

Evaluating the Clinical Validity of Gene-Disease Associations: An Evidence-Based Framework Developed by the Clinical Genome Resource

Natasha T. Strande,^{1,14} Erin Rooney Riggs,^{2,14} Adam H. Buchanan,³ Ozge Ceyhan-Birsoy,^{4,5,6,7} Marina DiStefano,⁴ Selina S. Dwight,⁸ Jenny Goldstein,¹ Rajarshi Ghosh,⁹ Bryce A. Seifert,¹ Tam P. Sneddon,⁸ Matt W. Wright,⁸ Laura V. Milko,¹ J. Michael Cherry,⁸ Monica A. Giovanni,³ Michael F. Murray,³ Julianne M. O'Daniel,¹ Erin M. Ramos,¹⁰ Avni B. Santani,^{11,12} Alan F. Scott,¹³ Sharon E. Plon,⁹ Heidi L. Rehm,^{4,5,6,7} Christa L. Martin,^{2,3,*} and Jonathan S. Berg^{1,*}

With advances in genomic sequencing technology, the number of reported gene-disease relationships has rapidly expanded. However, the evidence supporting these claims varies widely, confounding accurate evaluation of genomic variation in a clinical setting. Despite the critical need to differentiate clinically valid relationships from less well-substantiated relationships, standard guidelines for such evaluation do not currently exist. The NIH-funded Clinical Genome Resource (ClinGen) has developed a framework to define and evaluate the clinical validity of gene-disease pairs across a variety of Mendelian disorders. In this manuscript we describe a proposed framework to evaluate relevant genetic and experimental evidence supporting or contradicting a gene-disease relationship and the subsequent validation of this framework using a set of representative gene-disease pairs. The framework provides a semiquantitative measurement for the strength of evidence of a gene-disease relationship that correlates to a qualitative classification: “Definitive,” “Strong,” “Moderate,” “Limited,” “No Reported Evidence,” or “Conflicting Evidence.” Within the ClinGen structure, classifications derived with this framework are reviewed and confirmed or adjusted based on clinical expertise of appropriate disease experts. Detailed guidance for utilizing this framework and access to the curation interface is available on our website. This evidence-based, systematic method to assess the strength of gene-disease relationships will facilitate more knowledgeable utilization of genomic variants in clinical and research settings.

Introduction

The human genome comprises approximately 20,000 protein-coding genes (see OMIM website in [Web Resources](#)), of which about 3,000 have been reported in association with at least one Mendelian disease.¹ Roughly half¹ of these gene-disease relationships have been identified over the last decade, as technological advances have made it possible to use sequence information from small families or even single individuals to discover new candidate gene-disease relationships.^{2,3} However, there is substantial variability in the level of evidence supporting these claims, and a systematic method for curating and assessing evidence is needed.

Despite this variability, clinical laboratories may include genes with preliminary evidence of a gene-disease relationship on disease-targeted panels or in results returned from exome or genome sequencing. Some of the gene-disease relationships are either unable to be confirmed for many years or are ultimately proven wrong.⁴ Evaluating the clinical impact of variants identified in genes with an unclear

role in disease is exceedingly difficult and could lead to incorrect diagnoses, preventing further evaluations and/or resulting in errant management of the affected individual and their families. This scenario highlights the need for a standardized method to evaluate the evidence implicating a gene in disease and thereby determine the clinical validity² of a gene-disease relationship.

The NIH-funded Clinical Genome Resource (ClinGen)⁵ is creating an open-access resource to better define clinically relevant genes and variants based on standardized, transparent evidence assessment for use in precision medicine and research. Our group has developed a method that (1) qualitatively defines gene-disease clinical validity using a classification scheme based on the strength of evidence supporting the relationship and (2) provides a standardized semiquantitative approach to evaluate available evidence and arrive at such a classification. Currently, this framework is optimized for genes associated with monogenic disorders following autosomal dominant, autosomal-recessive, or X-linked inheritance. Future iterations will expand the framework to consider other modes of

¹Department of Genetics, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC 27514, USA; ²Autism & Developmental Medicine Institute, Geisinger Health System, Danville, PA 17837, USA; ³Genomic Medicine Institute, Geisinger Health System, Danville, PA 17822, USA; ⁴Laboratory for Molecular Medicine, Partners Personalized Medicine, Boston, MA 02139, USA; ⁵The Broad Institute of Harvard and MIT, Cambridge, MA 02142, USA; ⁶Harvard Medical School, Boston, MA 02115, USA; ⁷Department of Pathology, Brigham & Women's Hospital, Boston, MA 02115, USA; ⁸Department of Genetics, Stanford University, Stanford, CA 94305, USA; ⁹Department of Pediatrics, Baylor College of Medicine, Houston, TX 77030, USA; ¹⁰National Human Genome Research Institute, NIH, Bethesda, MD 20892, USA; ¹¹Department of Pathology and Laboratory Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA; ¹²Division of Genomic Diagnostics, Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA; ¹³McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD 21287, USA

¹⁴These authors contributed equally to this work

*Correspondence: clmartin1@geisinger.edu (C.L.M.), jonathan_berg@med.unc.edu (J.S.B.)

<http://dx.doi.org/10.1016/j.ajhg.2017.04.015>.

© 2017 American Society of Human Genetics.

inheritance, such as mitochondrial, and diseases with more complex genomic etiologies, including oligogenic or multifactorial conditions. Our approach is intended to neither define multifactorial disease risk nor to be a substitute for well-established statistical thresholds used for genome-wide association studies.^{6,7}

This novel framework classifies gene-disease relationships by the quantity and quality of the evidence supporting such a relationship. It builds on efforts to catalog gene-disease associations, such as the Online Mendelian Inheritance in Man (OMIM) and OrphaNet (see [Web Resources](#)), by systematically organizing the supporting and refuting evidence and then categorizing the strength of evidence supporting these relationships. The resulting clinical validity classifications are valuable to both clinicians and clinical laboratories. First, they provide insight into the strength of clinical associations for clinicians interpreting genetic test results for clinical care. Second, they serve to guide clinical genetic testing laboratories as they develop disease-specific clinical genetic testing panels or interpret genome-scale sequencing tests. By including only those genes with established clinical validity, the possibility of returning ambiguous, incorrect, or uninformative results is reduced, improving the quality of interpretation of genomic data.

