Comprehensive Genetic Analysis of Complement and Coagulation Genes in Atypical Hemolytic Uremic Syndrome


*Interdisciplinary PhD Program in Genetics, †Molecular Otolaryngology and Renal Research Laboratories, ‡Department of Biostatistics, College of Public Health, and §Rare Renal Disease Clinic, Departments of Pediatrics and Internal Medicine, Carver College of Medicine, University of Iowa, Iowa City, Iowa

ABSTRACT

Atypical hemolytic uremic syndrome (aHUS) is a thrombotic microangiopathy caused by uncontrolled activation of the alternative pathway of complement at the cell surface level that leads to microangiopathic hemolytic anemia, thrombocytopenia, and acute kidney failure. In approximately one half of affected patients, pathogenic loss-of-function variants in regulators of complement or gain-of-function variants in effectors of complement are identified, clearly implicating complement in aHUS. However, there are strong lines of evidence supporting the presence of additional genetic contributions to this disease. To identify novel aHUS-associated genes, we completed a comprehensive screen of the complement and coagulation pathways in 36 patients with sporadic aHUS using targeted genomic enrichment and massively parallel sequencing. After variant calling, quality control, and hard filtering, we identified 84 reported or novel nonsynonymous variants, 22 of which have been previously associated with disease. Using computational prediction methods, 20 of the remaining 62 variants were predicted to be deleterious. Consistent with published data, nearly one half of these 42 variants (19; 45%) were found in genes implicated in the pathogenesis of aHUS. Several genes in the coagulation pathway were also identified as important in the pathogenesis of aHUS. PLG, in particular, carried more pathogenic variants than any other coagulation gene, including three known plasminogen deficiency mutations and a predicted pathogenic variant. These data suggest that mutation screening in patients with aHUS should be broadened to include genes in the coagulation pathway.


Hemolytic uremic syndrome (HUS) is a thrombotic microangiopathy characterized by microangiopathic hemolytic anemia, thrombocytopenia, and acute kidney failure that is most frequently caused by Shiga-toxigenic Escherichia coli (STEC) serogroup O157. About 10% of HUS cases, however, are atypical and not caused by an antecedent STEC infection.¹ The prevalence of atypical HUS (aHUS) is estimated at 1 per 2 and 1 per 3.3 million children in the United States and Europe, respectively.²,³ Until the recent availability of Eculizumab, aHUS carried significant morbidity and mortality, with 25% of patients dying of disease, 50% of patients progressing to end stage renal failure, and 40%–80% of patients experiencing disease recurrence after transplantation.¹,⁴

Uncontrolled activation of the alternative pathway of complement at the level of the cell surface has been closely linked to disease.⁵,⁶ In about one half of aHUS patients, pathogenic mutations are identified in complement genes, including loss-of-function mutations in regulators of complement activity.
like complement factor H (CFH), complement factor I (CFI), and membrane cofactor protein (CD46) and gain-of-function mutations in effectors of complement like factor B (CFB) and complement component 3 (C3), in a recent large genetic screen of 794 aHUS patients, single mutations in CFH, C3, CFI, CFB, or CD46 were identified in 41% of patients, and combined mutations were found in 3% of patients. Additional genetic contributors include single nucleotide polymorphisms and haplotypes in CFH and CD46 copy number variation of the complement factor H-related region caused by nonallelic homologous recombination.

Although these studies clearly implicate mutations in complement genes in the pathogenesis of aHUS, additional genetic contributions to this disease are likely to exist. There are three lines of evidence that strongly support this hypothesis. First, in a large percentage of cases, no mutations are found in the commonly implicated complement genes. Second, there is an extremely high rate of incomplete penetrance in familial aHUS, which is consistent with the existence of additional modifying/contributory genetic factors. Third, current screens have narrowly focused on fewer than a dozen genes, although the complement system is substantially larger and integrated with numerous other signaling pathways.

