Published in final edited form as:
Am J Med Genet A. 2012 February ; 158A(2): 298–308. doi:10.1002/ajmg.a.34391.

Genome-wide SNP Genotyping Identifies the *Stereocilin* (*STRC*) Gene as a Major Contributor to Pediatric Bilateral Sensorineural Hearing Impairment

Lauren J Francey[1], **Laura K Conlin**[2], **Hanna E Kadesch**[1], **Dinah Clark**[1], **Donna Berrodin**[1], **Yi Sun**[1], **Joe Glessner**[3], **Hakon Hakonarson**[1,3], **Chaim Jalas**[4], **Chaim Landau**[4], **Nancy B Spinner**[2], **Margaret Kenna**[5], **Michal Sagi**[6], **Heidi L Rehm**[5,7], and **Ian D Krantz**[1,*]

[1]The Division of Human Genetics, The Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA

[2]The Department of Pathology and Laboratory Medicine, The Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA

[3]The Center for Applied Genomics, The Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA

[4]Bonei Olam Center for Rare Jewish Genetic Disorders, Brooklyn, NY, USA

[5]Dept of Otolaryngology and Communication Disorders, Children's Hospital Boston, Boston, MA 02115, USA

[6]Monique and Jacques Roboh Department of Genetic Research, The Department of Genetic and Metabolic Diseases, Hadassah-Hebrew University Medical Center, Jerusalem, Israel

[7]The Department of Pathology, Harvard Medical School, Boston, MA 02115, USA

Abstract

Hearing loss is the most prevalent sensory perception deficit in humans, affecting 1/500 newborns, can be syndromic or nonsyndromic and is genetically heterogeneous. Nearly 80% of inherited nonsyndromic bilateral sensorineural hearing loss (NBSNHI) is autosomal recessive. Although many causal genes have been identified, most are minor contributors, except for *GJB2*, which accounts for nearly 50% of all recessive cases of severe to profound congenital NBSNHI in some populations. More than 60% of children with a NBSNHI do not have an identifiable genetic cause. To identify genetic contributors, we genotyped 659 *GJB2* mutation negative pediatric probands with NBSNHI and assayed for copy number variants (CNVs). After identifying 8 mild-moderate NBSNHI probands with a Chr15q15.3 deletion encompassing the *Stereocilin* (*STRC*) gene amongst this cohort, sequencing of *STRC* was undertaken in these probands as well as 50 probands and 14 siblings with mild-moderate NBSNHI and 40 probands with moderately severe-profound NBSNHI who were *GJB2* mutation negative. The existence of a *STRC* pseudogene that is 99.6% homologous to the *STRC* coding region has made the sequencing interpretation complicated. We identified 7/50 probands in the mild-moderate cohort to have biallelic alterations in *STRC*, not including the 8 previously identified deletions. We also identified 2/40 probands to have biallelic alterations in the moderately severe-profound NBSNHI cohort, notably no large deletions in combination with another variant were found in this cohort. The data suggest that

*Correspondence: Ian D. Krantz, M.D., University of Pennsylvania School of Medicine, The Children's Hospital of Philadelphia, 1007C Abramson Research Center, 3615 Civic Center Boulevard, Philadelphia, PA 19104-4318, Tel. (215) 590-2931 Fax. (215) 590-3850, ian2@mail.med.upenn.edu.

STRC may be a common contributor to NBSNHI among *GJB2* mutation negative probands, especially in those with mild to moderate hearing impairment.

Keywords

Bilateral sensorineural hearing loss; SNHI; Chr15q15.3; *Stereocilin*; *STRC*; *DFNB16*; SNP genotyping array; copy number variation; CNV

INTRODUCTION

Hearing impairment is the most prevalent sensory perception deficit in humans affecting as many as 1 in 500 newborns with a pre-lingual hearing impairment [Hilgert et al., 2009]. Pre-lingual sensorineural hearing impairment (SNHI) refers to a hearing impairment before the development of speech. More than 50% of pre-lingual SNHI is attributed to genetic causes with approximately 80% of the genetic causes resulting in a non-syndromic form of SNHI [Peterson and Willems, 2006]. Autosomal recessive inheritance comprises the majority (nearly 80%) of nonsyndromic pre-lingual SNHI [Hilgert et al., 2009]. Among this category of autosomal recessive (AR) nonsyndromic bilateral sensorineural hearing impairment (NBSNHI) with pre-lingual age of onset there is abundant genetic heterogeneity. There have been 68 loci mapped for AR NBSNHI with upwards of 40 genes identified as a causative factor [Van Camp and Smith, 2011]. Although the list of AR loci mapped, and the number of genes involved is large, *DFNB1A/B* [OMIM 220290]/[OMIM 612645] harboring the *GJB2* [OMIM 121011] and *GJB6* [OMIM 604418] genes that encode the proteins Connexin26 and Connexin30 respectively, has been the only major etiological contributor to AR NBSNHI to date with nearly 50% of severe to profound congenital AR NBSNHI cases having causative *GJB2* mutations in some populations [Morton and Nance, 2006; Petit et al., 2001].

Recently, advances in genome-wide SNP genotyping array technologies have uncovered a new form of common genetic variation in the human genome, structural variations termed copy number variants (CNVs) [Conrad et al., 2006; Freeman et al., 2006; Redon et al., 2006; Sebat et al., 2004]. CNVs can take the form of small insertions, duplications or deletions that are greater than 1kb in size and often undetectable by non-array based cytogenetic technologies [Feuk et al., 2006a]. CNVs can influence gene expression, phenotypic variation and adaptation by disrupting genes and altering gene dosage, which may cause disease or confer risk to disease traits [Feuk et al., 2006b; McCarroll and Altshuler 2007; Redon et al., 2006; Wong et al., 2007]. With the advent of array-based technologies to assess CNVs in the human genome, it has been possible to look for CNVs that may have clinical relevance to the hearing impaired population. Deletion CNVs are known to affect a large number of genes and evidence also suggests that deletion CNVs are under negative selection [Conrad et al., 2006; Feuk et al., 2006b; Redon et al., 2006]. In this respect, rare deletion CNVs can be used as a starting point to uncover pathogenic genes implicated in individuals with a hearing loss. Finding deletion CNVs as a molecular mechanism of hearing loss may also implicate a gene within the deletion CNV where point mutations are the more prevalent cause in the hearing impaired population. When focusing primarily on an AR mode of inheritance, heterozygous CNVs in the population may represent carriers of a pathogenic CNV involving a gene causative of hearing loss. A commonly reported deletion CNV on Chr15q15.3 was recently identified as a cause of principally non-syndromic hearing loss, when present in a homozygous state in a proband with a slowly progressive, moderate hearing loss at all frequencies [Knijnenburg et al., 2009]. This deletion CNV on Chr15q15.3 was also found to be segregating in 3 unrelated consanguineous families with a moderate-to-severe hearing loss at all frequencies, asthenoteratozoospermia and a homozygous contiguous gene deletion

involving *Cation channel sperm-associated protein 2* (*CATSPER2* [MIM 607249]) and *Stereocilin (STRC* [MIM 606440]) termed the Deafness Infertility Syndrome (DIS [MIM 611102]) [Zhang et al., 2007]. *CATSPER2* encodes a testes specific protein that is responsible for sperm motility and has been shown to cause a nonsyndromic AR form of male infertility when mutated [Avidan et al., 2003]. *STRC* is an AR hearing loss gene that maps to *DFNB16* [MIM 603720] on Chr15q15.3 and was shown to be the cause of hearing loss in 2 unrelated consanguineous Middle Eastern families that had an early onset and non-progressive bilateral SNHI involving all frequencies and was more severe in the higher frequencies [Campbell et al., 1997; Verpy et al., 2001; Villamar et al., 1999]. Aside from these few reported cases *STRC* mutational analysis has not been investigated in a larger cohort of individuals with nonsyndromic hearing loss, due in large part to the complex genomic region in which *STRC* lies. This large segmental duplication includes a *STRC* pseudo gene that is 99.6% conserved. *STRC* encodes an extracellular structural protein found in the stereocilia of the outer hair cells (OHCs) in the inner ear called Stereocilin (STRC). STRC is associated with the OHCs horizontal top connectors and the tectorial membrane attachment crowns [Verpy et al., 2001; Verpy et al., 2011]. Horizontal top connectors are one of the permanent lateral links between the stereocilia that have been shown to contribute to the proper cohesion and positioning of stereociliary tips [Frolenkov et al., 2004; Nayak et al., 2007; Petit and Richardson, 2009; Verpy et al., 2011]. In the absence of horizontal top connectors, the *Strc* $-/-$ mouse loses the precise alignment stability within and between rows of stereocilia in OHCs resulting in a disorganized hair bundle [Verpy et al., 2008].