Material and Methods

Qualitative Description: Clinical Validity Classifications

The ClinGen Gene Curation Working Group (GCWG) is comprised of medical geneticists, clinical laboratory diagnosticians, genetic counselors, and biocurators with broad experience in both clinical and laboratory genetics. Over the course of 3 years, this group convened bi-monthly to develop the described framework for assessing gene-disease clinical validity through expert opinion and working group consensus. We first defined six classes to qualitatively describe the strength of evidence supporting a gene-disease association ([Figure 1](#)). The amount and type of evidence required for each clinical validity classification builds upon that of the previous classification level. Evidence used within this framework to assign a classification to a gene-disease pair is divided into two main types: genetic evidence and experimental evidence (described below). As evidence is likely to change over time, any given classification is representative only of the level of evidence at the time of curation.

The classification “No Reported Evidence” is used for genes that have not yet been asserted to have a causal relationship with a human monogenic disorder but may have some experimental data (e.g., model system data) suggesting a potential role for that gene in disease. The “Limited” classification requires at least one variant, asserted to be disease causing, to have plausible genetic evidence to support the association with human disease with or without gene-level experimental data. “Moderate” classification encompasses additional clinical evidence (e.g., multiple unrelated probands harboring variants with potential roles in disease) and supporting experimental evidence, all of which may be provided by multiple studies or a single robust study. Replication of the gene-disease association in subsequent independent publications and additional substantial genetic and experimental data are crit-

ical factors for the “Strong” classification. Finally, the hallmark of a “Definitive” gene-disease association is that, in addition to the accumulation of convincing genetic and experimental evidence, the relationship has been replicated and ample time has passed since the initial publication (in general, greater than 3 years) for any conflicting evidence to emerge. It is important to highlight that these classifications do not reflect the effect size or relative risk attributable to variants in a particular gene, but instead the strength of the evidence. For example, a definitive gene-disease association does not imply that a pathogenic variant in that gene confers 100% penetrance of the phenotype. This metric is not intended to assess the penetrance or risk to develop a disease outcome.

A gene-disease relationship can be determined to have one of the above classifications provided no substantial relevant and valid contradictory evidence exists to call the gene-disease relationship into question. If such evidence emerges, then the relationship is described as “Conflicting Evidence Reported.” Types of contradictory evidence may come from population studies (such as ExAC⁸), attempts to experimentally validate the gene-disease association, or re-analysis of the original family or cohort that was previously studied. Although the role of a specific *variant* in a given disease may be called into question by new evidence, this may not be sufficient to invalidate the role of the *gene* in that disease. Thorough evaluation by experts in the particular disease area is recommended to determine whether the contradictory evidence outweighs the existing supportive evidence to classify a gene into either a “Disputed” or “Refuted” category (see [Figure 1](#) for additional details).

Semi-Quantitative Assessment of Evidence

Assigning a clinical validity classification to a gene-disease pair requires assessment of the evidence supporting the association. We developed a semiquantitative approach to evaluate both genetic ([Figure 2](#)) and experimental ([Figure 3](#)) evidence in a standardized manner that promotes consistent collection and weighting of evidence (a detailed standard operating procedure is available on the ClinGen website; see [Web Resources](#)). Development of the quantitative aspect of this framework was based on the qualitative descriptions outlined in [Figure 1](#). Both the qualitative classifications and their quantitative counterparts were determined by consensus of the ClinGen Gene Curation Working Group members comprised of a diverse group of genetics experts and professionals with additional input from experts in multiple clinical domains. Throughout development of the framework, several gene-disease pairs (see [Table 1](#)) were iteratively curated as benchmarks with a known “anticipated classification” to determine appropriate scores and assigned ranges (e.g., *FGFR3* [MIM: 134934]:achondroplasia [MIM: 100800]).

Defined sub-categories of genetic and experimental evidence are given a suggested default “score.” However, given that evidence of the same general type may vary in its strength (particularly when considering different diseases), the scoring system also allows these scores to be adjusted within a set range of points, with final approval by experts within the particular disease domain. Finally, the maximum number of points allowed for the various types of genetic and experimental evidence is capped to prevent a preponderance of weak evidence from inappropriately inflating the gene-disease classification. Similarly, certain evidence categories are provided higher maximum scores, allowing key pieces of stronger evidence to proportionately influence the classification of a gene-disease pair.

Evidence Level		Evidence Description
Supportive Evidence	DEFINITIVE	The role of this gene in this particular disease has been repeatedly demonstrated in both the research and clinical diagnostic settings, and has been upheld over time (in general, at least 3 years). No convincing evidence has emerged that contradicts the role of the gene in the specified disease.
	STRONG	The role of this gene in disease has been independently demonstrated typically in at least two separate studies providing strong supporting evidence for this gene's role in disease, usually including both of the following types of evidence: <ul style="list-style-type: none"> Strong variant-level evidence demonstrating numerous unrelated probands with variants that provide convincing evidence for disease causality¹ as well as Compelling gene-level evidence from different types of supporting experimental data². In addition, no convincing evidence has emerged that contradicts the role of the gene in the noted disease.
	MODERATE	There is moderate evidence to support a causal role for this gene in this disease, typically including both of the following types of evidence: <ul style="list-style-type: none"> Several probands with variants that provide convincing evidence for disease causality¹ Moderate experimental data² supporting the gene-disease association The role of this gene in disease may not have been independently reported, but no convincing evidence has emerged that contradicts the role of the gene in the noted disease.
	LIMITED	There is limited evidence to support a causal role for this gene in this disease, such as: <ul style="list-style-type: none"> Fewer than three observations of variants that provide convincing evidence for disease causality¹ OR Variants have been observed in probands, but none have sufficient evidence for disease causality. Limited experimental data² supporting the gene-disease association The role of this gene in disease may not have been independently reported, but no convincing evidence has emerged that contradicts the role of the gene in the noted disease.
NO REPORTED EVIDENCE		Evidence for a causal role in disease has not been reported. These genes might be "candidate" genes based on linkage intervals, animal models, implication in pathways known to be involved in human diseases, etc., but no reports have directly implicated the gene in human disease cases.
Contradictory Evidence	CONFLICTING EVIDENCE REPORTED	Although there has been an assertion of a gene-disease association, conflicting evidence for the role of this gene in disease has arisen since the time of the initial report indicating a disease association. Depending on the quantity and quality of evidence disputing the association, the association may be further defined by the following two sub-categories: <ol style="list-style-type: none"> Disputed <ol style="list-style-type: none"> Convincing evidence <i>disputing</i> a role for this gene in this disease has arisen since the initial report identifying an association between the gene and disease. Refuting evidence need not outweigh existing evidence supporting the gene:disease association. Refuted <ol style="list-style-type: none"> Evidence refuting the role of the gene in the specified disease has been reported and significantly outweighs any evidence supporting the role. This designation is to be applied at the discretion of clinical domain experts after thorough review of available evidence
NOTES		
¹ Variants that disrupt function and/or have other strong genetic and population data (e.g. <i>de novo</i> occurrence, absence in controls, strong linkage to a small genomic interval, etc.) are considered convincing of disease causality in this framework. ² Examples of appropriate types of supporting experimental data based on those outlined in MacArthur et al. 2014.		