In this study, we sought to address these limitations by using targeted genomic enrichment and massively parallel sequencing (TGE+MPS) to screen the coding sequence and splice sites of all genes in both the complement and coagulation pathways. We selected the coagulation pathway because of its role in thrombotic thrombocytopenic purpura (TTP), a thrombotic microangiopathy in which extensive microscopic thrombi form in the microvasculature. aHUS and TTP share the clinical features of microangiopathic hemolysis, thrombocytopenia, and renal involvement, but in TTP patients also have fever and neurologic symptoms. The fact that the coagulation pathway warrants detailed scrutiny is also supported by the finding of pathogenic aHUS mutations in two anticoagulant-related genes, THBD and DGKE. THBD encodes thrombomodulin, which functions as a cofactor for thrombin to reduce blood coagulation and also regulates CFI-induced C3b inactivation. DGKE encodes diacylglycerol kinase-ε and is implicated in regulation of thrombus formation by modulating protein kinase C activity in endothelial cells and platelets.

In total, we sequenced 85 genes in 36 European-American sporadic aHUS patients, and we filtered and prioritized variants based on frequency and functional effects. As expected, we found novel deleterious variants in multiple complement genes, but we also identified deleterious nonsynonymous rare variants in several genes that play important roles in coagulation, definitely implicating this pathway in aHUS. The most frequently mutated gene in the coagulation pathway was PLG, which encodes plasminogen, a zymogen that is secreted by the liver and converted into plasmin by a variety of enzymes on binding to clots.

RESULTS

Patients
In total, 36 European-American patients with sporadic aHUS (24 men and 12 women) were recruited into this study from 2007 to 2011 (Figure 1). The diagnosis of aHUS was based on the presence of acute renal failure, thrombocytopenia, and microangiopathic hemolytic anemia after excluding antecedent STEC infection. All procedures were approved by the Institutional Review Board of the University of Iowa Carver College of Medicine.

Sequencing Quality Metrics and Variant Calling
We performed TGE+MPS on 36 European-American aHUS patients using an RNA-bait panel that targets coding sequence of genes in the complement and coagulation pathways (hereafter referred to as capture and sequencing of complement-associated disease exons [CasCADE]). The mean total reads generated per sample approached 9,000,000, with more than 97% of reads mapping to the reference genome and more than 50% of reads mapping to targeted regions. The average coverage per targeted base was 1670 reads; 99% of targeted regions were covered by more than 30 reads (Table 1). Of 7486 variants identified in 36 aHUS patients, 7071 (95.46%) variants passed quality control (Figure 2). Each subject carried a mean of 196 variants. After filtering by minor allele frequency and functional effect, 18 novel and 66 reported nonsynonymous rare variants (nsRVs) remained. All nsRVs were present in heterozygosity, with each patient carrying a mean of 2.33 nsRVs (range=0–5).

Figure 1. Diagnostic age distributes similarly in male and female aHUS patients. Although a trend toward higher diagnostic age was observed in women, this difference was not significant (Wilcoxon rank sum test P=0.067).
**Table 1. Quality of sequencing results**

<table>
<thead>
<tr>
<th>Sequencing Metrics</th>
<th>Mean ± SD</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total reads</td>
<td>9,195,907±1,294,156</td>
<td>9,176,081</td>
</tr>
<tr>
<td>Mapped reads</td>
<td>9,112,151±1,215,320</td>
<td>9,008,634</td>
</tr>
<tr>
<td>Mapped reads, %</td>
<td>97.84</td>
<td>98.39</td>
</tr>
<tr>
<td>Reads overlapping targets</td>
<td>5,025,594±875,441</td>
<td>5,053,139</td>
</tr>
<tr>
<td>Reads overlapping targets, %</td>
<td>54.84</td>
<td>55.92</td>
</tr>
<tr>
<td>Average coverage</td>
<td>1670.65±290.82</td>
<td>1677.48</td>
</tr>
<tr>
<td>Bases covered 1×, %</td>
<td>99.60</td>
<td>99.67</td>
</tr>
<tr>
<td>Bases covered 10×, %</td>
<td>99.17</td>
<td>99.35</td>
</tr>
<tr>
<td>Bases covered 30×, %</td>
<td>98.97</td>
<td>99.29</td>
</tr>
</tbody>
</table>