Here we present data from a cohort of probands with NBSNHI ranging in severity from a mild to profound hearing loss having biallelic alterations in *STRC* by SNP genotyping array or Sanger Sequencing. Mutations in *STRC* were found in 5/10 *GJB2* mutation negative AR NBSNHI families [2 or more hearing impaired siblings born to normal hearing parents], suggesting that STRC may be a significant etiological contributor to AR NBSNHI.

SUBJECTS AND METHODS

Study Cohort

The 669 probands enrolled in this hearing loss cohort presented at either The Genetics of Hearing Impairment Clinic at The Children's Hospital of Philadelphia (CHOP) for a genetic evaluation of their hearing loss [n = 395], the Laboratory of Molecular Medicine (LMM) at the Partner's Healthcare Center for Personalized Genetic Medicine [n = 272] or The Hadassah Hebrew University Medical Center [n=2]. Patients and family members presenting at CHOP were mostly referred by an otolaryngologist and/or audiologist at CHOP and evaluated both by a dysmorphologist and genetic counselor before enrollment in the hearing loss study under an institutional review board-approved protocol of informed consent. Representation from all pediatric ages (ranging from newborn to 18 years) was enrolled with an average age of 3–4 years at the time of enrollment and an equal gender distribution. Most patients (95%) had either a temporal bone MRI or CT performed. All patients with enlarged vestibular aqueducts (EVAs) were evaluated for mutations in *SLC26A4* (Pendred syndrome gene). Patients with clear dysmorphology or an established, or suspected, syndrome were not seen in this clinic but were evaluated/referred to a general genetics clinic. Patient samples from LMM were obtained through an anonymous discarded tissue IRB-approved protocol with limited clinical information available. All CHOP and LMM patients underwent molecular analysis for at least *GJB2* (complete coding region sequencing and exon 1) and *GJB6* deletion analysis as well as additional genes [driven by the clinical exam and family history] in some cases (e.g. *MTNR1, MTTS1, SLC26A4, MYO7A, OTOF, USH2A)*. Following this screening, all individuals with an identified molecular cause for their SNHI were excluded from this study cohort. Two Ashkenazi Jewish families with AR

NBSNHI were obtained from The Hadassah Hebrew University Medical Center under an approved protocol with informed consent. The two families were screened for *GJB2* and mapped for regions of homozygosity using the Affymetrix 250k SNP array. Included in our study cohort were 15 familial cases [2 or more affected siblings] without a known molecular cause for their NBSNHI. Severity of NBSNHI is defined by the degree of loss of the better ear. The cohort was further stratified by severity into two groups: mild-moderate (encompassing mild = 20–40 dB and moderate = 40–70 dB) and moderately severe to profound (encompassing severe = 70–90 dB and profound > 90 dB) with 254/669 of our cohort classified as having a mild to moderate hearing loss, 265/669 with a moderately severe to profound hearing loss and 150/669 were not able to be classified (Table I).

Sample preparation and array analysis

DNA was extracted from peripheral blood or from cultured lymphoblastoid cell lines if peripheral blood samples were not available using the Puregene genomic DNA isolation kit (Qiagen). The quality of the DNA was monitored by analysis of OD_{260}/OD_{280} and OD_{260}/OD_{230} ratios. Acceptable samples had values between 1.8 and 2.0 and ratios > 2.0, respectively. Thirty microliters of a 50 to 100ng/ul solution of genomic DNA was aliquoted into 96 well plates and genotyped on the Illumina BeadStation (Illumina, Inc.). The samples were whole genome-amplified, fragmented, hybridized, fluorescently tagged and scanned by the Center for Applied Genomics at CHOP according to recommended protocols. Research obtained patient analyses [n=509] were carried out using the Illumina HumanHap550v3 BeadChip(Illumina 550k), which contains 561,466 SNP probes, distributed genome-wide. Patients [n=150] obtained after 6/1/2009 were arrayed using the IlluminaQuad610 array (Illumina 610k) containing most of the SNP probes found on the Illumina 550k, an additional 37,355 SNP probes and 21,890 intensity-only CNV probes placed in regions where SNP coverage is poor. One patient arrayed on the Illumina 610k array platform was validated using the Illumina HumanOmni1-Quad BeadChip (Illumina Omni) containing 1,140,419 SNP probes to provide further resolution. All SNP array data were viewed in BeadStudio (Illumina, Inc.), the user interface that supports Illumina's SNP array data, for quality analysis. All samples had call rates greater than 98% with a standard deviation of $\log_2$R ratios below 0.35.

Copy number analysis

A total of 659 *GJB2* mutation negative probands with NBSNHI ranging in severity from mild to profound and 1,910 controls [mixed ethnicities representative of the US population] obtained from the Center for Applied Genomics at the Children's Hospital of Philadelphia were run on either the Illumina 550k SNP genotyping array [n=509] or the Illumina 610k SNP genotyping array [n=150] and normalized to the intersecting probes on the Illumina 550k SNP genotyping array before calling CNVs using the PennCNV detection algorithm [Wang et al., 2007]. One patient was also analyzed using the Illumina Omni SNP genotyping array for subsequent probe resolution. The 1,910 controls were arrayed using the Illumina 610k SNP array and normalized to the intersecting Illumina 550k SNP array probes before being used in the analysis. Further analysis was performed using the unfiltered probe set from the Illumina 610k SNP platform [n=150] and called using the CNV Workshop detection algorithm [Gai et al., 2010]. All CNV deletions called were manually assessed using BeadStudio.

A subset of 90 probands with varying degrees of hearing loss was analyzed using Sanger sequencing for copy number analysis. 80/90 probands had been previously genotyped as specified above and 10/90 probands were either not genotyped or their array data did not meet the quality thresholds for further analysis. Generic primer sets that would simultaneously amplify both the *STRC* and *pSTRC* gene copies were used in regions where

divergent base pairs existed in the amplimers (See Supporting Information onlin Supplemental eTable I). These divergent base pairs were analyzed by Mutation Surveyor v3.3 (Soft Genetics, LLC. Pennsylvania, USA) to determine the relative change in peak heights between the two gene copies following the user guide on mosaicism detection.

Mutational screening cohort

The mutational screening cohort included 90 probands-50 probands [and 14 siblings] having mild-moderate NBSNHI and 40 probands with moderately severe to profound NBSNHI-who were *GJB2* mutation negative and have an unknown molecular etiology for their hearing loss. The sequencing controls consisted of 96 individuals without a hearing loss. Genomic DNA was isolated from peripheral blood lymphocytes following the manufacturer's protocol (Gentra Puregene Blood kit, Qiagen).

