Functional Evaluation of Rare Variants in Complement Factor I using a Minigene Assay

Cobey JH Donelson[1,2], Richard JH Smith[1], Yuzhou Zhang[1]

[1]Molecular Otolaryngology and Renal Research Laboratories, Carver College of Medicine, University of Iowa, USA
[2]Graduate Biomedical Sciences Program, Carver College of Medicine, University of Iowa, USA

Introduction

The complement cascade is activated via one of three pathways: Classical, Lectin, and Alternative. Its activation converges at the cleavage of C3 into C3a and C3b. C3b can self-amplify or trigger activation of the terminal pathway and lead to the formation of the membrane attack complex (MAC, C5b-9). The complement cascade is an essential component of the innate immune system, however, over activation of the complement cascade can result in various complement-mediated diseases.

Complement Factor I (FI) is an important serine protease that directly regulates complement activity by inactivating C3b and C4b. Variants in *cis*-acting elements and canonical sequences in *CFI* can disrupt normal splicing and result in low expression of FI protein. Consequently, reduced FI can lead to the accumulation of C3b deposits on the cell surface and potentially drive the pathogenicity of complement-mediated diseases such as C3 glomerulopathy (C3G), atypical hemolytic uremic syndrome (aHUS), or various forms of recurrent infections.

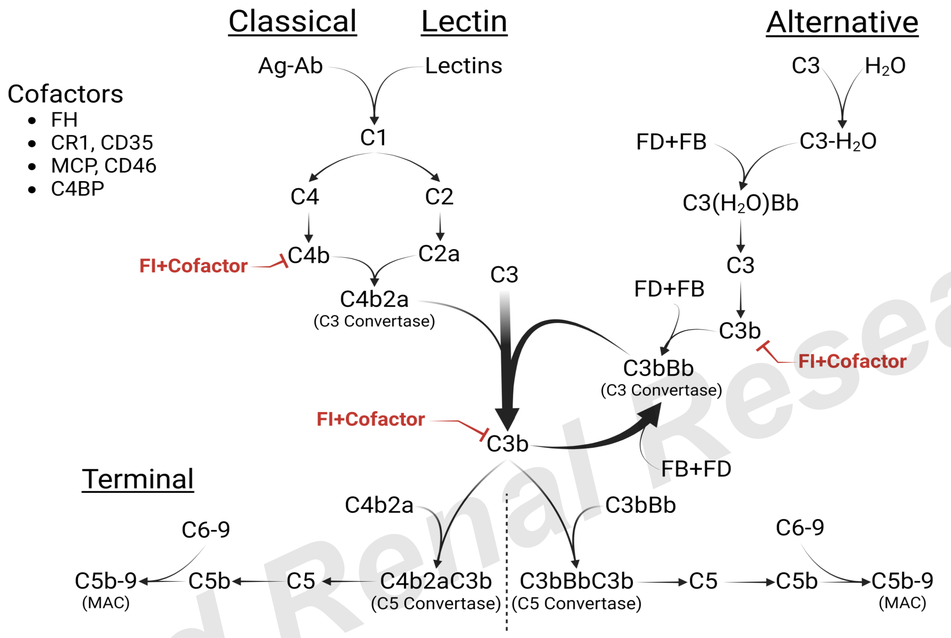

Figure 1. **Complement Cascade.** Complement is activated via the classical, lectin, and alternative pathways (AP). Activation converges at the cleavage of C3 to C3b, where the latter can enter the amplification loop of the AP or progress along the terminal pathway. FI inactivates C4b and C3b and prevents them from continuing in their respective pathway. FI and a cofactor regulate complement by inactivating C3b and C4b.