Figure 1. ClinGen Clinical Validity Classifications and Qualitative Descriptions

The suggested minimum criteria needed to obtain a given classification are described for each clinical validity classification. The types of evidence comprising these criteria are described in the text. The default classification for genes without a convincing human disease-causing variant is "No Reported Evidence." The level of evidence needed for each supportive gene-disease association category builds upon the previous category (e.g., "Limited" builds upon "Moderate"). Gene-disease associations classified as "Contradictory" likely have supporting evidence as well as opposing evidence, but are described separately from the classifications for supportive gene-disease associations.

Case-Level Data ^A	Evidence Type		Case Information			Suggested Points/Case		Points Given	Max Score
						Default	Range		
	Variant Evidence	Autosomal Dominant OR X-Linked Disorder ^B	Variant is <i>de novo</i> ^C			2	0-3		12
			Proband with predicted or proven null variant ^D			1.5	0-2		10
			Proband with other variant type with some evidence of gene impact ^E			0.5	0-1.5		7
		Autosomal Recessive	Two variants in <i>trans</i> and at least one <i>de novo</i> ^C or a predicted/proven null variant ^D			2	0-3		12
			Two variants (not predicted/proven null) with some evidence of gene impact ^E in <i>trans</i>			1	0-1.5		
	Segregation ^F Evidence	Evidence of segregation in one or more families		LOD Score Examples	3	5	0-7		7
					2	4			
					1.5	3			
1					1.5				
Case-Control Data ^G	Case-Control Study Type ^H	Case-Control Quality Criteria ^I			Suggested Points/Study		Points Given	Max Score	
	Single Variant Analysis ^{Ha}	• Variant Detection Methodology ^{Ia} • Power ^{Ib} • Bias and Confounding Factors ^{Ic} • Statistical Significance ^{Id}			0-6			12	
	Aggregate Variant Analysis ^{Hb}				0-6				
TOTAL ALLOWABLE POINTS for Genetic Evidence									12
General Notes									
<ul style="list-style-type: none">Detailed guidance for utilizing this scoring matrix is available on the ClinGen website in the standard operating procedure.All variants under consideration should be rare enough in the general population to be consistent with disease.Cohorts/cases should not be double counted. For example, individual cases included as part of case-control studies should not be given points from both the “Case Level Data” and “Case-Control Data” categories.Case-Level Data includes studies describing individuals or families with variation in the gene of interest.Case-Control studies are those in which statistical analysis is used to evaluate variation in cases compared to controls.Footnotes A-I are explained in the legend for Figure S2.									

Figure 2. Classes of Genetic Evidence and Their Relative Weights Used in the ClinGen Clinical Validity Framework

For additional points to consider when scoring genetic evidence, please see the standard operating procedure document available on our website. Genetic evidence is separated into two main categories: case-level data and case-control data. While a single publication may include both case-level and case-control data, individual cases should NOT be included in both categories. Each category is assigned a range of points with a maximum score that can be achieved. Case-level data are derived from studies describing individuals and/or families with qualifying variants in the gene of interest. Points should be assigned to each case based on the variant's inheritance pattern, molecular consequence, and evidence of pathogenicity in disease. In addition to variant evidence points, a gene-disease pair may also receive points for compelling segregation analysis (see Figure S1). Case-Control Data: Studies utilizing statistical analysis to evaluate variants in case subjects compared to control subjects. Case-control studies can be classified as either single-variant analysis or aggregate variant analysis, but the number of points allowable for either category is the same. Points should be assigned according to the overall quality of each study based on these criteria: variant detection methodology, power, bias and confounding factors, and statistical power. Note that the maximum total scores allowed for different types of case-level data are not intended to add up to the total points allowed for genetic evidence as a whole. This permits different combinations of evidence types to achieve the maximum total score.

Genetic Evidence

For the purposes of scoring, genetic evidence is divided into two categories: case-level data and case-control data (Figure 2). Studies describing individuals or families with genetic variants are scored as case-level data, while studies using statistical analyses to compare variants in case and control subjects are scored as case-control data. When case-level and case-control data are present in a single publication, points can be assigned in each category, but the same piece of evidence should not be counted more

than once. For example, an individual case that is also included within a case-control cohort should not be given points in both the "case-level data" and "case-control data" categories. In this scenario, points should be assigned to the most compelling and informative evidence.

Assessing case-level data requires consideration of the inheritance pattern and evaluation of the individual variants identified in each case. Within this framework, a case should be counted toward supporting evidence only if the reported variant has some

Evidence Category	Evidence Type	Suggested Points		Points Given	Max Score
		Default	Range		
Function	Biochemical Function	0.5	0-2		2
	Protein Interaction		0-2		
	Expression		0-2		
Functional Alteration	Cells from affected individual	1	0-2		2
	Engineered cells	0.5	0-1		
Models & Rescue	Animal model	2	0-4		4
	Cell culture model system	1	0-2		
	Rescue in animal model	2	0-4		
	Rescue in engineered equivalent	1	0-2		
Total Allowable Points for Experimental Evidence					6

