* **7**, 248–249 (2010).
13. MacArthur, D.G. *et al.* A systematic survey of loss-of-function variants in human protein-coding genes. *Science* **335**, 823–828 (2012).
14. Hicks, S., Wheeler, D.A., Plon, S.E. & Kimmel, M. Prediction of missense mutation functionality depends on both the algorithm and sequence alignment employed. *Hum. Mutat.* **32**, 661–668 (2011).
15. Bienaime, F. *et al.* Mutations in components of complement influence the outcome of Factor I–associated atypical hemolytic uremic syndrome. *Kidney Int.* **77**, 339–349 (2010).
16. Caprioli, J. *et al.* Genetics of HUS: the impact of *MCP*, *CFH*, and *IF* mutations on clinical presentation, response to treatment, and outcome. *Blood* **108**, 1267–1279 (2006).
17. Nilsson, S.C. *et al.* A mutation in factor I that is associated with atypical hemolytic uremic syndrome does not affect the function of factor I in complement regulation. *Mol. Immunol.* **44**, 1835–1844 (2007).
18. Geelen, J. *et al.* A missense mutation in factor I (*IF*) predisposes to atypical haemolytic uraemic syndrome. *Pediatr. Nephrol.* **22**, 371–375 (2007).
19. Sellier-Leclerc, A.-L. *et al.* Differential impact of complement mutations on clinical characteristics in atypical hemolytic uremic syndrome. *J. Am. Soc. Nephrol.* **18**, 2392–2400 (2007).
20. Fremeaux-Bacchi, V. *et al.* Complement factor I: a susceptibility gene for atypical haemolytic uraemic syndrome. *J. Med. Genet.* **41**, e84 (2004).
21. Kavanagh, D. *et al.* Characterization of mutations in complement factor I (*CFI*) associated with hemolytic uremic syndrome. *Mol. Immunol.* **45**, 95–105 (2008).
22. Li, M. *et al.* CFH haplotypes without the Y402H coding variant show strong association with susceptibility to age-related macular degeneration. *Nat. Genet.* **38**, 1049–1054 (2006).
23. Raychaudhuri, S. *et al.* Associations of CFHR1-CFHR3 deletion and a *CFH* SNP to age-related macular degeneration are not independent. *Nat. Genet.* **42**, 553–555; author reply 555–556 (2010).
24. Maller, J. *et al.* Common variation in three genes, including a noncoding variant in *CFH*, strongly influences risk of age-related macular degeneration. *Nat. Genet.* **38**, 1055–1059 (2006).
25. Lohmueller, K.E., Pearce, C.L., Pike, M., Lander, E.S. & Hirschhorn, J.N. Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nat. Genet.* **33**, 177–182 (2003).
26. Maller, J.B. *et al.* Variation in complement factor 3 is associated with risk of age-related macular degeneration. *Nat. Genet.* **39**, 1200–1201 (2007).
27. Yates, J.R.W. *et al.* Complement C3 variant and the risk of age-related macular degeneration. *N. Engl. J. Med.* **357**, 553–561 (2007).
28. Nishiguchi, K.M. *et al.* C9-R95X polymorphism in patients with neovascular age-related macular degeneration. *Invest. Ophthalmol. Vis. Sci.* **53**, 508–512 (2012).
29. Wu, J. *et al.* Structure of complement fragment C3b–factor H and implications for host protection by complement regulators. *Nat. Immunol.* **10**, 728–733 (2009).
30. Frémeaux-Bacchi, V. *et al.* Mutations in complement C3 predispose to development of atypical hemolytic uremic syndrome. *Blood* **112**, 4948–4952 (2008).
31. Sánchez-Corral, P. *et al.* Structural and functional characterization of factor H mutations associated with atypical hemolytic uremic syndrome. *Am. J. Hum. Genet.* **71**, 1285–1295 (2002).
32. Józsi, M. *et al.* Factor H and atypical hemolytic uremic syndrome: mutations in the C-terminus cause structural changes and defective recognition functions. *J. Am. Soc. Nephrol.* **17**, 170–177 (2006).
33. Manuelian, T. *et al.* Mutations in factor H reduce binding affinity to C3b and heparin and surface attachment to endothelial cells in hemolytic uremic syndrome. *J. Clin. Invest.* **111**, 1181–1190 (2003).
34. Montes, T., Tortajada, A., Morgan, B.P., Rodriguez de Cordoba, S. & Harris, C.L. Functional basis of protection against age-related macular degeneration conferred by a common polymorphism in complement factor B. *Proc. Natl. Acad. Sci. USA* **106**, 4366–4371 (2009).
35. Heurich, M. *et al.* Common polymorphisms in C3, factor B, and factor H collaborate to determine systemic complement activity and disease risk. *Proc. Natl. Acad. Sci. USA* **108**, 8761–8766 (2011).
36. Clark, S.J. *et al.* Tissue-specific host recognition by complement factor H is mediated by differential activities of its glycosaminoglycan-binding regions. *J. Immunol.* **190**, 2049–2057 (2013).
37. Yu, Y. *et al.* Common variants near FRK/COL10A1 and VEGFA are associated with advanced age-related macular degeneration. *Hum. Mol. Genet.* **20**, 3699–3709 (2011).
38. Neale, B.M. *et al.* Genome-wide association study of advanced age-related macular degeneration identifies a role of the hepatic lipase gene (*LIPC*). *Proc. Natl. Acad. Sci. USA* **107**, 7395–7400 (2010).
39. Chen, W. *et al.* Genetic variants near TIMP3 and high-density lipoprotein–associated loci influence susceptibility to age-related macular degeneration. *Proc. Natl. Acad. Sci. USA* **107**, 7401–7406 (2010).

[161]

Supplemental Manuscript 2

Yu, Y, Triebwasser, MP, Schramm, E, Wong, E. 2014

Human Molecular Genetics

Whole-exome sequencing identifies rare, functional *CFH* variants in families with macular degeneration

Yi Yu[1,†], Michael P. Triebwasser[2,†], Edwin K. S. Wong[3,†], Elizabeth C. Schramm[2,†], Brett Thomas[4], Robyn Reynolds[1], Elaine R. Mardis[5], John P. Atkinson[2], Mark Daly[4,6], Soumya Raychaudhuri[6,7,8,9,10], David Kavanagh[3] and Johanna M. Seddon[1,11,12,*]

[1]Ophthalmic Epidemiology and Genetics Service, New England Eye Center, Tufts Medical Center, Boston, MA, USA, [2]Division of Rheumatology, Department of Medicine, Washington University School of Medicine, St Louis, MO, USA, [3]Institute of Genetic Medicine, Newcastle University, International Centre for Life, Newcastle upon Tyne, UK, [4]Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA, USA, [5]The Genome Institute at Washington University, Saint Louis, MO, USA, [6]Program in Medical and Population Genetics, Broad Institute, Cambridge, MA, USA, [7]Partners HealthCare Center for Personalized Genetic Medicine, Boston, MA, USA, [8]Division of Genetics, Brigham and Women's Hospital, Boston, MA, USA, [9]Division of Rheumatology, Immunology and Allergy, Brigham and Women's Hospital, Boston, MA, USA, [10]Faculty of Medical and Human Sciences, University of Manchester, Manchester, UK, [11]Department of Ophthalmology, Tufts University School of Medicine, Boston, MA, USA and [12]Sackler School of Graduate Medical Sciences, Tufts University, Boston, MA, USA

We sequenced the whole exome of 35 cases and 7 controls from 9 age-related macular degeneration (AMD) families in whom known common genetic risk alleles could not explain their high disease burden and/or their early-onset advanced disease. Two families harbored novel rare mutations in *CFH* (R53C and D90G). R53C segregates perfectly with AMD in 11 cases (heterozygous) and 1 elderly control (reference allele) (LOD = 5.07, $P = 6.7 \times 10^{-7}$). In an independent cohort, 4 out of 1676 cases but none of the 745 examined controls or 4300 NHBLI Exome Sequencing Project (ESP) samples carried the R53C mutation ($P = 0.0039$). In another family of six siblings, D90G similarly segregated with AMD in five cases and one control (LOD = 1.22, $P = 0.009$). No other sample in our large cohort or the ESP had this mutation. Functional studies demonstrated that R53C decreased the ability of FH to perform decay accelerating activity. D90G exhibited a decrease in cofactor-mediated inactivation. Both of these changes would lead to a loss of regulatory activity, resulting in excessive alternative pathway activation. This study represents an initial application of the whole-exome strategy to families with early-onset AMD. It successfully identified high impact alleles leading to clearer functional insight into AMD etiopathogenesis.

INTRODUCTION

With the emerging understanding of the strong genetic underpinnings of age-related macular degeneration (AMD) (MIM 603075) based on genome-wide association studies and more recently targeted sequencing with the discovery of rare, penetrant variants (1,2), we are approaching an era of applying these genetic discoveries for personalized medicine. Common variants collectively allow us to predict increased risk of AMD and progression of disease on a population level (3–5), but may not be applicable for an individual family. Rare and penetrant variants, which are strongly related to AMD, can have a major impact on disease risk in an individual and a family. For example, the *CFH* R1210C variant, with a gene frequency of 0.02% in the European population, carries a 20-fold increased risk of AMD for an individual with the heterozygous mutation, the strongest genetic risk factor for AMD to date. These individuals are also diagnosed with advanced AMD at an earlier age (1).

*To whom correspondence should be addressed. Email: jseddon@tuftsmedicalcenter.org
[†]These authors contributed equally to this work.

More recently, rare variants in *C3*, *CFI* and *C9* have also been discovered that influence disease risk (2,6,7). These rare variants open up new avenues for family counseling, selective and possibly more informative clinical trials, and eventually for personalized targeted therapies.

We previously demonstrated that some families, despite a high density of affected individuals, have a lower than expected genotypic load based on known common variants, suggesting they may harbor undiscovered risk variants (8). It has been reported that rare, high penetrant mutations that co-segregate with early-onset or severe phenotypes may exist among the distribution of a common trait (e.g. high-density lipoprotein cholesterol) (9). They can be transmitted through generations like a dominant Mendelian disease, yet remain hidden in the population given the high frequency of AMD. Since very rare (private) mutations occur in only a few families, we performed state-of-the-art whole-exome sequencing in families most likely to harbor novel or rare, highly penetrant variants. Such variants in *CFH* were discovered to segregate with disease in two of the nine families reported here.

To understand the impact of these variants, we performed functional studies to characterize the effect(s) these mutations have on protein function. Functional studies are crucial to elucidate the relationship between rare variants and disease given the limited power to associate these variants in all but the largest families and cohorts.

RESULTS

We sequenced the exomes of 42 samples from 9 families with unusually low common variant AMD risk (Methods). The coverage was $>10\times$ for a median of 97.25% of the targeted regions (Supplementary Material, Fig. S1A). For the 18 783 RefSeq genes captured by the Nimblegen Human exome library, 16 717 genes were sequenced at $>10\times$ depth for 90% of the regions targeted per gene (Supplementary Material, Fig. S1B). After strict quality control, we identified 153 534 high quality variants with no missing genotypes. To test the overall quality of the sequence data, we compared the genotypes of variants found in the sequence data to variants derived from genotyping via an exome array (Illumina Infinium Human Exome BeadChip v1.0). There were 20 454 variants that were captured by sequencing and that were also genotyped using the exome array. The concordance

between sequencing and genotyping data was 99.76%. Allelic dosages were almost perfectly correlated ($r > 0.99$) for 93.5% of common variants ($f > 0.01$) and 99.2% of rare variants (Supplementary Material, Fig. S1C and D); only 1.3 and 0.7% of common and rare SNPs, respectively, had modestly correlated dosages ($r < 0.9$).

To screen for rare functional variants in the selected families, we filtered the variants in each step by several criteria (Table 1). First, we included only coding SNPs of high and moderate impact on protein sequence/structure (missense, nonsense, read-through or splice variants). We excluded SNPs with minor allele frequency greater than 0.1% in the 1000 genomes project or the NHBLI exome sequencing project (ESP). There was only one variant exome-wide (R53C in *CFH*) remaining in pedigree V that segregated with advanced AMD (Table 1). Remarkably, this was a missense variant, R53C, in the well known AMD gene, *CFH*.

For other families with multiple candidate SNPs remaining, we included SNPs that were annotated as loss of function or probably damaging by PolyPhen-2 or were assigned 'deleterious' by SIFT. Using functional prediction, pedigree II yielded a novel variant that segregated with disease and was predicted to be deleterious: D90G in *CFH* (Table 1). To further narrow down the list of candidate SNPs remaining after filtering (rare, segregating in exome sequenced individuals and predicted to be altered by either Sift or Polyphen-2), we recruited additional relatives and obtained their genotypes and analyzed segregation (Table 2 and Supplementary Material, Table S1). Table 2 shows the genomic locations, amino acid change and minor allele frequencies in ESP of the SNPs that passed filtering and segregated in each family with LOD scores > 1.0. The results for the candidate SNPs with no evidence of segregation when using data from sequenced and genotyped members of each family are shown in Supplementary Material, Table S1.

Because two of the new rare variants (R53C and D90G) are located in *CFH*, in which both a rare causal mutation (R1210C) (1) and common alleles have been previously associated (11–13), we conducted follow-up analyses by recruiting additional family members and a replication analysis in independent samples. We performed Sanger sequencing of the *CFH* gene in six additionally affected relatives in pedigree V using the same approach as our previous targeted sequencing project (2). These additional cases also carried the R53C mutation. There were a total of 11 cases heterozygous for R53C and 1 unaffected individual (90 years old) without the mutation in

Table 1. Number of SNPs identified in each family by filtering on their function and rareness in population databases

	I	II	III	IV	V	VI	VII	VIII	IX
Case: control ratio per family	3:1	5:1	4:1	3:1	5:1	4:0	4:1	4:0	3:1
High and moderate impact SNPs[a]	20397	18414	21313	19437	22984	17904	20853	17796	17603
MAF <0.1% in databases[b]	436	320	540	420	572	394	480	286	324
Shared between cases but not in controls	4	5	2	6	1	18	3	11	18
Probably damaging (Polyphen-2)	1	1	1	3	1	3	0	3	7
Deleterious (SIFT)	2	2	1	3	1	5	3	6	9
Probably damaging or deleterious	2	2	2	3	1	5	3	6	10
Probably damaging and deleterious	1	1	0	3	1	3	0	3	6

[a]High and moderate impact SNPs include missense, nonsense, read-through or splice variants.
[b]Variants with minor allele frequency (MAF) greater than 0.1% in the data sets of the 1000 genomes project or the NHLBI GO ESP were filtered out. Polyphen-2 is a tool that predicts possible impact of an amino acid substitution on the structure and function of a human protein using physical and comparative considerations. SIFT is a tool that predicts whether an amino acid substitution affects protein function. SNPs, single nucleotide polymorphisms.

