Improving hearing loss gene testing: a systematic review of gene evidence toward more efficient next-generation sequencing–based diagnostic testing and interpretation

Ahmad N. Abou Tayoun, PhD[1,2,3], Saeed H. Al Turki, PhD[1], Andrea M. Oza, MS, CGC[2],
Mark J. Bowser, MS, MPH[2], Amy L. Hernandez, MS, CGC[2], Birgit H. Funke, PhD[2,4],
Heidi L. Rehm, PhD[2,5], Sami S. Amr, PhD[2,5]

Purpose: With next generation sequencing technology improvement and cost reductions, it has become technically feasible to sequence a large number of genes in one diagnostic test. This is especially relevant for diseases with large genetic and/or phenotypic heterogeneity, such as hearing loss. However, variant interpretation remains the major bottleneck. This is further exacerbated by the lack in the clinical genetics community of consensus criteria for defining the evidence necessary to include genes on targeted disease panels or in genomic reports, and the consequent risk of reporting variants in genes with no relevance to disease.

Methods: We describe a systematic evidence-based approach for assessing gene–disease associations and for curating relevant genes for different disease aspects, including mode of inheritance, phenotypic severity, and mutation spectrum.

Results: By applying this approach to clinically available hearing loss gene panels with a total of 163 genes, we show that a significant number (45%) of genes lack sufficient evidence of association with disease and thus are expected to increase uncertainty and patient anxiety, in addition to intensifying the interpretation burden. Information about all curated genes is summarized. Our retrospective analysis of 539 hearing loss cases tested by our previous OtoGenomeV2 panel demonstrates the impact of including genes with weak disease association in laboratory wet-bench and interpretation processes.

Conclusion: Our study is, to our knowledge, the first to highlight the urgent need for defining the clinical validity of gene–disease relationships for more efficient and accurate clinical testing and reporting.

Genet Med advance online publication 12 November 2015

Key Words: clinical validity; gene–disease association; hearing loss; next-generation sequencing

With a prevalence of 2.5 in 1,000 births,[1] hearing loss represents the most common congenital sensory impairment. It is estimated that more than half of childhood hearing loss has a genetic cause,[2] and over 280 genes are claimed to be associated with either a nonsyndromic or syndromic form of the disease (based on a search for the term "Hearing loss OR Deafness" in the Human Gene Mutation Database (HGMD) in 2014 and information available at the Hereditary Hearing Loss Homepage at http://hereditaryhearingloss.org/). Over 100 genes have been associated with nonsyndromic hearing loss, and an even greater number of genes have been associated with syndromes in which hearing loss is a primary feature. While some of these genes have a well-established causal role in hearing loss, the evidence supporting the association of many of these genes to hearing loss has not been fully assessed and may be inadequate to assume causality.

With the advent of next-generation sequencing (NGS), it has become more feasible and less costly to increase the genes included on diagnostic genetic panels, albeit at the expense of increasingly complex data interpretation due to the large number of novel variants across many genes. In turn, many expanded gene panels do not offer the expected increase in clinical sensitivity thought to be gained through additional gene content.[3] Therefore, evaluating gene–disease associations should be a critical step for establishing diagnostic panels with the highest clinical sensitivity, while reducing the number of reported variants that are unrelated to the clinical manifestations of the patient. Indeed, this approach is recommended by the American College of Medical Genetics and Genomics in their practice guideline for clinical NGS.[4]

Despite the recent guidelines integrating gene-level and variant-level evidence for interpreting sequence variants in human disease,[5] the clinical genetics community lacks both detailed criteria for defining gene–disease associations and necessary cutoffs for including or excluding genes on diagnostic panels such that some genes included on clinical panels have inadequate evidence supporting their contribution to disease. As a

result, available clinical panels for the same disease tend to have highly variable numbers of genes.[6] As an example, five clinical genetics laboratories, registered in the Genetic Testing Registry as of October 2014, target a substantially different number of genes (27–150) in their NGS-based hearing loss testing.

While including only genes with the highest clinical validity will help to alleviate the burden of variant interpretation, curating those genes for different disease and molecular aspects will also facilitate ongoing clinical interpretation and prioritization of variants. For example, hearing loss diagnostic panels include a large number of genes, many of which have characteristic hearing loss features (e.g., severity, onset, laterality, presence or absence of other syndromic features), mode of inheritance (recessive, dominant, X-linked, mitochondrial), and mutation spectrum. Understanding the patient's family history and hearing loss characteristics can help in prioritizing variants in genes known to match the patient's inheritance pattern and disease features. Finally, defining the pathogenic mutation spectrum and any mutational hot spots in each gene can add another important filter in the variant interpretation process.

Here we describe a systematic approach to evaluate evidence for gene–disease associations through data that are available in the literature and in public databases. Using sensorineural hearing loss as a genetic disease model, we propose criteria for a four-tiered gene–disease association classification system, which can guide gene content decisions for diagnostic panels. We apply these criteria to 163 genes offered by five molecular diagnostic laboratories, including our own. We also present a retrospective evaluation of the data interpretation and laboratory "wet-bench" operational burden due to inclusion of weakly associated genes across 539 hearing loss probands tested by our laboratory. Finally, genes with more credible association with hearing loss were thoroughly curated; hearing loss–related information, such as the potential gene–disease mechanism, phenotypic description, inheritance patterns, and variation spectrum, is summarized.

METHODS

Evaluating gene–disease associations for nonsyndromic hearing loss genes

The evidence associating each gene with disease was manually assessed and quantified according to the criteria in **Table 1**.