**Figure 2.** Application of hard filtering uncovers pathogenic variants. Variants were filtered by focusing on nonsynonymous, frame shift, and indel changes. Rare variants are defined as variants with a minor allele frequency <1% in European-American and African-American populations based on data from the EVS database. Predicted deleterious mutations or rare variants are stop-gain mutations, indels, or missense variants with a high pathogenicity score (PS) (>4). PS for missense variants was calculated using PhyloP, SIFT, PolyPhen2, LRT, MutationTaster, and GERP++, and it ranged from zero to six. The mean PS of reported DRVs was 3.41; the mean PS of predicted deleterious missense variants was 4.57. *DRVs have been previously reported to be disease-causing or -associated. DRVs may or may not be recorded in public single nucleotide polymorphism databases. QC, quality control; UTR, untranslated region.

**Reported Disease-Related Variants**

Twenty-two of eighty-four nsRVs were reported disease-related variants (DRVs), most frequently being causally related to aHUS (15 DRVs in CD46, CFB, CHI, and CFI) (Table 2). Remarkably, however, there were three variants in PLG (c.112 A>G, p.Lys38Glu; c.2134 G>A, p.Gly712Arg; c.758 G>A, p.Arg253His) associated with plasminogen deficiency.26–28 which is a significant enrichment of DRVs in this gene compared with European-American data from the National Heart, Lung, and Blood Institute Exome Sequencing Project Exome Variant Server (EVS29; 3 DRV counts in 36 aHUS cases versus 76 DRV counts in 4300 EVS European-American subjects; Fisher exact test P=0.03). Other interesting ultrarare DRVs included a G-to-A transition that alters a canonical splice site in F12 (c.1681–1G>A) and leads to factor XII deficiency.20

**Novel nsRV Identification**

Among 18 novel variants that we identified, there were 15 missense, 2 frame-shift, and 1 splice site change variants (Supplemental Table 1). These variants included a G-to-C transversion in VWF, c.4165G>C (p.Glu1389Gln), that alters the same nucleotide as a DRV (c.4165G>A, p.Glu1389Lys), which is reported to cause type 2 von Willebrand disease by reducing amounts of high-molecular weight VWF and VWF-platelet binding.31 One novel variant, c.3221A>G, p.Asn1074Ser in F5, was shared by two aHUS patients and alters a potential N-link glycosylation site.

**Prediction of Deleterious Variants**

To assess possible pathogenicity of 62 variants not previously implicated in disease (58 missense, 1 splice site, and 3 frame-shift indels variants), we first used combined computational prediction methods (PhyloP, SIFT, PolyPhen2, LRT, MutationTaster, and GERP++) to calculate the mean pathogenicity score (PS) of the reported missense DRVs. Based on this PS, which was 3.81, we accepted as likely deleterious all novel and nsRVs with a PS of at least 4 (Table 2). The average PS of 16 missense novel and nsRVs meeting this minimal threshold was 4.57 (Table 3). Three frameshift indels (c.356_357insG in C3AR1, c.443_453del in complement component 2, and c.1770delG in mannann-binding lectin serine protease 1 [MASP1]) and a splice site change (c.1160–2A>G in CHI) were also considered possibly damaging.

In aggregate, 42 deleterious variants (22 DRVs and 20 predicted deleterious variants) were identified in 25 patients (69.4%) (Figure 3). Not surprisingly, of these variants, 20 variants were found in genes commonly implicated in aHUS (CHI, CD46, C3, CFH, F8, and THBD). However, 15 variants were found in other complement-related genes, and 7 variants were found in coagulation pathway genes. The greatest number of deleterious variants was found in CHI (11 variants in 10 patients), but the second greatest number was in PLG (4 variants in 4 patients), suggesting that this gene makes an important contribution to the aHUS phenotype.
In this study, we used TGE+MPS to screen 85 genes in the complement and coagulation pathways in 36 European-American patients with sporadic aHUS. After variant calling, quality control, and hard filtering, we identified 84 reported or novel nsRVs, 22 of which have been previously associated with disease. Using computational prediction methods (PhyloP, SIFT, PolyPhen2, LRT, MutationTaster, and GERP++), 20 additional variants were predicted to be deleterious. As expected and consistent with published data, nearly one half of these 42 variants (19; 45%) were found in genes implicated in the pathogenesis of aHUS (CFH, CD46, CFB, CFI, and C3). However, PLG, a gene in the coagulation pathway, also emerged as an important gene in the pathogenesis of aHUS.