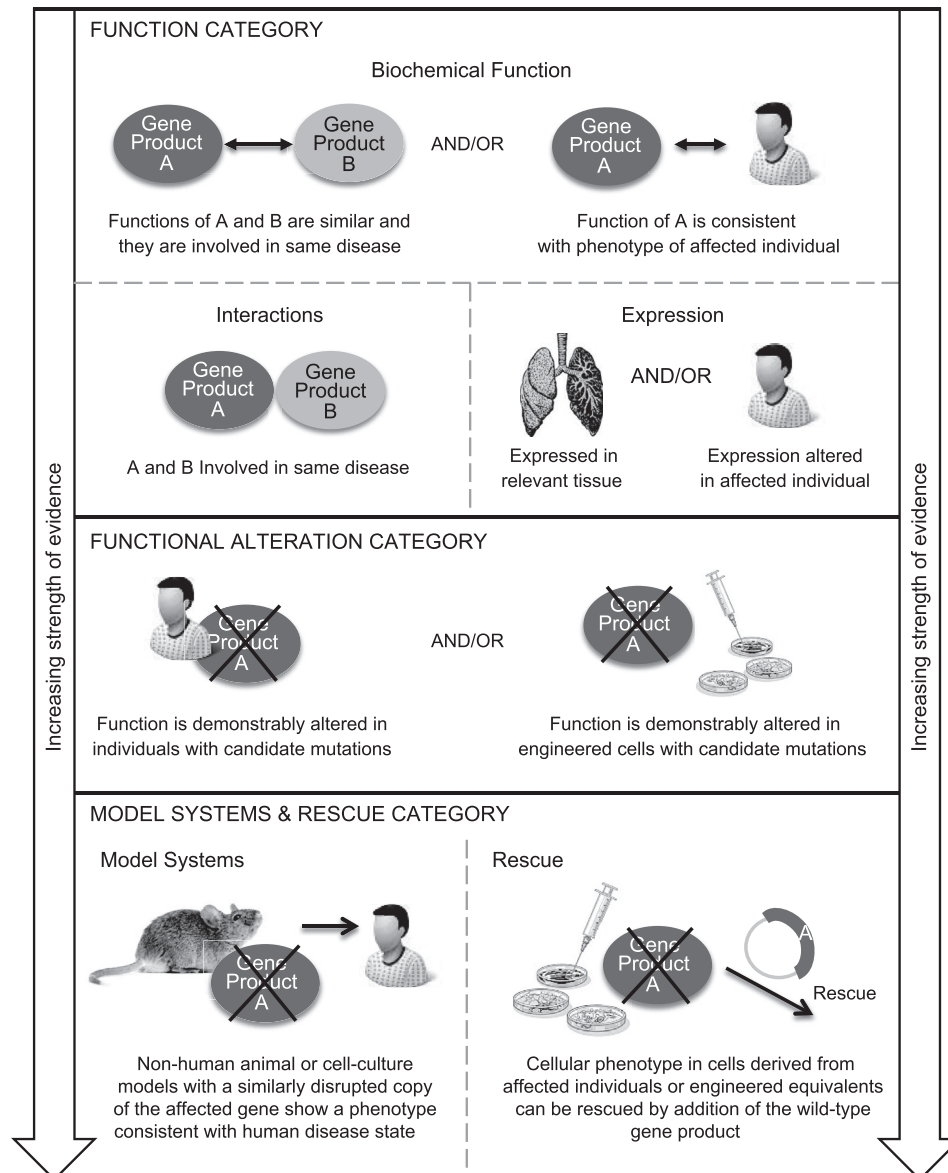


Figure 3. Types of Gene-Level Experimental Evidence and Their Relative Weights Used in the ClinGen Clinical Validity Framework Experimental evidence types used in the ClinGen gene curation framework are modified from MacArthur et al.⁹ Evidence types are divided into three categories based on their relative contribution to the overall clinical validity of a gene-disease pair, giving more weight to in vivo data. Each category is assigned a range of points with a maximum score that can be achieved, allowing more weight to be given

(legend continued on next page)

indication of a potential role in disease (e.g., impact on gene function, recurrence in affected individuals, etc.), does not have evidence that would contradict pathogenicity (e.g., population allele frequency), and is of the type consistent with the assumed disease mechanism (e.g., truncating variant for loss of function). Unless otherwise noted, the term “qualifying variant” implies that these criteria are met. In addition, points are assigned separately for segregation data to reflect the statistical probability that the locus is implicated in the disease. [Figures 2](#) and [S1](#) provide guidance on the number of points that should be considered for segregation evidence by LOD score; if a LOD score is not provided within the publication being evaluated, an estimated LOD score may be calculated in certain scenarios, as described in the standard operating procedure document provided on the ClinGen website.

Each study categorized as “case-control data” should be independently assessed to evaluate the quality of the study design (see [Figure 2](#)). Consultation with a clinical domain expert group (such as those affiliated with ClinGen) is recommended. For the purposes of this framework, studies are classified based on whether they include single-variant analysis or aggregate variant analysis. Single-variant analyses are those in which individual variants are evaluated for statistical enrichment in case subject compared to control subjects. More than one variant may be analyzed, but the variants have been independently assessed with appropriate statistical correction for multiple testing. Aggregate variant analyses are those in which the total number of variants is assessed for enrichment in case subjects compared with control subjects. This comparison is typically accomplished by sequencing the entire gene in both case and control subjects and demonstrating an increased “burden” of variants of one or more types.

Experimental Evidence

The experimental data scoring system is presented in [Figure 3](#). The gene-level experimental data used in this framework to assess a gene-disease association are consistent with those proposed by MacArthur and colleagues to implicate a gene in disease.⁹ The following experimental evidence types are used: biochemical function, experimental protein interactions, expression, functional alteration, phenotypic rescue, and model systems ([Figure 3](#), bottom). These categories capture the most relevant types of experimental information necessary to determine whether the function of the gene product is at least consistent with the disease with which it is associated, if not causally implicated.

Contradictory Evidence

While curators are encouraged to seek out and document (via qualitative description) conflicting evidence, no specific points are assigned to this category. The types of valid contradictory evidence and their relative weights will be unique to each gene-disease pair, and it would be misleading to attempt to uniformly quantify this type of negative evidence against the reported positive evidence. If there is substantial conflicting evidence, manual review and expert input is required to evaluate the strength of the contradictory evidence, determine whether it outweighs any available supporting evidence, and, if so, decide whether the gene-disease association should be classified as “Disputed” or “Refuted.”