STRC sequence analysis and primer design

Primers were designed to amplify all coding exons of *STRC* and the homologous *pSTRC* sequence using the Primer3 algorithm available through the UCSC Genome Browser's ExonPrimer application with the masking feature turned off (http://genome.ucsc.edu/). *STRC* sequence annotation corresponds to NM_153700.2 coding sequence in the UCSC genome build hg18. These 25 primers are termed generic primer sets and encompass exons 1–29 of *STRC* and *pSTRC*. Another set of 9 primers was designed manually using several *in silico* bioinformatics analysis tools. The BLAST (http://www.ncbi.nlm.nih.gov/) and BLAT (http://genome.ucsc.edu/) tools were used to perform multiple sequence alignments to determine areas of nonhomology around the coding sequence where *STRC* specific primers could then be designed to distinguish the two copies by PCR. These manually designed primer pairs were tested using *in silico* primer tests available in MacVector (MacVector, Inc.) and NCBI's PrimerBLAST algorithms (http://www.ncbi.nlm.nih.gov/). These 9 primers are termed unique primer sets and encompass exons 16 and 19–26. All primers were ordered from Integrated DNA Technologies and reconstituted at 200uM. Primer sequences, annealing temperatures and PCR product sizes are available upon request.

STRC mutational analysis

Each 25 ul PCR reaction contained 1x PCR Buffer, 1.5 mM MgCl2, 200uM of each dNTP, 1.5 U of AmpliTaq Gold Polymerase (Applied Biosystems), 0.4 uM of each of the forward and reverse primers and 30ng of DNA per sample. The PCR cycling parameters used for all PCR primer pairs were an initial denaturation at 94°C for 5 minutes followed by 36 cycles of denaturation at 94°C for 30s, annealing temperature between 55–58°C for 45s, and elongation at 72°C for 30s with a final extension at 72°C for 5 minutes. PCR products were analyzed on a 1% agarose gel for 30 minutes at 130V. PCR amplicons were enzymatically purified using 1.2ul of ExoSAP-IT (USB) per 6 ul of PCR product. The amplicons were sequenced with BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) and analyzed on a 3730 DNA Analyzer (Applied Biosystems) in the Nucleic Acid and Protein Core Facility (CHOP). Sequence was analyzed using Sequencher v4.9 software (Gene Codes Corporation, Ann Arbor, MI, USA). The potential pathogenicity for any missense polymorphisms/variants found was predicted by PolyPhen-2 (http://genetics.bwh.harvard.edu/pph2/) and SIFT (http://sift.jcvi.org/). All variants/mutations were confirmed by re-amplification and bidirectional sequencing of a new genomic DNA sample from the same patient. The mutations were then assessed in a control population and if DNA was available confirmed in the parents. Copy number was determined using the peak quantification function for mosaicsim in Mutation Surveyor v3.3 (Soft Genetics, LLC. Pennsylvania, USA) following the user manual and a previously described method [Minton et. al., 2011]. The ratio of *STRC* to p*STRC* peak heights for each divergent base pair was then generated using these raw values. A batch control with normal

copy number was used as the reference file as well as standard 1 and a proband with a homozygous *STRC* deletion (HLS433P) was used as standard 2.

RESULTS

STRC CNV deletion analysis

DNA samples from a total of 659 probands with NBSNHI ranging in severity from mild to profound who do not carry mutations in *GJB2* and DNA from 1,910 controls were run on either the Illumina 550k SNP genotyping array (n=509) or the Illumina 610k SNP genotyping array (n=150 probands; 1,910 controls) and normalized to the intersecting probes on the Illumina 550k SNP genotyping array before using the PennCNV detection algorithm to identify CNVs [Wang et al., 2007]. Further analysis by Sanger sequencing was undertaken in 80/659 probands with array data and 10 additional *GJB2* mutation negative probands without array data (Table I). The combined CNV deletion detection methods described here identified 7 homozygous and 10 heterozygous deletion CNVs (an additional STRC heterozygous deletion (family 4883) was identified by Affymetrix homozygosity mapping (see below) and therefore was not included in our totals identified by Illumina SNP arrays) encompassing the *STRC* gene in the *GJB2* mutation negative probands (Table II & Supplemental Table II) and 10/1,910 heterozygous deletions in the control cohort.

CNV analysis using Illumina SNP genotyping arrays

The PennCNV detection algorithm called 12,403 significant CNVs comprising 6,660 deletion CNVs and 5,743 duplication CNVs. Further analysis was performed on the 6,660 deletion CNVs to find potential candidate genes for AR hearing loss. Four different probands (HLS176P, HLS054P, HLS115P, HLS240P) and 1 sibling with hearing loss (HLS115AB) were found to have a 3 SNP deletion CNV (rs8038068, rs2927071, rs8042868) on Chr15q15.3 that was approximately 25kb in size (Figure 1A and Table II). One patient (HLS176P) had a homozygous deletion CNV while the three additional probands (HLS054P, HLS115P and HLS240P) along with 1 sibling (HLS115AB) were found to share the same 3 SNP deletion in the heterozygous state. Although the PennCNV detection algorithm defined a minimal 25kb deletion CNV consisting of 3 SNPs spanning the *CATSPER2* gene, the next nondeleted SNPs (rs2251844 & rs10438303) were spaced over 77kb both proximally and distally to the CNV detected (Fig 1A). The lack of sufficient SNP coverage in this region left over 150kb of sequence spanning 5 genes, *HISPPD2A, CKMT1B, STRC, CATSPER2, CKMT1A* and their highly homologous pseudocopies in which to define the deletion breakpoints in these four patients (Fig 1A).

To obtain higher probe coverage through this region, 150 patients genotyped using the Illumina 610k SNP Genotyping array were also called using the unfiltered probe set for the Illumina 610k array platform by the CNV Workshop detection algorithm [Gai et al., 2010]. Using this CNV detection method, an additional three patients (HLS433P, HLS493P, HLS576P) and one sibling with hearing loss (HLS493AB) were found to have a homozygous deletion and one patient (HLS461P) was found to have a heterozygous deletion (Fig 1A). This was a 6 SNP deletion (rs10518820, rs2260160, rs8038068, rs2927071, rs2920781, rs8042868) that now included coverage within *STRC*, intron 22. When manually called using Beadstudio, the deletion CNV spanned 11 SNPs and included 5 CNV probes (Fig 1A–D). One patient (HLS433P) was also run on the Illumina Omni SNP genotyping array and the deletion was validated as a homozygous 15 SNP deletion that involves exon 26 (rs2920791) and 4 additional intronic SNPs (rs2447196, rs12050645, rs2260160, rs2467437) within *STRC* (Fig 1B–D).

The 1,910 controls were all run on the Illumina 610k SNP Genotyping array and when normalized to the intersecting Illumina 550k SNP array probes and called by PennCNV, no deletion CNVs were detected within Chr15q15.3. However, upon further investigation of this region, the unfiltered Illumina 610k SNP array probes uncovered 10/1,910 controls with a heterozygous deletion CNV of this region (0.52%).