Methods

1. Identify patients with *CFI* variants in our cohort of C3G and/or aHUS
2. Determine FI expression by ELISA and/or RID
3. Clone DNA fragment into the pET01 plasmid
4. Transfect the plasmid into a mammalian cell line
5. Isolate transcribed RNA and perform RT-PCR
6. Perform Sanger sequencing

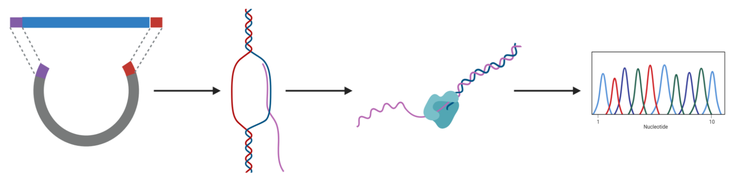

Figure 2. **Experimental design.** Variants in *CFI* were amplified and cloned into the pET01 plasmid via a recombination method. After transfection in mammalian cells, transcribed RNA was isolated and reverse transcribed into cDNA. cDNA was amplified, sequenced and analyzed for defects in splicing.

Results

Table 1. **Patients and Variants**

Patient	Sex	Variant	Amino Acid	Zygosity	FI Levels (mg/L)*	Diagnosis
1	F	c.355G>A	p.Gly119Arg	Het	22.1	aHUS
2	F	c.355G>A	p.Gly119Arg	Het	23.7	aHUS
3	F	c.355G>A	p.Gly119Arg	Het	-	aHUS
4	F	c.355G>A	p.Gly119Arg	Het	-	C3G
5	F	c.355G>A; c.859G>A	p.Gly119Arg; p.Gly287Arg	Comp Het	7.8	aHUS
6	M	c.355G>A	p.Gly119Arg	Het	16.6	C3G
7	F	c.355G>A	p.Gly119Arg	Het	-	aHUS
8	M	c.472G>A; c.1246A>C	p.Gly158Arg; p.Ile416Leu	Comp Het	10.2	aHUS
9	M	c.950G>A	p.Arg317Gln	Het	17.8	aHUS
10	F	c.950G>A	p.Arg317Gln	Het	-	Unknown
11	F	c.1429G>C	p.Asp477His	Het	16.8	aHUS
12	F	c.1429G>C	p.Asp477His	Het	26	C3G/C3GN
13	M	c.1429G>C	p.Asp477His	Het	15.8	aHUS
14	F	c.1429+1G>C		Hom	Undetectable	Recurrent Infection
15	M	c.1429+1G>C		Het	-	C3G/C3GN

*Reference range: 18-44 mg/L

Table 2. **Variant information**

Variant	Genomic Position	MAF*	ClinVar**	Franklin	SpliceAI Score
c.355G>A	4-110685820-C-T	$4.25x10^{-4}$	Conf. Interp.	LP	0.41
c.472G>A	4-109764547-C-T	$1.19x10^{-5}$	Not Reported	VUS	(-0.01) & 0.30
c.950G>A	4-110670749-C-T	$2.12x10^{-5}$	VUS	VUS	(-0.36)
c.1429G>C	4-110667378-C-G	$1.99x10^{-5}$	VUS	VUS	(-0.75) & 0.37
c.1429+1G>C	4-110667377-C-G	$2.83x10^{-5}$	P/LP	P	(-0.77) & 0.41

*collected from the gnomAD database (GRCh37/hg19; ENSG00000205403.8).
**Conf. Interp. = Conflicting Interpretations