This process involved evaluating the published functional (in vitro or in vivo) and/or genetic (linkage, case–control, de novo) evidence at the gene and variant levels. Variants reported in the literature or in public databases were used to support a gene–disease association and were assessed to determine pathogenicity based on criteria described by Duzkale et al.[7] In addition, the overall gene tolerance to common functional variants (missense or loss of function) in the control population also was considered as calculated by Petrovski and colleagues.[8]

Genes associated with multiple inheritance patterns and/or both syndromic and nonsyndromic forms of hearing loss were evaluated for each inheritance pattern and clinical presentation separately, and they also were given an overall evidence level equivalent to the highest level for any association (see **Supplementary Tables S1** and **S2** online).

Gene curation

Genes with an evidence level of two or higher were further curated based on the descriptive categories shown in **Table 2**. Data were collected from publications and publically available disease and variant databases, including the Human Gene Mutation Database (http://www.hgmd.cf.ac.uk/ac/index.php), OMIM (http://www.omim.org/), PubMed (http://www.ncbi.nlm.nih.gov/pubmed), GeneReviews (http://www.ncbi.nlm.nih.gov/books/NBK1116/), dbSNP (http://www.ncbi.nlm.nih.gov/SNP/), 1000 Genomes Project (http://www.1000genomes.org/), the National Heart, Lung, and Blood Institute Exome Sequencing Project (http://evs.gs.washington.edu/EVS/), the National Center for Biotechnology Information Reference Sequence Database (http://www.ncbi.nlm.nih.gov/refseq/), and the UCSC genome browser (http://genome.ucsc.edu/), as well hearing loss–specific databases such as the Hereditary Hearing Loss Homepage (http://hereditaryhearingloss.org/). Based on this curation, information about genes with strong and moderate association with hearing loss is summarized in **Supplementary Table S1** online. Information about the remaining genes is shown in **Supplementary Table S2** online.

Cost analysis

A total of 539 hearing loss probands were referred to the Laboratory for Molecular Medicine for genetic testing using the NGS-based hearing loss gene panel (OtoGenome V2).

Table 1 Criteria for gene–disease association

Evidence level	Description	Criteria
0	Undetermined association	No reported evidence
1	Weak association	One or few variants without functional or genetic (linkage, case–control, de novo) evidence OR linkage to a locus without gene-specific functional or genetic evidence OR animal model without human data OR conflicting evidence
2	Likely association	One or two pathogenic or likely pathogenic[a] variants with some, mostly in vitro, functional evidence
3	Strong association	≥2–3 Pathogenic or likely pathogenic[a] variants with strong, mostly in vivo, functional evidence

[a]See Duzkale et al.[7] for variant classification.

This gene panel was launched in 2012 and included the following 72 genes: *ACTG1*, *ATP6V1B1*, *BSND*, *CCDC50*, *CDH23*, *CLDN14*, *CLRN1*, *COCH*, *COL11A2*, *CRYM*, *DFNA5*, *DFNB31*, *DFNB59*, *DIAPH1*, *ESPN*, *ESRRB*, *EYA1*, *EYA4*, *GIPC3*, *GJB3*, *GJB2*, *GJB6*, *GPR98*, *GPSM2*, *GRHL2*, *GRXCR1*, *HGF*, *ILDR1*, *KCNE1*, *KCNQ1*, *KCNQ4*, *LHFPL5*, *LOXHD1*, *LRTOMT*, *MARVELD2*, *MIR182*, *MIR183*, *MIR96*, *MSRB3*, *MTTS1*, *MTRNR1*, *MYH14*, *MYH9*, *MYO15A*, *MYO1A*, *MYO3A*, *MYO6*, *MYO7A*, *OTOA*, *OTOF*, *PCDH15*, *PDZD7*, *POU3F4*, *POU4F3*, *PRPS1*, *RDX*, *SERPINB6*, *SLC17A8*, *SLC26A4*, *SLC26A5*, *TECTA*, *TMC1*, *TIMM8A*, *TJP2*, *TMIE*, *TMPRSS3*, *TPRN*, *TRIOBP*, *USH1C*, *USH1G*, *USH2A*, and *WFS1*.

After evaluating genes on this panel, some genes were determined to be weakly associated with disease based on the criteria in **Table 1**. The cumulative impact of these genes on laboratory "wet-bench" operations and data interpretation was estimated using data from the patient cohort ($n = 539$).

Wet-bench analysis focused mostly on the Sanger sequencing follow-up reactions that are necessary to confirm potential clinically significant variants and to fill in "uncallable" bases (defined as bases with <20× coverage) and/or "failed" regions (defined as regions where any 10 consecutive bases had <20× coverage). The average number of follow-up Sanger amplicons per gene was calculated across all 539 clinical samples (**Supplementary Table S3** online). For cost analysis, we estimated the cost per amplicon to be $8, which includes reagents, supplies, and labor. It is worth noting that this cost might be an underestimate of the burden posed by genes with weak evidence of disease association since it is based only on our laboratory's personnel and amplicon costs.

To estimate the interpretation burden, all variants identified and reported in weakly associated genes were counted and the impact on time and cost was calculated. Variant classification is a manual process that requires an in-depth assessment by a genetic counselor lasting between 22 and 120 min,[7] followed by a 15- to 30-min review by an American Board of Medical Genetics and Genomics–certified geneticist. Therefore, the cost was derived using the median hourly payment rate for a genetic counselor ($35 per hour) and geneticist ($56 per hour) from salary surveys by the National Society of Genetic Counselors (Boston, MA; Median Salary for 2014, available at http://nsgc.org/p/cm/ld/fid=68) and American College of Medical Genetics and Genomics (2011 Salary Survey Report; available at https://www.acmg.net/StaticContent/Survey/2011

ACMGSalarySurveyReportFINALMay2012.pdf), respectively. Furthermore, variants of uncertain clinical significance that contributed to the overall interpretation results were highlighted (**Supplementary Table S4** online). The impact on final report results of these genes was determined through the number of reports with inconclusive result due to these genes, either fully or in part. **Supplementary Table S5** online shows a breakdown of the interpretation time and cost per variant, and also takes into consideration time needed to include a patient's phenotype, age, and family history when preparing the final report. Finally, we describe diagnostic scenarios in which segregation analysis was performed for variants in these genes and the burden incurred, both from an operational and an interpretation perspective.