The 4 heterozygous CNV deletions found using the array platforms (HLS054P, HLS115P, HLS240P, HLS461P) were each found to have another variant on the presumed *STRC* trans allele upon further analysis by Sanger sequencing (Table II). HLS054P was found to have a heterozygous variant c.2667G>C (p.Q889H) in exon 8 on the presumed trans allele that was predicted by *in silico* analysis using SIFT and PolyPhen to affect protein function and possibly to be damaging to the protein (Table II). This variant was not seen in 85 ethnically matched controls screened. HLS240P, HLS115P and 1 sibling (HLS115AB) were found by Sanger sequencing to have an interstitial heterozygous deletion extending from exon 19 at least through exon 28 in addition to the heterozygous deletion CNV uncovered by PennCNV (Table II). The distal breakpoints of this interstitial deletion were not defined, as divergent base pairs between the *STRC* gene and pseudogene after exon 28 do not exist. The HLS115 parents were confirmed to each carry a heterozygous deletion by sequencing, the maternal allele had the larger Chr15q15.3 deletion whose breakpoints were not defined and the paternal allele had the smaller interstitial heterozygous deletion that begins upstream of exon 19 and extends through exon 28 but whose distal breakpoint has not been defined. The parents of HLS240P were each found to carry a heterozygous deletion of *STRC*, the paternal allele carried the larger Chr15q15.3 heterozygous deletion and the maternal allele carried the smaller interstitial heterozygous deletion. HLS461P was found to have two variants on the presumed trans allele. One is a novel heterozygous variant c.1873C>T (p.R625C) in exon 4 (Table II). This variant was found in the heterozygous state in 4/91 controls, but was not seen homozygously. SIFT and PolyPhen predicted this variant to be deleterious. The other variant, c.3898A>G (p.H1298R) in exon 19 was confirmed using a unique exon 19 primer set. This variant was found in the heterozygous state in 22/93 controls screened, but was not seen homozygously. Without parental testing the causality of these variants remains unclear, however, because c.3898A>G (p.H1298R) was observed at such a high frequency in controls and the minor allele frequency as determined by the 1000 Genomes project (rs2920780; 0.274) was also at a high frequency, this variant is likely benign.

Using the Illumina SNP Genotyping array data, 4 patients with heterozygous deletions, each found to have another potentially causal variant in trans by subsequent Sanger sequencing, and 4 patients with a homozygous deletion were identified among 659 *GJB2* mutation negative NBSNHI patients (Fig 1 and Table I).

CNV confirmation using Sanger Sequencing

The functional and pseudo copies of *STRC* have genomic sequence that is 98.9% identical and coding sequence that is 99.6% homologous as determined by multiple alignment tools, BLAST (http://blast.ncbi.nlm.nih.gov/Blast, NCBI36) and BLAT (http://genome.ucsc.edu/BLAT, hg18). We identified a total of 56 divergent base pairs between the two gene copies spanning exons 16–28 and roughly 80 base pairs of adjacent intronic sequence contained within the amplimers for these exons (See Supporting Information online Supplemental eTable I). We quantified the area under the peaks for each of the 56 divergent base pairs in a cohort of 90 patients and 14 siblings using Mutation Surveyor v3.3 (Soft Genetics LLC) (Figs 2, 3 and Supporting Information online Supplemental eTable I). This cohort included the 8/659 Chr15q15.3 CNV positive patients called by at least one of the CNV calling algorithms to train our sequencing CNV detection method, 72/659 patients from the array cohort with varying degrees of hearing loss that were not detected by either CNV calling algorithm and 10 patients without array data (Table I). In

the training set, 8/8 previously identified array CNV deletions and 2 siblings with hearing loss were confirmed to have a quantifiable peak height ratio at all of the 56 divergent base pairs used in the analysis that correlated with their array data (Fig 2 and Supporting Information online Supplemental eTable I). Patients with twocopies of each gene had relatively equal peak height ratios between the two gene copies, heterozygous deletion patients had peak height ratios roughly equal to half the intensity of *STRC* compared to the normal copy number sequence and homozygous deletion patients had a complete loss of *STRC* (Fig 2A–E). Using the peak height ratios established for these divergent base pairs using the array based deletion CNVs, the remaining 82 patients were analyzed. Three probands (HLS152P, HLS293P, HLS305P) and one sibling with hearing loss (HLS305AS) were found to have homozygous deletion CNVs spanning exons 16–28 (c.3540-?_5125+?del) that were not detected by the PennCNV algorithm (Table II and Fig 3). Analysis of the array data retrospectively confirmed the 3 SNP homozygous deletion of Chr15q15.3 in 3/3 patients that were missed by the PennCNV algorithm. The parental samples (HLS152M &HLS152F) available for testing were confirmed as heterozygous carriers of the Chr15q15.3 deletion that extends into *STRC* both by sequencing and array-based detection. Five heterozygous deletions, 4 extending from exons 16–28 (c.3540-?_5125+?del) and another smaller deletion extending from exons 16–24 (c.3540-?_4701+?del) (HLS234P, HLS518P, HLS277P, HLS364P and HLS421P respectively), were also uncovered through sequence-based analysis (Table II, Fig 3, and Supporting Information online Supplemental eTable II). HLS234P was found to have a novel heterozygous missense variant c.326T>A (p.M109K) in exon 2 that was not seen in 75 controls screened and was predicted by SIFT and PolyPhen to be damaging (Table II, and Fig 3). HLS234P did not have conclusive Illumina 550k SNP array data for the large Chr15q15.3 heterozygous deletion. HLS518P was found to have a novel heterozygous missense variant c.1021C>T (p.R341C) in exon 4 that was not seen in 47 controls and was predicted by SIFT and Polyphen to affect protein function and probably be damaging (Table II, and Fig 3). The large heterozygous deletion was maternally inherited while the missense variant in exon 4 was paternally inherited, although not confirmed to definitively lie within the *STRC* copy, the peak height ratio is suggestive of a 2:1 *pSTRC* to *STRC* ratio that correlates to a deletion of one allele and a mutation on the presumed trans allele of the *STRC* copy. Retrospectively, array analysis for this patient was confirmatory. HLS421P was found to have the same variant on the trans allele of exon 19, c.3898A>G (H1298R), as HLS461P that has been confirmed as being on the *STRC* functional copy, but is likely benign (Supporting Information online Supplemental eTable II and Fig 3). This patient did not have Illumina SNP array data available for analysis. Two patients (HLS277P, HLS364P) found to have heterozygous deletions spanning exons 16–28 and Illumina SNP array data consistent with a heterozygous Chr15q15.3 deletion were not found to have any other variants (Supporting Information online Supplemental eTable II and Fig 3). Notably these patients had a moderately severe to profound hearing loss, while all of the other deletion CNVs found in the cohort had a mild to moderate hearing loss, indicating that the heterozygous deletions in these patients may not be the sole cause of their hearing loss.

In our cohort of 669 probands, the combined CNV detection methods uncovered 7 homozygous *STRC* deletion CNVs and 9 heterozygous deletion CNVs (Table II and Supporting Information online Supplemental eTable II). 6/9 patients with a heterozygous deletion were confirmed to have another variant, however in only four of these were both parents available to allow for confirmation that the second variant was on the trans allele supporting a potential etiological contribution (Table II). All deletion CNV patients with biallelic changes were found to have a mild to moderate hearing loss.

