Human Genetics
https://doi.org/10.1007/s00439-021-02365-1

REVIEW

Evaluation of copy number variants for genetic hearing loss: a review of current approaches and recent findings

Wafaa Abbasi[1,2] · Courtney E. French[2] · Shira Rockowitz[2] · Margaret A. Kenna[2,3] · A. Eliot Shearer[2,3]

Received: 5 July 2021 / Accepted: 2 September 2021
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

Abstract
Structural variation includes a change in copy number, orientation, or location of a part of the genome. Copy number variants (CNVs) are a common cause of genetic hearing loss, comprising nearly 20% of diagnosed cases. While large deletions involving the gene *STRC* are the most common pathogenic CNVs, a significant proportion of known hearing loss genes also contain pathogenic CNVs. In this review, we provide an overview of currently used methods for detection of CNVs in genes known to cause hearing loss including molecular techniques such as multiplex ligation probe amplification (MLPA) and digital droplet polymerase chain reaction (ddPCR), array-CGH and single-nucleotide polymorphism (SNP) arrays, as well as techniques for detection of CNVs using next-generation sequencing data analysis including targeted gene panel, exome, and genome sequencing data. In addition, in this review, we compile published data on pathogenic hearing loss CNVs to provide an up-to-date overview. We show that CNVs have been identified in 29 different non-syndromic hearing loss genes. An understanding of the contribution of CNVs to genetic hearing loss is critical to the current diagnosis of hearing loss and is crucial for future gene therapies. Thus, evaluation for CNVs is required in any modern pipeline for genetic diagnosis of hearing loss.

Introduction

Sensorineural hearing loss (SNHL) affects 1–3 per 1000 newborns, with approximately 50% of congenital and prelingual deafness being caused by genetic factors (Smith et al. 2005; Morton and Nance 2006; Lieu et al. 2020). Individuals with non-syndromic hearing loss (NSHL), which represents the majority of genetic cases, have no extra-auditory phenotypes. There are hundreds of syndromic forms of hearing loss; some of the most common syndromic forms of hearing loss include Usher syndrome, Pendred Syndrome, and Waardenburg syndrome (Shearer et al. 1999).

Determining causal, or pathogenic, variants in individuals with genetic hearing loss informs management and treatment decisions, prevents redundant diagnostic testing, and provides invaluable information regarding recurrence risk and progression of HL (Lieu et al. 2020). Our knowledge of the pathological mechanisms underlying hearing loss has been made possible due to technological advances in molecular genetics. Over the past decade, single gene sequencing has given way to comprehensive genetic testing of all known hearing loss genes using new DNA sequencing technologies (Shearer and Smith 2015). To date, there are 123 identified non-syndromic hearing loss genes (https://hereditaryhearingloss.org) and more than 7000 reported pathogenic single-nucleotide variants (SNVs) in these genes (https://deafnessvariationdatabase.org) (Azaiez et al. 2018). New DNA sequencing technologies have simultaneously improved the overall diagnostic success of genetic testing for hearing loss while at the same allowing for improved detection of structural variation (SV).

SVs include alterations in copy number (copy number variants, CNVs), changes in genomic orientation (inversions), insertions of mobile elements of the genome, or movements of genomic location (translocations). SVs can also be complex and include combinations of CNVs, inversions, insertions, and translocations. This review will focus primarily on CNVs given that they are the most common cause of pathogenic SVs in humans. Duplications are CNVs which increase the

A. Eliot Shearer
eliot.shearer@childrens.harvard.edu

1 Boston University School of Medicine, Boston, MA, USA

2 Boston Children's Hospital, Children's Rare Disease Cohort Initiative, Boston, MA, USA

3 Department of Otolaryngology Head and Neck Surgery, Harvard Medical School, 300 Longwood Ave, BCH 3129, Boston, MA 02115, USA

Published online: 22 November 2021

number of the alleles to more than two at a particular autosomal genomic locus, while deletions are a reduction in the standard set of two alleles at an autosomal locus. CNVs are differentiated from small insertions or deletions (indels) in that a larger number of nucleotide bases are involved. Although there is no formal definition, an alteration of more than 50 bp is generally considered a CNV. Recurrent CNVs typically have endpoints consistently located in specific regions of the genome containing low copy repeats (LCRs). These recurrent CNVs are often formed by non-allelic homologous recombination (NAHR) involving LCRs. In contrast, non-recurrent CNVs are believed to occur due to non-homologous repair mechanisms that may randomly occur throughout the genome. Proposed mechanisms for formation of these non-recurrent CNVs include non-homologous end joining (NHEJ), microhomology-mediated end joining (MMEJ), and stress-induced, leading to errors in DNA replication (Hastings et al. 2009). While CNVs are believed to play a necessary role in genomic variability of a given population, they are also associated with pathological consequences by way of positional and gene dosage effects, gene disruption and gene fusion (Stranger et al. 2007).

Recent studies have shown that CNVs are a major cause of genetic NSHL. In one study, CNVs were found to be causative in 18.7% of 267 individuals with genetic hearing loss (Shearer et al. 2014). Deletions involving the gene *STRC* are the most common CNVs associated with hearing loss and are, in fact, the second most common pathogenic deafness-causing alleles in humans (causative in 11.2% of genetic diagnoses, second only to *GJB2* c.35delG in 14.4% of diagnoses) (Shearer et al. 2019). Variants in *STRC* were the second most common genetic cause of hearing loss in more than 2460 individuals, comprising 14.3% of all diagnoses (Shearer et al. 2019). But CNVs in *STRC* are not the only pathogenic CNVs in hearing loss: in 2016, we identified pathogenic CNVs in 16 different hearing loss genes (Shearer et al. 2014). Since that time there have been several other publications reporting new hearing loss CNVs as well as advances in technology for detecting these CNVs. The aims of this review are: (1) to provide a description of currently available technologies for detection of CNVs, and (2) to present an update on the genetic contribution of CNVs to genetic hearing loss. Based on these data, it is clear that CNV detection should be included in any diagnostic evaluation for genetic hearing loss.

Methods to detect CNVs

Molecular techniques: MLPA and ddPCR

MLPA is a modified version of multiplex PCR that detects CNVs by comparing the relative amplification of genomic loci through a unique hybridization and ligation method. A control DNA sample is required for complete evaluation, and amplification after hybridization and ligation is measured by fluorescence using a standard capillary electrophoresis instrument. MLPA is a low-cost, high-throughput assay. Assays may be developed in-house but this is labor-intensive. There is one ready-to-use commercial kit targeted towards genetic hearing loss CNVs (MRC Holland, the Netherlands) which includes probes for the *STRC* and *OTOA* regions, the two genes with the most prevalent pathogenic CNVs causing hearing loss (see below). MLPA has a high accuracy but is limited by the need for design of discrete functional probes. Because it is based on unique hybridization sites of the genome, MLPA can determine CNVs in small non-homologous regions of a gene but resolution is limited in regions that share sequence homology with other exons (e.g. pseudogenes). The assay itself requires a > 12-h hybridization step which is time-limiting for high-throughput workflows. The cost per assay is typically less than $50 USD per sample.

qPCR detects copy number changes through threshold cycles (Ct) between target genes and a reference sequence and is often used as a confirmatory assay for CNVs detected by other methods (Gu et al. 2015). It is a low cost, specific way to measure copy number variation, but the process is labor-intensive due to primer selection and multiple runs and results can be unreliable. ddPCR is a modified version of qPCR that maintains the simplicity of a PCR-based assay while detecting CNVs at higher resolution and sensitivity by partitioning the PCR mixture into thousands of discrete reactions. ddPCR has been shown to be effective as a primary assay to detect known deletions or to increase the diagnostic yield of NGS-based dosage analysis when used as a secondary assay, but is limited in its ability to detect novel deletions that affect the targeted region but do not overlap with sequences targeted by the probe (Amr et al. 2018; Tayoun et al. 2016). These assays are less time-consuming than an MLPA. The cost per assay is typically less than $50 USD per sample.

Microarrays

Array-based methods including array-comparative genomic hybridization (array-CGH) and SNP arrays are an efficient high-throughput approach to identifying known CNVs on a genome-wide scale. Unlike PCR-based methods, microarrays provide resolution across the genome, allowing for CNV detection from thousands of genes. In aCGH, CNVs are detected by comparing the quantity of differentially labeled target DNA in loci of interest to co-hybridized reference DNA (Tchinda and Lee 2006). Array-CGH is commonly used as a clinical screening tool to evaluate for suspected genomic deletion syndromes. It is also used to

evaluate genes involved in novel microdeletion and duplication syndromes and due to its high sensitivity and specificity can be used as a confirmatory method for NGS (Freitas et al. 2014; Sinajon et al. 2015; Sugiyama et al. 2019). However, the primary limitation of array-CGH is a low spatial resolution given the non-random distribution of probes around the genome. Most modern array-CGH platforms include more than 300,000 probes with a median of ~ 10,000 base pairs between probes. However, this varies significantly by gene which should be considered when evaluating the suitability of evaluation of hearing loss genes. For this reason, custom array-CGH platforms can be designed to better evaluate target genes (Boone et al. 2010). In addition, array-CGH is limited in its ability to effectively resolve gene conversions or balanced rearrangements.

Alternatively, SNP arrays compare the hybridization intensities of short base-pair DNA probes in a sample to a normalized reference set to identify SNPs across a genome. SNP data are then analyzed by algorithms such as QuantiSNP and PennCNV to infer data on duplications or deletions (Lahbib et al. 2019). This array offers higher quality and resolution and requires less sample DNA than aCGH (Vona et al. 2015). However, there are similar limitations to array designs in terms of spatial resolution being limited to on average ~ 10,000 base pairs which varies by gene. Current SNP arrays typically comprise > 100,000 SNPs across the genome. Informative SNPs in regions of complex CNVs are limited which impacts the comprehensiveness of this array. SNP arrays and array-CGH vary widely in cost depending on whether a standard or custom array is used, but typically cost less than $500 per sample.