RESULTS

Gene–disease association criteria: systemic evaluation of clinically available sensorineural hearing loss genes

Our goal was to establish a system for assessing evidence associating any gene with disease, and then using this assessment to define the criteria for including a gene on a diagnostic panel. Although we focused on hearing loss, this approach is applicable to other disease areas. We devised a four-tier system describing gene–disease association evidence levels ranging from 0 to 3, as shown in **Table 1**. Genes with no reported evidence were considered to have "evidence level 0," or an "undetermined association." A gene was proposed to have "weak evidence," or "evidence level 1," if one or a few variants in that gene were reported in patients but lacked supporting functional or genetic evidence. Similarly, a gene with only animal model studies supporting a disease association, but no human data, was classified as having evidence level 1. Genes with a "likely association" to disease, or "evidence level 2," had one or two variants that are classified as pathogenic or likely pathogenic according to Duzkale et al.[7] and had some—mostly in vitro—functional evidence. Finally, genes with more than two pathogenic or likely pathogenic variants as well as supporting in vivo functional evidence, such as mouse models, are classified as "strongly associated," or "evidence level 3" (**Table 1**).

For nonsyndromic hearing loss, we decided to use evidence level 2 as our cutoff for the inclusion of genes on a diagnostic disease panel (**Figure 1**). For syndromic hearing loss, the same cutoff was used for gene inclusion—however, only for genes that have been shown to cause "apparent" nonsyndromic disease in

Table 2 Disease and gene information obtained during curation process

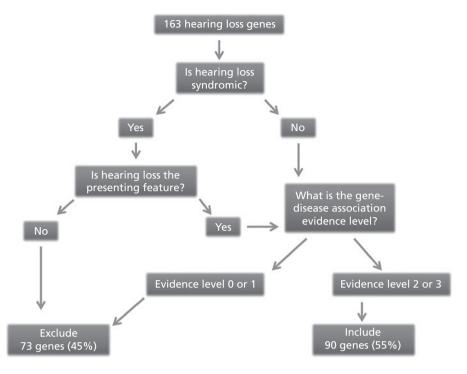| Association	Gene structure	Mutation spectrum	Phenotype	Functional studies
Strength	Exons	Types	Age at onset	In vivo
Disease	Promoters	Mechanism	Penetrance	In vitro
Inheritance	Transcripts	Clustering	Expressivity	
	Expression	Exome Sequencing Project	Key features	
	Protein domains	Database of Genome Variants	Audiological features	
			Syndromic features	

Figure 1 Inclusion criteria for nonsyndromic or "apparent" nonsyndromic hearing loss genes in commercially available diagnostic panels. Refer to **Supplementary Tables S1** and **S2** online for lists of genes that were either included or excluded and for the evidence strength or reason for exclusion.

which hearing loss is the initial presentation or a primary clinical feature. We excluded genes associated with clinically distinct syndromes in which patients have never been reported to present with only hearing loss and/or the other clinical features are very obvious (**Figure 1**).

Using these criteria, we evaluated the evidence for each of the 163 genes currently offered as part of hearing loss panels by five clinical laboratories listing hearing loss panels in the Genetic Testing Registry (last accessed October 2014). Two laboratories were private and offered hearing loss gene panels with 103 or 150 genes. The remaining three laboratories were in academic settings and offered panels consisting of 27, 72, or 90 genes. Although the range of tested genes was 27–150, the total was 163 unique genes. Surprisingly, we found that 73 genes (45%) did not meet our criteria for inclusion on a diagnostic hearing loss panel (**Figure 1**). Of these, 37 genes (~23%) lacked sufficient evidence for association with disease (evidence levels 0 or 1 according to **Table 1**), and 36 genes (~22%) were not qualified to be on a clinical panel because each leads to a distinct disorder in which hearing loss is a feature secondary to other, more prominent findings (**Supplementary Table S2** online). Overall, only 90 genes (55%) were considered to have moderate or strong evidence for association with nonsyndromic or "apparent" nonsyndromic hearing loss. Therefore, for patients with such etiology, and based on the available evidence we strongly recommend including those 90 genes on hearing loss testing panels. However, it is important to monitor the literature for any new evidence regarding the genes described here or any newer hearing loss genes that are yet to be identified.

Examples of genes with no or weak association

Below are a few examples of genes that were commercially available on diagnostic hearing loss panels; nearly all were included by at least three of the five surveyed laboratories. For these commonly sequenced genes, we describe the available evidence and rationale for their "weak" classification based on the criteria in **Table 1**. We do not rule out the potential involvement of some of these genes in hearing function and/or their contribution to hearing loss. Nevertheless, given the *current* available evidence for these genes (**Supplementary Table S2** online), their association to hearing loss is not clear, and therefore they cannot contribute to the clinical validity of the hearing loss test panel. In other words, variants detected in these genes are unlikely to identify variants that can be interpreted as being causal for hearing loss in patients with the present data. However, it is possible that *future* additional work on some of these genes may provide the evidence for a more credible association with disease.

