

UNIVERSITÀ DEGLI STUDI DI PADOVA
Corso di Laurea Magistrale in Medicina e Chirurgia

Dipartimento di Salute della Donna e del Bambino
Direttore: Prof. Eugenio Baraldi

U.O.C. Genetica ed Epidemiologia Clinica
Direttore: Prof. Leonardo Salviati

TESI DI LAUREA

**Analisi mutazionale del gene *STRC* mediante
tecniche di sequenziamento di seconda e terza
generazione in una coorte di pazienti affetti da
ipoacusia neurosensoriale**

Relatore: Prof. Matteo Cassina
Correlatore: Dott. Ugo Sorrentino

Laureando: Francesco Prevedello

Anno Accademico: 2022/2023

INDICE

1. RIASSUNTO	5
2. ABSTRACT	7
3. INTRODUZIONE	9
3.1 GENERALITA' SULLE IPOACUSIE	9
3.1.1 CLASSIFICAZIONE DELLE IPOACUSIE	10
3.1.2 VALUTAZIONE CLINICO-STRUMENTALE DELLE IPOACUSIE E PRESIDI TERAPEUTICI ASSOCIATI.....	12
3.2 IPOACUSIE AD EZIOLOGIA GENETICA CON PARTICOLARE RIFERIMENTO ALLE FORME STRC CORRELATE	14
3.2.1 IPOACUSIE GENETICHE NON SINDROMICHE	15
3.2.2 IPOACUSIE GENETICHE SINDROMICHE	23
3.2.3 FORME DI IPOACUSIA GENETICA STRC CORRELATE.....	26
3.3 PANORAMICA SULLE PRINCIPALI TECNOLOGIE DI INDAGINE GENETICA UTILIZZATE NEL CONTESTO DELLE IPOACUSIE	37
3.3.1 SEQUENZIAMENTI DI PRIMA GENERAZIONE	39
3.3.2 SEQUENZIAMENTI DI SECONDA GENERAZIONE: NEXT GENERATION SEQUENCING (NGS).....	41
3.3.3 SEQUENZIAMENTO DI TERZA GENERAZIONE MEDIANTE OXFORD NANOPORE	48
3.3.4 METODICHE DI INDAGINE DI CNV: ARRAY CGH	51
3.3.5 METODICHE DI INDAGINE DI CNV: MLPA	52
3.3.6 ANALISI BIOINFORMATICHE	53
4. SCOPO DELLO STUDIO	57
5. MATERIALI E METODI	59
5.1 SELEZIONE DEL CAMPIONE DELLO STUDIO	59
5.2 ANALISI MOLECOLARE MEDIANTE SEQUENZIAMENTO DI SECONDA GENERAZIONE (PANNELLO MULTIGENICO PER IPOACUSIE)	60
5.2.1 FASE DI PREPARAZIONE DELLE LIBRERIE	60
5.2.2 FASE DI GENERAZIONE DEI CLUSTERS.....	64
5.2.3 FASE DI SEQUENZIAMENTO.....	65
5.2.4 FASE DI ALLINEAMENTO E DI ANALISI DEI DATI	67
5.3 ANALISI MOLECOLARE DI CONFERMA DELLE VARIANTI DI SEQUENZA IN STRC: LR-PCR E SEQUENZIAMENTO	68
5.4 ANALISI BIOINFORMATICA DELLE CNV	70
5.5 ANALISI DI CONFERMA DELLE CNV: MLPA	70
5.6 SEQUENZIAMENTO DI TERZA GENERAZIONE MEDIANTE OXFORD NANOPORE	72
5.6.1 OXFORD NANOPORE: ARRICCHIMENTO MEDIANTE ADAPTIVE SAMPLING O LR-PCR	73
5.7 INTERPRETAZIONE DELLE VARIANTI	74
5.8 COMUNICAZIONE DELL'ESITO	74

5.9 STUDIO DEL CAMPIONE E CONFRONTO CON LA LETTERATURA ..	75
6. RISULTATI	77
6.1 TABELLE RIASSUNTIVE DEI RISULTATI	78
6.2 DESCRIZIONE DEI RISULTATI.....	85
6.2.1 GENERALITA' DEI PAZIENTI	85
6.2.2 SEDE DI CONSULENZA PRE-TEST	86
6.2.3 TECNICHE ANALITICHE UTILIZZATE.....	86
6.2.4 VARIANTI MUTAZIONALI INDIVIDUATE	87
6.2.5 SEGREGAZIONE PARENTALE	92
6.2.6 ETA' D'ESORDIO DELL'IPOACUSIA	93
6.2.7 GRADO DI IPOACUSIA.....	93
6.2.8 PROTESIZZAZIONE	94
6.2.9 DESCRIZIONE DI ULTERIORI PARAMETRI FENOTIPICI	95
6.3 GENOTIPO E FENOTIPO DEI PAZIENTI IN ANALISI.....	96
6.3.1 PAZIENTE 1	96
6.3.2 PAZIENTE 2	97
6.3.3 PAZIENTE 3	98
6.3.4 PAZIENTE 4	100
6.3.5 PAZIENTE 5	102
6.3.6 PAZIENTE 6	103
6.3.7 PAZIENTE 7	107
6.3.8 PAZIENTE 8	108
6.3.9 PAZIENTE 9	109
6.3.10 PAZIENTE 10	115
6.3.11 PAZIENTE 11	117
6.3.12 PAZIENTE 12	117
6.3.13 PAZIENTE 13	121
6.3.14 PAZIENTE 14	123
6.3.15 PAZIENTE 15	125
6.3.16 PAZIENTE 16	127
6.3.17 PAZIENTE 17	128
6.3.18 PAZIENTI 18A E 18B	130
7. DISCUSSIONE E CONCLUSIONI	133
7.1 FENOTIPO AUDIOLOGICO E GENOTIPO.....	133
7.2 PROTOCOLLO DI ANALISI	137
7.2.1 SEQUENZIAMENTO NGS E ANALISI DELLE CNV	138
7.2.2 SEQUENZIAMENTO DI TERZA GENERAZIONE OXFORD NANOPORE	140
7.3 RISVOLTI CLINICI DELLO STUDIO.....	142
7.4 CONCLUSIONI.....	143
8. BIBLIOGRAFIA	145

1. RIASSUNTO

Presupposti dello studio. I disordini genetici rappresentano uno dei principali fattori responsabili di perdita di udito, costituendo la causa di circa l'80% delle ipoacusie ad esordio prelinguale. Sono stati individuati centinaia di geni coinvolti nella patogenesi di ipoacusia ad eziologia genetica: tra questi rientra *STRC*, che si associa ad ipoacusia neurosensoriale isolata di grado lieve-moderato, ad esordio prelinguale, a trasmissione ereditaria autosomica recessiva. *STRC*, che è localizzato nella regione cromosomica 15q15.3, codifica per la proteina stereocilina, responsabile della connessione tra stereociglia adiacenti nelle cellule ciliate esterne dell'orecchio interno. A rendere complessa la diagnosi di ipoacusia associata al locus *STRC* contribuisce la presenza a valle di uno pseudogene, *pSTRC*, ad elevatissima omologia di sequenza (99%): ciò predispone all'insorgenza di riarrangiamenti genici strutturali quali Copy Number Variations (CNV) e conversioni geniche, e comporta notevoli criticità nell'analisi genetica della regione. Per tali ragioni, le tecniche di indagine utilizzate di routine quali il sequenziamento di seconda generazione risultano caratterizzate da efficienza diagnostica incompleta, favorendo l'introduzione e la sperimentazione di tecniche avanzate come il sequenziamento di terza generazione *Oxford Nanopore*.

Scopo dello studio. Lo studio è volto a caratterizzare fenotipicamente e genotipicamente una coorte di pazienti affetti da ipoacusia *STRC*-correlata, con particolare riferimento ai riarrangiamenti complessi del gene *STRC* (conversioni geniche), valutando al contempo l'accuratezza delle tecniche di indagine utilizzate.

Materiali e metodi. È stato selezionato un campione di 19 individui affetti da ipoacusia neurosensoriale recanti mutazioni in entrambi gli alleli del gene *STRC*. Ciascun probando è stato sottoposto ad indagine genetica presso il laboratorio di Genetica Medica ed Epidemiologia Clinica dell'Azienda Ospedale-Università di Padova. In particolare, sono stati eseguiti in ogni paziente il sequenziamento NGS di un pannello di geni associati ad ipoacusia neurosensoriale isolata e la ricerca bioinformatica di CNV con software *Agilent SureCall* a partire dai dati ottenuti con NGS; sono stati quindi effettuati approfondimenti specifici in base alle varianti osservate, tra cui MLPA con Kit apposito alla ricerca di CNV nel locus genico a cui appartiene *STRC*, sequenziamento NGS eseguito con protocollo di

arricchimento *Nextera* su ampliconi ottenuti tramite LR-PCR specifica per il gene *STRC*, e sequenziamento di terza generazione mediante *Oxford Nanopore*.

Risultati. Le caratteristiche fenotipiche dei pazienti studiati sono risultate coerenti con le informazioni presenti in letteratura, in quanto la maggioranza degli individui presentava ipoacusia neurosensoriale isolata bilaterale (100%), di grado lieve o lieve-moderato (58%), ed insorgenza prelinguale (54%).

L'analisi genetica ha invece evidenziato la presenza di 36 varianti alleliche, classificabili in base alla tipologia in piccole delezioni (6%) e sostituzioni nucleotidiche intrageniche (42%), ampie delezioni multiesoniche (44%) e verosimili conversioni geniche (8%).

Le varianti sono state poi classificate in base ai criteri dell'*American College of Medical Genetics* in patogenetiche (75%), verosimilmente patogenetiche (8%) e di significato incerto (17%).

Conclusioni. I risultati ottenuti hanno permesso di consolidare le informazioni disponibili in letteratura riguardo le manifestazioni fenotipiche e genotipiche più frequentemente associate all'ipoacusia *STRC* correlata; inoltre, grazie al sequenziamento di terza generazione, sono stati per la prima volta caratterizzati i punti di rottura di una conversione genica a carico di *STRC*.

Alla luce delle criticità associate all'indagine della regione genomica d'interesse, i risultati ottenuti hanno confermato la validità del protocollo di analisi impiegato. Va comunque sottolineato come si sia reso necessario l'utilizzo di metodologie d'analisi particolarmente avanzate quali il sequenziamento di terza generazione *Oxford Nanopore*, rivelatosi fondamentale nell'indagine delle conversioni geniche. Di conseguenza, non è ancora possibile introdurre il protocollo nella *routine* diagnostica, ma è necessario considerarne l'impiego nel caso gli esami di primo livello non si rivelino sufficienti alla formulazione di una corretta diagnosi.

2. ABSTRACT

Background. Genetic disorders represent one of the main factors responsible for hearing loss, constituting the cause of approximately 80% of prelingual onset hearing loss. Hundreds of genes involved in the pathogenesis of hearing loss of genetic etiology have been identified: among these is *STRC*, which is associated with isolated sensorineural hearing loss of mild-moderate degree, with prelingual onset, and autosomal recessive inheritance. *STRC*, which is located in the 15q15.3 chromosomal region, encodes the protein stereocilin, responsible for the connection between adjacent stereocilia in the outer hair cells of the inner ear. The presence downstream of a pseudogene, p*STRC*, with very high sequence homology (99%) makes the diagnosis of hearing loss associated with the *STRC* locus more complex: this predisposes to the onset of structural gene rearrangements such as Copy Number Variations (CNV) and conversions genes, and poses significant problems in the genetic analysis of the region. For these reasons, routinely used investigation techniques such as NGS second generation sequencing are characterized by incomplete diagnostic efficiency, favoring the introduction and experimentation of advanced techniques such as Oxford Nanopore third generation sequencing.

Objective. The study aims to carry out a phenotypic and genotypic characterization of a cohort of patients with *STRC*-related hearing loss, with particular reference to complex rearrangements of the *STRC* gene (gene conversions), and an evaluation of the accuracy of the investigation techniques used.

Materials and methods. A cohort of 19 individuals with sensorineural hearing loss with mutations in both alleles of the *STRC* gene was selected. Each proband underwent genetic investigation at the laboratory of Medical Genetics and Clinical Epidemiology of the Hospital of Padua. In particular, NGS sequencing of a panel of genes associated with isolated sensorineural hearing loss and bioinformatics research of CNV with *Agilent SureCall* software were performed in each patient starting from data obtained with NGS; specific investigations were then carried out based on the variants observed, including MLPA with a special kit to search for CNV in the *STRC* gene locus, NGS sequencing performed with *Nextera* enrichment protocol on amplicons obtained by LR-PCR specific for the *STRC* gene, and third-generation *Oxford Nanopore* sequencing.

Results. The phenotypic characteristics of the patients studied were consistent with the information present in the literature, as the majority of individuals presented isolated bilateral sensorineural hearing loss (100%), mild or mild-moderate (58%), and prelingual onset (54%).

The genetic analysis identified 36 allelic variants, classified by type in small deletions (6%) and intragenic nucleotide substitutions (42%), large multiexon deletions (44%) and probable gene conversions (8%).

The variants were then classified according to the ACMG criteria into pathogenic (75%), likely pathogenic (8%) and of uncertain significance (17%).

Conclusions. The results obtained allowed us to consolidate the information available in the literature regarding the phenotypic and genotypic manifestations most frequently associated with STRC-related hearing loss; furthermore, thanks to third-generation sequencing, the breakpoints of a gene conversion in STRC were characterized for the first time.

In light of the critical issues associated with the investigation of the genomic region of interest, the results obtained confirmed the validity of the analysis protocol used. However, it should be emphasized that the use of particularly advanced analysis methodologies has become necessary, such as Oxford Nanopore third-generation sequencing, which has proved to be fundamental in the investigation of gene conversions. Consequently, it is not yet possible to introduce the protocol into the diagnostic routine, but it is necessary to consider its employment whenever the first level tests are not sufficient to formulate a correct diagnosis.

3. INTRODUZIONE

3.1 GENERALITA' SULLE IPOACUSIE

L'ipoacusia è definita come l'indebolimento dell'apparato uditivo dovuto a un danno o alla degenerazione di uno o più dei suoi componenti, che si manifesta come un disturbo sensoriale caratterizzato dall'incapacità più o meno grave di percepire suoni.

L'ipoacusia è spesso associata alla sordità, ma è opportuno tener presente che non si tratta della medesima entità nosologica; Il termine sordità, infatti, deve essere applicato solo a quei soggetti che hanno un deficit uditivo così grave da non poter trarre vantaggio da nessuna applicazione acustica. (1)

L'organizzazione mondiale della sanità stima che nel mondo circa 1,5 miliardi di persone (quasi il 20% della popolazione) siano affette da ipoacusia, e che, nel 50% dei casi, questa sia prevenibile mediante adeguate misure di sanità pubblica. (2)

La prevalenza dell'ipoacusia risulta fortemente dipendente dall'età di esordio: la sua presenza alla nascita (ipoacusia congenita) interessa 1,3 bambini su 1000, ed incrementa a 2,7 bambini su 1000 entro i 5 anni di vita, 3,5 su 1000 durante l'adolescenza, fino ad arrivare al 40-50% della popolazione a 75 anni di età. (3)

È stato inoltre osservato come la proporzione di affetti risulti aumentata in determinate categorie di popolazione quali bambini nati prematuri e/o di peso molto basso alla nascita (4) (5) e bambini andati incontro a paralisi cerebrale infantile. (6) (7)

Considerata la frequenza con cui si manifesta l'ipoacusia congenita, ad oggi, nella maggior parte dei paesi sviluppati, viene offerta la possibilità di effettuare uno screening audiologico neonatale nel primo mese di vita, così da identificare i pazienti affetti, inquadrare il disturbo e permettere un intervento precoce. Una volta compiuti i tre anni di vita si raccomanda di ripetere lo screening, che viene effettuato mediante test delle otoemissioni acustiche OAE.

In Italia lo screening audiologico neonatale, pur essendo stato introdotto in molti centri già dai primi anni '00, è stato inserito all'interno dei Livelli essenziali di assistenza solo nel 2017, dopo che fu dimostrato come un intervento precoce e tempestivo risultasse fondamentale per facilitare lo sviluppo della parola e del

linguaggio in bambini non udenti e ipoudenti (8). Lo svolgimento sistematico del test di screening neonatale ha permesso di avere dati molto precisi riguardo l'epidemiologia delle ipoacusie in età infantile, mentre per quanto riguarda le ipoacusie nell'adulto i dati a disposizione risultano meno puntuali.

L'ipoacusia determina una condizione di disabilità, tanto che, se sufficientemente severa, rientra tra i criteri per l'ottenimento di un'indennità mensile di invalidità: il soggetto con deficit uditivo incontra infatti numerose difficoltà nel normale svolgimento delle attività quotidiane. I bambini piccoli (soprattutto nei primi tre anni di vita) non riescono ad ascoltare il parlato e conseguentemente risultano ostacolati nello sviluppo del linguaggio, negli apprendimenti scolastici e nella socialità relazionale; le persone adulte invece non riescono a rispondere al telefono, si trovano in difficoltà nell'ambiente lavorativo e in generale in contesti sociali.

3.1.1 CLASSIFICAZIONE DELLE IPOACUSIE

La classificazione delle ipoacusie viene effettuata sulla base di diversi criteri, quali sede della lesione, esordio, severità, lateralità, frequenza, ed eziologia. (9)

In base alla sede della lesione, si distinguono:

- Ipoacusia trasmissiva, caratterizzata da disfunzioni/anomalie dell'orecchio esterno o della catena ossiculare dell'orecchio medio.
- Ipoacusia neurosensoriale, caratterizzata da disfunzioni/anomalie delle strutture dell'orecchio interno. Può essere associata anche a manifestazioni vestibolari.
- Ipoacusia mista, in cui è presente una combinazione di componente trasmissiva e neurosensoriale.
- Ipoacusia centrale, legata a danno o disfunzione di strutture appartenenti al sistema nervoso periferico o centrale quali l'ottavo nervo cranico, il tronco encefalico, o la corteccia cerebrale.

In base all'età di esordio, si distinguono:

- Ipoacusia congenita, presente dalla nascita
- Ipoacusia pre-linguale, ad insorgenza antecedente all'acquisizione del linguaggio
- Ipoacusia post-linguale, ad insorgenza successiva all'acquisizione del linguaggio
- Ipoacusia dell'adulto, che si manifesta dopo i 18 anni

- Presbiacusia, perdita dell'udito età-correlata, manifestantesi usualmente dopo i 50 anni di età.

Per quanto riguarda la severità dei deficit uditivi, si prende in considerazione la minima intensità in Decibel in corrispondenza della quale l'individuo è in grado di percepire un suono, che fisiologicamente risulta essere tra i -10 e i 11 db. Sulla base di ciò, l'ipoacusia può essere:

- Lieve, se l'intensità è compresa tra i 26 e i 40 db
- Moderata se l'intensità è compresa tra i 41 e i 55 db
- Moderata-grave se l'intensità è compresa tra i 56 e i 70 db
- Grave se l'intensità è compresa tra 71 e i 90 db
- Profonda se l'intensità è maggiore di 90 db

In base alla lateralità, si individuano:

- Ipoacusia unilaterale, quando coinvolge un solo orecchio
- Ipoacusia bilaterale simmetrica, quando coinvolge entrambe le orecchie in maniera simmetrica
- Ipoacusia bilaterale asimmetrica, quando coinvolge entrambe le orecchie in maniera asimmetrica

Le ipoacusie possono poi essere distinte in base ai range di frequenza maggiormente coinvolti dal deficit uditivo, tenendo conto che fisiologicamente l'orecchio umano è in grado di percepire suoni di frequenza compresa tra i 20 e i 20000 Hertz. Esiste una correlazione tra questa classificazione e quella eziologica, in quanto la conoscenza delle frequenze maggiormente interessate, nonché la forma della curva audiografica che si può ottenere a partire dalle stesse, possono contribuire alla corretta definizione della diagnosi eziologica, come verrà discusso con maggior dettaglio nel paragrafo 3.1.2.

- Deficit di percezione delle basse frequenze (<500 Hz).
- Deficit di percezione delle medie frequenze (<501-2000 Hz).
- Deficit di percezione delle alte frequenze (>2000 Hz).
- Deficit di percezione completo.

In merito all'eziologia, lo sviluppo di ipoacusia può rispondere a molteplici e variegata cause, che possono anche sovrapporsi tra loro in un quadro di multifattorialità. Tra queste emergono in primis le cause degenerative, responsabili della stragrande maggioranza dei casi di ipoacusia acquisita in età adulta (10): si parla di presbiacusia, ossia perdita di udito correlata all'età e causata da fenomeni

di tipo degenerativo a carico delle strutture auricolari interne. Le forme ad esordio infantile e adolescenziale, invece, riconoscono come principali agenti eziologici i disordini genetici, responsabili di circa il 50-60% dei casi (11) e che saranno trattati approfonditamente nel paragrafo 3.2. Tra i fattori causativi di ipoacusia si annoverano infine fenomeni infettivi (*Citomegalovirus*, *Rubivirus*), meccanici (danno da rumore, traumi) o dovuti a malattie sistemiche (diabete, patologie autoimmuni).

3.1.2 VALUTAZIONE CLINICO-STRUMENTALE DELLE IPOACUSIE E PRESIDI TERAPEUTICI ASSOCIATI

La valutazione clinico-strumentale dell'ipoacusia è una pratica di pertinenza audiologica, e i professionisti sono formati al fine di adattare la visita alle caratteristiche degli individui testati.

Esistono diverse indagini diagnostiche che possono essere effettuate con l'intento di ottenere informazioni clinicamente rilevanti ai fini della gestione del paziente. Tra queste, si annoverano l'acquisizione di rilievi anamnestici, l'esame obiettivo del condotto uditivo, le misure audiometriche soggettive, le misure audiometriche oggettive, le indagini elettrofisiologiche, la timpanometria, l'imaging.

Rivestono particolare importanza le misure audiometriche soggettive: esse consistono in metodiche d'indagine in cui il paziente viene sottoposto ad erogazioni di suoni puri (cioè a una singola frequenza) ad intensità decrescente. Sulla base delle risposte del soggetto è possibile inferire la sua soglia uditiva, ovvero la minima intensità di un suono a cui corrisponde una percezione. I rilievi effettuati vengono riportati quindi in un grafico, denominato audiogramma, dove ad ogni frequenza in ascissa corrisponde l'intensità minima udibile dal paziente in ordinata.

La forma dell'audiogramma fornisce indicazioni riguardo le frequenze maggiormente coinvolte dall'ipoacusia, e quindi potenzialmente anche riguardo l'eziologia della stessa, come menzionato nel paragrafo 3.1.1:

- Curva audiometrica ascendente (*upsloping*): si tratta di un pattern tipicamente associato all'interessamento delle frequenze gravi. Può essere suggestivo di malattia di Ménière.
- Curva audiometrica a U (*cookie bite*): può essere osservata in caso di interessamento delle frequenze intermedie. Si associa ad ipoacusia da esposizione cronica a rumore o congenita.

- Curva audiometrica discendente (*downsloping*): associata al coinvolgimento delle frequenze acute, questa curva è tipica di ipoacusie ad eziologia virale, senile o da malattia sistemica.
- Curva audiometrica piatta (*flat*): indicativa di coinvolgimento pantonale, è suggestiva di atrofia della stria vascolare a livello cocleare.

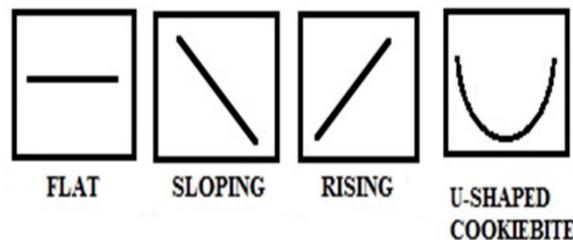


Figura 1. Rappresentazione iconografica delle curve audiometriche descritte in precedenza. (9)

Una volta definito con precisione il quadro clinico mediante l'impiego delle valutazioni clinico-strumentali, risulta necessaria l'impostazione di un trattamento adeguato. È cruciale che questo venga somministrato nelle tempistiche appropriate, specialmente nel contesto di pazienti pediatrici: è stato infatti osservato che interventi ritardati possono determinare deficit irreversibili nello sviluppo del linguaggio, nella lettura e nelle capacità comunicative. (12)

I dispositivi che trovano maggior impiego nella pratica clinica al fine di alleviare la sintomatologia dei soggetti ipoudenti sono le protesi acustiche e gli impianti cocleari.

Le protesi acustiche, o apparecchi acustici, sono apparecchi terapeutici applicabili nei casi di ipoacusia permanente di grado lieve, moderato e moderato-grave, sia di carattere neurosensoriale che trasmissivo. Esse sono assimilabili a un guscio, costituito da un microfono che riceve le onde acustiche, un amplificatore che ne aumenta l'intensità e ne migliora la qualità, e un altoparlante che le trasmette alle strutture dell'orecchio interno. Consentono quindi di stimolare maggiormente le cellule ciliate interne portando l'intensità di stimolazione uditiva sopra alla soglia minima di udibilità propria del paziente.

Nei casi di ipoacusia severa o profonda, le cellule ciliate interne non sono funzionanti, e quindi, per quanto vengano stimulate da eventuali protesi acustiche, non si genera una risposta neuronale da parte dell'orecchio interno e si rivela

necessario l'impianto cocleare, dispositivo elettrofisiologico che bypassa il sistema delle cellule ciliate interne, sostituendo quindi la funzionalità dell'organo di Corti, che risulta compromesso nel paziente.

L'impianto cocleare è costituito da due componenti separate: si distinguono infatti una porzione esterna, responsabile della captazione degli stimoli sonori ambientali, e una porzione interna, installata mediante un intervento chirurgico, preposta all'assorbimento degli impulsi elaborati dal ricevitore e alla trasmissione degli stessi al sistema nervoso centrale, per mezzo del nervo acustico.

3.2 IPOACUSIE AD EZIOLOGIA GENETICA CON PARTICOLARE RIFERIMENTO ALLE FORME STRC CORRELATE

I geni cui finora è stato riconosciuto un ruolo nel corretto sviluppo e funzionamento dell'apparato uditivo umano sono centinaia. Ne consegue che l'insorgenza o la trasmissione di una o più variazioni a carico di tali elementi sia in grado di determinare o predisporre l'insorgenza, precoce o tardiva, di un difetto della funzionalità uditiva. In questi termini, l'ipoacusia è da considerarsi certamente una condizione geneticamente eterogenea, in cui quadri clinici anche molto simili possono essere causati da mutazioni in geni diversi. Per contro, è possibile anche che mutazioni diverse nello stesso gene possano causare fenotipi differenti, sia in termini di caratteristiche cliniche sia di modalità di trasmissione.

La prevalenza delle forme genetiche nel contesto delle ipoacusie risulta strettamente età-dipendente. (13)

- Per quanto riguarda le ipoacusie congenite, circa il 50% riconosce eziologia genetica, mentre la restante metà è legata a condizioni non genetiche di vario tipo, più frequentemente di tipo infettivo, in particolare infezione congenita da *Citomegalovirus*
- Considerando le ipoacusie pre-linguali nel complesso, le forme ad eziologia genetica risultano ancor più frequenti, tanto che si stima rispondano di circa l'80% dei casi.
- La componente genetica risulta significativa anche nelle ipoacusie post-linguali, dell'adolescente e dell'adulto; non sono disponibili tuttavia dati precisi riguardo queste forme ed in generale il contributo delle cause

genetiche è meno chiaro. È comunque opportuno tenere sempre conto della multifattorialità di queste condizioni: è infatti molto frequente che l'ipoacusia insorga in soggetti con mutazioni genetiche predisponenti allo sviluppo del quadro clinico, che sono stati però esposti anche a fattori ambientali favorevoli.

Nel contesto delle ipoacusie ad eziologia genetica, il deficit uditivo si presenta nella gran parte dei casi (70%) come isolato, ovvero in assenza di anomalie malformative a carico dell'orecchio esterno e/o di anomalie morfologiche o funzionali a carico di altri apparati (possono comunque essere presenti manifestazioni associate quali acufeni, disordini dell'equilibrio, decadimento cognitivo, dolore auricolare); nel restante 30% dei casi invece l'ipoacusia risulta associata ad altre anomalie funzionali o malformative, contribuendo pertanto a definire un quadro clinico più complesso di tipo sindromico. (13)

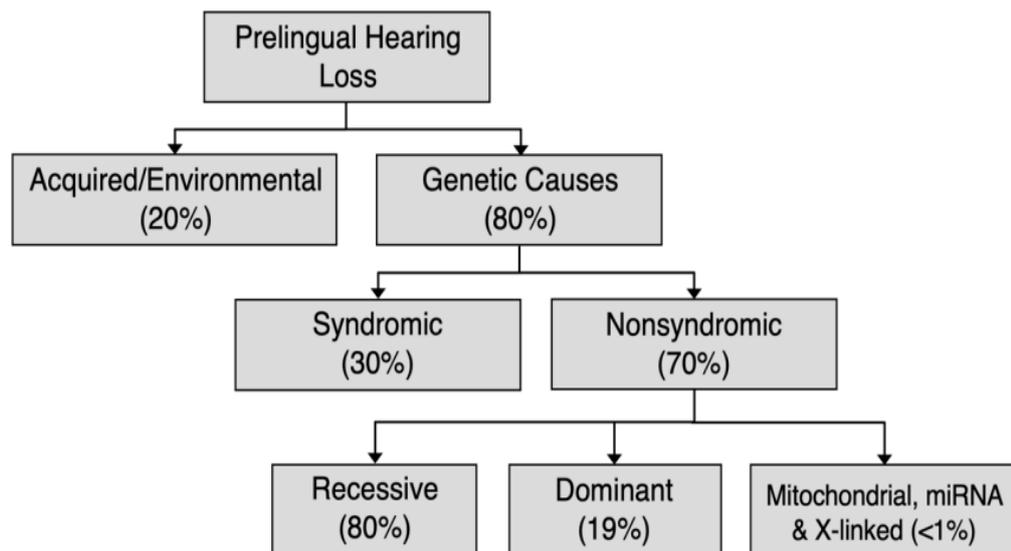


Figura 2. Cause di ipoacusia ad insorgenza prelinguale nei paesi sviluppati. (14)

3.2.1 IPOACUSIE GENETICHE NON SINDROMICHE

Sono stati finora individuati circa 125 geni coinvolti nella patogenesi delle ipoacusie genetiche non sindromiche (15). A tal proposito è utile ricordare che, accanto alla nomenclatura ufficiale HGNC (HUGO Gene Nomenclature Committee), esiste una nomenclatura tradizionale in cui i loci cromosomici associati a sordità genetica non sindromica sono indicati con la sigla DFN, dall'inglese DeaFNness (sordità), seguita da un numero che indica l'ordine cronologico con cui i loci cromosomici sono stati identificati: DFNA indica geni

responsabili di forme a trasmissione autosomica dominante, DFNB geni responsabili di forme a trasmissione autosomica recessiva, DFNBX geni responsabili di forme a trasmissione legata al cromosoma X.

Le forme caratterizzate da trasmissione ereditaria autosomica recessiva (AR) costituiscono circa il 75-80% delle ipoacusie genetiche non sindromiche. Un carattere patologico riconosce una modalità di trasmissione autosomica recessiva quando entrambe le copie del gene, sia quella di derivazione materna che quella di derivazione paterna, sono mutate: il fenotipo pertanto si manifesta in condizione di omozigosi (presenza della stessa mutazione in entrambe le copie del gene) oppure di eterozigosi composta (presenza di due mutazioni diverse nelle due copie del gene). Questa tipologia di trasmissione ereditaria implica la possibilità della nascita di un probando affetto a partire da genitori fenotipicamente sani, che presentano entrambi la mutazione in eterozigosi e trasmettono entrambi alla prole la copia mutata del gene. Il gene più frequentemente mutato in soggetti affetti da ipoacusia ad eziologia genetica è GJB2, che è responsabile di più del 50% delle ipoacusie non sindromiche a trasmissione autosomica recessiva (16). GJB2 codifica per la proteina connessina 26, proteina strutturale coinvolta nella formazione dei canali ionici presenti nella superficie delle cellule ciliate, il cui gradiente elettrico è fondamentale per la trasformazione dell'informazione acustica in impulso nervoso: un'eventuale perdita di funzione della connessina 26 si traduce in un quadro di ipoacusia neurosensoriale di grado severo o profondo.

Le forme di ipoacusia caratterizzate da eredità autosomica dominante (AD), d'altro canto, costituiscono circa il 20% delle ipoacusie genetiche non sindromiche. Si parla di trasmissione ereditaria autosomica dominante quando è sufficiente che una sola delle due copie del gene interessato sia mutata affinché sia presente il carattere patologico, che si manifesterà quindi in eterozigosi. La mutazione causativa può essere ereditata da un genitore affetto oppure insorgere *de novo* nel probando, che sarà quindi il primo nella famiglia a manifestare il quadro clinico.

Le forme caratterizzate da trasmissione ereditaria legata al cromosoma X riconoscono una prevalenza di circa il 2% di tutte le ipoacusie genetiche non sindromiche. Per quanto riguarda questo tipo di eredità, la trasmissione risulta più complessa in quanto i soggetti maschi con corredo cromosomico XY sono emizigoti per i geni presenti nel cromosoma X, ossia presentano una sola copia del gene: ciò significa che, nelle forme X-linked recessive, i maschi presentanti la mutazione

saranno sempre affetti, mentre le femmine, dotate di corredo cromosomico XX, saranno portatrici sane. In caso di forme X-linked dominanti invece sia i probandi maschio che i probandi femmina presenteranno il fenotipo patologico, che risulterà tuttavia clinicamente più grave nei maschi poiché emizigoti e quindi privi di una seconda copia funzionale del gene, a differenza delle femmine.

Esistono inoltre delle forme, ancor più rare e più spesso sindromiche, caratterizzate da trasmissione ereditaria di tipo mitocondriale, riguardante quindi il DNA localizzato nei mitocondri, organelli citoplasmatici deputati al processo di respirazione cellulare e caratterizzati dalla presenza di materiale genetico extranucleare. Poiché il citoplasma dello zigote, al momento della fecondazione è rappresentato quasi esclusivamente da quello dell'ovocita, anche i mitocondri e il DNA in esso contenuto sono di origine esclusivamente materna: mutazioni nel DNA mitocondriale vengono pertanto trasmesse per via matrilineare. Ogni cellula contiene centinaia di mitocondri e al momento della divisione cellulare, questi segregano casualmente, permettendo così la coesistenza di mitocondri a DNA mutato con mitocondri a DNA *wild type* (questa situazione è definita eteroplasmia): solo quando la proporzione di mitocondri con DNA mutato supera un certo limite, definito effetto soglia, o raggiunge la totalità (omoplasmia), si avrà la manifestazione del fenotipo patologico. Per tale motivo, anche all'interno della stessa famiglia potranno essere presenti quadri clinici molto variegati. Esempi notevoli, in termini di prevalenza di popolazione e di meccanismo fisiopatologico, di mutazioni del DNA mitocondriale sono le varianti m.1555°>G e m.1494C>T nel gene MR-RNR1 e le varianti m.7445A>C/T/G nel gene MT-TS1; in particolare, le varianti in MT-RNR1 sono una chiara rappresentazione della multifattorialità eziologica delle ipoacusie, in quanto è stato ampiamente dimostrato che la somministrazione di antibiotici aminoglicosidici a un soggetto portatore della variante patogena è in grado di scatenare l'insorgenza di una forma di ipoacusia grave bilaterale.

La tabella sottostante (Tabella I) illustra i geni più comunemente coinvolti nelle varie modalità di trasmissione dell'ipoacusie genetiche non sindromiche, con riferimenti alle manifestazioni cliniche più frequentemente associate. Si noti come le ipoacusie non sindromiche a trasmissione autosomica recessiva siano tendenzialmente severe, stabili e ad esordio prelinguale, mentre in quelle con modalità di trasmissione dominante l'esordio sia tendenzialmente post-linguale, il

quadro sia più lieve, il decorso progressivo e siano maggiormente interessate le frequenze acute.