Importantly, when considering incorporation of qPCR, ddPCR, or microarrays into a workflow for diagnostics, these assays cannot identify single-nucleotide variants (SNVs), which require a separate assay In contrast, CNV calling using NGS requires only a single assay and specific analysis tools for CNVs and SNVs.

CNV calling using next-generation sequencing (NGS)

Next-generation sequencing (NGS), also termed high-throughput sequencing, has dramatically improved the ability to generate genomic data for genetic diagnosis. This has allowed the advent of comprehensive genetic diagnosis for hearing loss and has now become the standard of care for evaluation of genetic hearing loss (Shearer and Smith 2015). In the new paradigm made available by NGS, a single wet-bench assay is performed leading to the generation of large amounts of data. These data can then be run through varying workflows which may include SNV analysis and also CNV analysis. Therefore, a separate analysis, not separate assay, is typically used for evaluation of CNVs. Crucially, further data analysis may be performed as technology advances without the need for repeat data generation.

Sequencing methods available for detecting CNVs and structural variants include targeted NGS, exome sequencing (ES), genome sequencing (GS), and long-read sequencing. Targeted NGS and ES are useful for calling exon deletions and duplications but have low resolution for determining exact breakpoints and will miss copy number neutral variants, e.g., inversions, where the breakpoints do not overlap the targeted regions. GS has increased power to identify non-exonic breakpoints and copy number changes, if a PCR-free protocol is used. Finally, other technologies such as long-read sequencing and optical mapping (described in the next section) are powerful tools for resolving complex structural variants and those involving repetitive regions of the genome.

There are four methods of structural variant calling generally used for NGS: read depth, paired read, split read and assembly-based (Fig. 1) (Escaramís et al. 2015). Read depth methods use changes in the read depth across deletions and duplications to accurately calculate absolute copy numbers. However, they cannot detect copy number neutral variants and have poor resolution of breakpoints. Paired read methods look for read pairs that map discordantly with regard to the expected insert size or orientation. They can detect copy number neutral variants but are also limited in breakpoint resolution. Split read methods use soft-clipped or split reads that map across breakpoints and can reach single-nucleotide resolution but are limited by short read lengths. Assembly based methods involve de novo assembly of contigs that are then aligned to the reference. They also have single-nucleotide resolution and can discover novel insertion sequences; however, they can be computationally expensive and are prone to errors. Many NGS variant calling algorithms integrate more than one method, such as Manta (read pair, split read, assembly) (Chen et al. 2016).

There are a number of variant calling algorithms available, each generally developed for a specific sequencing technology and/or type of structural variant. CNV callers using read depth methods have been developed to be used on ES (ExomeDepth) (Plagnol et al. 2012), WGS [Canvas (Roller et al. 2016), CNVnator (Abyzov et al. 2011)], or both [GATK-GCNV (Babadi et al. 2018), cn.MOPS (Klambauer et al. 2012)]. Callers for a broader range of structural variants are usually developed for WGS [Manta, Lumpy (Layer et al. 2014), Delly (Rausch et al. 2012), DRAGEN] or long-read sequencing [Sniffles (Sedlazeck et al. 2018), SVIM (Heller and Vingron 2019), NanoVar (Tham et al. 2020)]. Specialized callers for mobile element insertions (MEIs) and short tandem repeat expansions (STRs) have also been developed: MELT (Gardner et al. 2017), Scramble (Torene et al. 2020), TEI (Lee et al. 2012), and ExpansionHunter (Dolzhenko et al. 2019), GangSTR (Mousavi et al.

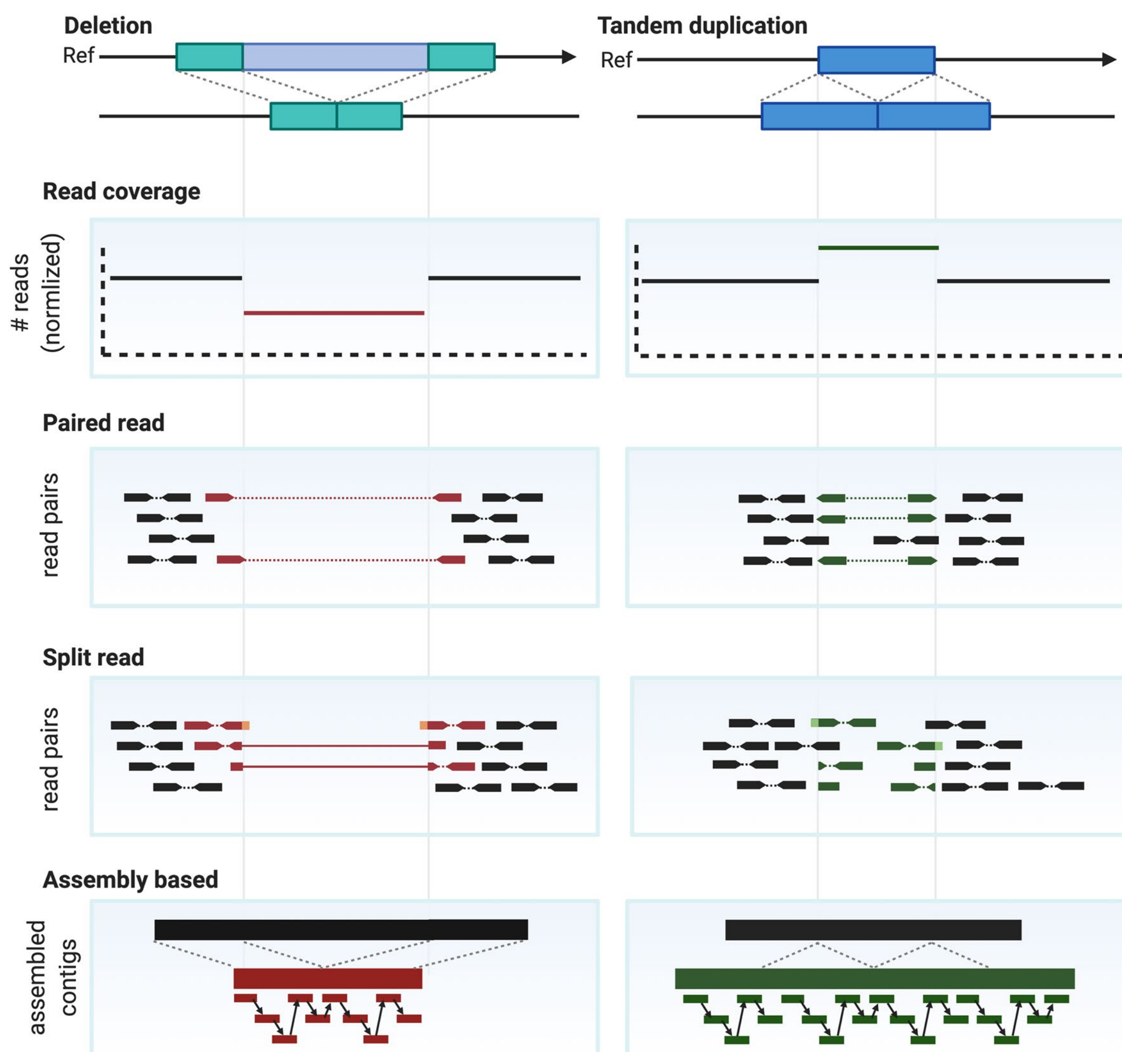

Fig. 1 Structural variant calling methods using next-generation sequencing. Read depth: normalized read depth decreases to 0.5 × at deletions (red) and increases to 1.5 × at duplications (green). Paired read: reads pairs align unexpectedly distant from each other when crossing deletions (red) and are in the wrong orientation when crossing tandem duplication breakpoints (green). Split read: reads that cross deletion (red) or duplication (green) breakpoints are either split (partially align to different regions) or soft-clipped. Assembly based: reads are used to assemble the novel sequence created by the structural variant and then aligned to the reference. Figure created with BioRender.com

2019), HipSTR (Willems et al. 2017), for MEIs and STRs, respectively. Some tools integrate multiple algorithms such as GATK-SV (Manta, Delly, MELT, cn.MOPS) (Collins et al. 2020).

The power to detect structural variants depends on the sequencing protocol, calling method and algorithm, and SV type and size, and these should be taken into consideration when planning or interpreting high-throughput sequencing for SVs in hearing loss genes. Features of particular genomic regions, such as extreme GC content or repetitive elements, can complicate SV detection. In particular, many genes have closely related pseudogenes that reduce read mapping quality and muddle copy number calling. Of note, seven genes implicated in hearing loss are located in highly repetitive regions of the genome and two genes have associated pseudogenes with high homology (*STRC* and *OTOA*) (Shearer et al. 2014). The diversity of SV calling algorithms available combined with few comprehensive benchmarking studies (Cameron et al. 2019; Kosugi et al. 2019; Zhao et al. 2020; Lepkes et al. 2021; Roca et al. 2019; Zhang et al. 2019) means that the choice needs to be decided on a case-by-case basis. Ideally, a combination of multiple technologies and calling algorithms can be deployed to comprehensively investigate structural variants in the genome (Chaisson et al. 2019). For example, while Manta generally rates highly for WGS analysis (Cameron et al. 2019), one may want to add

CNVnator for better detection of large deletions and duplications and MELT for MEIs (Kosugi et al. 2019).