Table 2. Potential pathogenic variants identified in each family with LOD score ≥1.0

FAM	Gene	Location	Ref	Alt	rsID	Amino acid	Allele frequency (ESP Europeans)	LOD score
II	CFH*	chr1:196643011	A	G		D90G	0	1.22
V	CFH*	chr1:196642206	C	T		R53C	0	5.07
VIII	CCDC75*	chr2:37315562	C	A		S9Y	0	1.50
IX	TARS2*	chr1:150477420	C	A		P620H	0	1.22
IX	WNT2	chr7:116918397	G	A	rs148046128	R299W	0.00093	1.22

Variants predicted to be both probably damaging by Polyphen-2 and deleterious by SIFT. FAM, family; Ref, reference allele; Alt, alternative allele; rsID, rs number; ESP, exome sequencing project. Asterisks indicate genes expressed in retina tissue (10).
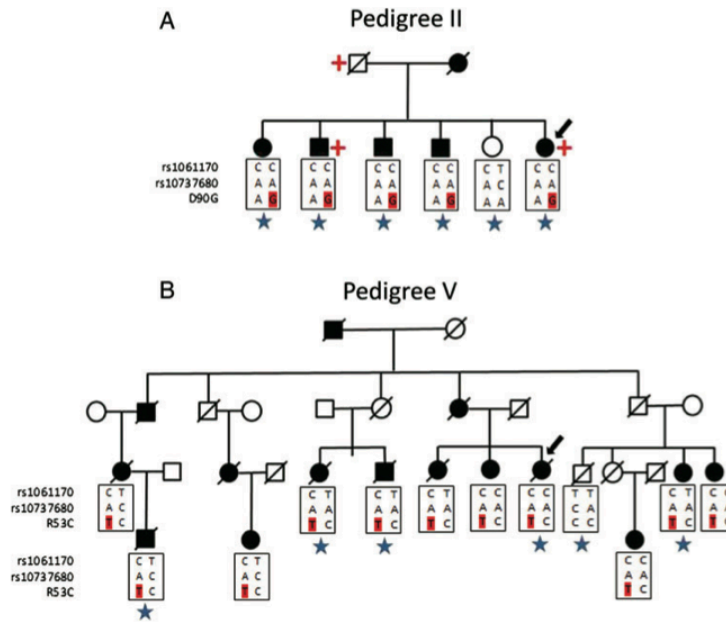
Figure 1. Rare *CFH* mutations segregate with AMD in families. (**A**) Pedigree of a family with AMD carrying the D90G mutation in *CFH*. (**B**) Pedigree of a large multigenerational family with AMD explained by the R53C mutation in *CFH*. Family members in the generation of the proband (arrow) and their descendants are displayed except for individuals not enrolled in the study including some unaffected individuals. Common antecedents are also displayed although deceased and not enrolled. Squares represent male and circles are female family members. Black symbols indicate affected persons and white symbols represent unaffected persons. Slashes indicate deceased family members. The genotype information of common risk SNPs in *CFH* and R53C or D90G mutation allele (in red) are shown below each genotyped individual. Blue star indicates a whole-exome sequenced sample. Red cross indicates individual with renal disease or low kidney function.

pedigree V (Fig. 1B). Complement Factor H R53C was significantly linked to advanced AMD in this pedigree (LOD = 5.07, $P = 6.7 \times 10^{-7}$). We separately sequenced *CFH* in our study targeting 681 genes in 1676 cases and 745 controls (2). We identified four cases with the R53C mutation (unrelated to individuals in pedigree V). Combining the ESP data set of European ancestry as shared controls (4300 samples) with our targeted sequencing data set, R53C was associated with AMD (Fisher's exact test, $P = 0.0039$).

Since some of the original families were selected for advanced AMD of early onset, we tested whether cases with R53C allele in the independent cohort also tended to be diagnosed with advanced AMD at an early age. We compared the age of diagnosis of advanced AMD in the four patients with R53C mutations who were unrelated to family V and in 1832 unrelated patients without rare mutations in *CFH, C3* or *CFI* genes. As seen previously with the rare R1210C variant of *CFH* (1), patients with R53C mutations had a significantly earlier onset of advanced AMD (median of 71 versus 76 years, $P = 0.03$, the rank-sum test).

Based on the exome sequencing of pedigree II, there were five affected individuals with D90G and one unaffected individual without this variant, resulting in a LOD score of 1.22 ($P = 0.009$). We did not find any additional samples with the D90G mutation in our targeted sequencing project and there were no additional relatives available for sequencing in this pedigree (Fig. 1A). The proband of this family has 'low renal function' and is on dialysis and her brother has chronic kidney disease

stage III with creatinine clearance of 41. Their father also had kidney disease (specific diagnosis unknown).

Although the R53C and D90G risk alleles reside on the *CFH* haplotype containing the rs1061170(Y402H) and rs10737680 common risk alleles (Fig. 1A and B), the common risk alleles cannot explain the extraordinary large effect size and the early-onset cases in these families. Therefore, we carried out functional studies to investigate the potential impact of R53C and D90G risk alleles on the activation of the complement pathway.

We measured serum FH levels in the 22 individuals carrying the R53C and D90G variants and the levels of secreted FH protein were normal. We then generated wild-type (FH1-4$_{WT}$) and mutant (FH1-4$_{R53C}$; FH1-4$_{D90G}$) in a recombinant fragment, comprising the functional domains CCPs 1–4 of FH. These proteins were expressed in both a yeast system and a mammalian

system. We saw the same effects in proteins produced in either system.

Effect of R53C and D90G on C3b binding

We used SPR to examine the binding interaction between FH1-4 (0.04–40 μM) to immobilized C3b. Overlaying sensograms show the steady-state response for the concentration range. We calculated the dissociation constant (K_D) using a steady-state affinity model. The affinity of FH1-4$_{WT}$ and FH1-4$_{D90G}$ for C3b was similar, 7.6 and 8.3 μM, respectively. The affinity of FH1-4$_{R53C}$ for C3b is slightly weaker than FH1-4$_{WT}$. The K_D was 12.2 μM for the FH1-4$_{R53C}$–C3b interaction versus 8.3 μM for FH1-4$_{WT}$ (Fig. 2). In addition, we identified a similar trend if the FH1-4 proteins were immobilized on a sensor chip. FH1-4$_{WT}$ and FH1-4$_{D90G}$ had a K_D of 8.6 and 8.3 μM,
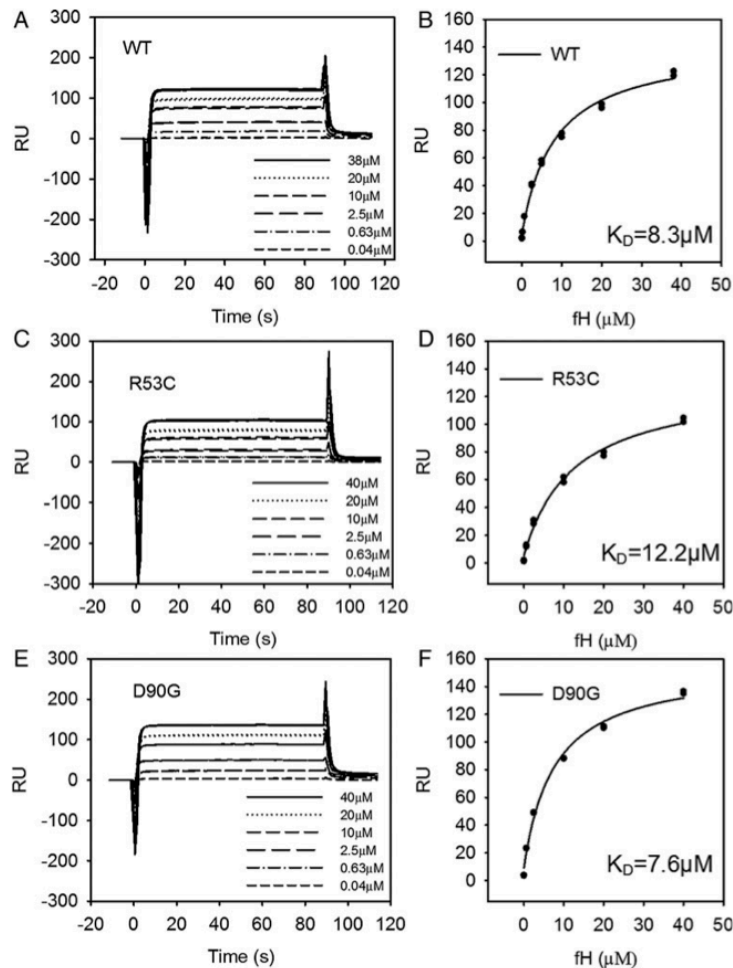
Figure 2. R53C binds C3b with a lower affinity than WT. Overlaying sensograms show the steady-state response for the binding of CFH1-4 (*Pichia*) (0.4–40 μM) to immobilized C3b on the surface of a CM5 sensor chip for (**A** and **B**) WT, (**C** and **D**) R53C and (**E** and **F**) D90G. The steady-state response was plotted against concentration and an affinity curve was fitted.

[166]

respectively, while the K_D of FH 1-4$_{R53C}$ for C3b was 34 μM (Supplementary Material, Fig. S2).

Decay accelerating activity (DAA)

We then examined decay accelerating activity (DAA) of FH1-4$_{R53C}$, FH1-4$_{D90G}$ and FH1-4$_{WT}$ using SPR (Fig. 3). In the SPR assay, the FH1-4$_{R53C}$ showed a marked loss of DAA, while FH1-4$_{D90G}$ exhibited normal DAA (Fig. 3). To further investigate the DAA defect in R53C, we used a sheep erythrocyte lysis assay. Fifty percent inhibition of lysis was achieved using 6.0 nM FH1-4$_{WT}$; however, 2000 nM (a >300-fold difference) was required to obtain the same inhibition using FH1-4$_{R53C}$. (Supplementary Material, Fig. S3).

Both R53C and D90G affect cofactor function

Using a fluid phase assay in which FH1-4 supplies cofactor activity (CA) for the protease CFI, we saw decreased CA with both mutants. As C3b is cleaved to iC3b, the C3 α' chain is consumed and α_1 and α_{40} fragments are generated (Supplementary Material, Fig. S4). Both mutants lag behind FH1-4$_{WT}$ in terms of α' remaining (Fig. 4) over a 30-min time course compared with the original amount of α'. FH1-4$_{D90G}$'s activity is significantly less than FH1-4$_{WT}$ at all time points (Fig. 4A, $P < 0.01$, the two-tailed t-test). FH1-4$_{R53C}$ also had a consistent and reproducible trend toward less activity than FH1-4$_{WT}$ (Fig. 4B, $P > 0.05$). To confirm that FH1-4$_{R53C}$'s CA was less than FH1-4$_{WT}$, we also used an hemolysis based assay; $\sim$8 μM of FH1-4$_{WT}$ was required to achieve 50% inhibition of lysis, while at any given concentration of FH1-4$_{R53C}$, less protection was seen (Fig. 4C).

DISCUSSION

Common variants explain a large part of the genetic burden related to AMD. However, the observation of densely affected families with early onset cases of advanced AMD lacking a significant contribution from common alleles suggested that rare, more Mendelian-like subtypes of disease exist, being hidden within the otherwise common phenotype. These rare, highly penetrant alleles can be detected only by sequencing, so we applied the strategy of whole-exome sequencing in these AMD families. To our knowledge, cohorts and families with AMD have never been thoroughly investigated previously on the whole-exome level. This approach successfully identified two rare mutations, R53C and D90G, in the *CFH* gene. These two newly identified *CFH* mutations are highly penetrant, similar to the previously identified R1210C mutation in the same gene (1) but with lower frequencies. The strong evidence of linkage and association with these very rare variants observed in this study indicates that our family selection strategy leveraged the inherent power of the family structure and effectively increased the power to detect such mutations. Our functional studies of these new high impact alleles provide clearer insight into the mechanisms underlying the development of AMD.

Other than these two new rare variants in *CFH*, several other variants passed our filtering analysis in selected families. Compared with the other families, the two pedigrees for whom a causal allele was identified in this report had more affected individuals as well as an unaffected individual sequenced, thus allowing identification and investigation of a smaller number of variants that followed the assumed dominant inheritance pattern. Other pedigrees did not have additional members available so we resorted to bioinformatic means to narrow the list of variants that could explain their disease. Such tools are not 100% accurate and additional causal variants may exist which are not predicted to be damaging or deleterious or that were not well covered by the exome sequencing. As more affected pedigrees are interrogated, we expect more overlap will occur between 'causal variants' in terms of genes and pathways. While availability of functional assays and known relationship of *CFH* to AMD made interpretation of new *CFH* variants more straightforward, some of these additional prioritized variants may also be relevant but await replication in additional families and unrelated cases as well as functional characterization.

Several rare risk alleles in genes of the complement pathway have been identified recently, including R1210C in *CFH* (1), G119R in *CFI* (14), K155Q in *C3* (2) and P167S in *C9* (2), as
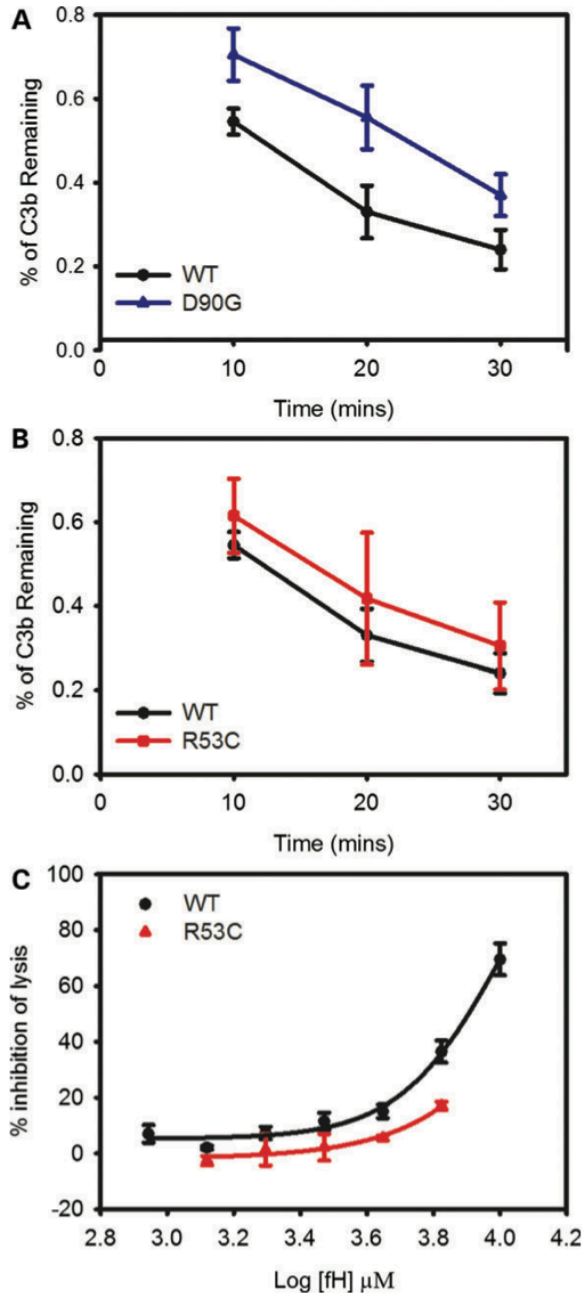
Figure 3. R53C has a profound decay accelerating defect. SPR decay accelerating assay for (**A**) R53C and (**B**) D90G. R53C fails to facilitate decay of the C3b:Bb complex in the same way that WT and D90G do (representative trace of three replicates).

Figure 4. Both D90G and R53C have defects in CA. (**A**) D90G has a significant defect in co-factor activity (CA) ($P < 0.01$ at all time points, two-tailed t-test). (**B**) R53C has lower CA in the fluid phase assay ($P > 0.05$ at all time points, two-tailed t-test). (**C**) R53C also has a significant defect in CA in a hemolysis based assay. For (A) and (B), quantitation of the α' produced over time relative to the amount of α' present originally. These are the average of four independent experiments; error bars represent one standard deviation. Quantification was done by densitometrically scanning the α' bands and normalizing to the amount of C3b in that lane (β chain).

well as multiple rare variants in *CFI* (2). Most of these rare alleles were found in genes where a common AMD risk locus had been previously associated, but they are independent of the common

risk haplotypes in the disease gene. In this family-based study, the two newly identified risk alleles, R53C and D90G, reside on the common risk haplotype but have larger effects. While these rare variants might be tagging a different causal mutation, we consider this possibility unlikely given the functional alterations conferred by these mutations. It is worthwhile to continue searching for other rare, causal, function-altering variants that are likely to unveil the underlying roles of the associated genes in AMD pathogenesis and to provide new targets for screening, prevention and treatment for AMD.

The molecular mechanisms of R53C and D90G are different from the previously identified R1210C substitution in *CFH*, which disrupts FH binding to C3b and heparin (15,16), and has been demonstrated to circulate in serum complexed covalently to albumin (15,17,18). R1210C and the common allele Y402H both affect localization of FH onto surfaces (19–22). These new rare variants illustrate that, even when localized, FH must have intact regulatory function to prevent AMD.

The R53C variant shows only a minor decrease in binding affinity for C3b. A co-crystal structure of C3b/FH demonstrates that R53 does not oppose C3b, but is on the exposed face of FH, in keeping with the minor differences in binding (Fig. 5). The R53C variant, however, severely impairs DAA of FH1-4 and reduces its CA.