Tabella I. Principali geni implicati nella ipoacusia neurosensoriale non sindromica. (14)

<i>Gene</i>	<i>Locus</i>	<i>Trasmissione</i>	<i>Fenotipo dell'ipoacusia</i>
<i>GJB2</i>	DFNB1A	AR	Preverbale, di gravità variabile, stabile, simmetrica
<i>SLC26A4</i>	DFNB4	AR	Preverbale, mista o neurosensoriale, di gravità variabile, stabile o progressiva, fluttuante
<i>MYO15A</i>	DFNB3	AR	Preverbale, da severa a profonda, riportati anche casi di ipoacusia più lieve, postverbale e progressiva
<i>OTOF</i>	DFNB9	AR	Neuropatia uditiva
<i>TMC1</i>	DFNB7/11	AR	Preverbale, da severa a profonda, stabile
<i>TMPRSS3</i>	DFNB8/10	AR	Preverbale oppure postverbale progressiva o stabile con una configurazione audiometrica ski-slope
<i>TECTA</i>	DFNB21	AR	Preverbale da moderata a severa, configurazione audiometrica piatta o downsloping.
<i>CDH23</i>	DFNB12	AR	Preverbale, da severa a profonda, stabile
<i>TRIOBP</i>	DFNB28	AR	Preverbale, da severa a profonda, stabile
<i>TMIE</i>	DFNB6	AR	Preverbale, di grado severo, stabile
<i>PJVK</i>	DFNB59	AR	Preverbale, da severa a profonda, stabile o progressiva; neuropatia uditiva
<i>ESPN</i>	DFNB36	AR	Preverbale, profonda, con o senza interessamento vestibolare
<i>PCDH15</i>	DFNB23	AR	Preverbale, da severa a profonda, stabile

Tabella I. Principali geni implicati nella ipoacusia neurosensoriale non sindromica. (14)

<i>ESRRB</i>	DFNB35	AR	Preverbale, di grado severo/profondo
<i>MYO7A</i>	DFNB2	AR	Preverbale o postverbale, gravità variabile
<i>GJB6</i>	DFNB1B	AR	Preverbale, di gravità variabile, stabile
<i>OTOG</i>	DFNB18B	AR	Preverbale, moderata, stabile, con o senza interessamento vestibolare
<i>OTOGL</i>	DFNB84B	AR	Preverbale, moderata, interessamento delle alte frequenze, stabile
<i>KCNJ10</i>	DFNB4	AR	Fenotipo come SLCA26A4, eredità digenica
<i>GRXCRI</i>	DFNB25	AR	Preverbale, di grado moderato/severo, progressiva
<i>TRIC</i>	DFNB49	AR	Preverbale, profonda, stabile
<i>GIPC3</i>	DFNB15	AR	Preverbale, da moderata a profonda, stabile
<i>CIB2</i>	DFNB48	AR	Preverbale, di grado severo/profondo, stabile
<i>TMHS</i>	DFNB67	AR	Preverbale, di grado severo/profondo
<i>STRC</i>	DFNB16	AR	Preverbale, di grado variabile, nella maggior parte dei casi moderata ed interessa le frequenze acute, può esserci interessamento vestibolare
<i>CLDN14</i>	DFNB29	AR	Preverbale, da severa a profonda, stabile
<i>RDX</i>	DFNB24	AR	Preverbale, da severa a profonda, stabile
<i>MYO6</i>	DFNB37	AR	Preverbale, di grado severo/profondo, possibile interessamento vestibolare
<i>MYO3A</i>	DFNB30	AR	Pre o postverbale, progressiva, interessa prima le alte frequenze

Tabella I. Principali geni implicati nella ipoacusia neurosensoriale non sindromica. (14)

<i>SLC26A5</i>	DFNB61	AR	Preverbale, di grado severo, stabile
<i>WHRN</i>	DFNB31	AR	Neurosensoriale, preverbale
<i>USH1C</i>	DFNB18	AR	Preverbale, da severa a profonda, stabile
<i>COL11A2</i>	DFNB53	AR	Preverbale, di grado severo/profondo, stabile
<i>OTOA</i>	DFNB22	AR	Neurosensoriale, preverbale, grado variabile, stabile
<i>HGF</i>	DFNB39	AR	Preverbale, profonda, configurazione audiometrica downsloping
<i>FOX11</i>	DFNB4	AR	Fenotipo come quello di SCL26A4, eredità digenica
<i>LOXHD1</i>	DFNB77	AR	Preverbale, severa/profonda oppure postverbale, progressiva, da lieve/moderata fino a severo/profondo
<i>MSRB3</i>	DFNB74	AR	Preverbale, profonda
<i>SYNE4</i>	DFNB76	AR	Preverbale, progressiva, interessa le frequenze acute
<i>TPRN</i>	DFNB77	AR	Preverbale, da severa a profonda, configurazione audiometrica down sloping o piatta, stabile
<i>TBC1D24</i>	DFNB86	AR	Preverbale, di grado severo/profondo
<i>COCH</i>	DFNB110	AR	Preverbale, profonda
<i>MYO7A</i>	DFNB2	AR	Preverbale o post verbale, gravità variabile
<i>WFS1</i>	DFNA6/14/38	AD	Preverbale, progressiva, interessa le basse frequenze
<i>KCNQ4</i>	DFNA2	AD	Postverbale (seconda decade), progressiva, interessa le alte frequenze, senza interessamento vestibolare
<i>COCH</i>	DFNA9	AD	Postverbale, progressiva, interessa le alte frequenze, può esserci interessamento vestibolare

Tabella I. Principali geni implicati nella ipoacusia neurosensoriale non sindromica. (14)

			ed una deiecenza del canale semicircolare superiore
GJB2	DFNA3	AD	Neurosensoriale, preverbale, progressiva o stabile
TECTA	DFNA8/12	AD	Preverbale, stabile o progressiva interessa le frequenze medieacute, curva audiometrica ad U
ACTG1	DFNA20/26	AD	Postverbale, progressiva, interessa le frequenze acute
EYA4	DFNA10	AD	Postverbale, progressiva, interessa tutte le frequenze
MYH14	DFNA4	AD	Postverbale, progressiva, interessa tutte le frequenze
SIX1	DFNA23	AD	Preverbale, configurazione audiometrica downsloping
MYO6	DFNA22	AD	Postverbale, progressiva, interessa le alte frequenze
MYO7A	DFNA11	AD	Postverbale (prima decade), progressiva, interessa tutte le frequenze
GSDME	DFNA5	AD	Postverbale, progressiva, interessa le alte frequenze
GJB3	DFNA2B	AD	Postverbale, progressiva, interessa le alte frequenze
POU4F3	DFNA15	AD	Postverbale, progressiva, interessa le alte frequenze
TMC1	DFNA36	AD	Postverbale, progressiva, configurazione audiometrica piatta o lievemente downsloping
COL11A2	DFNA13	AD	Neurosesnsoriale, postverbale (seconda decade), interessa le medie frequenze
CRYM	DFNA40	AD	Postverbale, progressiva
TFCP2L3	DFNA28	AD	Postverbale, progressiva
GJB6	DFNA3	AD	Preverbale, progressiva, interessa le alte frequenze
MYH9	DFNA17	AD	Postverbale, progressiva, interessa le alte frequenze

Tabella I. Principali geni implicati nella ipoacusia neurosensoriale non sindromica. (14)

<i>CCDC50</i>	DFNA44	AD	Postverbale, progressiva, basse e medie frequenze
<i>DIAPH1</i>	DFNA1	AD	Postverbale (prima decade), progressiva, interessa le basse frequenze
<i>CEACAM16</i>	DFNA4B	AD	Postverbale, progressiva, interessa tutte le frequenze
<i>DIABLO</i>	DFNA64	AD	Postverbale, progressiva, può essere associata ad acufeni
<i>DIAPH3</i>	AUNA1	AD	Neuropatia uditiva
<i>GRHL2</i>	DFNA28	AD	Postverbale, interessa tutte le frequenze, stabile o progressiva
<i>OSBPL2</i>	DFNA77	AD	Postverbale, progressiva, interessa le alte frequenze
<i>MIR96</i>	DFNA50	AD	Postverbale (seconda decade), progressiva, interessa tutte le frequenze
<i>P2RX2</i>	DFNA41	AD	Postverbale, progressiva, interessa tutte le frequenze, precoce presbiacusia
<i>SLC17A8</i>	DFNA25	AD	Postverbale, (seconda-sesta decade), progressiva, interessa le alte frequenze
<i>TBC1D24</i>	DFNA65	AD	Postverbale, progressiva, interessa le alte frequenze
<i>TNC</i>	DFNA56	AD	Postverbale, progressiva
<i>POU3F4</i>	DFNX2	X linked	Neurosensoriale o mista, preverbale, progressiva fino a profonda
<i>PRPS1</i>	DFNX1	X linked	Neurosensoriale, postverbale, progressiva, severo/profonda, interessa tutte le frequenze
<i>SMPX</i>	DFNX4	X linked	Neurosensoriale, postverbale, progressiva, da lieve a profonda, tutte le frequenze
<i>COL4A6</i>	DFNX6	X linked	Neurosensoriale, preverbale, progressiva, severo/profonda, interessa tutte le frequenze

Tabella I. Principali geni implicati nella ipoacusia neurosensoriale non sindromica. (14)

<i>MT-RNR1</i>	Mitocondriale	Severità variabile, Penetranza variabile, indotta da aminoglicosidi
<i>MT-TS1</i>	Mitocondriale	Severità variabile, Penetranza variabile
<i>MT-CO1</i>	Mitocondriale	Da severa a profonda, penetranza completa, indotta da aminoglicosidi

3.2.2 IPOACUSIE GENETICHE SINDROMICHE

Tra le ipoacusie ad eziologia genetica, al deficit uditivo si associano in circa il 30% dei casi anche anomalie, funzionali o strutturali, in altri organi o apparati, configurando così un quadro sindromico.

Sono state descritte oltre 400 sindromi associate a deficit uditivo: sebbene ciascuna di esse si distingua dalle altre in base alle modalità di trasmissione ereditaria, alle caratteristiche proprie dell'ipoacusia, alle manifestazioni cliniche associate e ad altri svariati elementi, si possono riconoscere dei tratti comuni alla maggior parte di queste condizioni patologiche, come ad esempio l'esordio preverbale.

Tra le forme a trasmissione autosomica dominante l'entità nosologica maggiormente prevalente è costituita dalla sindrome di Waardenburg, caratterizzata da ipoacusia neurosensoriale, piebaldismo e canizie precoce, seguita dal Disordine dello spettro branchio-oto-renale (BOR), che si associa ad ipoacusia di tipo trasmissivo, neurosensoriale o misto, malformazioni auricolari, anomalie renali e anomalie degli archi branchiali, quali fistole, cisti e schisi.

Tra le forme a trasmissione autosomica recessiva, le due sindromi genetiche di maggior rilievo sono la sindrome di Pendred e la sindrome di Usher, responsabili di quadri di ipoacusia neurosensoriale associata rispettivamente a disfunzione tiroidea e a retinite pigmentosa. Per quanto riguarda la sindrome di Usher va in particolare sottolineato che, potendo essere causata da mutazioni a carico di svariati geni differenti (eterogeneità genetica).

Tra le ipoacusie sindromiche a trasmissione ereditaria X-linked rientrano l'85% delle forme di sindrome di Alport, condizione caratterizzata da sordità neurosensoriale e manifestazioni renali.

La tabella sottostante (Tabella II) illustra i geni più comunemente coinvolti nelle varie modalità di trasmissione dell'ipoacusie genetiche sindromiche, con riferimenti alle manifestazioni cliniche più frequentemente associate.

Tabella II. Principali geni implicati nella ipoacusia neurosensoriale sindromica. (14)

<i>Gene/i</i>	<i>Sindrome</i>	<i>Trasmis- sione</i>	<i>Fenotipo dell'ipoacusia</i>	<i>Elementi sindromici</i>
<i>CHD7</i>	Sindrome CHARGE	AD	Può essere trasmissiva o neurosensoriale, ad esordio e grado variabili	Coloboma, anomalie cardiache, atresia delle coane, ritardo di crescita, malformazioni genito-urinarie, malformazioni auricolari
<i>COL2A1</i> <i>COL11A1</i> <i>COL11A2</i>	Sindrome di Stickler	AD	Può essere trasmissiva o neurosensoriale, ad esordio e grado variabili	Malformazioni auricolari, <i>facies</i> tipica, interessamento oculare, artropatia.
<i>EDN3</i> <i>EDNRB</i> <i>KITLG</i> <i>MITF</i> <i>PAX3</i> <i>SNAI2</i> <i>SOX10</i>	Sindrome di Waardenburg	AD	Neurosensoriale, esordio congenito e grado variabile	Piebaldismo, canuzie precoce
<i>EYA1</i> <i>SIX1</i> <i>SIX5</i>	Disordine dello spettro branchio-oto-renale (BOR)	AD	Può essere trasmissiva, neurosensoriale o mista, ad esordio e grado variabili	Malformazioni auricolari, anomalie renali, anomalie degli archi branchiali (fistole, cisti, schisi)
<i>NF2</i>	Schwannomatosi di tipo 2	AD	Neurosensoriale, esordio attorno ai 20 anni, grado variabile	Sviluppo di neoplasia nervosa quali Schwannomi vestibolari e meningioma

Tabella II. Principali geni implicati nella ipoacusia neurosensoriale sindromica.**(14)**

ADGRV1 USH2A WHRN	Sindrome di Usher di tipo 2	AR	Neurosensoriale, esordio congenito, grado moderato-severo	Retinite pigmentosa, cecità
BTD	Deficit di biotidinasi	AR	neurosensoriale, esordio e grado variabili	convulsioni, disturbi respiratori, ipotonia, rash cutaneo, alopecia, ritardo dello sviluppo
CDH23 CIB2 MYO7A PCDH15 USH1C USH1G	Sindrome di Usher di tipo 1	AR	Neurosensoriale, esordio congenito, grado severo-profondo	Retinite pigmentosa, cecità, interessamento vestibolare
CLRN1	Sindrome di Usher di tipo 3	AR	Neurosensoriale, esordio congenito, grado variabile, andamento progressivo	Retinite pigmentosa, cecità, interessamento vestibolare
COL9A1 COL9A2 COL9A3	Sindrome di Stickler	AR	Può essere trasmittiva o neurosensoriale, ad esordio e grado variabili	Malformazioni auricolari, <i>facies</i> tipica, interessamento oculare, artropatia.
KCNE1 KCNQ1	Sindrome di Jervell e Lange-Nielsen	AR	Neurosensoriale, esordio congenito, grado profondo	Anomalie cardiache, episodi sincopali
PEX1 PEX2 PEX3 PEX5 PEX6 PEX10 PEX11B PEX12 PEX13 PEX14	Sindrome di Zellweger	AR	Neurosensoriale, esordio e grado variabili	Difetti di migrazione neuronale nel cervello, dismorfismi craniofacciali, ipotonia profonda, convulsioni neonatali e disfunzione epatica

Tabella II. Principali geni implicati nella ipoacusia neurosensoriale sindromica. (14)

<i>PEX16</i>				
<i>PEX19</i>				
<i>PEX26</i>				
<i>PEX7</i> <i>PHYH</i>	Malattia di Refsum dell'adulto	AR	Neurosensoriale, esordio variabile, grado severo, andamento progressivo	polineuropatia periferica, degenerazione e atassia cerebellare, retinite pigmentosa
<i>SLC26A4</i>	Sindrome di Pendred	AR	Neurosensoriale, esordio congenito o infantile, grado severo-profondo	Gozzo eutiroideo, alterazioni vestibolari
<i>COL4A5</i>	Sindrome di Alport	X linked	Neurosensoriale, esordio in adolescenza, grado severo, andamento progressivo	Interessamento renale, anomalie oculari
<i>TIMM8A</i>	Sindrome di Mohr-Tranebjaerg	X linked	Neurosensoriale, esordio infantile, grado variabile, andamento progressivo	Distonia, atassia, demenza, anomalie oculari

3.2.3 FORME DI IPOACUSIA GENETICA STRC CORRELATE

Tra la moltitudine di geni deputati alla produzione di proteine coinvolte nel processo di percezione uditiva, uno dei più frequentemente alterati, e pertanto coinvolto nella patogenesi di una cospicua frazione di ipoacusie genetiche, è *STRC*, gene protagonista dello studio oggetto della trattazione.

STRC (numero OMIM *606440) mappa a livello del braccio lungo del cromosoma 15, nella regione 15q15.3; ha una dimensione di circa 19 kilobasi, contiene 29 esoni e codifica per la proteina stereocilina; variazioni loss-of-function in tale gene determinano ipoacusia neurosensoriale secondo una trasmissione ereditaria di tipo autosomico recessiva. (17)

Isolata per la prima volta nel 2001 in corrispondenza dell'orecchio interno del topo (18), e studiata con maggior approfondimento negli anni successivi, si è arrivati a caratterizzare la stereocilina come una voluminosa proteina strutturale

extracellulare costituita da 1809 aminoacidi presente nelle stereocilia delle cellule ciliate esterne, nell'orecchio interno. (19)

Dal punto di vista funzionale, la stereocilina è coinvolta nel mantenimento delle connessioni delle stereocilia: più specificatamente, essa mette in comunicazione stereocilia adiacenti mediante la formazione di strutture connettive orizzontali, e ancora le stereocilia più apicali alla membrana tectoria dell'organo di Corti. La formazione di suddette connessioni determina la corretta disposizione delle stereocilia nello spazio, consentendo pertanto alle cellule ciliate esterne di espletare la loro fisiologica funzione di amplificazione sonora in risposta al movimento attivo delle stereocilia, che con la loro deflessione attivano canali di meccanotrasduzione (MET), preposti alla conversione dell'informazione meccanica in impulso nervoso. (20)

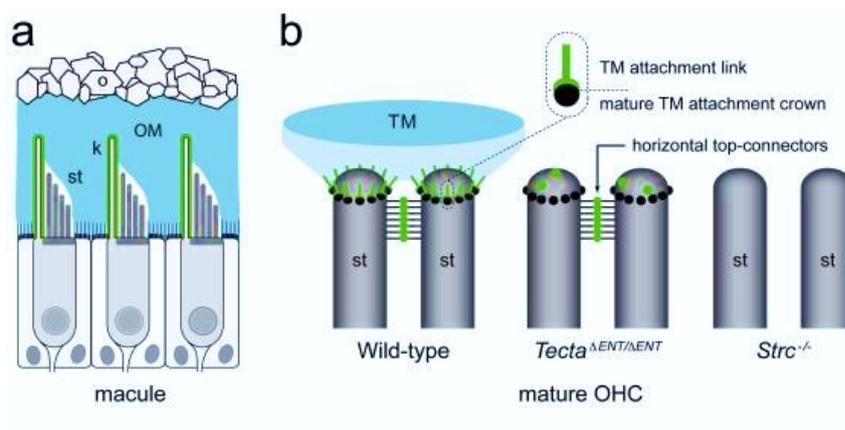


Figura 3. Ruolo della stereocilina nell'orecchio interno. (a) Rappresentazione schematica della disposizione delle cellule ciliate esterne nel contesto dell'orecchio interno. (b) Le maglie di attacco contenenti stereocilina, rappresentate come una serie di nervature a ombrello, attaccano la membrana tectoria a una piattaforma di ancoraggio composta da strutture dense di elettroni (cerchi neri) che formano una corona (corona di attacco della membrana tectoria). Nei topi mutanti *Tecta Δ ENT/ Δ ENT* con membrana tectoria staccata, la corona di attacco è presente, così come i connettori orizzontali superiori, mentre i collegamenti di attacco sono assenti. I topi mutanti *Strc $^{-/-}$* mancano sia la corona di attacco che i connettori orizzontali superiori. (19)

3.2.3.1 EPIDEMIOLOGIA DEI DISTURBI *STRC* CORRELATI

Le variazioni a carico del gene *STRC* rappresentano certamente una delle principali cause di ipoacusia, tuttavia non sono disponibili ad oggi dati epidemiologici definitivi a riguardo. Sono stati però effettuati molteplici studi e metanalisi orientati alla stima della frequenza mutazionale del gene in popolazioni provenienti da svariate aree geografiche.

In accordo con una metanalisi condotta nel 2021 da alcuni ricercatori cinesi sulla base di 37 pubblicazioni, sono state effettuate le seguenti stime. (21)

- La prevalenza globale di varianti bialleliche di *STRC* in popolazioni di pazienti affetti da sordità non *GJB2* correlata corrisponde a 4,08%.
- La prevalenza globale di varianti bialleliche di *STRC* in popolazioni di pazienti affetti da sordità genetica non *GJB2* correlata corrisponde a 9,63%
- La prevalenza globale di varianti bialleliche di *STRC* in popolazioni di pazienti affetti da sordità non *GJB2* di grado lieve-moderato correlata corrisponde a 14,36%
- La prevalenza globale di varianti bialleliche di *STRC* in popolazioni di pazienti affetti da sordità genetica non *GJB2* correlata di grado lieve-moderato corrisponde a 50,71%
- La prevalenza globale di varianti monoalleliche di *STRC* in popolazioni di pazienti affetti da sordità non *GJB2* correlata corrisponde a 4,84%
- La prevalenza globale di varianti monoalleliche di *STRC* nella popolazione generale non affetta da ipoacusia corrisponde a 1,36%

Si evidenzia come, in questo contesto, si siano considerate solo le sordità non associate a mutazioni di *GJB2*, gene codificante per la connessina 26, responsabile di più del 50% delle ipoacusie genetiche non sindromiche a trasmissione autosomica recessiva. Vista la sua elevatissima frequenza di mutazione, infatti, *GJB2* viene analizzato prioritariamente non appena si pone il sospetto di sordità genetica.

È importante poi puntualizzare che le stime elaborate considerano la prevalenza assoluta di determinate condizioni nonostante, sulla base dell'ampia variabilità genetica che sussiste tra diversi popoli ed etnie, questi dati siano fortemente soggetti a variazioni legate all'area geografica di provenienza della popolazione considerata. In particolare, si è osservato che le mutazioni del gene *STRC* sono di gran lunga

più frequenti nell'Asia continentale, raggiungendo picchi di prevalenza in Cina e Corea. (21)

3.2.3.2 PRINCIPALI VARIANTI MUTAZIONALI A CARICO DI *STRC*

Il gene *STRC*, come anticipato precedentemente, è associato allo sviluppo di sordità genetica caratterizzata da trasmissione ereditaria autosomica recessiva. Ciò implica che, affinché si manifesti il fenotipo clinico, è necessario che sussistano alterazioni genetiche su entrambi gli alleli del gene, sia quello di origine materna che quello di origine paterna: può trattarsi di mutazioni uguali, andando quindi a configurare un quadro di omozigosi, ma, più frequentemente, sono presenti anomalie differenti tra loro, individuando quindi una situazione denominata eterozigosi composta. Soggetti presentanti un solo allele mutato (eterozigoti) sono invece portatori sani.

Esistono svariate tipologie di alterazioni genetiche che possono interessare *STRC*: queste includono piccole variazioni di sequenza (sostituzioni, inserzioni, duplicazioni) che interessano singoli nucleotidi (mutazioni puntiformi) o comunque un numero limitato di basi, e riarrangiamenti più ampi e complessi, costituiti principalmente da *copy number variations* (CNV) e conversioni geniche. (22) (23) Nell'ambito delle alterazioni geniche di maggiori dimensioni, le più frequenti coinvolte nella patogenesi delle ipoacusie *STRC*-correlate, rappresentandone in senso assoluto la causa più prevalente, è la presenza di *Copy Number Variations* (CNV), ossia variazioni del numero di copie (24). Esse fanno riferimento alla condizione in cui il numero di copie di uno specifico segmento di DNA, di lunghezza compresa tra 100 bp e 3M bp, varia tra i genomi di diversi individui, individuando pertanto la presenza di delezioni e duplicazioni. Secondo definizioni più recenti, si possono considerare CNV anche variazioni di dimensione minore di 100 bp, purché includano almeno un esone del gene nella sua interezza.

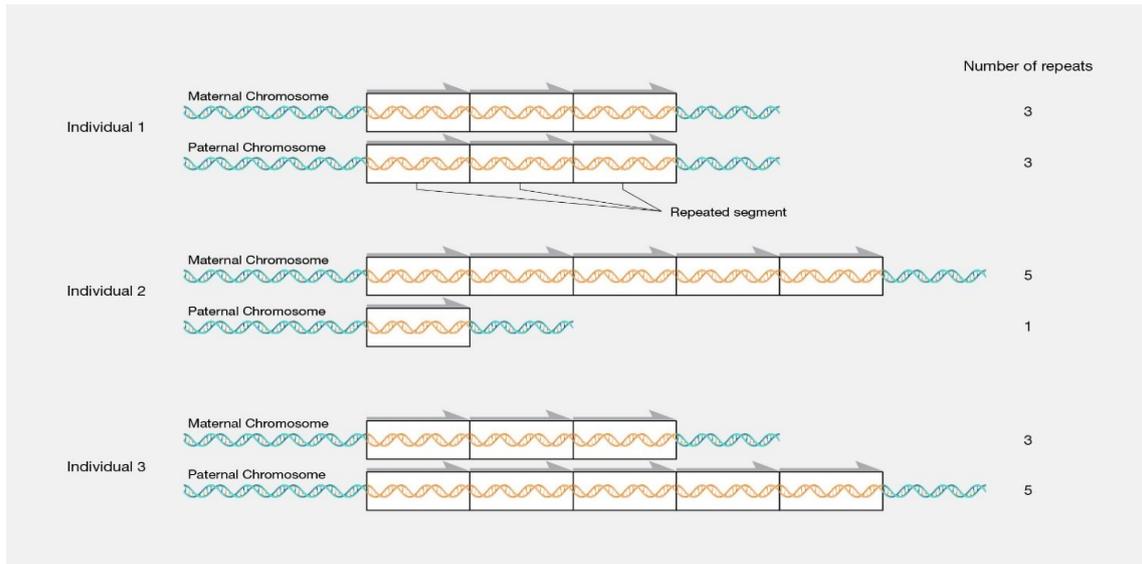


Figura 4. Rappresentazione schematica del fenomeno delle Copy Number Variations (75)

Alla base dello sviluppo di CNV possono esserci errori sia nel processo di ricombinazione genica che nel processo di riparazione del DNA (25). I principali meccanismi proposti sono:

- Ricombinazione omologa non allelica (*Non Allelic Homologous recombination*, NAHR). È una forma di ricombinazione omologa che si verifica tra due sequenze di DNA con alta omologia di sequenza, ma che non sono alleli dello stesso gene; può comportare lo sviluppo di delezioni, duplicazioni, inversioni, o, se si verifica tra sequenze appartenenti a cromosomi diversi, traslocazioni.
- Giunzione non omologa delle estremità (*Non homologous end joining*, NHEJ). Si tratta di un meccanismo di riparazione del DNA utilizzato in caso di rotture a doppio filamento, che consente di riunire le due estremità della rottura in assenza di una sequenza che possa fare da stampo, e può determinare l'insorgenza di CNV.
- Modello FoSTeS (*Fork stalling and Template switching*). Secondo questo modello, è possibile che CNV anche molto complesse, come duplicazioni interrotte da delezioni, siano frutto di fenomeni di arresto della forca replicativa durante il processo di replicazione del DNA, seguiti dall'aggancio del filamento neformato ad una nuova forca a monte o a valle.

- Retrotrasposizione (29). Una considerevole porzione del genoma umano è costituita da sequenze mobili definite trasposoni, in grado di trasciversi autonomamente in un intermedio a RNA per poi riconvertirsi a DNA ed integrarsi nel genoma in corrispondenza di diverse posizioni. Questo meccanismo rientra tra le cause di sviluppo di CNV.

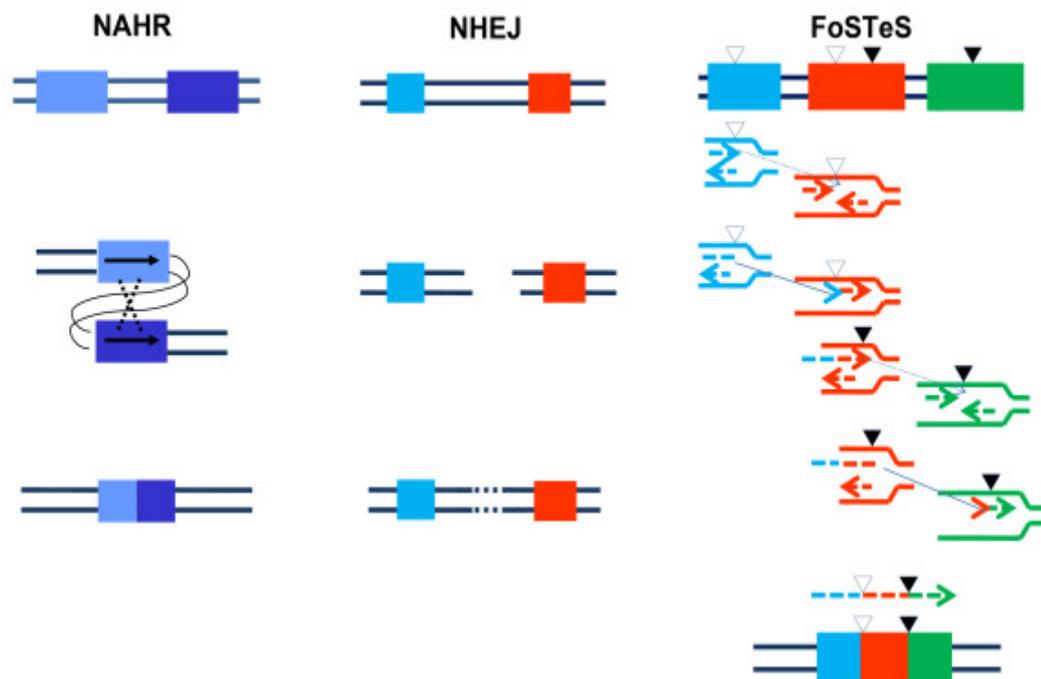


Figura 5. Confronto tra Ricombinazione omologa non allelica, giunzione non omologa delle estremità e Modello FoSTeS, risultanti in delezioni e duplicazioni del genoma. (25)

Lo sviluppo di CNV, in generale, può essere causativo di variazioni fenotipiche varie, sia fisiologiche che patologiche, oppure rappresentare varianti polimorfiche benigne; è stato inoltre individuato un loro ruolo anche in ambito evolutivo (30).

Per quanto riguarda specificatamente il gene *STRC*, invece, esso si trova in una regione del genoma sostanzialmente duplicata, come verrà specificato in maniera più approfondita nel paragrafo successivo, che risulta soggetta a ricorrenti CNV, più spesso esito di processi di ricombinazione omologa non allelica. Si stima addirittura che il 73% delle CNV coinvolte nello sviluppo di ipoacusia siano a carico del gene codificante la stereocilina. In particolare, le CNV maggiormente rappresentate e clinicamente rilevanti, sono rappresentate dalle delezioni. (23)

Un'altra tipologia di riarrangiamenti genici che sono stati osservati a carico di *STRC* sono le conversioni geniche. Nonostante sia presente poca letteratura scientifica

riguardo le conversioni geniche in *STRC*, esse sono state rilevate nella casistica di pazienti presa in considerazione dallo studio oggetto della trattazione, andando così a costituire uno dei maggiori punti di forza dello stesso.

Con conversione genica si intende un fenomeno per il quale si verifica un trasferimento non reciproco e unidirezionale di informazione genetica da una sequenza di DNA ad un'altra (31). Ne esistono due principali tipologie:

- Conversione genica allelica. Essa si verifica quando una sequenza allelica trasferisce la propria informazione genetica all'altra copia allelica dello stesso gene. Si verifica tipicamente durante il processo di ricombinazione omologa, tipico della profase I della meiosi, in seguito a tentativi di riparazione di fenomeni di anomali appaiamenti tra le basi.
- Conversione genica non allelica/ectopica (32). Parliamo di conversione non allelica quando l'informazione genetica viene trasferita da una sequenza di DNA ad un'altra non appartenente allo stesso gene, ma caratterizzata da elevata omologia con la sequenza di partenza. Essa può avvenire in *cis* se si verifica tra geni appartenenti allo stesso cromatide, o in *trans* se coinvolge geni appartenenti a cromatidi fratelli.

Il fenomeno può avvenire nel contesto della meiosi, analogamente alla conversione allelica, ma può anche essere frutto dei meccanismi preposti alla riparazione delle rotture del doppio filamento (DSB) causate dal danno al DNA.

Questo tipo di alterazione genica insorge di frequente nel contesto di geni dei quali esistono pseudogeni, ovvero frammenti di DNA non codificanti caratterizzati da elevata omologia di sequenza con il corrispettivo gene: ciò fa in modo che le eventuali mutazioni presenti nello pseudogene, di norma non rilevanti perché non espresse a livello proteico, vengano introdotte nel gene vero e proprio, che verrà quindi trascritto e tradotto in maniera anomala.

Tra i geni caratterizzati dalla presenza di pseudogeni associati vi è appunto *STRC*, che risulta dunque soggetto a fenomeni di conversione genica. È infatti verosimile che alcune varianti complesse di *STRC*, classificate in letteratura come CNV, siano in realtà conversioni geniche.

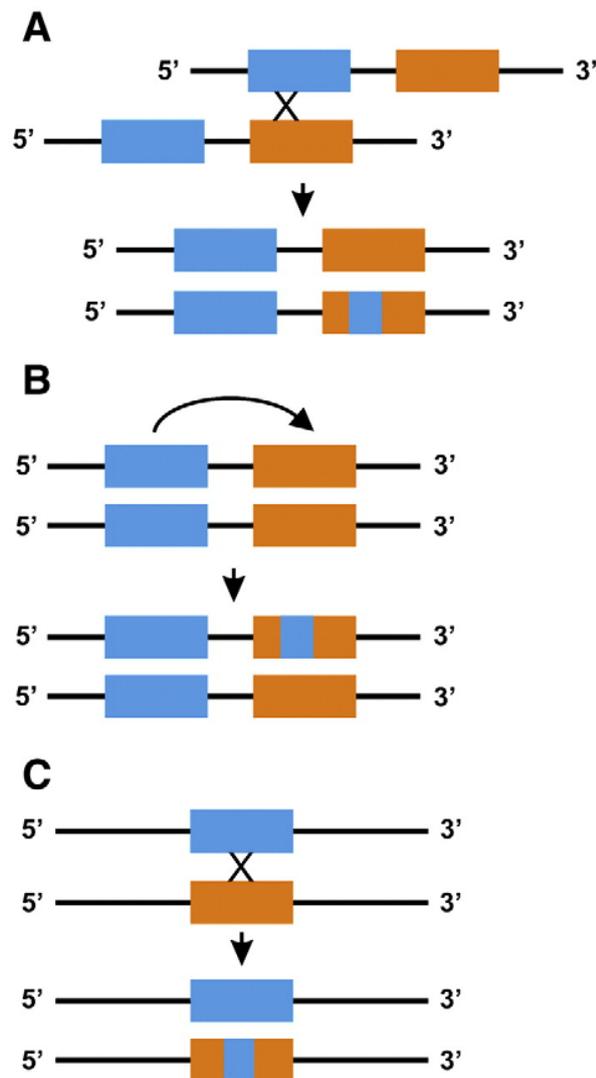


Figura 6. Tipologie di conversione genica. (A) conversione genica non allelica in trans. (B) conversione genica non allelica in cis. (C) conversione genica allelica. (76)

3.2.3.4 MANIFESTAZIONI FENOTIPICHE ASSOCIATE AD ANOMALIE DI *STRC*

L'ipoacusia da mutazione in *STRC* è di tipo neurosensoriale e caratterizzata da ampia variabilità fenotipica, conseguentemente all'eterogeneità mutazionale e ai co-fattori ambientali specifici di ogni paziente. Al netto di ciò, tuttavia, è possibile evidenziare alcune caratteristiche ricorrenti, che si ritrovano nella gran parte dei casi presenti in letteratura (33). Si tratta di un'ipoacusia neurosensoriale ad esordio preverbale, di grado variabile, nella maggior parte dei casi moderato, bilaterale, interessante le frequenze acute (curva audiometrica down-sloping), e non progressiva; possono essere associate manifestazioni vestibolari. (34)

L'ipoacusia da anomalie *STRC* correlate è generalmente isolata, tuttavia può verificarsi un quadro sindromico in caso di ampie delezioni in omozigosi a carico del gene *STRC* che coinvolgano anche l'adiacente gene *CATSPER2*. *CATSPER2* (Cation channel sperm-associated protein 2) è un gene, localizzato a monte di *STRC*, che codifica per una proteina canale di membrana coinvolta nella motilità degli spermatozoi (35). Il fenotipo patologico associato è denominato *Deafness Infertility Syndrome* (DIS – MIM #611102) ed è caratterizzato dall'associazione tra ipoacusia e infertilità maschile. (36)

Si ribadisce come la DIS sia un'entità nosologica associata ad infertilità specificatamente maschile: ne consegue che probandi femmine, pur presentanti un quadro mutazionale teoricamente causativo di DIS, manifesteranno ipoacusia isolata.

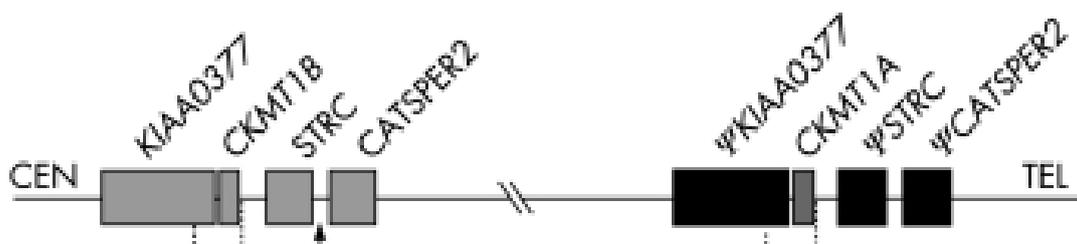


Figura 7. Organizzazione genomica del locus 15q15.3, si possono osservare i geni *STRC* e *CATSPER2* coi rispettivi pseudogeni a valle. Le linee tratteggiate indicano le duplicazioni segmentali che mediano CNV mediante NAHR. (36)

3.2.3.5 CRITICITA' ASSOCIATE ALL'ANALISI MUTAZIONALE DEL GENE *STRC*

Nell'ambito della diagnostica molecolare delle ipoacusie genetiche, una delle maggiori sfide che si riserva di affrontare la genetica medica contemporanea risiede proprio nell'analisi molecolare del gene *STRC*, che risulta particolarmente difficoltosa per svariate ragioni.

Tra le considerevoli criticità presentate dallo studio mutazionale di *STRC*, la principale consiste nella presenza, circa 100 kb a valle, di uno pseudogene *pSTRC*: si tratta di una copia del gene dotata di elevatissima omologia di sequenza, stimata pari al 99.6%, ma priva di funzionalità; non viene infatti né trascritta né tradotta. (17)

Gli pseudogeni sono quindi, in linea generale, delle sequenze caratterizzate da non funzionalità e omologia con un gene. Individuati per la prima volta nel 1977 (37),

ancora oggi non è del tutto chiaro se essi meritino l'etichetta di "DNA spazzatura" oppure rivestano un ruolo biologico, ed eventualmente quale esso sia. Nei primi anni 2000 sembrava che gli pseudogeni potessero essere coinvolti nella regolazione dell'espressione genica dei rispettivi geni omologhi ma il dibattito riguardo la loro funzione è ancora aperto (38).

In ottica classificativa, si individuano tre principali gruppi di pseudogeni, ciascuno dei quali si distingue per origine e caratteristiche.

- Pseudogeni processati o retrotrasposti (39). Si tratta di sequenze ottenute mediante un fenomeno di retrotrasposizione, in cui si verifica la trascrizione inversa di un RNA messaggero in DNA, a cui segue la reintegrazione del prodotto di reazione nel genoma. Si viene così a formare uno pseudogene che ricalca le sequenze esoniche del gene d'origine, ma risulta privo delle sequenze promotoriali e introniche.
- Pseudogeni non processati o duplicati (40). Originatisi in seguito a fenomeni di duplicazione genica, presentano ancor più elevata omologia di sequenza col gene primitivo ricalcandone anche le porzioni introniche; la perdita di funzionalità avviene attraverso l'acquisizione progressiva di mutazioni nel corso dell'evoluzione.
- Pseudogeni disabilitati o unitari (41). Questa categoria non rientra appieno nella definizione di pseudogene in quanto non è presente una copia funzionante nel genoma; si tratta infatti di sequenze divenute inattive nel corso dell'evoluzione in seguito ad un processo di accumulo mutazionale, facilitato da fattori propri della dinamica delle popolazioni, quali la deriva genetica o l'effetto collo di bottiglia (42).

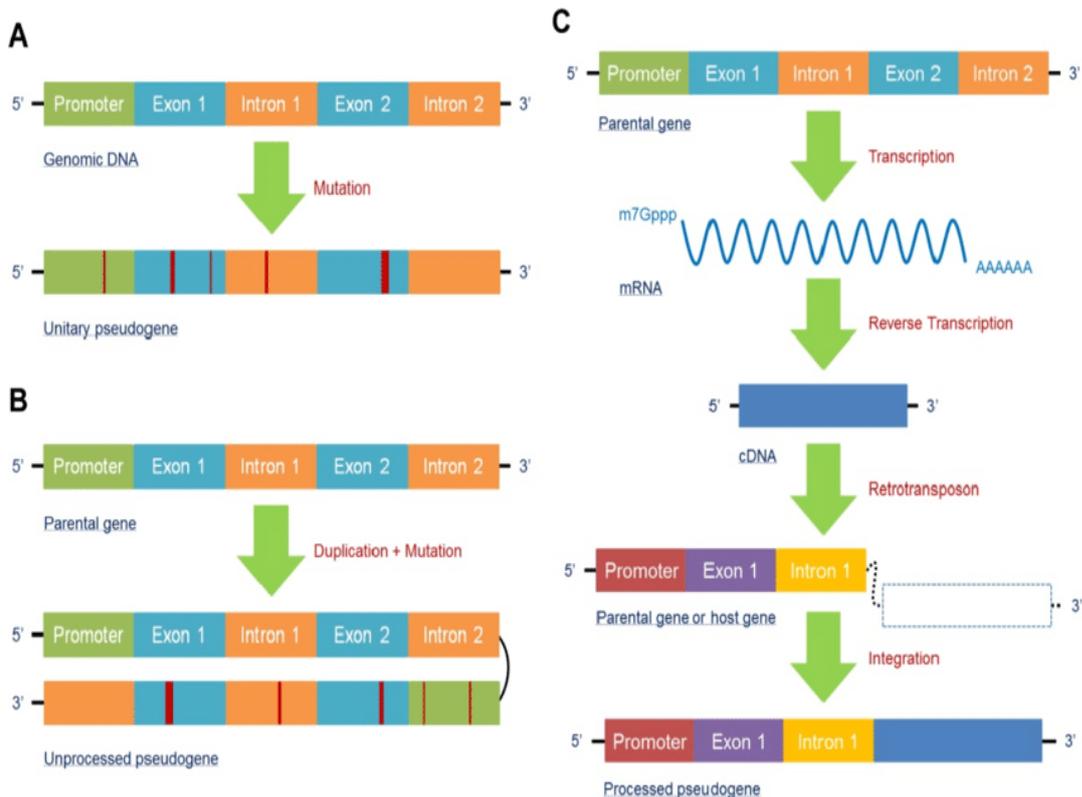


Figura 8. Rappresentazione schematica del processo di formazione degli pseudogeni: (A) pseudogene unitario, (B) pseudogene non processato, (C) pseudogene processato. (43)

Tornando alla trattazione specifica dello pseudogene pSTRC, esso appartiene alla seconda delle categorie elencate, ossia gli pseudogeni non processati o duplicati. Esso risulta infatti pressoché identico al gene corrispondente fino all'esone 15, mentre presenta maggiori differenze dall'introne 15 all'esone 29: queste differenze sono per lo più introniche o sono sostituzioni sinonime negli esoni 16 e 19; viene inoltre introdotto un codone di stop in corrispondenza dell'esone 20 dello pseudogene, mutazione responsabile della sua inattività (18).