CNV calling using long-read DNA sequencing technologies

Long-read technologies, or third generation sequencing, have the capability of generating reads with lengths ranging from 10 kb to several mb directly from native DNA, making them well adapted for detecting SVs in complex regions of the genome (Logsdon et al. 2020). Although NGS is cost-effective and has well established methods for CNV detection, it has low sensitivity (30–70%) when used to detect structural variations (Jenko Bizjan et al. 2019). Long-read technologies are more likely than NGS to detect complex SVs or CNVs in repetitive regions, using less reads (Norris et al. 2016). Current instruments include Oxford Nanopore Technologies' MinION device and Pacific Biosciences' single-molecule real-time (SMRT) sequencing. MinION is a portable device that translocates an individual DNA strand through a biological nanopore, determining the DNA sequence by measuring changes in ion currents (Jain et al. 2016). SMRT sequencing measures light emissions from fluorescently labeled nucleotides as they are incorporated into DNA (Roberts et al. 2013). Both sequence DNA in real time and provide single-nucleotide resolution. Limitations to broader application of long-read technologies include the higher relative cost, lower throughput and high error rate (Pollard et al. 2018). These technologies, however, are rapidly improving and are expected to play a pivotal role in DNA sequencing.

Optical mapping is another method that may be used for high-resolution evaluation of structural variation. During optical mapping, long segments of intact genomic DNA are directly imaged. This has been commercialized as the Bionano Saphyr (Bionano optical mapping, BOM) (Mak et al. 2016). The BOM method relies on ultra-high molecular weight DNA which is fluorescently labelled at specific sites. The DNA is converted into single strands, pulled through nanochannels, and imaged. Bioinformatics algorithms then create a structural map of the entire genome (essentially a de novo genome assembly), from which SVs may be detected. BOM has been used successfully for detection of several genetic disorders, but not yet for diagnosis of hereditary hearing loss (Chen et al. 2020). This technique has a high sensitivity for detection of SVs across the entire genome but is currently expensive (hundreds or thousands of dollars per sample). In addition, it relies on ultra-high molecular weight DNA, which typically must be isolated separately from standard DNA extraction methods. Therefore, this assay would typically fall outside of a standard molecular diagnostics workflow.

Comparison of CNV calling methods

Each of the methods described for detection of SVs in hearing loss genes have strengths and weaknesses with regards to workflow and provide a different spatial resolution for detection of SVs (Table 1). To demonstrate the different types of data provided by these methods, we performed evaluation of the *STRC-pSTRC* region on a single sample using four different techniques described here and the data is presented in Fig. 2. As described above, no method is perfect for detection of CNVs and Fig. 2 illustrates gaps in even the newest, most expensive technologies (NGS and BOM) at detecting CNVs. An understanding of the technical limitations of the methods used for detection of SVs is required for accurate interpretation of results.

Hearing loss genes affected by CNVs

NSHL displays extreme genetic and phenotypic heterogeneity, making diagnosis difficult. Approximately 70–80% of cases of NSHL are inherited in an autosomal recessive mode of inheritance, 20% as autosomal dominant, 1% as X-linked and < 1% as mitochondrial (Smith et al. 2005). The different types of NSHL are classified by both the affected gene and associated genomic locus. These loci are categorized by mode of inheritance (DFNA: autosomal dominant, DFNB: autosomal recessive, DFNX: X-linked) and numbered by order of discovery. The distinct phenotypic characteristics of a studied population and the different analytical methods used results in a high variability in reported diagnostic yields. Higher rates of diagnosis are shown in individuals with a family history of hearing loss, congenital hearing loss, or bilateral hearing loss cases (Sommen et al. 2016). In addition, methods that include CNV detection have a higher diagnostic yield than methods limited to SNV and indel detection alone (Bademci et al. 2014). To date, pathogenic CNVs in 29 genes, 15% of all known NSHL genes, have been identified (Table 2). The genes most commonly associated with HL include *STRC* (DFNB16), *OTOA* (DFNB22), and *GJB2/GJB6* (DFNB1).

STRC

The DFNB16 locus on chromosome 15q15.3 encompasses the *STRC* gene, the most common cause of mild to moderate HL in several populations. Biallelic mutations in *STRC* account for 5.5–14% of cases of genetic hearing loss, with a greater incidence in specific populations (Lahbib et al. 2019; Sommen et al. 2016; Sloan-Heggen et al. 2016; Moteki et al. 2016; Marková et al. 2018). The gene encodes stereocilin, a

Table 1 Techniques for evaluation of structural variants (SV) in known hearing loss genes

Technique	Cost	Resolution	Separate assay?	Strengths	Weaknesses
MLPA	$	Low	Yes	Commercially available probe-set, low cost	Resolution limited by probe design, time-consuming hybridization protocol, separate assay from SNV evaluation
ddPCR/Qpcr	$	Mid	Yes	Highly adaptable, low cost	Limited by regions targetable by PCR, separate assay from SNV evaluation
Microarray (array-CGH/SNP array)	$$	Mid	Yes	Genome-wide CNV evaluation	Low spatial representation of certain regions of the genome
SV calling from targeted panel or exome sequencing data	$$*	High	No	Can be integrated into SNV workflow (single assay)	Analysis methods vary widely, regions of high homology will be difficult to analyze, may be dependent on large numbers of samples evaluated simultaneously (read depth approach)
SV calling from short-read genome sequencing data	$$$*	Very high	No	High resolution, can be integrated into SNV workflow (single assay), genome-wide CNV evaluation	Regions of high homology will be difficult to analyze, may be dependent on large numbers of samples evaluated simultaneously (read depth approach)
Long-read DNA sequencing	$$$*	Highest	No	High resolution	Cost, may be separate assay
Oxford Bionano optical mapping	$$$	Highest	Yes	High resolution	Cost, separate assay that requires specific DNA extraction method

*If included as part of workflow, there is minimal additional cost for bioinformatics analysis of SVs

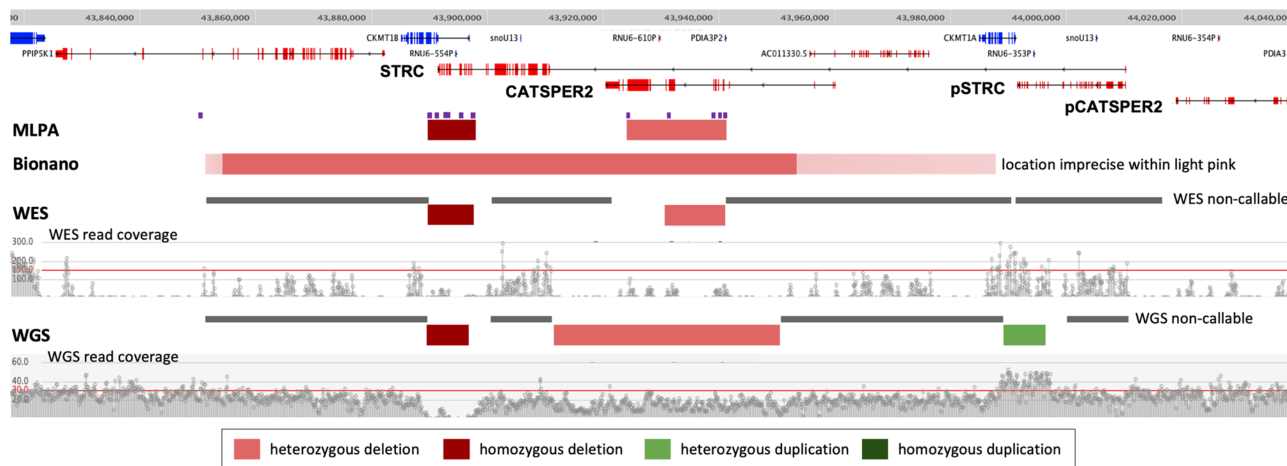

Fig. 2 Structural variant calling using four different methods on the same sample to detect variants in the *STRC-pSTRC* region on chr15q15.3. This individual has hearing loss due to a large hemizygous deletion of the region on one allele and a *pSTRC-STRC* gene conversion event on the other allele. Analysis shown here includes: MLPA with probes (purple) covering *STRC* and *CATSPER2*, Bionano Saphyr optical mapping, and the read depth calling method on both WES and WGS (short read). Read coverage for WES and WGS are also shown

protein that contributes to ciliary motility of outer hair cells in the inner ear. Stereocilin is essential in the formation of top connectors, which maintain longitudinal stiffness of stereocilia during mechanical transduction (Verpy et al. 2011). *STRC* is involved in a tandem duplication on chromosome 15 with three other genes—*CATSPER2*, *HISPPD2A*, and *CKMT1A*. The duplicated region contains *CKMT1B*, a functional homologue of *CKMT1A*, as well as pseudogenes *pPPIP5K1*, *pCATSPER2* and *pSTRC*, all nonfunctional due to inactivating mutations (Knijnenburg et al. 2009). *pSTRC*

Table 2 Previously reported pathogenic copy number variants (CNVs) in known hearing loss genes