MYO1A. This gene was originally implicated in autosomal-dominant nonsyndromic hearing loss based on a multigenerational Italian family that showed linkage to the DFNA38 locus, which encompasses the *MYO1A* gene.[9] Subsequent analysis of this family failed to identify a causative variant in *MYO1A*; however the authors proceeded to sequence this gene in other families with dominant sensorineural hearing loss and reported eight candidate variants in *MYO1A*.[10] The study lacked segregation or functional data to support pathogenicity of any of the variants identified, failing to support a gene–disease association between *MYO1A* and

dominant nonsyndromic hearing loss. In addition, four of these variants had allele frequencies ranging between 0.5 and 3.8% among Caucasian individuals (Exome Sequencing Project), contradicting the assumption of a causal role for these variants. Interestingly, while preparing this manuscript, two publications[11,12] showed that *MYO1A* is dispensable for normal hearing, consistent with the refutation of causality of the previously reported *MYO1A* variants as related to hearing loss.

PDZD7. This gene, encoding for a PDZ domain–containing scaffold protein, was originally implicated in autosomal-recessive deafness based on the identification of a homozygous reciprocal translocation, 46,XY,t(10;11)(q24.3;q23.3), detected in a boy with nonsyndromic congenital sensorineural hearing loss.[13] The chromosome 10 breakpoint of this translocation was expected to disrupt the open reading frame of one *PDZD7* transcript (one that lacks the PDZ domain; ENST00000393462) while sparing the open reading frames of all other PDZ domain–encoding transcripts. In addition, there was no functional evidence to confirm the impact of this disruption and rule out an effect on neighboring genes. Another report suggested a role for *PDZD7* as both a modifier of retinal disease in one patient with a causative homozygous variant in *USH2A*, and as a contributor to Usher syndrome through digenic interaction with *USH2A* or *GPR98* based on one patient for each gene.[14] Again, this study lacked supportive segregation and/or functional evidence for the two variants in the patients, and thus the association of *PDZD7* with nonsyndromic hearing loss or Usher syndrome based on these two studies is considered to be weak (evidence level 1).

GJB3. Several studies have provided sufficient evidence associating variants in *GJB3* with autosomal-dominant[15–17] and, to a more limited extent, with autosomal-recessive[18,19] erythrokeratodermia variabilis (OMIM 133200), a congenital skin disorder characterized by hyperkeratosis and transient erythema. However, evidence linking this gene with nonsyndromic hearing loss is limited. Several reports identified *GJB3* variants in patients with presumed autosomal-dominant,[20–22] autosomal recessive[23] and digenic[24] hearing loss. These studies, however, did not provide strong segregation or functional evidence to support the pathogenicity of candidate variants in any inheritance pattern, and the authors did not adequately screen for these variants in race-matched controls. For example, two of the purported causative variants, p.Arg32Trp and p.Tyr177Asp (NM_024009.2), have been separately detected at 2.4% (300/12706) and 0.6% (82/12924), respectively, in the Exome Sequencing Project cohort, suggesting that a causative role for these variants in patients with autosomal-dominant hearing loss is unlikely. This is further supported by a recent report that questioned the involvement of this gene in hearing loss through the identification of previously claimed pathogenic variants at high allele frequencies in a control data set.[12]

MIR182/183. These microRNAs occur as part of the microRNA 182/96/183 cluster (*MIR-182/96/183*) that is transcribed in tandem as a single polycistronic primary transcript expressed in the hair cells of the cochlea and the vestibule,[25] and they are thought to regulate sensorineural cell fates in the inner ear.[26] Although three variants in the *MIR96* gene have been shown to segregate with disease in three families with autosomal-dominant nonsyndromic hearing loss,[27,28] no variants were identified in either *MIR182* or *MIR183* in patients with hearing loss. A mouse model strongly supports the contribution of *MIR96* to hearing loss,[29] though no such evidence has been found for *MIR182* and *MIR183*; their expression as part of a polycistronic transcript in the inner ear is insufficient to assume a functional role for these microRNAs in the auditory process.

SLC26A5. The gene–disease association for *SLC26A5* was suggested in a study by Liu et al.,[30] in which a homozygous 5′-untranslated region splice acceptor variant in exon 3 of this gene was identified in two families with presumed recessive nonsyndromic deafness (a sibling pair in one family and a singleton in the second family). Although this variant is predicted to disrupt splicing, there are no functional data to support the prediction and limited data on the impact of the absence of the 5′-untranslated region on gene expression. Two subsequent studies identified three *SLC26A5* variants in two families with autosomal-recessive[31] or -dominant[32] hearing loss. In all three studies, however, segregation analysis was insufficient, and no other evidence was reported to support pathogenicity of any of the variants identified.

FOXI1. This gene was first implicated in Pendred syndrome (OMIM 274600) and nonsyndromic enlarged vestibular aqueduct based on the identification of five heterozygous missense variants in six sporadic cases.[33] In one of these cases, the patient also carried a heterozygous variant in *SLC26A4*, leading the authors to suggest a digenic interaction between *FOXI1* and *SLC26A4* (ref. 33). However, limited evidence for pathogenicity has been reported. Subsequent reports have questioned the association of *FOXI1* with Pendred syndrome and/or enlarged vestibular aqueduct[34,35] because of a lack of evidence and contradicting data. For example, functional analysis of a suspected heterozygous variant in a patient with hearing loss and enlarged vestibular aqueduct did not exhibit an impact on *SLC26A4* expression—the proposed mechanism of pathogenesis for *FOXI1* variants—and a variant in *SLC26A4* was not identified in this individual, casting doubt on the assertion of a digenic interaction between *FOXI1* and *SLC26A4* (ref. 34).

