Published in final edited form as:

Hum Mutat. 2018 November ; 39(11): 1494–1504. doi:10.1002/humu.23604.

The progression of the ClinGen gene clinical validity classification over time

Jennifer L. McGlaughon¹, Jennifer Goldstein¹, Courtney Thaxton¹, Sarah E. Hemphill², and Jonathan S. Berg¹

¹University of North Carolina at Chapel Hill, Chapel Hill, NC

²Laboratory for Molecular Medicine, Partners Healthcare Personalized Medicine, Cambridge, MA

Abstract

In order for ClinGen to maintain up-to-date gene-disease clinical validity classifications for use by clinicians and clinical laboratories, an appropriate timeline for re-evaluating curated gene-disease associations will need to be determined. To provide guidance on how often a gene-disease association should be re-curated, a retrospective analysis of 30 gene curation was performed. Curation were simulated at one-year intervals starting with the year of the first publication to assert disease-causing variants in the gene to observe trends in the classification over time, as well as factors that influenced changes in classification. On average, gene-disease associations spent the least amount of time in the “Moderate” classification before progressing to “Strong” or “Definitive”. In contrast, gene-disease associations that spent five or more years in the “Limited” classification were most likely to remain “Limited” or become “Disputed/Refuted”. Large population datasets contributed to the re-classification of several gene-disease associations from “Limited” to “Disputed/Refuted”. Finally, recent advancements in sequencing technology correlated with an increase in the quantity of case-level evidence that was curated per paper. This study provided a number of key points to consider when determining how often to re-curate a gene-disease association.

Keywords

ClinGen; Clinical Genome Resource; gene-disease association; clinical validity; curation; re-evaluation

Introduction

The objective of the NIH-funded Clinical Genome Resource (ClinGen) is to build a central resource that defines the clinical relevance of genes and variants within specific disease entities for use in precision medicine and research. In order to facilitate this effort, the ClinGen Gene Curation Working Group developed a framework for curators to semi-quantitatively define the clinical validity of a gene-disease association (Strande & Riggs et al., 2017). This process involves the curation and systematic evaluation of peer-reviewed

publicly available evidence to assign the strength of a gene-disease association into one of the following clinical validity classifications: Definitive, Strong, Moderate, Limited, No Reported Evidence, or Conflicting Evidence Reported.

As a resource for use by clinicians and clinical laboratories, it is important for the gene curation activities of ClinGen to reflect up-to-date information for gene-disease associations. Thus, the need for reassessment of clinical validity classification will be crucial, as additional information and evidence is published in the literature. This study examines how the clinical validity classification of a number of genes has changed over time and the factors that influenced those changes. By identifying those variables and trends in the clinical validity classifications over time, we can provide guidance for re-evaluating curated gene-disease associations.

Methods

A retrospective analysis of clinical validity classifications was performed on 22 gene-disease associations that were previously assigned a clinical validity classification using the ClinGen framework (Strande & Riggs et al., 2017) (Table 1). In order to identify information about the gene-disease association as it accumulated, PubMed searches were performed, limited to one-year intervals beginning with the first publication to assert disease-causing variants in a gene. For the 22 gene-disease associations that were chosen, the year of first assertion ranged from 1991 to 2016. The information gathered included genetic (case reports, segregation data, case-control studies), experimental, and contradictory evidence. Using the ClinGen gene curation framework Standard Operating Procedures (SOP), Version 5 (clinicalgenome.org), curators simulated clinical validity assessments at one year-increments. A “Limited” classification scored up to 6 total points, a “Moderate” classification scored 7–11 total points, a “Strong” classification scored 12–18 total points, and a “Definitive” classification scored 12–18 points with replication over time (>3 years). A “Disputed/Refuted” classification represents those gene-disease associations in which conflicting evidence for the role of the gene in disease was reported.

Population databases- the Exome Aggregation Consortium (ExAC; exac.broadinstitute.org) and the Genome Aggregation Consortium (gnomAD; gnomad.broadinstitute.org) (Lek et al, 2016)- were utilized to assess variant pathogenicity, but were only taken into account for scoring metrics based on the point in time of their public release (October, 2014 for ExAC and October, 2016 for gnomAD) in order to observe how these databases influence the interpretation of a variant’s pathogenicity. The allele frequency at which a variant was considered benign was determined by the disease domain experts for each gene-disease curation.

To assess how the clinical validity classifications would evolve in more recently asserted gene-disease associations, eight novel gene-disease associations reported in 2013 in the *American Journal of Human Genetics*, volume 93 (issues 1 and 2) were selected for analysis (Table 2). Curations were simulated at one-year increments from 2013 to 2016.

Statistical analysis was performed using an unpaired t-test (GraphPad). All current gene-disease validity summaries can be found on the ClinGen website (www.clinicalgenome.org) and hyperlinks of the classification summaries are provided in Tables 1 and 2.

Results

In order to observe any trends in clinical validity classifications over time, a diverse selection of 22 curated gene-disease associations (Table 1) were evaluated at one-year intervals using the ClinGen gene curation framework, starting with the year of the first publication to assert disease-causing variants in the gene of interest (Fig. 1). Of the 22 gene-disease associations, six are currently classified as “Limited”, three are “Moderate”, three are “Strong”, seven are “Definitive” and three are “Refuted”. The gene-disease associations were chosen to encompass a range of clinical domains, inheritance patterns, pathologies, molecular mechanisms, and time at first assertion (1991–2016) and were selected from available curations that had been performed by ClinGen Gene Curation Expert Panels

On average, over the entire lifetime of a gene-disease association (from first assertion to the latest assessment for this study, 2017), gene-disease pairs tended to spend a higher percentage of time at the “Limited” classification ($71.4 \pm 34.9\%$) compared to “Moderate” ($41.5 \pm 34.6\%$) (Fig. 2A). Given that the point range for the “Moderate” classification is 4 points (7–11), compared to “Limited” and “Strong/Definitive” which are 5.9 (0.1–6) and 6 points (12–18) respectively, it is logical that gene-disease associations would spend the least amount of time in the “Moderate” classification. Consistent with this observation, we found that 71% of gene-disease associations that spent more than a year at the “Moderate” classification progressed to “Strong” (1/7) or “Definitive” (4/7) by 2017. Of note, *JPH2*: hypertrophic cardiomyopathy was excluded from this analysis as it had only reached the “Moderate” classification within the last year of the curation time frame, and thus was not subject to the analysis. Only 2/7 gene-disease associations remain at “Moderate” (Fig. 2B), suggesting that these associations may spend a fractional amount of time at “Moderate” before advancing to a new classification. Furthermore, it should be noted that we did not encounter a gene-disease association that achieved a “Moderate” classification that then moved to a “Disputed” or “Refuted” classification, further supporting the suggestion that once a gene-disease association meets a classification of “Moderate” it will most likely advance to “Strong” or “Definitive.”