R53C is an example of the pleiotropic effects mutations in the alternative pathway (AP) can have. Generally, rare, highly penetrant variants in this pathway have been linked to a variety of kidney diseases, which can vary widely in their presentation and pathologic findings. The R53C variant has been detected in three such patients with glomerular disease. One patient was heterozygous for this variant and had moderately low C3 levels and C3 glomerulopathy (24). Another patient was heterozygous for R53C, had moderately low C3 levels and developed atypical hemolytic uremic syndrome (aHUS) (25). A third patient was homozygous for the variant and had very low C3 levels and mesangioproliferative glomerulonephritis type 1 (24). Interestingly, this amino acid has also been found mutated to histidine in another patient with aHUS (26). Functional analysis of the R53H variant demonstrated only a minor perturbation of binding to C3b, consistent with this amino acid not playing a major role in the interaction between C3b and FH. As with the R53C variant, the R53H variant also profoundly impaired DAA and reduced CA. None of the individuals carrying R53C in pedigree V reported renal disease. Both R53C and R1210C highlight the phenotypic heterogeneity possible with variants leading to dysregulation of the AP. There are no reports in the published literature of associations between D90G and kidney disease, to our knowledge. In our pedigree with the D90G mutation, two affected individuals with the variant reported having chronic kidney disease or reduced kidney function. It remains to be seen if individuals with such variants that manifest with renal disease will go on to develop AMD at higher rates.

A model of FI docked in a groove between FH and C3b would suggest that R53 would oppose FI in keeping with the impaired CA (27). Additionally, the loss of DAA suggests that an interaction with FB occurs on this face; overlapping binding sites for FI and FB have previously been suggested (26).

Despite residing in the interface between FH and C3b, the D90G variant does not impair the binding to C3b or disrupt DAA at levels detectable in our assays. D90G does, however,
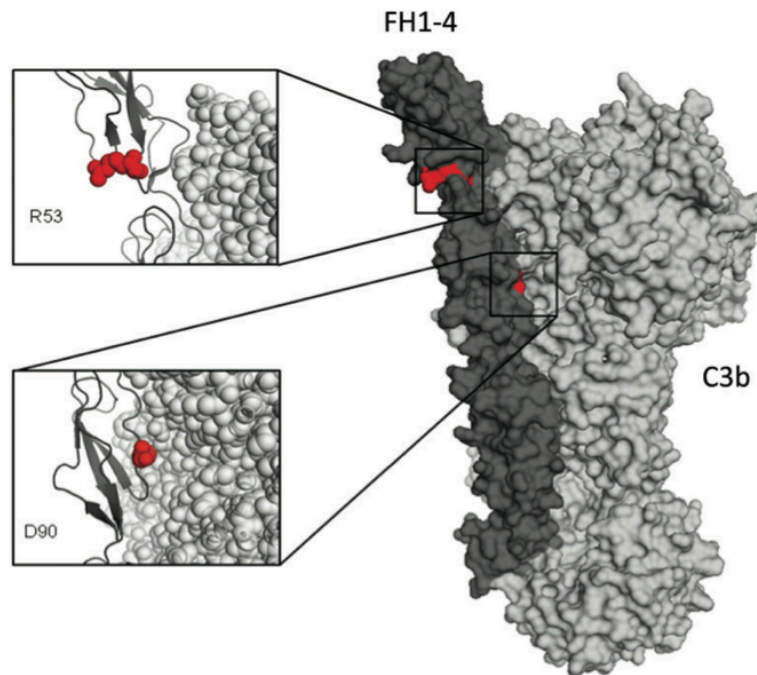
Figure 5. FH1-4-C3b co-crystal structure demonstrating the position of the R53C and D90G mutations. An X-ray-derived co-crystal structure of FH-C3b (23) was used to model the mutation and displayed with Pymol (Delano Scientific). The location of the two mutated amino acids (red spheres) are shown within the co-crystal structure of a FH1-4 (dark grey):C3b (light grey) complex. The R53 amino acid is on the opposite face to the FH-C3b interface, while the D90 amino acid is in direct opposition to C3b (Protein Data Base ID code 2WII) (23).

affect CA. In aHUS, CA is the critical modality of regulation required for AP homeostasis. This variant has not been identified previously in normal individuals (NHLBI ESP and 1000 Genome Project) or patient populations (aHUS or C3 glomerulopathies). The D90G phenotype of normal C3b binding to one of its regulators, but reduced CA has previously been described for complement inhibitors. Specifically, mutations in the C3b binding site of MCP (CD46) and CR1 (CD35) have been characterized with this type of a dysfunctional pattern (28). CA requires the transient formation of a tri-molecular complex (C3b:Cofactor:Factor I), which is dominated by low affinity interactions. While in most cases, dysfunction moves in parallel (i.e. alterations in C3b binding mimic those observed in CA), the requirement for FI to join the party indicates that some variants may solely affect CA (29–31).

Small losses in regulatory activity may lead to excessive AP activation over decades that manifest itself as the degenerative process of AMD. Thus, rare mutations such as D90G, with a more subtle loss of function, may not be fully penetrant for kidney disease, but are likely to be causative in AMD.

The identification of such rare variants with large effect has implications for clinical trial design and the medical care of carriers and their relatives. Currently, complement inhibitors are available for a variety of complementopathies (32,33) and promising results have come for a phase II trial in AMD (34). Carriers of variants such as those described here in *CFH*, and already described in *CFI*, would facilitate proving the concept that such a therapeutic intervention is efficacious.

MATERIALS AND METHODS

Diagnosis and family selection

All family members included in this study were evaluated by board-certified ophthalmologists. Individuals were evaluated (i) by clinical examination with visual acuity measurements, dilated slit-lamp biomicroscopy and stereoscopic color fundus photography or (ii) by reviewing ophthalmologic medical records. All individuals were graded using the Clinical Age-Related Maculopathy Grading System (CARMS) (35). Advanced AMD patients had either geographic atrophy (advanced dry AMD) or neovascular disease (wet AMD). Phenotypes were based on fundus photography or dilated fundus examination using standardized protocols.

Among a collection of 591 families we recruited over the past two decades, 364 families have at least two members with advanced AMD. Using an algorithm we previously published, we selected three families with a genotypic load that was less than expected given the density of disease in their family (3,8,36). We also included one family with more than two cases with a low genetic risk score (< -1.0) derived from a general linear model of the genotype dosages of known common alleles (37). Only 5% of the advanced cases in our GWAS cohort had such a low genetic risk score (< -1.0). The chance of observing two or more cases with low genetic risk score in a family by random chance is rare. Finally, we included five families with at least two advanced AMD cases diagnosed before age 60. Demonstrating the effect of the R1210C

variant, nine other families identified using these criteria had the high-risk *CFH* R1210C allele and were excluded from exome sequencing (1).

For each selected family, all advanced AMD cases with sufficient DNA and an unaffected relative with CARMS grade 1 (if available) were selected for whole exome sequencing. The number of cases and controls in each family are shown in Table 1.

Exome capture and sequencing

The samples were sequenced at the Genome Institute at Washington University. Genomic DNA extracted from blood samples was used to construct Illumina libraries with indexing barcodes to uniquely identify each sequenced individual. We sequenced family members by combining the indexed libraries at an equimolar mixture and then hybridizing to the Nimblegen 3.0 human exome reagent, designed to capture a total 60 Mb consisting of the coding exons and splice junctions of genes cataloged in the human RefSeq database, non-repetitive 3′ and 5′ UTRs, and selected non-coding RNA sequences. We isolated the hybridized library fragments, quantitated with qPCR and then sequenced with the Illumina HiSeq2000™ platform. We produced 2×100 bp reads on the HiSeq2000.

Read mapping, variant detection and annotation

Following deconvolution of the barcodes from each Illumina lane, we aligned individual reads for each sample to the human reference genome (NCBI build 37.3, hg19) using Burrows-Wheeler Aligner (BWA, v0.59) (38). We called the consensus genotypes in the target regions with SAMtools (39) and VarScan 2 (40) in the genomic regions targeted by the capture probe set. We filtered variants with 'PASS' and high quality (Phred-like quality score ≥ 30 and an allele balance ≥ 35) to remove systematic false positives followed by annotation with the VEP tool (41). All sequenced samples were required to have over $10\times$ coverage at greater than 90% of the targeted regions and over $20\times$ coverage at greater than 80% of the targeted regions.

Data analyses

To distinguish potentially pathogenic mutations from other variants, we examined variants that altered coding sequences (missense, nonsense, read-through or splice variants). We predicted that the variants responsible for advanced AMD in the selected families would be very rare and therefore likely to be previously unidentified or have low frequency in public databases. Consequently, we filtered out variants with allele frequency >0.1% in the data sets of the 1000 genomes project (42) or the NHLBI GO ESP (43). Since we were searching for mutations that have a high penetrance and follow a dominantly inherited pattern, we prioritized candidate variants that occur in all cases in a pedigree and none of the sequenced controls. Finally, when evaluating candidate variants, we applied one additional filter to include only variants predicted to be loss of function or probably damaging by PolyPhen-2 (44) or predicted to be deleterious by SIFT (45). These analyses were performed by xBrowse, developed at ATGU at the Massachusetts General Hospital (http://atgu.mgh.harvard.edu/xbrowse). We performed linkage analysis with the LINKAGE programs (46). Fisher's exact test was used to test the association in case-control samples.

TaqMan genotyping of selected SNPs

To confirm segregation within families, we genotyped candidate variants within additional samples within the same families at the Johns Hopkins Genotyping Core Laboratory using a custom made TaqMan genotyping assay by Applied Biosystems and with the ABI 7900 Real-Time PCR system. For the R53C and D90G variants in *CFH*, we sequenced the additional cases and controls using a targeted-sequencing approach described previously (2).

Production and purification of proteins

We generated and purified wild-type and mutant complement factor H (FH) in the setting of complement control protein (CCPs) 1-4 (FH1−4) using two distinct methods. The first production system utilized a *Pichia pastoris* system as previously described (26). Briefly, for wild-type we used a pPICZαB (Invitrogen) vector containing residues 19−263 of *CFH* (which encodes CCPs 1-4 of mature FH; residues 1−18 are the mammalian signal peptide) with an N-terminal myc tag and a C-terminal $6\times$ His tag. We generated the R53C and D90G point mutations using QuikChange site-directed mutagenesis kit (Stratagene) with the following primers: R53C—(f) cccaggctatctataaatgc <T> gccctggatatagatctcttgg; (r) ccaagagatctatatccagggc <A> gcatttatagatagcctggg and D90G—(f) ggccctgtggacatcctggag <G> tactccttttggtacttttaccc; (r) gggtaaaagtaccaaaaggagta <C> ctccaggatgtccacagggcc. We confirmed fidelity with bi-directional Sanger sequencing and transformed these vectors into *Pichia pastoris* (KM71H cells) by electroporation. Clones producing wild-type (FH1-4$_{WT}$) or mutant (FH1-4$_{MUT}$) FH1-4 were selected by zeocin and checked for protein expression.

We expressed protein in a 3 liter BioFlo 115 Biofermenter (New Brunswick) by transferring a starter culture in buffered minimal glycerol into 1 liter of basal fermentor salts [0.095% (w/v) calcium sulphate, 1.82% (w/v) potassium sulphate, 1.5% (w/v) magnesium sulphate heptahydrate, 0.42% (w/v) potassium hydroxide, 2.7% (v/v) phosphoric acid and 2.5% (v/v) glycerol, enriched with 1% (w/v) casein amino acids, 0.5% (w/v) PTM1 salts and 0.5% (v/v) antifoam A (Sigma)]. After the initial batch fed glycerol was exhausted, glycerol feeds were maintained for 24 h at 30°C. The cells were allowed to starve for 4 h before recombinant protein expression was induced with 0.75% methanol containing 0.5% (w/v) PTM1 salts. After 3 days at 15°C with methanol feeds, the supernatant was removed and filtered and its pH adjusted to 7.4.

We applied the supernatant to a 5 ml His trap column (GE-healthcare) at 4°C and eluted the protein with 500 mM imidazole followed by size exclusion chromatography on Superdex 200 (GE Healthcare). The protein concentrations were determined using absorbance at 280 nm and calculated extinction coefficients (47870 M.cm^{-1}).

The second method for production of FH$_{1-4}$ was performed in human 293-T cells. The first 265 amino acids, including the signal peptide of FH, were cloned by PCR and a C-terminal $6\times$ His Tag was added. This was inserted into a pSG5 vector using EcoR1. R53C was introduced using the following primers:
Forward: ACC CAG GCT ATC TAT AAATGC <T>GC CCT GGATAT AGATCT CTT

Reverse: AAG AGATCT ATATCC AGG GC<A>GCATTT ATA GAT AGC CTG GGT. D90G was introduced with forward: GGACATCCTGGAG<G>TACTCCTTTTGGT; reverse: ACCAAAAGGAGTA<C>CTCCAGGATGTCC. PCR with Phusion (NEB) in the presence of 1 M betaine was used to introduce mutations and they were confirmed with bidirectional Sanger sequencing.

Transfection of 293T cells was performed with XtremeGene-9 (Roche) in serum-free Optimem for 3 days. Supernatants were harvested and recombinant FH1-4 was purified using a His-Trap column (GE). Binding to the column was carried out at 50 mM imidazole with elution at 95 mM imidazole. Protein concentration was determined using an ELISA relative to full-length purified FH (CompTech). A monoclonal capture antibody (H2, #A254, Quidel) was coated in ELISA wells at 500 μg/ml. A goat polyclonal antibody (#A312, Quidel) was used for detection. Relative concentration was confirmed by subjecting FH 1-4 constructs to SDS−PAGE and subsequent silver stain or fluorescent stain with Krypton dye (Thermo).

Binding affinity for C3b by surface plasmon resonance

We monitored the binding affinity of FH1-4$_{WT}$, FH1-4$_{R53C}$ and FH1-4$_{D90G}$ by surface plasmon resonance (SPR) with Biacore X100 and Biacore 2000 instruments (GE Healthcare). These measurements were performed in two orientations. First, we immobilized 850 resonance units (RU) of human C3b (Comp-Tech) on a Biacore series CM5 sensor chip (GE Healthcare) using standard amine coupling. The reference surface of the chip was prepared by performing a mock coupling in the absence of any protein. We performed the experiments at 25°C and 30 μl/min flow rate in duplicate using FH1-4 produced in *Pichia* [concentrations 0.04−40 μM in 10 mM HEPES-buffered saline, 3 mM EDTA, 0.05% (v/v) surfactant p20 (GE Health-care)]. In the second method, we attached the FH1-4 recombinant proteins (produced in mammalian cells) to the CM5 sensor chips via standard amine coupling (~1000 RU). Purified C3b was injected in duplicate using multiple concentrations (500 nM−5.6 μM). In both cases, a contact time of 90 s was used (sufficient for achieving steady-state conditions) followed by a dissociation period of 45 s. We regenerated the chip surface between sample injections with one 45 s injection of 1 M NaCl (pH 7.0). We processed data using the BIAevaluation 4.1 software (GE Healthcare). The data from the reference cell and a blank (buffer) injection were subtracted and we calculated dissociation constants using a steady-state affinity model from the background-subtracted traces.

Measurement of decay accelerating activity by SPR

We measured DAA in real-time using a Biacore X100 instrument as described previously (26). Briefly, we immobilized 850 RU of C3b using standard amine coupling Biacore series CM5 sensor chip (GE Healthcare). We injected a mixture of 500 nM complement factor B (fB), and 60 nM complement factor D (fD), at 10 μl/min over the surface for 120 s to form the AP C3 convertase.

We allowed the convertase to decay naturally for 210 s and then injected 0.5 μM FH1-4$_{WT}$ or FH1-4$_{MUT}$ produced in *Pichia* [in running buffer, HEPES-buffered saline containing 0.5% (v/v) surfactant P20 and 1 mM MgCl$_2$, 25°C] across the surface for 60 s and C3 convertase decay was visualized in real time. Between injections, we regenerated the chip surface using a 45 s injection of 10 μM FH1-4$_{WT}$ followed by a 45 s injection of 1 M NaCl, pH 7.0. We evaluated data using BIAevaluation 4.1 (GE Healthcare). As a control, the construct was also flowed over the bare surface and binding data subtracted from the decay sensogram.