La presenza della sequenza pSTRC va ad inficiare fortemente l'attendibilità dell'analisi molecolare del gene, rendendo di difficile interpretazione le eventuali variazioni riscontrate nelle regioni con omologia del 100%, che potrebbero infatti appartenere tanto a STRC quanto a pSTRC. L'interpretazione dei dati ottenuti mediante tecnologie di sequenziamento o di analisi delle CNV risulta pertanto molto complessa poiché sussiste un alto rischio di attribuzione non corretta delle varianti tra gene e pseudogene: risulta conseguentemente spesso necessario servirsi di tecnologie più avanzate e complesse rispetto a quelle utilizzate routinariamente.

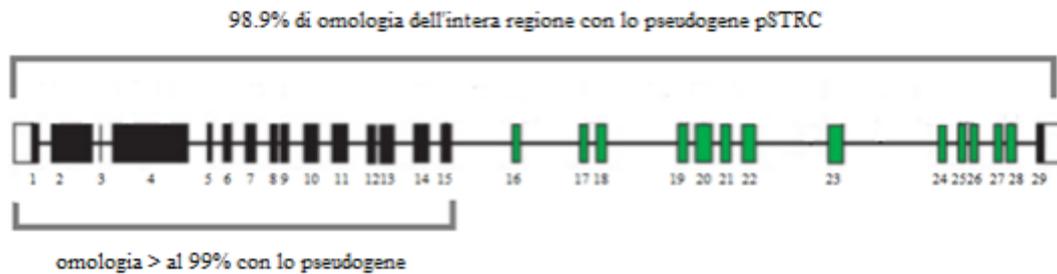


Figura 9. Rappresentazione schematica del gene STRC. I rettangoli neri rappresentano gli esoni con omologia di sequenza maggiore del 99% con lo pseudogene.

3.3 PANORAMICA SULLE PRINCIPALI TECNOLOGIE DI INDAGINE GENETICA UTILIZZATE NEL CONTESTO DELLE IPOACUSIE

Tra tutte le branche che vanno a comporre il vasto universo della medicina moderna, la genetica spicca in particolar modo per essere tra quelle andate incontro a maggior evoluzione negli ultimi pochi lustri.

La genetica nasce già nella seconda metà del XIX secolo grazie al contributo di Gregor Mendel, che pose le basi teoriche alla trasmissione ereditaria dei caratteri. Al tempo non erano tuttavia ancora disponibili le conoscenze di biologia molecolare necessarie a concettualizzare le osservazioni empiriche effettuate: ad oggi ci si riferisce a questo periodo embrionale della storia della genetica con l'espressione "genetica classica".

Un'importante svolta epistemologica nella transizione dalla genetica classica alla genetica contemporanea, o genetica del DNA, è rappresentata dagli esperimenti che vennero svolti a cavallo degli anni '40 e '50 del secolo scorso da parte di illustri scienziati, entrati ad oggi nella storia della medicina: gli avvenimenti maggiormente rilevanti furono gli esperimenti di Avery e di Hershey-Chase, che dimostrarono che la molecola contenente l'informazione genetica per la trasmissione dei caratteri ereditari è costituita dal DNA, e la descrizione della struttura tridimensionale a doppia elica del DNA stesso da parte di Watson e Crick. Solo in seguito a queste scoperte si è potuto iniziare ad affrontare la genetica con un approccio scientifico basato sull'evidenza, consci delle strutture biologiche molecolari coinvolte nei processi di trasmissione dei caratteri ereditari, che prima di quel momento erano stati solamente teorizzati in maniera speculativa (44).

La scoperta della struttura molecolare del DNA ha inoltre permesso, nei decenni successivi, una cospicua espansione degli studi di genetica, che sono andati incontro ad un'ulteriore brusca accelerazione nel 1970 con la scoperta, da parte di Arber, Nathans e Smith degli enzimi di restrizione, responsabili del taglio endonucleotidico del DNA in frammenti specifici, fintantoché nel 1977 venne sequenziato per la prima volta da parte di Fredrick Sanger un intero genoma, appartenente ad un virus batteriofago (45).

Negli anni successivi le tecniche di sequenziamento vengono perfezionate, grazie anche alla messa appunto nel 1985 da parte di Mullis della reazione a catena della polimerasi (PCR), che permette di amplificare frammenti di DNA in maniera esponenziale, fino a permettere, nel 1989, il sequenziamento del primo gene umano, ossia CTFR, codificante per una proteina canale coinvolta nella patogenesi della fibrosi cistica. È proprio in questi anni che la genetica inizia ad affermarsi nella medicina come branca specialistica.

A partire dagli anni '90 il rapido progresso tecnologico e l'avvento della digitalizzazione ha permesso un fiorente avanzamento delle tecniche di indagine del DNA, che sono andate incontro a costante innovazione, garantendo conseguentemente alla comunità scientifica l'acquisizione di innumerevoli nuove conoscenze riguardanti la composizione e il funzionamento del genoma umano, le quali sono state e vengono tutt'ora applicate in ambito clinico.

Il sequenziamento del genoma umano è stato completato per la prima volta nel 2003, coronando un meticoloso processo durato ben 13 anni di lavoro (46); ad oggi gli strumenti tecnologici a disposizione sono tuttavia estremamente più efficienti ed accurati (tra cui le tecniche di sequenziamento di nuova generazione), tanto da rendere possibile eseguire sequenziamenti di interi genomi o porzioni di esso in poche ore e con maggior precisione, facendo sì che si possa sfruttare tali tecniche nella quotidianità clinica delle varie aziende ospedaliere sparse nel territorio. Si stima inoltre che questa tendenza prosegua in futuro, rendendo ancor più agevole lo svolgimento di fini analisi molecolari che sembravano inimmaginabili solo qualche decennio fa.

Alla luce di questo breve excursus storico, appare evidente come sia fondamentale al giorno d'oggi la figura professionale del medico genetista, e come egli sia tenuto a mantenersi costantemente aggiornato riguardo le possibilità offerte dalle varie tecnologie disponibili al fine di supportare i processi di diagnosi e studio delle

moltissime patologie ereditarie conosciute, che sono ad oggi più di 6000 ed esercitano sulla medicina un impatto di tipo trasversale.

Nel presente capitolo viene effettuata una breve rassegna delle principali tecniche di indagine genetica, sulla base della loro importanza storica e del loro impiego nella pratica clinica, facendo particolare riferimento a quelle coinvolte nella ricerca di mutazioni di geni responsabili di ipoacusia, primariamente STRC.

Vengono descritte le tecnologie atte al sequenziamento, di prima, seconda e terza generazione, quelle designate alla ricerca di CNV, e gli strumenti bioinformatici a disposizione, incaricati di fornire supporto nei processi diagnostici ed analitici. Per ognuna di queste si evidenziano diverse caratteristiche, quali modalità operative, criticità, punti di forza, costi, applicazioni possibili, e rilevanza nell'analisi del gene STRC.

L'analisi molecolare e strutturale del gene STRC rappresenta infatti un'importante sfida per la genetica contemporanea in quanto le attuali tecnologie utilizzate nella routine laboratoristica spesso non risultano sufficientemente sensibili e specifiche nell'individuazione di mutazioni a causa della già citata presenza dello pseudogene pSTRC, dotato di elevatissima omologia di sequenza con STRC. La presenza di sequenze ad elevata omologia complica considerevolmente il processo di appaiamento tra le sonde e il materiale genetico che si desidera analizzare, comportando imprecisioni di lettura e di conseguenza elevata frequenza di falsi positivi e falsi negativi.

3.3.1 SEQUENZIAMENTI DI PRIMA GENERAZIONE

Il termine sequenziamento genetico indica il processo atto all'esatta determinazione della struttura primaria del DNA. Contenendo quest'ultimo tutte le informazioni geniche necessarie alla sintesi proteica e di conseguenza all'omeostasi e alla sopravvivenza di tutti gli organismi, appare evidente come la possibilità di conoscere la sua esatta struttura molecolare rivesta un traguardo rivoluzionario nel contesto della biologia molecolare e della medicina, costituendo un imprescindibile passo nella sfida atta a correlare il genotipo degli organismi viventi col proprio fenotipo.

La prima tecnica di sequenziamento genetico fu introdotta nel 1975 dallo scienziato duplice premio Nobel Frederick Sanger, e prende il nome di Metodo della Terminazione della Catena (47). Essa è basata sull'utilizzo di

dideossiribonucleotidi (ddNTP), ossia nucleotidi modificati privi del gruppo ossidrilico in 3', marcati radioattivamente, che determinano la cessazione del processo di replicazione qualora incorporati nella reazione di polimerizzazione da una DNA polimerasi ad alta processività. Vengono così prodotti vari filamenti tronchi, che verranno quindi sottoposti a migrazione elettroforetica e sequenziati con l'ausilio di lastre autoradiografiche.

Questo metodo riveste grandissima importanza storica, tuttavia ad oggi risulta superato poiché necessita di ingenti risorse in termini di lavoro e tempo. Esso richiede infatti l'utilizzo di grandi quantità gel di poliacrilammide, l'impiego di radioisotopi e lo sviluppo di lastre da esporre ai raggi X; la determinazione della sequenza inoltre viene effettuata manualmente, pertanto risulta possibile sequenziare solo 200 – 300 paia di basi in 2 – 3 giorni.

Maggiormente rilevante al giorno d'oggi, seppur ugualmente rimpiazzato da tecniche più avanzate, è il metodo Sanger automatizzato, introdotto per la prima volta nel 1986 (48), che prevede l'utilizzo di fluorocromi al posto dei radioisotopi. Ciascun ddNTP è coniugato ad una diversa sonda fluorescente; viene quindi effettuata una reazione di PCR asimmetrica in singola provetta, con un unico primer complementare alla sequenza da analizzare, che viene esteso da una DNA polimerasi utilizzando un mix di dNTP e ddNTP. I frammenti marcati ottenuti vengono separati tramite elettroforesi capillare e la lettura dei fluorofori mediante laser fornisce la sequenza del DNA del frammento analizzato. Attraverso questo metodo è possibile sequenziare un frammento di DNA che contiene dalle 700 alle 1000 paia di basi in poche ore.

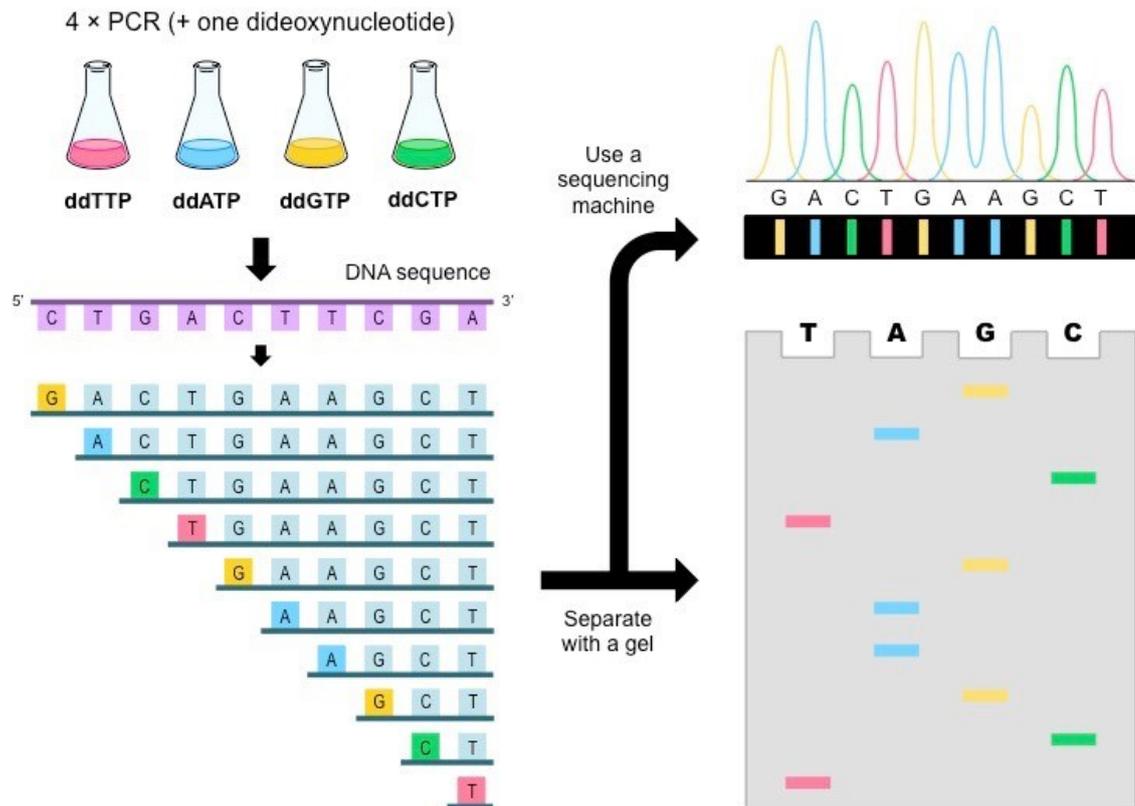


Figura 10. Rappresentazione schematica delle fasi del sequenziamento Sanger. (49)

Il sequenziamento Sanger è ancora utilizzato nella pratica clinica nel contesto della ricerca di singole mutazioni note o dell'analisi di piccoli geni costituiti da un numero esiguo di esoni. Nel campo delle ipoacusie, ad esempio, il sequenziamento Sanger trova largo impiego nell'analisi delle varianti a carico del gene *GJB2*, codificante per la proteina connessina 26 (50).

3.3.2 SEQUENZIAMENTI DI SECONDA GENERAZIONE: NEXT GENERATION SEQUENCING (NGS)

Il sequenziamento di seconda generazione, noto anche come sequenziamento massivo parallelo o Next Generation Sequencing (NGS), rappresenta uno strumento tecnologicamente molto avanzato ed efficiente, in grado di effettuare il sequenziamento in maniera rapida, accurata ed economica. Introdotto in commercio negli anni 2000 diviene nel giro di 15-20 anni parte integrante della quotidianità clinica e laboratoristica, soppiantando il meno performante sequenziamento Sanger, e costituendo tutt'oggi il metodo di indagine molecolare genetica di maggior rilievo in merito ad efficacia, precisione e tasso di utilizzo (51).

Naturalmente, rispetto a quando NGS è stato inizialmente concepito, sono stati introdotti diversi miglioramenti, che hanno permesso il perfezionamento della tecnica, esaltando ancor di più i suoi suddetti pregi. Correntemente, i sistemi adoperati nella tecnologia di seconda generazione sono molteplici, ognuno di questi sviluppato da una determinata azienda produttrice: pur differendo nelle configurazioni tecniche e nelle strumentazioni chimiche utilizzate, essi condividono il medesimo principio di base ovvero il sequenziamento parallelo massivo di molecole di DNA.

3.3.2.1 CARATTERISTICHE TECNICHE E OPERATIVE DEI SEQUENZIAMENTI DI SECONDA GENERAZIONE

La procedura di sequenziamento NGS presenta diverse sfumature in base alle tecnologie e ai protocolli utilizzati, ma prevede in ogni caso il susseguirsi di una serie di fasi, che verranno approfondite con maggior dettaglio nella sezione “Materiali e metodi”. (52)

- Preparazione delle librerie, che può eventualmente includere la fase di arricchimento. Dopo aver estratto il materiale genomico d’interesse da un campione biologico di partenza ed averlo frammentato in sequenze di lunghezza variabile (solitamente compresa tra 300 e 500 nucleotidi) mediante l’utilizzo di diversi mezzi quali ultrasuoni, enzimi di restrizione o nebulizzazione, avviene l’ibridizzazione tra i frammenti ottenuti (*clusters*) e delle sonde appositamente progettate per legarsi ad essi. A questo punto i frammenti vengono purificati dalle sonde, amplificati e posizionati in una *flow cell*. (53)

L’amplificazione delle sequenze di maggior interesse durante la preparazione delle librerie viene definita fase di arricchimento e può sfruttare diverse metodiche quale la PCR o la cattura mediante biglie magnetiche.

- Sequenziamento, con *base calling*. Il processo di sequenziamento vero e proprio viene attuato sfruttando l’impiego di una DNA polimerasi in grado di incorporare dideossinucleotidi fluorescenti: questa va a replicare i frammenti della libreria producendo dei nuovi filamenti (*reads*) la cui emissione luminosa è captata da un rilevatore che sarà in grado di fornire l’esatta sequenza dei nucleotidi. la lunghezza d’onda dell’emissione e

l'intensità del segnale determinano quello che viene chiamato *base calling*, ovvero la conversione per via computazionale dei dati ottenuti in forma di immagini in letture di sequenza.

Per ogni cluster, tutti i filamenti, identici tra loro, vengono letti simultaneamente: i *cluster* di una flow cell sono centinaia di milioni, il che spiega perché tale processo venga definito “sequenziamento massivo parallelo”.

- Allineamento e *variant calling*. Tale passaggio si avvale dell'utilizzo di programmi bioinformatici in grado di elaborare i file grezzi prodotti dal sequenziatore procedendo all'allineamento delle *reads* ottenute con un genoma umano di riferimento. Una volta allineate le sequenze, il software le confronta con il genoma di riferimento per identificare le varianti, procedendo in automatico alla loro caratterizzazione (*variant calling*).
- Annotazione. L'ultimo passaggio operativo del sequenziamento NGS consiste nel processo di annotazione delle varianti, ovvero la ricerca di tutti i dati disponibili nei database e/o in letteratura. Il risultato finale è una tabella riassuntiva delle varianti identificate ed annotate, con tanto di visualizzazione grafica delle *reads* allineate e della profondità di lettura.

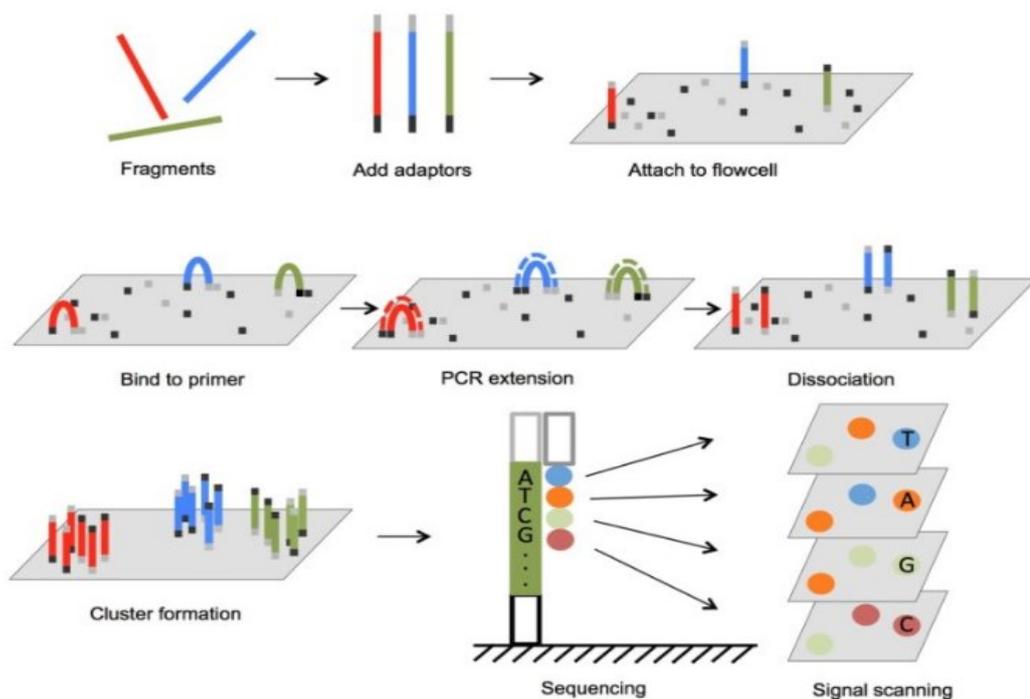


Figura 11. Sequenziamento di seconda generazione: tecnologia di sequenziamento Illumina. (77)

A differenza del metodo Sanger, in cui vengono eseguite singole reazioni di sequenza, nel NGS i cicli di reazione sono iterativi: il risultato è la generazione, in una singola corsa dello strumento, di sequenze di DNA che nel complesso sono nel *range* di centinaia di megabasi o gigabasi, ovvero un numero superiore di almeno tre o quattro ordini di grandezza rispetto alle tecnologie precedenti.

Il tasso di errore delle varie tecnologie NGS è superiore rispetto al metodo di Sanger, ma l'accuratezza dell'analisi è comunque garantita dal sequenziamento iterativo massivo, dal quale è possibile derivare una sequenza *consensus*. Per interpretare correttamente i dati ottenuti è quindi necessario analizzare un numero elevato di *reads* sovrapposte, ciascuna delle quali contribuisce a determinare un parametro denominato *depth of coverage* (profondità di lettura), che rappresenta il numero di letture presenti ad un determinato nucleotide nella sequenza target. Conoscendo le dimensioni del genoma target (G), il numero di letture (N) e la lunghezza media per lettura (L), è possibile ottenere il valore di coverage, che corrisponde a $N \times L / G$. Maggiore è la profondità di lettura, minore sarà il tasso d'errore dell'analisi, correlato ad errori del *base calling* (ovvero la conversione per via computazionale dei dati ottenuti in forma di immagini in letture di sequenza) e all'allineamento.

La profondità di lettura di una data regione sequenziata è variabile e dipende da numerosi fattori, tra cui la preparazione del campione ed il tipo di piattaforma utilizzata. Le varie piattaforme differiscono infatti dal punto di vista tecnico e forniscono risultati diversi in termini di lunghezza dei singoli frammenti sequenziali e di coverage.

3.3.2.2 CRITICITA' CORRELATE AI SEQUENZIAMENTI DI SECONDA GENERAZIONE

Dal punto di vista operativo l'NGS risulta molto simile al sequenziamento di prima generazione, descritto in precedenza; le differenze maggiormente rilevanti sono pertanto di carattere quantitativo e risiedono principalmente, come già anticipato, nella possibilità di analizzare quantità di materiale genomico superiori, oltre che nella maggior accuratezza, rapidità di analisi e risparmio economico. Al netto di tutto ciò, l'NGS non costituisce tuttavia un punto d'arrivo del tutto soddisfacente per quanto riguarda le tecnologie di sequenziamento del DNA, in quanto, pur

essendo di gran lunga superiore al metodo Sanger, presenta anch'esso una serie di criticità importanti.

Il sequenziamento NGS presenta infatti ancora dei limiti correlati all'imprecisione nella lettura di regioni ricche di citosine e guanine (sequenze C-G); tali sequenze sono abbondanti nelle regioni regolatorie del gene, localizzate in particolare a monte dello stesso e nel primo esone. Di conseguenza, la copertura del sequenziamento non è sempre ottimale, raggiungendo in tali regioni una profondità di lettura insufficiente (54).

Un altro limite di rilievo della metodica è rappresentato dalla ridotta lunghezza dei singoli frammenti sequenziati (*reads*), che oscilla tra le 50 e le 250 paia di basi. Questa caratteristica fa sì che il sequenziamento di seconda generazione risulti molto efficiente nella ricerca di varianti di piccole dimensioni, ma non sia ottimale nell'individuare riarrangiamenti genici di maggiori dimensioni, come le CNV, e di riarrangiamenti cromosomici strutturali come traslocazioni ed inversioni. Ad oggi, nella routine laboratoristica, l'analisi delle CNV si serve pertanto di metodiche alternative al sequenziamento NGS quali MLPA ed Array-CGH, che verranno approfondite nei prossimi paragrafi. La necessità di utilizzare tecniche differenti per l'analisi delle CNV risulta quindi dispendiosa in termini di tempo e denaro; per questo motivo si sta cercando di ovviare a tale criticità attraverso lo sviluppo di strumenti bioinformatici che ne migliorino la precisione diagnostica nel contesto delle CNV. (55)

Il sequenziamento NGS con *reads* corte impedisce inoltre l'analisi di sequenze ripetute (non è quindi utilizzabile nella diagnostica di patologie genetiche da espansione di triplette) e di sequenze che presentino una elevata omologia.

Queste criticità risultano quindi essere di particolare rilievo nel contesto del sequenziamento del gene *STRC*. Ciò è dovuto non solo all'elevata frequenza di alterazioni geniche strutturali quali CNV, ma anche alla presenza, a poche kilobasi di distanza, dello pseudogene *pSTRC* che presenta un elevatissimo grado di omologia di sequenza. (17)

Il sequenziamento di seconda generazione riveste comunque un ruolo di considerevole rilevanza nella ricerca di alterazioni genetiche in *STRC* e viene utilizzato routinariamente a tale proposito; il professionista incaricato della refertazione dell'esame eseguito deve tuttavia essere ben conscio del fatto che, per quanto la metodica sia molto precisa nella rilevazione di mutazioni puntiformi,

queste devono essere poi confermate, se necessario, utilizzando delle metodiche di analisi aggiuntive che permettano di discriminare se la variante sia localizzata a livello del gene (con possibili conseguenze cliniche) oppure dello pseudogene (senza conseguenze cliniche). Vanno inoltre tenuti presenti i limiti dell'analisi, anche utilizzando specifici software bioinformatici, nell'individuazione di riarrangiamenti maggiori, quali CNV o conversioni geniche.

La conferma delle varianti in regioni di omologia può essere effettuata mediante amplificazione selettiva della regione di interesse e successivo sequenziamento dell'amplificato. L'amplificazione selettiva necessita di una coppia di primer disegnati *ad hoc* in modo tale da sfruttare le differenze nella sequenza tra le due regioni con omologia. In caso di ampie regioni con elevata omologia di sequenza, come nel caso di *STRC* e del suo pseudogene, è necessario ricorrere ad una *Long Range PCR*; tale metodica permette infatti l'amplificazione di ampie sequenze di DNA di dimensione variabile da diverse migliaia di basi fino a 30 kilobasi, che tipicamente non possono essere amplificate sfruttando metodi e reagenti utilizzati per la PCR standard (56). Nel caso di *STRC*, è possibile amplificare selettivamente il gene mediante un'unica reazione di *Long Range PCR*, evitando quindi la "contaminazione" del campione da parte di sequenze dello pseudogene. È quindi possibile poi procedere con il sequenziamento dell'intero amplificato mediante NGS oppure della sola regione di interesse contenente la variante da confermare (in tal caso è necessaria una ulteriore amplificazione con PCR) mediante NGS o metodo Sanger.

Il sequenziamento NGS dell'intero amplificato di un determinato gene ottenuto mediante *Long Range PCR* può essere efficacemente effettuato utilizzando, nella fase di preparazione delle librerie, il protocollo Illumina "*Nextera Flex for Enrichment*". Tale protocollo si caratterizza per la rapidità del processo di preparazione delle librerie e per l'estrema precisione degli arricchimenti effettuati, che possono poi esitare in elevate profondità di lettura (anche 10mila *reads*). (57)

Le caratteristiche operative che distinguono questa metodologia risiedono nell'utilizzo di un'innovativa chimica basata sulle microsferi, che permette di eseguire sequenziamenti molto accurati sfruttando un'unica semplice fase di ibridazione tra le sonde e il campione biologico da analizzare. Nello specifico, viene effettuata una reazione denominata tagmentazione, che consiste nell'unione delle fasi di frammentazione del DNA e di ligazione degli adattatori: questa reazione, che

richiede mediamente soltanto pochi minuti (circa 15), viene effettuata direttamente sulle microsfele, in maniera uniforme e mediante l'utilizzo di trasposoni, facendo in modo che l'arricchimento del materiale genetico d'interesse avvenga in maniera estremamente precisa e rapida, arrivando a richiedere circa 6 ore, contro le oltre 48 impiegate dai sistemi di NGS tradizionale.

In alternativa al sequenziamento dopo Long Range PCR, l'identificazione delle varianti in regioni di omologia può essere effettuata utilizzando metodiche di sequenziamento di terza generazione, caratterizzate dalla produzione di *reads* più lunghe rispetto a quelle ottenibili con il sequenziamento di seconda generazione, come verrà approfondito nell'apposito paragrafo dedicato 3.3.3.

3.3.2.3 PRINCIPALI APPLICAZIONI DEL SEQUENZIAMENTO NGS IN GENETICA MEDICA

L'introduzione delle tecnologie NGS ha rappresentato una svolta cruciale nell'ambito del progresso della biologia molecolare e della medicina, garantendo la possibilità di eseguire indagini precedentemente inaccessibili.

In particolare, nel contesto clinico, il sequenziamento di seconda generazione permette di indagare agevolmente la presenza di specifiche mutazioni, specialmente di tipo puntiforme, in pazienti con sospetta patologia genetica o familiarità per essa. Per quanto la ricerca di mutazioni nel genoma di pazienti oggetto d'interesse fosse tecnicamente eseguibile già in precedenza, sfruttando il metodo di Sanger, di fatto ciò non si verificava con la stessa consistenza e sistematicità, sia per ragioni legate alle tempistiche e ai costi delle indagini, sia perché è proprio grazie a NGS che è stato possibile per la comunità medico-scientifica acquisire le conoscenze necessarie a definire le alterazioni rilevate e a determinarne l'eventuale patogenicità. Il sequenziamento di un unico genoma umano, portato a termine nel 2003 grazie all'applicazione del metodo Sanger, non risulta infatti sufficiente a distinguere quali alterazioni siano potenzialmente responsabili di un quadro patologico e quali semplici varianti individuali prive di significato clinico; per poter aspirare a riuscire a correlare il genotipo sequenziato col fenotipo osservato, traguardo che ancora oggi non può comunque considerarsi completamente raggiunto, è necessario infatti effettuare un numero elevatissimo di sequenziamenti, includendo sia l'uomo che altre specie animali, seguiti da processi di analisi

statistica e bioinformatica dei risultati: tutto ciò è divenuto possibile solo grazie all'introduzione delle metodologie di sequenziamento di seconda generazione.

È previsto in ambito clinico l'utilizzo di tre principali approcci d'indagine, ovvero l'analisi di pannelli multigenici specifici per un determinato fenotipo, l'analisi dell'esoma clinico (Mendelioma) e l'analisi dell'intero esoma (WES, Whole Exome Sequencing); il sequenziamento dell'intero genoma (WGS, Whole Genome Sequencing) non è ancora proponibile di routine nella diagnostica (per i costi associati e i limiti correlati all'interpretazione delle varianti in regioni non codificanti), ma è frequentemente utilizzato per finalità di ricerca.

I pannelli multigenici consistono nel sequenziamento, mediante tecniche di seconda generazione, di una serie di geni coinvolti nella patogenesi della specifica condizione d'interesse (58). Generalmente vengono analizzate le regioni codificanti e le giunzioni introni-esone: tali sono le regioni in cui si localizzano la maggior parte delle varianti patogenetiche. In casi specifici, il target di analisi viene esteso anche alle regioni UTR (UnTranslated Regions), alle regioni introniche profonde o alle regioni regolatorie del gene (quali il promotore). Esistono svariate tipologie di pannelli multigenici disponibili nella pratica clinica, ciascuno dei quali viene selezionato in base alla sintomatologia espressa dal paziente. Solo nel contesto delle ipoacusie esistono molteplici pannelli genici disponibili all'utilizzo nella pratica clinica, che vanno complessivamente a coprire i più di 100 geni noti che possono essere coinvolti nella patogenesi del quadro.

Un'altra tipologia di indagine utilizzata nel contesto dei sequenziamenti di seconda generazione è rappresentata dall'analisi dell'esoma clinico, anche noto come Mendelioma, oppure dell'intero esoma (WES): vengono sequenziate tutte le regioni codificanti (esoni) dei geni ad oggi noti associati a patologia (esoma clinico) o di tutti i geni del genoma (WES). Tali espedienti risultano maggiormente dispendiosi in termini economici e di tempistiche, ma sono caratterizzati da maggiore efficienza nell'individuazione di varianti non attese. (59)

3.3.3 SEQUENZIAMENTO DI TERZA GENERAZIONE MEDIANTE OXFORD NANOPORE

Nonostante l'introduzione delle tecnologie di sequenziamento genico di seconda generazione abbia rappresentato un fondamentale punto di svolta storico e culturale nel progresso scientifico volto all'analisi genetica della popolazione, esse

presentano comunque una serie di criticità, discusse nei precedenti paragrafi, che hanno determinato la necessità di sviluppare tecniche di nuova generazione ulteriormente avanzate. Tra queste, spiccano per rilevanza le tecnologie *Single Molecule Real Time Sequencing* (SMRT) di Pacific Biosciences (Life Technologies) e *Oxford Nanopore*.

Oxford Nanopore, tecnologia immessa in commercio nel 2014 dall'omonima azienda britannica, consiste in un insieme di prodotti volti al sequenziamento di DNA tramite l'utilizzo di un sistema di riconoscimento basato sui nanopori. (60)

Dal punto di vista operativo, il sequenziamento Nanopore sfrutta la presenza di nanopori inseriti in una membrana elettro-resistente, attraversata da una corrente ionica. Quando una molecola di DNA attraversa un poro, causa una specifica alterazione nel flusso di corrente: durante il passaggio dell'acido nucleico vengono quindi operate numerose misurazioni della corrente ionica, le quali vengono poi convertite da un software di *basecalling* nella corrispettiva sequenza genomica.

Il *nanopore sequencing* prevede la preparazione di una libreria in cui il DNA viene riparato alle estremità tramite l'aggiunta di una piccola coda di adenina che creerà delle terminazioni compatibili con l'adattatore di sequenziamento che contiene delle timidine. A questo punto attraverso una reazione di ligazione vengono uniti gli adattatori ad entrambe le estremità dei frammenti di DNA in modo da poter sequenziare entrambi i filamenti e ridurre il *single strand bias*.

Gli adattatori sono complessi DNA-proteina, in cui un'elicasi assicura il movimento graduale del DNA attraverso il poro. Quindi il DNA a doppio filamento viene srotolato da una proteina motrice nei pressi del poro, attraverso il quale quindi passa un filamento alla volta dotato di velocità di traslocazione pari a circa 400 basi al secondo. Questa tecnologia permette l'analisi in tempo reale di lunghi frammenti di DNA nativo evitando quindi il bias della PCR presente nelle tecnologie NGS attualmente in uso e permettendo di sequenziare anche regioni del genoma più complesse.

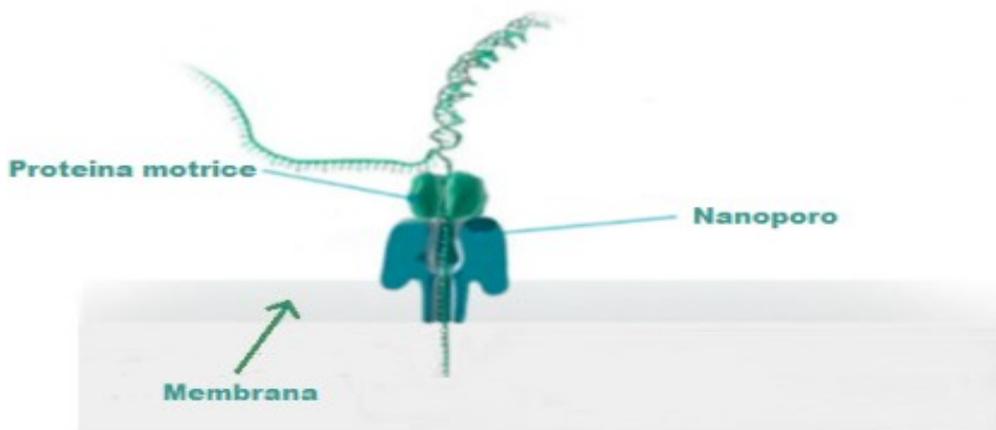


Figura 12. Il processo del Nanopore sequencing. (78)

La tecnologia di Oxford Nanopore permette di analizzare il DNA nativo; nella maggior parte dei casi è però necessario effettuare un arricchimento per ottenere un numero di *reads* sufficiente da permettere un'analisi accurata. I protocolli di arricchimento proposti dalla ditta sono: l'utilizzo di sonde Cas9, l'adaptive sampling, oppure il sequenziamento di ampliconi. (61) (62) (63)

Nonostante i vantaggi correlati alla generazione di *reads* lunghe ed alla possibilità di utilizzare DNA nativo, il sequenziamento Oxford Nanopore presenta alcuni limiti: in particolare, il sistema è caratterizzato da un elevato tasso di errore che è stimato pari a circa 10-15%, soprattutto nella rilevazione di inserzioni/delezioni; l'analisi bioinformatica, necessaria per l'interpretazione dell'output del sequenziamento, utilizza inoltre tools open source molto complessi, che richiedono dispositivi informatici estremamente potenti per essere supportati. (64)

Alla luce delle caratteristiche elencate, e viste le criticità ben note correlate al sequenziamento di numerosi geni ostili, quali ad esempio il gene *STRC*, caratterizzato dalla presenza a valle di uno pseudogene dotato di elevatissima omologia di sequenza, appare evidente come le tecniche di terza generazione, che permettono la produzione di *reads* molto più lunghe e possono non necessitare dello step di amplificazione del DNA, risultino di fondamentale importanza.

Ad oggi il sequenziamento di terza generazione non risulta ancora utilizzato di routine per scopi di diagnosi clinica, pertanto il suo utilizzo è limitato prevalentemente a contesti di ricerca; è tuttavia verosimile che prossimamente esso entrerà ad essere parte integrante del processo diagnostico di numerose patologie

ad eziologia genetica per le quali non risultano sufficienti le tecniche di seconda generazione.

3.3.4 METODICHE DI INDAGINE DI CNV: ARRAY CGH

Le due tecniche maggiormente rilevanti per l'identificazione delle Copy Number Variants (CNV) sono l'analisi Array-CGH e l'analisi MLPA.