Gene	Phenotype	NSHL locus	CNV types	CNV description	Pubmed ID
CDH23	ARNSHL USH	DFNB12	Duplication	Partial duplication	29986705
CIB2	ARNSHL	DFNB48	Deletion	Partial gene deletion	29112224
DFNA5	ADNSHL	DFNA5	Complex	Large deletion and small insertion	9771715
ELMOD3	ARNSHL	DFN88	Deletion	Partial deletion	30284680
EYA4	ADNSHL	DFNA10	Deletion, duplication	Partial gene deletion, partial gene duplication, upstream regulatory region deletion	15735644, 24963352, 32107406
GJB2	ARNSHL ADNSHL	DFNB1 DFNA3	Deletion	Partial, whole gene and upstream regulatory region deletion	19101659, 20236118, 15994881, 25491636, 29595809
GJB6	ARNSHL ADNSHL	DFNB1 DFNA3	Deletion	Partial, whole gene, and upstream regulatory deletion	11896458, 11807148, 11668644, 24963352, 32107406, 25989237, 28405014, 23700267, 24720341, 24774219, 26444186, 26969326, 28012540, 27480936
GRHL2	ADNSHL	DFNA28	Deletion	Whole gene deletion	26455667
LOXDH1	ARNSHL	DFNB77	Deletion	Partial gene deletion	32747562
MARVELD2	ARNSHL	DFNB49	Deletion	Partial gene deletion	33597575
MYH9	ADNSHL	DFNA17	Deletion, duplication	Partial and whole gene deletion, partial gene duplication	18284620, 24963352, 26969326
MYO6	ARNSHL ADNSHL	DFNB37 DFNA22	Deletion, duplication	Partial gene deletion, whole gene duplication	24963352, 31785455
MYO7A	ARNSHL ADNSHL USH	DFNB2 DFNA11	Deletion	Partial gene deletion	9382091, 31598937
OTOA	ARNSHL	DFNB22	Deletion, duplication, conversion	Partial or whole gene deletion, partial gene duplication, pseudogene conversion	19888295, 24963352, 25062256, 26226137, 26969326, 27068579, 28000701, 28501645, 29986705, 31204719, 31527525, 32681043, 33492714, 32747562, 26445815
OTOF	ARNSHL	DFNB9	Deletion	Partial gene deletion	20211493
PCDH15	ARNSHL USH	DFNB23	Deletion, duplication	Partial gene deletion, partial gene duplication	20538994, 16679490, 17277737, 26226137, 26969326, 27344577
PNPT1	ARNSHL	DFNB70	Duplication	Partial gene duplication	24963352
POU3F4	XLNSHL	DFNX3	Deletion, complex rearrangements	Deletions and rearrangements of upstream regulatory regions	2412083, 16365218, 8872461, 19930154, 20668882, 25491636, 26600195
POU4F3	ADNSHL	DFNA15	Deletion	Whole gene deletion	24556497
RDX	ARNSHL	DFNB24	Deletion	Partial gene deletion	29986705
SERPINB6	ARNSHL	DFNB91	Deletion	Partial gene deletion	24963352
SLC26A4	ARNSHL PDS	DFNB4	Deletion	Partial gene deletion	17443271, 18285825, 19287372, 12676893, 24963352, 30268946, 30896630, 28000701, 26445815

Table 2 (continued)

Gene	Phenotype	NSHL locus	CNV types	CNV description	Pubmed ID
STRC	ARNSHL DIS	DFNB16	Deletion, duplication, conversion	Partial, upstream regulatory elements and whole gene deletion, whole gene duplication, pseudogene conversion	11687802, 17098888, 24963352, 24853665, 25062256, 26011646, 26226137, 26969326, 27068579, 27469136, 27480936, 28984810, 29196752, 29339441, 29425068, 29595809, 29986705, 30867468, 31218851, 31552524, 32747562, 28000701, 26445815
TECTA	ARNSHL ADNSHL	DFNB21 DFNA8/12	Deletion	Partial gene deletion	17431902
TJP2	ADNSHL	DFNA51	Duplication	Tandem inverted duplication of entire gene	20602916
TMC1	ARNSHL ADNSHL	DFNB7/11 DFNA36	Deletion	Partial gene deletion	11850618, 19187973, 24963352, 26969326, 31138263
TMPRSS3	ARNSHL	DFNB8/10	Deletion, complex	Partial gene deletion, complex microsatellite insertion	1137999, 24963352, 26969326, 31016883, 32747562
TRIOBP	ARNSHL	DFNB28	Deletion	Whole gene deletion	24963352
WFS1	ADNSHL WFS	DFNA6/14	Deletion, duplication	Partial gene deletion, partial gene duplication	15277431, 24963352

This table was adapted and expanded from Stranger et al. (2007). Note that there were multiple genes with gene deletions and duplications identified in Nishio et al. (2018), but the pathogenicity of these CNVs was not described and so these CNVs are not included in this table (Nishio et al. 2018)

ADNSHL autosomal dominant non-syndromic hearing loss (loci are termed DFNA), *ARNSHL* autosomal recessive non-syndromic hearing loss (loci are termed DFNB), *DIS* deafness infertility syndrome, *PDS* Pendred syndrome, *WFS* Wolfram syndrome, *USH* Usher syndrome, *XLNSHL* X-linked non-syndromic hearing loss

has 98.9% genomic and 99.6% coding sequence homology with *STRC*, with the first 15 exons being 100% identical (Freitas et al. 2014; Lahbib et al. 2019). The majority of *STRC*-associated HL is associated with CNV events, including promoter, partial, whole and multi gene deletions, as well as gene-pseudogene conversions (Shearer et al. 2014; Sloan-Heggen et al. 2016; Moteki et al. 2016; Baux et al. 2017). A multigene deletion that encompasses both *STRC* and *CATSPER2*, which encodes a protein required for sperm motility, results in the deafness infertility syndrome.

OTOA

The DFNB22 locus on chromosome 16p12.2 encompasses *OTOA* which encodes otoancorin, a 120-kda glycoprotein expressed exclusively in the inner ear that contributes to the attachment of acellular gels to the apical surface of sensory epithelia in the inner ear for the maintenance of proper conditions of inner hair cells (Zwaenepoel et al. 2002). The region encompassing *OTOA* on chromosome 16 includes several LCRs (BP1, BP2, and BP3), increasing the likelihood of NAHR events, suggesting this region may be a CNV hotspot (Tassano et al. 2019). *OTOA* is the second most common gene associated with CNVs after *STRC* (Shearer et al. 2014; Sugiyama et al. 2019). CNVs in this gene, including partial and whole gene deletions, have been associated with progressive, mid-frequency, mild to profound hearing loss in several populations, ranging from congenital to childhood deafness (Sugiyama et al. 2019; Bademci et al. 2014; Alkowari et al. 2017; Ramzan et al. 2020). A pseudogene, *OTOAP1*, located 820 kb downstream of *OTOA*, has 99% sequence similarity with exons 22–28 of the gene. Pseudogene mediated gene conversions of OTOA that introduce a premature stop codon in exon 22 (p.Glu 787) have been reported as a cause of DFNB22, most likely through non-allelic homologous double recombination and double strand break repair mediated gene conversion events (Laurent et al. 2021).

GJB2/GJB6

The DFNB1 locus on chromosome 13q12 was the first mapped gene locus associated with ARNSHL and accounts for up to 50% of cases of genetic hearing loss in some populations. Mutations at the DFNB1 locus are the most common cause of severe to profound hearing loss worldwide as well as the second most common cause of mild to moderate hearing loss, after *STRC* (Sloan-Heggen et al. 2016; Mehta et al. 2016). DFNB1 includes the gene *GJB2* (Gap Junction Protein Beta 2) which encodes the 26 kDa Connexin 26 Transmembrane Protein (Cx26), involved in intercellular gap formation and potassium homeostasis necessary for endocochlear potential generation (Tayoun et al. 2016). Large genomic deletions may occur upstream of *GJB2* and may involve the gene *GJB6*. The *GJB6* (Gap Junction Beta Protein 6) gene is located 35 kb upstream of *GJB2*, and encodes Connexin 30 Transmembrane protein (Cx30) and cis-regulatory elements that affect *GJB2* expression. The 309 kb *GJB6* deletion (del GJB6-D13S1830) is the most frequent pre-lingual hearing impairment in the Hispanic population and second most prevalent in Spain with high frequencies of this deletion found in Brazil, France, Czech Republic, Russia and Argentina (Varga et al. 2014; Pandya et al. 2020). del GJB6-D13S1830 deletes the first five exons of *GJB6* and extends distally but keeps *GJB2* intact (Castillo et al. 2002). A similar smaller 232 kb deletion (del (GJB6-D13S1854) has been shown in high frequency in Spain and the UK (Castillo et al. 2005). Other reported CNVs include a 131 kb deletion telomeric to *GJB2* and *GJB6* reducing mRNA expression of both genes, a 101 Kb deletion in GJB2 (del GJB2-13S175) found in Russia among people of Ingush ancestry, a 179.4 kb deletion with undefined breakpoints detected by ddPCR, and a heterozygous 8 kb deletion in GJB2 encompassing the upstream regulatory region of GJB2 (Tayoun et al. 2016; Rehman et al. 2015; Wilch et al. 2010; Bliznetz et al. 2017; Abe et al. 2018). In 8–30% of patients with a heterozygous mutation found in the DFNB1 locus, another variant is not detected. Comparing heterozygous mutation frequencies in the deaf population with those in the normal hearing population shows a higher monoallelic mutation rate in the deaf population, suggesting that most have DFNB1-related deafness with an undetected deletion and are not incidental carriers of *GJB2* variants (Tayoun et al. 2016).

Conclusion

A better understanding of the prevalence, mechanisms, and detection of SVs is critical to improving our ability to understand the pathogenic mechanisms of genetic hearing loss. With newly available technologies, we are only now understanding the large contribution of these complicated mutations to human hearing loss. SVs, and CNVs in particular, contribute to up to 20% of genetic diagnoses and have been identified in 29 different NSHL genes. Deletions are the most common pathogenic mechanism, and the most commonly involved genes are *STRC, OTOA,* and *GJB2*. Duplications are more rarely reported and the pathogenicity of duplications is often difficult to determine.