STRC sequence variants

After excluding the previously identified 16 patients who were CNV positive, 74 probands and 12 siblings with varying degrees of hearing loss and mixed ethnicities were also screened for point mutations using Sanger sequencing. Five of 74 probands (HLS147P, HLS149P, HLS213P, HLS253P, HLS513P) were found to have a homozygous c.3898A>G (p.H1298R) missense variant (Supporting Information online Supplemental eTable II and Fig 3) with no apparent deletions, however, using the 1000 Genomes data a Yoruban individual was found to be homozygous for this variant and thus may not contribute to the molecular etiology of these patients. Parental samples were not available for confirmation in these patients. Two of 74 probands (HLS180P and HLS440P) were found to have compound heterozygous missense variants. HLS180P was found to have a heterozygous missense variant c.2640 G>T (p.E880D) in exon 8 [heterozygous in 1/85 controls with no homozygotes while SIFT and PolyPhen predict this variant to likely be tolerated and possibly/probably damaging respectively] and a second heterozygous missense variant c. 3898A>G (p.H1298R) in exon 19 described previously (Supporting Information online Supplemental eTable II & Fig 3). Parents were not available for testing. HLS440P had three heterozygous missense variants including c.3436G>A (p.D1146N), [not seen in 89 controls and SIFT and PolyPhen predict this variant to likely be tolerated and possibly/probably damaging respectively] c.3898A>G (p.H1298R) in exon 19 and c.4433 C>T (p.T1478I) [not seen in 74 controls and both SIFT and PolyPhen predict this variant to likely be benign] which was verified with unique *STRC* primers for exon 23 (Table II, Supplemental Table II & Figure 3). As no parents were available for testing, interpretation of these variants remains unclear. One patient (HLS222P) was found to have a novel presumably homozygous deletion of 11 bps in exon 6 (c.2303_2313+1del) that is predicted to result in the disruption of a splice site [not seen in 41 controls] (Table II and Fig 3). It is unknown in which gene copy (*STRC* or *pseudoSTRC*) this variant is found, but the peak height ratios suggest that 2 copies are affected. No parental samples were available for testing, although it is known that the parents are consanguineous and of Middle Eastern decent.

A large Ashkenazi Jewish family (3883) with six siblings having a mild hearing loss and four siblings without a hearing loss born to two parents without a hearing loss and no known consanguinity was screened after obtaining mapping information from the Affymetrix 250k SNP array that indicated a block of homozygosity on Chr15q15.3 was segregating with the siblings who had a mild hearing loss. After sequencing DNA obtained from this family, a homozygous missense variant c.4171 C>G (p.R1391G) was found in all 6 siblings with a mild hearing loss who share the block of homozygosity (Table II and Fig 3). The parents were confirmed to be heterozygous carriers of the missense variant and this variant was verified to be on the *STRC* copy using unique primer sets for exon 21. This variant was not seen in the homozygous state in 91 control samples screened but was seen in 1/91 controls heterozygously. This variant has not been previously reported. Another unrelated Ashkenazi Jewish family (4883) with two siblings sharing the same haplotype block and a mild to moderate hearing loss were found to be compound heterozygotes for this variant and a large gene deletion (Table II). SIFT/PolyPhen predict this variant c.4171 C>G (p.R1391G) to be benign/probably damaging respectively.

Screening for point mutations identified 10/74 probands with varying degrees of hearing loss with biallelic variants in *STRC*. Six probands were found to be homozygous or compound heterozygous for c.3898A>G (p.H1298R) in exon 19 (Supporting Information online Supplemental eTable II). This variant may be benign as the 1000 Genomes data report a Yoruban individual with this homozygous variant, although no clinical information is available to rule out the possibility of a mild hearing loss in this individual. After removing these individuals and stratifying the cohort by severity of hearing loss, 2/36 probands with a mild to moderate hearing loss were found to have point mutations in *STRC*

as a possible etiologic contributor to their hearing loss; while 2/38 probands with a moderately severe to profound hearing loss were found to have potential *STRC* causative mutations that have not been previously reported in the 1000 Genomes data.

The combined array and sequencing data identified 7 patients with homozygous gene deletions, seven patients with heterozygous gene deletions and potentially causative variants on their trans allele and one patient with a homozygous missense variant among probands with mild to moderate hearing loss (Table II). Notably there were no large *STRC* deletions with another variant in trans found among the probands with moderately severe to profound hearing loss [HLS277P, HLS364P] (Supplemental Table II). However, one proband was found to be compound heterozygous for potential *STRC* mutations and one proband was homozygous for an 11 bp deletion that may contribute etiologically among probands with moderately severe to profound hearing loss (Table II).

DISCUSSION

We analyzed a total of 669 probands with NBSNHI ranging in severity from a mild to profound hearing loss for CNVs using three different Illumina genome-wide SNP genotyping array platforms and Sanger sequencing. Collectively, the different CNV calling methods identified 17 *STRC* deletions, seven homozygous and nine heterozygous deletions among the 659 probands with both array genotyping information and sequencing data and one heterozygous deletion in a patient with sequencing data, but without array genotyping information (Figs 1, 2, Table II, and Supporting Information online Supplemental eTable II). After sequencing for point mutations, 6/9 probands with heterozygous deletion CNVs were found to have a novel variant; four confirmed on the trans allele and two presumed to be on the trans allele but in a region that we are unable to differentiate the *STRC* gene form the pseudogene (Table II and Fig 3). In total, 17/669 (~2.5%) probands were found to have a *STRC* deletion by at least one CNV detection method that may contribute to their hearing loss. However, three probands were not found to carry another novel variant in addition to their heterozygous CNV deletion indicating that for these three probands *STRC* may not be the only contributing factor for these probands molecular etiology (Supporting Information online Supplemental eTable II). All of the probands with bilallic *STRC* deletions [n=9] and those probands with a deletion on one allele and a putative point mutation on the other allele [n=4] have a mild to moderate hearing loss (Table II). The CNV cohort was subsequently stratified into subgroups based on severity of hearing loss to determine the prevalence of *STRC* CNV deletions in the mild to moderate hearing impaired population. Of the total CNV cohort 254/669 probands were classified as having a mild to moderate hearing loss, 265/669 probands had a moderately severe to profound hearing loss and 150/669 were not classified by severity (Table I). In this *GJB2* mutation negative cohort the prevalence of *STRC* deletions in the mild to moderate hearing loss population was ~5.5% (14/254).

A subset of the CNV cohort (n=80) and 10 additional *GJB2* mutation negative probands without array genotyping information were screened for point mutations in *STRC*. This subset included 50 unrelated probands (plus 14 siblings) with mild to moderate hearing loss and 40 probands with a moderately severe to profound hearing loss. After removing the previously identified probands with *STRC* deletion CNVsdescribed above, 4/74 probands (~5.4%) were found to have potentially causative novel variants in the *STRC* gene (Table II). Combining both data sets together (probands with CNV deletions and probands with point mutations), although complicated by different denominators, *STRC* could etiologically explain up to 6.3% of all biallelic *GJB2* mutation negative mild to moderate hearing impaired probands. *STRC* was screened inb 10 *GJB2* mutation negative AR hearing loss families [two or more siblings with hearing loss born to normal hearing parents] and found to be the putative cause of hearing loss in five (50%). While the numbers of AR hearing

impaired families studied here are small these findings suggest that in proven recessive, *GJB2* mutation negative families with NBSNHI mutations in *STRC* may be an important etiologic contributor.

Due to the segmental duplication within Chr15q15.3 encompassing the *STRC* gene, the Illumina 550k SNP genotyping array did not have sufficient SNP coverage across this region (Fig 1A). For this reason, it is difficult to detect this deletion using current clinical CNV calling algorithms and manual analysis of this region will likely be more effective. When the 1,910 controls run on the Illumina 610k array platform were normalized to the Illumina 550k array platform probes, the PennCNV algorithm did not detect any CNVs in this region. However, if all probes were left in and analyzed using BeadStudio, 12/1,910 controls were found to have CNVs in this region, 10 heterozygous deletions (0.52%) and 2 duplications (0.10%). From these data the expected incidence of homozygously deleted *STRC* related hearing loss able to be detected by array would be about 1 in 145,000. This is lower than the previously reported frequency for this genomic region where the heterozygous deletion carrier frequency was determined to be ~1.6% and the duplication frequency was found to be ~0.7% with an extrapolated incidence of about 1 in 16,000 for *STRC* specific hearing loss [Knijnenburg et al., 2009]. The reason for this discrepancy is likely due to the use of different detection methodologies, different array platforms and different CNV calling algorithms used to identify these CNVs. Our data support the previous findings that this region is a hotspot for CNV formation mediated by nonallelic homologous recombination (NAHR) [Knijnenburg et al., 2009; Zhang et al., 2007].