L'Array-CGH (*Array-Comparative Genomic Hybridization*, Ibridazione genomica comparativa su microarray) è una tecnica largamente utilizzata nella pratica clinica soprattutto nel contesto della diagnosi prenatale, che si basa sulla comparazione quantitativa tra DNA in esame (DNA test) e DNA genomico di riferimento (reference DNA) (65). Questi DNA sono frammentati e quindi marcati in maniera differenziale con molecole fluorescenti (generalmente si utilizza un fluorocromo rosso per il DNA test ed un fluorocromo verde per il reference DNA). Successivamente i campioni vengono mescolati in parti uguali e fatti incubare (Ibridazione) su un microarray, costituito da un supporto di vetro la cui superficie è coperta da oligonucleotidi (sonde), ciascuno dei quali rappresenta una specifica regione del genoma umano. Al termine della suddetta incubazione, sia il DNA in esame che quello di controllo ibrideranno in maniera competitiva alle sonde presenti sull'array, determinando così l'emissione di segnali fluorescenti che sono rilevati da uno specifico scanner e quindi elaborati da appositi software al fine di evidenziare la presenza di CNV nel campione in studio.

Il potere risolutivo della tecnologia, che varia in base alla tipologia di array utilizzata, è generalmente compreso tra 30 e 100 Kilobasi nella diagnostica clinica; l'analisi permette quindi di identificare agilmente sbilanciamenti multigenici ma non può essere impiegata nella ricerca di piccole CNV intrageniche (esoniche o multiesoniche); a tal fine si utilizzano quindi tecnologie alternative, quali l'analisi MLPA.

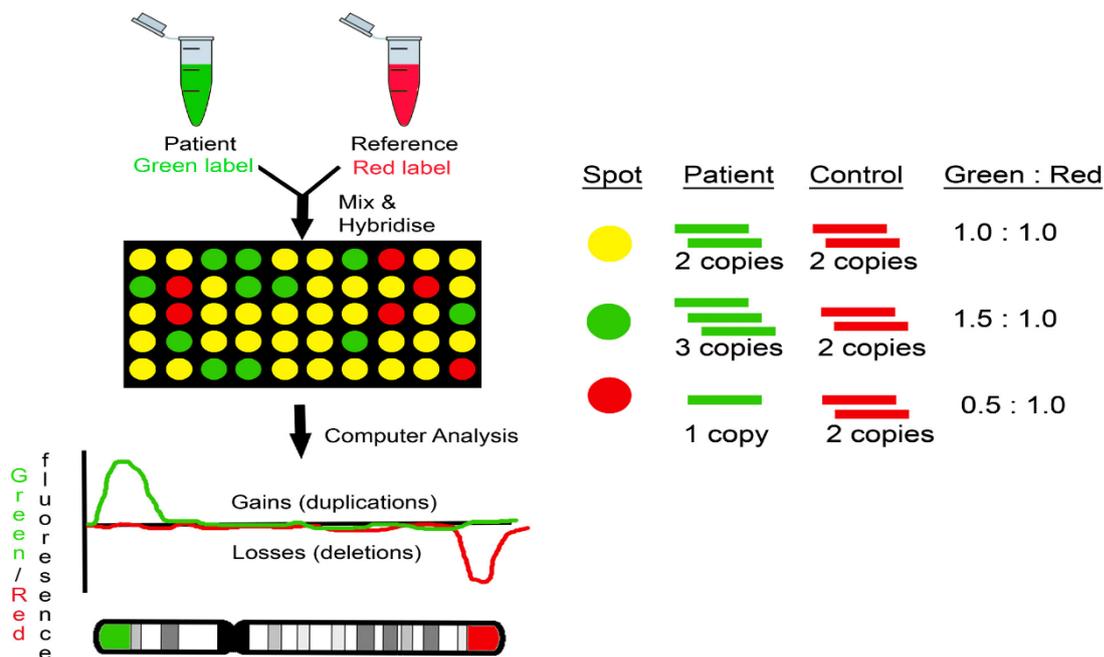


Figura 13. Rappresentazione iconografica del protocollo Array-CGH. (79)

3.3.5 METODICHE DI INDAGINE DI CNV: MLPA

L'amplificazione legatura-dipendente multipla della sonda (*Multiplex Ligation-dependent Probe Amplification*, MLPA) è una tecnica di biologia molecolare descritta per la prima volta nel 2002 e ad oggi ampiamente utilizzata per scopi di ricerca e diagnostici in tutto il mondo.

Si tratta di una tecnologia volta a rilevare anomalie del numero di copie (CNV), ossia delezioni e amplificazioni, in specifiche e predeterminate regioni del genoma associabili a disturbi genetici; è considerata ad oggi il metodo più affidabile e dotato del miglior rapporto costo/efficacia ai fini del rilevamento di CNV specifiche e accertate. In aggiunta a ciò, MLPA può essere utilizzata anche per la valutazione dello stato di metilazione delle isole CpG del DNA, per l'identificazione di mutazioni puntiformi e/o polimorfismi di singolo nucleotide (SNPs), e per la quantificazione dell'mRNA. (66)

Dal punto di vista operativo, la MLPA è una metodica basata sul legame di sonde oligonucleotidiche in corrispondenza di sequenze target. Tali sonde sono costituite da coppie di oligonucleotidi che, se correttamente legate e allineate sulla sequenza bersaglio, consentono l'amplificazione di un prodotto PCR complementare alla sequenza. Costruendo in modo opportuno il kit, è possibile riconoscere e successivamente amplificare contemporaneamente fino a 50 sequenze, sfruttando il

principio della PCR multipla; vengono così ottenuti diversi frammenti di PCR, che sono quindi separati mediante elettroforesi capillare e visualizzati come un elettroferogramma. L'area e l'altezza di ogni picco sono proporzionali al numero di copie del locus genico che si sta analizzando: in caso di delezioni o alterazioni qualitative della sequenza bersaglio si osserva la riduzione o assenza del picco corrispondente, mentre nel caso di sbilanciamenti in eccesso si assiste ad un suo aumento.

La MLPA è utile anche al fine di rilevare mutazioni puntiformi già note: in questa occasione le sonde devono essere disegnate in modo che il sito di legame sia corrispondente al sito della mutazione puntiforme, che verrà eventualmente segnalata tramite una riduzione del segnale e dell'area del picco all'elettroferogramma. (67)

Per un'analisi corretta, il profilo dei picchi di un campione deve essere sempre confrontato con quelli ottenuti da almeno due/tre campioni di DNA privo dell'anomalia in oggetto (controlli negativi) e di un campione in cui essa è presente (controllo positivo).

La MLPA riveste in particolare un ruolo di fondamentale importanza nell'indagine genetica di quelle patologie in cui il principale meccanismo patogenetico non è rappresentato da mutazioni puntiformi o di piccole dimensioni, evidenziabili mediante tecniche di sequenziamento, bensì da variazioni esoniche o multiesoniche del numero di copie del gene malattia. Esempi notevoli in tal senso sono la Distrofia Muscolare di Duchenne (DMD), la sindrome di Prader-Willi (PWS) e l'Atrofia Muscolare Spinale (SMA). Esistono peraltro numerosissime altre patologie in cui la proporzione di difetti causativi riconducibili a variazioni quantitative è comunque sufficientemente rilevante dal punto di vista epidemiologico da richiedere frequentemente un'integrazione con la tecnica MLPA come indagine di secondo livello dopo le tecniche di sequenziamento, come ad esempio l'ipoacusia *STRC* correlata. (68)

3.3.6 ANALISI BIOINFORMATICHE

Il rapido e progressivo sviluppo delle conoscenze e delle possibilità acquisite nella sfera dell'informatica che si è verificato nel recente passato, che si protrae ad oggi e che verosimilmente caratterizzerà ancor più significativamente il futuro più prossimo, ha consentito l'introduzione di svariate tipologie di analisi bioinformatica

coinvolte nella diagnosi dei disturbi ad eziologia genetica, architettate appositamente al fine di fornire supporto nelle diverse fasi del processo di diagnostica molecolare. (69)

“Analisi bioinformatica” è un termine ombrello che descrive un insieme di tecniche che si servono di sistemi computerizzati al fine di raccogliere e analizzare grandi quantità di dati biologici. Il denominatore comune che caratterizza queste tecniche di indagine consiste in un approccio di tipo multidisciplinare ai problemi biologici, tradizionalmente affrontati esclusivamente con metodiche di chimica e biologia molecolare, affiancandovi quindi un utilizzo estensivo dell’informatica, della matematica applicata, della statistica e dell’intelligenza artificiale.

È pertanto implicata la raccolta di una cospicua quantità di dati, generalmente estrapolati a partire dai molti esperimenti descritti in letteratura, e la ricerca di tendenze statistiche e correlazioni significative: le applicazioni dell’analisi bioinformatica riguardano non solo l’analisi della sequenza del DNA, ma anche la regolazione dell’espressione genica e il confronto tra genomi di diversi organismi. Nel contesto dell’analisi bioinformatica possono essere utilizzate svariate tecniche differenti in base agli specifici problemi biologici che devono essere risolti: in alcuni casi viene semplicemente effettuata la raccolta di grandi quantità di informazioni in un unico database seguita dall’esecuzione di analisi statistiche di base sui dati raccolti; altre situazioni implicano una manipolazione molto più sfumata delle informazioni raccolte e possono coinvolgere programmi maggiormente complessi dal punto di vista informatico e matematico. Al netto di ciò, è possibile distinguere i vari strumenti bioinformatici in base alle rispettive potenzialità operative ed al ruolo svolto nella risoluzione dei quesiti biologici.

Tra le tecniche disponibili, si prendono in particolare considerazione, sia per la rilevanza generale nella pratica clinica che per le applicazioni nello studio oggetto della tesi, le analisi bioinformatiche volte alla ricerca di CNV a partire da sequenziamenti NGS e le analisi di tipo interpretativo e predittivo delle mutazioni individuate mediante la diagnostica di laboratorio.

3.3.6.1 ANALISI BIOINFORMATICA DELLE COPY NUMBER VARIATIONS

Tra le principali applicazioni della bioinformatica nell’ambito della diagnostica molecolare di disturbi ad eziologia genetica emerge l’analisi delle CNV eseguita a

partire dagli esiti di un sequenziamento di seconda generazione: ciò consiste nell'estrapolazione di un dato indiretto riguardo la presenza di eventuali delezioni o amplificazioni geniche sulla base della numerosità di *reads* ottenute dal sequenziamento di una determinata porzione di genoma. Si tratta di una metodica molto rapida e semplice, ampiamente utilizzata nella pratica clinica e caratterizzata da elevato valore predittivo negativo, ma non sufficientemente precisa da essere considerata diagnostica: ne è quindi raccomandata l'esecuzione in prima battuta in caso i sospetto di CNV, ma, se il responso è positivo, è necessario confermare quanto osservato mediante l'utilizzo di tecniche di analisi delle CNV maggiormente specifiche, quali ad esempio MLPA, metodica discussa in precedenza. (70)

Esistono svariati programmi utilizzati in bioinformatica al fine di rilevare la presenza di CNV a partire da sequenziamenti di nuova generazione; uno dei più comuni, adottato anche nel contesto dello studio oggetto dell'elaborato di tesi, consiste nella versione 4.2.2 di *Sure Call*, realizzato dall'azienda statunitense *Agilent technologies*.

Il programma è in grado di conteggiare, allineare e confrontare le *reads* NGS ottenute a partire dal campione di DNA del soggetto analizzato, in cui si sospetta la presenza di CNV, con quelle ottenute da un controllo putativamente privo di alterazioni geniche in corrispondenza della regione d'interesse. In presenza di CNV, ci si aspetta che *Sure Call* evidenzi una discrepanza particolarmente significativa per quanto riguarda il numero di *reads* svolte durante il sequenziamento NGS nell'uno e nell'altro paziente, indicando così la presenza e l'esatta posizione di un eventuale squilibrio genico. Questo passaggio risulta particolarmente rilevante nel caso del riscontro di una singola variante in un gene associato a patologia recessiva o per svelare, in caso di apparente omozigosità di una variante, la presenza di una delezione a carico di un allele.

Le strumentazioni bioinformatiche permettono pertanto di ottenere un dato di tipo indiretto riguardo la presenza di riarrangiamenti genici maggiori a partire da un sequenziamento NGS, sistema di indagine molecolare tipicamente designato per la ricerca di mutazioni puntiformi o riguardanti un numero esiguo di nucleotidi.

3.3.6.2 ANALISI BIOINFORMATICA DELLE MUTAZIONI RILEVATE TRAMITE L'UTILIZZO DI SOFTWARE INTERPRETATIVI E PREDITTIVI

Gli strumenti bioinformatici rivestono un ruolo di rilievo nel contesto della diagnosi di disturbi ad eziologia genetica non solo per quanto riguarda la ricerca di alterazioni geniche ma anche nell'interpretazione delle stesse: esistono infatti numerosi programmi in grado di analizzare le mutazioni individuate, considerando diversi fattori quali la frequenza allelica, i dati presenti in letteratura, la conservazione evolutiva delle sequenze, il tipo di mutazione e molti altri; a partire da queste informazioni i programmi restituiscono quindi una stima dell'eventuale patogenicità della mutazione presa in esame.

Più nello specifico, sulla base dei parametri elencati in precedenza, i programmi in questione sono in grado di classificare le varianti in base alla stima della loro patogenicità in una serie di categorie, in accordo con le linee guida stilate dall'American College of Medical Genetics (ACMG). (71) (72)

- Varianti benigne/verosimilmente benigne: la probabilità che queste varianti siano causative del fenotipo osservato nel paziente è minore del 10%
- Varianti di significato incerto (VUS): la probabilità che esse siano patogene è compresa tra 10% e 90%. Si tratta di una categoria dentro la quale ricadono la maggioranza delle varianti individuate con le metodologie di sequenziamento disponibili, e risultano piuttosto complicate da refertare in quanto non è chiaro il loro ruolo biologico. Spesso in presenza di VUS è indicato effettuare la segregazione parentale in modo tale da verificarne la presenza nei genitori del probando e correlare il genotipo al fenotipo
- Varianti verosimilmente patogenetiche: la probabilità che siano causative del disturbo è superiore al 90%
- Varianti francamente patogenetiche: la probabilità che esse siano causative di malattia è superiore al 99%.

4. SCOPO DELLO STUDIO

La popolazione oggetto del presente studio è costituita da una coorte di 21 pazienti affetti da ipoacusia neurosensoriale non sindromica da mutazione in *STRC* analizzati presso il laboratorio della UOC Genetica ed Epidemiologia Clinica dell'Azienda Ospedale Università di Padova.

Nello specifico, gli obiettivi dello studio sono i seguenti:

- Analisi delle mutazioni individuate nei soggetti appartenenti alla casistica oggetto dello studio, classificandole dal punto di vista strutturale e molecolare e valutandone la frequenza.
- Descrizione delle manifestazioni cliniche dei diversi pazienti oggetto di studio, ricercando un'eventuale correlazione fra il loro genotipo ed il fenotipo presentato.
- Valutazione dell'accuratezza dell'analisi bioinformatica dei dati di sequenziamento NGS per l'identificazione di CNV e riarrangiamenti complessi di *STRC*.
- Caratterizzazione molecolare di un riarrangiamento complesso del gene *STRC* (conversione genica) mediante l'utilizzo di tecniche di sequenziamento di seconda e terza generazione.

5. MATERIALI E METODI

5.1 SELEZIONE DEL CAMPIONE DELLO STUDIO

Ai fini dello svolgimento dello studio, è stato selezionato un campione di 21 soggetti affetti da ipoacusia neurosensoriale recanti mutazioni di STRC in entrambi gli alleli del gene.

La casistica selezionata comprende 9 pazienti che hanno effettuato consulenza genetica presso l'unità operativa complessa di Genetica Medica dell'Azienda Ospedale Università di Padova, mentre i restanti 12 pazienti hanno svolto la consulenza presso altre sedi, che esulano dall'ospedale di Padova; le indagini molecolari e strutturali volte alla ricerca di mutazioni del gene STRC considerate nello studio sono invece state interamente effettuate presso il laboratorio di Genetica di Padova.

In particolare, le suddette indagini genetiche sono state svolte tra il 2017 e il 2022, e al momento del test i probandi presentavano un'età compresa tra 1 e 69 anni.

Ognuno dei soggetti facenti parte della casistica, o i corrispettivi genitori o tutori legali in caso di pazienti minorenni, hanno esplicitamente espresso l'informata disponibilità a partecipare in forma anonima a studi o ricerche scientifiche, firmando un apposito modulo alla presenza di un medico.

A partire dai dati disponibili è stato quindi stilato un database in cui, per ogni paziente, venivano specificate, quando note, le seguenti informazioni: data di nascita, indicazione all'analisi genetica, anno di esecuzione del test genetico, pannello o pannelli eseguiti, risultato dell'analisi molecolare, eventuale esecuzione di segregazione familiare delle

varianti riscontrate e/o presenza di altri familiari sottoposti all'indagine genetica, fenotipo clinico.

Relativamente al fenotipo clinico, sono state segnalate, quando note, le caratteristiche del quadro ipoacusico specificando l'insorgenza dell'ipoacusia (preverbale/periverbale vs postverbale), la gravità (lieve/moderata vs severa/profonda) e l'andamento (progressivo vs stabile). Nel caso in cui fosse disponibile la documentazione clinica dello studio neuroradiologico, in particolare TC dell'orecchio e RM cerebrale, questa è stata riportata. Sono stati riportati inoltre

altri segni clinici noti e l'età d'esordio degli stessi. Per tutti i pazienti è stato inoltre segnalato se fosse stata eseguita una consulenza genetica presso il Servizio di Genetica ed Epidemiologia Clinica dell'Azienda Ospedale Università di Padova e, nei casi in cui è stata effettuata, le informazioni in essa contenute sono state utilizzate per la compilazione del database.

5.2 ANALISI MOLECOLARE MEDIANTE SEQUENZIAMENTO DI SECONDA GENERAZIONE (PANNELLO MULTIGENICO PER IPOACUSIE)

L'intera coorte di pazienti presa in considerazione ai fini dello studio svolto è stata studiata in prima battuta mediante indagini di sequenziamento di seconda generazione NGS, mirate alla ricerca di mutazioni genetiche coinvolte nella patogenesi del quadro clinico di ipoacusia neurosensoriale.

A tale scopo, sono stati presi in esame diverse tipologie di pannelli genici associati ad ipoacusia, i quali si differenziano tra di loro a seconda della quantità di sequenze geniche analizzate: tutti i pannelli genici utilizzati nei pazienti dello studio includevano il gene *STRC*.

Al fine di disegnare i pannelli oggetto di discussione, selezionando pertanto quali geni sia opportuno considerare e quali no nel contesto di diagnosi di ipoacusia ad eziologia genetica, ci si è serviti di un software fornito dall'azienda *Agilent technologies*, denominato *SureDesign*. Esso consiste in un programma che permette di definire le regioni di maggior interesse e di disegnare, in base ai parametri selezionati, delle specifiche sonde deputate al processo di arricchimento delle regioni target nel successivo step di preparazione delle librerie. Le regioni target consistono solitamente in tutte le sequenze codificanti dei geni selezionati +/-10 nucleotidi fiancheggianti.

5.2.1 FASE DI PREPARAZIONE DELLE LIBRERIE

Il DNA da analizzare è stato estratto dai pazienti oggetto dello studio a partire da materiale biologico ematico o salivare grazie ad un sistema automatico (*Magpurix*, *Resnova*) che sfrutta il legame degli acidi nucleici a biglie magnetiche.

Una volta estratto il DNA, si è proceduto con la preparazione delle librerie. Nel corso degli anni sono stati utilizzati due diversi sistemi per l'esecuzione di questo passaggio: *Haloplex Target Enrichment System* e *SureSelect Target Enrichment System*; entrambi permettono un'analisi rapida, semplice ed efficiente delle regioni genomiche di interesse.

Il processo *Haloplex* è suddivisibile in quattro fasi principali:

1. *Digestione del campione.* Il campione di DNA viene frammentato sfruttando 16 enzimi di restrizione diversi e quindi denaturato, permettendo così l'ottenimento di una libreria di frammenti. La corretta digestione dei frammenti viene verificata tramite l'utilizzo di un *Bioanalyzer*, ovvero un sistema che sfrutta il principio dell'elettroforesi su gel applicato ad un chip, in cui i vari frammenti migrano a seconda delle loro dimensioni e vengono confrontati con una serie di frammenti di dimensioni note contenute nel cosiddetto ladder di riferimento.
2. *Ibridizzazione delle sonde.* La libreria di sonde *Haloplex*, costituita da oligonucleotidi progettati specificatamente per legarsi ad entrambe le estremità di un frammento di restrizione contenente una sequenza target, viene aggiunta al campione e ibridizzata con i frammenti bersaglio: in questo modo i frammenti che contengono la sequenza d'interesse sono portati a formare una molecola di DNA circolare, in quanto la sonda porta all'avvicinamento delle due estremità. Le sonde contengono nella parte centrale una sequenza specifica tipica di questo metodo, utile per il successivo sequenziamento, che viene così incorporata nella molecola circolare, tra le due estremità del frammento, assieme ad una coppia di *adapters* che serviranno per la *bridge PCR* durante la fase di generazione dei *clusters*. Inoltre, a ciascun campione viene legato un indice, o *barcode*, ovvero una breve sequenza nucleotidica che, identificando in maniera univoca tutte le *reads* di uno stesso campione, permette l'analisi in parallelo di più pazienti.
3. *Purificazione e ligazione.* Poiché ciascuna sonda contiene una molecola di biotina, i frammenti desiderati possono essere recuperati tramite biglie magnetiche contenenti streptavidina, sfruttando la reazione enzimatica che si verifica tra le due molecole. Le provette vengono quindi inserite in piastre magnetiche che attirano le biglie verso i bordi della provetta, ancorando solo

le molecole di DNA circolare, e rimuovendo i frammenti non circolari. A questo punto viene aggiunta una ligasi che permette la chiusura dei frammenti circolarizzati, i quali possono ora essere liberati da sonda e biglie magnetiche.

4. *Amplificazione dei frammenti tramite PCR.* Solamente le molecole di DNA completamente circolari vengono amplificate dalla PCR, utilizzando primer universali dotati di *barcode*. Al termine di questo passaggio si ha a disposizione un campione di DNA target amplificato, pronto per il sequenziamento: l'amplificazione è indispensabile perché il sistema di *base calling*, basato su emissioni luminose, non è abbastanza sensibile da captare singole emissioni di fluorofori. Come nel passaggio 1, viene quindi effettuato un controllo tramite *Bionalyzer*, per assicurarsi che i frammenti siano idonei al sequenziamento. Durante la fase di preparazione delle librerie vengono aggiunti anche degli *adapters*, ovvero degli oligonucleotidi che, alla fine del processo di amplificazione tramite PCR, si troveranno alle due estremità di ogni frammento. Questi adattatori sono utili per la generazione dei clusters, in quanto la libreria viene caricata su una *flow-cell* dove i frammenti sono catturati su una superficie su cui sono legati degli oligonucleotidi complementari proprio agli *adapters*

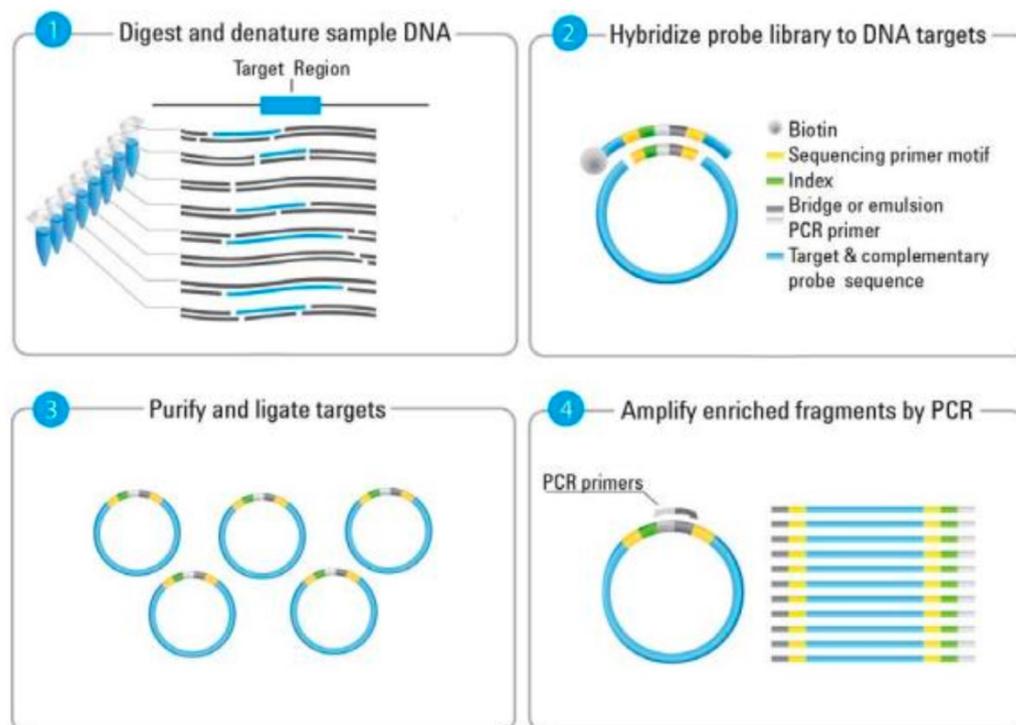


Figura 14. Preparazione delle librerie secondo metodo Haloplex.

Il processo *Sureselect*, d'altro canto, pur condividendo alcuni concetti di base, prevede una sequenza di passaggi diversi, suddivisibili in tre fasi.

1. *Preparazione del campione.* il campione di DNA va incontro ad un processo di tagmentazione, che consiste nella contemporanea frammentazione e aggiunta alle sue estremità di frammenti adattatori attraverso l'azione enzimatica di trasposasi ingegnerizzate. I frammenti di DNA così ottenuti vengono quindi amplificati mediante PCR, purificati, ed isolati mediante l'utilizzo di sfere magnetizzate.
2. *Ibridazione.* I frammenti amplificati e purificati vengono controllati in termini di quantità e qualità mediante *bioanalyzer*. Si procede quindi all'ibridazione del DNA con le sonde specifiche della libreria desiderata mediante termociclature e successivamente all'eliminazione dei frammenti di DNA non legati alle sonde di cattura mediante l'utilizzo di biglie magnetizzate con streptavidina.
3. *Indicizzazione.* I frammenti ottenuti, legati alle sonde, sono quindi sottoposti ad indicizzazione, mediante l'aggiunta di sequenze specifiche tramite PCR. Il prodotto di tale reazione viene nuovamente purificato e quantificato ed è quindi pronto per prendere parte alla generazione dei cluster mediante bridge PCR, su cui poi avverrà il vero e proprio sequenziamento.

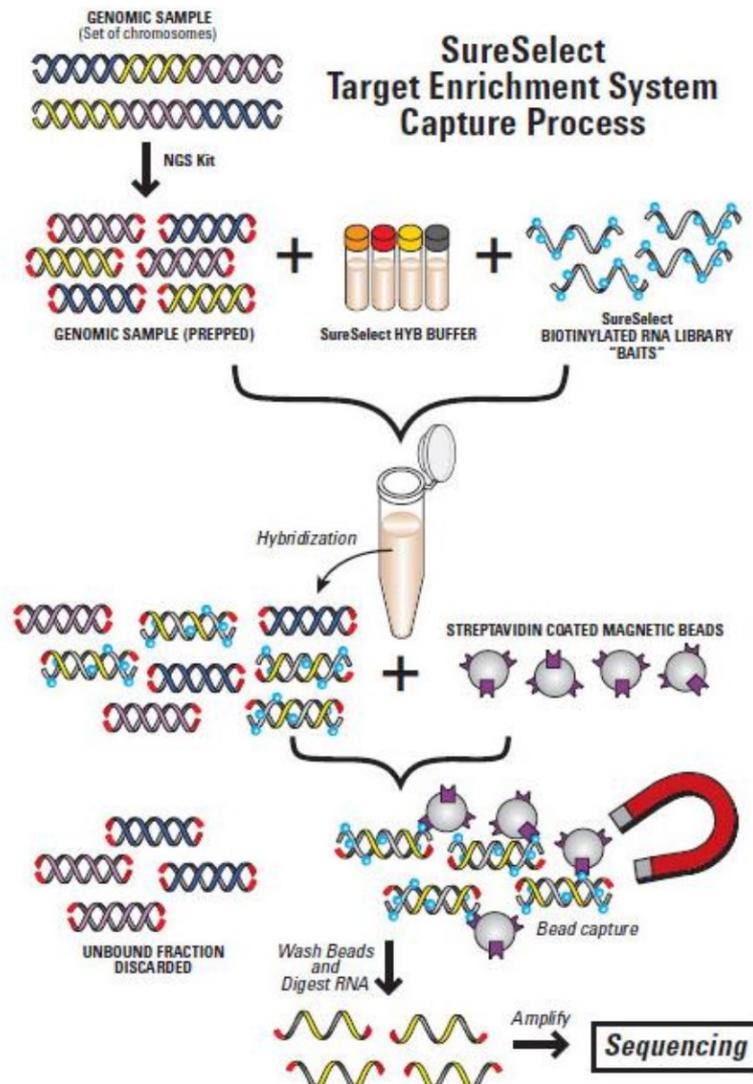


Figura 15. Preparazione delle librerie secondo metodo SureSelect

5.2.2 FASE DI GENERAZIONE DEI CLUSTERS

Una volta ottenute le librerie, il processo continua con la fase di generazione dei *clusters*, svolto su un dispositivo denominato *flow-cell*. La *flow-cell* è costituita da un vetrino avente diverse corsie (*lanes*), sulla cui superficie sono legati due differenti tipologie di oligonucleotidi, ognuna delle quale riconosce solo uno dei due *adapters* legati alle estremità dei frammenti. L'ibridazione è quindi consentita dal legame tra il primo tipo di oligonucleotide e uno degli adattatori; a questo punto una DNA-polimerasi estende la sequenza dell'oligonucleotide usando il frammento come stampo, fintantoché viene prodotta una molecola di DNA a doppio filamento che verrà quindi denaturata e lavata dal frammento originale. Il filamento ottenuto procede quindi a piagarsi "a ponte" in quanto l'adapter situato sull'estremità libera

va a riconoscere il secondo tipo di oligonucleotide legato alla superficie, all'interno dello stesso cluster: ha inizio così un processo di *solid-phase amplification* tramite *bridge PCR* (PCR "a ponte"). Una volta formato il ponte, la polimerasi duplica il filamento generando un ponte a doppio filamento, che viene poi denaturato, ottenendo quindi due filamenti (*forward e reverse*) saldamente ancorati alla *flow-cell*. Ciascun filamento fungerà a sua volta da stampo per i successivi cicli di *bridge PCR*, generando quindi clusters che contengono circa 1000 copie identiche dello stesso frammento target. In seguito i filamenti *reverse* vengono rimossi, in modo da lasciare solo le sequenze *forward*: a questo punto nella *flow-cell* sono rappresentati tutti i frammenti della libreria amplificati e divisi ordinatamente in clusters.

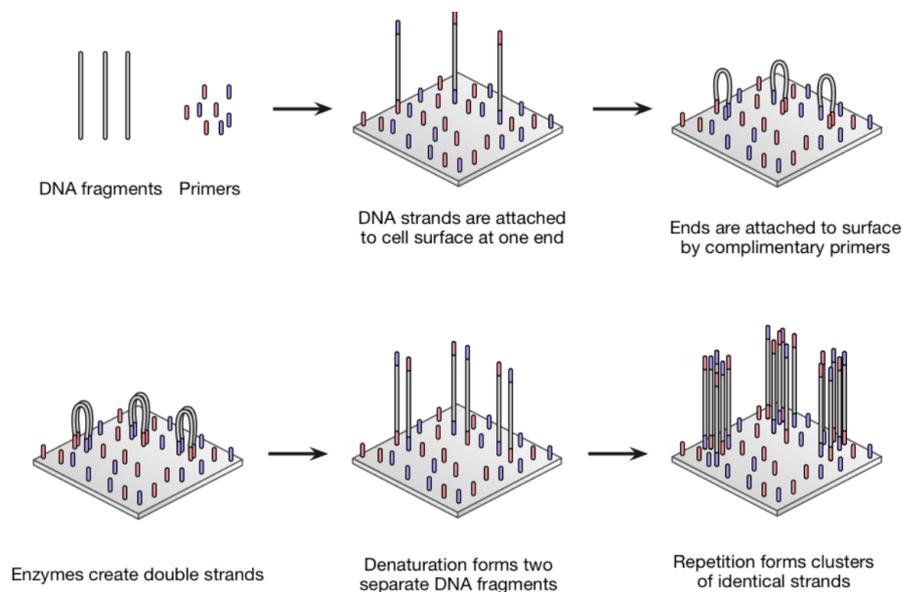


Figura 16. Illustrazione del processo di bridge PCR e clustering

5.2.3 FASE DI SEQUENZIAMENTO

In seguito alla costruzione della libreria è finalmente possibile procedere con la fase di sequenziamento vero e proprio, che nel contesto dello studio è stato svolto utilizzando strumentazione *Illumina*. Essa permette l'esecuzione di un sequenziamento per sintesi, il quale consiste nell'utilizzo di dNTPs marcati con quattro diversi fluorocromi, che sfruttano la chimica a terminazione di catena reversibile per l'identificazione delle singole basi del frammento di DNA in esame.

Il processo ha inizio con l'inserimento del primo primer, che si appaia con la sequenza specifica per il sequenziamento introdotta nel passaggio di preparazione dei frammenti di DNA: alla reazione vengono quindi aggiunti, oltre alla DNA polimerasi, i quattro dNTPs opportunamente modificati, cosicché può avvenire l'estensione del *primer* e la produzione della prima *read*. Ognuno dei 4 tipi di dNTP è legato ad un diverso marcatore fluorescente e all'estremità 3'-OH presenta un gruppo terminatore di catena reversibile, il quale impedisce che venga aggiunto più di un nucleotide alla volta; Il nucleotide aggiunto dipende ovviamente dalla sequenza del frammento, in accordo con il principio di complementarità delle basi (A-T; G-C). Ad ogni ciclo un laser eccita lo specifico gruppo fluorescente legato al dNTP incorporato determinando l'emissione di una fonte luminosa che ne permette l'identificazione sulla base della lunghezza d'onda dell'emissione e dell'intensità del segnale. Una volta avvenuta l'assegnazione della base, il gruppo terminatore presente al 3'-OH e il fluorocromo vengono rimossi, in modo tale da consentire l'appaiamento e la continuazione del processo con il sequenziamento della base successiva.

Completata la prima lettura, il nuovo filamento complementare prodotto viene rimosso ed è introdotto un particolare *primer* che permette il sequenziamento del *barcode*, così da indentificare il frammento sequenziato e facilitare il successivo passaggio di allineamento.

A questo punto rientrano in gioco gli oligonucleotidi fissati alla superficie alla fine del passaggio di generazione dei *clusters*: questi vengono infatti sbloccati venendo così permesso loro di legare l'estremità libera del filamento sequenziato, formando nuovamente una struttura a ponte (a singolo filamento). Viene introdotto quindi un *primer* che permette la lettura del secondo *barcode*; il frammento generato viene successivamente rilasciato.

A questo punto, la polimerasi opera un'estensione dell'oligonucleotide usando il filamento come stampo, formando nuovamente un ponte a doppio filamento, che viene successivamente linearizzato e privato dell'originale filamento *forward*, lasciando solamente quello *reverse*. In questo modo, reintroducendo la polimerasi, i 4 nucleotidi marcati e un secondo primer, può avere inizio la lettura anche del filamento *reverse*.

Questa metodologia prende il nome di "*paired-end sequencing*" e risulta maggiormente sensibile ed accurata rispetto a quella "*single-end*" poiché ciascun

frammento viene sequenziato da entrambe le estremità e in direzioni opposte. Tale espediente facilita notevolmente le successive operazioni di allineamento con la sequenza di riferimento, specialmente in riferimento al problema delle regioni costituite da una stessa base ripetuta più di 3 volte.

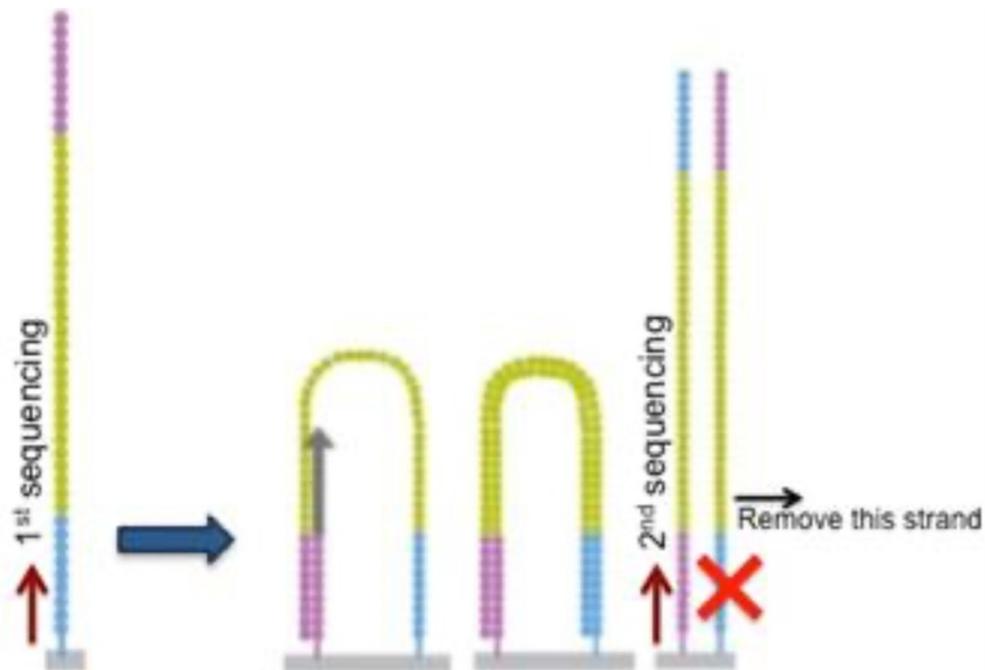


Figura 17. Rappresentazione schematica del processo di Paired-end sequencing

5.2.4 FASE DI ALLINEAMENTO E DI ANALISI DEI DATI

Le sequenze ottenute attraverso le procedure appena discusse vengono suddivise a seconda del genoma di provenienza utilizzando i *barcodes* unici che erano stati aggiunti durante la fase di preparazione dei frammenti.

Per quanto riguarda il processo di allineamento e analisi dei dati è stato invece utilizzato il software *SureCall*, prodotto dall'azienda statunitense *Agilent Technologies*: il programma è incaricato dell'elaborazione dei dati grezzi ottenuti tramite sequenziamento, dell'allineamento degli stessi col genoma umano di riferimento e con le sonde del pannello utilizzato, e della fase di chiamata delle varianti (*variant calling*).

Durante l'allineamento vengono selezionate le *reads* che presentano un certo grado di omologia di sequenza, in modo da radunarle nello stesso gruppo; le letture *forward* e *reverse* vengono quindi appaiate, creando così delle sequenze contigue che vengono poi allineate seguendo il genoma di riferimento. Come già accennato

prima, il sistema *paired-ends* permette un allineamento più accurato in quanto risulta in grado di risolvere eventuali ambiguità quali la presenza di zone ricche di una stessa base o comunque con molte ripetizioni.

Una volta allineate le sequenze, il software le confronta con il genoma di riferimento al fine di identificare le varianti, procedendo in automatico alla loro caratterizzazione. Segue poi la cosiddetta “annotazione” delle varianti, ovvero la ricerca di tutti i dati disponibili nei database e/o in letteratura. Il risultato finale consiste in una tabella riassuntiva delle varianti identificate ed annotate, con tanto di visualizzazione grafica delle *reads* allineate, della *depth of coverage* e del *coverage*.