While some consensus has emerged with regards to use of comprehensive genetic testing for hearing loss with NGS (either ES or targeted panels), there is not yet consensus on an optimal workflow for evaluation of possible pathogenic CNVs in these genes. Methods for CNV detection vary by laboratory and include separate ddPCR testing (Rentas

and Abou 2021), separate microarray testing (Guan et al. 2018), ES bioinformatics analysis (Wang et al. 2021; Zazo Seco et al. 2017), and targeted NGS bioinformatics analysis (Sloan-Heggen et al. 2016). There are clearly differences between the methods and the best method for incorporation varies by laboratory (Table 1). Regardless of the method used, it is critical to evaluate for CNVs in genetic hearing loss. In particular, a significant proportion of genetic mild/moderate hearing loss is due to pathogenic CNVs because *GJB2, STRC,* and *OTOA* are frequent causes of this type of hearing loss and all three genes are commonly involve pathogenic CNVs (Yokota et al. 2019).

New technologies such as GS, long-read sequencing, and optical mapping of the genome will continue to expand our understanding of the contribution of SVs to genetic human hearing loss. These advances, and improved diagnosis, will lay the groundwork for future molecular therapies for hearing loss, given that gene replacement strategies make genes affected by CNVs a prime target for therapeutics.

Author contributions AES: Study conception and reviewed literature. WA: Reviewed literature. All authors wrote, edit, and approved the final manuscript.

Funding No external funding.

Declarations

Conflict of interest The authors declare no conflicts of interest.

References

Abe S, Nishio SY, Yokota Y, Moteki H, Kumakawa K, Usami SI (2018) Diagnostic pitfalls for *GJB2*-related hearing loss: a novel deletion detected by Array-CGH analysis in a Japanese patient with congenital profound hearing loss. Clin Case Rep 6(11):2111–2116. https://doi.org/10.1002/ccr3.1800 (**PMID: 30455902; PMCID: PMC6230644**)

Abyzov A, Urban AE, Snyder M, Gerstein M (2011) CNVnator: an approach to discover, genotype, and characterize typical and atypical CNVs from family and population genome sequencing. Genome Res 21(6):974–984. https://doi.org/10.1101/gr.114876.110 (**Epub 2011 Feb 7. PMID: 21324876; PMCID: PMC3106330**)

Alkowari MK, Vozzi D, Bhagat S, Krishnamoorthy N, Morgan A, Hayder Y, Logendra B, Najjar N, Gandin I, Gasparini P, Badii R, Girotto G, Abdulhadi K (2017) Targeted sequencing identifies novel variants involved in autosomal recessive hereditary hearing loss in Qatari families. Mutat Res 800–802:29–36. https://doi.org/10.1016/j.mrfmmm.2017.05.001 (**Epub 2017 May 4. PMID: 28501645**)

Amr SS, Murphy E, Duffy E, Niazi R, Balciuniene J, Luo M, Rehm HL, Abou Tayoun AN (2018) Allele-specific droplet digital PCR combined with a next-generation sequencing-based algorithm for diagnostic copy number analysis in genes with high homology: proof of concept using stereocilin. Clin Chem 64(4):705–714. https://doi.org/10.1373/clinchem.2017.280685 (**Epub 2018 Jan 16. PMID: 29339441**)

Azaiez H, Booth KT, Ephraim SS, Crone B, Black-Ziegelbein EA, Marini RJ, Shearer AE, Sloan-Heggen CM, Kolbe D, Casavant T, Schnieders MJ, Nishimura C, Braun T, Smith RJH (2018) Genomic landscape and mutational signatures of deafness-associated genes. Am J Hum Genet 103(4):484–497. https://doi.org/10.1016/j.ajhg.2018.08.006 (**Epub 2018 Sep 20. PMID: 30245029; PMCID: PMC6174355**)

Babadi M et al (2018) Abstract 2287: precise common and rare germline CNV calling with GATK. Cancer Res 78(13 Supplement):2287. https://doi.org/10.1158/1538-7445.AM2018-2287

Bademci G, Diaz-Horta O, Guo S, Duman D, Van Booven D, Foster J 2nd, Cengiz FB, Blanton S, Tekin M (2014) Identification of copy number variants through whole-exome sequencing in autosomal recessive nonsyndromic hearing loss. Genet Test Mol Biomarkers 18(9):658–661. https://doi.org/10.1089/gtmb.2014.0121 (**Epub 2014 Jul 25. PMID: 25062256; PMCID: PMC4150376**)

Baux D, Vaché C, Blanchet C, Willems M, Baudoin C, Moclyn M, Faugère V, Touraine R, Isidor B, Dupin-Deguine D, Nizon M, Vincent M, Mercier S, Calais C, García-García G, Azher Z, Lambert L, Perdomo-Trujillo Y, Giuliano F, Claustres M, Koenig M, Mondain M, Roux AF (2017) Combined genetic approaches yield a 48% diagnostic rate in a large cohort of French hearing-impaired patients. Sci Rep 7(1):16783. https://doi.org/10.1038/s41598-017-16846-9 (**PMID: 29196752; PMCID: PMC5711943**)

Bliznetz EA, Lalayants MR, Markova TG, Balanovsky OP, Balanovska EV, Skhalyakho RA, Pocheshkhova EA, Nikitina NV, Voronin SV, Kudryashova EK, Glotov OS, Polyakov AV (2017) Update of the GJB2/DFNB1 mutation spectrum in Russia: a founder Ingush mutation del(GJB2-D13S175) is the most frequent among other large deletions. J Hum Genet 62(8):789–795. https://doi.org/10.1038/jhg.2017.42 (**Epub 2017 Apr 13. PMID: 28405014; PMCID: PMC5584515**)

Boone PM, Bacino CA, Shaw CA, Eng PA, Hixson PM, Pursley AN, Kang SH, Yang Y, Wiszniewska J, Nowakowska BA, del Gaudio D, Xia Z, Simpson-Patel G, Immken LL, Gibson JB, Tsai AC, Bowers JA, Reimschisel TE, Schaaf CP, Potocki L, Scaglia F, Gambin T, Sykulski M, Bartnik M, Derwinska K, Wisniowiecka-Kowalnik B, Lalani SR, Probst FJ, Bi W, Beaudet AL, Patel A, Lupski JR, Cheung SW, Stankiewicz P (2010) Detection of clinically relevant exonic copy-number changes by array CGH. Hum Mutat 31(12):1326–1342. https://doi.org/10.1002/humu.21360 (**Epub 2010 Nov 2. PMID: 20848651; PMCID: PMC3158569**)

Cameron DL, Di Stefano L, Papenfuss AT (2019) Comprehensive evaluation and characterisation of short read general-purpose structural variant calling software. Nat Commun 10(1):3240. https://doi.org/10.1038/s41467-019-11146-4 (**PMID: 31324872; PMCID: PMC6642177**)

Chaisson MJP, Sanders AD, Zhao X, Malhotra A, Porubsky D, Rausch T, Gardner EJ, Rodriguez OL, Guo L, Collins RL, Fan X, Wen J, Handsaker RE, Fairley S, Kronenberg ZN, Kong X, Hormozdiari F, Lee D, Wenger AM, Hastie AR, Antaki D, Anantharaman T, Audano PA, Brand H, Cantsilieris S, Cao H, Cerveira E, Chen C, Chen X, Chin CS, Chong Z, Chuang NT, Lambert CC, Church DM, Clarke L, Farrell A, Flores J, Galeev T, Gorkin DU, Gujral M, Guryev V, Heaton WH, Korlach J, Kumar S, Kwon JY, Lam ET, Lee JE, Lee J, Lee WP, Lee SP, Li S, Marks P, Viaud-Martinez K, Meiers S, Munson KM, Navarro FCP, Nelson BJ, Nodzak C, Noor A, Kyriazopoulou-Panagiotopoulou S, Pang AWC, Qiu Y, Rosanio G, Ryan M, Stütz A, Spierings DCJ, Ward A, Welch AE, Xiao M, Xu W, Zhang C, Zhu Q, Zheng-Bradley X, Lowy E, Yakneen S, McCarroll S, Jun G, Ding L, Koh CL, Ren B, Flicek P, Chen K, Gerstein MB, Kwok PY, Lansdorp PM, Marth GT, Sebat J, Shi X, Bashir A, Ye K, Devine SE, Talkowski ME, Mills RE, Marschall T, Korbel JO, Eichler EE, Lee C (2019)

Multi-platform discovery of haplotype-resolved structural variation in human genomes. Nat Commun 10(1):1784. https://doi.org/10.1038/s41467-018-08148-z (**PMID: 30992455; PMCID: PMC6467913**)

Chen X, Schulz-Trieglaff O, Shaw R, Barnes B, Schlesinger F, Källberg M, Cox AJ, Kruglyak S, Saunders CT (2016) Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications. Bioinformatics 32(8):1220–1222. https://doi.org/10.1093/bioinformatics/btv710 (**Epub 2015 Dec 8. PMID: 26647377**)

Chen M, Zhang M, Qian Y, Yang Y, Sun Y, Liu B, Wang L, Dong M (2020) Identification of a likely pathogenic structural variation in the LAMA1 gene by Bionano optical mapping. NPJ Genom Med 5:31. https://doi.org/10.1038/s41525-020-0138-z (**PMID: 33083009; PMCID: PMC7538933**)