TJP2. A tandem inverted genomic duplication of 270 kb that encompassed the *TJP2* gene was found to segregate in affected members of an Israeli kindred with dominant, adult-onset, nonsyndromic hearing loss,[36] which led to the association of this gene with hearing loss. Although the *TJP2* expression level was shown to be higher in affected individuals carrying the duplication, a positional effect, especially on genes in the vicinity of this region cannot be ruled out. In addition, only

two sequence variants in two small families with autosomal-dominant hearing loss[37] have been reported, and these reports lacked segregation and functional data to support the pathogenicity of either variant. By contrast, a recent report showed an excess of loss-of-function *TJP2* variants in individuals with progressive cholestatic liver disease, strongly associating *TJP2* with this alternate disease.[38]

CRYM. In one study, 192 patients with nonsyndromic hearing loss were screened for causative variants in the *CRYM* gene based on its high expression in the inner ear.[39] Two candidate variants were identified: a missense variant (c.941A>C, p.Lys314Thr, NM_001888.3) in one proband and her affected mother, and a de novo stop-loss variant (c.945A>T, p.X315**Tyr**extX5, NM_001888.3) in the second proband. Interestingly, another variant affecting the stop codon in this gene (c.943T>C, p.X315**Gln**extX5, NM_001888.3) , leading to a similar stop loss effect, has been found in the general population at an allele frequency of 0.2% (7/4,400 African American chromosomes; Exome Sequencing Project). The high frequency of this variant, which is almost equal to the incidence of autosomal-dominant hearing loss (1/500 or 0.2%)[12] argues against a role for the variant in pathogenicity and questions a causal impact for stop-loss variants in this gene. In addition, the fact that *CRYM* has an overall tolerance to common (minor allele frequency >0.1%) functional (missense, nonsense, splice site) variants in the control population[8] is inconsistent with what would be expected for a gene associated with dominant hearing loss. The frequency data, along with the limited evidence for pathogenicity for the two variants identified in the two probands mentioned above, casts uncertainty on the association between *CRYM* and dominant hearing loss.

GJB6. Large deletions in the DFNB1 locus, which includes the *GJB2* and *GJB6* genes, have been strongly associated with autosomal-recessive sensorineural hearing loss.[40,41] These deletions often include the promoter and coding regions of *GJB6*, which lie upstream of *GJB2*; they were first thought to cause hearing loss through biallelic disruption of *GJB6* when homozygous, or through a digenic effect when occurring with a second pathogenic variant in *GJB2*. However, the two most common deletions—*GJB6*-D13S1830 and *GJB6*-D13S1854—were recently shown to significantly downregulate *GJB2* expression. This implies that a loss of *GJB2* expression through disruption of *cis*-regulatory elements, and not a digenic interaction between *GJB2* and *GJB6*, is the underlying mechanism causing hearing loss due to these deletions.[42,43] Therefore, despite the moderate evidence associating missense variants in *GJB6*—namely, Gly11Arg and Ala88Val, NM_006783.4—with autosomal-dominant hidrotic ectodermal dysplasia (Clouston syndrome; OMIM 129500),[44] the evidence associating sequence variants in this gene with autosomal-dominant or digenic (with *GJB2*) nonsyndromic hearing loss[45,46] remains questionable given the limited segregation and functional evidence for reported variants. In summary,

it remains to be determined whether variants affecting the function or expression of *GJB6* are causative for hearing loss.

Curating genes with sufficient evidence for association with disease

To further facilitate ongoing clinical interpretation of variants found in the 90 genes with sufficient evidence for disease association, these genes were thoroughly curated for relevant data to inform the pathogenic variant spectrum across each gene; the related clinical presentation (type of hearing loss, onset, severity and affected frequencies, syndromic features); the inheritance patterns; and the mechanism of disease associated with each gene (**Supplementary Table S1** online; information about the remaining 73 genes is available in **Supplementary Table S2** online). To further delineate gene–disease associations, we provided separate evidence levels for genes reported to cause multiple inheritance patterns and/or clinical presentations. An illustrative example is *COL11A2*, which is associated with autosomal-dominant, recessive, syndromic, and/or nonsyndromic hearing loss. Although the overall evidence level associating *COL11A2* with disease is strong, the evidence associating it with autosomal-recessive nonsyndromic hearing loss is weak compared with that associating this gene with autosomal-dominant nonsyndromic hearing loss or with autosomal-recessive otospondylomegaepiphyseal dysplasia.

When combined with patient-specific features and family history, the gene-specific information in **Supplementary Table S1** can be highly informative for prioritizing relevant variants over others detected through sequencing of all 90 genes. For example, variants in *WFS1* may be more suspect in a patient with low-frequency hearing loss and a dominant family history. Ultimately, automated variant prioritization or pathogenicity prediction tools can potentially be developed by combining patient-specific clinical data with the disease-specific and gene-specific information provided in **Supplementary Table S1** online.

Impact and cost analysis

To assess the impact of including genes with limited evidence on laboratory workflow, we focused our analysis on genes included in an early NGS-based hearing loss diagnostic panel (OtoGenome V2) offered by the Partners Laboratory for Molecular Medicine. This panel consisted of a total of 72 genes thought to underlie nonsyndromic or "apparent" nonsyndromic sensorineural hearing loss. Based on the criteria described here, eight genes (11%) were found to have insufficient evidence for disease causality and were excluded from the newly released and updated OtoGenome V3 panel. In addition, although included in our panel for copy-number analysis, sequence variants in *GJB6* were considered to have weak association with hearing loss and therefore were excluded from interpretation. A total of 539 probands with hearing loss were tested using the OtoGenome V2 panel since 2012, giving us the opportunity to assess retrospectively the impact the eight weakly associated genes had on laboratory wet-bench operations and data interpretation.

Although additional hybridization probes to capture all coding regions in the eight genes are used in our NGS library preparation, the additional cost for such probes is minimal such that we excluded it from our analysis. In addition, there is no extra labor involved in including more probes since a capture kit is run on each sample regardless of its probe content. We also assumed no extra effort or cost on the wet-bench NGS runs because of the high capacity of these types of sequencing platforms. It might be argued, however, that reducing gene content will lead to increased coverage across targeted regions and hence allow for more sample multiplexing per batch/run, ultimately lowering the sequencing cost. Nonetheless, given each laboratory's relatively fixed test volume and the already high coverage in targeted gene panels, the impact of removing those genes on sequencing cost might actually be insignificant. Therefore, for this analysis we considered only the impact on wet-bench processes that come from the accompanying Sanger sequencing needed to ensure "test completeness" as well as to confirm identified variants.