Summary and Final Matrix

The scores assigned to both genetic and experimental evidence are tallied to generate a total score (ranging from 1 to 18) that corresponds to a preliminary clinical validity classification ([Figure 4](#)). The system provides a transparent method for summarizing and assessing all curated evidence for a gene-disease pair, encouraging consistency between curators. While the summary matrix facilitates a preliminary assessment of the gene-disease relationship, the initial curator or expert reviewer may adjust the classification, supplying a specific rationale for the change. Final classifications are determined in collaboration with disease experts, who review the preliminary classification and supporting evidence and work to come to a consensus with the preliminary curators. In the event that the disease experts and preliminary curators disagree on a final classification, a senior member of the ClinGen Gene Curation Working Group may be brought in to facilitate a final classification, erring toward the more conservative classification if consensus cannot be achieved. It should be noted that experimental data alone cannot justify a clinical validity classification beyond “No Reported Evidence,” and at least one human genetic variant with a plausible causal association must be present to attain “Limited” classification. The difference between “Limited,” “Moderate,” and “Strong” gene-disease classifications is justified by the quality and quantity of evidence; it is expected that valid gene-disease associations will gradually accumulate enough supporting evidence and be replicated over time to attain a “definitive” classification. This framework relies predominantly on evidence obtained from published primary literature, identified through resources such as PubMed and OMIM (see [Web Resources](#)), and independently assessed by curators; however, if necessary, unpublished information available from publicly accessible resources, such as variant databases,^{10,11} may be used as long as some supporting evidence is provided.

Results

With this framework, we evaluated 33 gene-disease pairs representing a variety of disease domains and spanning the spectrum of clinical validity classifications (see [Table 1](#)). These pairs were intentionally chosen to be representative of the diversity in monogenic disorders with regards to inheritance patterns, disease prevalence, and levels of evidence to support a relationship. To assess the reproducibility of our scoring metric, each gene-disease pair was evaluated by two independent curators; paired curators reached concordant clinical validity classifications in 29 of the 31 (93.5%) gene-disease pairs with available published evidence ([Figure 5](#); associations classified as “No Reported Evidence” were excluded). All major discrepancies between curators were discussed and resolved when possible prior to review by clinical domain experts (either ClinGen Clinical Domain Working Group [CDWG] members or ad hoc disease experts mentioned

to in vivo data (e.g., Models & Rescue) over in vitro experimental data. Evidence within the function category is given the least weight and is comprised of the following types of evidence: biochemical function, interactions, and expression. Functional alteration experiments in cells from affected individuals carrying candidate pathogenic variants are given more weight than the function category. Finally, model systems and phenotypic rescue experiments are given the most weight in our framework. Note that the maximum total scores allowed for different categories of experimental evidence are not intended to add up to the total allowable points. This permits different combinations of evidence types to achieve the maximum total score.

Table 1. Categorization of Gene-Disease Pairs Used to Validate the Gene-Validity Framework

Disease Category	HGNC Gene Symbol	Gene MIM ID	Disease Curated	Inheritance Pattern	Orphanet ID, Phenotype MIM ID	Expert Reviewed Classification ^a
Bone marrow failure	<i>NHP2</i>	606470	dyskeratosis congenita	recessive	ORPHA1775, MIM: 613987	limited
	<i>RAD51C</i>	602774	Fanconi anemia	recessive	ORPHA84, MIM: 613390	moderate
	<i>RPS10</i>	603632	Diamond-Blackfan anemia	dominant	ORPHA124, MIM: 613308	definitive
	<i>RPS24</i>	602412	Diamond-Blackfan anemia	dominant	ORPHA124, MIM: 610629	definitive
	<i>TSR2</i>	300945	Diamond-Blackfan anemia with mandibulofacial dysostosis	X-linked	ORPHA124, MIM: 300946	limited
	<i>WRAP53</i>	612661	dyskeratosis congenita	recessive	ORPHA1775, MIM: 613988	moderate
Cardiovascular disorders	<i>AKAP9</i>	604001	Romano-Ward syndrome	dominant	ORPHA101016, MIM: 611820	limited
	<i>SCN4B</i>	608256	long QT syndrome	dominant	ORPHA768, MIM: 611819	limited
	<i>SMAD3</i>	603109	Loeys-Dietz type 3	dominant	ORPHA284984, MIM: 613795	definitive
	<i>TMPO</i>	188380	familial or idiopathic dilated cardiomyopathy	dominant	ORPHA154, MIM: 613740 ^b	contradictory (refuted)
Hereditary cancer	<i>DICER1</i>	606241	pleuropulmonary blastoma	dominant	ORPHA64742, MIM: 601200	definitive
	<i>PALB2</i>	610355	hereditary breast cancer	dominant	ORPHA227535, MIM: 114480	definitive
	<i>PMS2</i>	600259	hereditary pancreatic cancer	N/A	N/A	no reported evidence
	<i>RAD51D</i>	602954	hereditary breast cancer	dominant	ORPHA227535, MIM: 614291	limited
Immune disorders	<i>CIQB</i>	120570	immunodeficiency due to CIQ deficiency	recessive	ORPHA169147, MIM: 613652	definitive
	<i>CD3E</i>	186830	severe combined immunodeficiency	recessive	ORPHA183660, MIM: 615615	definitive
Skeletal dysplasia	<i>ARSD</i>	300002	chondrodysplasia punctata	N/A	N/A	no reported evidence
	<i>COL2A1</i>	120140	spondyloepiphyseal dysplasia (Stanescu type)	dominant	ORPHA94068, MIM: 616583	moderate
	<i>FGFR3</i>	134934	achondroplasia	dominant	ORPHA15, MIM: 100800	definitive
	<i>LBR</i>	600024	anadysplasia-like, spontaneously remitting spondylometaphyseal dysplasia	recessive	ORPHA448267, none	moderate
Neuromuscular disorders	<i>BAG3</i>	603883	myofibrillar myopathy	dominant	ORPHA593, MIM: 612954	definitive
	<i>MYO9A</i>	604875	arthrogryposis	recessive	ORPHA109007, none	limited
	<i>PSD3</i>	614440	antecubital pterygium syndrome	dominant	ORPHA2987, none	limited
	<i>VPS8</i>	N/A	arthrogryposis	recessive	ORPHA109007, none	limited

(Continued on next page)