Collins RL, Brand H, Karczewski KJ, Zhao X, Alföldi J, Francioli LC, Khera AV, Lowther C, Gauthier LD, Wang H, Watts NA, Solomonson M, O'Donnell-Luria A, Baumann A, Munshi R, Walker M, Whelan CW, Huang Y, Brookings T, Sharpe T, Stone MR, Valkanas E, Fu J, Tiao G, Laricchia KM, Ruano-Rubio V, Stevens C, Gupta N, Cusick C, Margolin L, Genome Aggregation Database Production Team; Genome Aggregation Database Consortium, Taylor KD, Lin HJ, Rich SS, Post WS, Chen YI, Rotter JI, Nusbaum C, Philippakis A, Lander E, Gabriel S, Neale BM, Kathiresan S, Daly MJ, Banks E, MacArthur DG, Talkowski ME (2020) A structural variation reference for medical and population genetics. Nature 581(7809):444–451. https://doi.org/10.1038/s41586-020-2287-8 (**Epub 2020 May 27. Erratum in: Nature. 2021 Feb;590(7846):E55. PMID: 32461652; PMCID: PMC7334194**)

del Castillo I, Villamar M, Moreno-Pelayo MA, del Castillo FJ, Alvarez A, Tellería D, Menéndez I, Moreno F (2002) A deletion involving the connexin 30 gene in nonsyndromic hearing impairment. N Engl J Med 346(4):243–249. https://doi.org/10.1056/NEJMoa012052 (**PMID: 11807148**)

del Castillo FJ, Rodríguez-Ballesteros M, Alvarez A, Hutchin T, Leonardi E, de Oliveira CA, Azaiez H, Brownstein Z, Avenarius MR, Marlin S, Pandya A, Shahin H, Siemering KR, Weil D, Wuyts W, Aguirre LA, Martín Y, Moreno-Pelayo MA, Villamar M, Avraham KB, Dahl HH, Kanaan M, Nance WE, Petit C, Smith RJ, Van Camp G, Sartorato EL, Murgia A, Moreno F, del Castillo I (2005) 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 42(7):588–594. https://doi.org/10.1136/jmg.2004.028324 (**PMID: 15994881; PMCID: PMC1736094**)

Dolzhenko E, Deshpande V, Schlesinger F, Krusche P, Petrovski R, Chen S, Emig-Agius D, Gross A, Narzisi G, Bowman B, Scheffler K, van Vugt JJFA, French C, Sanchis-Juan A, Ibáñez K, Tucci A, Lajoie BR, Veldink JH, Raymond FL, Taft RJ, Bentley DR, Eberle MA (2019) ExpansionHunter: a sequence-graph-based tool to analyze variation in short tandem repeat regions. Bioinformatics 35(22):4754–4756. https://doi.org/10.1093/bioinformatics/btz431 (**PMID: 31134279; PMCID: PMC6853681**)

Escaramís G, Docampo E, Rabionet R (2015) A decade of structural variants: description, history and methods to detect structural variation. Brief Funct Genomics 14(5):305–314. https://doi.org/10.1093/bfgp/elv014 (**Epub 2015 Apr 15. PMID: 25877305**)

Freitas ÉL, Oiticica J, Silva AG, Bittar RS, Rosenberg C, Mingroni-Netto RC (2014) Deletion of the entire POU4F3 gene in a familial case of autosomal dominant non-syndromic hearing loss. Eur J Med Genet 57(4):125–128. https://doi.org/10.1016/j.ejmg.2014.02.006 (**Epub 2014 Feb 18. PMID: 24556497**)

Gardner EJ, Lam VK, Harris DN, Chuang NT, Scott EC, Pittard WS, Mills RE, 1000 Genomes Project Consortium, Devine SE (2017) The Mobile Element Locator Tool (MELT): population-scale mobile element discovery and biology. Genome Res 27(11):1916–1929. https://doi.org/10.1101/gr.218032.116 (**Epub 2017 Aug 30. PMID: 28855259; PMCID: PMC5668948**)

Gu X, Guo L, Ji H, Sun S, Chai R, Wang L, Li H (2015) Genetic testing for sporadic hearing loss using targeted massively parallel sequencing identifies 10 novel mutations. Clin Genet 87(6):588–593. https://doi.org/10.1111/cge.12431 (**Epub 2014 Aug 7. PMID: 24853665**)

Guan Q, Balciuniene J, Cao K, Fan Z, Biswas S, Wilkens A, Gallo DJ, Bedoukian E, Tarpinian J, Jayaraman P, Sarmady M, Dulik M, Santani A, Spinner N, Abou Tayoun AN, Krantz ID, Conlin LK, Luo M (2018) AUDIOME: a tiered exome sequencing-based comprehensive gene panel for the diagnosis of heterogeneous nonsyndromic sensorineural hearing loss. Genet Med 20(12):1600–1608. https://doi.org/10.1038/gim.2018.48 (**Epub 2018 Mar 29. PMID: 29595809**)

Hastings PJ, Ira G, Lupski JR (2009) A microhomology-mediated break-induced replication model for the origin of human copy number variation. PLoS Genet 5(1):e1000327. https://doi.org/10.1371/journal.pgen.1000327 (**Epub 2009 Jan 30. PMID: 19180184; PMCID: PMC2621351**)

Heller D, Vingron M (2019) SVIM: structural variant identification using mapped long reads. Bioinformatics 35(17):2907–2915. https://doi.org/10.1093/bioinformatics/btz041 (**PMID: 30668829; PMCID: PMC6735718**)

Jain M, Olsen HE, Paten B, Akeson M (2016) The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. Genome Biol 17(1):239. https://doi.org/10.1186/s13059-016-1103-0 (**Erratum in: Genome Biol. 2016 Dec 13;17 (1):256. PMID: 27887629; PMCID: PMC5124260**)

Jenko Bizjan B, Katsila T, Tesovnik T, Šket R, Debeljak M, Matsoukas MT, Kovač J (2019) Challenges in identifying large germline structural variants for clinical use by long read sequencing. Comput Struct Biotechnol J 18:83–92. https://doi.org/10.1016/j.csbj.2019.11.008 (**PMID: 32099591; PMCID: PMC7026727**)

Klambauer G, Schwarzbauer K, Mayr A, Clevert DA, Mitterecker A, Bodenhofer U, Hochreiter S (2012) cn.MOPS: mixture of poissons for discovering copy number variations in next-generation sequencing data with a low false discovery rate. Nucleic Acids Res 40(9):e69. https://doi.org/10.1093/nar/gks003

Knijnenburg J, Oberstein SA, Frei K, Lucas T, Gijsbers AC, Ruivenkamp CA, Tanke HJ, Szuhai K (2009) A homozygous deletion of a normal variation locus in a patient with hearing loss from non-consanguineous parents. J Med Genet 46(6):412–417. https://doi.org/10.1136/jmg.2008.063685 (**Epub 2009 Feb 25. PMID: 19246478**)

Kosugi S, Momozawa Y, Liu X, Terao C, Kubo M, Kamatani Y (2019) Comprehensive evaluation of structural variation detection algorithms for whole genome sequencing. Genome Biol 20(1):117. https://doi.org/10.1186/s13059-019-1720-5 (**PMID: 31159850; PMCID: PMC6547561**)

Lahbib S, Leblond CS, Hamza M, Regnault B, Lemée L, Mathieu A, Jaouadi H, Mkaouar R, Youssef-Turki IB, Belhadj A, Kraoua I, Bourgeron T, Abdelhak S (2019) Homozygous 2p11.2 deletion supports the implication of ELMOD3 in hearing loss and reveals the potential association of CAPG with ASD/ID etiology. J Appl Genet 60(1):49–56. https://doi.org/10.1007/s13353-018-0472-3 (**Epub 2018 Oct 4. PMID: 30284680**)

Laurent S, Gehrig C, Nouspikel T, Amr SS, Oza A, Murphy E, Vannier A, Béna FS, Carminho-Rodrigues MT, Blouin JL, Van Cao H, Abramowicz M, Paoloni-Giacobino A, Guipponi M (2021) Molecular characterization of pathogenic OTOA gene conversions in hearing loss patients. Hum Mutat 42(4):373–377. https://doi.org/10.1002/humu.24167 (**Epub 2021 Mar 14. PMID: 33492714**)

Layer RM, Chiang C, Quinlan AR, Hall IM (2014) LUMPY: a probabilistic framework for structural variant discovery. Genome Biol 15(6):R84. https://doi.org/10.1186/gb-2014-15-6-r84 (**PMID: 24970577; PMCID: PMC4197822**)

Lee E, Iskow R, Yang L, Gokcumen O, Haseley P, Luquette LJ 3rd, Lohr JG, Harris CC, Ding L, Wilson RK, Wheeler DA, Gibbs RA, Kucherlapati R, Lee C, Kharchenko PV, Park PJ, Cancer Genome Atlas Research Network (2012) Landscape of somatic retrotransposition in human cancers. Science 337(6097):967–971. https://doi.org/10.1126/science.1222077 (**Epub 2012 Jun 28. PMID: 22745252; PMCID: PMC3656569**)

Lepkes L, Kayali M, Blümcke B, Weber J, Suszynska M, Schmidt S, Borde J, Klonowska K, Wappenschmidt B, Hauke J, Kozlowski P, Schmutzler RK, Hahnen E, Ernst C (2021) Performance of in silico prediction tools for the detection of germline copy number variations in cancer predisposition genes in 4208 female index patients with familial breast and ovarian cancer. Cancers (basel) 13(1):118. https://doi.org/10.3390/cancers13010118 (**PMID: 33401422; PMCID: PMC7794674**)

Lieu JEC, Kenna M, Anne S, Davidson L (2020) Hearing loss in children: a review. JAMA 324(21):2195–2205. https://doi.org/10.1001/jama.2020.17647 (**PMID: 33258894**)