In assessing gene-disease associations with a “Limited” classification, we found that out of the ten genes that spent at least three or more years at “Limited,” 50% of them currently remain at “Limited.” Only 2/10 progressed to a higher classification: one each to “Moderate” and “Definitive,” while 3/10 became “Disputed/Refuted” (Fig. 2C). This observation is consistent with the finding that gene-disease associations spend more time at the “Limited” classification, compared to “Moderate” and “Strong” (Fig. 2A). This trend was even more significant when we looked at gene-disease associations that spent at least five or more years at “Limited”: only 1/8 progressed to a higher classification, while 4/8 stayed at “Limited” and 3/8 became “Disputed/Refuted” (Fig. 2D). It is interesting to note that gene-disease associations that began as “Limited” and were re-classified as “Disputed/Refuted” only reached a maximum of two points in genetic and experimental evidence

before conflicting evidence was reported (*TMPO*: dilated cardiomyopathy, *MYO1A*: sensorineural hearing loss, and *PMS1*: Lynch syndrome). In contrast, “Limited” gene-disease associations that progressed to a higher classification (*LBR*: anadysplasia-like, spondylometaphyseal dysplasia, *PDZD7*: sensorineural hearing loss, *JPH2*: hypertrophic cardiomyopathy, *RPS24*: Diamond-Blackfan anemia, and *COX15*: Leigh syndrome) initially scored more than two points at the time of the first assertion. These findings suggest that gene-disease associations at the low end of the “Limited” point range (<2 points) are more likely to remain “Limited” or become “Disputed/Refuted” than gene-disease associations that initially scored in mid to upper “Limited” point range (2–6 points).

Disputed/Refuted gene-disease associations

Three out of 22 genes analyzed spent between 9 and 20 years at “Limited” before being reclassified as “Disputed” or “Refuted” between 2013 and 2014 (Fig. 1). For each of these gene-disease associations, the disputing or refuting factor that contributed most to the reclassification was the high maximum allele frequencies recorded for the asserted disease-causing variant(s) in the healthy control populations, as assessed by the use of ExAC and gnomAD. For the “Disputed” or “Refuted” categories, it is worth noting that the classifications are based on a specific claim of a monogenic gene-disease association. Thus, any dispute/refute is specific to that association and does not exclude other claims as being valid, nor does it suggest that the gene is not involved in any other disease by means of a different mechanism and/or mode of inheritance.

In 2005, *TMPO*: dilated cardiomyopathy was classified as “Limited” based on one case report of a proband with a missense variant (Taylor et al., 2005). However, with the public release of the ExAC database in 2014, this variant was reported at an overall frequency of 0.01508 in the population and was found in the homozygous state in 141 individuals, suggesting the variant is benign. Therefore, *TMPO*: dilated cardiomyopathy was classified as “Refuted” (Strande & Riggs et al., 2017).

PMS1: Lynch syndrome was originally given a classification of “Limited” in 1994 based on the report of a proband harboring a variant that resulted in exon skipping (Nicolaidis et al., 1994). In 2001, when the patient’s genome was further analyzed, it was demonstrated that she also carried a pathogenic mutation in *MSH2* (which is also associated with Lynch syndrome) that segregated with disease in the family, indicating that the *MSH2* variant was the probable causative variant (Liu et al.). There was only one additional report in 1999 of a proband carrying two missense variants in *PMS1* (Wang et al., 1999), which were later found at an overall frequency of 0.04513 and 0.04594 in ExAC, each with 11 homozygotes, suggesting that they are both benign. This led to a “Refuted” classification for *PMS1*: Lynch syndrome (ClinGen Colon Cancer and Polyposis Gene Curation Expert Panel, clinicalgenome.org).

Finally, *MYO1A*: Sensorineural hearing loss was classified as “Limited” in 2003 (Donaudy, et al., 2003) based on the identification of six missense variants, one in-frame deletion, and one nonsense variant in *MYO1A* in probands with hearing loss. However, four of the missense variants were later found at an allele frequency greater than 0.5% in gnomAD, and the nonsense variant was found to be inherited from an unaffected mother. A 2014 report

identified alternate genetic causes of hearing loss in two probands with novel variants in *MYO1A* and one proband that was originally reported by Donaudy et al. (2003) (Eisenberger et al., 2014). Finally, a 2017 report identified 12 individuals without hearing loss that harbored either a loss-of-function variant in *MYO1A* or one of the variants that had been reported by Donaudy et al. (2003) (Patton et al., 2017). Each of these factors led to a final classification of “Refuted” (ClinGen Hereditary Hearing Loss Gene Curation Expert Panel, clinicalgenome.org).

Impact of new sequencing technologies on clinical validity classifications

In recent years, advancements in sequencing technologies, such as next-generation sequencing (NGS) and whole exome sequencing (WES), have allowed for rapid sequencing of DNA from large numbers of individuals at a fraction of the cost, which has led to increased use of genetic testing in healthcare and research to facilitate genetic variant identification. Additionally, these technologies have led to the development of large genetic population data sets, such as ExAC, which was publicly released in October, 2014 (Lek et al., 2016). In order to determine how these recent advancements in genomics affect the discovery and clinical validity of novel gene-disease associations, we simulated clinical validity assessments of eight novel gene-disease associations reported in 2013 in the *American Journal of Human Genetics* at one-year intervals from 2013 through 2016 (Table 2).

Interestingly, besides advancing from “Strong” to “Definitive,” only one gene-disease association changed classification during the 2013–2016 time period—*KLHL40*: nemaline myopathy (Fig.3). Although the total points for the initial gene-disease association in 2013 were within the range for a “Strong” classification, it was constrained to a “Moderate” level because there was only one study with evidence to support the role of *KLHL40* in nemaline myopathy (Ravenscroft et al., 2013). According to the ClinGen Clinical Validity Classification system, a gene-disease association is classified as “Strong” if the role of the gene in disease has been demonstrated independently in at least two separate studies. As additional studies were reported in 2014, it was upgraded to “Strong.” Conversely, *PIK3R1*: SHORT syndrome reached the “Strong” classification in the first year (2013) because three independent studies implicating the gene’s role in disease were published in the same issue (Thauvin-Robinet et al, 2013; Chudasama et al, 2013; Dymment et al, 2013). Both *KLHL40*: nemaline myopathy and *PIK3R1*: SHORT syndrome reached a final classification of “Definitive” in 2016 due to replication over time.

In order to further investigate how the next generation of sequencing technologies may have impacted the clinical validity classification over time, we focused our analysis to the ten “Definitive” gene-disease associations (Tables 1 and 2, Figures 1 and 3). The ten gene-disease pairs were divided into two groups: those that were discovered using traditional methods (linkage analysis, direct sequencing) prior to 2010 (*GAA*: Pompe disease, *TPM1*: hypertrophic cardiomyopathy, *SMAD4*: juvenile polyposis syndrome, *USH1C*: Usher syndrome type 1C, *RPS24*: Diamond-Blackfan anemia, and *PDZD7*: sensorineural hearing loss) and those that were discovered using modern genomic technology (WES) after 2010

(*ABCC9*: Cantu syndrome, *KLHL40*: nemaline myopathy, and *PIK3R1*: SHORT syndrome).

