Evaluating the strength of evidence for gene-condition relationships: impact on multi-gene panel testing design

Christopher Tan, Darlene Riethmaier, Michelle Hogue, Nila Patil, Kristin McDonald Gibson, Tom Winder, Jody Westbrook, Scott Topper

Invitae Corporation, San Francisco, CA

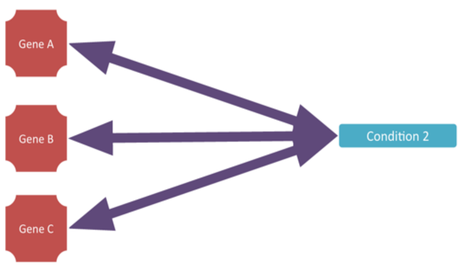
Disclosure statement: All authors are employees and stockholders of Invitae Corporation.

Introduction

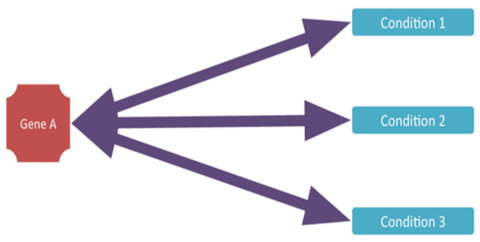
As multi-gene panels have become routine in evaluating patients for genetically heterogeneous conditions, growing inter-laboratory variability has occurred in the number of genes offered in panels for many disorders. Some of this variability can be attributed to the relationship between the timing of assay design and the rapid discovery of new gene–condition relationships; however understanding the clinical validity of the diverse multi-gene panels offered in today's molecular diagnostic setting is becoming increasingly important. Establishing the clinical validity of a multi-gene panel depends on an accurate and detailed understanding of the validity of each included gene. We proposed a method for establishing the clinical validity of genes and evaluated that method with a set of 93 pediatric and neuromuscular gene–condition relationships.

From a strict Mendelian perspective, pathogenic variants in a gene can cause a clinical condition.

A single condition can be caused by mutations in any one of many genes (genotypic heterogeneity).



Different mutations in a single gene can lead to different conditions (phenotypic heterogeneity).

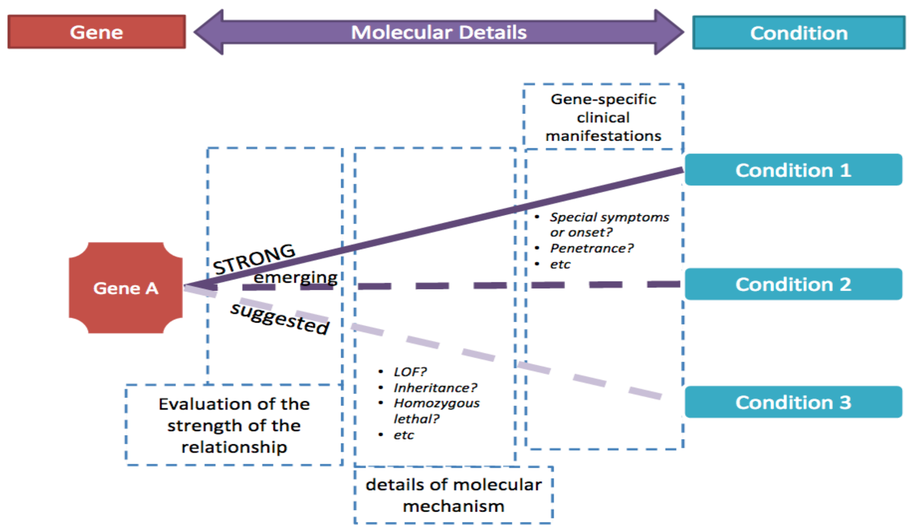


Understanding the strength of evidence supporting the causality of the gene-condition relationship is core to clinical molecular diagnostics.

Methods

A robust variant classification method based on ACMG guidelines was used to evaluate the pathogenicity of published, clinically observed variants. Gene–condition relationships were categorized as follows:

- **Strong** indicates the existence of at least one clinically observed variant supported by sufficient evidence to classify that variant as pathogenic. Strong indicates that the relationship has been proven.
- **Suggested** indicates the existence of some preliminary evidence suggesting a causal relationship; in this scenario there is insufficient evidence to classify clinically observed variants as pathogenic.
- **Emerging** indicates that an *additional* condition purportedly caused by a gene already determined to have a **strong** relationship with a different condition.



Results

We categorized 85 genes as having a strong association with at least one neuromuscular condition (muscular dystrophy, myopathy, congenital myasthenic syndrome). Eight genes were categorized as having only a suggested association with a neuromuscular condition. Three of those genes (SCN4A, SYNE1 and TMEM43) have a strong association with a non-neuromuscular condition.