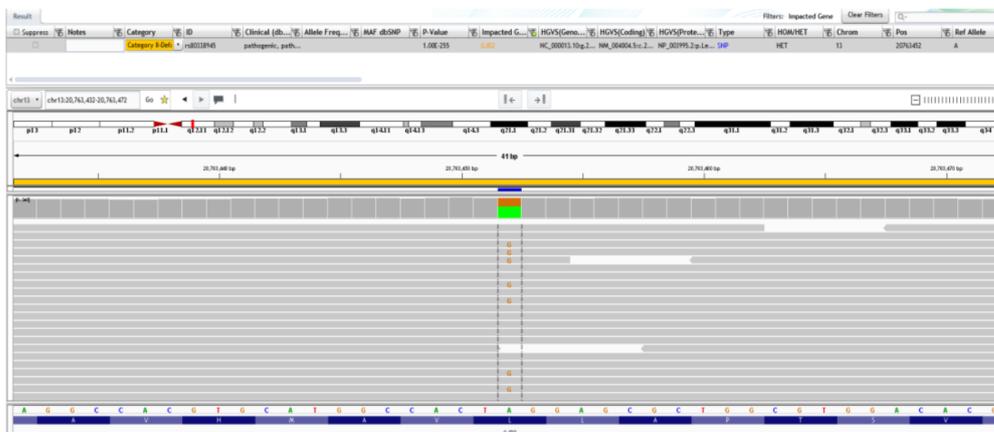


Figura 18. Schermata del software SureCall

5.3 ANALISI MOLECOLARE DI CONFERMA DELLE VARIANTI DI SEQUENZA IN *STR*: LR-PCR E SEQUENZIAMENTO

Una metodologia di indagine di secondo livello utilizzata nello studio in pazienti selezionati consiste nel sequenziamento NGS effettuato dopo preparazione delle librerie con protocollo *Nextera*, il quale è stato eseguito su materiale genomico amplificato tramite *Long Range PCR*.

Tradizionalmente, la *Long Range PCR* viene effettuata utilizzando una miscela di Taq DNA polimerasi combinata con una piccola quantità di DNA polimerasi *proofreading* (correttrice di bozze): la presenza della DNA polimerasi correttrice di bozze permette di prevenire l’insorgenza di errori nell’accoppiamento delle basi, garantendo quindi alla Taq DNA polimerasi di svolgere la propria funzione molto

più a lungo di quanto farebbe altrimenti, e risultando quindi nella produzione di segmenti di DNA più lunghi.

Per il presente studio, è stata utilizzata una coppia di primer specifici che permettessero l'amplificazione del solo gene, annullando o riducendo al minimo la contaminazione da parte dello pseudogene pSTRC nella fase di sequenziamento; i primer utilizzati sono quelli precedentemente pubblicati da Mandelker et al. (73).

Le caratteristiche operative che distinguono il protocollo *Nextera* risiedono nell'utilizzo di un'innovativa chimica basata sulle microsfere, che permette di eseguire sequenziamenti molto accurati sfruttando un'unica semplice fase di ibridazione tra le sonde e il campione biologico da analizzare. Nello specifico, viene effettuata una reazione denominata tagmentazione, che consiste nell'unione delle fasi di frammentazione del DNA e di ligazione degli adattatori: questa reazione, che richiede mediamente soltanto pochi minuti (circa 15), viene effettuata direttamente sulle microsfere, in maniera uniforme e mediante l'utilizzo di trasposoni, facendo in modo che l'arricchimento del materiale genetico d'interesse avvenga in maniera estremamente precisa e rapida, arrivando a richiedere circa 6 ore, contro le oltre 48 impiegate dai sistemi di NGS tradizionale.

Una volta ottenute le librerie da analizzare, è possibile quindi procedere con la fase di sequenziamento vero e proprio, che ricalca i passaggi tipici del NGS trattati precedentemente.

Nel presente studio, il regione target di sequenziamento era costituita dall'intero amplicone generato mediante LR-PCR, che comprendeva l'intero gene *STRC* (erano quindi incluse tutte le sequenze introniche).

La metodica è stata utilizzata per confermare la presenza delle varianti identificate mediante l'analisi NGS del pannello multigenico a carico del gene *STRC* e definirne quindi la zigosità; tale approccio era necessario per tutte le varianti localizzate in una regione genica non distinguibile dallo pseudogene.

In alcuni pazienti, la conferma delle varianti è stata effettuata mediante il tradizionale sequenziamento Sanger; in tal caso, è stato necessario effettuare una ulteriore PCR sull'amplicone ottenuto mediante LR-PCR (*nested PCR*), utilizzando primer che permettessero l'amplificazione della regione genica interessata dalla variante.

5.4 ANALISI BIOINFORMATICA DELLE CNV

Un'ulteriore procedura effettuata in tutti i pazienti facenti parte della coorte oggetto di studio è l'analisi bioinformatica delle CNV, essendo queste un fenomeno molto frequente nel contesto delle ipoacusie ad eziologia genetica. Sfruttando le potenzialità proprie del software *SureCall* si è infatti andati a ricercare, partendo dai dati ottenuti da NGS, la presenza di alterazioni geniche di tipo quantitativo sotto forma di variazioni del numero di copie, siano esse delezioni o duplicazioni.

Questo passaggio risulta possibile grazie all'elevata profondità di lettura che caratterizza il sequenziamento di nuova generazione: applicando determinati algoritmi che mettono a confronto tra loro i campioni analizzati con dei controlli apparentemente sani e verosimilmente privi di alterazioni a livello dei geni responsabili di ipoacusia, e le relative profondità di lettura, il programma è infatti in grado di suggerire la presenza di eventuali CNV nei geni considerati.

Il più grande limite di questa analisi consiste nel fatto che, non essendone ad oggi nota la sensibilità assoluta (in particolare in caso di CNV che interessino singoli esoni), non è possibile utilizzarla come strumento diagnostico, ma solo come ausilio per suggerire la presenza di alterazioni altrimenti non evidenziabili, che vanno comunque confermate con metodiche alternative, come ad esempio MLPA, PCR quantitativa, array-CGH e SNP-array. Nel contesto dello studio condotto, è stata selezionata MLPA come indagine di secondo livello atta a confermare la presenza di CNV.

5.5 ANALISI DI CONFERMA DELLE CNV: MLPA

Un'ulteriore tecnica di analisi utilizzata in pazienti selezionati MLPA (*Multiplex Ligation-dependent Probe Amplification*) è una metodica volta a rilevare la presenza di CNV con maggior accuratezza e precisione rispetto all'analisi bioinformatica, tanto da essere in grado di individuare alterazioni a carico di singoli esoni e di restituire le esatte coordinate genomiche delle stesse. Nel contesto dello studio MLPA è stata applicata in pazienti con sospetta delezione in corrispondenza del gene *STRC*, suggerita dall'analisi bioinformatica.

L'algoritmo procedurale tipico di MLPA risulta facilmente riproducibile grazie alla presenza nel mercato di specifici Kit, prodotti dall'azienda *MRC Holland*, e denominati SALSA MLPA Kit, ciascuno dei quali risulta specifico per una determinata regione genomica di interesse.

Nel contesto dello studio preso in esame, ci si è serviti di SALSA MLPA Kit specifici per il locus contenente i geni *STRC*, *CATSPER2* e *OTOA*, individuati con la numerazione P461.

Dal punto di vista operativo, MLPA è una metodica basata sul legame di sonde oligonucleotidiche in corrispondenza di sequenze target, seguito dall'amplificazione del prodotto di tale reazione. Costruendo in modo opportuno le sonde, è possibile riconoscere e successivamente amplificare contemporaneamente fino a 50 sequenze, sfruttando il principio della PCR multipla; vengono così ottenuti diversi frammenti di PCR, che sono quindi separati mediante elettroforesi capillare e visualizzati come un elettroferogramma. L'area e l'altezza di ogni picco sono proporzionali al numero di copie del locus genico che si sta analizzando: in caso di delezioni si osserva riduzione o assenza del picco corrispondente, viceversa in caso di amplificazioni.

Per un'analisi corretta, il profilo dei picchi di un campione deve essere sempre confrontato con quelli ottenuti da almeno due/tre campioni di DNA privo dell'anomalia in oggetto (controlli negativi) e di un campione in cui essa è presente (controllo positivo).

Le sonde presenti nel Kit SALSA MLPA P461 ricalcano sequenze presenti nei geni *STRC*, *OTOA* e *CATSPER2*, ma sono presenti anche sequenze appartenenti allo pseudogene *pSTRC*, importante confondente da tenere sempre in considerazione qualora si svolga una qualsiasi analisi genetica a carico del gene *STRC*. In particolare, per quanto riguarda *STRC*, si tratta di sei sequenze geniche differenti, ciascuna delle quali fa riferimento ad un differente esone, nello specifico esone 19, 20, 23, 24, 25, 28; per quanto riguarda *pSTRC* invece ci sono quattro sequenze, associate rispettivamente agli esoni 19, 20, 23, 28. Ciò significa che verranno individuate CNV di pertinenza degli esoni appena elencati, i quali rappresentano aree geniche a minor omologia di sequenza. Per quanto riguarda le aree geniche a maggior omologia di sequenza, non essendo disponibili sonde specifiche che differenzino gene da pseudogene, la tecnica MLPA non è ad oggi in grado di

valutare la presenza di eventuali CNV; è pertanto necessario utilizzare altre metodologie di analisi al fine di diagnosticare suddette anomalie.

5.6 SEQUENZIAMENTO DI TERZA GENERAZIONE MEDIANTE OXFORD NANOPORE

Il sequenziamento di terza generazione, effettuato mediante tecnologia *Oxford Nanopore*, costituisce una tecnica di indagine di terzo livello, impiegata nello studio nei pazienti in cui sono state sospettate conversioni geniche a carico di *STRC*.

La procedura applicata ha previsto l'utilizzo di una *flowcell* in cui sono presenti due compartimenti riempiti di soluzione ionica e separati da una membrana con incorporati migliaia di nanopori. Un costante bias di voltaggio produce una corrente ionica attraverso i nanopori, per cui al passaggio di una molecola di DNA attraverso il poro, si crea una perturbazione della corrente elettrica che viene registrata da un sensore come modificazioni in picoampere (pA): l'identificazione di ogni nucleotide è stata pertanto effettuata rilevando la caratteristica modulazione della corrente prodotta dalle diverse basi, che passano attraverso il nanoporo.

È stata preparata una libreria in cui il DNA è andato incontro all'aggiunta di una coda di adenina alle estremità, che ha permesso la formazione di terminazioni compatibili con l'adattatore di sequenziamento che contiene delle timidine. Successivamente, attraverso una reazione di ligazione, sono stati uniti gli adattatori ad entrambe le estremità dei frammenti di DNA in modo da poter sequenziare entrambi i filamenti e ridurre il *single strand bias*.

Il DNA a doppio filamento è stato quindi srotolato da una proteina motrice nei pressi del poro, attraverso il quale quindi passa un filamento alla volta dotato di velocità di traslocazione pari a circa 400 basi al secondo.

Questa tecnologia permette l'analisi in tempo reale di lunghi frammenti di DNA nativo evitando quindi il bias della PCR presente nelle tecnologie NGS attualmente in uso e permettendo di sequenziare anche regioni del genoma più complesse.

5.6.1 OXFORD NANOPORE: ARRICCHIMENTO MEDIANTE ADAPTIVE SAMPLING O LR-PCR

La tecnologia di *Oxford Nanopore* permette di analizzare il DNA nativo; nella maggior parte dei casi è però necessario effettuare un arricchimento per ottenere un numero di *reads* sufficiente da permettere un'analisi accurata.

I protocolli di arricchimento proposti dalla ditta sono: l'utilizzo di sonde Cas9, l'*adaptive sampling*, oppure il sequenziamento di ampliconi. Nel contesto dello studio sono stati messi a confronto gli ultimi due sistemi.

La tecnologia dell'*adaptive sampling* permette di selezionare in modo bioinformatico le regioni che devono essere sequenziate. Per avere un'adeguata cinetica di sequenziamento i protocolli Oxford Nanopore consigliano di identificare tra lo 0.1% e il 10% del genoma di riferimento. Per questo è stato creato un file in cui è stata riportata in formato FASTA la sequenza della regione di interesse comprendente circa 5Mb e contenente i geni *STRC*, *CATSPER2*, *STRCP1*.

In questo modo, grazie al sequenziamento in tempo reale, il software sarà in grado di capire se il frammento sequenziato rientra nella regione target. In tal caso continuerà a passare attraverso il poro, altrimenti ci sarà una depolarizzazione locale, il frammento verrà espulso e il poro si libererà per accogliere un nuovo frammento; questo permetterà di arricchire solo le regioni genomiche di interesse. Quindi non è più presente una fase di arricchimento mediante ibridazione con sonde tipica dei sistemi NGS di seconda generazione e soprattutto non è più necessario uno step di PCR.

Per quanto riguarda invece l'arricchimento mediante ampliconi, il procedimento è simile a quello di *Illumina* con la differenza che in questo caso non sono previsti step di frammentazione e di amplificazione per legare gli adattatori. Inoltre, il campione deve essere manipolato molto delicatamente senza mai usare il *vortex* in modo da mantenere un DNA ad alto peso molecolare che poi consentirà di produrre *reads* molto lunghe.

Una volta ottenuto l'amplicone mediante LR-PCR vengono attaccati gli adattatori e la proteina motrice mediante una reazione di ligazione. Vengono successivamente eseguiti una serie di passaggi di purificazione mediante biglie ed etanolo in modo da eliminare frammenti corti aspecifici.

5.7 INTERPRETAZIONE DELLE VARIANTI

Le varianti ottenute mediante l'impiego di tecniche di indagine di primo, secondo e terzo livello sono state interpretate in accordo con le linee guida stilate dall'American College of Medical Genetics (ACMG). (71) (72)

Sulla base di diversi criteri quali, elencati in ordine di importanza, la frequenza allelica, la tipologia di mutazione (troncante, delezione multiesonica, sostituzione missenso etc.), la conservazione evolutiva delle sequenze, le informazioni disponibili in letteratura, e la stima restituita dall'impiego di software interpretativi e predittivi, ciascuna variante individuata è stata classificata in base alla presunta patogenicità in:

- Varianti benigne/verosimilmente benigne: la probabilità che queste varianti siano causative del fenotipo osservato nel paziente è minore del 10%
- Varianti di significato incerto (VUS): la probabilità che esse siano patogene è compresa tra 10% e 90%. Si tratta di una categoria dentro la quale ricadono la maggioranza delle varianti individuate con le metodologie di sequenziamento disponibili, e risultano piuttosto complicate da refertare in quanto non è chiaro il loro ruolo biologico. Spesso in presenza di VUS è indicato effettuare la segregazione parentale in modo tale da verificarne la presenza nei genitori del probando e correlare il genotipo al fenotipo
- Varianti verosimilmente patogenetiche: la probabilità che siano causative del disturbo è superiore al 90%
- Varianti francamente patogenetiche: la probabilità che esse siano causative di malattia è superiore al 99%.

5.8 COMUNICAZIONE DELL'ESITO

Una volta esaurite le procedure diagnostiche attuabili al fine di decretare la tipologia di alterazione da cui è affetto ciascun paziente facente parte della coorte, è stato necessario comunicare agli stessi, ai tutori legali, e alle rispettive famiglie l'esito delle analisi. In particolare, si è provveduto a:

- Inviare il risultato dell'indagine al medico richiedente, consigliando, a prescindere dall'esito, una consulenza genetica per commentarlo;
- Convocare i pazienti per una consulenza genetica post-test, durante la quale è stato discusso il risultato dell'analisi. in caso di positività, sono state ridiscusse con la famiglia le caratteristiche della condizione clinica diagnosticata, la relativa prognosi e il rischio di ricorrenza.

Quando possibile, l'indagine molecolare è stata estesa anche ai genitori, così da poter confermare la fase allelica delle varianti identificate in *STRC*: in questo modo è stato possibile ottenere una migliore interpretazione delle varianti riscontrate e una più precisa stima del rischio di ricorrenza.

5.9 STUDIO DEL CAMPIONE E CONFRONTO CON LA LETTERATURA

Una volta individuate le mutazioni responsabili del quadro clinico patologico e comunicato ai soggetti l'esito delle indagini genetiche effettuate, è stata eseguita una valutazione più approfondita della documentazione clinica disponibile, tenendo conto del fatto che questa era molto più ampia nel caso di pazienti che hanno eseguito consulenza genetica pre-test presso il servizio di Genetica Medica dell'Azienda Ospedale Università di Padova. Sono state quindi raccolte e riportate in un documento Excel molteplici informazioni riguardanti i pazienti oggetto di studio, allo scopo di caratterizzare in maniera più dettagliata il fenotipo e definire un'eventuale correlazione genotipo-fenotipo. I vari parametri presi in considerazione includono:

- Codice identificativo della consulenza genetica effettuata
- Sesso
- Data di nascita
- Informazioni anamnestiche riguardo l'albero genealogico del soggetto
- Età e modalità di insorgenza dell'ipoacusia
- Bilateralità o unilateralità dell'ipoacusia
- Grado e andamento dell'ipoacusia, prendendo in considerazione, se disponibile, la forma della cura audiometrica
- Installazione di dispositivi protesici e età in cui ciò è stato eventualmente effettuato

- Esami di imaging neuroradiologici svolti nel paziente, con particolare attenzione a RM encefalo e TC rocche cerebrali
- Presenza di elementi sindromici

È stata successivamente effettuata una revisione della letteratura per determinare se i fenotipi clinici dei nostri pazienti fossero concordi con quanto già descritto in precedenti lavori.

6. RISULTATI

Per ciascuno dei pazienti appartenenti al campione d'interesse selezionato, costituito da 19 individui, sono state effettuate una serie di analisi, che, assieme alla documentazione clinica ed anamnestica disponibile, e l'eventuale segregazione parentale, hanno permesso di classificare i soggetti secondo vari parametri di riferimento, ottenendo così diverse informazioni riguardo la condizione patologica d'interesse, ossia l'ipoacusia ad eziologia genetica *STRC* correlata.

È doveroso inoltre specificare come, riguardo molti dei parametri analizzati, non sono disponibili dati riguardanti l'intera casistica, ed è stato quindi necessario ridurre il campione di riferimento. In particolare, molte informazioni cliniche ed anamnestiche non sono risultate accessibili, soprattutto per quanto riguarda i soggetti che hanno eseguito la consulenza esternamente all'Azienda Ospedaliera di Padova.

Va infine evidenziato che, per quanto il campione sia costituito da 19 individui, due di essi, identificati come paziente 18A e paziente 18B, sono sorella e fratello, e condividono le medesime varianti mutazionali: ciò significa che nell'analisi statistica di determinati parametri concernenti il genotipo i due pazienti sono stati considerati come un unico caso, definendo così un campione di 18 soggetti anziché 19.

I risultati ottenuti nel corso dello studio scientifico svolto riguardano sia il genotipo che il fenotipo dei pazienti facenti parte del campione, oltre che le informazioni anagrafiche degli stessi.

Per ogni individuo sono stati considerati le seguenti variabili, di cui alcune sono risultate tuttavia disponibili solo in un numero ristretto di soggetti.

- Sesso
- Anno di nascita
- Anno in cui sono state eseguite le indagini genetiche
- Svolgimento o meno della consulenza genetica presso AOPD
- Tecniche di indagine genetica utilizzate
- Varianti mutazionali individuate, almeno due per ogni soggetto
- Patogenicità di ciascuna variante
- Segregazione parentale

- Grado dell'ipoacusia
- Modalità d'insorgenza dell'ipoacusia (pre-linguale o post-linguale)
- Età d'insorgenza dell'ipoacusia in caso di forme post-linguali
- Installazione di protesi auricolare
- Presenza di elementi sindromici
- Informazioni anamnestiche di rilievo
- Familiarità per ipoacusia

6.1 TABELLE RIASSUNTIVE DEI RISULTATI

Le caratteristiche fenotipiche dei pazienti inclusi nello studio, la descrizione delle analisi genetiche a cui sono stati sottoposti, il genotipo patogenetico o potenzialmente patogenetico identificato a carico del gene *STRC* sono riassunti nelle seguenti tabelle:

- Tabella III: Tecniche analitiche utilizzate. Prende in considerazione per ciascun elemento del campione le tecniche di indagine genetica impiegate.
- Tabella IV: Varianti identificate in *STRC*, patogenicità e segregazione parentale. Le varianti individuate su ciascun allele, escluse le CNV, sono state riportate secondo le raccomandazioni della Human Genome Variation Society (HGVS) (74); tali varianti sono state descritte utilizzando la sequenza di riferimento NCBI NM_153700.2 relativa al trascritto di *STRC*. La classificazione in termini di patogenicità delle variazioni è stata effettuata secondo i criteri stabiliti dall'American College of Medical Genetics. (71) (72)
- Tabella V: Fenotipo audiologico dei pazienti. Vengono inoltre considerati il sesso di ciascun paziente, l'età al momento dell'analisi, la sede di esecuzione della consulenza pre-test, il grado di severità dell'ipoacusia, le modalità d'esordio dell'ipoacusia, e l'eventuale utilizzo di protesi acustiche. Molte voci della tabella risultano non note, e sono state pertanto indicate con la sigla NN (non noto).

Tabella III. Tecniche analitiche utilizzate.

ID paziente	Tecniche analitiche utilizzate
1	NGS + LRPCR + SANGER, CNV + MLPA
2	NGS + LRPCR + SANGER, CNV + MLPA
3	NGS, CNV + MLPA
4	NGS + LRPCR + Nextera, CNV + MLPA
5	NGS + LRPCR + Nextera, CNV + MLPA
6	NGS + LRPCR + Nextera, CNV + MLPA
7	NGS, CNV + MLPA
8	NGS, CNV + MLPA
9	NGS + LRPCR + Nextera + Nanopore, CNV + MLPA
10	NGS, CNV
11	NGS + LRPCR + Nextera, CNV
12	NGS + LRPCR + Nextera, CNV
13	NGS + LRPCR + Nextera, CNV
14	NGS, CNV
15	NGS, CNV
16	NGS, CNV + MLPA
17	NGS + LRPCR + Nextera, CNV
18 A	NGS + LRPCR + Nextera, CNV
18 B	NGS + LRPCR + Nextera, CNV

Note. NGS: sequenziamento di seconda generazione di pannello geni associato a ipoacusia; CNV: analisi bioinformatica delle Copy Number Variations con software *Agilent SureCall*; MLPA: MLPA con kit apposito per rilevare delezioni in *STRC*, *CATSPER2*, *OTOA*; LR-PCR+Nextera: sequenziamento NGS del gene *STRC* effettuato con protocollo di arricchimento *Nextera* su materiale genomico amplificato tramite *Long Range PCR*; Nanopore: sequenziamento di terza generazione *Oxford Nanopore* del gene *STRC*

Tabella IV. Varianti identificate in STRC, patogenicità e segregazione parentale

ID Paziente	Allele1	Patogenicità allele 1 (ACMG)	Allele 2	Patogenicità allele 2 (ACMG)	Segregazione parentale
1	c.2914C>T p.(Arg972Trp)	P	delezione multiesonica <i>STRC + CATSPER2</i>	P	Sì
2	delezione multiesonica <i>STRC</i>	P	c.2356delC p.(Leu786Cysfs*29)	LP	Sì
3	c.1631A>G p.(Tyr544Cys)	VUS	delezione multiesonica <i>STRC + CATSPER2</i>	P	No
4	delezione multiesonica <i>STRC + CATSPER2</i>	P	c.4128-2A>T	LP	Sì
5	c.3217C>T p.(Arg1073*)	P	delezione multiesonica <i>STRC + CATSPER2</i>	P	No
6	delezione multiesonica <i>STRC</i>	P	delezione multiesonica <i>STRC</i>	P	No
7	c.877G>A p.(Gly293Ser)	VUS	delezione multiesonica <i>STRC + CATSPER2</i>	P	No

Tabella IV. Varianti identificate in STRC, patogenicità e segregazione parentale

ID Paziente	Allele1	Patogenicità		Patogenicità allele 2 (ACMG)	Segregazione parentale
		allele 1 (ACMG)	Allele 2		
8	delezione multiesonica <i>STRC + CATSPER2</i>	P	delezione multiesonica <i>STRC + CATSPER2</i>	P	Sì
9	delezione multiesonica <i>STRC + CATSPER2</i>	P	conversione genica	P	No
10	delezione multiesonica <i>STRC + CATSPER2</i>	P	delezione multiesonica <i>STRC + CATSPER2</i>	P	No
11	c.4784G>A p.(Gly1595Asp)	VUS	c.4784G>A p.(Gly1595Asp)	VUS	No
12	c.4917_4918delinsCT p.(Leu1640Phe)	P	conversione genica in fase di accertamento	P	No
13	c.2494C>T p.(Arg832Trp)	LP	c.1368_1380del p.(Glu457Leufs*75)	P	Sì
14	delezione multiesonica <i>STRC + CATSPER2</i>	P	c.4917_4918delinsCT p.(Leu1640Phe)	P	No

Tabella IV. Varianti identificate in STRC, patogenicità e segregazione parentale

ID Paziente	Allele1	Patogenicità allele 1 (ACMG)	Allele 2	Patogenicità allele 2 (ACMG)	Segregazione parentale
15	c.4219-1G>A	P	conversione genica in fase di accertamento	P	No
16	delezione multiesonica <i>STRC + CATSPER2</i>	P	delezione multiesonica <i>STRC + CATSPER2</i>	P	No
17	c.1673G>A p.(Cys558Tyr)	VUS	c.1447C>T p.(Arg483*)	P	Si
18 A	c.1583C>T p.(Pro528Leu)	VUS	c.1002G>A p.(=)	VUS	No
18 B	c.1583C>T p.(Pro528Leu)	VUS	c.1002G>A p.(=)	VUS	No

Note. P: variante patogenetica; LP: variante verosimilmente patogenetica; VUS: variante di significato incerto.

Tabella V. Fenotipo audiologico dei pazienti

Paziente	Sesso	Età analisi	consulenza AOPD	Grado ipoacusia	Esordio ipoacusia	Protesi
1	F	13	Sì	moderato	post (6 anni)	No
2	M	1	Sì	lieve-moderato	pre	NN
3	M	13	Sì	severo	pre	Sì
4	F	32	Sì	lieve	post (10 anni)	sì, 11 anni
5	F	6	No	NN	NN	NN
6	F	69	Sì	lieve	post	NN
7	M	8	Sì	lieve-moderato	post (8 anni)	No
8	F	10	No	NN	NN	NN
9	F	51	No	NN	pre	NN
10	F	9	No	NN	pre	NN

Tabella V. Fenotipo audiologico dei pazienti

Paziente	Sesso	Età analisi	consulenza AOPD	Grado ipoacusia	Esordio ipoacusia	Protesi
11	F	20	Sì	lieve	pre	Sì
12	M	15	No	NN	NN	NN
13	M	6	No	NN	pre	NN
14	M	6	No	lieve-moderato	post (3 anni)	sì, 5 anni
15	F	15	No	NN	NN	NN
16	M	23	Sì	moderato	post	NN
17	M	4	No	lieve	pre	Consigliata
18 A	F	18	No	severo	NN	NN
18 B	M	14	No	severo	NN	NN

Note. NN: non noto; pre: esordio prelinguale; post: esordio postlinguale

6.2 DESCRIZIONE DEI RISULTATI

6.2.1 GENERALITA' DEI PAZIENTI

Il campione oggetto dello studio in questione è risultato essere costituito da 19 pazienti: tra questi, 9 soggetti erano di sesso maschile (47%), 10 di sesso femminile (53%).

L'età che ciascun soggetto aveva al momento delle analisi genetiche, svolte tra il 2017 e il 2022, oscilla tra 1 e 69 anni. È possibile distribuire la popolazione d'interesse in base alla fascia d'età in cui essi rientravano al momento dell'analisi: 2 pazienti avevano età compresa tra 1 e 5 anni (10%), 6 tra 6 e 10 anni (32%), 7 tra 11 e 20 anni (37%), 4 maggiore di 20 anni (21%).

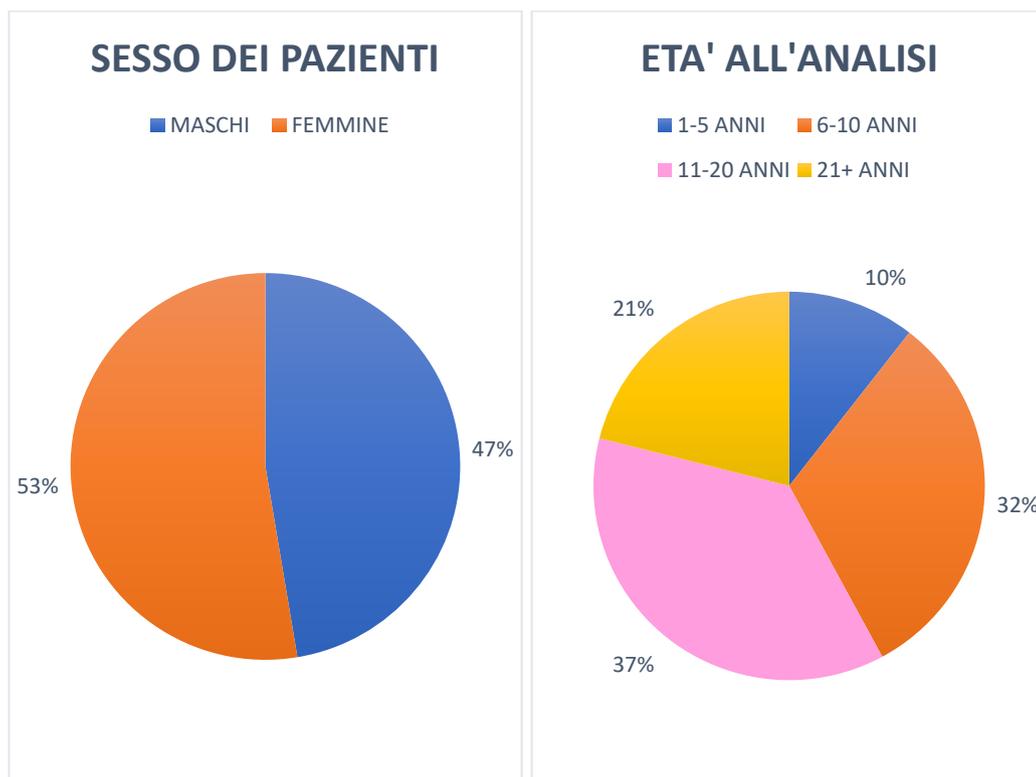


Figura 19. Areogramma rappresentante la distribuzione dei pazienti per sesso.

Figura 20. Areogramma rappresentante la distribuzione dei pazienti per età

6.2.2 SEDE DI CONSULENZA PRE-TEST

Nel contesto della popolazione oggetti di studio, costituita da 19 pazienti, 8 di questi (42%) avevano eseguito la consulenza genetica pre-test presso l'unità operativa complessa di Genetica Medica ed Epidemiologia dell'Azienda Ospedaliera di Padova, mentre i restanti 11 (58%) la avevano svolta in separata sede.

L'interesse del campione è stata tuttavia indagata dal punto di vista molecolare presso il Laboratorio di Genetica dell'Università degli studi di Padova.

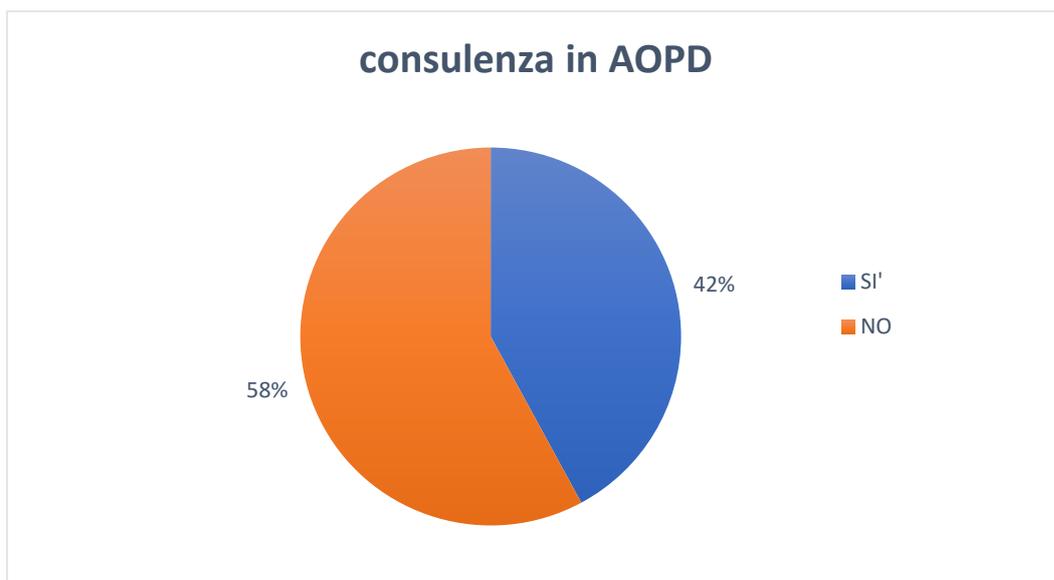


Figura 21. Areogramma rappresentante la distribuzione dei pazienti per luogo della consulenza pre-test.

6.2.3 TECNICHE ANALITICHE UTILIZZATE

Ciascun paziente oggetto di studio è stato indagato tramite molteplici metodologie di analisi genetica. In particolare, sono stati utilizzati:

- NGS di pannello multigenico per geni associati ad ipoacusia. Esso è stato utilizzato nella totalità dei pazienti (100%).
- Ricerca bioinformatica di CNV con software *Agilent SureCall* a partire dai dati ottenuti con NGS. Essa è stata utilizzata nella totalità dei pazienti (100%).
- MLPA con Kit apposito alla ricerca di CNV di pertinenza del locus genico a cui appartiene *STRC*. Essa è stata utilizzata in 10 pazienti (53%).
- LR-PCR associata a NGS con protocollo di arricchimento *Nextera* specifica per gene *STRC*. Essa è stata utilizzata in 10 pazienti (53%).

- LR-PCR associata a sequenziamento Sanger specifica per gene *STRC*. Essa è stato utilizzato in 2 pazienti (11%).
- Sequenziamento di terza generazione Oxford Nanopore specifico per il gene *STRC*. Esso è stato utilizzato in 1 paziente (6%).

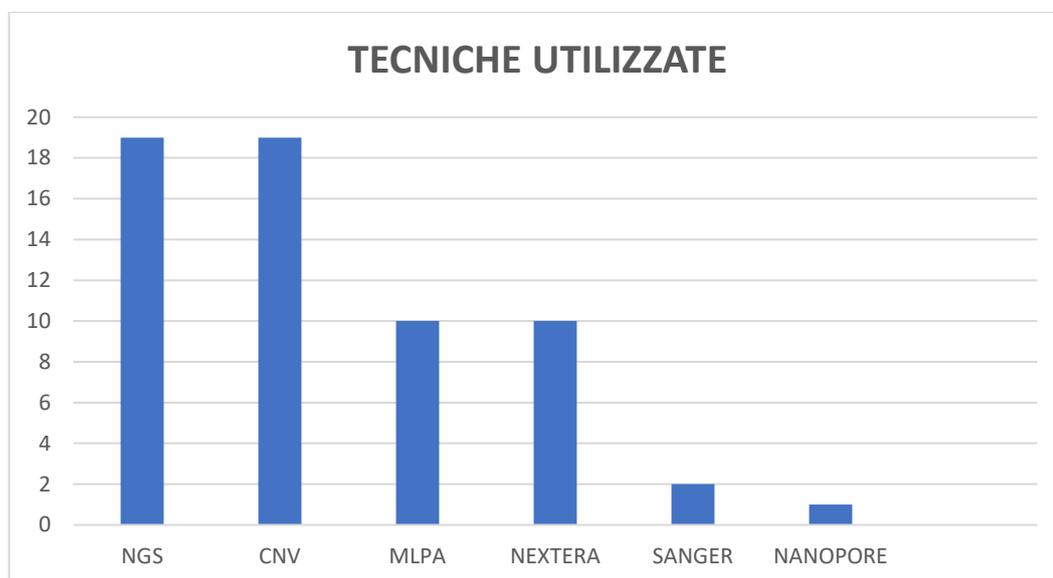


Figura 22. Istogramma rappresentante le varie tecniche di indagine genetica utilizzate in ciascun paziente.

I 13 pazienti in cui è stata rilevata mediante analisi bioinformatica la possibile presenza di CNV a carico del gene *STRC* sono sempre stati approfonditi con metodica MLPA, eccetto che in tre casi (pazienti 10,14 e 15) in cui ciò non è stato possibile per ragioni logistiche.

In ognuno di essi l'analisi MLPA ha restituito conferma di quanto ipotizzato, rivelando pertanto una concordanza del 100% tra analisi bioinformatica e MLPA.

6.2.4 VARIANTI MUTAZIONALI INDIVIDUATE

Per quanto riguarda questo parametro, si fa riferimento ad un campione costituito da 18 individui, poiché si considerano i pazienti 18A e 18B come un unico caso, essendo sorella e fratello portatori del medesimo genotipo.

Per ciascun paziente sono state individuate due varianti mutazionali del gene *STRC*, verosimilmente una per ciascun allele (non si ha invero la certezza di ciò nei pazienti presentanti varianti puntiformi in cui non è stata eseguita la segregazione parentale), pertanto l'analisi statistica verterà su 36 elementi. Tra le varianti individuate, si annoverano:

- Delezioni multiesoniche. Ne sono state individuate 16 (44%).
- Riarrangiamenti genici complessi, verosimilmente conversioni geniche. Ne sono state individuate verosimilmente 3 (8%).
- Sostituzioni nucleotidiche. Ne sono state individuate 15 (42%).
- Delezioni nucleotidiche. Ne è stata individuata 2 (6%).

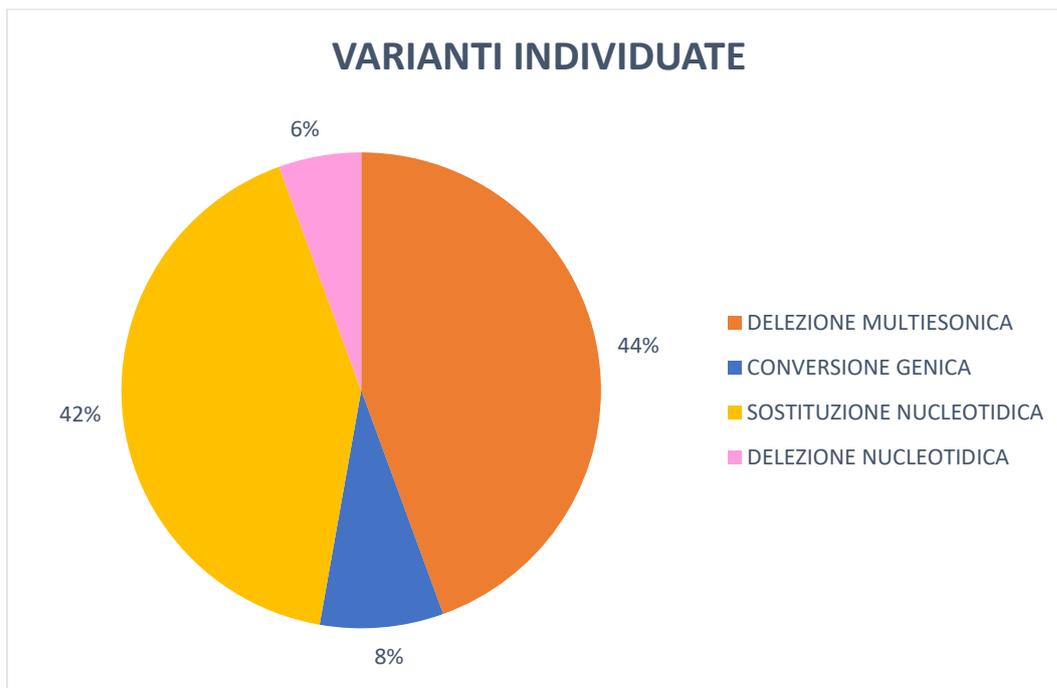


Figura 23. Areogramma rappresentante la distribuzione delle varianti mutazionali individuate.