There was no significant difference in the average amount of time it took for the gene-disease associations to reach “Definitive” between the pre-2010, traditional methods group (4.33 ± 1.37 years) vs the post-2010, WES group (3.0 ± 0 years, $p = 0.1465$) (Fig. 4A). However, there was a significant increase in the average number of probands curated per paper for the post-2010, WES group (1.98 ± 0.55 vs 8.5 ± 3.29 , $p = 0.0015$) (Fig. 4B), suggesting that next generation sequencing technologies may allow researchers to analyze a greater number of cases, while those gene-disease associations that were discovered using traditional methods relied on more papers to reach “Definitive” status. In fact, this trend was observed when we looked at the number of papers curated for both groups (Fig. 4C), though it did not reach significance, which is likely due to a small sample size. For the pre-2010 group, an average of 9.7 ± 3.7 papers were curated to reach “Definitive” status vs 5.0 ± 1.0 papers curated for the post-2010 group ($p = 0.0742$).

Discussion

As the number of completed gene curations continues to grow, ClinGen gene curation groups will need to determine how often gene-disease associations will be re-evaluated in order to stay up-to-date with new information as it becomes available in the literature. Analysis of the 30 gene-disease associations discussed here can begin to provide insight into how the clinical validity classification changes over time. This analysis is not only relevant to ClinGen curation activities, but can also be applied by any groups offering diagnostic testing to determine how often genes on panels are assessed. A systematic approach for re-analysis of genes on diagnostic testing panels will help to ensure the most accurate results are reported.

The first thing to consider when determining how often to re-evaluate a gene-disease association is the initial clinical validity classification. In effect, the “Limited” classification was intended to represent a relatively weak single publication with few cases and little supporting functional evidence, while the “Moderate” classification was intended to represent a stronger initial publication providing more cases and ample functional evidence supporting the gene’s role in the disease. A “Strong” classification implies additional independent corroboration of an initial gene-disease association, increasing both the case-level and gene-level evidence.

There are several reasons why gene-disease associations may stay at “Limited” for a longer period of time. First, the disease may be rare, in which case there are very few probands reported in the literature. It is important to note, however, that with the ClinGen framework, it is possible for even rare gene-disease associations to reach a “Strong” or “Definitive” classification based on the strength of available variant evidence and supporting experimental evidence. Secondly, the gene of interest may be only one of many genes known to cause the same disease and therefore variants in that gene only account for a small number of cases. For example, *TCAP*, *AKAP9*, *NHP2*, and *KLF10* account for fewer than 1% of cases of their respective diseases, which are caused by variants in more than ten genes

each (see notes on Table 1). Finally, a gene-disease association may spend a greater time at “Limited” due to disparities in research and funding for different clinical domains. For these reasons, a gene-disease association may remain at “Limited” for several years, but with the availability of technologies such as WES, and resources such as GeneMatcher (www.genematcher.org) and GenomeConnect (www.genomeconnect.org) that allow clinicians and researchers who are interested in the same gene to connect, it may eventually progress to a new classification with time.

When determining how often to re-evaluate a “Limited” gene-disease association, it is appropriate to consider the length of time spent at “Limited” and the total number of points. Gene-disease associations with greater than five years at “Limited” and points within the lower range of the “Limited” classification (e.g. <2 points) should be re-evaluated less frequently than newly discovered “Limited” gene-disease associations that have points in the upper range of the “Limited” classification. For example, out of the 30 gene-disease associations examined here, there are several “Limited” gene-disease associations at the higher end of the classification point range, and thus are likely to change classification with one or two additional studies, including: *KLF10*: hypertrophic cardiomyopathy, *NHP2*: dyskeratosis congenita, *RAB28*: cone-rod dystrophy, and *ARL2BP*: retinitis pigmentosa (Tables 1 and 2). Conversely, “Limited” gene-disease associations, such as *AKAP9*: long QT syndrome and *TCAP*: hypertrophic cardiomyopathy, that have fewer than two points and have remained “Limited” for more than ten years (Table 1, Fig. 1) should be re-evaluated less frequently, and importantly they should be monitored for newly published data that could dispute or refute the gene-disease association.

Not surprisingly, gene-disease associations tend to spend less time in the “Moderate” category compared to “Limited.” This is likely due to the small point range for this category compared to “Limited” and “Strong/Definitive” and the fact that any corroborating publications would be likely to push the supporting evidence into the “Strong” category. Similar to gene-disease associations at the high end of the point range for “Limited,” curation groups may want to re-evaluate “Moderate” associations more frequently, and automated systems for alerting curation groups to new publications about the gene or disease would help to keep categorizations up to date. Gene-disease associations also spend very little time at “Strong” due to the fact that they can reach “Definitive” once the association has been upheld for at least three years (Strande & Riggs et al., 2017). “Strong” gene-disease associations would therefore only need to be re-evaluated once they reach the three year mark for replication over time.

Given the availability of more advanced genomic sequencing technologies, large population databases, and resources like GeneMatcher and GenomeConnect, it was surprising that most of the novel gene-disease associations from 2013 did not change classification between 2013 and 2016 (Fig. 3). It was observed that newer gene-disease associations tend to have more probands per paper (Fig. 4B), perhaps due to the use of relatively inexpensive genomic sequencing technologies, match-making and global communication that enables groups to gather cohorts of patients more easily, or the journal’s selectivity for the number of probands in a publication. Therefore, the absence of classification changes between 2013–2016 could be due to the rarity of the disease, or that variants in the gene only account for a small

fraction of cases. However, it is important to note that all four gene-disease associations that fell within the “Moderate” point range in 2013 (*LRPAP1*: myopaia, *AQP5*: palmoplantar keratoderma, *ELAC2*: combined oxidative phosphorylation deficiency, and *CTCF*: intellectual disability) were within two points of the “Strong/Definitive” classification by 2016. Similarly, both gene-disease associations that fell within the “Limited” point range in 2013 (*RAB28*: cone-rod dystrophy and *ARL2BP*: retinitis pigmentosa) were within one point of the “Moderate” classification by 2016, suggesting that they might change classification with an additional publication or two.

In summary, the results from this retrospective study provide several key points to consider when determining how often to re-evaluate a gene-disease association. First, “Limited” classifications should be considered carefully. Gene-disease associations that have been at the low end of the “Limited” point range (e.g. <2 points) for an extended period of time (e.g. >5 years) and have not been disputed/refuted with the release of ExAC or gnomAD, may not need to be evaluated as frequently as those at the higher end of the “Limited” point range and those that have been discovered more recently. Low “Limited” gene-disease associations may only need to be evaluated every three to five years, while high “Limited” gene-disease associations every one to two years. Eventually, consensus groups may need to reassess the original evidence and determine whether failure to replicate the initial finding constitutes a justification for disputing the gene-disease association. “Moderate” gene-disease associations should be re-evaluated more frequently, possibly every one to two years, given the small point range for the category. Finally, “Strong” gene-disease associations should be re-evaluated at the three-year mark to reflect replication over time, allowing the association to reach the “Definitive” classification. These recommendations are based on our initial review. However, re-analysis and adjustment of these time points may be required in the future once more gene-disease associations have been analyzed.

Overall, this study has provided insight into the process of biocuration and has given key consideration points in determining the most effective re-evaluation process for gene-disease association classifications.