The four variants that we identified in PLG included three reported plasminogen deficiency-related variants: c.112 A>G (p.Lys38Glu), c.758 G>A (p.Arg253His), and c.2134 G>A (p.Gly712Arg). Of these known variants, c.112 A>G, p.Lys38Glu (rs73015965; also referred to as Lys19Glu27) is noted for being the first variant identified in sporadic plasminogen-deficient patients (two men and one woman) and being associated with low plasminogen levels in the unaffected parental carrier.26 Subsequent studies have shown that this missense variant is found in one third of patients with plasminogen deficiency type I and associated with less severe reduction in plasma plasminogen levels and milder disease.27,28 The fourth variant that we identified, c.505 C>T, p.Pro169Ser, is novel and predicted to be disease-causing based on genetic conservation, bioinformatics (PS=4), and frequency data (Figure 5 and Table 2).

### DISCUSSION

In this study, we used TGE+MPS to screen 85 genes in the complement and coagulation pathways in 36 European-American patients with sporadic aHUS. After variant calling, quality control, and hard filtering, we identified 84 reported or novel nsRVs, 22 of which have been previously associated with disease. Using computational prediction methods (PhyloP, SIFT, PolyPhen2, LRT, MutationTaster, and GERP++), 20 additional variants were predicted to be deleterious. As expected and consistent with published data, nearly one half of these 42 variants (19; 45%) were found in genes implicated in the pathogenesis of aHUS (CFH, CD46, CFB, CFI, and C3). However, PLG, a gene in the coagulation pathway, also emerged as an important gene in the pathogenesis of aHUS.

The four variants that we identified in PLG included three reported plasminogen deficiency-related variants: c.112 A>G (p.Lys38Glu), c.758 G>A (p.Arg253His), and c.2134 G>A (p.Gly712Arg). Of these known variants, c.112 A>G, p.Lys38Glu (rs73015965; also referred to as Lys19Glu27) is noteworthy for being the first variant identified in sporadic plasminogen-deficient patients (two men and one woman) and being associated with low plasminogen levels in the unaffected parental carrier.26 Subsequent studies have shown that this missense variant is found in one third of patients with plasminogen deficiency type I and associated with less severe reduction in plasma plasminogen levels and milder disease.27,28 The fourth variant that we identified, c.505 C>T, p.Pro169Ser, is novel and predicted to be disease-causing based on genetic conservation, bioinformatics (PS=4), and frequency data (Figure 5 and Table 2).

### Copy Number Variants in CFHRs

We identified copy number variants (CNVs) in the CFHR genomic region in 15 patients (41.7%). Homozygosity for delCFHR3-CFHR1 was found in three patients who carried no deleterious nsRVs. Ten patients were heterozygous for this deletion, and in two other patients, we identified heterozygous duplications in the CFHR1 and CFHR4 genes. The frequency of these variations was not statistically different than the frequency seen in 314 healthy controls (unpublished data; Fisher exact test P=0.16; validated Hardy–Weinberg Equilibrium in patients and controls).

### Multiple Variants

Twelve patients (33.3%) carried two or more deleterious nsRVs (Supplemental Table 2). In six of these patients, the deleterious nsRVs were present only in complement genes, but in the remaining six patients, deleterious nsRVs were found in both complement and coagulation genes. Although not statistically significant, there was a trend to negative correlation between CNVs of the CFHR region and deleterious nsRVs (Figure 4). In five patients (13.9%), we were unable to identify genetic alterations in any complement or coagulation gene.