Table 1. Continued

Disease Category	HGNC Gene Symbol	Gene MIM ID	Disease Curated	Inheritance Pattern	Orphanet ID, Phenotype MIM ID	Expert Reviewed Classification ^a
Miscellaneous	<i>AGTR2</i>	300034	X-linked non-syndromic intellectual disability	X-linked	ORPHA777, none	contradictory (disputed)
	<i>ATF6</i>	605537	achromatopsia	recessive	ORPHA49382, MIM: 616517	strong
	<i>CHD1L</i>	613039	renal or urinary tract malformation	dominant	ORPHA93545, none	limited
	<i>HNRNPK</i>	600712	Au-Kline syndrome	dominant	ORPHA453504, MIM: 616580	moderate
	<i>LAMB1</i>	150240	lissencephaly 5	recessive	ORPHA352682, MIM: 615191	moderate
	<i>NGLY1</i>	610661	x	recessive	ORPHA404454, MIM: 615273	definitive
	<i>SMARCA1</i>	300012	syndromic intellectual disability with Coffin-Syris-like features	dominant	none, none	moderate
	<i>SKI</i>	164780	Shprintzen-Goldberg	dominant	ORPHA311140, MIM: 182212	definitive
	<i>SOS2</i>	601247	Noonan syndrome	dominant	ORPHA648, MIM: 616559	moderate

Abbreviations: N/A, not applicable.

^aAll gene-disease classifications are accurate as of January 2017.

^bPhenotype MIM was associated with *TMPO* at the time of curation, but has since been removed due to updated information.

in the [Acknowledgments](#)); experts agreed with the preliminary classifications for 87.1% (27/31) of the gene-disease pairs with published evidence ([Figure 5](#)). The four discrepancies between the expert and curator classifications were each different by only a single category (e.g., limited versus moderate). Of note, the original classifications for *HNRNPK* (MIM: 600712) and *SMARCA1* (MIM: 300012) were at the border between limited and moderate (6.5 points); in each case, the preliminary curators' lack of specific clinical expertise led to uncertainty regarding the scoring of evidence requiring such knowledge. Consulting with clinical experts in the disease resolved these issues, resulting in both genes being upgraded to moderate. In the case of *WRAP53* (MIM: 612661), the expert was aware of additional published experimental evidence that when included increased the classification from limited to moderate. Upon reviewing the curated evidence for *RAD51D* (MIM: 602954) and breast cancer (MIM: 614291), the domain expert upgraded the classification from disputed to limited (with the approval of the GCWG) due to the specificity of the experimental evidence and insufficient power of the current studies to rule out a role for *RAD51D* in breast cancer ([Figure 5](#)). Details and references for each curation are provided in supplemental figures ([Figures S2–S65](#)).

Discussion

The evidence-based framework described here qualitatively defines clinical validity classifications for gene-dis-

ease associations in monogenic conditions and provides a systematic framework for evaluating key criteria required for these classifications. This method is intentionally flexible to accommodate curation of a wide spectrum of genes and conditions by curators with varying levels of expertise. The semiquantitative scoring system combined with the qualitative classification scheme guides curators through the preliminary decision-making process, while the expert-level review provides disease-specific experience to weigh in on the final classification.

This effort to create a generalized framework may result in some specific challenges due to the heterogeneity of genetic conditions, in both phenotype and prevalence. For example, conditions that span a large phenotypic spectrum may pose a challenge when defining what constitutes a condition and what is most relevant for curation purposes. In general, ClinGen encourages its expert curation groups to focus on disease associations that have been asserted in the literature or in other authoritative sources (e.g., OMIM, Orphanet Disease Ontology). Expert reviewers may find it useful in certain scenarios to curate both a syndromic disease association as well as an isolated/non-syndromic disease association limited to a particular sub-phenotype, for example, when a disease entity encompasses sub-phenotypes that are caused by different mutational mechanisms. This is a topic of continued discourse within the ClinGen working groups and will be incorporated into future manuscripts that will focus on the curation approach for individual ClinGen disease-focused expert groups.

Clinical Validity Summary Matrix

GENE/DISEASE PAIR:				
Assertion criteria	Genetic Evidence (0-12 points)	Experimental Evidence (0-6 points)	Total Points (0-18)	Replication Over Time (Y/N)
Description	Case-level, family segregation, or case-control data that support the gene-disease association	Gene-level experimental evidence that support the gene-disease association	Sum of Genetic & Experimental Evidence	> 2 publications with convincing evidence over time (>3 yrs)
Assigned Points				
CALCULATED CLASSIFICATION		LIMITED	1-6	
		MODERATE	7-11	
		STRONG	12-18	
		DEFINITIVE	12-18 & Replicated Over Time	
Valid contradictory evidence (Y/N)*	List references and describe evidence:			
CURATOR CLASSIFICATION				
FINAL CLASSIFICATION				

Figure 4. Final Summary Matrix Used to Provisionally Classify Gene-Disease Associations

A summary matrix was designed to generate a “provisional” clinical validity assessment using a point system consistent with the qualitative descriptions of each classification. Genetic evidence: total number of points (not exceeding 12) obtained using the scoring metric in Figure 2. If no human variants associated with disease have been reported in the literature, then the default classification is “No Reported Evidence.” Experimental evidence: total number of points (not exceeding 6) derived from each of the experimental categories in Figure 3. Replication over time: yes, if more than 3 years has passed since the publication of the first paper reporting the gene-disease relationship AND more than two publications with human mutations exist. Contradictory evidence: no points are assigned to this category; instead, the curator should provide a summary of contradictory information. Scoring: the sum of the quantified evidence from each category can be used to determine a “provisional” classification using the scale at the bottom of the figure. If a curator does not agree with this classification, he/she may provide a different suggested classification along with appropriate justification.

Ultra-rare disorders may have a relatively small number of probands described in the medical literature, thus limiting their potential to achieve a high genetic evidence score within this matrix. This obstacle is mostly circumvented by allowing compelling pieces of genetic evidence to score the maximum number of points (for example, see *CD3E* [MIM: 186830] and severe combined immunodeficiency [MIM: 615615] in Figures S14 and S15). When substantial experimental evidence is also available, these conditions can attain a “Strong” or “Definitive” classification. On the opposite end of the spectrum are conditions that occur commonly in the general population, such as cancer, where the predominant etiology is multifactorial rather than monogenic. In the less common Mendelian cancer predisposition syndromes, incomplete penetrance is a typical feature that can lead to confounding factors in family genetic studies such as apparently non-penetrant family members who carry a disease-associated variant and phenocopies among family members without a disease-associated

variant. For such conditions, case-control data may provide more compelling evidence to support the gene-disease association (see Figures S36 and S37 for *PALB2* [MIM: 610355] and hereditary breast cancer [MIM: 114480] as an example).