Acknowledgements

We wish to thank NHGRI for financial support (U41HG009650) and Erin Riggs, Ahmad Abou Tayoun, Heather Baudet, Christian Schaff, and the following ClinGen groups for their contributions: Gene Curation Working Group, Hypertrophic Cardiomyopathy Gene Curation Expert Panel, Colon Cancer and Polyposis Gene Curation Expert Panel, and Hereditary Hearing Loss Gene Curation Expert Panel.

Web Resources

ClinGen, <https://www.clinicalgenome.org/>
ClinGen Gene Curation SOP, <https://www.clinicalgenome.org/curation-activities/gene-disease-validity/educational-and-training-materials/standard-operating-procedures/>
OMIM, <https://omim.org/>
Orphanet, <https://www.orpha.net/consor/cgi-bin/index.php>

References

- Aldahmesh MA, Khan AO, Alkuraya H, Adly N, Anazi S, Al-Saleh AA, Alkuraya FS (2013). Mutations in LRPAP1 are associated with severe myopia in humans. *Am J Hum Genet*, 93(2), 313–320. doi: 10.1016/j.ajhg.2013.06.002 [PubMed: 23830514]
- Alders M, Bikker H, & Christiaans I (2003; updated 2018). Long QT Syndrome In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Stephens K, & Amemiya A (Eds.), *GeneReviews®*. Seattle (WA): University of Washington, Seattle; 1993–2018. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK1129/>
- Antonicka H, Mattman A, Carlson CG, Glerum DM, Hoffbuhr KC, Learly SC, Kennaway NG, & Shoubbridge EA (2003). Mutations in COX15 produce a defect in the mitochondrial heme biosynthetic pathway, causing early-onset fatal hypertrophic cardiomyopathy. *Am J Hum Genet*, 72(1), 101–114. doi: 10.1086/345489 [PubMed: 12474143]
- Blaydon DC, Lind LK, Plagnol V, Linton KJ, Smith FJ, Wilson NJ, Kelsell DP (2013). Mutations in AQP5, encoding a water-channel protein, cause autosomal-dominant diffuse nonepidermolytic palmoplantar keratoderma. *Am J Hum Genet*, 93(2), 330–335. doi: 10.1016/j.ajhg.2013.06.008 [PubMed: 23830519]
- Borovik L, Modaff P, Waterham HR, Krentz AD, & Pauli RM (2013). Pelger-huet anomaly and a mild skeletal phenotype secondary to mutations in LBR. *Am J Med Genet A*, 161A(8), 2066–2073. doi: 10.1002/ajmg.a.36019 [PubMed: 23824842]
- Bos JM, Subramaniam M, Hawse JR, Christiaans I, Rajamannan NM, Maleszewski JJ, Ackerman MJ (2012). TGFβ-inducible early gene-1 (TIEG1) mutations in hypertrophic cardiomyopathy. *J Cell Biochem*, 113(6), 1896–1903. doi: 10.1002/jcb.24058 [PubMed: 22234868]
- Chen L, Marquardt ML, Tester DJ, Sampson KJ, Ackerman MJ, & Kass RS (2007). Mutation of an A-kinase-anchoring protein causes long-QT syndrome. *Proc Natl Acad Sci U S A*, 104(52), 20990–20995. doi: 10.1073/pnas.0710527105 [PubMed: 18093912]
- Chudasama KK, Winnay J, Johansson S, Claudi T, König R, Haldorsen I, Njølstad PR (2013). SHORT syndrome with partial lipodystrophy due to impaired phosphatidylinositol 3 kinase signaling. *Am J Hum Genet*, 93(1), 150–157. doi: 10.1016/j.ajhg.2013.05.023 [PubMed: 23810379]
- Cirino AL & Ho C (2008; updated 2014). Hypertrophic Cardiomyopathy Overview In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Stephens K, & Amemiya A (Eds.), *GeneReviews®*. Seattle (WA): University of Washington, Seattle; 1993–2018. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK1768/>
- Clinton C & Gazda HT (2009; updated 2016). Diamond-Blackfan Anemia In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Stephens K, & Amemiya A (Eds.), *GeneReviews®*. Seattle (WA): University of Washington, Seattle; 1993–2018. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK7047/>
- Davidson AE, Schwartz N, Zelinger L, Stern-Schneider G, Shoemark A, Spitzbarth B, Webster AR (2013). Mutations in ARL2BP, encoding ADP-ribosylation-factor-like 2 binding protein, cause autosomal-recessive retinitis pigmentosa. *Am J Hum Genet*, 93(2), 321–329. doi: 10.1016/j.ajhg.2013.06.003 [PubMed: 23849777]
- Donaudy F, Ferrara A, Esposito L, Hertzano R, Ben-David O, Bell RE, Gasparini P (2003). Multiple mutations of MYO1A, a cochlear-expressed gene, in sensorineural hearing loss. *Am J Hum Genet*, 72(6), 1571–1577. doi: 10.1086/375654 [PubMed: 12736868]
- Dymant DA, Smith AC, Alcantara D, Schwartzentruber JA, Basel-Vanagaite L, Curry CJ, Innes AM (2013). Mutations in PIK3R1 cause SHORT syndrome. *Am J Hum Genet*, 93(1), 158–166. doi: 10.1016/j.ajhg.2013.06.005 [PubMed: 23810382]
- Eisenberger T, Di Donato N, Baig SM, Neuhaus C, Beyer A, Decker E, Bolz HJ (2014) Targeted and genome-wide NGS data disqualify mutations in MYO1A, the “DFNA48 gene”, as a cause of deafness. *Hum Mutat*, 35(5), 565–570. doi: 10.1002/humu.22532 [PubMed: 24616153]
- Esteban-Jurado C, Vila-Casadesús M, Garre P, Lozano JJ, Pristoupilova A, Beltran S, Castellví-Bel S (2015). Whole-exome sequencing identifies rare pathogenic variants in new predisposition genes for familial colorectal cancer. *Genet Med*, 17(2), 131–142. doi: 10.1038/gim.2014.89 [PubMed: 25058500]