Many divergent base pairs exist between the *STRC* gene and the *STRC* pseudo gene including the introduction of a stop codon in exon 20 (c.4057 C>T, p.Q1353X) in the pseudogene (the critical change that renders this copy a pseudogene) [Verpy et al., 2001]. In this study we identified a total of 56 divergent base pairs between the two gene copies in the 3′ region of the gene (intron 15 through exon 28) including the c.4057 C>T, p.Q1353X variant (Supporting Information online Supplemental eTable I and Fig 2). In a recent report, the c.4057 C>T, p.Q1353X divergent sequence variant was reported as a variant of unknown significance (VUS) in a family with profound hearing loss [Shearer et al., 2010]. This would not likely contribute to this case or the hearing loss population, as it is consistently present in any Chr15q15.3 CNV negative population regardless of a hearing loss. The c.3898A>G (p.H1298R) variant found in this study is likely benign. The observed frequency of heterozygosity in controls was high (22/93) and the minor allele frequency as determined by the 1000 Genomes project (0.274) was also at a high frequency. The 1000 Genomes data also indicates that a Yoruban individual was found to carry this variant in the homozygous state, although we do not have clinical information on this individual and therefore cannot rule out the possibility of a mild hearing loss in this individual. Functional analyses of all the variants found in this study are currently underway to clarify the impact of these variants to each individuals hearing loss.

Hearing loss is genetically heterogeneous and typically monogenic. To date, 40 causative AR genes have been identified with 68 loci mapped and many genes yet to be identified [Van Camp and Smith, 2011]. Although the list of causal AR hearing loss genes is quite extensive, *GJB2* has been the only significant etiological contributor to the nonsyndromic hearing loss population to date. Our data suggest that *STRC* is another common cause of nonsyndromic AR NBSNHI in the pediatric population, especially in probands with a mild to moderate hearing loss.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to acknowledge the support and involvement of the families with hearing loss as well as Professor Orly Elpeleg, Shamir Zenvirt, Avraham Shaag, Rachel Dahan, Lital Sheva and Adina Quint from the Monique and Jacques Roboh Department of Genetic Research, The Department of Genetic and Metabolic Diseases, Hadassah, Hebrew University Medical Center, 91120 Jerusalem for all of their efforts in mapping regions of homozygosity in the Ashkenazi population analyzed. We are also grateful to Dr. Matt Deardorff for the use of Mutation Surveyor. This work is supported by NIDCD R33DC008630 & R33DC008630-03S1 (IDK) and CHOP Institutional Development Funds (IDK).

References

Avidan N, Tamary H, Dgany O, Cattan D, Pariente A, Thulliez M, Borot N, Moati L, Barthelme A, Shalmon L, Krasnov T, Ben-Asher E, Olender T, Khen M, Yaniv I, Zaizov R, Shalev H, Delaunay J, Fellous M, Lancet D, Beckmann J. CATSPER2, a human autosomal nonsyndromic male infertility gene. Eur J Hum Genet. 2003; 11(7):497–502. [PubMed: 12825070]

Campbell DA, McHale DP, Brown KA, Moynihan LM, Houseman M, Karbani G, Parry G, Janjua AH, Newton V, al-Gazali L, Marksham AF, Lench NJ, Mueller RF. A new locus for non-syndromal, autosomal recessive, sensorineural hearing loss (DFNB16) maps to human chromosome 15q21-q22. J Med Genet. 1997; 34(12):1015–1017. [PubMed: 9429146]

Conrad DF, Andrews TD, Carter NP, Hurles ME, Pritchard JK. A high-resolution survey of deletion polymorphism in the human genome. Nat Genet. 2006; 38(1):75–81. [PubMed: 16327808]

Feuk L, Carson AR, Scherer SW. Structural variation in the human genome. Nat Rev Genet. 2006a; 7(2):85–97. [PubMed: 16418744]

Feuk L, Marshall CR, Wintle RF, Scherer SW. Structural variants: changing the landscape of chromosomes and design of disease studies. Hum Mol Genet. 2006b; 15(Spec No 1):R57–66. [PubMed: 16651370]

Freeman JL, Perry GH, Feuk L, Redon R, McCarroll SA, Altshuler DM, Aburatani H, Jones KW, Tyler-Smith C, Hurles ME, Carter NP, Scherer SW, Lee C. Copy number variation: new insights in genome diversity. Genome Res. 2006; 16(8):949–961. [PubMed: 16809666]

Frolenkov GI, Belyantseva IA, Friedman TB, Griffith AJ. Genetic insights into the morphogenesis of inner ear hair cells. Nat Rev Genet. 2004; 5(7):489–498. [PubMed: 15211351]

Gai X, Perin JC, Murphy K, O'Hara R, D'Arcy M, Wenocur A, Xie HM, Rappaport EF, Shaikh TH, White PS. CNV Workshop: an integrated platform for high-throughput copy number variation discovery and clinical diagnostics. BMC Bioinformatics. 2010; 11:74. [PubMed: 20132550]

Hilgert N, Smith RJ, Van Camp G. Forty-six genes causing nonsyndromic hearing impairment: which ones should be analyzed in DNA diagnostics? Mutat Res. 2009; 681(2–3):189–196. [PubMed: 18804553]

Knijnenburg J, Oberstein SA, Frei K, Lucas T, Gijsbers AC, Ruivenkamp CA, Tanke HJ, Szuhai K. A homozygous deletion of a normal variation locus in a patient with hearing loss from non-consanguineous parents. J Med Genet. 2009; 46(6):412–417. [PubMed: 19246478]

McCarroll SA, Altshuler DM. Copy-number variation and association studies of human disease. Nat Genet. 2007; 39(7 Suppl):S37–42. [PubMed: 17597780]

Minton JAL, Flanagan SE, Ellard S. Mutation Surveyor: Software for DNA Sequence Analysis. Methods Mol Biol. 2011; 688:143–53. [PubMed: 20938837]

Morton CC, Nance WE. Newborn hearing screening--a silent revolution. N Engl J Med. 2006; 354(20):2151–2164. [PubMed: 16707752]

Nayak GD, Ratnayaka HS, Goodyear RJ, Richardson GP. Development of the hair bundle and mechanotransduction. Int J Dev Biol. 2007; 51(6–7):597–608. [PubMed: 17891720]

Peterson M, Willems P. Non-syndromic, autosomal-recessive deafness. Clinical Genetics. 2006; 69:371–392. [PubMed: 16650073]

Petit C, Levilliers J, Hardelin JP. Molecular genetics of hearing loss. Annu Rev Genet. 2001; 35:589–646. [PubMed: 11700295]

Petit C, Richardson GP. Linking genes underlying deafness to hair-bundle development and function. Nat Neurosci. 2009; 12(6):703–710. [PubMed: 19471269]

Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, Fiegler H, Shapero MH, Carson AR, Chen W, Cho EK, Dallaire S, Freeman JL, Gonzalez JR, Gratacos M, Huang J, Kalaitzopoulos D, Komura D, MacDonald JR, Marshall CR, Mei R, Montgomery L, Nishimura K, Okamura K, Shen F, Somerville MJ, Tchinda J, Valsesia A, Woodwark C, Yang F, Zhang J, Zerjal T, Armengol L, Conrad DF, Estivill X, Tyler-Smith C, Carter NP, Aburatani H, Lee C, Jones KW, Scherer SW, Hurles ME. Global variation in copy number in the human genome. Nature. 2006; 444(7118):444–454. [PubMed: 17122850]

