ATYPICAL HUS RESEARCH GRANT

Annual Progress Report

August 2011

Project Title: Targeted-Genomic Capture and High-Throughput Sequencing for Genetic Testing and New Gene Discovery in Atypical Hemolytic Uremic Syndrome

Principle Investigator: Dr. Richard Smith

Institutional address: The University of Iowa
5270 CBRB
Iowa City, IA 52242
Phone number: 319-335-6623
FAX number: 319-353-5869
E-mail address: richard-smith@uiowa.edu

INTRODUCTION

The pathogenesis of aHUS is linked to dysregulation of the alternative pathway of the complement cascade. Loss-of-function mutations in the complement regulators factor H (*CFH*), membrane cofactor protein (*MCP*), factor I (*CFI*) and thrombomodulin (*THBD*), and gain-of-function mutations in the complement activators factor B (*CFB*) and complement component 3 (*C3*) have been implicated in the disease. In nearly 40% of patients, however, mutations are not identified in these genes raising the possibility that there may be other genetic causes of aHUS. Identifying these others genes is an important initiative as it may be relevant to the clinical course of aHUS in these patients.

In this focus proposal, we sought to discover novel genes that are causally related to aHUS. To achieve this goal, we targeted aHUS patients negative for mutations in the known aHUS-related genes and screened them for genetic mutations in a large number of other genes using targeted-genome capture and massively parallel sequencing. The targeted-genome platform included all complement genes in addition to genes in pathways that impact the complement cascade. We have called this platform CASCADE (Capture And Sequencing of Complement-Associated Disease Exons).

In this report, we describe our experience with several patients using CASCADE.

METHODS

Patient Population

DNA from patients with the diagnosis of atypical hemolytic uremic syndrome was obtained using established protocols. Only patients who tested negative for mutations in *CFH*, *CFHR5*, *CD46*, *CFI*, *C3*, *CFB*, and *THBD* were included in the first analysis. All patients submitted blood samples to the Molecular Otolaryngology and Renal Research Laboratories (MORL, Iowa City, IA) after providing informed consent according to University of Iowa Institutional Review Board guidelines.

Targeted Capture and Next Generation Sequencing

Genetic testing was completed using custom-made solution-based sequence capture followed by sequencing on the Illumina GA2 system. We designed two different

platforms: CASCADEv1 and CASCADEv2. In version two we improved upon version one by omitting the Toll-Like Receptors, which were not covered well using this technology; we also added several more genes in the coagulation pathway (Table 1).

Category	**CASCADEv1**	**CASCADE v2**
Number of genes	68	85
Number of exons	755	1071
Number of base pairs	325,322	198,457
Genes in the complement cascade	51	58
Genes in the coagulation pathway	7	26
Toll-like receptor genes	10	0

Table 1: Improvements made to the original CASCADE.

Exons were targeted and enriched using Agilent SureSelect solution-based sequence capture (Gnirke, et al., 2009). In brief, this massively parallel technique simultaneously amplifies targeted regions using engineered complementary biotinylated RNA baits (Agilent). The RNA baits are designed to be complementary to the exons of interest. Once hybridization occurs, the regions of interest are captured using magnetic 'beads' that attract the baits and their complementary DNA out of solution. Lastly, the beads are washed and the RNA is digested.

Once the exons of interest were amplified using targeted sequence capture they were sequenced at the Baylor College of Medicine (Drs. Richard Gibbs and Steve Scherer) using the Illumina GA2 (CASCADEv1) or Hi-seq (CASCADEv2) platform (Bentley, et al., 2008). In brief, single-stranded DNA libraries hybridized to complementary oligonucleotides are covalently bound to each of the eight lanes of a flow cell. By attaching molecular barcodes to individual samples, we were able to include 10-12 samples on each of the eight lanes allowing us to sequence multiple individuals with each Illumina sequencing run (a technique currently being validated by our lab).

Once hybridized to the flow cell, the DNA was amplified *in situ* and then used as a template for synthetic sequencing using fluorescent reversible terminator deoxyribonucleotides, allowing an imager to read the fluorescence signals.

Analysis of the Data

In this study, we used a minimal depth-of-coverage of 10X for variant calling. Sequence reads were aligned to the reference genome generated by the human genome project using Bowtie and BWA. Once the sequencing reads were aligned, variations between the aligned sequencing reads and the human genome were identified using Samtools. We prioritized variants that were heterozygous, found in less then 10% of healthy individuals (according to dbSNP or 1000 genomes), and found in only one or two aHUS patients. Candidate variations were verified by Sanger sequencing. All novel variants were screened against 150 ethnically matched controls.

RESULTS

Sample outcomes for CASCADEv1 and CASCADEv2

Seventeen aHUS patients were run on the first two versions of CASCADE. Three aHUS patients with no known pathogenetic variant were run on CASCADEv1. Each of the samples was run on one lane of the Illumina GA2X and in all three samples we were able to verify variants that likely contribute to disease. In our second study, 14 aHUS patients negative for pathogenic mutations were captured using CASCADEv2. To decrease costs, these samples were bar-coded so that multiple samples could be run on a single lane of the Illumina HI-seq. Of 14 samples, 12 were sequenced successfully. (Two samples did not sequence due to their low DNA concentrations after capture.) In seven patients we identified variants that potentially contribute to the pathogenesis of aHUS.