One advantage of focused gene panels over whole-exome or whole-genome sequencing is the ability to ensure complete sequencing coverage across all tested genes by using Sanger sequencing to fill in uncallable bases or failed regions as a result of the low coverage by NGS (see the Methods section). Such sequencing gaps are overwhelming in whole-exome or whole-genome sequencing, rendering Sanger "fill-in" impractical and cost-prohibitive. In our current clinical testing protocol, Sanger sequencing is used to confirm clinically significant variants identified by NGS. On average, 25.4 follow-up Sanger assays were performed for every OtoGenome case ($N = 539$). Of those, 1.4 assays per case were in the eight excluded genes, accounting for a total of 773 unnecessary follow-up Sanger assays among this patient cohort. **Supplementary Table S3** online shows the number of amplicons per gene divided into ones that were the result of failed, uncallable, or variant confirmation reactions. Given the estimated cost per amplicon ($8/amplicon; see Methods), the total extra cost due to Sanger analysis in those genes is estimated to be around $6,184 ($N = 539$ patients), or $11.47 per sample (**Table 3**). This cost analysis might be an underestimate of the actual burden posed by the excluded genes since it was based on labor and amplicon estimates from our laboratory only.

We also investigated the impact of including the weakly associated genes in the interpretation of genetic results, which has now become the major challenge facing clinical laboratories performing expanded gene panels. A total of 197 variants in the 8 genes were identified and required manual assessment. Of these variants, 53 were classified as variants of unknown significance (VUSs) and were included in reports for 51 cases (9% of all 539 reports). The rest were classified as benign or likely benign ($n = 144$).

In 15 of these 51 cases, the hearing loss was explained by the presence of variants in other known hearing loss genes, thus ruling out the involvement of the excluded genes. In 28 cases VUSs were reported in combination with VUSs in genes with strong disease evidence (5%). Eight inconclusive reports (1.5%) were solely due to VUSs detected in the excluded genes. The last group of reports would otherwise have been reported as negative, thereby avoiding the ambiguities in communicating these results to clinicians, patients, and their families, and reducing anxiety for families surrounding the recurrence risk for carriers of these variants. Furthermore, in 6 of the 51 cases with VUSs from excluded genes, segregation analysis was performed by testing other family members to aid in interpretation. This generated an additional 10 reports for 10 family members, with no impact on the overall result in each family, contributing to the uncertainty and potentially increasing anxiety for these families.

Using the time needed to assess and interpret each variant while taking the patient's family history, age, and clinical context into consideration (Methods and **Supplementary Table S5** online), we estimated the total extra time and cost burden ($N = 539$) to be 201 h and $8,269, respectively. This translates to an extra 22 minutes, on average, or $15.34 for interpretation per case (**Table 3**). Overall, with sequencing and interpretation burdens combined, we estimate a cumulative extra cost of $27 per case and a total of $14,450 for all 539 cases.

DISCUSSION

In this work, we provide a systematic approach and a scoring system for assessing gene–disease association. We also present our cutoff for including any gene on a diagnostic NGS panel; by applying this approach to genes that are clinically available for hearing loss testing, we show that a significant number of genes do not have sufficient clinical validity to be included in clinical testing. Using over 500 clinical cases tested for hearing loss using an NGS panel, we retrospectively assessed the impact of including genes with a weak association with hearing loss on laboratory workflow and interpretation. In addition, for genes with sufficient evidence for disease causality and inclusion on panels ($n = 90$), a description of the gene–disease association is presented (**Supplementary Table S1** online), which can be a useful reference to facilitate the variant interpretation process and provides an overview of clinical manifestations (inheritance, onset, severity, and other features) associated with each gene.

Table 3 Total and per-sample extra cost and time associated with sequencing and interpretation of variants in weak genes ($N = 539$ cases)

		Time (h)		Cost ($)	
	Total	**Total**	**Per sample**	**Total**	**Per sample**
Sequencing	773 amplicons	NA	NA	6,184	11.47
Interpretation	197 variants	200.75	0.37	8,269.4	15.34
Total	NA	NA	NA	14,453.4	26.82

NA, not applicable.

Our study clearly highlights the need for performing a thorough review of evidence for gene–disease associations before deciding on gene content for clinical testing. Our experience with hearing loss panels emphasizes the urgent need for defining the strength of gene–disease associations to help guide gene content for diagnostic panels. In fact, this challenge has been recognized by the clinical genetics community and, during the preparation of this article, the Gene Curation Working Group of the Clinical Genome Resource released a seven-tier classification system to organize gene–disease associations (https://www.clinicalgenome.org/knowledge-curation/gene-curation/clinical-validity-classifications/).

The information provided for the 90 hearing loss genes evaluated in this study (**Supplementary Table S1** online) is expected to facilitate variant interpretation in clinical laboratories that offer genetic testing for nonsyndromic hearing loss. Combining patient phenotypic and family history information with current knowledge about the gene's mode of inheritance, expected phenotype, and variant spectrum is useful for variant prioritization. Combined with other parameters, such as conservation and protein domain information, we expect this information to assist in variant assessment. A major challenge, and a potential limitation of this work, is the ability to maintain an up-to-date knowledge base about genes in hearing loss (or any other disease area). One approach would be to standardize the gene assessment process and to create user-friendly interfaces that facilitate gathering or updating information about any gene of interest. In fact, this is one of the goals of the ClinGen Project, and tools to support this activity will soon be available.