Sebat J, Lakshmi B, Troge J, Alexander J, Young J, Lundin P, Maner S, Massa H, Walker M, Chi M, Navin N, Lucito R, Healy J, Hicks J, Ye K, Reiner A, Gilliam TC, Trask B, Patterson N, Zetterberg A, Wigler M. Large-scale copy number polymorphism in the human genome. Science. 2004; 305(5683):525–528. [PubMed: 15273396]

Shearer AE, DeLuca AP, Hildebrand MS, Taylor KR, Gurrola J 2nd, Scherer S, Scheetz TE, Smith RJ. Comprehensive genetic testing for hereditary hearing loss using massively parallel sequencing. Proc Natl Acad Sci U S A. 2010; 107(49):21104–21109. [PubMed: 21078986]

Van Camp, G.; Smith, RJ. Hereditarty Hearing Loss Homepage. 2011. http://hereditaryhearinglossorg

Verpy E, Leibovici M, Michalski N, Goodyear RJ, Houdon C, Weil D, Richardson GP, Petit C. Stereocilin connects outer hair cell stereocilia to one another and to the tectorial membrane. J Comp Neurol. 2011; 519(2):194–210. [PubMed: 21165971]

Verpy E, Masmoudi S, Zwaenepoel I, Leibovici M, Hutchin TP, Del Castillo I, Nouaille S, Blanchard S, Laine S, Popot JL, Moreno F, Mueller RF, Petit C. Mutations in a new gene encoding a protein of the hair bundle cause non-syndromic deafness at the DFNB16 locus. Nat Genet. 2001; 29(3): 345–349. [PubMed: 11687802]

Verpy E, Weil D, Leibovici M, Goodyear RJ, Hamard G, Houdon C, Lefevre GM, Hardelin JP, Richardson GP, Avan P, Petit C. Stereocilin-deficient mice reveal the origin of cochlear waveform distortions. Nature. 2008; 456(7219):255–258. [PubMed: 18849963]

Villamar M, del Castillo I, Valle N, Romero L, Moreno F. Deafness locus DFNB16 is located on chromosome 15q13-q21 within a 5-cM interval flanked by markers D15S994 and D15S132. Am J Hum Genet. 1999; 64(4):1238–1241. [PubMed: 10090914]

Wang K, Li M, Hadley D, Liu R, Glessner J, Grant SF, Hakonarson H, Bucan M. PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. Genome Res. 2007; 17(11):1665–1674. [PubMed: 17921354]

Wong KK, deLeeuw RJ, Dosanjh NS, Kimm LR, Cheng Z, Horsman DE, MacAulay C, Ng RT, Brown CJ, Eichler EE, Lam WL. A comprehensive analysis of common copy-number variations in the human genome. Am J Hum Genet. 2007; 80(1):91–104. [PubMed: 17160897]

Zhang Y, Malekpour M, Al-Madani N, Kahrizi K, Zanganeh M, Lohr NJ, Mohseni M, Mojahedi F, Daneshi A, Najmabadi H, Smith RJH. Sensorineural deafness and male infertility: a contiguous gene deletion syndrome. J Med Genet. 2007; 44(4):233–240. [PubMed: 17098888]

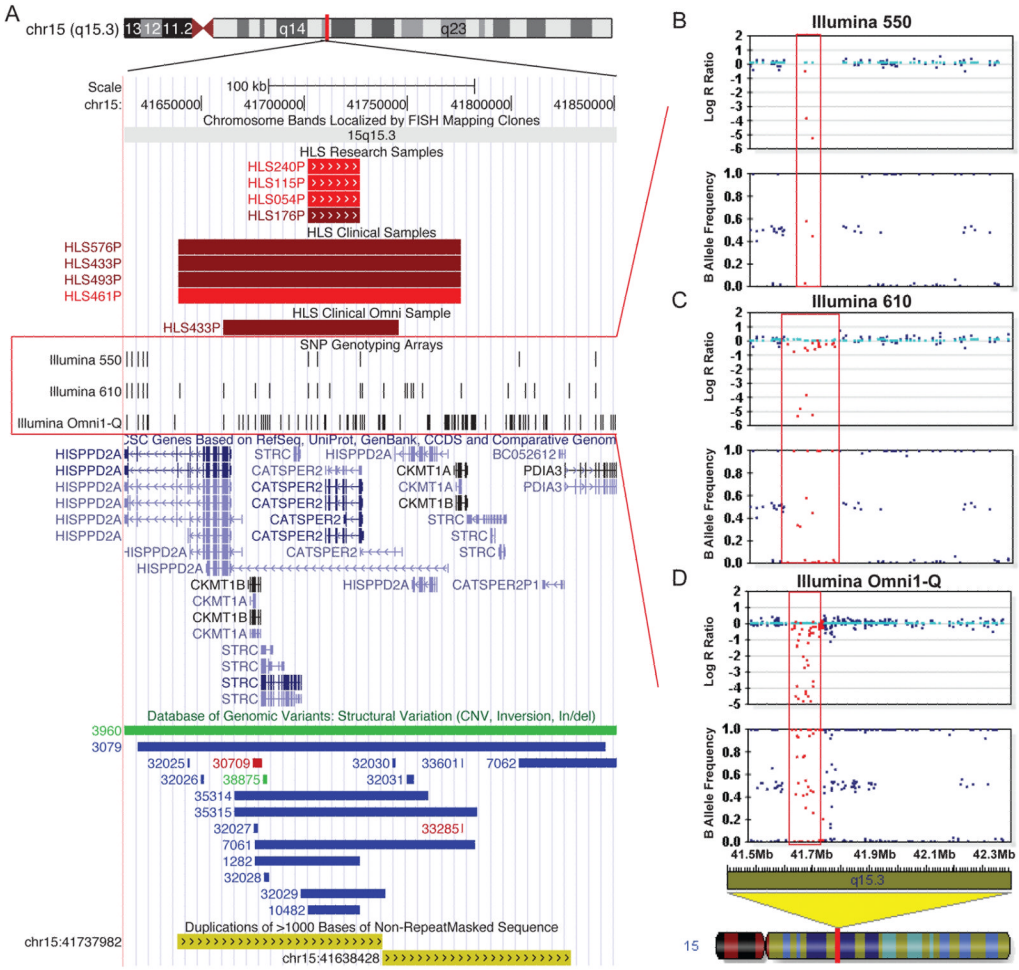

Figure 1. SNP genotyping array data of *GJB2* mutation negative cohort for Chr15q15.3 with successive Illumina array platform resolution for HLS433P

(A) SNP array data for all probands in the array cohort visually depicted in the UCSC Genome Browser, Build hg18 with a genomic window of 41.6Mb to 41.85Mb on Chr15q15.3. The light red horizontal bars represent the heterozygous deletions and the dark red bars represent the homozygous deletions separated by array platform and calling algorithm used to detect the CNV. Research = PennCNV and Clinical = CNV Workshop. The SNP probes for each Illumina SNP Genotyping array platform are depicted as vertical black bars below the deletion CNVs followed by the UCSC gene track, Database of Genomic Variants track and the Segmental Duplication track from the UCSC Genome Browser, Build hg18. (B–D) Increased probe resolution for the Illumina SNP genotyping platform using HLS433P genotyping data visualized in Beadstudio. (B) Illumina 610k SNP array data filtered to the intersecting markers on the Illumina 550k SNP array that was not called by PennCNV. (C) Illumina 610K SNP array data unfiltered reveals a larger deletion window that was detected by CNV Workshop. (D) Illumina Omni1-Quad SNP array confirms a homozygous deletion and further defines the deletion CNV breakpoints for HLS433P.