Coverage Statistics

Baits were designed and covered 99% of the requested regions (the coding regions of all 85 genes). After the captured regions were sequenced and realigned to the reference genome, 99% of regions were covered by 10 sequencing reads (10X) and 96% were covered at 40X. The average depth-of-coverage for this set of samples was 1855 reads per base pair (1855X).

Variants Identified by CASCADE

We discovered several novel variants and rare polymorphisms that are predicted to be important in the pathogenesis of aHUS (Table 3). CASCADE identified a total of 24 variants in ten patients. Of these variant, 37% are in genes in the alternative pathway and 25% are in genes in the terminal pathway of the complement cascade. There were no variants in either the classical or lectin pathways.

The single gene with the most variants was *ADAMTS13*, with a total of six variants (25% of the total). The CFH-related genes as a group had the second highest number of variations with a total of five. The other 13% of variants were located in other genes in the coagulation pathway (Figure1).

Patient	Known	Protein	Amino Acid Change	1000 Genomes	dbSNP Frequency	Control Frequency (n=150)
Positive Control	CFH, G1194D MCP, F242C	CFH	G1194D	-	-	0%
		MCP	F242C	-	-	0%
		CFHR2	T71M	-	-	0%
		CFHR5	R356H	-	1.4%	-
Patient 1	None	ADAMTS13	P618A	6.7%	6.7%	-
Patient 2	None	CFHR2	C72Y	0.6%	5%	-
		C8B	D382Y	-	-	1%
Patient 3	None	CFI	R406H	-	1%	-
Patient 4	C3, P63L	ADAMTS13	A1033T	1%	3%	-
		ADAMTS13	R1060W	-	-	0.6%

		PLAT	R164W	0.8%	5%	
Patient 5	None	CFHR2	C72Y	0.6%	5%	-
		ADAMTS13	A747V	-	-	0%
Patient 6	FH auto antibodies	C5AR1	R79W	-	-	0%
		ADAMTS13	A732V	1%	3%	-
		VTN	E126V	0.2%	-	-
Patient 7	Deletion	GPR77	H329R	0.3%	2%	-
Patient 8	Deletion	C9	D127Y	0.4%	2%	-
		CFHR4	Y35H	0.6%	-	-
Patient 9	None	C8B	V536D	-	-	0%
Patient 10	CFB, L533R	PLG	A494V	1%	1%	-
		PLG	S460R	2%	-	-
		CFI	H183R	0.2%	-	-
		ADAMTS13	R386C	-	-	0%

Table 3: Interesting variants identified using CASCADE. Seven of these variants are novel mutations that are not present in controls (1000 Genomes, dbSNP, or our in house controls). The rest of the changes are rare polymorphisms that are seen in healthy individuals.

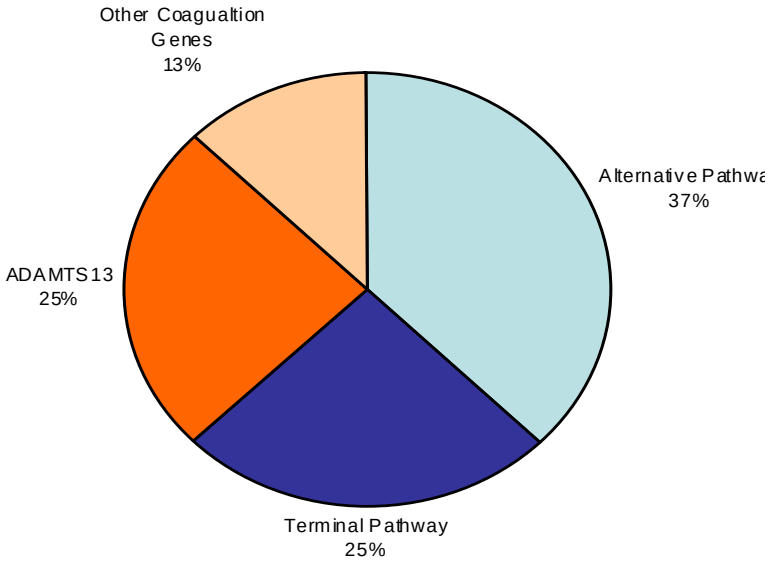

Figure 1: Functional distribution of variants identified with CASCADE.

Functional Significance of ADAMTS13 Variants

Mutations in *ADAMTS13* are known to cause Thrombotic Thrombocytopenic Purpura (TTP). TTP is characterized by a pentad of symptoms including: neurologic symptoms, kidney failure, fever, thrombocytopenia, and microangiopathic hemolytic anemia. aHUS

and TTP both fall under the umbrella of Thrombotic Microangiopathic Anemia (TMA) and at times can be difficult to distinguish from one another.