Measuring DAA on sheep erythrocytes

We prepared C3b-coated sheep erythrocytes (EA-C3b) as described previously (26). The cells were resuspended to 2% (v/v) in AP buffer (5 mM sodium barbitone, pH 7.4, 150 mM NaCl, 7 mM MgCl$_2$, 10 mM EGTA) and then the AP C3 convertase was formed on the cell surface by incubating 50 μl of the cell preparation with an equal volume of AP buffer containing fB (40 μg/ml; purified from serum by Dr C. Harris, Cardiff) and fD (0.4 μg/ml; CompTech) at 37°C for 15 min. We incubated cells with 50 μl of a concentration range of CFH1-4$_{WT}$ and CFH1-4$_{R53C}$ in PBS/20 mM EDTA for 15 min (both produced in *Pichia*). We developed lysis by adding 50 μl of 4% (v/v) normal human serum depleted of fB and CFH (NHS-ΔBΔH; prepared by Dr C. Harris, Cardiff) (47) in PBS/20 mM EDTA and incubating at 37°C for 60 min. To determine the amount of lysis, we pelleted cells by centrifugation and hemoglobin release was measured at 410 nm (A$_{410}$). We used buffer only (no FH1-4) to define maximum lysis by the addition of serum-A$_{410}$ (buffer only). Percentage of inhibition of lysis in the presence of increasing concentrations of FH was defined as [A$_{410}$(buffer only)−A$_{410}$(CFH)/A$_{410}$(buffer only)*100].

Cofactor assay in fluid phase

Fluid phase assays were carried out with the following concentrations of components: FH1-4 (produced in mammalian cells), 23 nM (0.67 ng/μl); C3b, (Comptech), 35.1 nM (6.67 ng/μl) and factor I (FI) (Comptech), 25 nM (2.2 ng/μl). All dilutions were made in 150 mM NaCl Tris-buffered saline. A positive control used an equivalent mass of full-length purified FH (Comptech). A negative control used full-length FH but no CFI. Reactions were prepared on ice and with a refrigerated centrifuge. Time began when reactions were placed in a 37°C water bath. Samples were removed at 10, 20 and 30 min. Reactions were stopped with 3× reducing Laemmli buffer, mixed and heated at 95° C for 5 min. A volume of reaction equivalent to 45 ng of C3b was electrophoresed by SDS−PAGE and stained with Krypton dye. Gels were scanned on a Typhon variable mode imager with excitation at 530 and a 580 nm filter. Quantification was performed in ImageQuantTL. Krypton stain was found to be equally sensitive and more robust than silver stain or western blot using polyclonal anti-C3 antibodies.

Measuring CA on sheep erythrocytes

To test for CA, we resuspended washed EA-C3b cells to 2% (v/v) in AP buffer and incubated with an equal volume of a range of concentrations of CFH1-4$_{WT}$ and FH1-4$_{R53C}$ (produced in *Pichia*) and 2.5 μg/ml FI (CompTech) for 8 min at 25°C. After three washes in AP buffer, we mixed a 50 μl aliquot of cells

(2%) with 50 µl AP buffer containing FB (40 µg/ml) and fD (0.4 µg/ml) and then incubated for 15 min at 25°C to form the AP C3 convertase on the remaining C3b. We developed lysis by adding 50 µl of 4% (v/v) NHS-ΔBΔH in PBS/20 mM EDTA and incubating at 37°C for 10 min. We again calculated the percentage of inhibition of lysis in the presence of increasing concentrations of FH using $[A_{410}(\text{buffer only}) - A_{410}(\text{FH})/A_{410}(\text{buffer only}) * 100]$.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* online.

Conflict of Interest statement. None declared.

FUNDING

REFERENCES

1. Raychaudhuri, S., Iartchouk, O., Chin, K., Tan, P.L., Tai, A.K., Ripke, S., Gowrisankar, S., Vemuri, S., Montgomery, K., Yu, Y. *et al.* (2011) A rare penetrant mutation in CFH confers high risk of age-related macular degeneration. *Nat. Genet.*, **43**, 1232–1236.
2. Seddon, J.M., Yu, Y., Miller, E.C., Reynolds, R., Tan, P.L., Gowrisankar, S., Goldstein, J.I., Triebwasser, M., Anderson, H.E., Zerbib, J. *et al.* (2013) Rare variants in CFI, C3 and C9 are associated with high risk of advanced age-related macular degeneration. *Nat. Genet.*, **45**, 1366–1370.
3. Seddon, J.M., Reynolds, R., Maller, J., Fagerness, J.A., Daly, M.J. and Rosner, B. (2009) Prediction model for prevalence and incidence of advanced age-related macular degeneration based on genetic, demographic, and environmental variables. *Invest. Ophthalmol. Vis. Sci.*, **50**, 2044–2053.
4. Seddon, J.M., Reynolds, R., Yu, Y., Daly, M.J. and Rosner, B. (2011) Risk models for progression to advanced age-related macular degeneration using demographic, environmental, genetic, and ocular factors. *Ophthalmology*, **118**, 2203–2211.
5. Yu, Y., Reynolds, R., Rosner, B., Daly, M.J. and Seddon, J.M. (2012) Prospective assessment of genetic effects on progression to different stages of age-related macular degeneration using multistate Markov models. *Invest. Ophthalmol. Vis. Sci.*, **53**, 1548–1556.
6. Helgason, H., Sulem, P., Duvvari, M.R., Luo, H., Thorleifsson, G., Stefansson, H., Jonsdottir, I., Masson, G., Gudbjartsson, D.F., Walters, G.B. *et al.* (2013) A rare nonsynonymous sequence variant in C3 is associated with high risk of age-related macular degeneration. *Nat. Genet.*, **45**, 1371–1374.
7. Zhan, X., Larson, D.E., Wang, C., Koboldt, D.C., Sergeev, Y.V., Fulton, R.S., Fulton, L.L., Fronick, C.C., Branham, K.E., Bragg-Gresham, J. *et al.* (2013) Identification of a rare coding variant in complement 3 associated with age-related macular degeneration. *Nat. Genet.*, **45**, 1375–1379.
8. Sobrin, L., Maller, J.B., Neale, B.M., Reynolds, R.C., Fagerness, J.A., Daly, M.J. and Seddon, J.M. (2010) Genetic profile for five common variants associated with age-related macular degeneration in densely affected families: a novel analytic approach. *Eur. J. Hum. Genet.*, **18**, 496–501.
9. Morrison, A.C., Voorman, A., Johnson, A.D., Liu, X., Yu, J., Li, A., Muzny, D., Yu, F., Rice, K., Zhu, C. *et al.* (2013) Whole-genome sequence-based analysis of high-density lipoprotein cholesterol. *Nat. Genet.*, **45**, 899–901.
10. Farkas, M.H., Grant, G.R., White, J.A., Sousa, M.E., Consugar, M.B. and Pierce, E.A. (2013) Transcriptome analyses of the human retina identify unprecedented transcript diversity and 3.5 Mb of novel transcribed sequence via significant alternative splicing and novel genes. *BMC Genomics*, **14**, 486.
11. Klein, R.J., Zeiss, C., Chew, E.Y., Tsai, J.Y., Sackler, R.S., Haynes, C., Henning, A.K., SanGiovanni, J.P., Mane, S.M., Mayne, S.T. *et al.* (2005) Complement factor H polymorphism in age-related macular degeneration. *Science*, **308**, 385–389.
12. Haines, J.L., Hauser, M.A., Schmidt, S., Scott, W.K., Olson, L.M., Gallins, P., Spencer, K.L., Kwan, S.Y., Noureddine, M., Gilbert, J.R. *et al.* (2005) Complement factor H variant increases the risk of age-related macular degeneration. *Science*, **308**, 419–421.
13. Edwards, A.O., Ritter, R. 3rd, Abel, K.J., Manning, A., Panhuysen, C. and Farrer, L.A. (2005) Complement factor H polymorphism and age-related macular degeneration. *Science*, **308**, 421–424.
14. van de Ven, J.P., Nilsson, S.C., Tan, P.L., Buitendijk, G.H., Ristau, T., Mohlin, F.C., Nabuurs, S.B., Schoenmaker-Koller, F.E., Smailhodzic, D., Campochiaro, P.A. *et al.* (2013) A functional variant in the CFI gene confers a high risk of age-related macular degeneration. *Nat. Genet.*, **45**, 813–817.
15. Manuelian, T., Hellwage, J., Meri, S., Caprioli, J., Noris, M., Heinen, S., Jozsi, M., Neumann, H.P., Remuzzi, G. and Zipfel, P.F. (2003) Mutations in factor H reduce binding affinity to C3b and heparin and surface attachment to endothelial cells in hemolytic uremic syndrome. *J. Clin. Invest.*, **111**, 1181–1190.
16. Ferreira, V.P., Herbert, A.P., Cortes, C., McKee, K.A., Blaum, B.S., Esswein, S.T., Uhrin, D., Barlow, P.N., Pangburn, M.K. and Kavanagh, D. (2009) The binding of factor H to a complex of physiological polyanions and C3b on cells is impaired in atypical hemolytic uremic syndrome. *J. Immunol.*, **182**, 7009–7018.
17. Sanchez-Corral, P., Perez-Caballero, D., Huarte, O., Simckes, A.M., Goicoechea, E., Lopez-Trascasa, M. and de Cordoba, S.R. (2002) Structural and functional characterization of factor H mutations associated with atypical hemolytic uremic syndrome. *Am. J. Hum. Genet.*, **71**, 1285–1295.
18. Jozsi, M., Heinen, S., Hartmann, A., Ostrowicz, C.W., Halbich, S., Richter, H., Kunert, A., Licht, C., Saunders, R.E., Perkins, S.J. *et al.* (2006) Factor H and atypical hemolytic uremic syndrome: mutations in the C-terminus cause structural changes and defective recognition functions. *J. Am. Soc. Nephrol.*, **17**, 170–177.
19. Clark, S.J., Ridge, L.A., Herbert, A.P., Hakobyan, S., Mulloy, B., Lennon, R., Wurzner, R., Morgan, B.P., Uhrin, D., Bishop, P.N. *et al.* (2013) Tissue-specific host recognition by complement factor H is mediated by differential activities of its glycosaminoglycan-binding regions. *J. Immunol.*, **190**, 2049–2057.
20. Clark, S.J., Perveen, R., Hakobyan, S., Morgan, B.P., Sim, R.B., Bishop, P.N. and Day, A.J. (2010) Impaired binding of the age-related macular degeneration-associated complement factor H 402H allotype to Bruch's membrane in human retina. *J. Biol. Chem.*, **285**, 30192–30202.
21. Weismann, D., Hartvigsen, K., Lauer, N., Bennett, K.L., Scholl, H.P., Charbel Issa, P., Cano, M., Brandstatter, H., Tsimikas, S., Skerka, C. *et al.* (2011) Complement factor H binds malondialdehyde epitopes and protects from oxidative stress. *Nature*, **478**, 76–81.
22. Sjoberg, A.P., Trouw, L.A., Clark, S.J., Sjolander, J., Heinegard, D., Sim, R.B., Day, A.J. and Blom, A.M. (2007) The factor H variant associated with age-related macular degeneration (His-384) and the non-disease-associated

form bind differentially to C-reactive protein, fibromodulin, DNA, and necrotic cells. *J. Biol. Chem.*, **282**, 10894–10900.

23. Wu, J., Wu, Y.Q., Ricklin, D., Janssen, B.J., Lambris, J.D. and Gros, P. (2009) Structure of complement fragment C3b-factor H and implications for host protection by complement regulators. *Nat. Immunol.*, **10**, 728–733.

24. Servais, A., Noel, L.H., Roumenina, L.T., Le Quintrec, M., Ngo, S., Dragon-Durey, M.A., Macher, M.A., Zuber, J., Karras, A., Provot, F. *et al.* (2012) Acquired and genetic complement abnormalities play a critical role in dense deposit disease and other C3 glomerulopathies. *Kidney Int.*, **82**, 454–464.

25. Fakhouri, F., Roumenina, L., Provot, F., Sallee, M., Caillard, S., Couzi, L., Essig, M., Ribes, D., Dragon-Durey, M.A., Bridoux, F. *et al.* (2010) Pregnancy-associated hemolytic uremic syndrome revisited in the era of complement gene mutations. *J. Am. Soc. Nephrol.*, **21**, 859–867.

26. Pechtl, I.C., Kavanagh, D., McIntosh, N., Harris, C.L. and Barlow, P.N. (2011) Disease-associated N-terminal complement factor H mutations perturb cofactor and decay-accelerating activities. *J. Biol. Chem.*, **286**, 11082–11090.

27. Roversi, P., Johnson, S., Caesar, J.J., McLean, F., Leath, K.J., Tsiftsoglou, S.A., Morgan, B.P., Harris, C.L., Sim, R.B. and Lea, S.M. (2011) Structural basis for complement factor I control and its disease-associated sequence polymorphisms. *Proc. Natl Acad. Sci. USA*, **108**, 12839–12844.

28. Fang, C.J., Fremeaux-Bacchi, V., Liszewski, M.K., Pianetti, G., Noris, M., Goodship, T.H. and Atkinson, J.P. (2008) Membrane cofactor protein mutations in atypical hemolytic uremic syndrome (aHUS), fatal Stx-HUS, C3 glomerulonephritis, and the HELLP syndrome. *Blood*, **111**, 624–632.

29. Liszewski, M.K., Leung, M., Cui, W., Subramanian, V.B., Parkinson, J., Barlow, P.N., Manchester, M. and Atkinson, J.P. (2000) Dissecting sites important for complement regulatory activity in membrane cofactor protein (MCP; CD46). *J. Biol. Chem.*, **275**, 37692–37701.

30. Krych, M., Hauhart, R. and Atkinson, J.P. (1998) Structure-function analysis of the active sites of complement receptor type 1. *J. Biol. Chem.*, **273**, 8623–8629.

31. Krych, M., Clemenza, L., Howdeshell, D., Hauhart, R., Hourcade, D. and Atkinson, J.P. (1994) Analysis of the functional domains of complement receptor type 1 (C3b/C4b receptor; CD35) by substitution mutagenesis. *J. Biol. Chem.*, **269**, 13273–13278.

32. Legendre, C.M., Licht, C., Muus, P., Greenbaum, L.A., Babu, S., Bedrosian, C., Bingham, C., Cohen, D.J., Delmas, Y., Douglas, K. *et al.* (2013) Terminal complement inhibitor eculizumab in atypical hemolytic-uremic syndrome. *N. Engl. J. Med.*, **368**, 2169–2181.

33. Hillmen, P., Hall, C., Marsh, J.C., Elebute, M., Bombara, M.P., Petro, B.E., Cullen, M.J., Richards, S.J., Rollins, S.A., Mojcik, C.F. *et al.* (2004) Effect of eculizumab on hemolysis and transfusion requirements in patients with paroxysmal nocturnal hemoglobinuria. *N. Engl. J. Med.*, **350**, 552–559.

34. Yehoshua, Z., Rosenfeld, P.J. and Albini, T.A. (2011) Current clinical trials in dry AMD and the definition of appropriate clinical outcome measures. *Semin. Ophthalmol.*, **26**, 167–180.

35. Seddon, J.M., Sharma, S. and Adelman, R.A. (2006) Evaluation of the clinical age-related maculopathy staging system. *Ophthalmology*, **113**, 260–266.

36. Fritsche, L.G., Chen, W., Schu, M., Yaspan, B.L., Yu, Y., Thorleifsson, G., Zack, D.J., Arakawa, S., Cipriani, V., Ripke, S. *et al.* (2013) Seven new loci associated with age-related macular degeneration. *Nat. Genet.*, **45**, 433–439, 439e431–432.

37. Yu, Y., Bhangale, T.R., Fagerness, J., Ripke, S., Thorleifsson, G., Tan, P.L., Souied, E.H., Richardson, A.J., Merriam, J.E., Buitendijk, G.H. *et al.* (2011) Common variants near FRK/COL10A1 and VEGFA are associated with advanced age-related macular degeneration. *Hum. Mol. Genet.*, **20**, 3699–3709.

