Assessment of unclassified genetic variants in the Canadian inherited bleeding disorder population using pathogenicity interpretation software

Laura L. Swystun,[1,2] Jayne Leggo,[1,2] Shawn Tinlin,[1,2] Paula James,[2,3] David Lillicrap[1,2]

[1]The National Inherited Bleeding Disorder Genotyping Laboratory
[2]Department of Pathology and Molecular Medicine, Queen's University, Kingston, Ontario
[3]Department of Medicine, Queen's University, Kingston, Ontario

Introduction:
While molecular diagnostics for inherited bleeding disorders is a routine part of clinical practice, the ascription of an uncharacterized variant to a disease phenotype remains an ongoing challenge. Publicly accessible registries that catalogue disease-specific variant data is available for many inherited bleeding disorders including von Willebrand disease and the hemophilias. However, additional lines of evidence are essential for the accurate assessment of variant pathogenicity. This can include the absence of the variant in the normal population, *in silico* analysis of pathogenicity for missense variants, or splice-site analysis for variants localized to intron-exon boundaries. Variant interpretation software integrates applications that query population frequency assessment and *in silico* analytical tools into one interface that allow for rapid characterization of variant significance.

Methods: We evaluated the utility of Alamut™ software for variant pathogenicity interpretation in cases where the evidence for pathogenicity was ambiguous. Population frequency was considered for all variants. Pathogenicity assessments of missense variants were performed using Poly-Phen2, SIFT, MutationTaster, and Align GVGD, which predict pathogenicity by nucleotide/amino acid conservation and amino acid physiochemical differences (Grantham distances). Variants localized to intron/exon boundaries was assessed for their influence on splicing with SpliceSiteFinder, MaxEntScan, NNSPLICE, GeneSplicer and Human Splice Finder programs, with a reduction of 10% in the efficiency of a splice site necessary to consider an effect deleterious. Variants were classified as "pathogenic", "likely pathogenic", "uncertain significance", "likely benign" and "benign" based on the strength of this evidence in accordance with current American College of Medical Genetics guidelines.

Results:

Case No.	Gene	Variant (cDNA)	Variant (protein)	Database Report(s)	Pop Freq.	Missense Score	Splicing Score	Interpretation
1	*F13A1*	c.232C>T	p.(Arg78Cys)	1	–	Likely damaging	–	Likely pathogenic
2	*F5*	c.1837T>C	p.(Cys613Arg)	0	–	Likely damaging	–	Likely pathogenic
		c.4210C>T	p.(Pro1404Ser)	0	7.29%	Tolerated	–	Benign
		c.878C>G	p.(Ser293Cys)	0	–	Likely damaging	–	Likely pathogenic
3	*F8*	c.5012G>A	p.(Arg1671His)	1	–	Uncertain	–	Uncertain significance
		c.4999C>T	p.(Arg1667Trp)	0	–	Likely damaging	–	Likely pathogenic
4	*F8*	c.601+5G>C	N/A	0	–	N/A	-11.8	Likely pathogenic
5	*F9*	c.[839-20insA] +[=]	N/A	6	1.5%	Tolerated	-1.85	Benign
6	*F9*	c.685G>A	p.(Gly229Ser)	0	–	Likely damaging	–	Likely pathogenic
7	*VWF*	c.3251G>A	p.(Cys1804Tyr)	0	–	Likely damaging	–	Likely pathogenic
8	*F7*	c.329G>T	p.(Cys110Phe)	0	–	Likely damaging	–	Likely pathogenic
9	*F7*	c.1123C>T	p.(Arg375Trp)	1	0.031%	Uncertain	–	Uncertain significance

Conclusion: Variant interpretation software can distinguish polymorphisms from rare variants, provide evidence of pathogenicity, and assist in the identification of the causative variant when more than one variants are identified within an individual. These tools may also provide complementary information about the influence of missense and silent variants on splicing. Systematized variant interpretation may also assist in identifying variants that have been previously misidentified as pathogenic in disease-specific databases. Integration of variant interpretation protocols into post-analytic assessments will allow for the provision of more accurate variant assessments reported in the Canadian inherited bleeding disorder patient population.