Two of the polymorphisms identified using CASCADE (A618 and V732) had been previously characterized in TTP patients and were found to decrease both the antigen levels and activity levels of ADAMTS13 (Plaimauer B et al. Blood 2006) (Figure 2). This is the first time these variants have been identified in aHUS patients.

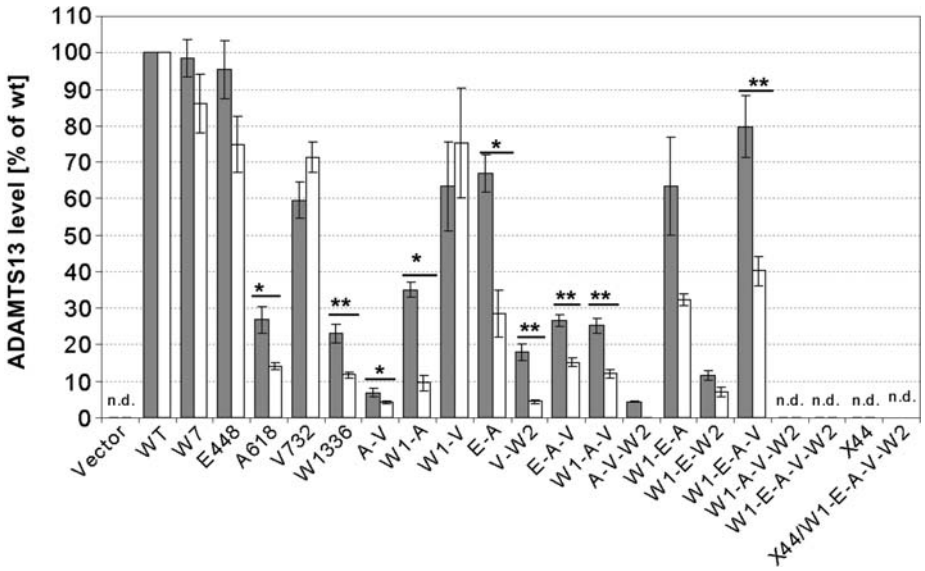

Figure 2: ADAMTS13 antigen and activity levels obtained in the conditioned medium of cells expressing ADAMTS13 variants. (Plaimauer B et al. Blood 2006; 107:118-125)

DISCUSSION

Dysregulation of the AP has been well documented in aHUS, consistent with results of this study showing that most of the identified variants were found in genes in the AP. Several variants were also seen in the terminal pathway but interestingly, no variants were found in either the lectin or classical pathway.

CFH is the most important regulator of the AP and the most frequently mutated gene in aHUS. Interestingly, we found five variants in the CFH related genes (Table 2). The CFH related gene family regulates functional activity of the AP and terminal complement cascade. In earlier work, our lab identified *CFHR5* as a gene that should be screened in all patients with aHUS. Based on this study, *CFHR2* and *CFHR4* should be added to that list. Our on-going work is to complete functional studies on *CFHR2* and *CFHR5*.

The most commonly mutated gene was *ADAMTS13*, carrying 13% of the total variants. Two of the variants we identified are known to affect protein function, supporting the hypothesis that mutations in this gene contribute to aHUS. This finding is significant because it is the first time *ADAMTS13* has been implicated in aHUS. However, mutations in *ADAMTS13* are known to cause TTP, and as mentioned earlier, aHUS and TTP are both thrombic microangiopathies and may be difficult to distinguish clinically. TTP patients usually have no ADAMTS13 activity. Our data suggest that patients with homozygous mutations in *ADAMTS13* develop TTP, while patients with heterozygous variants are predisposed to develop aHUS. (All of the variants in *ADAMTS13* found in our aHUS patients were heterozygous.) This hypothesis is consistent with the pathogenesis of aHUS: endothelial cell damage caused by complement attack leads to

the formation of one of the hallmarks of aHUS - thrombosis. Thrombi form in the microvascular of the glomerulus and shred erythrocytes leading to microangiopathic anemia, thrombocytopenia and eventually end stage renal failure.

In total, 38% of variants found by CASCADE were located in proteins of the coagulation pathway including: ADAMTS13, PLAT, and PLG. These proteins function together to prevent thrombosis. Based on this finding, we propose that variants in these genes act as modifiers in patients with mutations in genes in the complement cascade predisposing them to developing an aHUS phenotype instead of other complement-mediated diseases such as dense deposit disease (DDD). This finding also helps to explain why one pathway, the alternative pathway, can be implicated in different pathologies.

In conclusion this study has identified novel aHUS genes, in particularly *ADAMTS13*, and has implicated a new pathway in the pathogenesis of aHUS. Sequence capture technology itself is uniquely suited to complex disease. By definition complex diseases like aHUS are not only multigenic but also triggered by environmental causes. When sequence capture technology is paired with next generation sequencing we can screen large numbers of genes in a time and cost efficient manner. Using CASCADE we have been able to generate a genetic complement profile or “complotype” for aHUS patients. This complotype will be vital explore the genotype-phenotype relationships in aHUS.