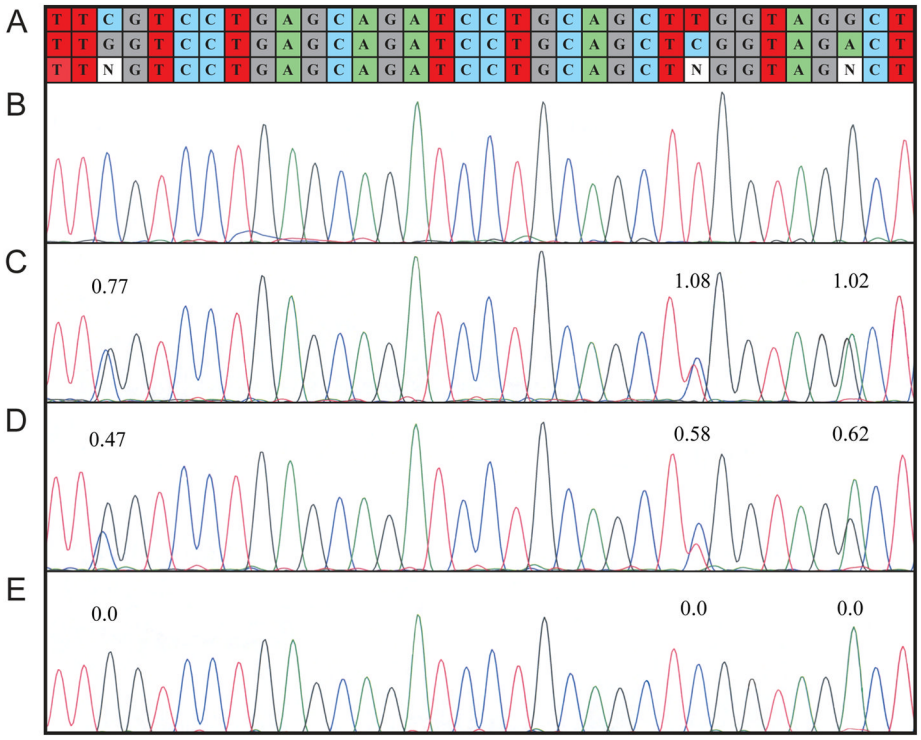

Figure 2. Copy number determination using Sanger sequencing in regions of the gene with divergent base pairs

(A) Alignment of *STRC* and *pSTRC* for a subset of the exon 24 coding sequence displaying 3/56 divergent base pairs used for copy number determination. The white squares with Ns depict the divergent base pair sequences between the two copies used to determine copy number. (B–C) Chromatograms of control sample with normal copy number by array used as reference file and batch control-baseline *STRC* specific peak heights for each base pair. (B) Control sample amplified using unique primers for exon 24. (C) Control sample amplified using generic primers for exon 24 and the peak height ratios for *STRC* to *pSTRC* as quantified using Mutation Surveyor's area under the peak function described in methods. (D) Heterozygous deletion by array (HLS461P) has a decrease in *STRC* sequence to about 50% of control at each of the divergent base pairs. (D) Homozygous deletion by array (HLS433P) has a complete absence of *STRC* sequence.

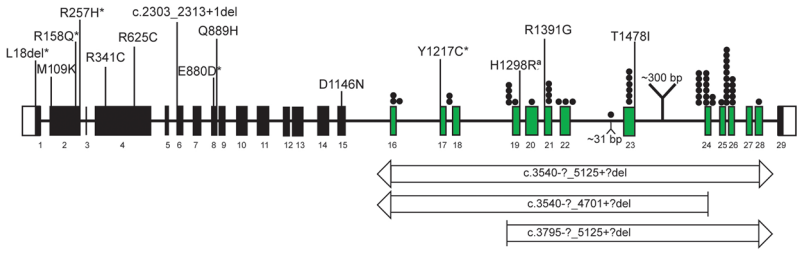

Figure 3. The schematic representation of *STRC* and *pSTRC* sequence variation and mutational sequencing results

Black boxes represent coding exons with 100% homology between the two gene copies while the green boxes represent exons with at least 1 base pair divergence between the two gene copies in or around the coding region. The individual divergent base pairs (n=56) in or around the coding sequence that were used to determine copy number by Sanger sequencing are represented above the corresponding exons by black circles. White boxes are the untranslated regions and horizontal black lines are the introns. Variants not seen in the homozygous state in controls screened here are listed above the corresponding exons where they were found. The white arrows below the gene graphic represent the various sized deletions that were detected by Sanger sequencing. Defined deletion breakpoints are indicated with a vertical black bar and undefined breakpoints are indicated with an open arrowhead. * Indicates a sequence variant that was not seen in a homozygous or compound heterozygous state and so the contribution of these variants to hearing loss cannot be determined. [a] Indicates a variant that was present in a homozygous or compound heterozygous state in one or more probands but was also seen in one Yoruban control from the 1000 genome database and so its potential pathogenicity cannot be clearly defined at this time.

Table I

Study Cohort and Analyses

Total Cohort Studied	**669**
CHOP	395
LMM	272
Hadassah	2
Severity of BSNHI	
Mild to Moderate	254
Severe to Profound	265
Unable to Classify	150
	669
Sanger Sequencing Cohort	**90**
CHOP	88
LMM	0
Hadassah	2
Severity of BSNHI	
Mild to Moderate	50
Severe to Profound	40
Unable to Classify	0
	90
SNP Genotyping Array Cohort	**659**
CHOP	387
LMM	272
Hadassah	0
Mild to Moderate	246
Severe to Profound	261
Unable to Classify	150
	659
With SNP Array Data	
CHOP	80
LMM	0
Hadassah	0

Table II

Homozygous or compound heterozygous *STRC* novel variants in *GJB2* mutation negative cohort

Patient	Allele 1	Allele 2	Control Chromosomes		1000 Genomes MAF	Ethnicity	Parents
CNV Analysis (13/669)							
HLS176P	*STRC* gene deletion [a]	*STRC* gene deletion [a]	10/3820	10/3820	Novel	Caucasian	N/A
HLS433P						Asian	
HLS493P						Caucasian	
HLS576P						Asian/Caucasian	
HLS152P	c.3540-?_5125+?del(exons 16–28)	c.3540-?_5125+?del(exons 16–28)	N.D.			Caucasian	Confirmed
HLS293P							N/A
HLS305P							N/A
HLS240P	*STRC* gene deletion [a]	c.3795-?_5125+?del(exons 19–28)	10/3820	N.D.			Confirmed
HLS115P							Confirmed
HLS054P		c.2667G>C (p.Q889H)	10/3820	0/170			N/A
HLS461P		c.1873C>T(p.R625C)[b]	10/3820	4/182			
HLS518P	c.3540-?_5125+?del(exons 16–28)	c.1021C>T (p.R341C)[b]	N.D.	0/94			Confirmed
HLS234P		c.326T>A (p.M109K)[b]		0/150			N/A
Point Mutation Analysis (4/74)							
Mild to Moderate (2/36)							
3883	c.4171C>G(p.R1391G)	c.4171C>G(p.R1391G)	1/182	1/182	Novel	Ashkenazi	Confirmed
4883		*STRC* gene deletion [c]		10/3820			Confirmed
Moderately Severe to Profound (2/38)							
HLS440P	c.3436G>A(p.D1146N) [b]	c.4433C>T (p.T1478I)	0/178	0/148	Novel	Hispanic	N/A
HLS222P	c.2303_2313+1 del [b]	c.2303_2313+1 del [b]	0/82			Middle Eastern	Consanguineous

N/A, not available for testing.

N.D., not determined

Nomenclature is based on NCBI reference sequence NM_153700.2

[a] Presumed whole gene deletion identified using an Illumina SNP genotyping array platform

[b] Variants not confirmed as *STRC* copy specific

[c] Presumed whole gene deletion found using Affymetrix 250k SNP genotyping array platform