Our retrospective analysis shows that several aspects of laboratory testing, including wet-bench processes, variant interpretation, and follow-up familial testing can be more efficient by excluding genes with weak disease associations. This estimation does not take into consideration the extra time and effort that a genetic counselor might spend explaining variants in those genes while delivering the test results to patients. A brief survey of 14 genetic counselors in Partners Healthcare–affiliated hospitals and laboratories showed that, on average, and depending on different clinical and test result scenarios, a genetic counselor would spend extra time communicating variants in genes of unknown clinical significance to patients (unpublished data). We also expect that VUSs from such genes may have a significant psychosocial impact on individuals and their family members; however, this has not been examined in our analysis.

In summary, our study demonstrates that evaluating the clinical validity of all genes on a diagnostic panel can reduce uncertainty and inefficiencies that affect everyone in the testing process, including providers, genetic counselors, laboratory personnel, and, most important, patients.

SUPPLEMENTARY MATERIAL
Supplementary material is linked to the online version of the paper at http://www.nature.com/gim

REFERENCES
1. White KR, Muñoz K. Screening. http://www.infanthearing.org/ncham/publications/screening.pdf. 2008. Accessed October 2015.
2. Smith RJH, Shearer AE, Hildebrand MS, Van Camp G. Deafness and Hereditary Hearing Loss Overview. In: Pagon RA, Adam MP, Ardinger HH, Wallace SE, Amemiya A, Bean LJH, Bird TD, Dolan CR, Fong CT, Smith RJH, Stephens K (eds). *GeneReviews*. University of WashingtonSeattle (WA): Seattle, WA University of Washington, Seattle, 1993–2015.
3. Alfares AA, Kelly MA, McDermott G, et al. Results of clinical genetic testing of 2,912 probands with hypertrophic cardiomyopathy: expanded panels offer limited additional sensitivity. *Genet Med* 2015;17:880–888.
4. Rehm HL, Bale SJ, Bayrak-Toydemir P, et al.; Working Group of the American College of Medical Genetics and Genomics Laboratory Quality Assurance Commitee. ACMG clinical laboratory standards for next-generation sequencing. *Genet Med* 2013;15:733–747.
5. MacArthur DG, Manolio TA, Dimmock DP, et al. Guidelines for investigating causality of sequence variants in human disease. *Nature* 2014;508:469–476.
6. Rehm HL. Disease-targeted sequencing: a cornerstone in the clinic. *Nat Rev Genet* 2013;14:295–300.
7. Duzkale H, Shen J, McLaughlin H, et al. A systematic approach to assessing the clinical significance of genetic variants. *Clin Genet* 2013;84:453–463.
8. Petrovski S, Wang Q, Heinzen EL, Allen AS, Goldstein DB. Genic intolerance to functional variation and the interpretation of personal genomes. *PLoS Genet* 2013;9:e1003709.
9. D'Adamo P, Pinna M, Capobianco S, et al. A novel autosomal dominant nonsyndromic deafness locus (DFNA48) maps to 12q13-q14 in a large Italian family. *Hum Genet* 2003;112:319–320.
10. Donaudy F, Ferrara A, Esposito L, et al. Multiple mutations of MYO1A, a cochlear-expressed gene, in sensorineural hearing loss. *Am J Hum Genet* 2003;72:1571–1577.
11. Eisenberger T, Di Donato N, Baig SM, et al. Targeted and genomewide NGS data disqualify mutations in MYO1A, the "DFNA48 gene", as a cause of deafness. *Hum Mutat* 2014;35:565–570.
12. Shearer AE, Eppsteiner RW, Booth KT, et al. Utilizing ethnic-specific differences in minor allele frequency to recategorize reported pathogenic deafness variants. *Am J Hum Genet* 2014;95:445–453.
13. Schneider E, Märker T, Daser A, et al. 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* 2009;18:655–666.
14. Ebermann I, Phillips JB, Liebau MC, et al. PDZD7 is a modifier of retinal disease and a contributor to digenic Usher syndrome. *J Clin Invest* 2010;120:1812–1823.
15. Morley SM, White MI, Rogers M, et al. A new, recurrent mutation of GJB3 (Cx31) in erythrokeratodermia variabilis. *Br J Dermatol* 2005;152:1143–1148.
16. Richard G, Smith LE, Bailey RA, et al. Mutations in the human connexin gene GJB3 cause erythrokeratodermia variabilis. *Nat Genet* 1998;20:366–369.
17. Wilgoss A, Leigh IM, Barnes MR, et al. Identification of a novel mutation R42P in the gap junction protein beta-3 associated with autosomal dominant erythrokeratoderma variabilis. *J Invest Dermatol* 1999;113:1119–1122.
18. Fuchs-Telem D, Pessach Y, Mevorah B, Shirazi I, Sarig O, Sprecher E. Erythrokeratoderma variabilis caused by a recessive mutation in GJB3. *Clin Exp Dermatol* 2011;36:406–411.

19. Gottfried I, Landau M, Glaser F, et al. A mutation in GJB3 is associated with recessive erythrokeratodermia variabilis (EKV) and leads to defective trafficking of the connexin 31 protein. *Hum Mol Genet* 2002;11:1311–1316.

20. Alexandrino F, Oliveira CA, Reis FC, Maciel-Guerra AT, Sartorato EL. Screening for mutations in the GJB3 gene in Brazilian patients with nonsyndromic deafness. *J Appl Genet* 2004;45:249–254.

21. Kelsell DP, Wilgoss AL, Richard G, Stevens HP, Munro CS, Leigh IM. Connexin mutations associated with palmoplantar keratoderma and profound deafness in a single family. *Eur J Hum Genet* 2000;8:141–144.

22. Xia JH, Liu CY, Tang BS, et al. Mutations in the gene encoding gap junction protein beta-3 associated with autosomal dominant hearing impairment. *Nat Genet* 1998;20:370–373.