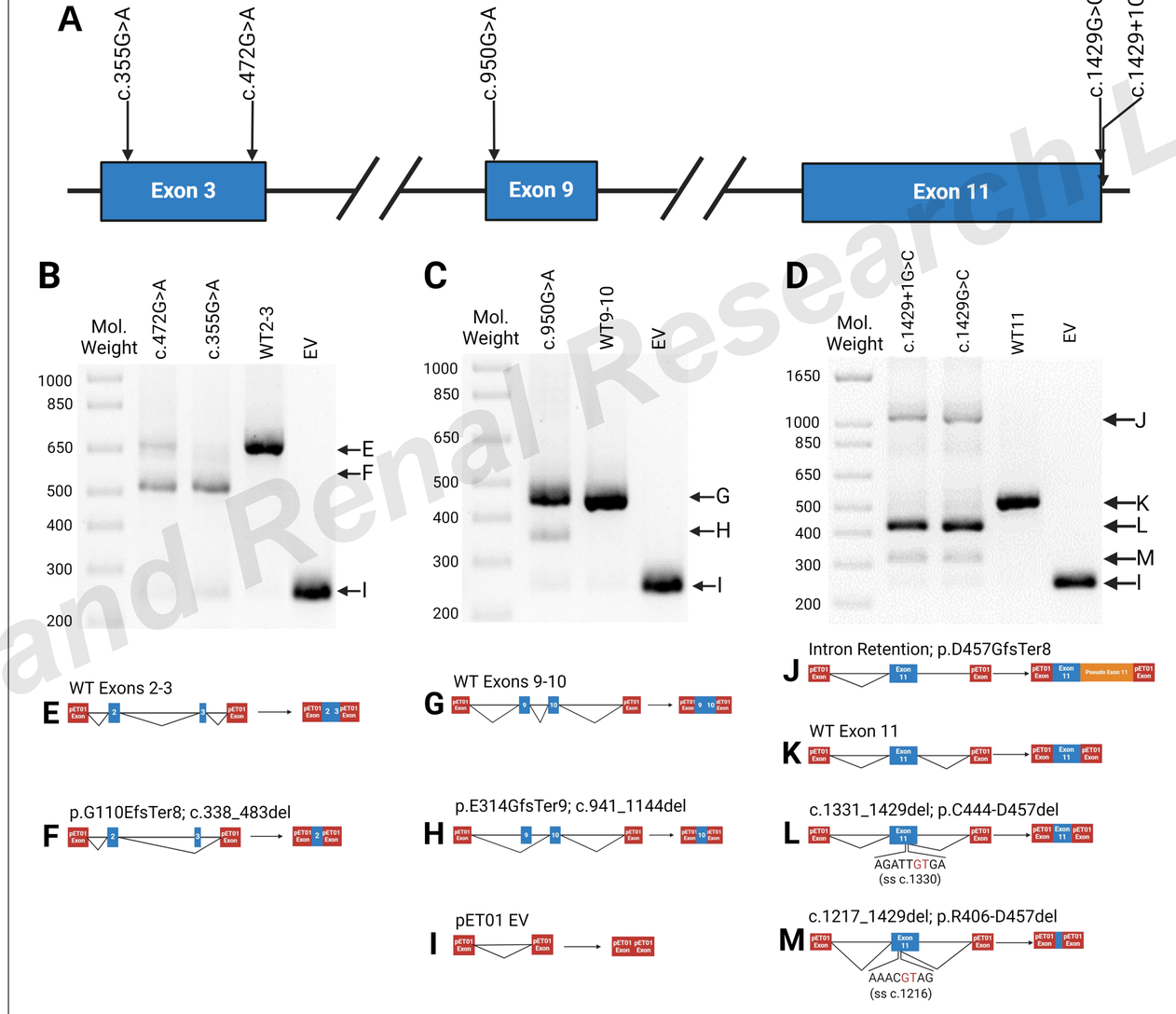

Figure 4. **Impact of Variants. (A)** Variant location. **(B-D)** RT-PCR results. A 1.5% agarose gels depicting the impact of c.472G>A, c.355G>A, c.950G>A, c.1429+1G>C, & c.1429G>C. **(E-F & I)** Visual depictions show Sanger sequencing displaying the affects of the exonic variants, c.472G>A and & c.355G>A, arranged from the uppermost to lowermost band. **(G-I)** Visual depictions show Sanger sequencing displaying the affect of the exonic variant, c.950G>C, arranged from the uppermost to lowermost band. **(J-M & I)** Visual depictions show Sanger sequencing displaying the affects of the splice site variants, c.1429+1G>C & c.1429G>C, arranged from the uppermost to lowermost band. The amino acid consequences are shown above the Sanger sequencing results.