38. Li, H. and Durbin, R. (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, **25**, 1754–1760.

39. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G. and Durbin, R. (2009) The sequence alignment/map format and SAMtools. *Bioinformatics*, **25**, 2078–2079.

40. Koboldt, D.C., Zhang, Q., Larson, D.E., Shen, D., McLellan, M.D., Lin, L., Miller, C.A., Mardis, E.R., Ding, L. and Wilson, R.K. (2012) VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res.*, **22**, 568–576.

41. McLaren, W., Pritchard, B., Rios, D., Chen, Y., Flicek, P. and Cunningham, F. (2010) Deriving the consequences of genomic variants with the Ensembl API and SNP effect predictor. *Bioinformatics*, **26**, 2069–2070.

42. Abecasis, G.R., Auton, A., Brooks, L.D., DePristo, M.A., Durbin, R.M., Handsaker, R.E., Kang, H.M., Marth, G.T. and McVean, G.A. (2012) An integrated map of genetic variation from 1,092 human genomes. *Nature*, **491**, 56–65.

43. Tennessen, J.A., Bigham, A.W., O'Connor, T.D., Fu, W., Kenny, E.E., Gravel, S., McGee, S., Do, R., Liu, X., Jun, G. *et al.* (2012) Evolution and functional impact of rare coding variation from deep sequencing of human exomes. *Science*, **337**, 64–69.

44. Adzhubei, I.A., Schmidt, S., Peshkin, L., Ramensky, V.E., Gerasimova, A., Bork, P., Kondrashov, A.S. and Sunyaev, S.R. (2010) A method and server for predicting damaging missense mutations. *Nat. Methods*, **7**, 248–249.

45. Sim, N.L., Kumar, P., Hu, J., Henikoff, S., Schneider, G. and Ng, P.C. (2012) SIFT web server: predicting effects of amino acid substitutions on proteins. *Nucleic Acids Res.*, **40**, W452–W457.

46. Lathrop, G.M., Lalouel, J.M., Julier, C. and Ott, J. (1984) Strategies for multilocus linkage analysis in humans. *Proc. Natl Acad. Sci. USA*, **81**, 3443–3446.

47. Tortajada, A., Montes, T., Martinez-Barricarte, R., Morgan, B.P., Harris, C.L. and de Cordoba, S.R. (2009) The disease-protective complement factor H allotypic variant Ile62 shows increased binding affinity for C3b and enhanced cofactor activity. *Hum. Mol. Genet.*, **18**, 3452–3461.

[173]

Whole exome sequencing identifies variants in families with macular degeneration

Yi Yu, [1, 13] Michael P. Triebwasser, [2, 13] Edwin K. S. Wong, [3, 13]

Elizabeth C. Schramm, [2, 13] Brett Thomas, [4] Robyn Reynolds, [1] Elaine Mardis, [5] John P. Atkinson,
[2] Mark Daly, [4,6] Soumya Raychaudhuri, [6-10] David Kavanagh, [3] Johanna M. Seddon[1,11, 12]

[1]Ophthalmic Epidemiology and Genetics Service, New England Eye Center,
    Tufts Medical Center, Boston, Massachusetts, USA.
[2]Division of Rheumatology, Department of Medicine, Washington University
    School of Medicine, St. Louis, MO, USA
[3]Institute of Genetic Medicine , Newcastle University, International Centre for Life, Newcastle upon
    Tyne, UK
[4]Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, Massachusetts,
    USA.
[5]The Genome Institute at Washington University, Saint Louis, MO, USA
[6]Program in Medical and Population Genetics, Broad Institute, Cambridge, Massachusetts, USA
[7]Partners HealthCare Center for Personalized Genetic Medicine, Boston, Massachusetts, USA.
[8]Division of Genetics, Brigham and Women's Hospital, Boston, Massachusetts, USA.
[9]Division of Rheumatology, Immunology, and Allergy, Brigham and Women's Hospital, Boston,
    Massachusetts, USA.
[10]Faculty of Medical and Human Sciences, University of Manchester, Manchester, UK
[11]Department of Ophthalmology, Tufts University School of Medicine, Boston,
    Massachusetts, USA.
[12]Sackler School of Graduate Medical Sciences, Tufts University, Boston, MA, USA
[13]These authors contributed equally to this work.

Correspondence: jseddon@tuftsmedicalcenter.org

Supplementary Table S1-Other potential pathogenic variants suggested by XBrowse.

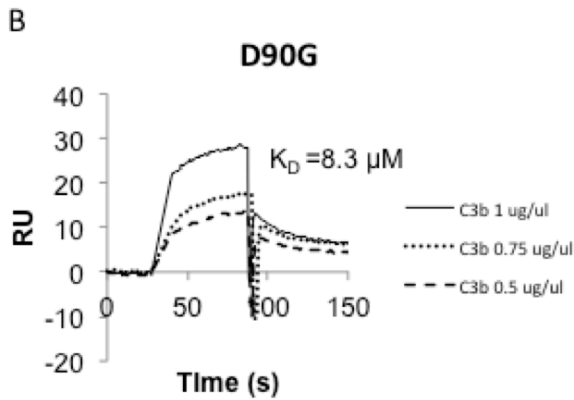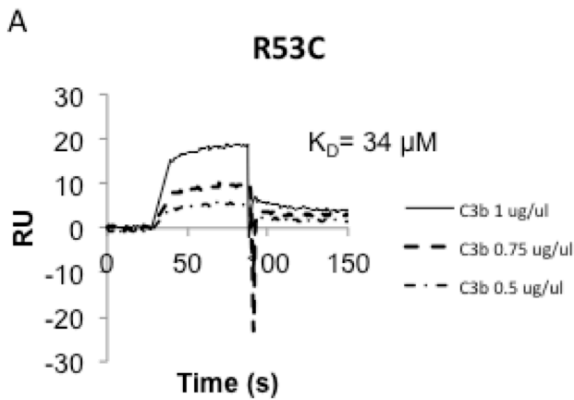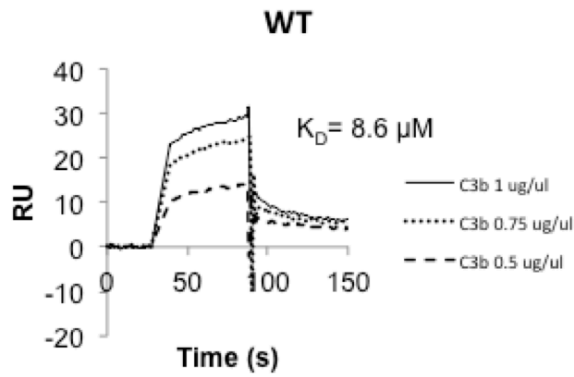| FAM	Genes	Location	REF	ALT	RSID	Amino Acid	Allele Frequencies in ESP	LOD score	Polyphen2	SIFT
					Variants predicted to be both probably damaging by PolyPhen2 and deleterious by SIFT					
I	MUC4	chr3:195490456	T	G	.	N4573H	0	-0.90	prob.dam.	deleterious
IV	CCT4	chr2:62106083	A	C	.	L148W	0.0002	-5.40	prob.dam.	deleterious
IV	CNTN3	chr3:74316451	G	A	.	A928V	0	-7.89	prob.dam.	deleterious
IV	ALDH1L1	chr3:125865723	G	A	rs144099397	S344F	0.0005	-1.80	prob.dam.	deleterious
VI	MFN2	chr1:12067297	A	G	.	Q687R	0	0.65	prob.dam.	deleterious
VI	TAB2	chr6:149700430	C	T	.	T460M	0	0.65	prob.dam.	deleterious
VI	NOL11	chr17:65716047	A	T	.	D94V	0	0.65	prob.dam.	deleterious
VIII	AGT	chr1:230841679	C	T	rs74315283	R375Q	0	-1.80	prob.dam.	deleterious
VIII	EPSTI1	chr13:43543310	G	A	.	A84V	0.0001	-1.80	prob.dam.	deleterious
IX	GRM7	chr3:6903211	G	C	.	E46Q	0	-0.30	prob.dam.	deleterious
IX	MTMR14	chr3:9730694	C	T	.	P454L	0	-0.30	prob.dam.	deleterious
IX	CDAN1	chr15:43020878	C	T	rs200401359	G926R	0	-0.30	prob.dam.	deleterious
IX	ZNF836	chr19:52659758	T	A	.	K393M	0	-0.30	prob.dam.	deleterious
					Variants predicted to be either probably damaging by PolyPhen2 or deleterious by SIFT					
I	KIF20B	chr10:91477454	G	A	rs149456198	V416I	0.0005	1.20	benign	deleterious
II	PDLIM7	chr5:176917023	C	T	.	R217H	0	1.22	NA	deleterious
III	CLASRP	chr19:45561058	T	C	.	V172A	0	-0.80	prob.dam.	tolerated
III	TULP4	chr6:158923961	C	T	rs140116628	T1089M	0.0003	-4.40	benign	deleterious
VI	ARHGEF10L	chr1:17961406	G	C	.	V608L	0	0.90	poss.dam.	deleterious
VI	TPO	chr2:1546296	C	T	.	R908C	0.0001	0.90	poss.dam.	deleterious
VII	OR1C1	chr1:247920937	C	T	.	V258I	0.0001	-3.06	poss.dam.	deleterious
VII	DDHD2	chr8:38097798	C	A	.	P210T	0.0002	-2.65	poss.dam.	deleterious
VII	SYT8	chr11:1857324	C	A	.	A156D	0	-8.40	poss.dam.	deleterious
VIII	ABCA9	chr17:67045529	G	C	rs150105567	R67G	0.0006	0.65	poss.dam.	deleterious
VIII	DNAH17	chr17:76482076	G	A	rs181353842	P2414L	0.0008	0.65	benign	deleterious
VIII	IGLC7	chr22:23265006	C	T	.	S81F	0	0.65	poss.dam.	deleterious
IX	GRM7	chr3:6903211	G	C	.	E46Q	0	0.65	prob.dam.	tolerated
IX	HLX	chr1:221053633	G	A	.	G1935D	0.0001	0.65	benign	deleterious
IX	SLC25A47	chr14:100795151	C	T	rs201454370	S139L	0.0004	0.65	benign	deleterious
IX	LRBA	chr4:151357949	T	A	rs147096866	D2294V	0.0002	0.65	poss.dam.	deleterious

FAM, family; Ref, reference allele; Alt, alternative allele; rsID, rs number; ESP, exome sequencing project; prob.dam., probably damaging; poss.dam., possibly damaging.

[175]
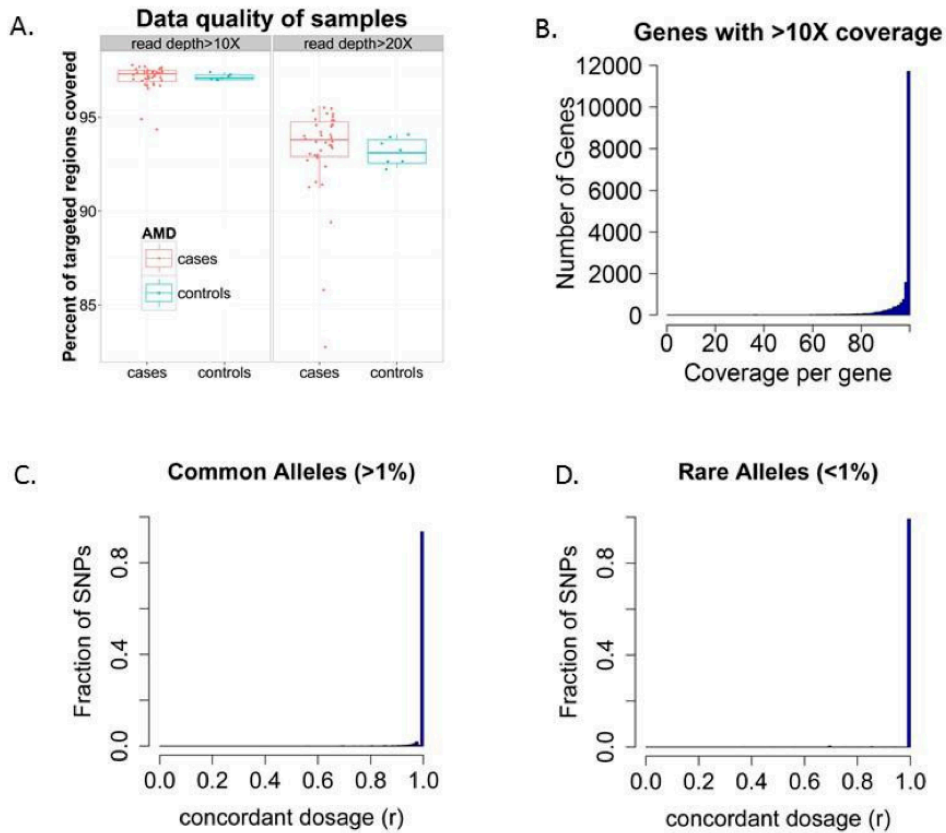
**Supplemental Figure 1. Exome sequencing coverage and genotypes concordance.**

(A) All sequenced samples were required to have over 10X coverage at greater than 90% of the targeted regions(median = 97.25%) and over 20X coverage at greater than 80% of the targeted regions(median = 93.7%); (B) Distribution of the percent regions covered at >10X depth for each gene; (C,D) Histogram of correlations between minor allele dosages at 2,426 SNPs as determined by sequence-based and exome-array-based genotyping for common alleles (>1% frequency) and rare alleles (<1% frequency).

WT

$K_D = 8.6\ \mu M$

— C3b 1 ug/ul
······· C3b 0.75 ug/ul
– – – C3b 0.5 ug/ul

A

R53C

$K_D = 34\ \mu M$

— C3b 1 ug/ul
– – C3b 0.75 ug/ul
– · – C3b 0.5 ug/ul

B

D90G

$K_D = 8.3\ \mu M$

— C3b 1 ug/ul
······· C3b 0.75 ug/ul
– – – C3b 0.5 ug/ul

C

Supplemental Figure 2. CFHR53C, but not CFHD90G, demonstrates a weaker affinity for C3b compared to CFHWT.