Le diverse varianti individuate possono essere a loro volta suddivise in molteplici sottoclassi, in base a differenti parametri quali le regioni geniche interessate e gli effetti che esse hanno sulla sequenza proteica codificata dal gene.

Più dettagliatamente, le delezioni multiesoniche possono essere sottoclassificate in base alle regioni coinvolte in:

- Delezioni multiesoniche interessanti esclusivamente *STRC*. Essa rappresentano 3 delle 16 delezioni multiesoniche individuate (19%)
- Delezioni multiesoniche interessanti *STRC* e *CATSPER2*. Essa rappresentano le restanti 13 delezioni multiesoniche individuate (81%)

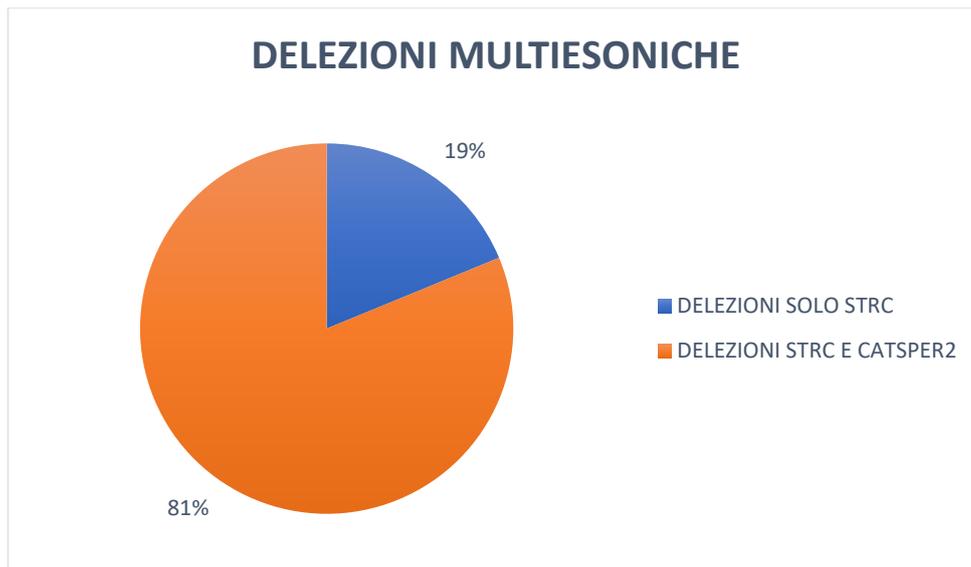


Figura 24. Areogramma rappresentante la caratterizzazione delle delezioni multiesoniche individuate.

Le piccole mutazioni (sostituzioni e delezioni nucleotidiche) possono essere invece sottoclassificate, in base alla regione genica in corrispondenza della quale si localizzano, in:

- Mutazioni introniche. Si tratta di varianti che interessano regioni non codificanti del gene e determinano verosimilmente alterazioni nel processo di *splicing*. Esse rappresentano 2 delle 17 mutazioni individuate (12%)
- Mutazioni esoniche. Si tratta di varianti che interessano le regioni codificanti del gene. Esse rappresentano le restanti 15 mutazioni individuate (88%). Va invero specificato come alcune varianti esoniche vadano verosimilmente ugualmente ad agire sul processo di *splicing*, qualora siano localizzate in corrispondenza della giunzione introne-esone: nella casistica in esame, esse costituiscono 2 elementi (12%).

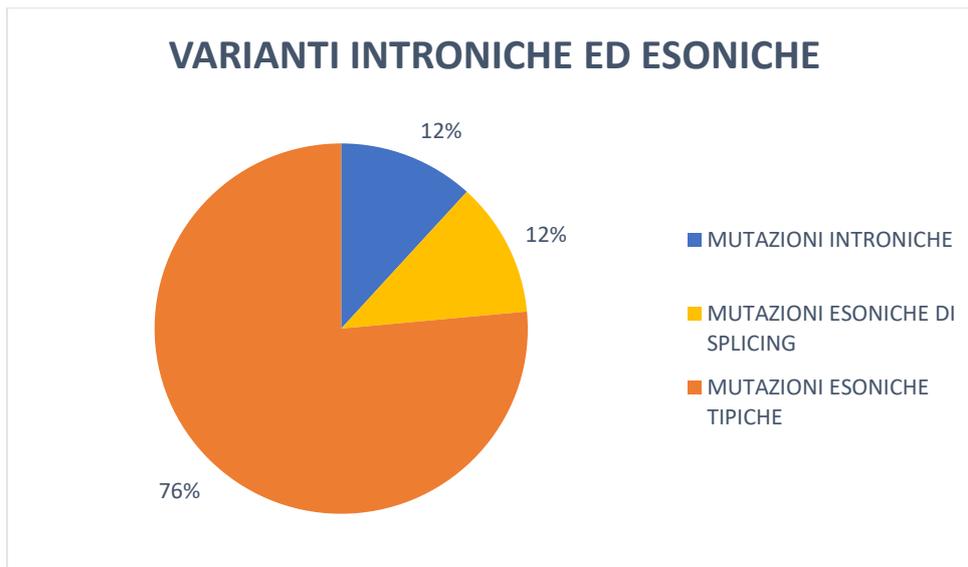


Figura 25. Areogramma rappresentante la caratterizzazione delle piccole variazioni nucleotidiche introniche ed esoniche individuate.

Le piccole mutazioni esoniche possono infine essere suddivise in funzione del loro predetto effetto sulla sequenza aminoacidica in:

- Sostituzioni missenso, qualora modifichino la sequenza aminoacidica in corrispondenza di un singolo elemento. Esse rappresentano 12 tra le 15 sostituzioni individuate (80%)
- Sostituzioni nonsense, qualora determinino l'introduzione precoce di un codone di stop, causando la produzione di una proteina tronca. Esse rappresentano 2 tra le 15 sostituzioni individuate (13%)
- Sostituzioni sinonime, qualora apparentemente non modifichino la sequenza aminoacidica; risultano patologiche nel caso in cui alterino il processo di *splicing*. Esse rappresentano 1 tra le 15 sostituzioni individuate (7%).

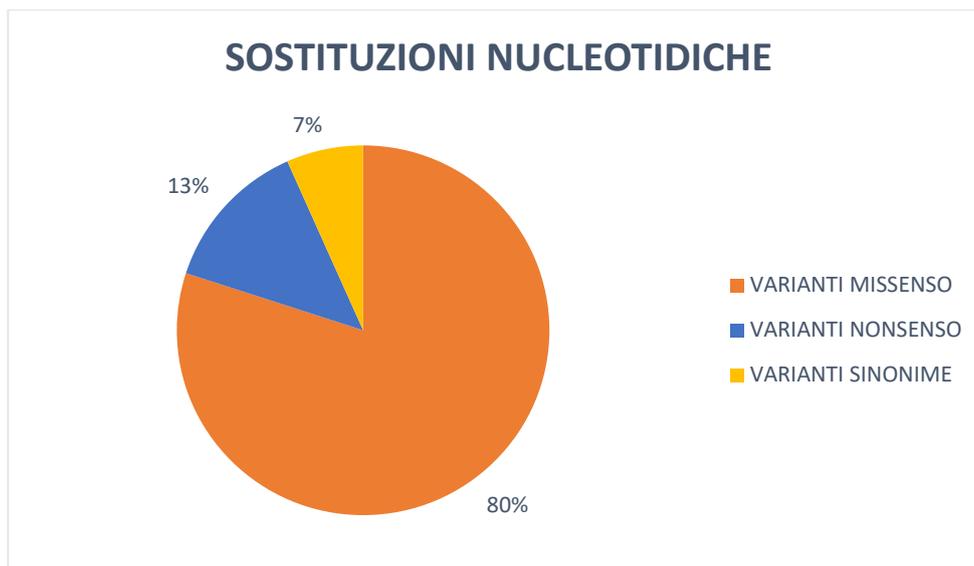


Figura 26. Areogramma rappresentante il predetto effetto a livello proteico delle piccole variazioni nucleotidiche esoniche.

Per quanto riguarda la patogenicità delle mutazioni, sono state considerate ai fini dello studio varianti che, in accordo con quanto stabilito dall'impiego di software predittivi, possono risultare patogenetiche (probabilità di essere causative di malattia maggiore o uguale a 99%), verosimilmente patogenetiche (probabilità di essere causative di malattia compresa tra 10% e 99%), oppure di significato incerto (probabilità di essere causative di malattia compresa tra 10% e 99%).

- Varianti patogenetiche. Sono risultate patogenetiche 27 tra le 36 varianti oggetto di studio (75%)
- Varianti verosimilmente patogenetiche. Sono risultate verosimilmente patogenetiche 3 tra le 36 varianti oggetto di studio (8%)
- Varianti di significato incerto (VUS). Sono risultate di significato incerto 6 tra le 36 varianti oggetto di studio (17%).

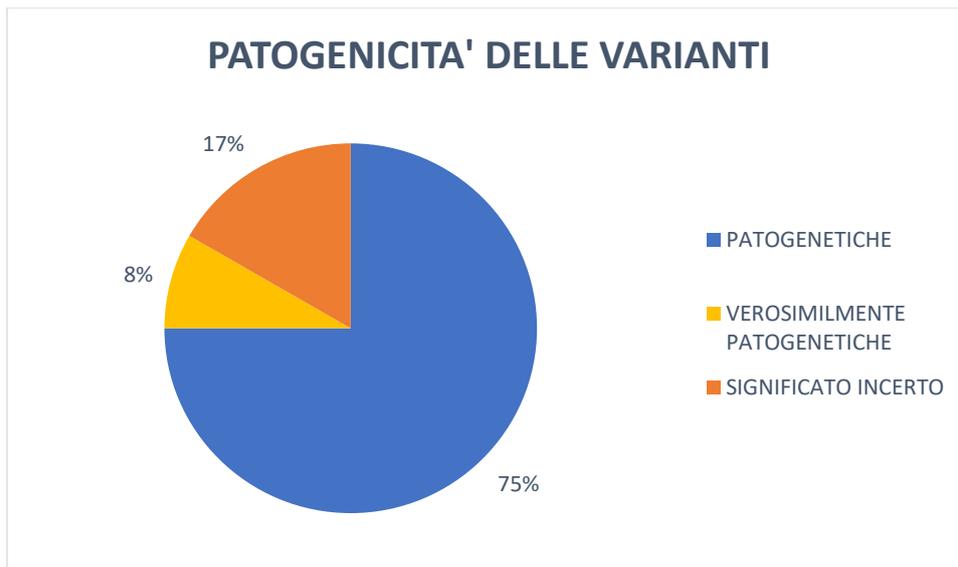


Figura 27. Areogramma rappresentante la patogenicità delle varianti individuate.

6.2.5 SEGREGAZIONE PARENTALE

La segregazione parentale consiste nell'estendere l'analisi ai genitori del probando in modo tale da poter definire le modalità di trasmissione delle varianti individuate e determinare l'appartenenza allelica di ciascuna delle stesse.

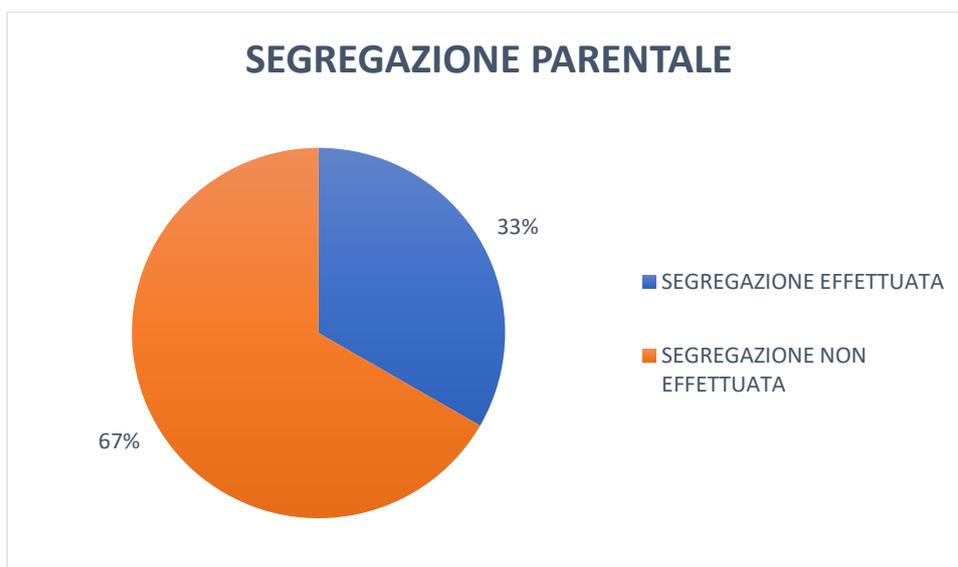


Figura 28. Areogramma rappresentante lo svolgimento della segregazione parentale nel campione analizzato.

Nel contesto dello studio effettuato, è stato possibile effettuare la segregazione parentale in solo 6 casi su 18 (33%. i pazienti 18A e 18B vengono nuovamente considerati come un unico caso a sé stante in quanto fratelli). Nei restanti 12 la segregazione parentale non è stata effettuata.

6.2.6 ETA' D'ESORDIO DELL'IPOACUSIA

I dati relativi a questo parametro sono stati reperiti in 13 soggetti: tra questi, 7 (37%) hanno presentato un'insorgenza dell'ipoacusia di tipo prelinguale, i restanti 6 (32%) di tipo postlinguale. Tra le ipoacusie ad esordio postverbale, l'età di insorgenza è stata individuata in 4 pazienti, che sono stati riconosciuti come ipoacusici rispettivamente a 3, 6, 8 e 10 anni.

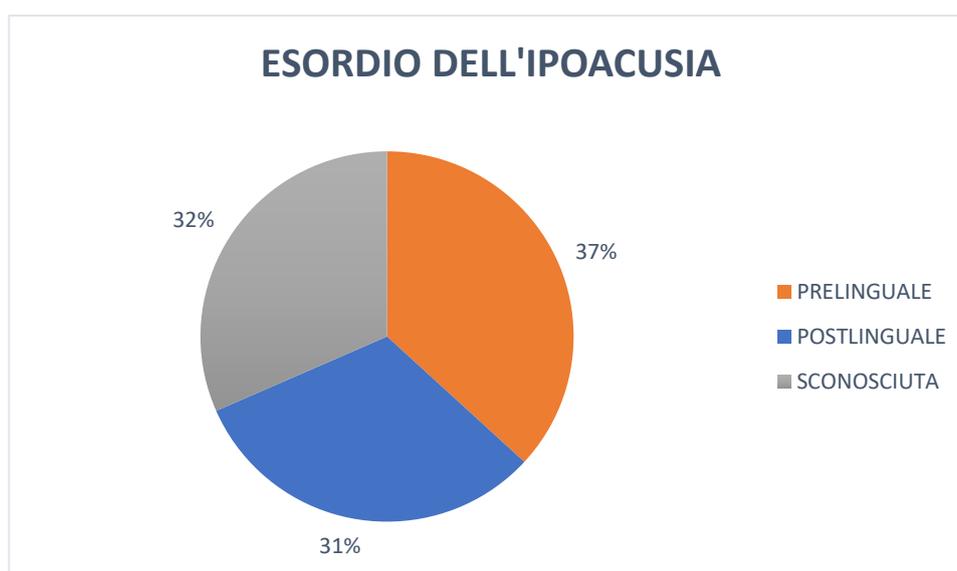


Figura29. Areogramma rappresentante la distribuzione dei pazienti in base all'esordio dell'ipoacusia.

6.2.7 GRADO DI IPOACUSIA

In 12 soggetti sono stati raccolti dati riguardanti la severità dell'ipoacusia: essa è risultata lieve in 4 di essi (21%), lieve-moderata in 3 (16%), moderata in 2 (11%), severa in 3 (21%). Si sottolinea come nel paziente 14, classificato come portatore di ipoacusia di grado lieve-moderato, si è osservata in realtà la presenza di ipoacusia di grado lieve in corrispondenza delle frequenze acute e lieve-moderato in corrispondenza delle frequenze gravi.

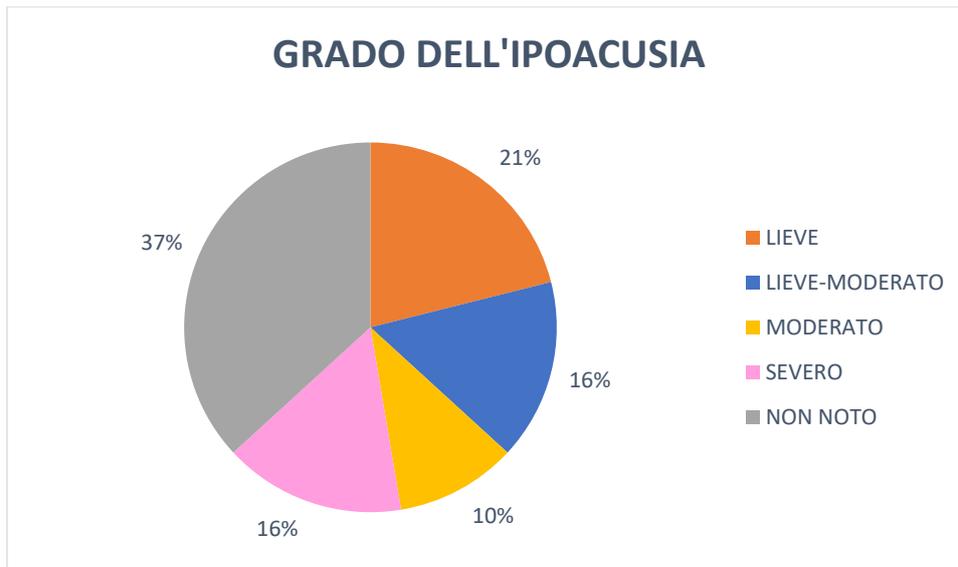


Figura 30. Areogramma rappresentante la distribuzione dei pazienti in base al grado dell'ipoacusia.

6.2.8 PROTESIZZAZIONE

Informazioni riguardo l'eventuale utilizzo di protesi acustiche sono risultate disponibili in 9 pazienti e in un numero ancor più ristretto sono state reperite informazioni aggiuntive quali l'età alla quale sono state prescritte per la prima volta o i risultati funzionali delle stesse. Nel complesso, 5 pazienti (26%) utilizzavano protesi acustiche; l'età della prima prescrizione è nota per due di questi, rispettivamente 5 e 11 anni. Tre pazienti (16%) non utilizzavano protesi, di cui un soggetto per volere dei genitori nonostante le indicazioni mediche. In un ulteriore paziente infine è stata indicata la protesizzazione ma non sono disponibili ulteriori dati a riguardo. Il trattamento ha dato generalmente buoni risultati in termini di outcome audiologico, eccetto che in un individuo.



Figura 31. Areogramma rappresentante la distribuzione dei pazienti in base alla protesizzazione.

6.2.9 DESCRIZIONE DI ULTERIORI PARAMETRI FENOTIPICI

Oltre ai parametri fenotipici appena analizzati ed inseriti nella tabella V, sono stati presi in considerazione, ove possibile, ulteriori caratteristiche cliniche ed anamnestiche atte a caratterizzare con maggior precisione il fenotipo di ciascun individuo oggetto di studio.

- Lateraltà del deficit uditivo. In tutti i pazienti ove chiaramente specificato (12, 63%), l'ipoacusia è risultata essere bilaterale, in accordo con quanto già noto dalla letteratura su *STRC*.
- Studio neuroradiologico associato. Sono stati reperiti gli esiti degli esami di imaging neuroradiologico effettuati in 2 soggetti appartenenti alla coorte di studio. Nello specifico, sono stati eseguiti RM encefalo e TC rocche mastoidi, risultate entrambe nella norma in entrambi i casi.
- Informazioni anamnestiche riguardanti l'albero genealogico dei soggetti. Si è riusciti ad accedere all'anamnesi familiare di 11 tra i soggetti indagati nello studio (58%): di questi, in 4 individui (21%) non è stato evidenziato nulla di rilevante, mentre nei restanti 7 (37%) è stata riscontrata familiarità per ipoacusia, nonostante in 3 di questi le manifestazioni cliniche e le modalità di insorgenza della condizione nei familiari rimanda ad un'eziologia acquisita o comunque multifattoriale piuttosto che monogenica.

6.3 GENOTIPO E FENOTIPO DEI PAZIENTI IN ANALISI

6.3.1 PAZIENTE 1

Si tratta di un soggetto di sesso femminile, sottoposta ad analisi genetica all'età di 13 anni, dopo aver eseguito consulenza genetica in AOPD.

È stata sottoposta alle seguenti tecniche di indagine: NGS di pannello multigenico per geni associati ad ipoacusia, ricerca bioinformatica di CNV con software *Agilent SureCall* a partire dai dati ottenuti con NGS, LR-PCR associata a sequenziamento Sanger specifica per gene *STRC*, MLPA con Kit apposito alla ricerca di CNV di pertinenza del locus genico a cui appartiene *STRC*.

Esse hanno permesso di individuare le seguenti variazioni:

- NM_153700.2: c.2914C>T p.(Arg972Trp). Essa consiste in una sostituzione nucleotidica puntiforme di una citosina con una timina in posizione 2914 (si evidenzia come, nel conteggio dei nucleotidi, si considerino soltanto quelli appartenenti a regioni esoniche), che determina nella proteina codificata dal gene *STRC* una sostituzione aminoacidica di un'arginina con un triptofano in posizione 792.

In accordo con quanto stabilito secondo i criteri di classificazione ACMG, si tratta di una mutazione patogenetica.

- Delezione multiesonica interessante sia il gene *STRC* che il gene *CATSPER2*.

In accordo con quanto stabilito secondo i criteri di classificazione ACMG, si tratta di una mutazione patogenetica.

È stata inoltre effettuata la segregazione parentale, che ha permesso di definire la tipologia di trasmissione delle variazioni individuate e l'appartenenza delle stesse a differenti alleli: la sostituzione puntiforme è risultata essere di origine paterna, la delezione multiesonica di origine materna.

determina il *frameshift* del trascritto, ossia uno slittamento della cornice di lettura, con inserimento di un codone di stop prematuro.

In accordo con quanto stabilito secondo i criteri di classificazione ACMG, si tratta di una mutazione verosimilmente patogenetica.

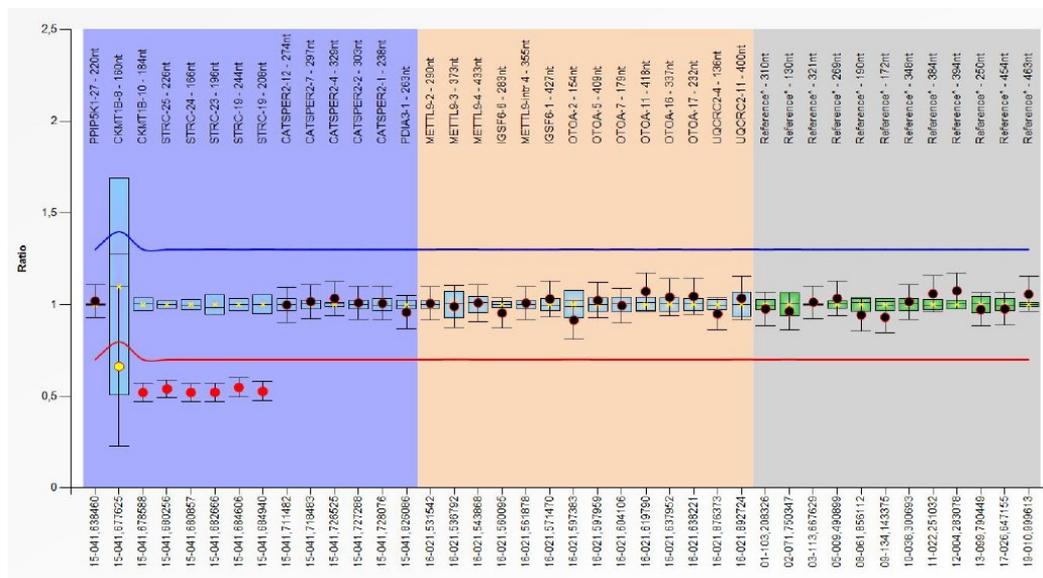


Figura 33. Risultato dell'indagine MLPA eseguita sul paziente 2. Si può osservare la presenza di una delezione in eterozigosi, interessante il gene STRC.

È stata inoltre effettuata la segregazione parentale, che ha permesso di definire la tipologia di trasmissione delle variazioni individuate e l'appartenenza delle stesse a differenti alleli: la delezione multiesonica è risultata essere di origine paterna, la delezione puntiforme di origine materna.

Per quanto riguarda il fenotipo, il paziente manifestava ipoacusia neurosensoriale bilaterale di grado lieve-moderato, interessante in particolar modo le frequenze centrali.

L'esordio è stato prelinguale.

All'anamnesi familiare è emersa la presenza di neuropatia uditiva retrococleare in zia materna nata prematura, insorta in seguito a enterocolite necrotizzante e relativa terapia antibiotica.

6.3.3 PAZIENTE 3

Si tratta di un soggetto di sesso maschile, sottoposto ad analisi genetica all'età di 13 anni, dopo aver eseguito consulenza genetica in AOPD.

È stata sottoposto alle seguenti tecniche di indagine: NGS di pannello multigenico per geni associati ad ipoacusia, ricerca bioinformatica di CNV con software *Agilent SureCall* a partire dai dati ottenuti con NGS, MLPA con Kit apposito alla ricerca di CNV di pertinenza del locus genico a cui appartiene *STRC*.

Esse hanno permesso di individuare le seguenti variazioni:

- Delezione multiesonica interessante sia il gene *STRC* che il gene *CATSPER2*.

In accordo con quanto stabilito secondo i criteri di classificazione ACMG, si tratta di una mutazione patogenetica.

- NM_153700.2: c.1631A>G p.(Tyr544Cys). Essa consiste in una sostituzione nucleotidica puntiforme di una adenina con una guanina in posizione 1631, che determina nella proteina codificata dal gene *STRC* una sostituzione aminoacidica di una tirosina con una cisteina in posizione 544. In accordo con quanto stabilito secondo i criteri di classificazione ACMG, si tratta di una variante di significato incerto (VUS)

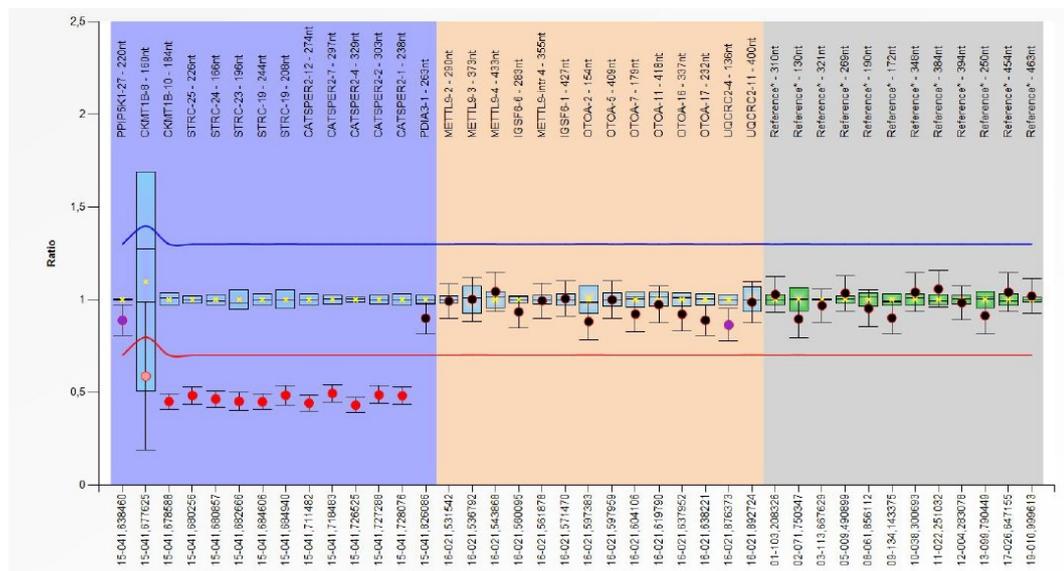


Figura 34. Risultato dell'indagine MLPA eseguita sul paziente 3. Si può osservare la presenza di una delezione in eterozigosi, interessante i geni *STRC* e *CATSPER2*.

Nel paziente in esame non è stata eseguita la segregazione parentale, pertanto non è possibile determinare le modalità di trasmissione ereditaria delle variazioni individuate.

Per quanto riguarda il fenotipo, il paziente manifestava ipoacusia neurosensoriale bilaterale di grado severo.

L'esordio è stato pre-linguale, congenito, individuato già allo screening neonatale effettuato con otoemissioni acustiche OAE.

È stata eseguita protesizzazione, accompagnata da trattamento logopedico.

Sono state eseguite risonanza magnetica dell'encefalo e tomografia computerizzata delle rocche mastoidee, risultate entrambe nella norma.

6.3.4 PAZIENTE 4

Si tratta di un soggetto di sesso femminile, sottoposta ad analisi genetica all'età di 32 anni, dopo aver eseguito consulenza genetica in AOPD.

È stata sottoposta alle seguenti tecniche di indagine: NGS di pannello multigenico per geni associati ad ipoacusia, ricerca bioinformatica di CNV con software *Agilent SureCall* a partire dai dati ottenuti con NGS, LR-PCR associata a NGS con protocollo di arricchimento *Nextera* specifica per gene *STRC*, MLPA con Kit apposito alla ricerca di CNV di pertinenza del locus genico a cui appartiene *STRC*. Esse hanno permesso di individuare le seguenti variazioni:

- Delezione multiesonica interessante sia il gene *STRC* che il gene *CATSPER2*.

In accordo con quanto stabilito secondo i criteri di classificazione ACMG, si tratta di una mutazione patogenetica.

- NM_153700.2: c.4128-2A>T. Essa consiste in una sostituzione nucleotidica puntiforme di una adenina con una timina in posizione 4128-2. Ciò implica l'interessamento di una regione intronica, che non essendo numerata in maniera convenzionale, poiché la numerazione nucleotidica considera specificatamente le sequenze esoniche, viene indicata in base alla distanza dal nucleotide esonico più prossimo: in questo caso si ha una sostituzione in corrispondenza di un nucleotide intronico che dista due basi dall'esone a valle. Verosimilmente, vista la posizione prossima alla giunzione introne-esone, questa variante costituisce una mutazione che, pur non interessando una porzione codificante del gene, ne altera la funzionalità andando ad inficiare il corretto processo di *splicing*.

In accordo con quanto stabilito secondo i criteri di classificazione ACMG, si tratta di una mutazione verosimilmente patogenetica.

È stata inoltre effettuata la segregazione parentale, che ha permesso di definire la tipologia di trasmissione delle variazioni individuate e l'appartenenza delle stesse

a differenti alleli: la delezione multiesonica è risultata essere di origine paterna, la sostituzione puntiforme di origine materna.

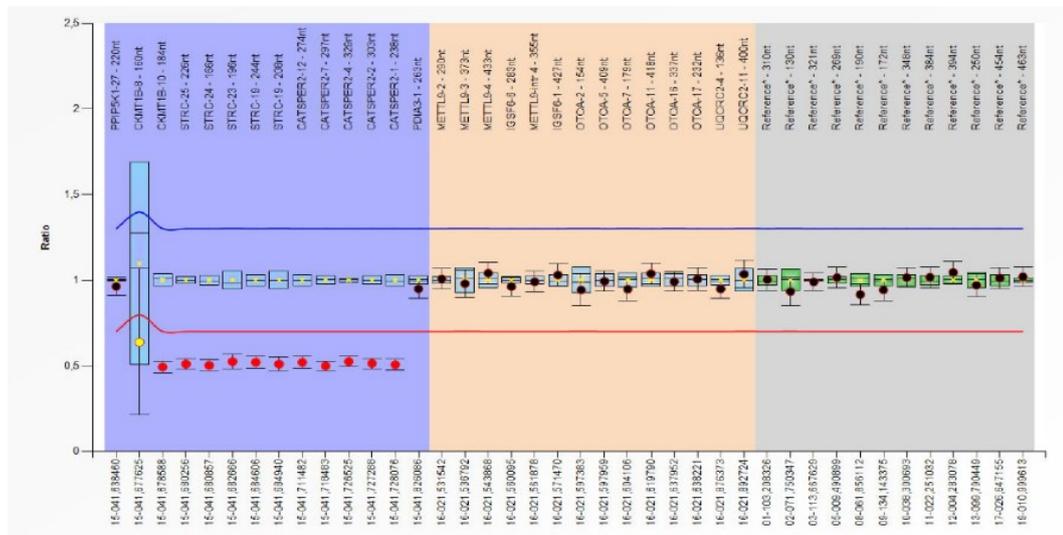


Figura 35. Risultato dell'indagine MLPA eseguita nel paziente 4. Si può osservare la presenza di una delezione in eterozigosi, interessante i geni STRC e CATSPER2.

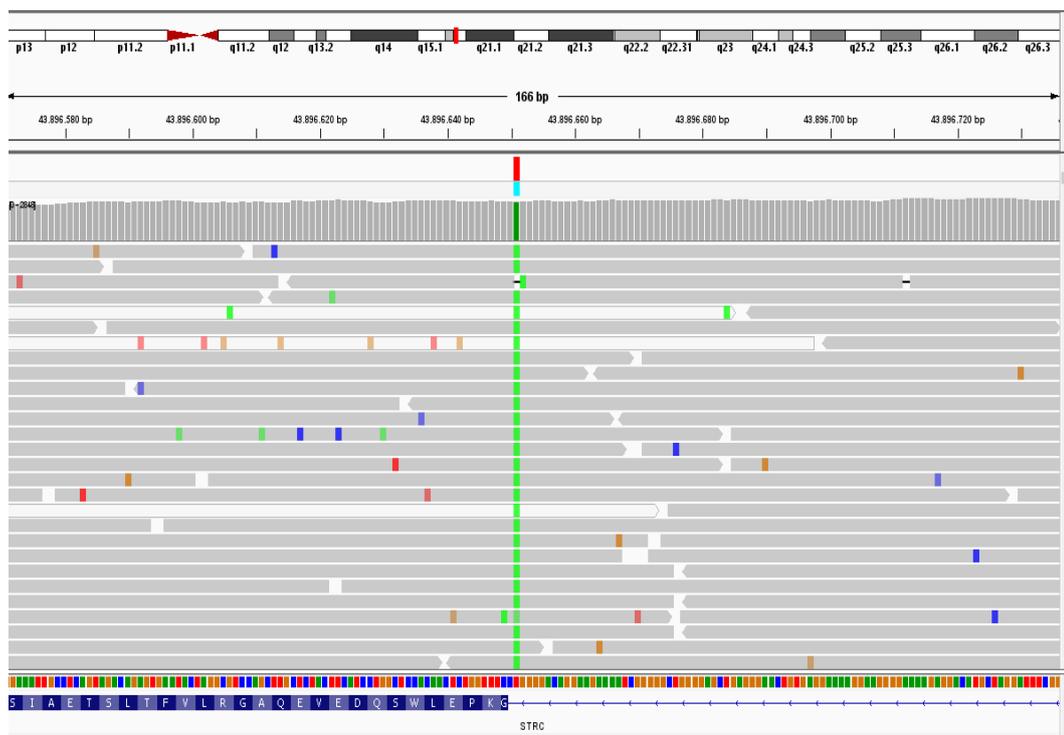


Figura 36 Risultato del sequenziamento NGS effettuato con protocollo di arricchimento Nextera eseguito nel paziente 4. Si può osservare la presenza di una sostituzione nucleotidica in apparente omozigosi: in realtà si tratta di una condizione di emizigosi in quanto uno dei due alleli risulta deleta.

Per quanto riguarda il fenotipo, la paziente manifestava ipoacusia neurosensoriale bilaterale di grado lieve.

L'esordio è stato post-linguale, verificatosi all'età di 10 anni. È stata eseguita protesizzazione all'età di 11 anni.

6.3.5 PAZIENTE 5

Si tratta di un soggetto di sesso femminile, sottoposta ad analisi genetica all'età di 6 anni, dopo aver eseguito consulenza genetica in sede esterna ad AOPD.

È stata sottoposta alle seguenti tecniche di indagine: NGS di pannello multigenico per geni associati ad ipoacusia, ricerca bioinformatica di CNV con software *Agilent SureCall* a partire dai dati ottenuti con NGS, LR-PCR associata a NGS con protocollo di arricchimento *Nextera* specifica per gene *STRC*, MLPA con Kit apposito alla ricerca di CNV di pertinenza del locus genico a cui appartiene *STRC*. Esse hanno permesso di individuare le seguenti variazioni:

- Delezione multiesonica interessante sia il gene *STRC* che il gene *CATSPER2*.

In accordo con quanto stabilito secondo i criteri di classificazione ACMG, si tratta di una mutazione patogenetica.

- NM_153700.2: c.3217C>T p.(Arg1073*). Essa consiste in una sostituzione nucleotidica puntiforme di una citosina con una timina in posizione 3217, che determina nella proteina codificata dal gene *STRC* l'inserimento di un codone di stop prematuro in posizione 1073 a sostituzione di un'arginina.

In accordo con quanto stabilito secondo i criteri di classificazione ACMG, si tratta di una mutazione patogenetica.

Nella paziente in esame non è stata eseguita la segregazione parentale, pertanto non è possibile determinare le modalità di trasmissione ereditaria delle variazioni individuate.

Non è stato possibile risalire ad informazioni cliniche ed anamnestiche approfondite a riguardo.

