"Phenotypes of Distinction": When and how to integrate unique phenotypic information into variant interpretation

B. Johnson, S. Topper, S. Aradhya, A. Entezam, Y. Kobayashi, K. Mitchel, C. Morgan, R. Nussbaum, J. Tahiliani, K. Nykamp

Invitae Corporation, San Francisco, CA

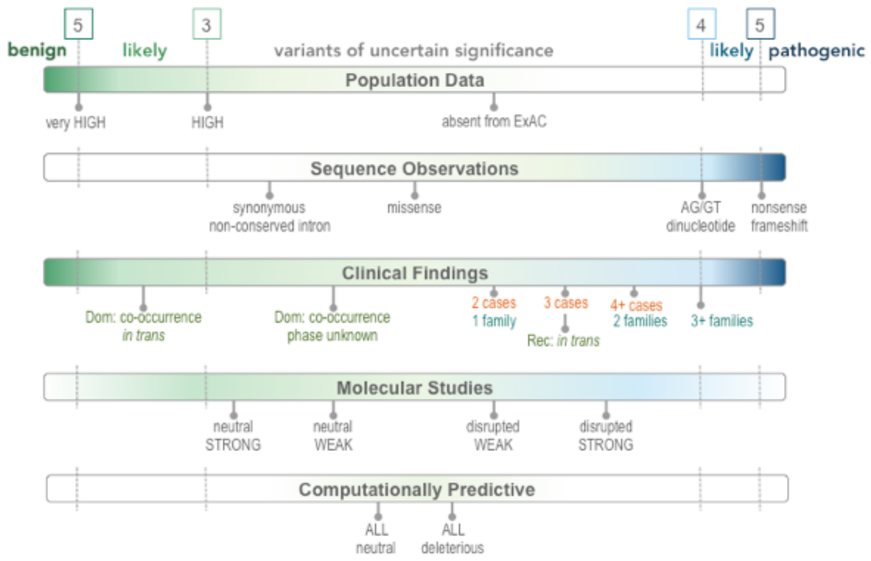
Disclosure statement: All authors are employees and stockholders of Invitae Corporation.

Introduction

A common error made in the course of sequence variant classification is to conclude incorrectly that a novel VUS found in a patient with a disease must be causative. This incorrect conclusion arises from an overestimate of the prior probability that any variants found in a gene associated with the patient's disorder must be causative. In fact, the likelihood that any variants detected in a gene cause disease depends on how distinctive the phenotype of the individual is, the degree of locus heterogeneity, the fraction of locus heterogeneity that is accounted for by the genes being tested, and the prevalence of phenocopies in the population. The likelihood is also modulated by the observed genotype in the patient. The ACMG Interpretation of Sequence Variants guidelines address this subject with a single, somewhat vague criterion: PP4, "Patient's phenotype or family history is highly specific for a disease with a single genetic etiology". Because the presence of a distinctive phenotype in a patient can provide a powerful line of evidence for variant classification, we set out to establish a systematic approach for integrating unique phenotypic data into variant interpretation with more detailed, gene-level guidelines. For well-described hereditary diseases with a specific phenotype, there is a relatively high prior probability that pathogenic variant(s) will be detected, if the appropriate genes have been sequenced.

Methods

- We use a point system for variant interpretation that is based on the ACMG guidelines[1,2]:



- We defined a new set of evidence-based criteria that can be applied during variant interpretation when the following criteria are met: (1) our diagnostic yield is >75% for the gene(s) tested, (2) the clinical features described in a given patient (literature or Invitae patient) must be so specific that they are essentially pathognomonic for the disorder, and (3) the patient's genotype must match the expected inheritance of the disease. The relative weight of the criteria is determined by the likelihood that a patient's observed genotype explains disease:

Inheritance	Path Points	Description
AR, XR	2	Homozygous or hemizygous variant in pathognomonic gene
AR, XR	1.5	Rare het variant co-occurring w/ LP/P variant in pathognomonic gene
AR, XR	1	Rare het variant co-occurring w/ another rare het variant in pathognomonic gene

Results

- To use this criteria, we first pre-curate information on our diagnostic yield of the test given the genes tested and the unique features of the disorder which must be present. An example is shown below:

Test	Genes	Diagnostic yield for the defined disorder (must be ≥ 75%)	Diagnostic Guidelines (aka Minimum REQUIRED features)	Reference
X-linked L1 syndrome	L1CAM	85% if criteria are met (note: yield is only 50% if no affected relative)	1 affected family member (affected = L1 syndrome (3 of 4 features below)) AND 3 of the following 4 features: Hydrocephalus, Adducted thumbs, Agenesis of the corpus callosum, Aqueduct stenosis (MRI)	PMID: 9846429 Describes required features and clinical sensitivity

Results

- As of September 16, 2016, we have used this criteria to aid in the interpretation of 84 samples (77 unique variants)
- As a result of these new criteria, we have been able to classify 30 unique rare variants as likely pathogenic or pathogenic
- In 42 instances, these criteria were applied for variant interpretation and the variant remained a VUS. However, the use of these criteria make it easier to get to LP/P in the future via family variant testing and/or additional cases.
- 21 patients have received a positive genotypic diagnosis for primary ciliary dyskinesia, L1 syndrome, pyridoxine responsive epilepsy, DHPR deficiency, MCAD deficiency, and SCID - they would have received uncertain findings without the use of these criteria:

						New evidence criteria			
Test ordered	Case #	Clinical features	Gene	Variant	Zygos	homozygous/ hemizygous	rare VUS with LP/P	2 rare VUS	Criteria effect/ Interp
ACADM gene	1	Plasma acylcarnitines and urine organic acids are diagnostic for MCAD deficiency including elevated hexanoylglycine and suberylglycine, 4 older siblings with MCAD deficiency	ACADM	c.388-3T>G (Intronic)	Het		X		VUS to LP
				c.985A>G (p.Lys329Glu)	Het				P
ALDH7A1 gene	2	Clinical diagnosis of pyridoxine responsive epilepsy, elevated urine alpha-aminoadipic semialdehyde	ALDH7A1	c.1193G>T (p.Gly398Val)	Het			X	LP to P
				c.986G>A (p.Arg329Lys)	Het			X	VUS to LP
Hyperphe panel	3	Positive NBS for DHPR deficiency with low DHPR enzyme activity on confirmatory testing	QDPR	c.344C>T (p.Ser115Leu)	Homo	X			VUS to P
L1CAM gene	4	Hydrocephalus, aqueductal stenosis, agenesis of corpus callosum, clasped thumbs	L1CAM	c.806+5G>A (Intronic)	Hemi	X			VUS to LP
PCD panel	5	Primary Ciliary Dyskinesia, laterality defect	ARMC4	c.3022-?_*344+?del, Deletion (Exon 20)	Homo	X			VUS to P
	6	Bronchiectasis	CCDC39	c.1167+1261A>G (Intronic)	Homo		X		VUS to LP
	7	Situs inversus	CCDC39	c.1167+1261A>G (Intronic)	Het		X		VUS to LP
				c.357+1G>C (Splice donor)	Het				P
	8	Bronchiectasiss, restrictive lung disease, asthma	CCDC39	c.1167+1261A>G (Intronic)	Homo		X		VUS to LP
	9	Primary Ciliary Dyskinesia, laterality defect	CCDC40	c.2597A>G (p.Asn866Ser)	Het		X		VUS to P
				c.2712-1G>T (Splice Acceptor)	Het				P
	10	Primary Ciliary Dyskinesia	CCDC40	c.2597A>G (p.Asn866Ser)	Het		X		VUS to P
				c.2440C>T (p.Arg814*)	Het				P
	11	Primary Ciliary Dyskinesia	CCNO	c.538dupG (p.Val180Glyfs*55)	Het		X		VUS to P
				c.454G>T (p.Glu152*)	Het				P
	12	Primary Ciliary Dyskinesia, laterality defect	DNAAF3	c.868-?_*401+?del, Deletion (Exons 7-12)	Het		X		VUS to LP
				c.378delG (p.His127Thrfs*34)	Het				P
	13	Primary Ciliary Dyskinesia	DNAH11	c.11062-?_12195+?del , Deletion (Exons 68-74)	Het		x		VUS to LP
				c.6130C>T (p.Arg2044*)	Het				P
	14	Primary Ciliary Dyskinesia	DNAH11	c.7441-?_7914+?del , Deletion (Exons 46-48)	Het		X		VUS to LP
				c.4333C>T (p.Arg1445*)	Het				P
	15	Neonatal respiratory distress, chronic pneumonias, recurrent ear infections, chronic rhinorrhea, situs inversus totalis, chronic cough and sputum; strong family Hx of situs inversus and suspected PCD	DNAH5	c.10226G>C (p.Trp3409Ser)	Het		X		VUS to LP
				c.7034G>A (p.Trp2345*)	Het				P
	16	Primary Ciliary Dyskinesia, laterality defect	DNAH5	c.8030G>A (p.Arg2677Gln)	Het	X	X		VUS to LP
				c.8440_8447del (p.Glu2814*)	Het				P
	17	Primary Ciliary Dyskinesia, laterality defect	DNAH5	c.8010+3A>G (Intronic)	Het		X		VUS to LP
				c.12107G>A (p.Trp4036*)	Het				P
	18	Primary Ciliary Dyskinesia	DNAH5	c.8642C>G (p.Ala2881Gly)	Het		X		VUS to P
				c.8167C>T (p.Gln2723*)	Het				P
	19	Primary Ciliary Dyskinesia	DYX1C1	c.1048-?_*362+?del, Deletion (Exons 9-10)	Het		X		VUS to LP
				c.784-?_893+?del, Deletion (Exon 7)	Het				P
	20	Primary Ciliary Dyskinesia	RSPH9	c.1A>T (Initiator codon)	Het		X		VUS to P
				c.123T>A (p.Tyr41*)	Het				P
SCID panel	21	0 Trecs on newborn screen, T cell - B cell + NK cell - phenotype	IL2RG	c.344G>A (p.Cys115Tyr)	Hemi	X			VUS to LP

Conclusions

- We have developed a systematic framework for the inclusion of highly distinctive phenotypic information when performing variant analysis.
- Careful curation of the disorders for which these criteria can be used is necessary, including the required distinctive phenotypes along with the diagnostic yield for testing individuals with the distinctive phenotypes.
- Each of the new evidence types on their own are insufficient to reach a likely pathogenic interpretation if the variant has only been seen in 1 affected individual; population frequencies, functional studies, and other clinical findings are necessary to reach a likely pathogenic classification.
- This framework provides a mechanism to account for the increased prior probabilities for testing in rare disorders with highly distinctive phenotypes.

References

1. PMID: 25741868
2. https://www.invitae.com/en/variant-classification/

The following transcripts were used in this analysis: ACADM (NM_000016.5), ALDH7A1 (NM_001182.4), ARMC4 (NM_018076.2), CCDC39 (NM_181426.1), CCDC40 (NM_017950.3), CCNO (NM_021147.3), DNAAF3 (NM_001256714.1), DNAH11 (NM_001277115.1), DNAH5 (NM_001369.2), DYX1C1 (NM_130810.3), RSPH9 (NM_152732.4), IL2RG (NM_000206.2), L1CAM (NM_000425.4), QDPR (NM_000320.2)