Overlaying sensograms show the steady state response for the binding of C3b (1-5 µM) to CFH1-4 proteins immobilized on a CM5 sensor chip. Response was plotted against concentration and the KD was calculated using the 1:1 binding model in the BIAeval software.
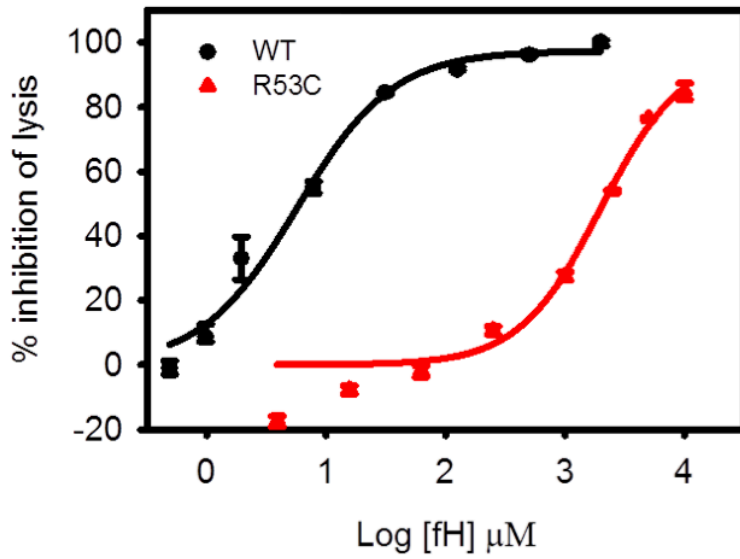
Supplemental Figure 3. Hemolytic experiments confirm R53C's d e c a y accelerating activity. In assays using sheep erythrocyte lysis, the decay defect of R53C is clear. 50% inhibition of lysis was achieved using 6.0 nM of WT. In contrast, 2000 nM (>300-fold more) was required using R53C.
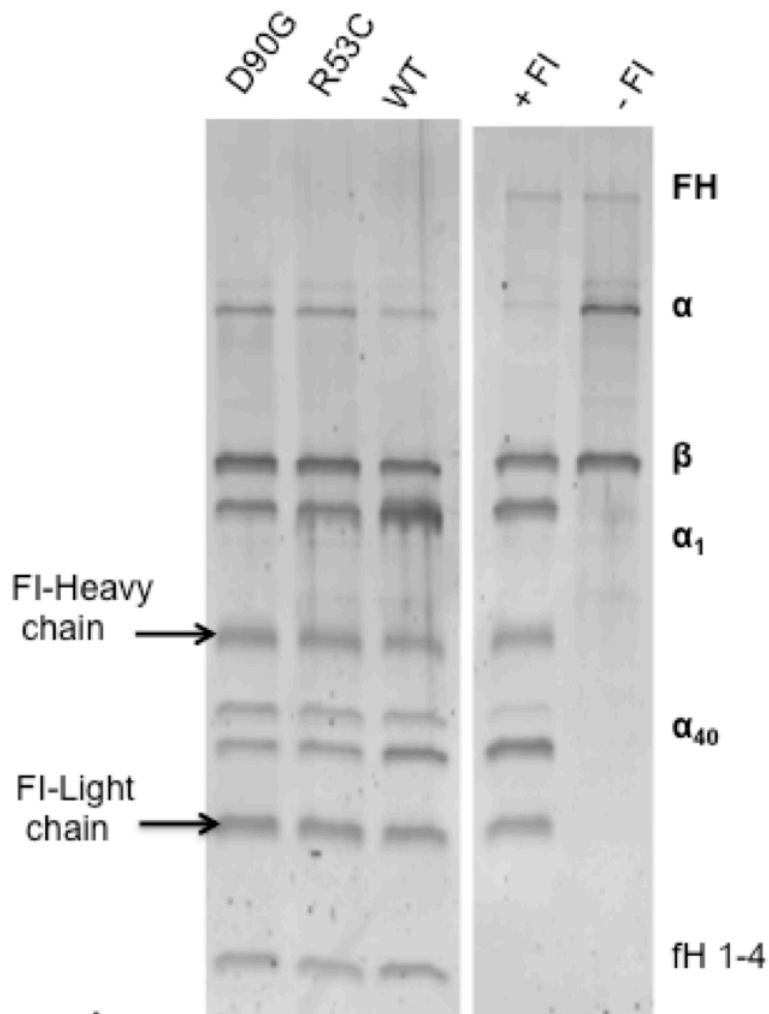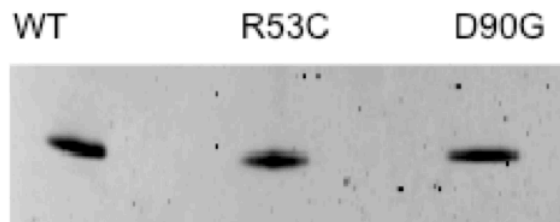
Supplemental Figure 4. Representative gel of fluid phase cofactor assay.

(A) Representative gel shows that both mutants clearly fail to cleave α' at the same rate as WT and that as a result more α'remains and less α_1 and less α_{40} are generated (10% Tris-Gly SDS). (B) Equal amounts of each cofactor protein was present in each reaction (separate 12% Tris-Gly SDS).

Supplemental Manuscript 3

Salmon, JE, Heuser, C, Triebwasser, MP, et. al. 2012

PLoS Medicine

PLoS MEDICINE

Mutations in Complement Regulatory Proteins Predispose to Preeclampsia: A Genetic Analysis of the PROMISSE Cohort

Jane E. Salmon[1]*, Cara Heuser[2], Michael Triebwasser[3], M. Kathryn Liszewski[3], David Kavanagh[4], Lubka Roumenina[5], D. Ware Branch[2], Tim Goodship[4], Veronique Fremeaux-Bacchi[5], John P. Atkinson[3]

1 Autoimmunity and Inflammation Program, Hospital for Special Surgery, Cornell Weill Medical College, New York, New York, United States of America, 2 Department of Obstetrics and Gynecology, University of Utah, Salt Lake City, Utah, United States of America, 3 Department of Medicine/Division of Rheumatology, Washington University School of Medicine, St. Louis, Missouri, United States of America, 4 Institute of Human Genetics, Newcastle University, Newcastle upon Tyne, United Kingdom, 5 Assistance Publique-Hopitaux de Paris, Höpital Européen Georges-Pompidou, Service d'Immunologie Biologique, Paris, France

Abstract

Background: Pregnancy in women with systemic lupus erythematosus (SLE) or antiphospholipid antibodies (APL Ab)—autoimmune conditions characterized by complement-mediated injury—is associated with increased risk of preeclampsia and miscarriage. Our previous studies in mice indicate that complement activation targeted to the placenta drives angiogenic imbalance and placental insufficiency.

Methods and Findings: We use PROMISSE, a prospective study of 250 pregnant patients with SLE and/or APL Ab, to test the hypothesis in humans that impaired capacity to limit complement activation predisposes to preeclampsia. We sequenced genes encoding three complement regulatory proteins—membrane cofactor protein (MCP), complement factor I (CFI), and complement factor H (CFH)—in 40 patients who had preeclampsia and found heterozygous mutations in seven (18%). Five of these patients had risk variants in MCP or CFI that were previously identified in atypical hemolytic uremic syndrome, a disease characterized by endothelial damage. One had a novel mutation in MCP that impairs regulation of C4b. These findings constitute, to our knowledge, the first genetic defects associated with preeclampsia in SLE and/or APL Ab. We confirmed the association of hypomorphic variants of MCP and CFI in a cohort of non-autoimmune preeclampsia patients in which five of 59 were heterozygous for mutations.

Conclusion: The presence of risk variants in complement regulatory proteins in patients with SLE and/or APL Ab who develop preeclampsia, as well as in preeclampsia patients lacking autoimmune disease, links complement activation to disease pathogenesis and suggests new targets for treatment of this important public health problem.

Study Registration: ClinicalTrials.gov NCT00198068

Please see later in the article for the Editors' Summary.

[179]

Introduction

Preeclampsia complicates 4% 5% of all pregnancies worldwide, causing significant maternal and neonatal mortality, and claims the lives of >60,000 mothers each year in developing countries [1]. Although typically diagnosed with the onset of hypertension and proteinuria after 20 weeks' gestation, the syndrome begins earlier in pregnancy with abnormal placental development. The specific "anatomical" defect, the failure of uterine spiral arteries to remodel into dilated, flaccid vessels, leads to underperfusion of the intervillous space and placental hypoxia. The clinical manifestations of preeclampsia represent the maternal response to an excess of antiangiogenic factors released by the hypoperfused placenta. These factors include vasculopathic factors such as soluble fms-like tyrosine kinase 1 (sFlt-1), a potent vascular endothelial growth factor (VEGF) antagonist, and soluble endoglin, an inhibitor of TGF-β signaling [2,3]. Disease manifestations range from mild blood pressure elevations to severe hypertension, the HELLP syndrome (hemolysis, elevated liver enzymes, and low platelets), or eclampsia (seizures). The molecular basis for placental dysregulation is unknown and treatment is limited. Higher rates of preeclampsia in sisters, daughters, and mothers of affected women suggests a genetic contribution to risk for disease, and evidence for paternal components of predisposition supports the hypothesis that the genotype of the fetus contributes to the overall risk of preeclampsia [4].

Immunologic maladaption has been proposed as a pathogenic mechanism for preeclampsia, but remains an unproven theory. We have previously established a link between the complement system and angiogenic factor imbalance associated with placental dysfunction (Figure 1). Studies in mouse models of pregnancy indicate that complement activation targeted to the placenta drives angiogenic imbalance, placental insufficiency, and endothelial injury [5]. In our experimental model, complement C5a-C5a receptor interactions trigger release of sFlt-1. This antiangiogenic factor is associated with hypertension, proteinuria, and glomerular endotheliosis in rodents [2,3,6] and is elevated in the circulation of pregnant women destined for preeclampsia. Further, in mice with abnormal pregnancies, the alternative complement pathway amplifies fetoplacental injury [5]. Also, uncontrolled complement activation due to deficiency of a membrane complement regulator analogous to human MCP is embryonic lethal because of maternal complement-triggered damage, and pregnancies are rescued in mice with reduced alternative pathway function [7]. Elevated levels of the alternative pathway complement activation fragment Bb in the first 20 weeks of pregnancy was recently found to be independently associated with preeclampsia later in pregnancy [8] and underscores the importance of complement, particularly the alternative pathway, in this syndrome.

Pregnancy in women with systemic lupus erythematosus (SLE) or the antiphospholipid syndrome (APS), two autoimmune diseases characterized by complement-mediated injury, is associated with an increased risk of preeclampsia, placental insufficiency, fetal growth restriction, and miscarriage [9]. In these patients, autoantibodies targeted to the placenta initiate local complement activation which, if unabated, leads to abnormal placental development (Figure 1). Indeed, complement activation products have been demonstrated in deciduas, chorionic villi, and vessel walls in the placentas of patients with APL Ab and in patients with preeclampsia [10,11]. Moreover, evidence of alternative pathway activation has been found in the circulation of pregnant SLE patients. In normal pregnancies, excessive complement activation is prevented by complement regulatory proteins that are highly expressed on trophoblast membranes (membrane cofactor protein

[MCP (CD46)], decay accelerating factor [DAF (CD55)], and CD59) as well as circulating complement regulatory proteins (complement factor H [CFH], C4b binding protein, and complement factor I [CFI]) (Figure 1) [11]. We hypothesized that impaired capacity to limit complement activation predisposes pregnant women with SLE or APL Ab to preeclampsia.

Loss-of-function mutations in CFH, MCP, and CFI or gain-of-function mutations in complement components factor B and C3 lead to undesirable complement activation in patients with atypical hemolytic uremic syndrome (aHUS), a microangiopathy characterized by microvascular endothelial activation, cell injury, and thrombosis [12,13,14,15]. Complement dysregulation (specifically, enhanced function of the alternative pathway) contributes to the pathology of aHUS, a disease associated with glomerular endothelial cell injury and microthrombi, not unlike the pathologic findings in preeclampsia, APS, and SLE. Complement mutations in aHUS demonstrate incomplete penetrance (~50%), indicating that additional genetic and environmental factors are required to manifest disease. Pregnancy, SLE, and/or APL Ab may provide these additional factors. The aims of this study are to translate our findings in mouse models to patients and determine whether variations in complement regulatory proteins leading to impaired
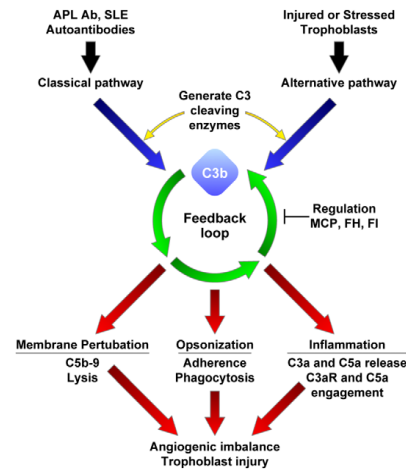

Figure 1. Complement activation in preeclampsia. Black and blue arrows indicate triggering of complement cascades by autoantibodies or damaged self-tissue; green arrows, deposition of C3b on a target sets in motion the powerful amplification loop of the alternative pathway; red arrows, effector activities of complement are generated by C3b deposition and C3a release and the downstream mediators C5b-9 and C5a. Regulation of feedback loop of alternative pathway on the placental trophoblasts (or endothelial cells) occurs through limited proteolytic cleavage of C3b to generate iC3b. This reaction is carried out by a serine protease, factor I (FI), along with membrane cofactor protein (MCP) or CFH (FH). Because SLE and APS are characterized by autoantibodies that trigger the classical pathway, defective regulation of C4b (a component of the classical pathway C3 convertase) by MCP is also likely to influence the severity of tissue injury and risk for preeclampsia. If regulators such as MCP, CFI (FI), or CFH (FH) are dysfunctional, excessive complement activation occurs. This may result in placental damage, thrombobosis and release of antiangiogenic factors, culminating in preeclampsia.
doi:10.1371/journal.pmed.1001013.g001

[180]

function are associated with preeclampsia in patients with SLE and/or APL Ab and whether such variants are also present in women who develop preeclampsia in the absence of autoimmune disease.

Methods

Participants

The first group of patients studied were enrolled in The PROMISSE Study (**P**redictors of p**R**egnancy **O**utcome: bio**M**arkers **I**n antiphospholipid antibody **S**yndrome and **S**ystemic lupus **E**rythematosus), a prospective multi-center observational study to identify predictors of pregnancy outcome initiated in 2003. Patients with SLE (defined as ≥4 ACR criteria) and/or APL Ab (defined as at least one of following documented twice between 6 wk and 5 y apart: ACL [IgG or IgM ≥40 units], lupus anticoagulant, or anti-β2GPI [IgG or IgM ≥40 units]) and disease controls were recruited by 12 wk gestation. At screening, patients with renal disease (proteinuria >1,000 mg/24 h, RBC casts, or serum creatinine >1.2 mg/dl), taking prednisone >20 mg/d, diabetes mellitus (Type I and Type II antedating pregnancy), hypertension (blood pressure ≥140/90 mmHg), or multiple gestations were excluded. Detailed medical and obstetrical information and serial blood specimens were obtained monthly during the course of pregnancy. Preeclampsia was defined as new onset of elevated systolic blood pressure (≥140 mmHg) and/or elevated diastolic blood pressure (≥90 mmHg) after 20 wk gestation on two occasions at least 4 h apart and proteinuria of 300 mg or greater in a 24 h urine specimen or ≥1+ on dipstick at least 4 h apart in the absence of pyelonephritis or hematuria. Severe preeclampsia and HELLP are defined by standard criteria [16]. In women with proteinuria present in early gestation, preeclampsia was defined as an increase in proteinuria to double their baseline level. For patients with a history of preeclampsia in previous pregnancies, medical charts were reviewed to document that patients met the above criteria above.

Of 250 enrolled patients with SLE and/or APL Ab who completed their pregnancies before 31 May 2008, 30 patients developed preeclampsia during the study and 10 additional patients had documented preeclampsia in a previous pregnancy (Table 1). Three of the 30 patients with preeclampsia during the PROMISSE Study pregnancy had documented preeclampsia in a previous pregnancy. None of the disease-free controls developed preeclampsia. The racial/ethnic distribution of the 40 patients with preeclampsia was 67% white, 17% black, 7% Asian, 5% Hispanic, and 4% other.

The Institutional Review Board at each of the PROMISSE Study sites (Hospital for Special Surgery, NY, NY, US; NYU School of Medicine, NY, NY, US; Johns Hopkins University, Department of Rheumatology, Johns Hopkins Medical Institu-

tions, Baltimore, MD, US; Mount Sinai Hospital, Toronto, Canada; University of Utah, Intermountain Health Care, Salt Lake City, UT, US; Oklahoma Medical Research Foundation, University of Oklahoma Health Sciences Center, Oklahoma City, OK, US) gave approval for human participation for patients. Written informed consent that included the use of material for genetic research was obtained from all PROMISSE participants who contributed samples toward this study. All study samples were deidentified.

The second group of patients studied was from the Universal Samples Database at University of Utah, which stores maternal blood and clinical data from women who present to labor and delivery. Charts were reviewed to confirm the diagnosis of severe preeclampsia and HELLP defined according to established criteria, and 59 patients meeting criteria and with available DNA were identified [16]. The racial/ethnic distribution of patients with severe preeclampsia and/or HELLP was 74% white (non-Hispanic), 25% Hispanic, and 1% other. Control pregnancies defined as uncomplicated term deliveries in women with no more than one pregnancy loss and at least one live birth were identified in the Universal Samples Database at University of Utah. The Institutional Review Board at University of Utah, Salt Lake City, UT, gave approval for human participation for the participants whose samples are stored at the Universal Samples Database at University of Utah. Samples from 143 ethnically matched controls were available.

The allele frequency of MCP A304V and CFI IVS12+5 G>T variants in white individuals were determined using DNA from The UK Blood Services collection of Common Controls (UKBS collection) which forms part of the Wellcome Trust Case Control Consortium (WTCCC) (http://www.wtccc.org.uk) [17,18].