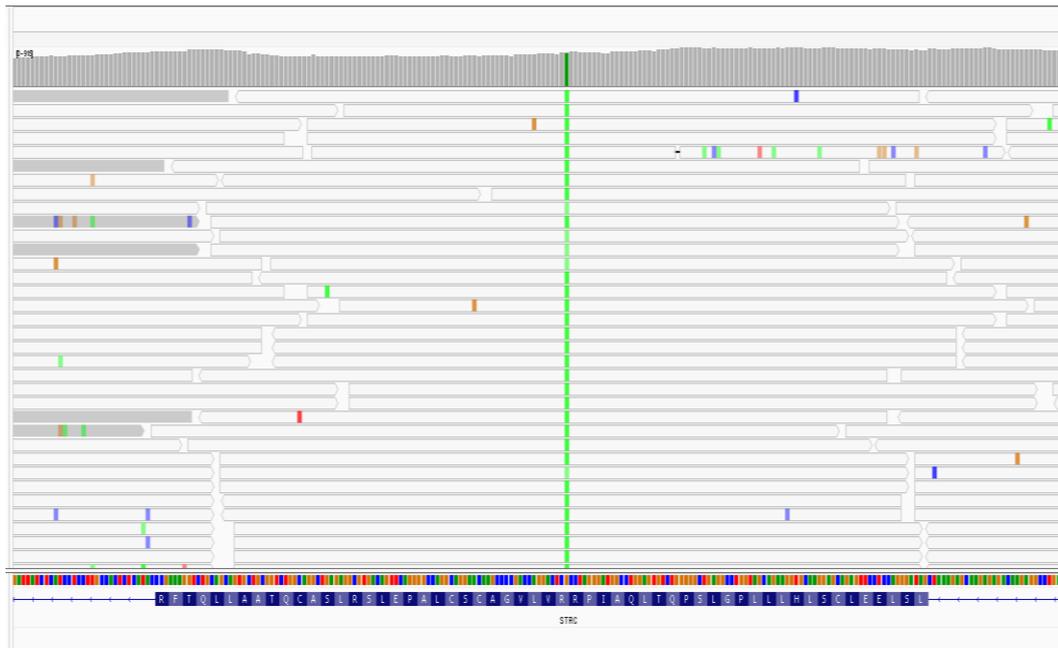


Figura 37. Risultato del sequenziamento NGS effettuato con protocollo di arricchimento Nextera eseguito nel paziente 5. Si può osservare la presenza di una sostituzione nucleotidica in apparente omozigosi: in realtà si tratta di una condizione di emizigosi in quanto uno dei due alleli risulta deletato.

6.3.6 PAZIENTE 6

Si tratta di un soggetto di sesso femminile, sottoposta ad analisi genetica all'età di 69 anni, dopo aver eseguito consulenza genetica in AOPD.

È stata sottoposta alle seguenti tecniche di indagine: NGS di pannello multigenico per geni associati ad ipoacusia, ricerca bioinformatica di CNV con software *Agilent SureCall* a partire dai dati ottenuti con NGS, LR-PCR associata a NGS con protocollo di arricchimento *Nextera* specifica per gene *STRC*, MLPA con Kit apposito alla ricerca di CNV di pertinenza del locus genico a cui appartiene *STRC*. Il sequenziamento NGS di pannello multigenico associato ad ipoacusia non ha restituito la presenza di varianti puntiformi, mentre l'analisi bioinformatica delle CNV ha suggerito la presenza di una delezione completa di *STRC*, non interessante *CATSPER2*, in apparente eterozigosi.

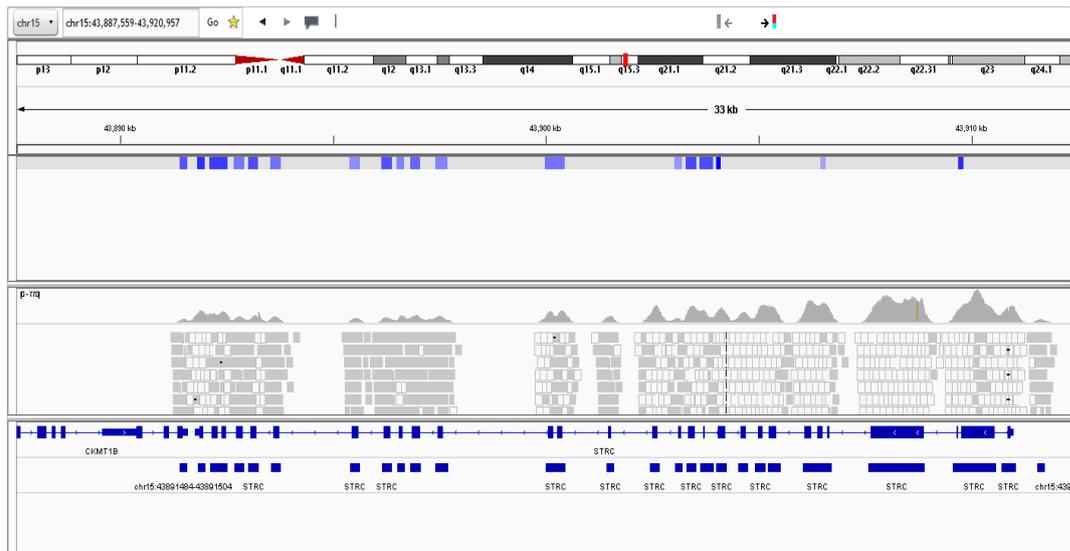


Figura 38. Risultato dell'analisi bioinformatica delle CNV effettuata nel paziente 6 a partire dai dati del sequenziamento NGS eseguito con metodiche tradizionali nello stesso (protocollo Sureselect). Il software suggerisce la presenza di una delezione completa del gene *STRC* in corrispondenza di un allele.

Successivamente, è stata eseguita indagine MLPA, che ha confermato la presenza di delezione completa di *STRC* in eterozigosi. Trattandosi tuttavia di una condizione recessiva, la variazione identificata non era sufficiente per giustificare il fenotipo osservato.

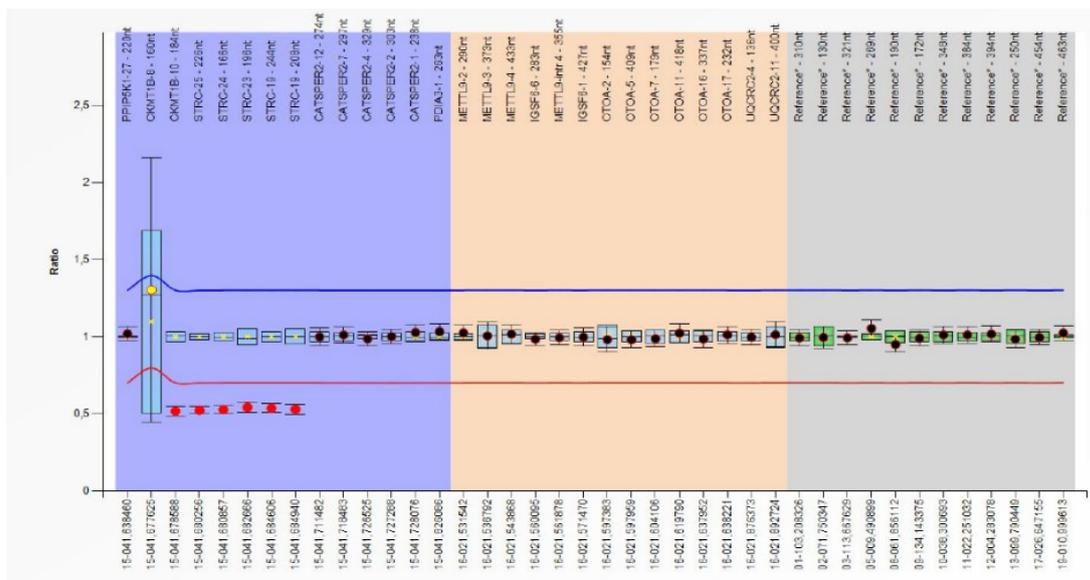


Figura 39. Risultato dell'indagine MLPA eseguita nel paziente 6. Si può osservare la presenza di una delezione in apparente eterozigosi, interessante il gene *STRC*.

L'analisi genetica è stata pertanto approfondita mediante l'impiego di sequenziamento NGS effettuato con protocollo di arricchimento *Nextera* su materiale genomico amplificato mediante LR-PCR, al fine di individuare eventuali varianti non riconosciute con le tecniche precedentemente utilizzate. Essendo stata confermata in uno dei due alleli la presenza di una delezione interessante l'intero gene, ci si aspetta di amplificare tramite LR-PCR esclusivamente l'allele non deleto.

Con tale metodica è stato tuttavia possibile sequenziare solamente la porzione distale del gene, mentre la porzione prossimale ha presentato una copertura insufficiente: viene quindi ipotizzata la presenza, nell'allele d'interesse, di un riarrangiamento con delezione della porzione prossimale del gene.

Né l'analisi bioinformatica né la metodica MLPA sono state in grado di suggerire questo tipo di sbilanciamento, in quanto interessante una regione con omologia quasi completa con lo pseudogene.

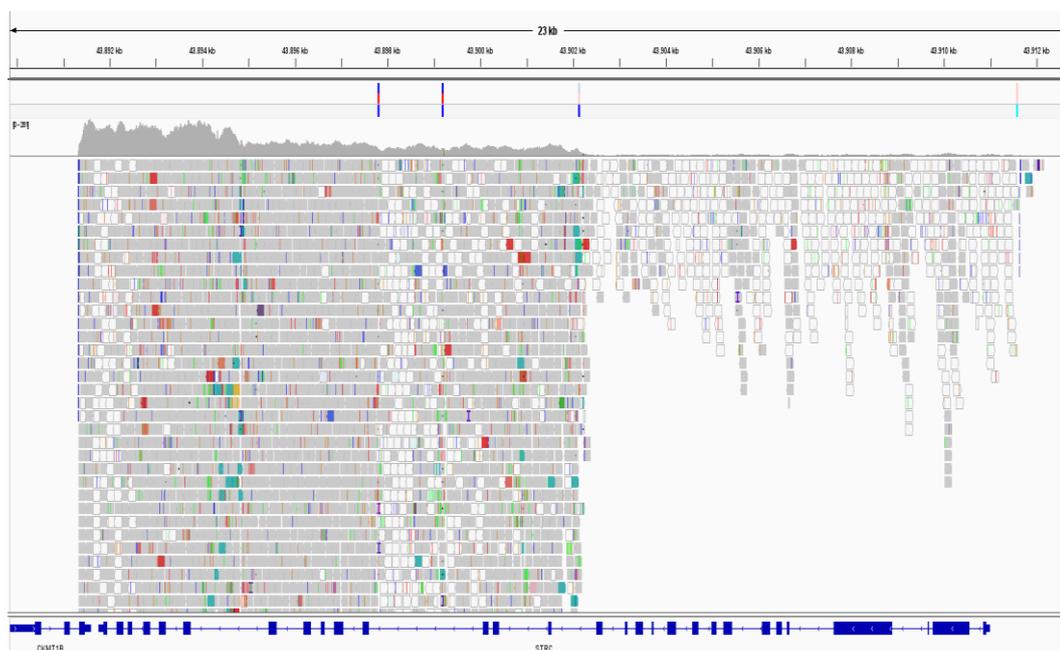


Figura 40. Risultato del sequenziamento NGS effettuato con protocollo di arricchimento *Nextera* eseguito nel paziente 6. È stato verosimilmente sequenziato un solo allele, presentante delezione della regione prossimale di *STRC*, mentre l'altro allele, caratterizzato dalla presenza di una delezione completa del gene, non è stato sequenziato perché la delezione interessava anche la regione di attacco dei primer per LR-PCR.

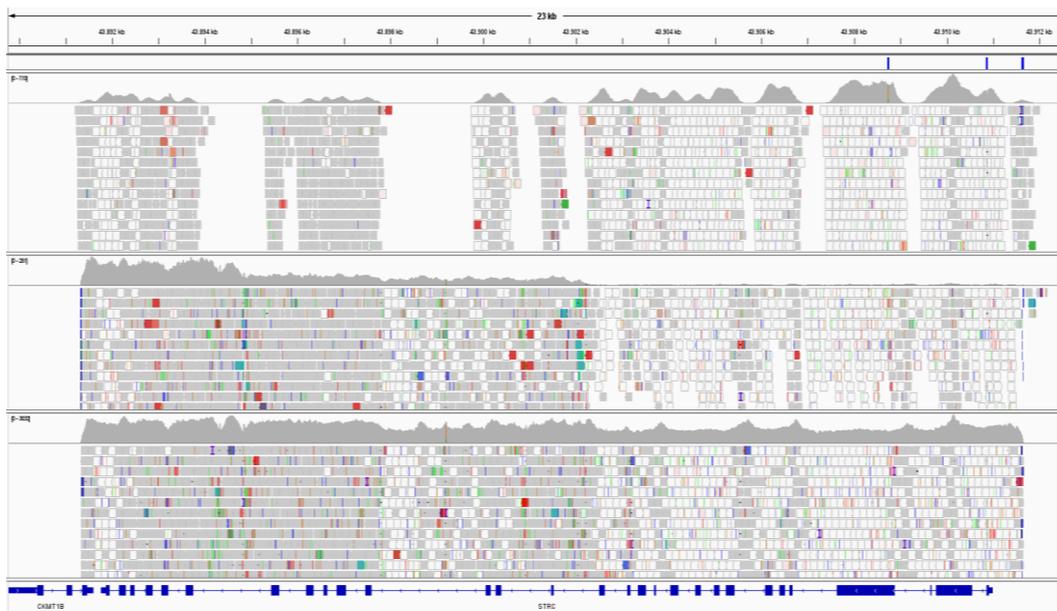


Figura 41. Confronto tra i sequenziamenti effettuati nel paziente 6, sfruttando rispettivamente il protocollo SureSelect (in alto) e l'arricchimento con LR-PCR+Nextera (in centro), e un sequenziamento Nextera effettuato su un controllo non presentante CNV in corrispondenza del gene esaminato (in basso). L'analisi NGS con Nextera individua la delezione prossimale a differenza di Sureselect poiché privo di contaminazione da pseudogene.

Le varie metodologie d'indagine impiegate hanno quindi permesso di individuare le seguenti variazioni:

- Delezione multiesonica interessante esclusivamente il gene *STRC* nella sua interezza.
In accordo con quanto stabilito secondo i criteri di classificazione ACMG, si tratta di una mutazione patogenetica.
- Delezione multiesonica interessante la sola regione prossimale di *STRC*.
In accordo con quanto stabilito secondo i criteri di classificazione ACMG, si tratta di una mutazione patogenetica.

Nella paziente in esame non è stata eseguita la segregazione parentale, pertanto non è possibile determinare le modalità di trasmissione ereditaria delle variazioni individuate.

Per quanto riguarda il fenotipo, la paziente manifestava ipoacusia neurosensoriale bilaterale di grado lieve.

L'esordio è stato post-linguale, la paziente non è stata tuttavia in grado di definire con precisione l'età di insorgenza.

6.3.7 PAZIENTE 7

Si tratta di un soggetto di sesso maschile, sottoposto ad analisi genetica all'età di 8 anni, dopo aver eseguito consulenza genetica in AOPD.

È stata sottoposto alle seguenti tecniche di indagine: NGS di pannello multigenico per geni associati ad ipoacusia, ricerca bioinformatica di CNV con software *Agilent SureCall* a partire dai dati ottenuti con NGS, MLPA con Kit apposito alla ricerca di CNV di pertinenza del locus genico a cui appartiene *STRC*.

Esse hanno permesso di individuare le seguenti variazioni:

- Delezione multiesonica interessante sia il gene *STRC* che il gene *CATSPER2*.

In accordo con quanto stabilito secondo i criteri di classificazione ACMG, si tratta di una mutazione patogenetica.

- NM_153700.2: c.877G>A p.(Gly293Ser). Essa consiste in una sostituzione nucleotidica puntiforme di una guanina con un'adenina in posizione 877, che determina nella proteina codificata dal gene *STRC* una sostituzione aminoacidica di una glicina con una serina in posizione 293. Per quanto la mutazione sia stata interpretata come una variante missenso, va evidenziato come, in realtà, interessando il secondo nucleotide di un esone, potrebbe verosimilmente costituire una variante di *splicing*.

In accordo con quanto stabilito secondo i criteri di classificazione ACMG, si tratta di una variante di significato incerto (VUS)

È stata inoltre effettuata la segregazione parentale, che ha permesso di definire la tipologia di trasmissione delle variazioni individuate e l'appartenenza delle stesse a differenti alleli: la delezione multiesonica è risultata essere di origine paterna, la sostituzione puntiforme di origine materna.

Per quanto riguarda il fenotipo, il paziente manifestava ipoacusia neurosensoriale bilaterale di grado lieve-moderato.

L'esordio è stato post-linguale, manifestandosi all'età di 8 anni.

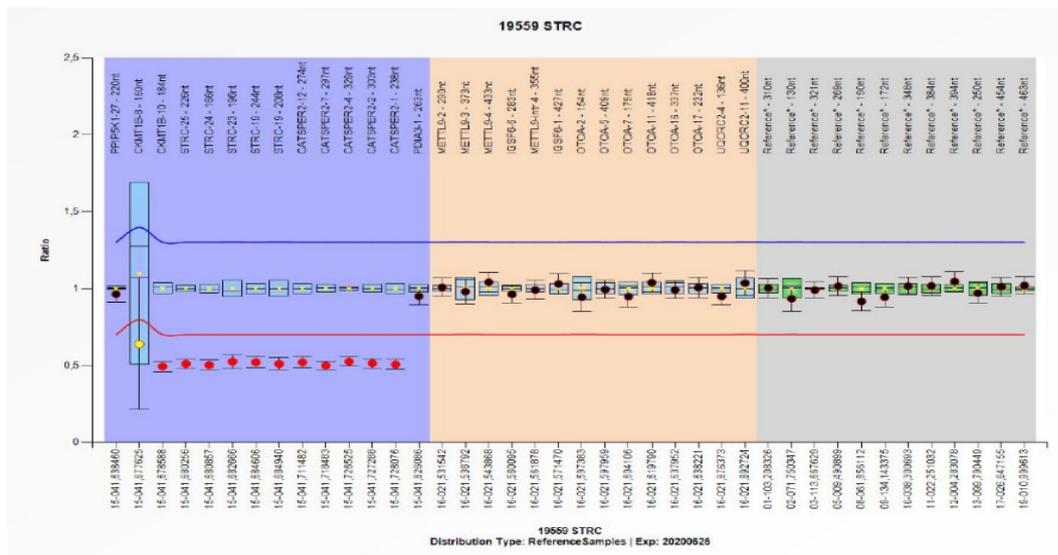


Figura 42. Risultato dell'indagine MLPA eseguita nel paziente 7. Si può osservare la presenza di una delezione in eterozigosi, interessante i geni *STRC* e *CATSPER2*.

6.3.8 PAZIENTE 8

Si tratta di un soggetto di sesso femminile, sottoposta ad analisi genetica all'età di 10 anni, dopo aver eseguito consulenza genetica in sede esterna ad AOPD.

È stata sottoposta alle seguenti tecniche di indagine: NGS di pannello multigenico per geni associati ad ipoacusia, ricerca bioinformatica di CNV con software *Agilent SureCall* a partire dai dati ottenuti con NGS, MLPA con Kit apposito alla ricerca di CNV di pertinenza del locus genico a cui appartiene *STRC*.

Il sequenziamento NGS ha evidenziato una profondità di lettura pari a zero in tutte le regioni in cui le *reads* allineano in modo specifico al gene, permettendo così di ipotizzare la presenza di una delezione multiesonica in omozigosi; la successiva analisi MLPA, ha confermato l'ipotesi, evidenziando la presenza in omozigosi della seguente CNV:

- Delezione multiesonica interessante sia il gene *STRC* che il gene *CATSPER2*.

In accordo con quanto stabilito secondo i criteri di classificazione ACMG, si tratta di una mutazione patogenetica.

Non è stata effettuata la segregazione parentale, che sarebbe comunque risultata superflua in quanto la paziente presentava una delezione in omozigosi.

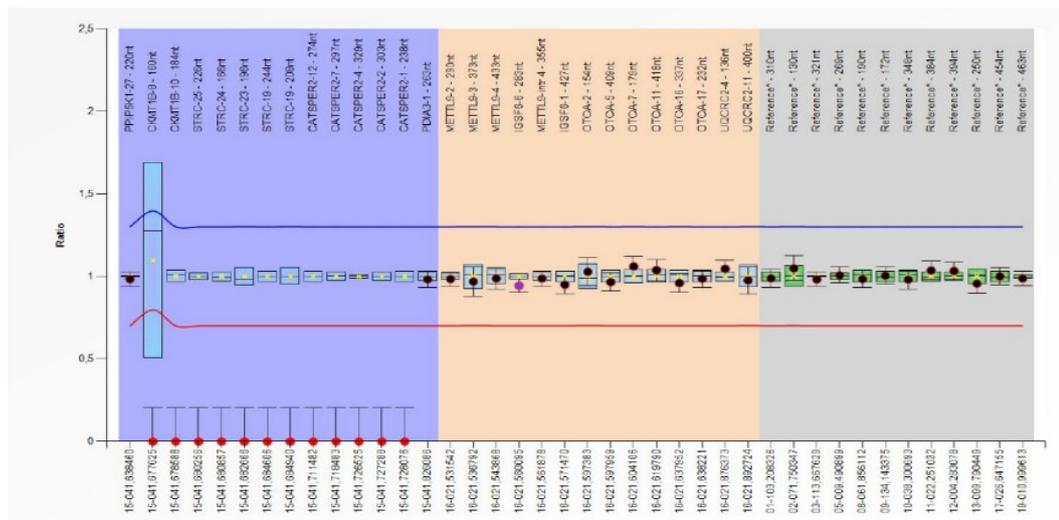


Figura 43. Risultato dell'indagine MLPA eseguita nel paziente 8. Si può osservare la presenza di una delezione in omozigosi, interessante i geni *STRC* e *CATSPER2*.

Non è stato possibile risalire ad informazioni cliniche ed anamnestiche approfondite a riguardo.

Si specifica tuttavia come la presenza di delezione omozigote interessante i geni *STRC* e *CATSPER2* determini un fenotipo sindromico nei soggetti di sesso maschile, che si manifesta con ipoacusia e infertilità (sindrome DIS): essendo la paziente in questione di sesso femminile, ella manterrà una normale fertilità.

6.3.9 PAZIENTE 9

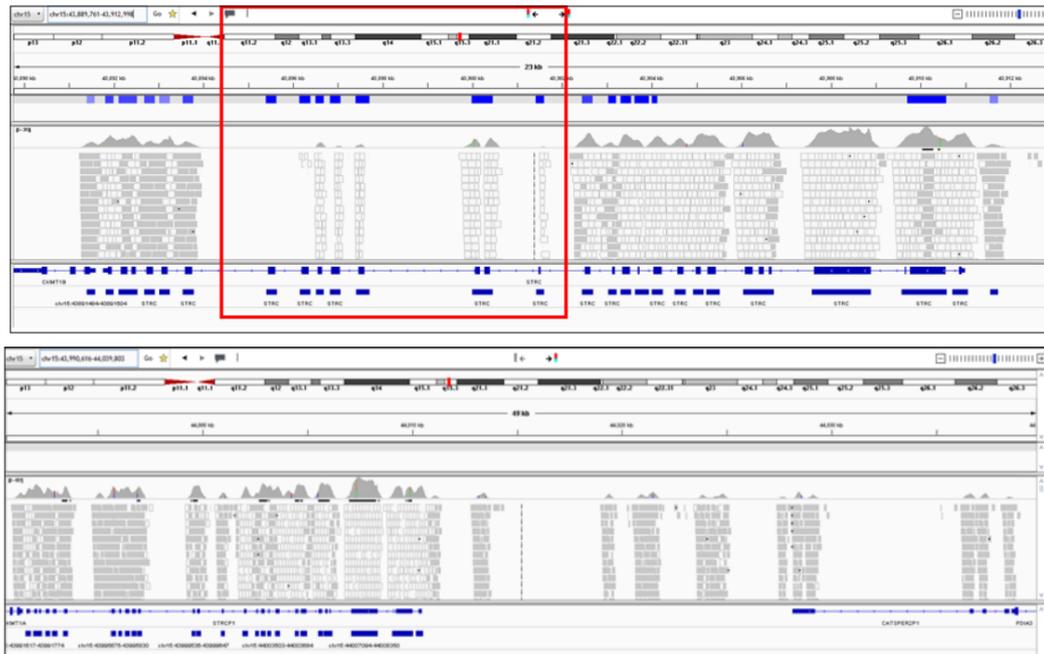
Si tratta di un soggetto di sesso femminile, sottoposta ad analisi genetica all'età di 51 anni, dopo aver eseguito consulenza genetica in sede esterna ad AOPD.

È stata sottoposta alle seguenti tecniche di indagine: NGS di pannello multigenico per geni associati ad ipoacusia, ricerca bioinformatica di CNV con software *Agilent SureCall* a partire dai dati ottenuti con NGS, LR-PCR associata a NGS con protocollo di arricchimento *Nextera* specifica per gene *STRC*, sequenziamento di terza generazione Oxford Nanopore specifico per il gene *STRC*, MLPA con Kit apposito alla ricerca di CNV di pertinenza del locus genico a cui appartiene *STRC*. Il sequenziamento NGS di un pannello di geni associati a ipoacusia non ha identificato la presenza di varianti puntiformi, tuttavia ha presentato una profondità di lettura prossima allo zero in corrispondenza della regione compresa tra gli esoni 16 e 23 del gene *STRC*: ciò suggerisce la presenza di un'apparente delezione in omozigosi di parte del gene.



Figura 44. Confronto tra i risultati di sequenziamento NGS eseguito con metodiche tradizionali (protocollo Sureselect) nel paziente 9, in basso, e in un controllo verosimilmente privo di varianti strutturali a carico del locus d'interesse, in alto. Si nota l'assenza di coverage nel paziente 9 in corrispondenza della regione compresa tra gli esoni 16 e 23 di STRC.

L'analisi bioinformatica delle CNV ha quindi suggerito la presenza di un'ampia delezione multiesonica interessante nella loro interezza i geni *STRC* e *CATSPER2*, ed un ulteriore sbilanciamento a carico di parte di *STRC* (esoni 16-23), visualizzabile per assenza di copertura. I risultati sono stati poi confermati mediante MLPA.



*Figura 45. Analisi bioinformatica delle CNV eseguita nel paziente 9 a partire dai dati ottenuti con NGS. Il riquadro superiore evidenzia una delezione che interessa i geni *STRC* e *CATSPER2* nella loro interezza, la regione con il riquadro rosso è verosimilmente deleta in omozigosi. Il riquadro inferiore riguarda la sequenza dello pseudogene, e non sembra presentare alterazioni.*

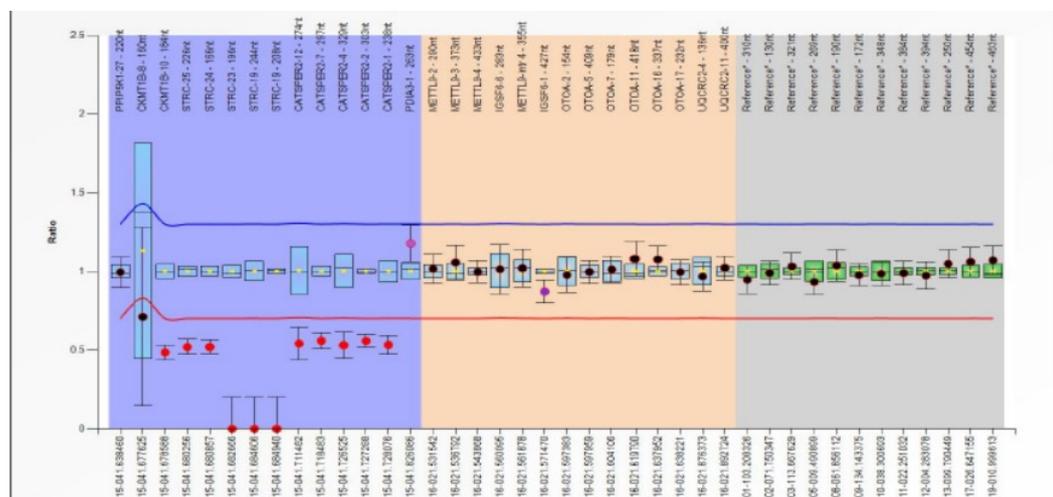


Figura 46. Risultato dell'indagine MLPA eseguita nel paziente 9.

Al fine di caratterizzare i punti di rottura della delezione di dimensioni minori è stato quindi eseguito sequenziamento NGS effettuato con protocollo di arricchimento *Nextera* su materiale genomico amplificato con LR-PCR specifica per *STRC*. Essendo verosimilmente presente in uno dei due alleli una delezione

dell'intero gene, ci si aspetta di amplificare, mediante LR-PCR, esclusivamente l'allele interessato dalla delezione più piccola.

Il sequenziamento sembra confermare quanto osservato tramite le metodologie d'indagine applicate precedentemente; tuttavia, l'allineamento dell'amplicone ottenuto tramite LR-PCR con la sequenza dello pseudogene *pSTRC1* evidenzia la presenza di *reads* che mappano in maniera specifica con lo pseudogene in corrispondenza della regione ritenuta deleta: si ipotizza quindi la presenza di un riarrangiamento genico complesso, verosimilmente costituito da una conversione genica, in cui del materiale di partenza pseudogenica è stato trasferito nel gene.

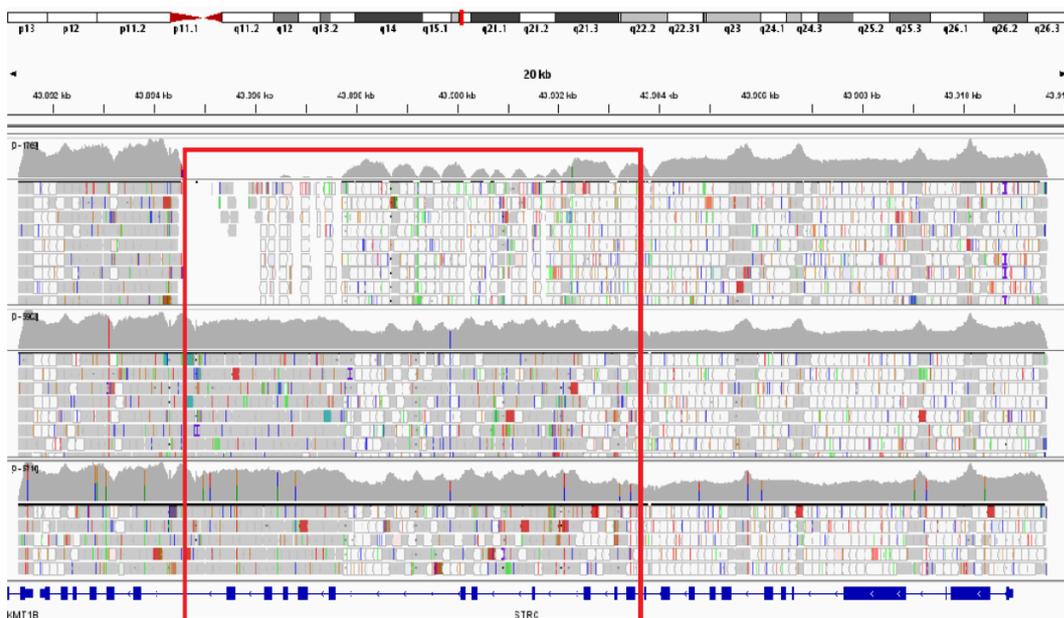


Figura 47. Confronto tra il risultato del sequenziamento NGS effettuato con protocollo di arricchimento Nextera eseguito nel paziente 9 (in alto) e in due controlli verosimilmente privi di varianti strutturali a carico del locus d'interesse. Il paziente 9 mostra nella regione evidenziata col riquadro rosso un coverage diverso rispetto agli altri controlli per i quali non sono state evidenziate delezioni a carico di STRC.



Figura 48. Allineamento dell'amplicone ottenuto mediante LR-PCR e processato con protocollo Nextera con la regione dello pseudogene STRCP1. In rosso è evidenziata la regione in cui nel paziente 9 (in alto) sono presenti delle reads che mappano in modo specifico con lo pseudogene. Nei controlli non sono invece presenti reads che mappano in modo specifico allo pseudogene.

Al fine di confermare l'ipotesi di conversione genica, è stato eseguito sequenziamento di terza generazione mediante *Oxford Nanopore*.

In questo contesto sono stati utilizzati due approcci diversi: l'*adaptive sampling* e il sequenziamento dell'amplicone dopo LR-PCR. Mentre il primo non ha consentito il raggiungimento di una copertura adeguata, il secondo ha permesso l'ottenimento di risultati soddisfacenti.

Le *reads* ottenute, caratterizzate da lunghezza tale da riuscire a coprire l'intera regione di circa 20 Kb, sono state allineate sia con la sequenza del gene che con la sequenza dello pseudogene. L'allineamento con il gene evidenzia nella parte centrale compresa tra gli esoni 11 e 23 *coverage* nettamente diminuito e presenza di numerose variazioni di sequenza rispetto a quanto ci si aspetta in un allele privo di riarrangiamenti strutturali; allineando invece le *reads* con lo pseudogene il *coverage* nella regione risulta omogeneo e la sequenza pressoché priva di variazioni. Queste osservazioni hanno permesso di comprendere come la zona apparentemente deleta fosse in realtà costituita da sequenze di origine pseudogenica, ed è stata così verificata l'ipotesi di conversione genica.

Sono stati inoltre caratterizzati i punti di rottura della stessa in corrispondenza degli introni 11 e 23 di *STRC*.

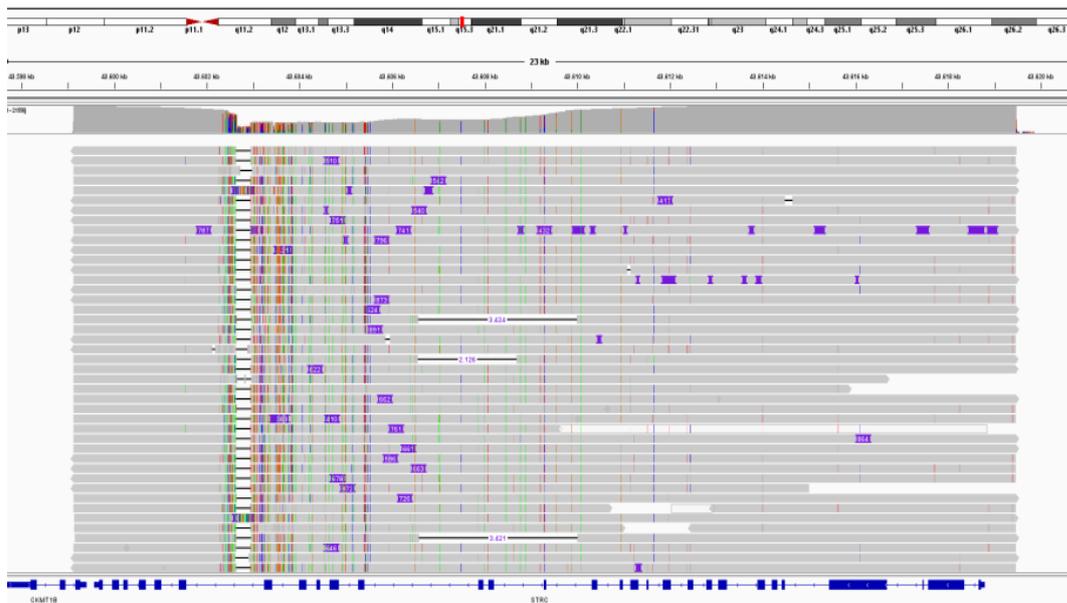


Figura 49. Risultato del sequenziamento di terza generazione Oxford Nanopore eseguito nel paziente 9. Nell'immagine viene rappresentato il sequenziamento dell'allele del gene *STRC* verosimilmente oggetto di conversione genica: si osserva, in corrispondenza della regione centrale del gene, un'area caratterizzata da coverage diminuito e numerose variazioni di sequenza rispetto a quanto ci si aspetta in un allele privo di riarrangiamenti strutturali.

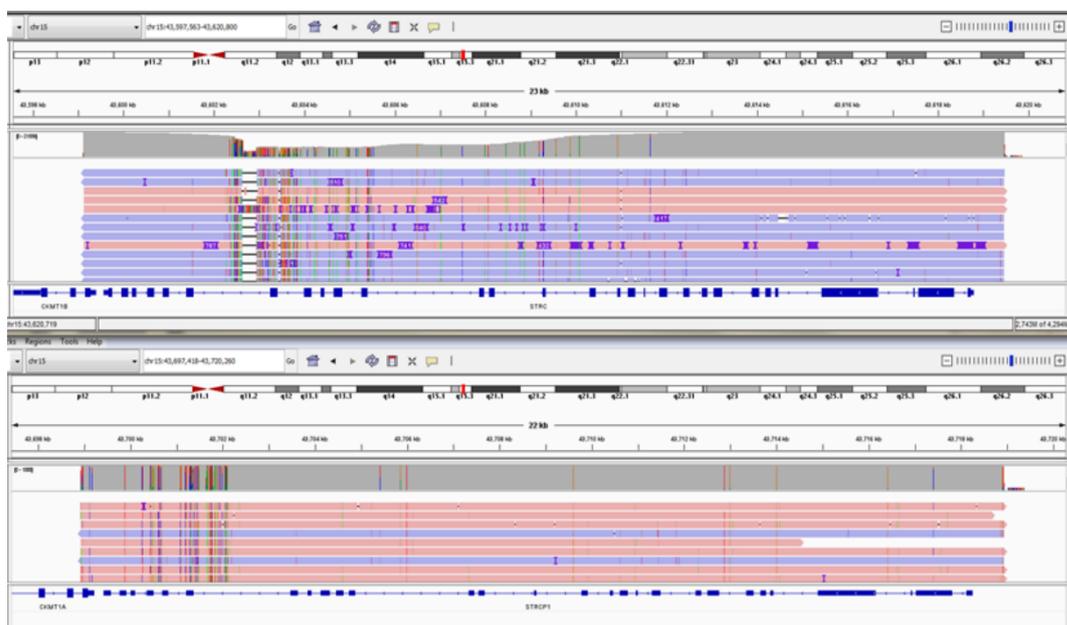


Figura 50. Risultato del sequenziamento di terza generazione Oxford Nanopore eseguito nel paziente 9. Viene rappresentato l'allineamento delle reads ottenute dal sequenziamento dell'allele d'interesse con la regione del gene, in alto, e con la regione dello pseudogene, in basso: si può osservare come la regione ad alta varianza di sequenza e basso coverage in corrispondenza della regione centrale gene mappa invece correttamente con la sequenza dello pseudogene. Questo ci permette ricostruire una sequenza consensus e di verificare l'ipotesi di conversione genica

L'utilizzo delle varie tecniche d'indagine ha quindi permesso di individuare le seguenti variazioni:

- Delezione multiesonica interessante sia il gene *STRC* che il gene *CATSPER2*.

In accordo con quanto stabilito secondo i criteri di classificazione ACMG, si tratta di una mutazione patogenetica.

- Conversione genica con delezione di parte del gene *STRC*.

In accordo con quanto stabilito secondo i criteri di classificazione ACMG, si tratta di una mutazione patogenetica.

Nella paziente in esame non è stata eseguita la segregazione parentale, pertanto non è possibile determinare le modalità di trasmissione ereditaria delle variazioni individuate.

Non avendo la paziente eseguito consulenza genetica presso AOPD, non è stato possibile risalire ad informazioni cliniche ed anamnestiche approfondite a riguardo, tuttavia sappiamo che l'esordio è stato pre-linguale.

6.3.10 PAZIENTE 10

Si tratta di un soggetto di sesso femminile, sottoposta ad analisi genetica all'età di 9 anni, dopo aver eseguito consulenza genetica in sede esterna ad AOPD.

È stata sottoposta alle seguenti tecniche di indagine: NGS di pannello multigenico per geni associati ad ipoacusia, ricerca bioinformatica di CNV con software *Agilent SureCall* a partire dai dati ottenuti con NGS.

Il sequenziamento NGS ha evidenziato una profondità di lettura pari a zero in tutte le regioni in cui le *reads* allineano in modo specifico al gene, permettendo così di ipotizzare la presenza di una delezione multiesonica in omozigosi. Nonostante non sia stata effettuata MLPA per ragioni logistiche, è stata diagnosticata la presenza in omozigosi della seguente CNV:

- Delezione multiesonica interessante sia il gene *STRC* che il gene *CATSPER2*.