- Fahim AT, Daiger SP, & Weleber RG (2000; updated 2017). Nonsyndromic Retinitis Pigmentosa Overview In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Stephens K, & Amemiya A (Eds.), GeneReviews®. Seattle (WA): University of Washington, Seattle; 1993–2018. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK1417/>
- Gazda HT, Grabowska A, Merida-Long LB, Latawiec E, Schneider HE, Lipton JM, Sieff CA (2006). Ribosomal protein S24 gene is mutated in Diamond-Blackfan anemia. *Am J Hum Genet*, 79(6), 1110–1118. doi: 10.1086/510020 [PubMed: 17186470]
- Grange DK, Nichols CG, & Singh GK (2014). Cantú Syndrome and Related Disorders In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Stephens K, & Amemiya A (Eds.), GeneReviews®. Seattle (WA): University of Washington, Seattle; 1993–2018. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK246980/>
- Gregor A, Oti M, Kouwenhoven EN, Hoyer J, Sticht H, Ekici AB, Zweier C (2013). De novo mutations in the genome organizer CTCF cause intellectual disability. *Am J Hum Genet*, 93(1), 124–131. doi: 10.1016/j.ajhg.2013.05.007 [PubMed: 23746550]
- Haack TB, Kopajtich R, Freisinger P, Wieland T, Rorbach J, Nicholls TJ, Prokisch H (2013). ELAC2 mutations cause a mitochondrial RNA processing defect associated with hypertrophic cardiomyopathy. *Am J Hum Genet*, 93(2), 211–223. doi: 10.1016/j.ajhg.2013.06.006 [PubMed: 23849775]
- Hayashi T, Arimura T, Itoh-Satoh M, Ueda K, Hohda S, Inagaki N, ... Kimura A (2004). Tcap gene mutations in hypertrophic cardiomyopathy and dilated cardiomyopathy. *J Am Coll Cardiol*, 44(11), 2192–2201. doi: 10.1016/j.jacc.2004.08.058 [PubMed: 15582318]
- Hershberger RE & Morales A (2007; updated 2015). Dilated Cardiomyopathy Overview In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Stephens K, & Amemiya A (Eds.), GeneReviews®. Seattle (WA): University of Washington, Seattle; 1993–2018. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK1309/>
- Howe JR, Roth S, Ringold JC, Summers RW, Järvinen HJ, Sistonen P, Aaltonen LA (1998). Mutations in the SMAD4/DPC4 gene in juvenile polyposis. *Science*, 280(5366), 1086–1088. [PubMed: 9582123]
- Innes AM & Dymant DA (2014). SHORT Syndrome In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Stephens K, & Amemiya A (Eds.), GeneReviews®. Seattle (WA): University of Washington, Seattle; 1993–2018. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK201365/>
- Jiang D, Li J, Xiao X, Li S, Jia X, Sun W, Guo X, & Zhang Q (2014). Detection of mutations in LRPAP1, CTSB, LEPREL1, ZNF644, SLC39A5, and SCO2 in 298 families with early-onset high myopia by exome sequencing. *Invest Ophthalmol Vis Sci*, 56(1), 339–345. doi: 10.1167/iovs.14-14850 [PubMed: 25525168]
- Kohl S, Jägle H, & Wissinger B (2004; updated 2016). Achromatopsia In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Stephens K, & Amemiya A (Eds.), GeneReviews®. Seattle (WA): University of Washington, Seattle; 1993–2018. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK1418/>
- Kohl S, Zobor D, Chiang WC, Weisschuh N, Staller J, Gonzalez Menendez I, Lin JH (2015). Mutations in the unfolded protein response regulator ATF6 cause the cone dysfunction disorder achromatopsia. *Nat Genet*, 47(7), 757–765. doi: 10.1038/ng.3319 [PubMed: 26029869]
- Kohlman W & Gruber SB (2004; updated 2018). Lynch Syndrome In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Stephens K, & Amemiya A (Eds.), GeneReviews®. Seattle (WA): University of Washington, Seattle; 1993–2018. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK1211/>
- Landstrom AP, Weisleder N, Batalden KB, Bos JM, Tester DJ, Ommen SR, Ackerman MJ (2007). Mutations in JPH2-encoded junctophilin-2 associated with hypertrophic cardiomyopathy in humans. *J Mol Cell Cardiol*, 42(6), 1026–1035. doi: 10.1016/j.yjmcc.2007.04.006 [PubMed: 17509612]
- Larson Haidle J & Howe JR (2003; updated 2017). Juvenile Polyposis Syndrome In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Stephens K, & Amemiya A (Eds.), GeneReviews®. Seattle (WA): University of Washington, Seattle; 1993–2018. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK1469/>

- Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, MacArthur DG, & Exome Aggregation Consortium (2016). Analysis of protein-coding genetic variation in 60,706 humans. *Nature*, 536(7616), 285–291. doi: 10.1038/nature19057 [PubMed: 27535533]
- Lentz J & Keats BJB (1999; updated 2016). Usher Syndrome Type I In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Stephens K, & Amemiya A (Eds.), *GeneReviews®*. Seattle (WA): University of Washington, Seattle; 1993–2018. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK1265/>
- Leslie N & Bailey L (2007; updated 2017). Pompe Disease In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Stephens K, & Amemiya A (Eds.), *GeneReviews®*. Seattle (WA): University of Washington, Seattle; 1993–2018. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK1261/>
- Lin Z, Li S, Feng C, Yang S, Wang H, Ma D, Tan X (2016). Stabilizing mutations of KLHL24 ubiquitin ligase cause loss of keratin 14 and human skin fragility. *Nat Genet*, 48(12), 1508–1516. doi: 10.1038/ng.3701 [PubMed: 27798626]
- Liu T, Yan H, Kuismanen S, Percesepe A, Bisgaard ML, Pedroni M, Lindblom A (2001). The role of hPMS1 and hPMS2 in predisposing to colorectal cancer. *Cancer Res*, 61(21), 7798–7802. [PubMed: 11691795]
- Lynch HT & de la Chapelle A (2003). Hereditary colorectal cancer. *N Engl J Med*, 348(10), 919–923. doi: 10.1056/NEJMra012242 [PubMed: 12621137]
- Nicolaides NC, Papadopoulos N, Liu B, Wei YF, Carter KC, Ruben SM, Kinzler KW (1994). Mutations of two PMS homologues in hereditary nonpolyposis colon cancer. *Nature*, 371(6492), 75–80. doi: 10.1038/371075a0 [PubMed: 8072530]
- North KN & Ryan MM (2002; updated 2015). Nemaline Myopathy In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Stephens K, & Amemiya A (Eds.), *GeneReviews®*. Seattle (WA): University of Washington, Seattle; 1993–2018. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK1288/>
- Patton J, Brewer C, Chien W, Johnston JJ, Griffith AJ, & Biesecker LG (2016). A genotypic ascertainment approach to refute the association of MYO1A variants with non-syndromic deafness. *Eur J Hum Genet*, 25(1), 147–149. doi: 10.1038/ejhg.2016.140 [PubMed: 27759032]
- Pfender EG, & Bruckner AL (1998; updated 2016). Epidermolysis Bullosa Simplex In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Stephens K, & Amemiya A (Eds.), *GeneReviews®*. Seattle (WA): University of Washington, Seattle; 1993–2018. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK1369/>
- Pigors M, Sarig O, Heinz L Plagnol V, Fischer J, Mohamad J, Blaydon DC (2016). Loss-of-Function Mutations in SERPINB8 Linked to Exfoliative Ichthyosis with Impaired Mechanical Stability of Intercellular Adhesions. *Am J Hum Genet*, 99(2), 430–436. doi: 10.1016/j.ajhg.2016.06.004 [PubMed: 27476651]
- Ravenscroft G, Miyatake S, Lehtokari VL, Todd EJ, Vornanen P, Yau KS, Laing NG (2013). Mutations in KLHL40 are a frequent cause of severe autosomal-recessive nemaline myopathy. *Am J Hum Genet*, 93(1), 6–18. doi: 10.1016/j.ajhg.2013.05.004 [PubMed: 23746549]
- Roosing S, Rohrschneider K, Beryozkin A, Sharon D, Weisschuh N, Staller J, den Hollander AI (2013). Mutations in RAB28, encoding a farnesylated small GTPase, are associated with autosomal-recessive cone-rod dystrophy. *Am J Hum Genet*, 93(1), 110–117. doi: 10.1016/j.ajhg.2013.05.005 [PubMed: 23746546]
- Savage SA (2009; updated 2016). Dyskeratosis Congenita In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Stephens K, & Amemiya A (Eds.), *GeneReviews®*. Seattle (WA): University of Washington, Seattle; 1993–2018. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK22301/>
- Schneider E, Märker T, Daser A, Frey-Mahn G, Beyer V, Farkas R, Haaf T (2009). Homozygous disruption of PDZD7 by reciprocal translocation in a consanguineous family: a new member of the Usher syndrome protein interactome causing congenital hearing impairment. *Hum Mol Genet*, 18(4), 655–666. doi: 10.1093/hmg/ddn395 [PubMed: 19028668]
- Strande NT, Riggs ER, Buchanan AH, Ceyhan-Birsoy O, DiStefano M, Dwight SS, Berg JS (2017). Evaluating the Clinical Validity of Gene-Disease Associations: An Evidence-Based Framework