Logsdon GA, Vollger MR, Eichler EE (2020) Long-read human genome sequencing and its applications. Nat Rev Genet 21(10):597–614. https://doi.org/10.1038/s41576-020-0236-x (**Epub 2020 Jun 5. PMID: 32504078; PMCID: PMC7877196**)

Mak AC, Lai YY, Lam ET, Kwok TP, Leung AK, Poon A, Mostovoy Y, Hastie AR, Stedman W, Anantharaman T, Andrews W, Zhou X, Pang AW, Dai H, Chu C, Lin C, Wu JJ, Li CM, Li JW, Yim AK, Chan S, Sibert J, Džakula Ž, Cao H, Yiu SM, Chan TF, Yip KY, Xiao M, Kwok PY (2016) Genome-wide structural variation detection by genome mapping on nanochannel arrays. Genetics 202(1):351–362. https://doi.org/10.1534/genetics.115.183483 (**Epub 2015 Oct 28. PMID: 26510793; PMCID: PMC4701098**)

Marková SP, Brožková DŠ, Laššuthová P, Mészárosová A, Krůtová M, Neupauerová J, Rašková D, Trková M, Staněk D, Seeman P (2018) STRC gene mutations, mainly large deletions, are a very important cause of early-onset hereditary hearing loss in the czech population. Genet Test Mol Biomarkers 22(2):127–134. https://doi.org/10.1089/gtmb.2017.0155 (**PMID: 29425068**)

Mehta D, Noon SE, Schwartz E, Wilkens A, Bedoukian EC, Scarano I, Crenshaw EB 3rd, Krantz ID (2016) Outcomes of evaluation and testing of 660 individuals with hearing loss in a pediatric genetics of hearing loss clinic. Am J Med Genet A 170(10):2523–2530. https://doi.org/10.1002/ajmg.a.37855 (**Epub 2016 Aug 2. PMID: 27480936**)

Morton CC, Nance WE (2006) Newborn hearing screening–a silent revolution. N Engl J Med 354(20):2151–2164. https://doi.org/10.1056/NEJMra050700 (**PMID: 16707752**)

Moteki H, Azaiez H, Sloan-Heggen CM, Booth K, Nishio SY, Wakui K, Yamaguchi T, Kolbe DL, Iwasa YI, Shearer AE, Fukushima Y, Smith RJ, Usami SI (2016) Detection and confirmation of deafness-causing copy number variations in the STRC gene by massively parallel sequencing and comparative genomic hybridization. Ann Otol Rhinol Laryngol 125(11):918–923. https://doi.org/10.1177/0003489416661345 (**Epub 2016 Jul 28. PMID: 27469136; PMCID: PMC5537730**)

Mousavi N, Shleizer-Burko S, Yanicky R, Gymrek M (2019) Profiling the genome-wide landscape of tandem repeat expansions. Nucleic Acids Res 47(15):e90. https://doi.org/10.1093/nar/gkz501 (**PMID: 31194863; PMCID: PMC6735967**)

Nishio SY, Moteki H, Usami SI (2018) Simple and efficient germline copy number variant visualization method for the Ion AmpliSeq™ custom panel. Mol Genet Genomic Med 6(4):678–686. https://doi.org/10.1002/mgg3.399 (**Epub ahead of print. PMID: 29633566; PMCID: PMC6081219**)

Norris AL, Workman RE, Fan Y, Eshleman JR, Timp W (2016) Nanopore sequencing detects structural variants in cancer. Cancer Biol Ther 17(3):246–253. https://doi.org/10.1080/15384047.2016.1139236 (**Epub 2016 Jan 19. PMID: 26787508; PMCID: PMC4848001**)

Pandya A, O'Brien A, Kovasala M, Bademci G, Tekin M, Arnos KS (2020) Analyses of del(GJB6-D13S1830) and del(GJB6-D13S1834) deletions in a large cohort with hearing loss: caveats to interpretation of molecular test results in multiplex families. Mol Genet Genomic Med 8(4):e1171. https://doi.org/10.1002/mgg3.1171 (**Epub 2020 Feb 17. PMID: 32067424; PMCID: PMC7196463**)

Plagnol V, Curtis J, Epstein M, Mok KY, Stebbings E, Grigoriadou S, Wood NW, Hambleton S, Burns SO, Thrasher AJ, Kumararatne D, Doffinger R, Nejentsev S (2012) A robust model for read count data in exome sequencing experiments and implications for copy number variant calling. Bioinformatics 28(21):2747–2754. https://doi.org/10.1093/bioinformatics/bts526 (**Epub 2012 Aug 31. PMID: 22942019; PMCID: PMC3476336**)

Pollard MO, Gurdasani D, Mentzer AJ, Porter T, Sandhu MS (2018) Long reads: their purpose and place. Hum Mol Genet 27(R2):R234–R241. https://doi.org/10.1093/hmg/ddy177 (**PMID: 29767702; PMCID: PMC6061690**)

Ramzan M, Bashir R, Salman M, Mujtaba G, Sobreira N, Witmer PD, Baylor-Hopkins Center for Mendelian Genomics, Naz S (2020) Spectrum of genetic variants in moderate to severe sporadic hearing loss in Pakistan. Sci Rep 10(1):11902. https://doi.org/10.1038/s41598-020-68779-5 (**PMID: 32681043; PMCID: PMC7368073**)

Rausch T, Zichner T, Schlattl A, Stütz AM, Benes V, Korbel JO (2012) DELLY: structural variant discovery by integrated paired-end and split-read analysis. Bioinformatics 28(18):i333–i339. https://doi.org/10.1093/bioinformatics/bts378 (**PMID: 22962449; PMCID: PMC3436805**)

Rehman AU, Santos-Cortez RL, Drummond MC, Shahzad M, Lee K, Morell RJ, Ansar M, Jan A, Wang X, Aziz A, Riazuddin S, Smith JD, Wang GT, Ahmed ZM, Gul K, Shearer AE, Smith RJ, Shendure J, Bamshad MJ, Nickerson DA, University of Washington Center for Mendelian Genomics, Hinnant J, Khan SN, Fisher RA, Ahmad W, Friderici KH, Riazuddin S, Friedman TB, Wilch ES, Leal SM (2015) Challenges and solutions for gene identification in the presence of familial locus heterogeneity. Eur J Hum Genet 23(9):1207–1215. https://doi.org/10.1038/ejhg.2014.266 (**Epub 2014 Dec 10. PMID: 25491636; PMCID: PMC4538203**)

Rentas S, Abou TA (2021) Utility of droplet digital PCR and NGS-based CNV clinical assays in hearing loss diagnostics: current status and future prospects. Expert Rev Mol Diagn 21(2):213–221. https://doi.org/10.1080/14737159.2021.1887731 (**Epub 2021 Feb 26. PMID: 33554673**)

Roberts RJ, Carneiro MO, Schatz MC (2013) The advantages of SMRT sequencing. Genome Biol 14(7):405. https://doi.org/10.1186/gb-2013-14-6-405 (**Erratum in: Genome Biol. 2017;18(1):156. PMID: 23822731; PMCID: PMC3953343**)

Roca I, González-Castro L, Fernández H, Couce ML, Fernández-Marmiesse A (2019) Free-access copy-number variant detection tools for targeted next-generation sequencing data. Mutat Res 779:114–125. https://doi.org/10.1016/j.mrrev.2019.02.005 (**Epub 2019 Feb 23. PMID: 31097148**)

Roller E, Ivakhno S, Lee S, Royce T, Tanner S (2016) Canvas: versatile and scalable detection of copy number variants. Bioinformatics 32(15):2375–2377. https://doi.org/10.1093/bioinformatics/btw163 (**Epub 2016 Mar 24. PMID: 27153601**)

Sedlazeck FJ, Rescheneder P, Smolka M, Fang H, Nattestad M, von Haeseler A, Schatz MC (2018) Accurate detection of complex structural variations using single-molecule sequencing.

Nat Methods 15(6):461–468. https://doi.org/10.1038/s41592-018-0001-7 (**Epub 2018 Apr 30. PMID: 29713083; PMCID: PMC5990442**)

Shearer AE, Smith RJ (2015) Massively parallel sequencing for genetic diagnosis of hearing loss: the new standard of care. Otolaryngol Head Neck Surg 153(2):175–182. https://doi.org/10.1177/0194599815591156 (**Epub 2015 Jun 17. PMID: 26084827; PMCID: PMC4743024**)

Shearer AE, Hildebrand MS, Smith RJH (1999) Hereditary Hearing Loss and Deafness Overview. [updated 2017 Jul 27]. In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Mirzaa G, Amemiya A (eds) GeneReviews® [Internet]. Seattle (WA): University of Washington, Seattle; 1993–2021. PMID: 20301607

Shearer AE, Kolbe DL, Azaiez H, Sloan CM, Frees KL, Weaver AE, Clark ET, Nishimura CJ, Black-Ziegelbein EA, Smith RJ (2014) Copy number variants are a common cause of non-syndromic hearing loss. Genome Med 6(5):37. https://doi.org/10.1186/gm554 (**PMID: 24963352; PMCID: PMC4067994**)

Shearer AE, Shen J, Amr S, Morton CC, Smith RJ, Newborn Hearing Screening Working Group of the National Coordinating Center for the Regional Genetics Networks (2019) A proposal for comprehensive newborn hearing screening to improve identification of deaf and hard-of-hearing children. Genet Med 21(11):2614–2630. https://doi.org/10.1038/s41436-019-0563-5 (**Epub 2019 Jun 7. Erratum in: Genet Med. 2019 Jun 18;: PMID: 31171844; PMCID: PMC6831511**)