### DISCUSSION

In this study, we used TGE+MPS to screen 85 genes in the complement and coagulation pathways in 36 European-American patients with sporadic aHUS. After variant calling, quality control, and hard filtering, we identified 84 reported or novel nsRVs, 22 of which have been previously associated with disease. Using computational prediction methods (PhyloP, SIFT, PolyPhen2, LRT, MutationTaster, and GERP++), 20 additional variants were predicted to be deleterious. As expected and consistent with published data, nearly one half of these 42 variants (19; 45%) were found in genes implicated in the pathogenesis of aHUS (CFH, CD46, CFB, CFI, and C3). However, PLG, a gene in the coagulation pathway, also emerged as an important gene in the pathogenesis of aHUS.

The four variants that we identified in PLG included three reported plasminogen deficiency-related variants: c.112 A>G (p.Lys38Glu), c.758 G>A (p.Arg253His), and c.2134 G>A (p.Gly712Arg). Of these known variants, c.112 A>G, p.Lys38Glu (rs73015965; also referred to as Lys19Glu27) is noteworthy for being the first variant identified in sporadic plasminogen-deficient patients (two men and one woman) and being associated with low plasminogen levels in the unaffected parental carrier.26 Subsequent studies have shown that this missense variant is found in one third of patients with plasminogen deficiency type I and associated with less severe reduction in plasma plasminogen levels and milder disease.27,28 The fourth variant that we identified, c.505 C>T, p.Pro169Ser, is novel and predicted to be disease-causing based on genetic conservation, bioinformatics (PS=4), and frequency data (Figure 5 and Table 2).
The enrichment of PLG for plasminogen deficiency DRVs is statistically significant (P<0.05) and suggests that subnormal activity of plasminogen or plasmin may compromise the degradation of thrombi in aHUS. We also identified other DVRs linked to thrombosis and fibrinolysis, including a variant in THBD (c.1502 C>T [p.Pro501Leu]) that has been associated with aHUS and thromboembolic disease. In vitro, this missense change reduces cell surface expression of thrombomodulin, which may contribute in vivo to suppression of the alternative pathway through CFI-mediated C3b inactivation and venous thrombosis. Another variant, c.1681–1 G>A in F12, alters a canonical splice acceptor, leading to a truncated transcript, reduced protein expression, increased partial thromboplastin time, and recurrent superficial venous thrombosis.

The role of coagulation activity in aHUS patients has been evaluated at the mRNA level in one study, in which gene expression was quantitated in glomeruli from archival paraffin-embedded renal biopsies. Expression of antifibrinolytic, prothrombotic plasminogen activator inhibitor-1 and antithrombotic thrombomodulin were increased, and expression of profibrinolytic, antithrombotic tissue plasminogen activator was decreased compared with controls, suggesting that this patient also carries a deleterious variant in MASP2 (c.1633 A>G or p.N545D, PS=4).

Multiple variants in common aHUS genes were identified in two patients, which is in agreement with expected frequencies (3%–12%) (Supplemental Table 2). However, these numbers are almost certainly an underestimate, because one third of patients carried multiple deleterious nsRVs. We also found that del(CFHR3-CFHR1) is disease-contributory, whereas patients with aHUS who carry more mutations are less likely to have CFHR deletions and vice versa.

Figure 6 shows a modified model of aHUS pathogenesis that integrates our results and recent findings in heme as a secondary hit. The key concept in this model is the dysregulation loop, which can be activated spontaneously or by trigger
thus closing the loop. Dysregulation of complement and/or coagulation typically caused by genetic abnormalities maintains the cycle.

We noted two important limitations in this study. First, the relatively small sample size reduces statistical power to uncover minor genetic contributions. Second, we focused on exonic variants and indels at the exclusion of intronic and untranslated region variants, which could be used to identify haplotypes that may contribute to aHUS. Although a more comprehensive genetic screen in large cohort of aHUS patients will remediate these shortcomings, this study is noteworthy for implicating the coagulation pathway and PLG in particular as important contributors to aHUS.

CONCISE METHODS

DNA Extraction, Targeted Genomic Enrichment, and Massively Parallel Sequencing

Genomic DNA was extracted from peripheral blood using the Gentra Puregene Kit (Qiagen Inc., Valencia, CA), and quality was verified by 1% agarose gel electrophoresis and NanoDrop 1000 spectrophotometry (Thermo Scientific Inc., Wilmington, DE). Before sequence capture, DNA concentration was determined using the Qubit dsDNA BR Assay Kit (Invitrogen Inc., Carlsbad, CA).