23. Liu XZ, Xia XJ, Xu LR, et al. Mutations in connexin31 underlie recessive as well as dominant non-syndromic hearing loss. *Hum Mol Genet* 2000;9:63–67.

24. Liu XZ, Yuan Y, Yan D, et al. Digenic inheritance of non-syndromic deafness caused by mutations at the gap junction proteins Cx26 and Cx31. *Hum Genet* 2009;125:53–62.

25. Weston MD, Pierce ML, Rocha-Sanchez S, Beisel KW, Soukup GA. MicroRNA gene expression in the mouse inner ear. *Brain Res* 2006;1111:95–104.

26. Li H, Kloosterman W, Fekete DM. MicroRNA-183 family members regulate sensorineural fates in the inner ear. *J Neurosci* 2010;30:3254–3263.

27. Mencía A, Modamio-Høybjør S, Redshaw N, et al. Mutations in the seed region of human miR-96 are responsible for nonsyndromic progressive hearing loss. *Nat Genet* 2009;41:609–613.

28. Soldà G, Robusto M, Primignani P, et al. A novel mutation within the MIR96 gene causes non-syndromic inherited hearing loss in an Italian family by altering pre-miRNA processing. *Hum Mol Genet* 2012;21:577–585.

29. Lewis MA, Quint E, Glazier AM, et al. An ENU-induced mutation of miR-96 associated with progressive hearing loss in mice. *Nat Genet* 2009;41:614–618.

30. Liu XZ, Ouyang XM, Xia XJ, et al. Prestin, a cochlear motor protein, is defective in non-syndromic hearing loss. *Hum Mol Genet* 2003;12:1155–1162.

31. Mutai H, Suzuki N, Shimizu A, et al. Diverse spectrum of rare deafness genes underlies early-childhood hearing loss in Japanese patients: a cross-sectional, multi-center next-generation sequencing study. *Orphanet J Rare Dis* 2013;8:172.

32. Toth T, Deak L, Fazakas F, Zheng J, Muszbek L, Sziklai I. A new mutation in the human pres gene and its effect on prestin function. *Int J Mol Med* 2007;20:545–550.

33. Yang T, Vidarsson H, Rodrigo-Blomqvist S, Rosengren SS, Enerback S, Smith RJ. Transcriptional control of SLC26A4 is involved in Pendred syndrome and nonsyndromic enlargement of vestibular aqueduct (DFNB4). *Am J Hum Genet* 2007;80:1055–1063.

34. Cirello V, Bazzini C, Vezzoli V, et al. Molecular and functional studies of 4 candidate loci in Pendred syndrome and nonsyndromic hearing loss. *Mol Cell Endocrinol* 2012;351:342–350.

35. Landa P, Differ AM, Rajput K, Jenkins L, Bitner-Glindzicz M. Lack of significant association between mutations of KCNJ10 or FOXI1 and SLC26A4 mutations in Pendred syndrome/enlarged vestibular aqueducts. *BMC Med Genet* 2013;14:85.

36. Walsh T, Pierce SB, Lenz DR, et al. Genomic duplication and overexpression of TJP2/ZO-2 leads to altered expression of apoptosis genes in progressive nonsyndromic hearing loss DFNA51. *Am J Hum Genet* 2010;87:101–109.

37. Kim MA, Kim YR, Sagong B, et al. Genetic analysis of genes related to tight junction function in the Korean population with non-syndromic hearing loss. *PLoS One* 2014;9:e95646.

38. Sambrotta M, Strautnieks S, Papouli E, et al.; University of Washington Center for Mendelian Genomics. Mutations in TJP2 cause progressive cholestatic liver disease. *Nat Genet* 2014;46:326–328.

39. Abe S, Katagiri T, Saito-Hisaminato A, et al. Identification of CRYM as a candidate responsible for nonsyndromic deafness, through cDNA microarray analysis of human cochlear and vestibular tissues. *Am J Hum Genet* 2003;72:73–82.

40. del Castillo FJ, Rodríguez-Ballesteros M, Alvarez A, et al. A novel deletion involving the connexin-30 gene, del(GJB6-d13s1854), found in trans with mutations in the GJB2 gene (connexin-26) in subjects with DFNB1 non-syndromic hearing impairment. *J Med Genet* 2005;42:588–594.

41. Del Castillo I, Moreno-Pelayo MA, Del Castillo FJ, et al. Prevalence and evolutionary origins of the del(GJB6-D13S1830) mutation in the DFNB1 locus in hearing-impaired subjects: a multicenter study. *Am J Hum Genet* 2003;73:1452–1458.

42. Rodriguez-Paris J, Schrijver I. The digenic hypothesis unraveled: the GJB6 del(GJB6-D13S1830) mutation causes allele-specific loss of GJB2 expression in cis. *Biochem Biophys Res Commun* 2009;389:354–359.

43. Rodriguez-Paris J, Tamayo ML, Gelvez N, Schrijver I. Allele-specific impairment of GJB2 expression by GJB6 deletion del(GJB6-D13S1854). *PLoS One* 2011;6:e21665.

44. Lamartine J, Munhoz Essenfelder G, Kibar Z, et al. Mutations in GJB6 cause hidrotic ectodermal dysplasia. *Nat Genet* 2000;26:142–144.

45. Gardner P, Oitmaa E, Messner A, Hoefsloot L, Metspalu A, Schrijver I. Simultaneous multigene mutation detection in patients with sensorineural hearing loss through a novel diagnostic microarray: a new approach for newborn screening follow-up. *Pediatrics* 2006;118:985–994.

46. Grifa A, Wagner CA, D'Ambrosio L, et al. Mutations in GJB6 cause nonsyndromic autosomal dominant deafness at DFNA3 locus. *Nat Genet* 1999;23:16–18.