Gene	Muscular dystrophy	Myopathy	Congenital myasthenic syndrome	MedGen UIDs	Gene	Muscular dystrophy	Myopathy	Congenital myasthenic syndrome	MedGen UIDs
ACTA1		Strong		371799, 108177	KBTBD13		Strong		373095
AGRN			Strong	815069	KLHL40		Strong		815539
ALG2			Suggested	831266	KLHL41		Strong		816714
ANO5	Strong			370102, 331575	LAMA2	Strong			224728
B3GALNT2	Strong			767552	LARGE	Strong			461764, 373284
B4GAT1	Strong			815372	LDB3		Strong		322840
BAG3		Strong		414119	LIMS2	Suggested			25589244
BIN1		Strong		98049	LMNA	Strong			98048, 413212, 320400, 413043
CAPN3	Strong			358391	LMOD3		Strong		830573
CAV3	Strong	Strong		433151	MATR3		Strong		342950
CCDC78		Strong		766623	MTM1		Strong		98374
CFL2		Strong		343979	MUSK			Strong	833690
CHAT			Strong	140751	MYF6		Suggested		482333
CHKB	Strong			355943	MYH7		Strong		449370, 374868
CHRNA1			Strong	373259, 199759	MYOT	Strong	Strong		322957, 331802
CHRNB1			Strong	833647, 833664, 373251	NEB		Strong		342534
CHRND			Strong	833694, 833685, 833675	PLEC	Strong		Strong	21263134, 462339
CHRNE			Strong	373251, 344169, 833673	PNPLA2		Strong		339913
CNTN1		Strong		393406	POMGNT1	Strong			462869, 461762, 461767
COL6A1	Strong	Strong		468393	POMGNT2	Strong			766727
COL6A2	Strong	Strong		468393	POMK	Strong			815294, 808099
COL6A3	Strong	Strong		468393	POMT1	Strong			75553, 461765, 332193
COLQ			Strong	400481	POMT2	Strong			461761, 461766, 461768
CRYAB		Strong		324735	PREPL			Suggested	24610330
DAG1	Strong			851332, 462534	RAPSN			Strong	323066
DES	Strong	Strong		330449, 815467	RYR1		Strong		199773, 108177, 388775, 799613
DMD	Strong			3925, 182959	SCN4A			Suggested	349046
DNAJB6	Strong			460114	SEPN1		Strong		388775, 108177
DNM2		Strong		322437	SGCA	Strong			334108
DOK7			Strong	376880	SGCB	Strong			347674
DPAGT1			Strong	766559	SGCD	Strong			331308
DPM1	Strong			324784	SGCG	Strong			98045
DPM2	Strong			767299	STAC3		Strong		340586
DPM3	Strong			414534	SYNE1	Suggested			414476
DYSF	Strong			419874	TCAP	Strong			400895
EMD	Strong			148284	TIA1		Strong		67441
FHL1	Strong			395525	TMEM43	Suggested			765974
FKRP	Strong			461763, 335764, 339580	TMEM5	Strong			767295
FKTN	Strong			140820, 413465, 370585	TNNT1		Strong		344273
FLNC		Strong		372186	TNPO3	Strong			333983
GAA	Strong			5340	TOR1AIP1	Suggested			891994
GFPT1			Strong	350478	TPM2		Strong		324513, 108177
GMPPB	Strong		Suggested	815546, 815551, 811507	TPM3		Strong		373089, 108177
GNE		Strong		322174	TRAPPC11	Strong			815566
ISPD	Strong			766244, 807556	TRIM32	Strong			78750
ITGA7	Strong			413044	TTN	Strong	Strong		333047, 324741, 350930
					VCP		Strong		322251

Example of establishing a strong relationship between a gene and a single condition.

The COL6A1 gene has long been considered a cause of type VI collagenopathies, a relationship illustrated by an often observed pathogenic p.Gly290Arg variant that is absent from control populations and has been shown to segregate strongly with type VI collagenopathies in affected families (PMID:15689448, 24038877, 18825676). The glycine residue (p.Gly290) affected by this missense change lies within the triple helix region. Glycine residues within this region are crucial to maintaining fibrillar collagen structure and stability (PMID:7695699, 19344236). In COL6A1, missense substitutions that affect glycine residues within this region are reported in many affected individuals (PMID:24038877). This pathogenic variant in COL6A1 causes type VI collagenopathies, and the link between gene and condition is therefore established.

Example of establishing a suggested relationship between a gene and a single condition.

LIMS2 is suspected but not yet proven to cause limb-girdle muscular dystrophy. A single publication reports two afffected siblings carrying compound heterozygous missense variants in LIMS2 (PMID:25589244); however, a detailed evaluation of the underlying evidence regarding for each of these variants leads us to conclude that they should be classified as variants of uncertain significance. Currently no clinically observed variants in LIMS2 can be classified as pathogenic, and therefore we cannot be certain that pathogenic variants in LIMS2 cause LGMD. Therefore, we classify the relationship between LIMS2 and LGMD as 'suggested'.

Conclusions

We created a framework for categorizing gene–condition relationships, that establishes a method for distinguishing between genes proven to cause a condition and genes for which only preliminary evidence suggests an association. Although testing a gene before its clinical validity is conclusively established has legitimate benefits, understanding the rationale for its inclusion on a panel is critical for clinicians. The clinical utility of the findings in any gene ultimately depends on the strength of the evidence linking that gene to disease.