- Developed by the Clinical Genome Resource. *Am J Hum Genet*, 100(6), 895–906. doi: 10.1016/j.ajhg.2017.04.015 [PubMed: 28552198]
- Taylor MR, Slavov D, Gajewski A, Vlcek S, Ku L, Fain PR, Mestroni L; Familial Cardiomyopathy Registry Research Group (2005). Thymopoietin (lamin-associated polypeptide 2) gene mutation associated with dilated cardiomyopathy. *Hum Mutat*, 26(6), 566–574. doi: 10.1002/humu.20250 [PubMed: 16247757]
- Thauvin-Robinet C, Auclair M, Caron-Debarle M, Avila M, St-Onge J, Le Merrer M, Rivière JB (2013). PIK3R1 mutations cause syndromic insulin resistance with lipodystrophy. *Am J Hum Genet*, 93(1), 141–149. doi: 10.1016/j.ajhg.2013.05.019 [PubMed: 23810378]
- Thierfelder L, Watkins H, MacRae C, Lamas R, McKenna W, Vosberg HP, Seidman JG, & Seidman CE (1994). Alpha-tropomyosin and cardiac troponin T mutations cause familial hypertrophic cardiomyopathy: a disease of the sarcomere. *Cell*, 77(5), 701–712. [PubMed: 8205619]
- van Bon BW, Gilissen C, Grange DK, Hennekam RC, Kayserili H, Engels H, Hoischen A (2012). Cantú syndrome is caused by mutations in ABCC9. *Am J Hum Genet*, 90(6), 1094–1101. doi: 10.1016/j.ajhg.2012.04.014 [PubMed: 22608503]
- Vervy E, Leibovici M, Zwaenepoel I, Liu XZ, Gal A, Salem N, Petit C (2000). A defect in harmonin, a PDZ domain-containing protein expressed in the inner ear sensory hair cells, underlies Usher syndrome type 1C. *Nat Genet*, 26(1), 51–55. doi: 10.1038/79171 [PubMed: 10973247]
- Vulliamy T, Beswick R, Kirwan M, Marrone A, Digweed M, Walne A, & Dokal I (2008). Mutations in the telomerase component NHP2 cause the premature ageing syndrome dyskeratosis congenita. *Proc Natl Acad Sci U S A*, 105(23), 8073–8078. doi: 10.1073/pnas.0800042105 [PubMed: 18523010]
- Wang Q, Lasset C, Desseigne F, Saurin JC, Maugard C, Navarro C, ... Puisieux A (1999). Prevalence of germline mutations of hMLH1, hMSH2, hPMS1, hPMS2, and hMSH6 genes in 75 French kindreds with nonpolyposis colorectal cancer. *Hum Genet*, 105(1–2), 79–85. [PubMed: 10480359]
- Zhong F, Savage SA, Shkreli M, Giri N, Jessop L, Myers T, Chen R, Alter BP, & Artandi SE (2011). Disruption of telomerase trafficking by TCAB1 mutation causes dyskeratosis congenita. *Genes Dev*, 25(1), 11–16. doi: 10.1101/gad.2006411 [PubMed: 21205863]
- Zhong N, Martiniuk F, Tzall S, Hirschhorn R (1991). Identification of a missense mutation in one allele of a patient with Pompe disease, and use of endonuclease digestion of PCR-amplified RNA to demonstrate lack of mRNA expression from the second allele. *Am J Hum Genet*, 49(3), 635–645. [PubMed: 1652892]

Figure 1. Clinical validity classifications over time for 22 gene-disease associations. One exception to the point and classification assignments was in the gene-disease association *COX15*: Leigh Syndrome, which due to replication over time could be “Definitive” but remained at “Strong” after expert review (see “Methods” for classification point ranges).

Figure 2.

Trends in the clinical validity over time across the 22 gene-disease associations. A. Percentage of time spent at each clinical validity classification over the lifetime of a gene-disease association. For each gene-disease association, the year(s) spent at a classification were divided by the total number of years the gene-disease pair has been classified (Table 1). *KLHL24*: epidermal bullosa simplex and *SERPINB8*: exfoliative ichthyosis were excluded from this analysis as they have only been classified at one time point and have not had time to change classification. Mean percentages of time spent at each classification are as follows: $71.4 \pm 34.9\%$ at "Limited," $41.5 \pm 34.6\%$ at "Moderate," $33.4 \pm 34.9\%$ at "Strong," $64.0 \pm 26.2\%$ at "Definitive," and $19.8 \pm 6.2\%$ at "Disputed/Refuted." B. Current clinical validity classifications of genes that spent more than one year at "Moderate." *JPH2*: hypertrophic cardiomyopathy was excluded from this analysis as it has only been classified at one time point and has not had time to change classification. Two out of seven of those gene-disease associations remain at "Moderate," 1/7 progressed to "Strong," and 4/7 progressed to "Definitive." C. Current clinical validity classifications of genes that spent at

least three years at “Limited.” Five out of ten gene-disease associations remain at “Limited,” 1/10 progressed to “Moderate,” 1/10 progressed to “Definitive,” and 3/10 are now “Disputed/Refuted.” D. Current clinical validity classifications of genes that spend at least five years at “Limited.” Four out of eight gene-disease associations remain at “Limited,” 1/8 progressed to “Moderate,” and 3/8 are now “Disputed/Refuted.”

Figure 3.

Clinical validity classifications over time for eight novel gene-disease associations from 2013–2016. *Although the total points for *KLHL40*: nemaline myopathy reached the point range for a “Strong” classification in 2013, it was downgraded to “Moderate” as a “Strong” classification requires two or more independent studies to reach this classification per the ClinGen gene curation framework. In 2016, both *KLHL40*: nemaline myopathy and *PIK3RI*: SHORT syndrome were upgraded to “Definitive” due to replication over time (indicated by r/t).

Figure 4.