Genetic and Statistical Analyses

Genomic DNA was extracted from peripheral blood leukocytes, and coding sequence and intronic flanking regions of MCP, CFH, and CFI were screened as described [19,20]. Mutations considered for our analysis were either missense or nonsense coding mutations and mutations in conserved positions of splice sites and promoters. To further restrict our study to mutations that are likely deleterious, we ignored mutations present in dbSNP (http://www.ncbi.nlm.nih.gov/projects/SNP/), except for those previously observed in aHUS. This left novel variants and known aHUS alleles with the described properties. Mutations that occur in less than 5% of the population are termed low-frequency variants, and those in <0.5% are rare variants [21].

The rare variants could not be tested for Hardy-Weinberg equilibrium (HWE) because they do not appear in controls. The A304V variant was tested and found to be in HWE in 40 PROMISSE and 59 Utah cases ($p = 1.0$) and in the WTCCC

Table 1. Preeclampsia in autoimmune PROMISSE patients.

Clinical Characteristics	No. Enrolled	No. with Preeclampsia in PROMISSE	No. with Preeclampsia in Previous Pregnancy[a]	Total Patients with Preeclampsia[b]	% Patients with Preeclampsia
SLE only	149	19	5	23	15%
APL Ab only	59	6	5	10	17%
SLE with APL Ab	32	5	3	7	22%
Total	**250**	**30**	**13**	**40**	**17%**

[a]The PROMISSE Study pregnancy was the first pregnancy in 40% of the SLE patients (with and without APL Ab) and in 9% for APL Ab patients without SLE. The APL Ab patients differed from SLE patients with regard to multiparity because pregnancy loss is a criterion for testing for APL Ab.
[b]Three patients had preeclampsia in the PROMISSE Study and also had preeclampsia in previous pregnancies.
doi:10.1371/journal.pmed.1001013.t001

[181]

controls ($p = 0.15$). The 34 PROMISSE controls could not be assessed for HWE because no alleles were present, and the 143 Utah controls were found not to be in HWE, largely because of a single homozygote that was sampled ($p = 0.02$).

To test the hypothesis that A304V is associated with pree-clampsia, we combined cases and controls from the PROMISSE and Utah studies. These two populations were combined because the frequencies of A304V were similar, and Fisher's exact test was used to test for enrichment of a deleterious, dominant allele.

Expression and Functional Assessment of MCP

Substitution mutagenesis was performed using the QuikChange site-directed mutagenesis kit (Stratagene). The MCP template was the MCP isoform BC1 (GenBank accession no. X59405) cloned into the EcoRI site of plasmid pCDNA3.1+ (Stratagene). Transient transfections were performed with Fugene-6 (Roche Molecular Biochemicals) into Chinese hamster ovary (CHO) K1 cells. After cell lysis, MCP was quantified in an ELISA as described [22]. Briefly, MCP mAb GB24 was coated at 2.5 µg/ml on microtiter wells followed by blocking in PBS containing 4% BSA and 0.05% Tween 20. Cell lysates and standards were incubated in wells followed by reaction with a rabbit anti-MCP antiserum, horseradish- peroxidase labeled donkey anti-rabbit IgG (Jackson ImmunoResearch), and ImmunoPure TMB Sub-strate kit (Pierce).

Ligand binding and cofactor assays for C4b and C3b have been described [22]. Briefly, an ELISA format was used for character-izing ligand binding. C3b or C4b (Complement Technologies) was coated onto wells of a microtiter plate. Next, dilutions of MCP wild-type or mutant CHO lysates were added. Cofactor assays employed biotinylated ligands in low- or physiologic salt buffer (i.e., 10 mM Tris [pH 7.4], with 25 or 150 mM NaCl), CFI (100 ng), and cell lysates (25 pg of wild-type or mutant MCP). After 20 and 90 min, cleavage fragments were analyzed by a 10% size fractionation on reducing SDS/PAGE followed by transfer, Western blotting, and densitometric scanning [22].

Results

Patients with Preeclampsia

The PROMISSE Study is a prospective multi-center observa-tional study in which patients with SLE and/or APL Ab were enrolled at 12 wk gestation and followed through pregnancy. Patients with renal insufficiency, diabetes mellitus (antedating pregnancy), hypertension, or multiple gestations (risk factors predisposing to preeclampsia) were excluded. Between 2003 and 2008, 250 patients with SLE and/or APL Ab completed their pregnancies, 40 of whom developed preeclampsia during the study or had preeclampsia in a previous pregnancy (Table 1). All patients with preeclampsia were screened for mutations in the genes encoding the complement regulatory proteins MCP, CFI, and CFH. Mutations that occur in less than 5% of the population (low-frequency alleles) are termed uncommon variants [21].

We identified heterozygous mutations in complement system pathway genes in seven of 40 patients with preeclampsia (18%) four patients with mutations in MCP, two with CFI mutations, and one with a CFH mutation. Complement C3 levels were normal at enrollment in these patients. We did not find any mutations in 34 patients without preeclampsia (15 matched for disease [SLE and/or APL Ab], age, and ethnicity, and 19 healthy control patients matched for age and ethnicity) (7/40 versus 0/34, $p < 0.01$; also see results of mutation screening described for each gene in the sections to follow). We reject the null hypothesis of equal or fewer rare, deleterious alleles in controls at the 0.02 level

($p = 0.0105$ for the one-tailed test). This approach represents a variation of existing collapsing methods for the analysis of low-frequency variants [23]. The frequency of these mutations and variants in other healthy populations is described below. The clinical characteristics of patients with mutations are presented in Table 2.

To replicate the relationship of complement system variants with preeclampsia and to determine whether mutations in complement regulatory proteins are present in non-autoimmune patients with severe preeclampsia and/or HELLP, we screened 59 patients and 143 ethnically matched healthy controls with normal pregnancies from the Universal Samples Database at the University of Utah for mutations in MCP, CFI, and CFH. We identified heterozygous mutations in the genes of five patients four in MCP and one in CFI. No significant sequence variants were detected in CFH. The clinical characteristics of the patients from this cohort are presented in Table 3. Three individuals among the healthy controls had the A304V variant (5/59 versus 3/143, $p < 0.05$). Details of the mutations and variants and their distribution are described below.

Mutations and Variants of MCP

MCP is a widely expressed transmembrane protein that binds C3b and C4b deposited on cell surfaces and serves as a cofactor for their cleavage by CFI. The resulting attached fragments, iC3b and C4d, are not capable of forming convertases. By cleaving C3b, MCP is a key regulator on host cells of the amplification loop of the alternative pathway. Four patients with preeclampsia (10%) had mutations in MCP. Three patients were heterozygous for MCP A304V, a low-frequency hypomorphic variant previously shown to be deficient in control of alternative pathway activation on the cell surface [14]. The allele occurred at a frequency of 3.7% in the PROMISSE Study patients with preeclampsia. All three patients with A304V were white, and one had SLE only, one APL Ab only, and one both SLE and APL Ab. Two patients had severe preeclampsia and two patients reported prior pregnancy compli-cations, including preeclampsia (Table 2). This variant has been identified in four of 238 patients from the French multi-center aHUS cohort and in two of 181 healthy French controls, and in none of 120 white participants from another study [12]. Because the frequency of MCP A304V was not reported in publically available databases, including HapMap, we genotyped a large population of whites. In 1,529 individuals from the WTCCC, the allele frequency was 2.1% (61 heterozygotes and two homozy-gotes). Of note, no pregnancy histories or other medical information was available for these individuals.

We confirmed the association of MCP A304V with preeclamp-sia in a second cohort identified from the Universal Samples Database at the University of Utah. In 59 non-autoimmune patients with severe preeclampsia and/or HELLP, we identified four individuals (6.8% of patients) heterozygous for MCP A304V (allele frequency 3.4%). The details of their pregnancies are presented on Table 3. In 144 ethnically and geographically matched controls with uncomplicated pregnancies, two individuals were heterozygous and one homozygous for MCP A304V (allele frequency 1.4%), none of whom had histories of thrombotic microangiopathy or preeclampsia. The frequency of MCP A304V was consistent with our findings in the WTCCC. Combining our autoimmune and non-autoimmune cohorts with preeclampsia ($n = 99$), we identified seven individuals heterozygous for MCP A304. In both groups separately and when combined (99 cases versus 178 controls), A304V confers risk of preeclampsia. The variant acts in a dominant fashion with an odds ratio of 4.4 (95% confidence interval 1.1 17.6; $p = 0.027$).

[182]

Table 2. Clinical characteristics of autoimmune patients with complement regulatory protein mutations.

Mutation/Variant	Autoimmune Disease	GA at Delivery (wk)	Other Pregnancies
MCP/A304V	SLE	35.4[a]	3 early losses, 1 live birth
MCP/A304V	SLE, APL Ab	29.0[b]	None
MCP/A304V	APL Ab	37.3	2 live births, 1 with preeclampsia
MCP/K32N	SLE	25.0	2 early losses, 1 live birth
CFI/IVS12+5 G>T	APL Ab	29.0[c]	1 live birth, preterm at 24 weeks
CFI/IVS12+5 G>T	SLE	37.5	1 miscarriage; 1 elective termination
CFH/S40A	SLE	37.4	None

Patients were white, except the patient with MCP K32N, who was black. Characteristics of preeclampsia [16].
[a]Severe preeclampsia with visual disturbance.
[b]Preeclampsia with oligohydramnios and intrauterine growth restriction (<5th percentile).
[c]Severe preeclampsia, HELLP syndrome.
GA, gestational age.
doi:10.1371/journal.pmed.1001013.t002

We have previously studied in detail the complement inhibitory profile of A304V employing permanent CHO cell lines expressing an equal copy number of wild-type or A304V. While there was no detectable difference in ELISA type assays of solubilized protein, in situ the A304V was ~50% as effective at controlling C3b deposition by the alternative pathway [14]. In addition, a 25%–50% reduction in levels of wild-type complement regulatory proteins expressed by CHO cells impairs regulation of the alternative pathway, showing that haploinsufficiency has functional consequences [14,24].

The potential role of complement dysregulation in preeclampsia is underscored by our discovery of a novel mutation in exon 2 of the gene encoding MCP (K32N) in a patient from the PROMISSE Study who had SLE. Her pregnancy ended with preeclampsia and fetal death at 25 wk. She had a history of two early pregnancy losses. MCP K32N is a missense mutation that alters a charged solvent-exposed amino acid in complement control protein repeat (CCP) 1 [25]. To examine its functional consequences, the K32N mutant construct was transiently transfected in CHO cells. The mutant protein was quantified in cellular lysates and binding to human C4b and C3b was assessed (Figure 2A). While binding to C3b was similar to wild-type MCP, the mutant showed an approximately 50% decrease in C4b binding ($p<0.001$) compared to wild-type MCP. Given the lower binding to C4b, we suspected that the K32N mutant form of MCP would have a decreased capacity to serve as a cofactor for CFI-mediated cleavage of C4b. To examine this possibility, we performed cofactor assays comparing the ability of wild-type MCP and rare mutant K32N to serve as cofactors for cleavage of C3b and C4b by monitoring loss of the α' chain and generation of cleavage fragments (C4d in the case of C4b and α1 in the case of C3b) (Figure 2B) [24]. As predicted by the binding studies, cofactor activity for C3b was similar for the K32N mutant and wild-type MCP, whereas the K32N mutant showed decreased CFI-mediated cleavage activity for C4b. At 20 min in the C4b cofactor activity assay, no C4d was detected in the K32N lane, but an easily discernable band in wild-type MCP. We performed densitometric scanning of the C4d after 90 min and found that cleavage of C4b to C4d in wild-type was four times greater than in the K32N mutant (Figure 2C). Our results define the functional defect in K32N and identify the initial disease-related mutation in MCP that results exclusively in impaired ability to bind C4b and mediate C4b cofactor activity. Because SLE is characterized by autoantibodies that trigger the classical pathway, defective regulation of C4b is likely to influence the severity of tissue injury and risk for preeclampsia (Figure 1). Indeed, the presence of C4 degradation products in placentas from patients with APL Ab and in kidneys from

Table 3. Clinical characteristics of non-autoimmune preeclampsia patients with complement regulatory protein mutations.

Mutation/Variant	Maternal Age (y)	Preeclampsia	GA at Delivery (wk)	Fetal Weight in Grams (Percentile weight for GA)
MCP/A304V	24	Severe	33	2,102 (27)
MCP/A304V	23	Severe[a]	34	2,807 (48)
MCP/A304V	41	HELLP[b]	22	Neonatal demise
MCP/A304V	33	Severe[c]	33	2,013 (22)
CFI/I398L	24	Severe[d]	37	2,940 (35); 2,145 (2)

All patients were white.
[a]Previous pregnancy preeclampsia delivered at 36 wk.
[b]Pregnancy complicated by sphenoid sinus thrombosis; thrombophilia evaluation negative; four normal previous pregnancies; history of idiopathic thrombocytopenia purpura; mother with SLE.
[c]Oligohydramnios.
[d]Twins.
GA, gestational age.
doi:10.1371/journal.pmed.1001013.t003

[183]

lupus patients underscores the importance of C4 in disease pathogenesis [10,26].

Mutation in CFI

CFI, a plasma glycoprotein, is a serine protease that cleaves the α' chain of C3b and C4b in the presence of a cofactor protein, such as CFH or MCP. Inactivation of C3b and C4b prevents the formation of C3 and C5 convertases and thereby down-modulates the classical, lectin, and alternative pathways. We found identical heterozygous mutations in the gene encoding CFI in two unrelated white women with preeclampsia in the PROMISSE Study (Table 2). Similar to three of the four patients with MCP variants, those with CFI mutations had a history of previous pregnancy complications, specifically preterm deliveries at 24 and 28 wk, respectively. The mutation identified in CFI, an IVS12+5 G>T change, has been reported in patients with aHUS and was shown to affect the donor site of intron 12 reducing the splice score from 93 to 86 [27] but did not alter CFI serum levels [12]. The CFI IVS12+5 G>T variant has been identified in six of 325 French control individuals (0.9% allele frequency) and in eight of 238 patients from the French multi-center aHUS cohort (1.7% allele frequency) (alone in four patients and in association with a second mutation in four cases; V. Fremeaux-Bacchi, unpublished results). The frequency of this variant is not reported in publically available databases. Our findings suggest that CFI IVS12+5 G>T is a low-frequency variant that may be enriched in aHUS. Its clinical impact cannot be assessed in the control patients because pregnancy histories and other medical information were not available for these individuals.

The two patients presented here and other patients with the IVS12+5 G>T variant have had normal CFI levels (66 and 63 mg/l, nl 42-78 mg/l), but CFI levels within the normal range does not exclude alterations in local production [28,29]. In the 12 Eutherian mammals that have a similar CFI gene architecture to human CFI (Ensemble release 52), conservation is perfect at the +5 intron position. The lack of variation at this position relative to the surrounding bases argues that a G at this base is functionally important.

We screened CFI in 57 non-autoimmune patients with severe preeclampsia and/or HELLP in our second cohort from the Universal Samples Database at University of Utah. In one patient, we identified a heterozygous nonsynonymous sequence variant in exon 11 (c.1246A>G, p.Ile398Leu). We did not find this change in 117 white British controls. Bienaime et al. and Sellier-Leclerc et al. have previously demonstrated this variant in three aHUS patients (two heterozygous, one homozygous) but not in 100 controls [29,30]. CFI was undetectable in the individual with the homozygous change [29]. Recombinant expression of the I398L mutant confirmed nearly complete absence of production compared to wild-type CFI, which was associated with altered subcellular localization detected by immunofluorescence and altered endoglycosidase H digestion [29]. We identified two other nonsynonymous changes in this cohort: c. 1217G>A, p.Arg388His [19,28] and c.1322A>G, p.Lys423Arg [rs41278047]. Both changes have previously been described in control populations and are classed as SNPs in our study. Identification of CFI I398L mutation, a complete loss of function leading to haploinsufficiency, supports a role for defects in complement regulation by CFI in the pathogenesis of preeclampsia. This association also suggests that enrichment of IVS12+5G>T variants in patients with preeclampsia and aHUS [12] is functionally important, although its mechanism is as yet undefined.