Discussion

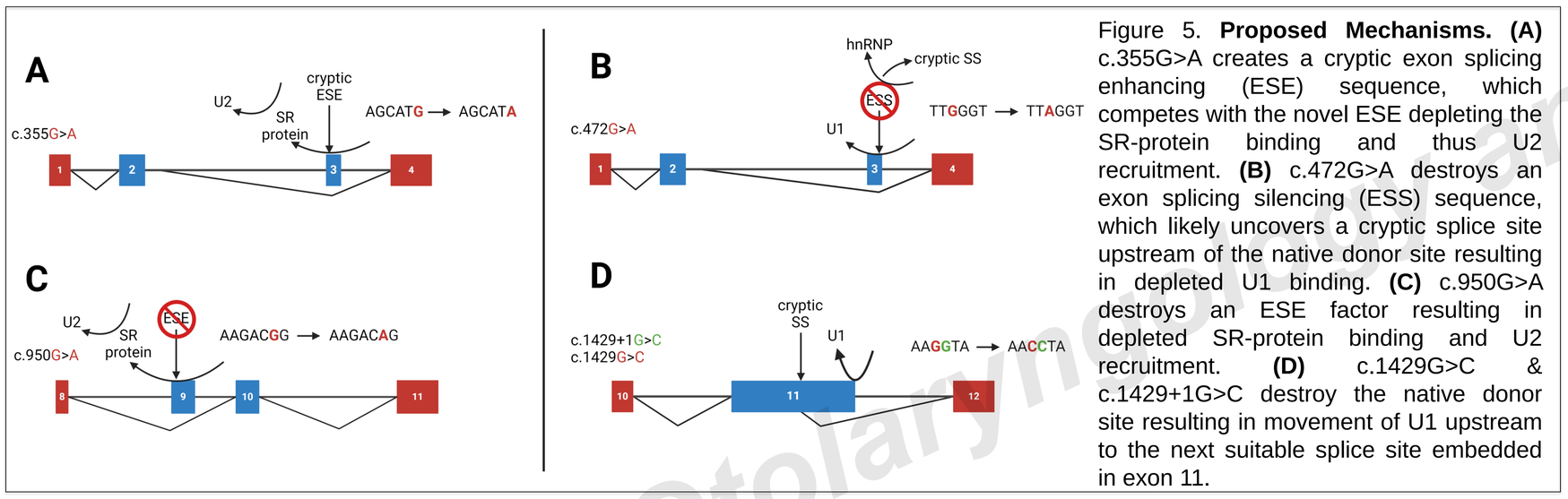

Figure 5. **Proposed Mechanisms. (A)** c.355G>A creates a cryptic exon splicing enhancing (ESE) sequence, which competes with the novel ESE depleting the SR-protein binding and thus U2 recruitment. **(B)** c.472G>A destroys an exon splicing silencing (ESS) sequence, which likely uncovers a cryptic splice site upstream of the native donor site resulting in depleted U1 binding. **(C)** c.950G>A destroys an ESE factor resulting in depleted SR-protein binding and U2 recruitment. **(D)** c.1429G>C & c.1429+1G>C destroy the native donor site resulting in movement of U1 upstream to the next suitable splice site embedded in exon 11.

Conclusion

This study shows how five ultrarare variants in *CFI*–c.355G>A, c.472G>A, c.950G>A, c.1429G>C, c.1429+1G>C–result in aberrant splicing. The splice site variants—c.1429G>C & c.1429+1G>C—destroy the native splice site resulting in improper RNA splicing, while the remaining variants—c.355G>A, c.472G>A, and c.950G>A—disrupt native *cis*-acting elements, and their corresponding *trans*-acting elements, and disrupt normal pre-mRNA splicing functionally resulting in null alleles and haploinsufficiency of *CFI,* the consequence of which is less control of complement cascade. These findings highlight the importance of functional studies to understand the impact of rare genetic variants in *CFI* in patients diagnosed with complement-mediated diseases.

Acknowledgements

National Institutes of Health R01 DK110023 to RJHS. We are grateful to the patients who entrust us with their care. Images were created using BioRender.com