Advancements in genomic sequencing technologies correlate with the amount of genetic evidence curated per paper. A. Average amount of time to “Definitive” for gene-disease associations asserted before and after 2010. There is no significant difference in the amount of time it took to reach a “Definitive” classification for gene-disease associations asserted prior to 2010 (4.33 ± 1.37 years) using traditional sequencing methods (*GAA*: Pompe disease, *TPM1*: hypertrophic cardiomyopathy, *SMAD4*: juvenile polyposis syndrome, *USH1C*: Usher syndrome type 1C, *RPS24*: Diamond-Blackfan anemia, *PDZD7*: sensorineural hearing loss) compared to gene-disease associations asserted after 2010 (3.0 ± 0 years, $p = 0.1465$) using WES (*ABCC9*: Cantu syndrome, *KLHL40*: nemaline myopathy, *PIK3R1*: SHORT syndrome). B. Average number of probands per paper curated for “Definitive” gene-disease associations asserted before and after 2010. For “Definitive” gene disease associations asserted prior to 2010, there was an average of 1.98 ± 0.55 probands curated per paper. For “Definitive” gene-disease associations asserted after 2010, there was an average of 8.5 ± 3.29 probands curated per paper (**, $p = 0.0015$). C. Average number of papers curated for “Definitive” gene-disease associations asserted before and after 2010. For “definitive gene-disease associations asserted prior to 2010, there was an

average of 9.7 ± 3.7 papers curated vs 5.0 ± 1.0 papers curated for those asserted after 2010 ($p = 0.0742$).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1.

List of gene-disease pairs used to assess clinical validity classifications over time.

HGNC Gene Symbol	Disease (Inheritance)	Orphanet ID, OMIM Phenotype	Disease Prevalence [†]	Year of 1 st assertion in the literature (Total years curated)	Expert Reviewed Classification (Points)	Notes
<i>KLHL24</i>	Epidermolysis bullosa simplex (AD)	ORPHA N/A OMIM #617294	1:30,000 to 1:50,000 (Pfender & Bruckner, 1998)	2016 (1) (Lin et al.)	Strong^a (12 GE, 4 EE)	<i>EXPH5</i> , <i>KRT5</i> , <i>KRT14</i> , <i>TGCM5</i> account for ~81% of cases (Pfender & Bruckner, 1998).
<i>SERPINB8</i>	Exfoliative ichthyosis (AR)	ORPHA 289586 OMIM #617115	N/A	2016 (1) (Pigors et al.)	Limited^a (1.75 GE, 1 EE)	
<i>ATF6</i>	Achromatopsia (AR)	ORPHA 49382 OMIM# 616517	< 1:30,000 (Kohl et al, 2004)	2015 (2) (Kohl et al.)	Strong^b (12 GE, 2 EE)	6 genes are associated with the disease. Variants in <i>ATF6</i> have been found in at least 12 families in the literature (Kohl et al., 2004).
<i>XRCC4</i>	Hereditary colorectal cancer (AD)	ORPHA N/A OMIM N/A	5–10% of CRC cases are hereditary (Lynch & de la Chapelle, 2003)	2014 (3) (Esteban-Jurado et al.)	Limited^c (1.5 GE, 0 EE)	<i>XRCC4</i> is also involved in short stature, microcephaly, and endocrine dysfunction (OMIM #616541).
<i>LBR</i>	Regressive spondylometa-physal dysplasia (AR)	ORPHA 448267 OMIM N/A [‡]	< 1:1,000,000 (Orphanet)	2013 (4) (Borovik et al.)	Moderate^b (4 GE, 2.5 EE)	
<i>ABCC9</i>	Cantú syndrome (AD)	ORPHA 1517 OMIM #239850	N/A	2012 (5) (van Bon et al.)	Definitive^a (12 GE, 1 EE, r/t)	Variants in <i>ABCC9</i> account for ~28/35 cases in the literature (Grange et al., 2014).
<i>WRAP53</i>	Dyskeratosis congenita (AR)	ORPHA 1775 OMIM# 613988	1–9:1,000,000 (Orphanet)	2010 (7) (Zhong et al.)	Moderate^b (3.5 GE, 6 EE)	11 genes are associated with the disease. <i>WRAP53</i> variants account for <1% of cases (Savage, 2009).
<i>KLF10</i>	Hypertrophic cardiomyopathy (AD)	ORPHA 217569 OMIM N/A	1:500 (Cirino & Ho, 2008)	2010 (7) (Bos et al.)	Limited^d (1.25 GE, 4.5 EE)	>16 genes are associated with HCM. <i>KLF10</i> variants account for <1% of cases (Cirino & Ho, 2008).

HGNC Gene Symbol	Disease (Inheritance)	Orphanet ID, OMIM Phenotype	Disease Prevalence [†]	Year of 1 st assertion in the literature (Total years curated)	Expert Reviewed Classification (Points)	Notes
<i>PDZD7</i>	Sensorineural hearing loss (AR)	ORPHA 90636 OMIM #618003	1:1,000 to 1:700 (Orphanet)	2009 (8) (Schneider et al.)	Definitive^e (12 GE, 4 EE, r/t)	60–80% of cases are of genetic origin. ~86 causative genes for this disease have been identified (Orphanet).
<i>NHP2</i>	Dyskeratosis congenita (AR)	ORPHA 1775 OMIM #613987	1–9:1,000,000 (Orphanet)	2008 (9) (Vulliamy et al.)	Limited^b (2 GE, 4 EE)	11 genes are associated with the disease. <i>NHP2</i> variants account for <1% of cases (Savage, 2009).
<i>JPH2</i>	Hypertrophic cardiomyopathy (AD)	ORPHA 217569 OMIM #613873	1:500 (Cirino & Ho, 2008)	2007 (10) (Landstrom et al.)	Moderate^d (4.2 GE, 4.5 EE)	>16 genes are associated with HCM. <i>JPH2</i> variants account for <1% of cases (Cirino & Ho, 2008).
<i>AKAP9</i>	Long QT syndrome (AD)	ORPHA 101016 OMIM #611820	1–5:10,000 (Orphanet)	2007 (10) (Chen et al.)	Limited^b (0.5 GE, 1.5 EE)	15 genes are associated with LQTS. Variants in <i>AKAP9</i> account for <1% of cases (Alders et al., 2003).
<i>RPS24</i>	Diamond-Blackfan anemia (AD)	ORPHA 124 OMIM #610629	1:100,000 to 1:200,000 (Clinton & Gazda, 2009)	2006 (11) (Gazda et al.)	Definitive^b (10.5 GE, 3.5 EE, r/t)	~2% of cases are attributed to variants in <i>RPS24</i> (Clinton & Gazda, 2009).
<i>TMPO</i>	Dilated cardiomyopathy (AD)	ORPHA 154 OMIM N/A	1:250 (Hershberger & Morales, 2007)	2005 (12) (Taylor et al.)	Refuted^b	Only 1 proband with a <i>TMPO</i> variant reported in the literature, which was later refuted (Strande & Riggs et al., 2017).
<i>TCAP</i>	Hypertrophic cardiomyopathy (AD)	ORPHA 217569 OMIM #607487	1:500 (Cirino & Ho, 2008)	2004 (13) (Hayashi et al.)	Limited^d (0.2 GE, 1 EE)	>16 genes are associated with HCM. <i>TCAP</i> variants account for <1% of cases (Cirino & Ho, 2008).
<i>MYO1A</i>	Sensorineural hearing loss (AD)	ORPHA 90635 OMIM N/A		2003 (14) (Donaudy et al.)	Refuted^e	
<i>COX15</i>	Leigh syndrome (AR)	ORPHA 255241/1561 OMIM #256000/615119	1:40,000 (Rahman & Thorburn, 2015)	2002 (15) (Antonicka et al.)	Strong^{d,f} (8 GE, 5 EE)	<5% of complex IV-deficient Leigh syndrome is caused by variants in <i>COX15</i> (Rahman & Thorburn, 2015).