Targeted genomic enrichment was performed using CasCADE, an RNA-bait panel based on the Agilent SureSelect Target Enrichment System (Agilent Technologies Inc., Santa Clara, CA) that targets coding sequence of genes in the complement and coagulation pathways. CasCADE was developed and optimized to capture 1071 exons of 85 complement and coagulation genes representing 250 kB genomic DNA (Supplemental Figure 1). In CasCADE versions 1 and 2, baits were balanced based on sequencing results of test samples to optimize efficiency and accuracy of targeted enrichment, and in this study, we used CasCADE version 3 to perform capture and enrichment of 2 µg genomic DNA using a solution-based targeted genomic enrichment protocol as previously described.42 Multiplexing indexes were used to barcode prepared libraries. Quality and concentration were evaluated using a Bioanalyzer 2100 (Agilent Technologies Inc., Santa Clara, CA). After pooling, libraries were sequenced with a 2× 100-bp paired-end module in one lane on a HiSeq 2000 (Illumina Inc., San Diego, CA) at the DNA Core Facility of the University of Iowa’s Institute of Human Genetics.
Variant Calling and Filtering

Sequence data were analyzed using a local installation of the open-source Galaxy software running on a high-performance computing cluster at the University of Iowa. A dedicated compute queue for Galaxy as well as shared compute resources were leveraged for minimized job wait time and continuous availability. We use a combination of custom and publicly available tools on the pipeline running in Galaxy: read mapping was performed with Burrows-Wheeler Alignment, duplicate removal was performed with Picard, local realignment and variant calling were performed with GATK, enrichment statistics were performed with NGSRich, and variant annotation was performed with ANNOVAR.43,44

During hard filtering, we used quality control metrics that included a phred-scaled quality score $\geq 30$, depth $\geq 10$, and Q/D $\geq 5$, and then focused on the following variants: nonsynonymous, frame shift, splicing site changes, and indels. Common variants from either the EVS29 or the 1000 Genomes Project (April of 2012)45 with a minor allele frequency value $\leq 1\%$ in any population were excluded. We defined as novel those variants not reported in EVS, 1000 Genomes, National Center for Biotechnology Information dbSNP (build 137), aHUS mutation database (www.fh-hus.org), or our in-house database. To assign pathogenicity, we combined multiple functional prediction methods, including PhyloP, SIFT, PolyPhen2, LRT, MutationTaster, and GERP++. Prediction results for each variant were integrated to generate a pathogenicity score, which was the sum of tools predicting the variant to be deleterious. Missense nonsynonymous...
single nucleotide variants were considered likely deleterious when the PS≥4. Splicing site variants, nonsense mutations, and indels were not assigned a PS.

**Multiplex Ligation-Dependent Probe Amplification**
CNVs were evaluated over the contiguous **CFH-CFHR** genomic region using multiplex ligation-dependent probe amplification. Probes were designed according to MRC-Holland’s guidelines (MRC-Holland, Amsterdam, The Netherlands) or obtained from the literature. CNVs were assigned by data comparison of normalized peak areas generated in five control samples.

**Statistical Analyses**
Data were analyzed using R (version 2.15.1). The Fisher exact test was used to compare allelic frequencies in cases and controls. All tests were two-tailed, and $P$ values<0.05 were considered significant.

**ACKNOWLEDGMENTS**
This study was supported in part by the Foundation for Children with Atypical HUS.

**DISCLOSURES**
None.

**REFERENCES**


32. 1514, 2008

33. 1073, 2006

34. 50: 5070–5079, 2009

35. 261, 2008

36. 261, 2008

37. 205, 2001

38. 332–344, 1999

39. 201–205, 2001

40. 2057, 1992

41. 1177–1177, 2006

42. 5067, 2010

43. 1073, 2006

44. 345–357, 2009

45. 1512

46. 17: 2017

47. 111: 1512

48. 2012

49. 1166–1666, 1993


This article contains supplemental material online at http://jasn.asnjournals.org/lookup/suppl/doi:10.1681/ASN.2013050453/-/DCSupplemental.