One limitation of any such system is the challenge of balancing thorough literature curation and practical time commitment. This system can accommodate an exhaustive literature review, but in most cases will require curating only the amount of information sufficient to reach the maximum number of points in the matrix. In some scenarios this method may fail to include pertinent information, which could impact the classification (e.g., omission of contradictory evidence). Another potential limitation is the subjective nature of certain evidence types (e.g., experimental), which may lead to variability between different groups assessing evidence. However, due to the transparency of the evidence base, the incorporation of expert review, and the ability to reassess classifications over time, such drawbacks are likely to be self-limiting.

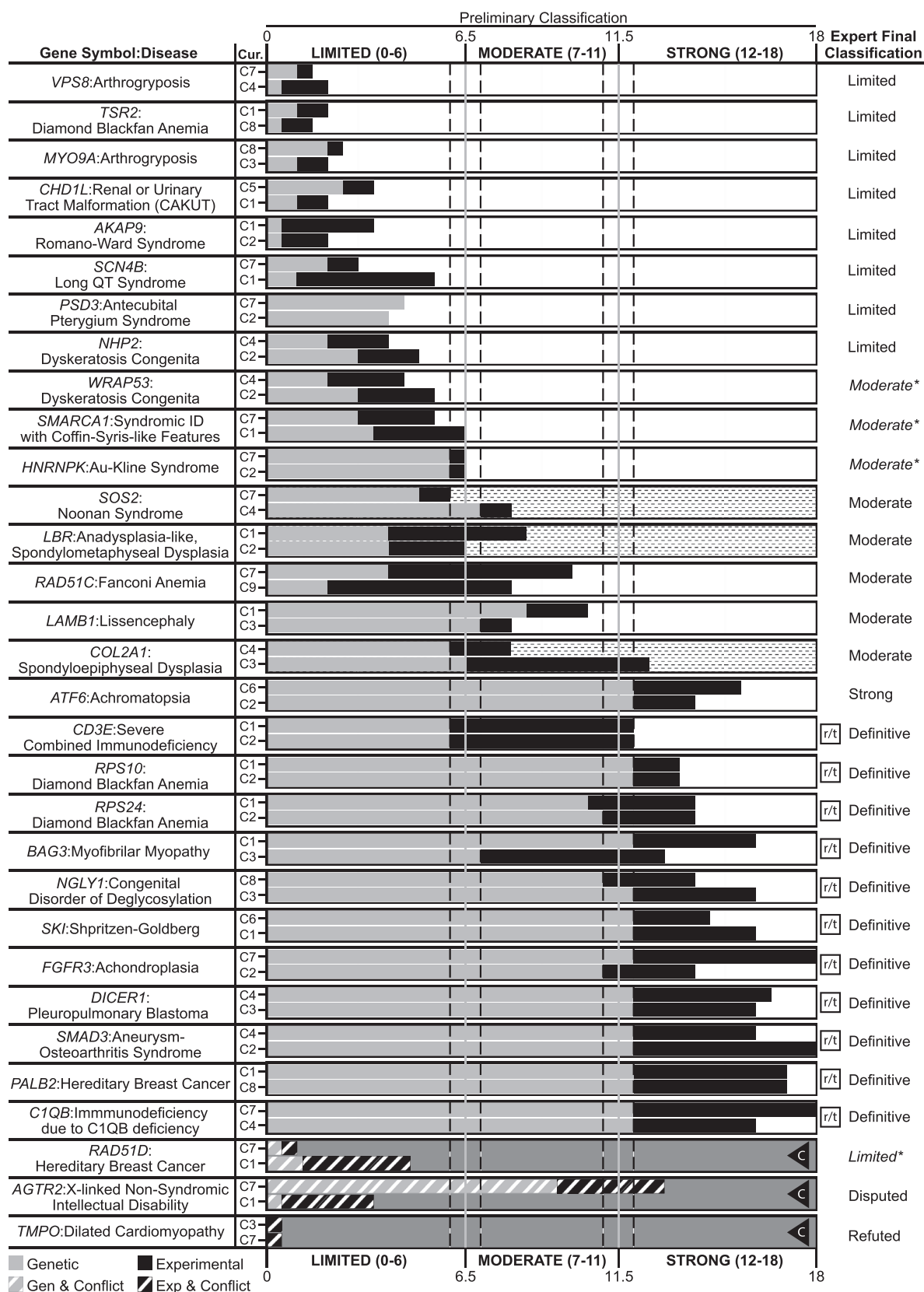


Figure 5. Comparison of Provisional Clinical Validity Classifications and Associated Matrix Scores for Selected Gene-Disease Pairs Evaluated by Multiple Curators

Of the 33 gene-disease pairs (y axis) curated to validate the clinical validity curation framework, 31 were classified using the summary matrix (two gene-disease pairs, *PMS2*:pancreatic cancer and *ARSD*:chondrodysplasia punctata, were classified as “No evidence reported”

(legend continued on next page)

ClinGen's ultimate goal is to enhance the incorporation of genomic information into clinical care, an important component of the Precision Medicine Initiative.¹² The implementation of this framework will be supported by an open-access ClinGen curation interface (under development) that will guide curators through the curation process and will serve as a platform for extension to the community. In essence, this framework aims to provide a systematic, transparent method to evaluate a gene-disease relationship in an efficient and consistent manner suitable for a diverse set of users. A detailed standard operating procedure for this framework is available on the ClinGen website. All curated evidence, including clinical validity assessments, will also be made readily accessible to clinical laboratories, clinicians, researchers, and the community via our website. Additionally, for community members that wish to contribute papers of interest and/or request curation of a gene-disease pair, a "reporter" form is available on the ClinGen website.