HGNC Gene Symbol	Disease (Inheritance)	Orphanet ID, OMIM Phenotype	Disease Prevalence [†]	Year of 1 st assertion in the literature (Total years curated)	Expert Reviewed Classification (Points)	Notes
<i>USH1C</i>	Usher syndrome type 1C (AR)	ORPHA 231169 OMIM #276904	1–9:100,000 (Orphanet)	2000 (17) (Verpy et al.)	Definitive^e (12 GE, 6 EE, r/t)	6 genes are associated with the disease. 1–15% of cases are attributed to variants in <i>USH1C</i> (Lentz & Keats, 1999).
<i>SMAD4</i>	Juvenile polyposis syndrome (AD)	ORPHA 329971 OMIM #174900	1:16,000 to 1:100,000 (Haidle & Howe, 2003)	1998 (19) (Howe et al.)	Definitive^c (12 GE, 6 EE, r/t)	~27% of cases are attributed to variants in <i>SMAD4</i> (Haidle & Howe, 2003).
<i>PMS1</i>	Lynch syndrome (AD)	ORPHA 144 OMIM N/A	1:440 (Kohlmann & Gruber, 2004)	1994 (23) (Nicolaidis et al.)	Refuted^c	Variants in <i>MLH1</i> , <i>MSH2</i> , <i>MSH6</i> , <i>PMS2</i> and <i>EPCAM</i> account for most cases (Kohlmann & Gruber, 2004).
<i>TPMI</i>	Hypertrophic cardiomyopathy (AD)	ORPHA 217569 OMIM #115196	1:500 (Cirino & Ho, 2008)	1994 (23) (Thierfelder et al.)	Definitive^d (12 GE, 6 EE, r/t)	~2% of HCM cases are caused by variants in <i>TPMI</i> (Cirino & Ho, 2008).
<i>GAA</i>	Pompe disease (AR)	ORPHA 365 OMIM #232300	1:14,000 to 1:100,000 (Leslie & Bailey, 2007)	1991 (26) (Zhong et al.)	Definitive^a (12 GE, 6 EE, r/t)	<i>GAA</i> is the only gene associated with Pompe disease (Leslie & Bailey, 2007).

Abbreviations- AD: Autosomal Dominant, AR: Autosomal Recessive, N/A: Not Applicable, r/t: replication over time, GE: Genetic Evidence, EE: Experimental Evidence

[†] Overall disease prevalence, not specific to gene.

[‡] An OMIM entry for Pelger-Huet anomaly with mild skeletal anomalies (OMIM #618019) has been created since the original curation was performed (Strande & Riggs et al., 2017) and may represent the most appropriate disease entity. This should be considered and updated upon re-analysis of the gene-disease association.

Classifications approved by:

^a Jonathan Berg, MD, PhD, ClinGen

^b Reported in Strande & Riggs et al., 2017

^c ClinGen Colon Cancer and Polyposis Gene Curation Expert Panel

^d ClinGen Hypertrophic Cardiomyopathy Gene Curation Expert Panel

^e ClinGen Hereditary Hearing Gene Curation Expert Panel

^f Heather Baudet, MD, PhD, ClinGen

Table 2.

List of novel gene-disease associations from Volume 93 (Issues 1 and 2) of the *American Journal of Human Genetics* used to assess clinical validity classifications from 2013–2016.

HGNC Gene Symbol	Disease (Inheritance)	Orphanet ID, OMIM Phenotype	Disease Prevalence [†]	Expert Reviewed Classification (Points)	Notes	Reference
<i>AQP5</i>	Palmoplantar keratoderma, Bothnian type (AD)	ORPHA 2337 OMIM #600231	1:40,000, 0.3–0.55% in northern Sweden (Orphanet)	Moderate ^a (7.25 GE, 3.5 EE)		Blaydon et al., 2013
<i>ELAC2</i>	Combined oxidative phosphorylation deficiency (AR)	ORPHA 369913 OMIM #615440	N/A	Moderate ^a (6 GE, 4 EE)	>20 genes are involved in complex I deficiency (MIM252010).	Haack et al., 2013
<i>KLHL40</i>	Nemaline myopathy, severe congenital (AR)	ORPHA 171430 OMIM #615348	1:50,000 (North & Ryan, 2002)	Definitive ^a (12 GE, 6 EE, r/t)	Classified as moderate in 2013, strong in 2014, definitive in 2016 due to r/t. There are ~10 genes involved in the disease (North & Ryan, 2002).	Ravenscroft et al., 2013
<i>LRPAP1</i>	Myopia (AR)	ORPHA 98619 OMIM #615431	~25% (Aldamesh, 2013)	Moderate ^a (9.5 GE, 1.5 EE)	Linkage analysis has identified at least 13 loci for high myopia (Jiang et al., 2015).	Aldahmesh et al., 2013
<i>PIK3R1</i>	SHORT syndrome (AD)	ORPHA 3163 OMIM #269880	<50 cases in the literature (Innes & Dymont, 2014)	Definitive ^a (12 GE, 3 EE, r/t)	Upgraded to Definitive in 2016 due to replication over time. <i>PIK3R1</i> is the only gene currently associated with the disease (Innes & Dymont, 2014).	Chudasama et al., 2013; Dymont et al., 2013; Thauvin-Robinet et al., 2013
<i>RAB28</i>	Cone-rod dystrophy (AR)	ORPHA 1872 OMIM #615374	1:30,000 to 1:40,000 (Roosing et al., 2013)	Limited ^a (3.75 GE, 1.5 EE)	Seven genes have been previously implicated in the AR form of the disease (Roosing et al., 2013).	Roosing et al., 2013
<i>ARL2BP</i>	Retinitis pigmentosa (AR)	ORPHA 791 OMIM #268000	1–5:10,000 (Orphanet)	Limited ^a (4.5 GE, 1.5 EE)	AR cases make up 5–20% of cases. There are more than 50 genes associated with the AR condition (Fahim et al., 2000).	Davidson et al., 2013
<i>CTCF</i>	Intellectual disability (AD)	ORPHA 363611 OMIM #615502	1:1,000,000 (Orphanet)	Moderate ^b (4.5 GE, 6 EE)		Gregor et al., 2013

Abbreviations- AD: Autosomal Dominant, AR: Autosomal Recessive, N/A: Not Applicable, r/t: replication over time, GE: Genetic Evidence, EE: Experimental Evidence

[†] Overall disease prevalence, not specific to gene.

Classifications approved by:

^a Jonathan Berg, MD, PhD, ClinGen

Author Manuscript
Christian Schaaf, MD, PhD, FACMG, ClinGen

Author Manuscript

Author Manuscript

Author Manuscript