Sinajon P, Gofine T, Ingram J, So J (2015) Microdeletion 8q22.2-q22.3 in a 40-year-old male. Eur J Med Genet 58(11):569–572. https://doi.org/10.1016/j.ejmg.2015.10.004 (**Epub 2015 Oct 9. PMID: 26455667**)

Sloan-Heggen CM, Bierer AO, Shearer AE, Kolbe DL, Nishimura CJ, Frees KL, Ephraim SS, Shibata SB, Booth KT, Campbell CA, Ranum PT, Weaver AE, Black-Ziegelbein EA, Wang D, Azaiez H, Smith RJH (2016) Comprehensive genetic testing in the clinical evaluation of 1119 patients with hearing loss. Hum Genet 135(4):441–450. https://doi.org/10.1007/s00439-016-1648-8 (**Epub 2016 Mar 11. PMID: 26969326; PMCID: PMC4796320**)

Smith RJ, Bale JF Jr, White KR (2005) Sensorineural hearing loss in children. Lancet 365(9462):879–890. https://doi.org/10.1016/S0140-6736(05)71047-3 (**PMID: 15752533**)

Sommen M, Schrauwen I, Vandeweyer G, Boeckx N, Corneveaux JJ, van den Ende J, Boudewyns A, De Leenheer E, Janssens S, Claes K, Verstreken M, Strenzke N, Predöhl F, Wuyts W, Mortier G, Bitner-Glindzicz M, Moser T, Coucke P, Huentelman MJ, Van Camp G (2016) DNA diagnostics of hereditary hearing loss: a targeted resequencing approach combined with a mutation classification system. Hum Mutat 37(8):812–819. https://doi.org/10.1002/humu.22999 (**Epub 2016 May 6. PMID: 27068579**)

Stranger BE, Forrest MS, Dunning M, Ingle CE, Beazley C, Thorne N, Redon R, Bird CP, de Grassi A, Lee C, Tyler-Smith C, Carter N, Scherer SW, Tavaré S, Deloukas P, Hurles ME, Dermitzakis ET (2007) Relative impact of nucleotide and copy number variation on gene expression phenotypes. Science 315(5813):848–853. https://doi.org/10.1126/science.1136678 (**PMID:17289997;PMCID:PMC2665772**)

Sugiyama K, Moteki H, Kitajiri SI, Kitano T, Nishio SY, Yamaguchi T, Wakui K, Abe S, Ozaki A, Motegi R, Matsui H, Teraoka M, Kobayashi Y, Kosho T, Usami SI (2019) Mid-frequency hearing loss is characteristic clinical feature of *OTOA*-associated hearing loss. Genes (basel) 10(9):715. https://doi.org/10.3390/genes10090715 (**PMID: 31527525; PMCID: PMC6770988**)

Tassano E, Ronchetto P, Calcagno A, Fiorio P, Gimelli G, Capra V, Scala M (2019) "Distal 16p12.2 microdeletion" in a patient with autosomal recessive deafness-22. J Genet 98(2):56 (**PMID: 31204719**)

Tayoun AN, Mason-Suares H, Frisella AL, Bowser M, Duffy E, Mahanta L, Funke B, Rehm HL, Amr SS (2016) Targeted droplet-digital PCR as a tool for novel deletion discovery at the DFNB1 locus. Hum Mutat 37(1):119–126. https://doi.org/10.1002/humu.22912 (**Epub 2015 Oct 29. PMID: 26444186**)

Tchinda J, Lee C (2006) Detecting copy number variation in the human genome using comparative genomic hybridization. Biotechniques 41(4):385, 387, 389 passim. https://doi.org/10.2144/000112275 (**PMID: 17068952**)

Tham CY, Tirado-Magallanes R, Goh Y, Fullwood MJ, Koh BTH, Wang W, Ng CH, Chng WJ, Thiery A, Tenen DG, Benoukraf T (2020) NanoVar: accurate characterization of patients' genomic structural variants using low-depth nanopore sequencing. Genome Biol 21(1):56. https://doi.org/10.1186/s13059-020-01968-7 (**PMID: 32127024; PMCID: PMC7055087**)

Torene RI, Galens K, Liu S, Arvai K, Borroto C, Scuffins J, Zhang Z, Friedman B, Sroka H, Heeley J, Beaver E, Clarke L, Neil S, Walia J, Hull D, Juusola J, Retterer K (2020) Mobile element insertion detection in 89,874 clinical exomes. Genet Med 22(5):974–978. https://doi.org/10.1038/s41436-020-0749-x (**Epub 2020 Jan 22. PMID: 31965078; PMCID: PMC7200591**)

Varga L, Mašindová I, Hučková M, Kabátová Z, Gašperíková D, Klimeš I, Profant M (2014) Prevalence of DFNB1 mutations among cochlear implant users in Slovakia and its clinical implications. Eur Arch Otorhinolaryngol 271(6):1401–1407. https://doi.org/10.1007/s00405-013-2559-0 (**Epub 2013 May 23. PMID: 23700267**)

Verpy E, Leibovici M, Michalski N, Goodyear RJ, Houdon C, Weil D, Richardson GP, Petit C (2011) Stereocilin connects outer hair cell stereocilia to one another and to the tectorial membrane. J Comp Neurol 519(2):194–210. https://doi.org/10.1002/cne.22509 (**PMID: 21165971; PMCID: PMC3375590**)

Vona B, Hofrichter MA, Neuner C, Schröder J, Gehrig A, Hennermann JB, Kraus F, Shehata-Dieler W, Klopocki E, Nanda I, Haaf T (2015) DFNB16 is a frequent cause of congenital hearing impairment: implementation of STRC mutation analysis in routine diagnostics. Clin Genet 87(1):49–55. https://doi.org/10.1111/cge.12332 (**Epub 2014 Jan 21. PMID: 26011646; PMCID: PMC4302246**)

Wang J, Xiang J, Chen L, Luo H, Xu X, Li N, Cui C, Xu J, Song N, Peng J, Peng Z (2021) Molecular diagnosis of non-syndromic hearing loss patients using a stepwise approach. Sci Rep 11(1):4036. https://doi.org/10.1038/s41598-021-83493-6 (**PMID: 33597575; PMCID: PMC7889619**)

Wilch E, Azaiez H, Fisher RA, Elfenbein J, Murgia A, Birkenhäger R, Bolz H, Da Silva-Costa SM, Del Castillo I, Haaf T, Hoefsloot L, Kremer H, Kubisch C, Le Marechal C, Pandya A, Sartorato EL, Schneider E, Van Camp G, Wuyts W, Smith RJ, Friderici KH (2010) A novel DFNB1 deletion allele supports the existence of a distant cis-regulatory region that controls GJB2 and GJB6 expression. Clin Genet 78(3):267–274. https://doi.org/10.1111/j.1399-0004.2010.01387.x (**Epub 2010 Mar 1. PMID: 20236118; PMCID: PMC2919588**)

Willems T, Zielinski D, Yuan J, Gordon A, Gymrek M, Erlich Y (2017) Genome-wide profiling of heritable and de novo STR variations. Nat Methods 14(6):590–592. https://doi.org/10.1038/nmeth.4267 (**Epub 2017 Apr 24. PMID: 28436466; PMCID: PMC5482724**)

Yokota Y, Moteki H, Nishio SY, Yamaguchi T, Wakui K, Kobayashi Y, Ohyama K et al (2019) Frequency and clinical features of hearing loss caused by STRC deletions. Sci Rep 9(1):1–9

Zazo Seco C, Wesdorp M, Feenstra I, Pfundt R, Hehir-Kwa JY, Lelieveld SH, Castelein S, Gilissen C, de Wijs IJ, Admiraal RJ, Pennings RJ, Kunst HP, van de Kamp JM, Tamminga S, Houweling AC, Plomp AS, Maas SM, de Koning Gans PA, Kant SG, de Geus CM, Frints SG, Vanhoutte EK, van Dooren MF, van den Boogaard MH, Scheffer H, Nelen M, Kremer H, Hoefsloot L,

Schraders M, Yntema HG (2017) The diagnostic yield of whole-exome sequencing targeting a gene panel for hearing impairment in The Netherlands. Eur J Hum Genet 25(3):308–314. https://doi.org/10.1038/ejhg.2016.182 (**Epub 2016 Dec 21. PMID: 28000701; PMCID: PMC5315517**)

Zhang L, Bai W, Yuan N, Du Z (2019) Comprehensively benchmarking applications for detecting copy number variation. PLoS Comput Biol 15(5):e1007069. https://doi.org/10.1371/journal.pcbi.1007069 (**Erratum in: PLoS Comput Biol. 2019;15(9):e1007367. PMID: 31136576; PMCID: PMC6555534**)

Zhao L, Liu H, Yuan X, Gao K, Duan J (2020) Comparative study of whole exome sequencing-based copy number variation detection tools. BMC Bioinformatics 21(1):97. https://doi.org/10.1186/s12859-020-3421-1 (**PMID: 32138645; PMCID: PMC7059689**)

Zwaenepoel I, Mustapha M, Leibovici M, Verpy E, Goodyear R, Liu XZ, Nouaille S, Nance WE, Kanaan M, Avraham KB, Tekaia F, Loiselet J, Lathrop M, Richardson G, Petit C (2002) Otoancorin, an inner ear protein restricted to the interface between the apical surface of sensory epithelia and their overlying acellular gels, is defective in autosomal recessive deafness DFNB22. Proc Natl Acad Sci USA 99(9):6240–6245. https://doi.org/10.1073/pnas.082515999 (**Epub 2002 Apr 23. PMID: 11972037; PMCID: PMC122933**)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.