Carefully evaluated gene-disease clinical validity classifications, as provided by this framework, will be useful to clinical laboratories as they evaluate genes for inclusion on disease-targeted panels, or as they decide how to categorize, prioritize, and return results from exome/genome sequencing. Clinicians may choose to use these types of gene-disease classifications as they interpret laboratory results for the individuals they care for; for instance, they may choose not to adjust medical management based on variants in genes of limited clinical validity. Researchers could also utilize this framework to evaluate the clinical validity of their own newly discovered associations and identify promising target genes for future work in order to augment the currently available evidence and attain a "Strong" or "Definitive" classification. In addition, professional societies and regulatory bodies may utilize these clinical validity assessments when making recommendations or guidelines for clinical genetic testing. Ultimately, our systematic, evidence-based method for evaluating gene-disease associations will provide a strong foundation for genomic medicine.

Supplemental Data

Supplemental Data include 65 figures and can be found with this article online at <http://dx.doi.org/10.1016/j.ajhg.2017.04.015>.

Acknowledgments

This work was supported by grants U41 HG006834-01A1, U01 HG007437-01, and U01 HG007436-01 from the National Human

Genome Research Institute (NHGRI), the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD), and through contract HHSN261200800001E from the National Cancer Institute (NCI). ClinVar is supported by the Intramural Research Program of the NIH, National Library of Medicine. We would like to thank the following groups and individuals for contributing their disease expertise to review the examples included in this manuscript: Alan Beggs, Alison Bertuch, Rebecca H. Buckley, Eugene Chung, Bill Craigen, Jennifer M. Puck, Sharon A. Savage, Fergus J. Couch, the ClinGen Hereditary Breast and Ovarian Cancer gene curation working group, Birgit H. Funke and the ClinGen Cardiomyopathy gene curation working group, and the ClinGen RASopathy curation working group. Input on the framework was also provided by the ClinGen Hereditary Breast and Ovarian Cancer gene curation working group and Ray Hershberger, Mike Gollob, and the ClinGen Channelopathy gene curation working group. We would also like to thank Scott Goehringer for his invaluable help in preparing the curated examples for the ClinGen website and supplemental figures.

Received: February 22, 2017

Accepted: April 26, 2017

Published: May 25, 2017

Web Resources

ClinGen, <https://www.clinicalgenome.org/>

ClinGen Gene Curation, <https://www.clinicalgenome.org/working-groups/gene-curation/>

ClinGen Gene Curation SOP, <https://www.clinicalgenome.org/working-groups/gene-curation/projects-initiatives/gene-disease-clinical-validity-sop/>

ClinGen Knowledge Base, https://search.clinicalgenome.org/kb/agents/sign_up

OMIM, <http://www.omim.org/>

Orphanet, <http://www.orpha.net/consor/cgi-bin/index.php>

References

1. Chong, J.X., Buckingham, K.J., Jhangiani, S.N., Boehm, C., Sobreira, N., Smith, J.D., Harrell, T.M., McMillin, M.J., Wiszniewski, W., Gambin, T., et al.; Centers for Mendelian Genomics (2015). The genetic basis of Mendelian phenotypes: discoveries, challenges, and opportunities. *Am. J. Hum. Genet.* 97, 199–215.
2. Haddow, J., and Palomacki, G. (2003). ACCE: a model process for evaluating data on emerging genetic tests. In *Human Genome Epidemiology: A Scientific Foundation for Using Genetic Information to Improve Health and Prevent Disease*, M. Khoury, J. Little, and W. Burke, eds. (Oxford University Press), pp. 217–233.

and are not shown). Genetic evidence (gray bars) and experimental evidence (black bars) were evaluated by two independent curators (C1–C9) to arrive at a provisional classification (x axis). Gene-disease relationships scoring between 12 and 18 points can be "Strong" or "Definitive," depending on whether the association has been replicated over time (indicated by the squared "r/t"), in which case the preliminary classification is "Definitive." Clinical validity classifications that were discordant between preliminary curators are represented with a dashed background. Gene-disease pairs in which conflicting evidence was reported are represented by diagonal lines through the evidence bars and a gray background. The letter "C" in a triangle indicates that the curators classified the gene-disease pair as "Conflicting Evidence Reported." Each gene-disease pair was ultimately evaluated by an expert in the field for a final classification (far right column). Final expert classifications that differed from the preliminary classification are indicated by italics and asterisks.

3. Wilfert, A.B., Chao, K.R., Kaushal, M., Jain, S., Zöllner, S., Adams, D.R., and Conrad, D.F. (2016). Genome-wide significance testing of variation from single case exomes. *Nat. Genet.* 48, 1455–1461.
4. Eisenberger, T., Di Donato, N., Baig, S.M., Neuhaus, C., Beyer, A., Decker, E., Mürbe, D., Decker, C., Bergmann, C., and Bolz, H.J. (2014). Targeted and genomewide NGS data disqualify mutations in MYO1A, the “DFNA48 gene”, as a cause of deafness. *Hum. Mutat.* 35, 565–570.
5. Rehm, H.L., Berg, J.S., Brooks, L.D., Bustamante, C.D., Evans, J.P., Landrum, M.J., Ledbetter, D.H., Maglott, D.R., Martin, C.L., Nussbaum, R.L., et al.; ClinGen (2015). ClinGen—the Clinical Genome Resource. *N. Engl. J. Med.* 372, 2235–2242.
6. Lander, E., and Kruglyak, L. (1995). Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat. Genet.* 11, 241–247.
7. Sham, P.C., and Purcell, S.M. (2014). Statistical power and significance testing in large-scale genetic studies. *Nat. Rev. Genet.* 15, 335–346.
8. Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., O’Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B., et al.; Exome Aggregation Consortium (2016). Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 536, 285–291.
9. MacArthur, D.G., Manolio, T.A., Dimmock, D.P., Rehm, H.L., Shendure, J., Abecasis, G.R., Adams, D.R., Altman, R.B., Antonarakis, S.E., Ashley, E.A., et al. (2014). Guidelines for investigating causality of sequence variants in human disease. *Nature* 508, 469–476.
10. Landrum, M.J., Lee, J.M., Riley, G.R., Jang, W., Rubinstein, W.S., Church, D.M., and Maglott, D.R. (2014). ClinVar: public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Res.* 42, D980–D985.
11. Fokkema, I.F., Taschner, P.E., Schaafsma, G.C., Celli, J., Laros, J.F., and den Dunnen, J.T. (2011). LOVD v.2.0: the next generation in gene variant databases. *Hum. Mutat.* 32, 557–563.
12. Collins, F.S., and Varmus, H. (2015). A new initiative on precision medicine. *N. Engl. J. Med.* 372, 793–795.