Mutations in Factor H

The single factor H mutation (S40A) in the PROMISSE cohort was extensively investigated and not shown to have a functional defect (normal secretion in a transfection system, normal plasma levels, and normal to increased C3b binding and cofactor activity) (unpublished data). No factor H mutations were found in the Utah cohort.

Discussion

MCP and CFI mutations are, to our knowledge, the first genetic defects to be associated with preeclampsia in pregnant patients with SLE and/or APL Ab. Our hypothesis that mutations affecting function (missense, nonsense, and conserved splice site and promoter mutations) in negative regulators of the amplification loop of complement would be enriched in preeclampsia as compared with controls is supported by our finding that such mutations in these genes carries an odds ratio of 8.2 (95% CI 1 69) in the PROMISSE study cohort. We replicated the association of mutations in genes encoding complement inhibitors with pre-eclampsia in a second cohort of patients with severe preeclampsia and/or HELLP, but without underlying autoimmune disease. Taken together, our results suggest that dysregulation of complement activation can be a risk for preeclampsia.

Mutations in *MCP* and *CFI* appear to predispose to the development of disease, and as observed in aHUS, mutations are heterozygous and penetrance is incomplete [12]. In both preeclampsia and aHUS, a second hit that initiates the complement cascade or endothelial injury is required to manifest disease. In PROMISSE Study patients, autoantibodies may lower the threshold for preeclampsia, resulting in clinically apparent disease. These autoantibodies trigger the classical pathway of complement locally and systemically. Complement activation is then amplified through the feedback loop of the alternative pathway, leading to further injury of the placenta and stimulation of endothelial cells to express a prothrombotic, proadhesive phenotype (Figure 1). In patients with obesity or diabetes, conditions that predispose to preeclampsia and are characterized by systemic inflammation, the inflammatory process may sensitize vascular endothelium to injury. In aHUS, infection or pregnancy may trigger complement and precipitate disease [12,31]. Indeed, in four patients with impaired complement regulatory protein function, pregnancy precipitated thrombotic microangiopathy in the form of atypical HELLP with severe renal involvement [32,33]. Our findings in patients in the PROMISSE Study, taken together with case series of aHUS and HELLP patients, argue that pregnancy per se provokes or amplifies thrombotic microangiopathy in those already at risk due to inadequate control of the complement system.

In sequencing only three complement regulatory genes, we found heterozygous mutations in 10% 20% of PROMISSE patients with preeclampsia, which likely underestimates the importance of genetic defects in complement regulation, because we analyzed only the maternal genes, although trophoblasts are of fetal origin and express both maternal and paternal genes. Also, we have yet to analyze CFB and C3, which account for 10% 15% of the mutations in aHUS [13,15]. Our findings, in this prospective, longitudinal study of patients at risk for preeclampsia (in contrast to a biased sample from a case series), demonstrate that mutations or uncommon variants in complement regulatory proteins are not rare in those who develop disease and support the possibility that they increase disease risk. Confirmation of the presence of hypomorphic uncommon variants and mutations in a cohort of non-autoimmune patients with severe preeclampsia (the
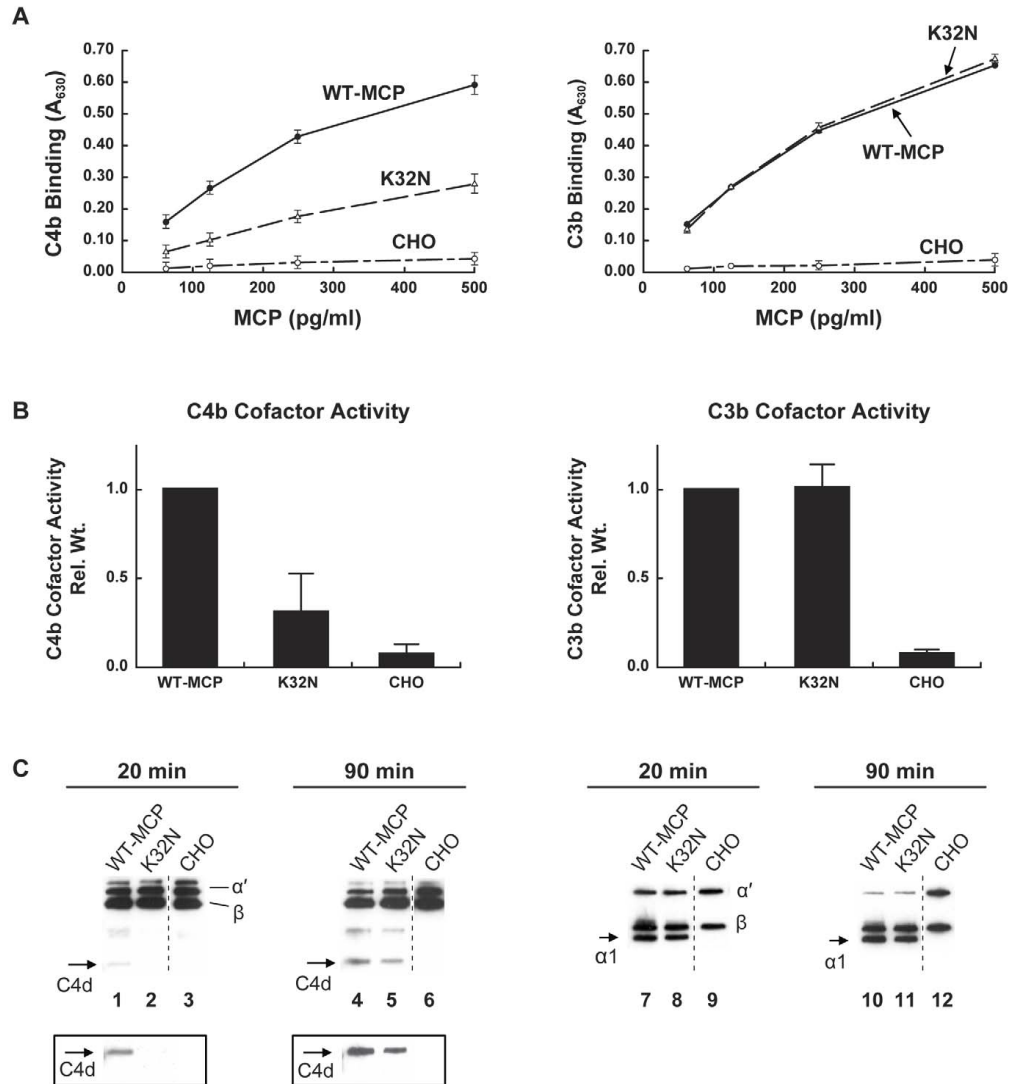
Figure 2. C3b and C4b binding and cofactor activity of K32N compared to wild type MCP. (A) C4b and C3b binding. CHO cell lysates from transient transfections were incubated with C4b- or C3b-coated wells and binding was assessed using a polyclonal antibody to MCP in an ELISA format. CHO, a mock transfected control. Data represent the mean ± standard deviation at each concentration from three independent experiments. (B) Cofactor activity for CFI-mediated cleavage of C4b and C3b. Cell lysates containing 25 pg of MCP (wild-type or K32N) were incubated with biotinylated C4b or C3b in the presence of purified human CFI in 25 mM NaCl and cleavage was assessed by Western blot. C4b and C3b peptide chains and cleavage fragments are identified. For C4b cofactor activity, C4d fragment generated by cleavage of C4b after 90 min was quantified by densitometric scanning, compared with quantity of β chain, and normalized to wild type. For C3b cofactor activity, α1 fragment generated by cleavage was compared to β chain by densitometry, and normalized to wild type. Data represent the mean ± standard deviation of three independent experiments. (C) Representative Western blots of MCP activity for CFI-mediated cleavage of C4b and C3b. Left panel: MCP K32N has deficient C4b cofactor activity at 20 and 90 min. The C4d fragment generated by cleavage of the α′ chain is not visible at 20 min and is diminished relative to wild type at 90 min. The boxed inserts represent 5-fold longer exposure of the blot. The CHO negative control was a nonadjacent lane (broken line). Right panel: MCP K32N has similar C3b cofactor activity to wild type. Equal amounts of α1 fragments are generated after 20 and 90 min. The CHO negative control was a nonadjacent lane (broken line). In two additional experiments, similar results were obtained in cofactor assays at 150 mM NaCl concentration (unpublished data).
doi:10.1371/journal.pmed.1001013.g002

[185]

Utah cohort) provides further evidence for the importance of complement activation in disease pathogenesis.

Based on findings in our patients and prior experiments in mouse models, we propose the following mechanism for the genesis of preeclampsia. Defective regulation of the complement system allows for the excessive complement activation that leads to placental damage, abnormal placental development, generalized endothelial activation, and the release of antiangiogenic factors toxic to the fenestrated endothelium of glomeruli, the choroid plexus, and liver sinusoids a sequence of events that culminates in clinical preeclampsia (Figure 1). Future studies of the PROMISSE cohort will determine whether elevations of split products generated by activation of the alternative or classical complement pathway predict preeclampsia in patients with APL Ab and/or SLE, and the information in the current report will allow us to refine these analyses and compare those with and without mutations leading to excessive complement activation. Our findings underscore the important role of complement activation in preeclampsia, define mutations and likely mechanisms for increased risk in patients with SLE and/or APL Ab, and suggest new targets for treatment of this important public health problem that, thus far, has defied reliable prediction and satisfactory intervention.

References

1. Ilekis JV, Reddy UM, Roberts JM (2007) Preeclampsia a pressing problem: an executive summary of a National Institute of Child Health and Human Development workshop. Reprod Sci 14: 508 523.
2. Maynard S, Epstein FH, Karumanchi SA (2008) Preeclampsia and angiogenic imbalance. Annu Rev Med 59: 61 78.
3. Young BC, Levine RJ, Karumanchi SA (2010) Pathogenesis of preeclampsia. Annu Rev Pathol 5: 173 192.
4. Esplin MS, Fausett MB, Fraser A, Kerber R, Mineau G, et al. (2001) Paternal and maternal components of the predisposition to preeclampsia. N Engl J Med 344: 867 872.
5. Girardi G, Yarilin D, Thurman JM, Holers VM, Salmon JE (2006) Complement activation induces dysregulation of angiogenic factors and causes fetal rejection and growth restriction. J Exp Med 203: 2165 2175.
6. Qing X, Redecha PB, Burmeister MA, Tomlinson S, D'Agati VD, et al. (2010) Targeted inhibition of complement activation prevents features of preeclampsia in mice. Kidney Int 2010 Oct 13 [Epub ahead of print].
7. Mao D, Wu X, Deppong C, Friend LD, Dolecki G, et al. (2003) Negligible role of antibodies and C5 in pregnancy loss associated exclusively with C3-dependent mechanisms through complement alternative pathway. Immunity 19: 813 822.
8. Lynch AM, Gibbs RS, Murphy JR, Byers T, Neville MC, et al. (2008) Complement activation fragment Bb in early pregnancy and spontaneous preterm birth. Am J Obstet Gynecol 199: 354 e351 358.
9. Tincani A, Bazzani C, Zingarelli S, Lojacono A (2008) Lupus and the antiphospholipid syndrome in pregnancy and obstetrics: clinical characteristics, diagnosis, pathogenesis, and treatment. Semin Thromb Hemost 34: 267 273.
10. Shamonki JM, Salmon JE, Hyjek E, Baergen RN (2007) Excessive complement activation is associated with placental injury in patients with antiphospholipid antibodies. Am J Obstet Gynecol 196: 167 e161 165.
11. Girardi G, Bulla R, Salmon JE, Tedesco F (2006) The complement system in the pathophysiology of pregnancy. Mol Immunol 43: 68 77.
12. Caprioli J, Noris M, Brioschi S, Pianetti G, Castelletti F, et al. (2006) Genetics of HUS: the impact of MCP, CFH, and IF mutations on clinical presentation, response to treatment, and outcome. Blood 108: 1267 1279.
13. Noris M, Remuzzi G (2009) Atypical hemolytic-uremic syndrome. N Engl J Med 361: 1676 1687.
14. Fang CJ, Fremeaux-Bacchi V, Liszewski MK, Pianetti G, Noris M, et al. (2008) Membrane cofactor protein mutations in atypical hemolytic uremic syndrome (aHUS), fatal Stx-HUS, C3 glomerulonephritis, and the HELLP syndrome. Blood 111: 624 632.
15. Kavanagh D, Richards A, Atkinson J (2008) Complement regulatory genes and hemolytic uremic syndromes. Annu Rev Med 59: 293 309.
16. (2002) ACOG practice bulletin. Diagnosis and management of preeclampsia and eclampsia. Number 33, January 2002. Obstet Gynecol 99: 159 167.
17. (2007) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447: 661 678.
18. Burton PR, Clayton DG, Cardon LR, Craddock N, Deloukas P, et al. (2007) Association scan of 14,500 nonsynonymous SNPs in four diseases identifies autoimmunity variants. Nat Genet 39: 1329 1337.
19. Fremeaux-Bacchi V, Dragon-Durey MA, Blouin J, Vigneau C, Kuypers D, et al. (2004) Complement factor I: a susceptibility gene for atypical haemolytic uraemic syndrome. J Med Genet 41: e84.
20. Fremeaux-Bacchi V, Moulton EA, Kavanagh D, Dragon-Durey MA, Blouin J, et al. (2006) Genetic and functional analyses of membrane cofactor protein (CD46) mutations in atypical hemolytic uremic syndrome. J Am Soc Nephrol 17: 2017 2025.
21. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, et al. (2009) Finding the missing heritability of complex diseases. Nature 461: 747 753.
22. Liszewski MK, Leung M, Cui W, Subramanian VB, Parkinson J, et al. (2000) Dissecting sites important for complement regulatory activity in membrane cofactor protein (MCP; CD46). J Biol Chem 275: 37692 37701.
23. Li B, Leal SM (2008) Methods for detecting associations with rare variants for common diseases: application to analysis of sequence data. Am J Hum Genet 83: 311 321.
24. Liszewski MK, Leung MK, Schraml B, Goodship TH, Atkinson JP (2007) Modeling how CD46 deficiency predisposes to atypical hemolytic uremic syndrome. Mol Immunol 44: 1559 1568.
25. Casasnovas JM, Larvie M, Stehle T (1999) Crystal structure of two CD46 domains reveals an extended measles virus-binding surface. Embo J 18: 2911 2922.
26. Cohen D, Koopmans M, Kremer Hovinga IC, Berger SP, Roos van Groningen M, et al. (2008) Potential for glomerular C4d as an indicator of thrombotic microangiopathy in lupus nephritis. Arthritis Rheum 58: 2460 2469.
27. Rogozin IB, Milanesi L (1997) Analysis of donor splice sites in different eukaryotic organisms. J Mol Evol 45: 50 59.
28. Kavanagh D, Richards A, Noris M, Hauhart R, Liszewski MK, et al. (2008) Characterization of mutations in complement factor I (CFI) associated with hemolytic uremic syndrome. Mol Immunol 45: 95 105.
29. Bienaime F, Dragon-Durey MA, Regnier CH, Nilsson SC, Kwan WH, et al. (2010) Mutations in components of complement influence the outcome of Factor I-associated atypical hemolytic uremic syndrome. Kidney Int 77: 339 349.
30. Sellier-Leclerc AL, Fremeaux-Bacchi V, Dragon-Durey MA, Macher MA, Niaudet P, et al. (2007) Differential impact of complement mutations on clinical characteristics in atypical hemolytic uremic syndrome. J Am Soc Nephrol 18: 2392 2400.
31. Fakhouri F, Roumenina L, Provot F, Sallee M, Caillard S, et al. (2010) Pregnancy-associated hemolytic uremic syndrome revisited in the era of complement gene mutations. J Am Soc Nephrol 21: 859 867.
32. Fakhouri F, Jablonski M, Lepercq J, Blouin J, Benachi A, et al. (2008) Factor H, membrane cofactor protein, and factor I mutations in patients with hemolysis, elevated liver enzymes, and low platelet count syndrome. Blood 112: 4542 4545.
33. Fang CJ, Richards A, Liszewski MK, Kavanagh D, Atkinson JP (2008) Advances in understanding of pathogenesis of aHUS and HELLP. Br J Haematol 143: 336 348.

[186]