In accordo con quanto stabilito secondo i criteri di classificazione ACMG, si tratta di una mutazione patogenetica.

Non è stata effettuata la segregazione parentale, che sarebbe comunque risultata superflua in quanto la paziente presentava una delezione in omozigosi.

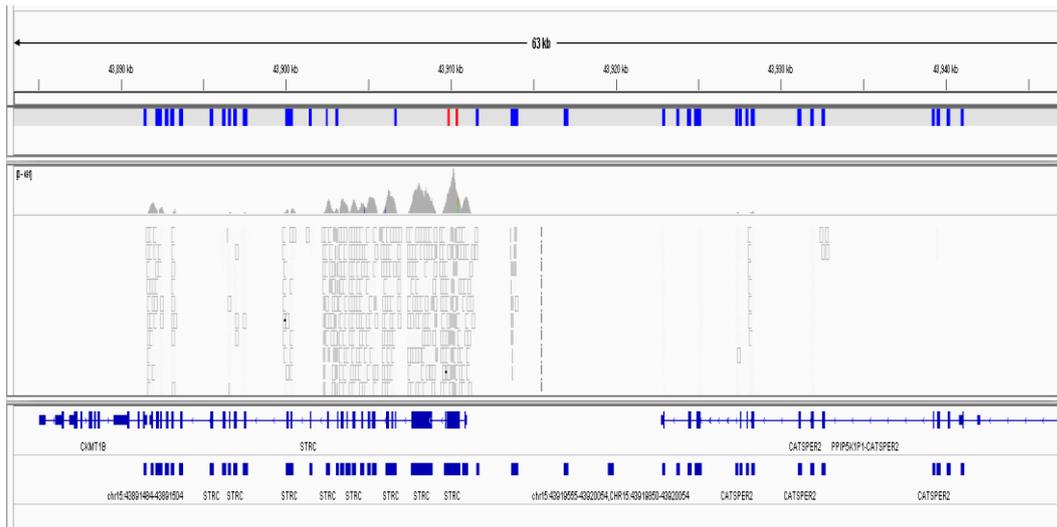


Figura 51. Risultato dell'analisi bioinformatica delle CNV effettuata nel paziente 10 a partire dai dati del sequenziamento NGS eseguito nello stesso. Il software suggerisce la presenza in omozigosi di una delezione multiesonica interessante l'interezza dei geni *STRC* e *CATSPER2*

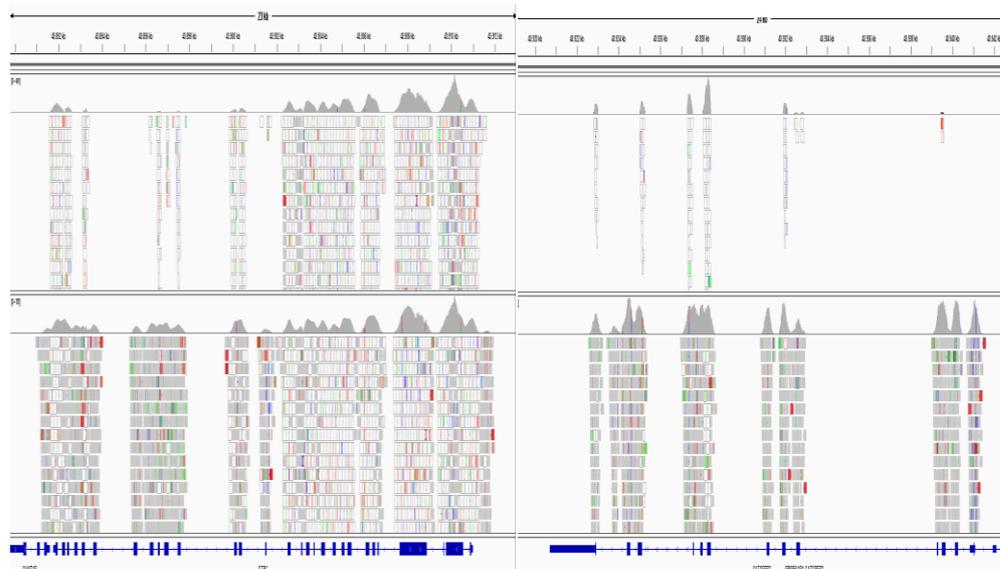


Figura 52. Confronto tra il sequenziamento NGS effettuato nel paziente 10 e il sequenziamento svolto in un controllo privo di delezioni a carico dei geni *STRC* (a sinistra) e *CATSPER2* (a destra).

Non è stato possibile risalire ad informazioni cliniche ed anamnestiche approfondite a riguardo, tuttavia sappiamo che l'esordio dell'ipoacusia è stato prelinguale.

Si specifica inoltre come la presenza di delezione omozigote interessante i geni *STRC* e *CATSPER2* determini un fenotipo sindromico nei soggetti di sesso maschile, che si manifesta con ipoacusia e infertilità (sindrome DIS): essendo la paziente in questione di sesso femminile, ella manterrà fertilità normale.

6.3.11 PAZIENTE 11

Si tratta di un soggetto di sesso femminile, sottoposta ad analisi genetica all'età di 20 anni, dopo aver eseguito consulenza genetica in AOPD.

È stata sottoposta alle seguenti tecniche di indagine: NGS di pannello multigenico per geni associati ad ipoacusia, ricerca bioinformatica di CNV con software *Agilent SureCall* a partire dai dati ottenuti con NGS, LR-PCR associata a NGS con protocollo di arricchimento *Nextera* specifica per gene *STRC*.

Esse hanno permesso di individuare la seguente variante in omozigosi:

- NM_153700.2: c.4784G>A p.(Gly1595Asp). Essa consiste in una sostituzione nucleotidica puntiforme di una guanina con un'adenina in posizione 4784, che determina nella proteina codificata dal gene *STRC* una sostituzione aminoacidica di una glicina con un aspartato in posizione 1595. In accordo con quanto stabilito secondo i criteri di classificazione ACMG, si tratta di una variante di significato incerto (VUS)

Non è stata effettuata la segregazione parentale, che avrebbe permesso di distinguere il quadro di omozigosi da un'eventuale emizigosi; l'analisi bioinformatica delle CNV non ha tuttavia suggerito la presenza di delezioni.

Viene inoltre sottolineato come l'esecuzione di sequenziamento con protocollo di arricchimento *Nextera* non fosse in realtà necessario, in quanto la mutazione individuata in omozigosi era localizzata in corrispondenza di un'area di non omologia tra gene e pseudogene.

Per quanto riguarda il fenotipo, la paziente manifestava ipoacusia neurosensoriale bilaterale di grado lieve, con interessamento principalmente a carico delle frequenze acute.

L'esordio è stato pre-linguale, l'andamento della condizione clinica stabile.

È stata eseguita protesizzazione.

All'anamnesi sono emersi consanguineità genitoriale, che rende ragione del quadro di omozigosi della variante individuata.

6.3.12 PAZIENTE 12

Si tratta di un soggetto di sesso maschile, sottoposto ad analisi genetica all'età di 15 anni, dopo aver eseguito consulenza genetica esternamente ad AOPD.

È stata sottoposto alle seguenti tecniche di indagine: NGS di pannello multigenico per geni associati ad ipoacusia, ricerca bioinformatica di CNV con software *Agilent SureCall* a partire dai dati ottenuti con NGS, LR-PCR associata a NGS con protocollo di arricchimento *Nextera* specifica per gene *STRC*.

Il sequenziamento NGS di un pannello di geni associati a ipoacusia ha permesso di individuare la presenza di due alleli differenti, di cui uno presenta una sostituzione dinucleotidica, mentre l'altro presenta variazioni nucleotidiche tipiche dello pseudogene.

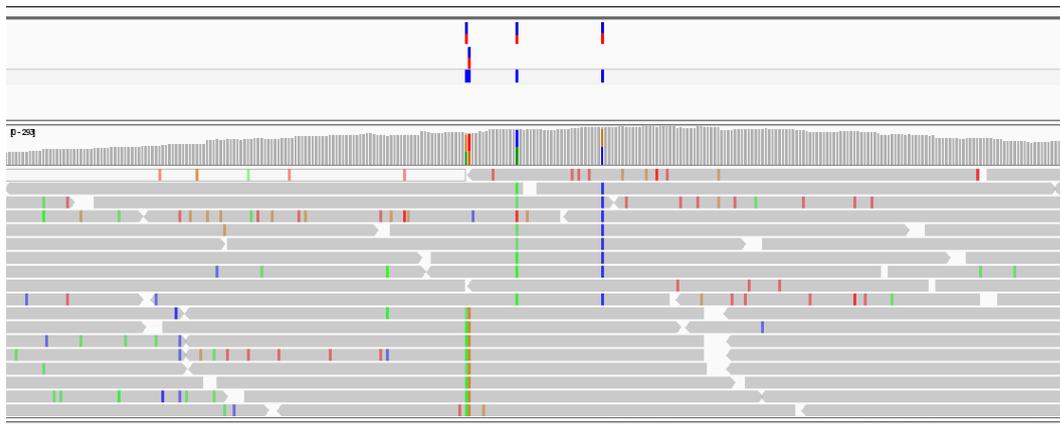


Figura 53. Risultato del sequenziamento NGS eseguito nel paziente 12 con protocollo *SureSelect*. È possibile osservare in un allele una sostituzione dinucleotidica; nell'altro diverse variazioni, che corrispondono alla sequenza dello pseudogene.

La presenza di variazioni nucleotidiche tipicamente pseudogeniche evidenziate nel gene tramite sequenziamento NGS ha sollevato l'ipotesi della presenza di un riarrangiamento strutturale complesso a carico di un allele; è stato quindi eseguito sequenziamento NGS effettuato con protocollo di arricchimento *Nextera* su materiale genetico amplificato mediante LR-PCR specifica per il gene *STRC*. Esso non è riuscito tuttavia a sequenziare l'allele d'interesse, restituendo la presenza della sostituzione dinucleotidica in apparente omozigosi: viene di conseguenza ipotizzata, a carico dell'allele non sequenziato, la presenza di un riarrangiamento cromosomico che interessa anche le sequenze di attacco dei *primer* della LR-PCR, impedendone pertanto l'amplificazione.

La verosimile conversione genica dovrà essere confermata mediante sequenziamento di terza generazione *Oxford Nanopore*, attualmente in fase di accertamento.



Figura 54. Risultato del sequenziamento NGS eseguito nel paziente 12. In alto si può osservare un sequenziamento eseguito con metodiche tradizionali (Sureselect), in basso un sequenziamento eseguito con protocollo di arricchimento Nextera. In Nextera è stato sequenziato solo uno dei due alleli in quanto l'altro presenta un riarrangiamento strutturale che interessa la regione di attacco dei primer per la LR-PCR; in Sureselect invece si nota la presenza di due alleli, di cui uno presentante diverse variazioni di sequenza tipiche dello pseudogene (si vedano anche le Figure 53 e 55), facendo pertanto ipotizzare la presenza di una conversione genica a carico dello stesso.

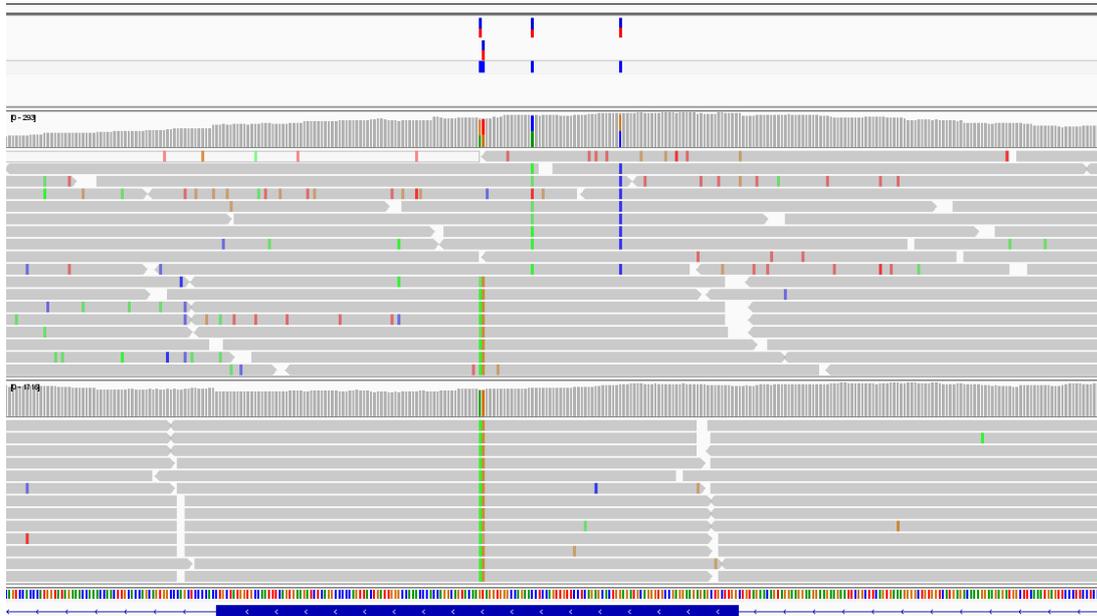


Figura 55. *Ingrandimento dell'immagine precedente. Si può osservare la presenza della sostituzione dinucleotidica d'interesse, che risulta presente in eterozigosi nel sequenziamento eseguito con protocollo Sureselect ed in apparente omozigosi nel sequenziamento eseguito con protocollo Nextera dopo LR-PCR. Ciò rende ragione del fatto che la LR-PCR è stata in grado di amplificare un unico allele a causa del riarrangiamento strutturale a carico dell'altro allele, verosimilmente interessante la regione di attacco dei primer.*

Le tecniche d'indagine utilizzate hanno quindi permesso di individuare le seguenti variazioni:

- NM_153700.2: c.4917_4918delinsCT p.(Leu1640Phe). Essa consiste in una sostituzione che interessa due nucleotidi, rispettivamente in posizione 4917 e 4918, che vengono sostituiti con una citosina e una timina. Tale variazione determina nella proteina codificata dal gene *STRC* con la sostituzione di una leucina con una fenilalanina in posizione 1640. Questa variante risulta analoga a quella presente nel paziente 12.

In accordo con quanto stabilito secondo i criteri di classificazione ACMG, si tratta di una mutazione patogenetica.

- Possibile conversione genica in fase di accertamento. È stata infatti individuata la presenza in trans di una possibile delezione che interessa almeno gli ultimi tre esoni del gene *STRC* e che potrebbe essere conseguenza di un più complesso riarrangiamento cromosomico.

In accordo con quanto stabilito secondo i criteri di classificazione ACMG, si tratta di una mutazione patogenetica.

Nel paziente in esame non è stata eseguita la segregazione parentale, pertanto non è possibile determinare le modalità di trasmissione ereditaria delle variazioni individuate.

Non è stato possibile risalire ad informazioni cliniche ed anamnestiche approfondite a riguardo, tuttavia sappiamo che anche il fratello è affetto.

6.3.13 PAZIENTE 13

Si tratta di un soggetto di sesso maschile, sottoposto ad analisi genetica all'età di 6 anni, dopo aver eseguito consulenza genetica esternamente ad AOPD.

È stata sottoposto alle seguenti tecniche di indagine: NGS di pannello multigenico per geni associati ad ipoacusia, ricerca bioinformatica di CNV con software *Agilent SureCall* a partire dai dati ottenuti con NGS, LR-PCR associata a NGS con protocollo di arricchimento *Nextera* specifica per gene *STRC*.

Esse hanno permesso di individuare le seguenti variazioni:

- NM_153700.2: c.2494C>T p.(Arg832Trp). Essa consiste in una sostituzione nucleotidica puntiforme di una citosina con una timina in posizione 2494, che determina nella proteina codificata dal gene *STRC* una sostituzione aminoacidica di un'arginina con un triptofano in posizione 832. In accordo con quanto stabilito dall'impiego di software predittivi, si tratta di una mutazione patogenetica.

In accordo con quanto stabilito secondo i criteri di classificazione ACMG, si tratta di una mutazione verosimilmente patogenetica.

- NM_153700.2: c.1368_1380del p.(Glu457Leufs*75). Essa consiste in una delezione di 13 nucleotidi, dalla posizione 1368 alla posizione 1380, che determina il *frameshift* del trascritto, ossia uno slittamento della cornice di lettura, con inserimento di un codone di stop prematuro.

In accordo con quanto stabilito secondo i criteri di classificazione ACMG, si tratta di una mutazione patogenetica.

È stata inoltre effettuata la segregazione parentale, che ha permesso di definire la tipologia di trasmissione delle variazioni individuate e l'appartenenza delle stesse a differenti alleli: la sostituzione puntiforme è risultata essere di origine paterna, la delezione di origine materna.

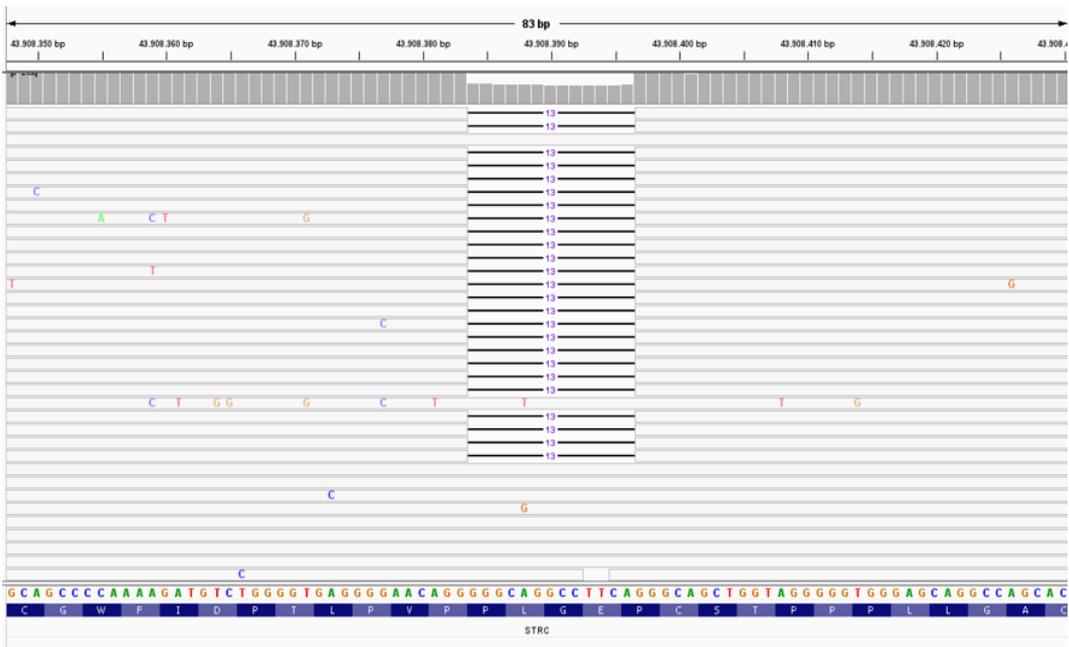


Figura 56. Risultato del sequenziamento NGS effettuato con protocollo di arricchimento Nextera eseguito nel paziente 13. Si può osservare la delezione di 13 nucleotidi in eterozigosi.

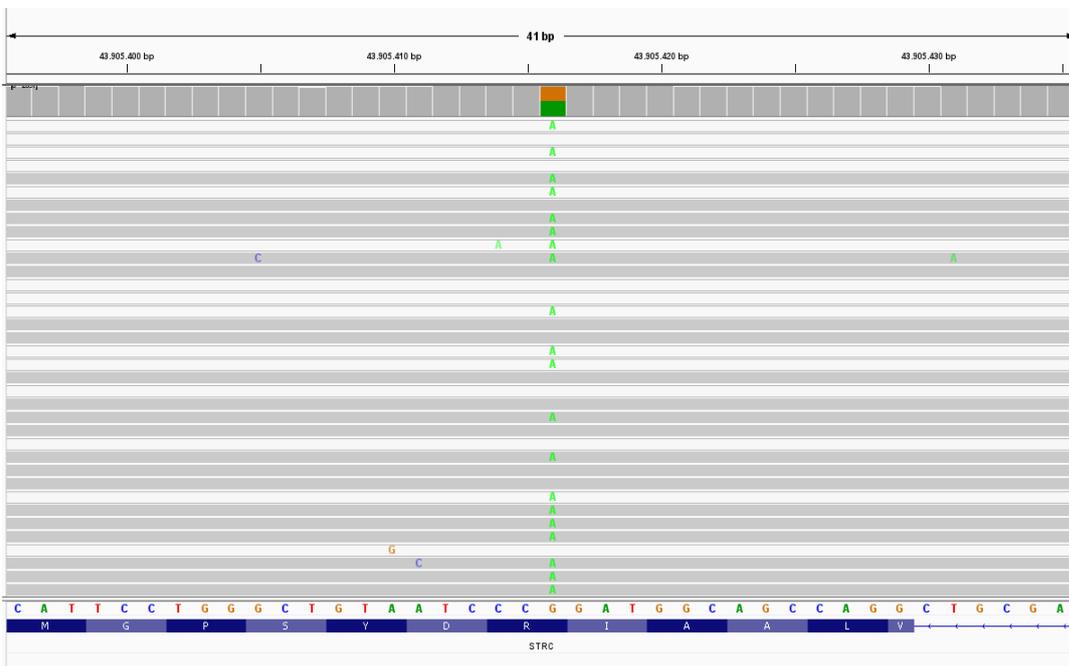


Figura 57. Risultato del sequenziamento NGS effettuato con protocollo di arricchimento Nextera eseguito nel paziente 13. Si può osservare la sostituzione puntiforme in eterozigosi.

Non è stato possibile risalire ad informazioni cliniche ed anamnestiche approfondite a riguardo, tuttavia sappiamo che il grado dell'ipoacusia è lieve-moderato e che l'esordio è stato pre-linguale.

6.3.14 PAZIENTE 14

Si tratta di un soggetto di sesso maschile, sottoposto ad analisi genetica all'età di 6 anni, dopo aver eseguito consulenza genetica esternamente ad AOPD.

È stata sottoposto alle seguenti tecniche di indagine: NGS di pannello multigenico per geni associati ad ipoacusia, ricerca bioinformatica di CNV con software *Agilent SureCall* a partire dai dati ottenuti con NGS.

Esse hanno permesso di individuare le seguenti variazioni:

- NM_153700.2: c.4917_4918delinsCT p.(Leu1640Phe). Essa consiste in una sostituzione che interessa due nucleotidi, rispettivamente in posizione 4917 e 4918, che vengono sostituiti con una citosina e una timina. Tale variazione determina nella proteina codificata dal gene *STRC* la sostituzione di una leucina con una fenilalanina in posizione 1640. Questa variante risulta analoga a quella presente nel paziente 12 (si tratta di una variante nota già precedentemente descritta in letteratura).

In accordo con quanto stabilito secondo i criteri di classificazione ACMG, si tratta di una mutazione patogenetica.

- Delezione multiesonica interessante sia il gene *STRC* che il gene *CATSPER2*.

In accordo con quanto stabilito secondo i criteri di classificazione ACMG, si tratta di una mutazione patogenetica.

Nel paziente in esame non è stata eseguita la segregazione parentale, pertanto non è possibile determinare le modalità di trasmissione ereditaria delle variazioni individuate.

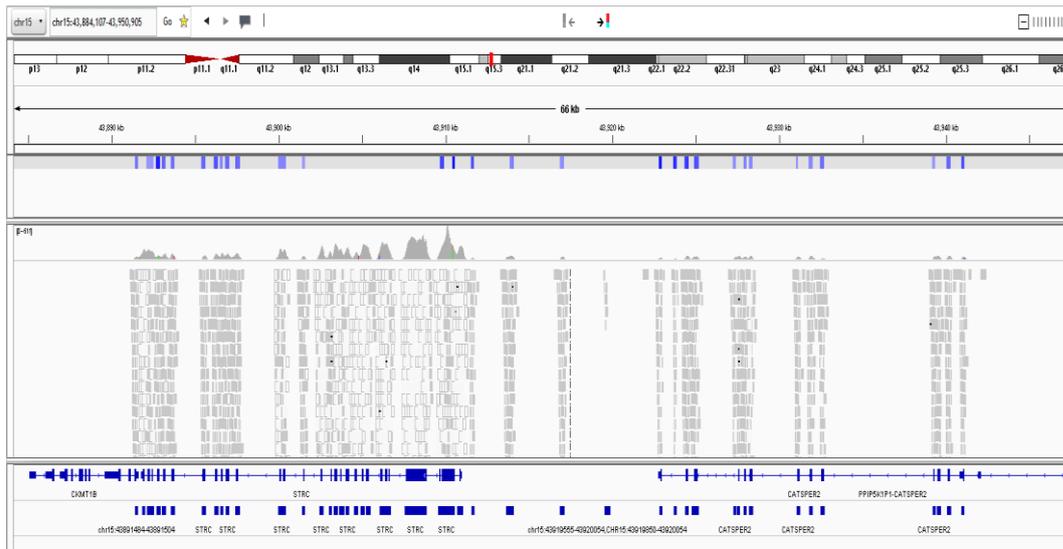


Figura 58. Risultato dell'analisi bioinformatica delle CNV effettuata nel paziente 14 a partire dai dati del sequenziamento NGS eseguito nello stesso. È possibile osservare la presenza di una delezione multiesonica che interessa una copia dei geni STRC e CATSPER2.

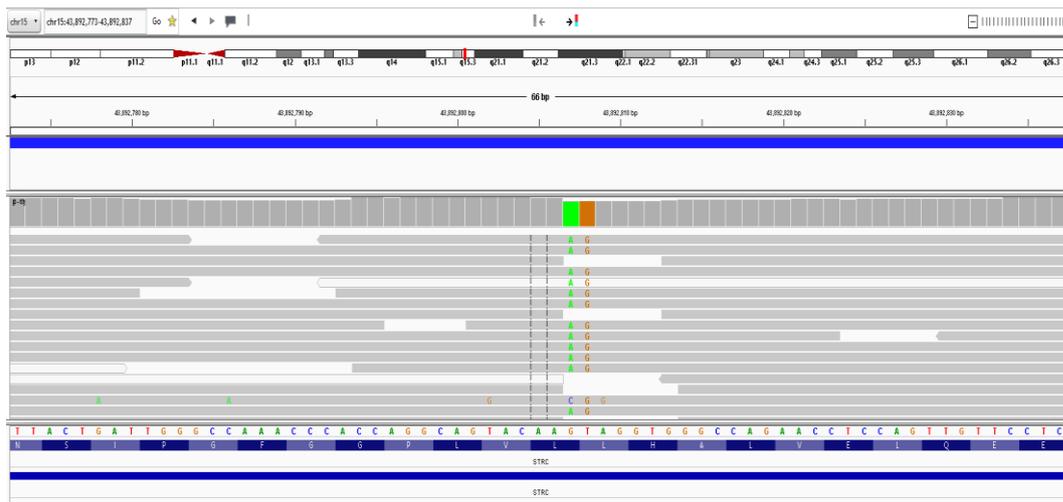


Figura 59. Risultato del sequenziamento NGS eseguito nel paziente 14. È possibile osservare la presenza di una sostituzione dinucleotidica in apparente omozigosi.

L'ipoacusia è caratterizzata da grado lieve in corrispondenza delle frequenze acute e lieve-moderato in corrispondenza delle frequenze gravi.

L'esordio è stato post-linguale, all'età di 3 anni.

È stata eseguita protesizzazione all'età di 5 anni, con risultati non brillanti nei primi due mesi.

6.3.15 PAZIENTE 15

Si tratta di un soggetto di sesso femminile, sottoposta ad analisi genetica all'età di 15 anni, dopo aver eseguito consulenza genetica in sede esterna ad AOPD.

È stata sottoposta alle seguenti tecniche di indagine: NGS di pannello multigenico per geni associati ad ipoacusia, ricerca bioinformatica di CNV con software *Agilent SureCall* a partire dai dati ottenuti con NGS.

Il sequenziamento NGS di un pannello di geni associato a ipoacusia ha identificato in *STRC* la presenza di una sostituzione nucleotidica in apparente omozigosi.

È stata quindi effettuata l'analisi bioinformatica delle CNV, che ha suggerito la presenza di una delezione multiesonica a carico della porzione distale di *STRC* e di una concomitante duplicazione dello pseudogene in corrispondenza della regione deleta nel gene: questo tipo di pattern permette di ipotizzare la presenza di un riarrangiamento genico complesso, possibilmente una conversione genica, in uno degli alleli del gene. La sostituzione identificata con NGS è anch'essa localizzata in corrispondenza dell'area deleta nell'allele riarrangiato, pertanto è verosimile che la mutazione sia presente in emizigosi.

La possibile conversione genica dovrà essere confermata mediante l'impiego di sequenziamento di terza generazione *Oxford Nanopore*, indagine attualmente in corso di accertamento.

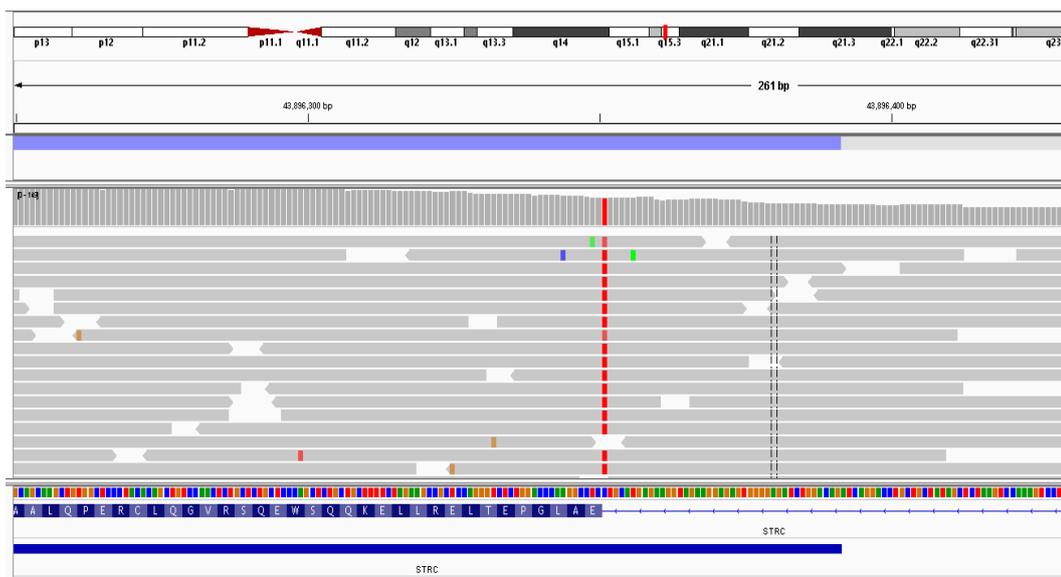


Figura 60. Risultato del sequenziamento NGS eseguito nel paziente 15. E' possibile osservare la presenza di una sostituzione nucleotidica in apparente omozigosi.

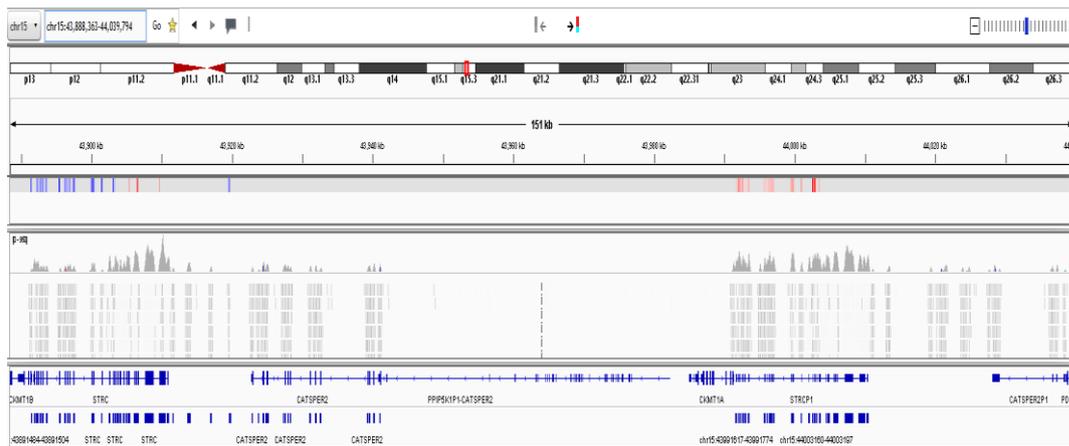


Figura 61. Risultato dell'analisi bioinformatica delle CNV effettuata nel paziente 15 a partire dai dati del sequenziamento NGS eseguito nello stesso. E' possibile osservare la presenza di una delezione multiesonica di parte di *STRC* in corrispondenza della regione distale del gene, e una duplicazione nello pseudogene *STRCP1*: si tratta verosimilmente di un riarrangiamento genico complesso (possibile conversione genica).

Le tecniche utilizzate hanno quindi permesso di individuare le seguenti variazioni:

- NM_153700.2: c.4219-1G>A. Essa consiste in una sostituzione nucleotidica puntiforme di una guanina con una adenina in posizione 4219-1. Ciò implica l'interessamento di una regione intronica, indicata in base alla distanza dal nucleotide esonico più prossimo: in questo caso si ha una sostituzione in corrispondenza di un nucleotide intronico che dista una base dall'esone a valle. Verosimilmente, vista la posizione prossima alla giunzione introne-esone, questa variante costituisce una mutazione che, pur non interessando una porzione codificante del gene, ne altera la funzionalità andando ad inficiare il corretto processo di *splicing*.

In accordo con quanto stabilito secondo i criteri di classificazione ACMG, si tratta di una mutazione patogenetica.

- Possibile conversione genica in fase di accertamento. Sono state infatti individuate una delezione multiesonica di parte di *STRC* in corrispondenza della regione distale del gene, e una duplicazione nello pseudogene *STRCP1* dell'intera porzione deleta nel gene.

In accordo con quanto stabilito secondo i criteri di classificazione ACMG, si tratta di una mutazione patogenetica.

Nella paziente in esame non è stata eseguita la segregazione parentale, pertanto non è possibile determinare le modalità di trasmissione ereditaria delle variazioni individuate.

Non è stato possibile risalire ad informazioni cliniche ed anamnestiche approfondite a riguardo.

6.3.16 PAZIENTE 16

Si tratta di un soggetto di sesso maschile, sottoposto ad analisi genetica all'età di 23 anni, dopo aver eseguito consulenza genetica in AOPD.

È stata sottoposto alle seguenti tecniche di indagine: NGS di pannello multigenico per geni associati ad ipoacusia, ricerca bioinformatica di CNV con software *Agilent SureCall* a partire dai dati ottenuti con NGS, MLPA con Kit apposito alla ricerca di CNV di pertinenza del locus genico a cui appartiene *STRC*.

Il sequenziamento NGS ha evidenziato una profondità di lettura pari a zero in tutte le regioni in cui le *reads* allineano in modo specifico al gene, permettendo così di ipotizzare la presenza di una delezione multiesonica in omozigosi; la successiva analisi MLPA, ha confermato l'ipotesi, evidenziando la presenza in omozigosi della seguente CNV:

- Delezione multiesonica interessante sia il gene *STRC* che il gene *CATSPER2*.

In accordo con quanto stabilito secondo i criteri di classificazione ACMG, si tratta di una mutazione patogenetica.

Non è stata effettuata la segregazione parentale, che sarebbe comunque risultata superflua in quanto il paziente presentava una delezione in omozigosi, pertanto la segregazione parentale non avrebbe consentito l'acquisizione di informazioni aggiuntive.

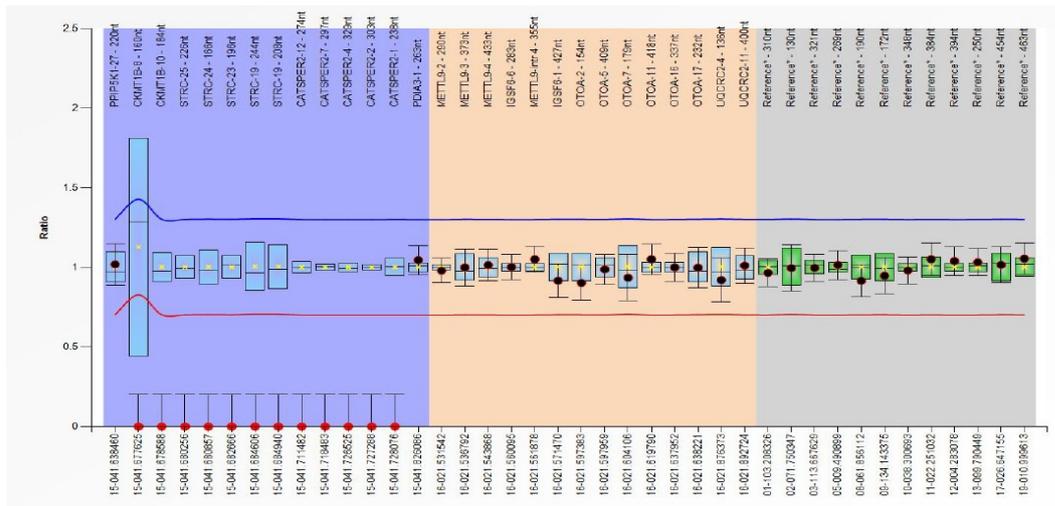


Figura 62. Risultato dell'indagine MLPA eseguita nel paziente 16. Si può osservare la presenza di una delezione in omozigosi, interessante i geni *STRC* e *CATSPER2*.

Per quanto riguarda il fenotipo, il paziente manifestava ipoacusia neurosensoriale bilaterale di grado moderato.

L'esordio è stato post-linguale, non è stato possibile risalire con esattezza all'età in cui si è manifestato il quadro clinico.

Si specifica inoltre come la presenza di delezione omozigote interessante i geni *STRC* e *CATSPER2* determini un fenotipo sindromico nei soggetti di sesso maschile, che si manifesta con ipoacusia e infertilità (sindrome DIS).

6.3.17 PAZIENTE 17

Si tratta di un soggetto di sesso maschile, sottoposto ad analisi genetica all'età di 4 anni, dopo aver eseguito consulenza genetica esternamente ad AOPD.

È stata sottoposto alle seguenti tecniche di indagine: NGS di pannello multigenico per geni associati ad ipoacusia, ricerca bioinformatica di CNV con software *Agilent SureCall* a partire dai dati ottenuti con NGS, LR-PCR associata a NGS con protocollo di arricchimento *Nextera* specifica per gene *STRC*.

Esse hanno permesso di individuare le seguenti variazioni:

- NM_153700.2: c.1673G>A p.(Cys558Tyr). Essa consiste in una sostituzione nucleotidica puntiforme di una guanina con un'adenina in posizione 1673, che determina nella proteina codificata dal gene *STRC* una sostituzione aminoacidica di una cisteina con una tirosina in posizione 558. In accordo con quanto stabilito secondo i criteri di classificazione ACMG, si tratta di una variante di significato incerto (VUS)

- NM_153700.2: c.1447C>T p.(Arg483*). Essa consiste in una sostituzione nucleotidica puntiforme di una citosina con una timina in posizione 1447, che determina nella proteina codificata dal gene *STRC* l'inserimento di un codone di stop prematuro al posto di un codone codificante un'arginina in posizione 483.

In accordo con quanto stabilito secondo i criteri di classificazione ACMG, si tratta di una variante patogenetica.

È stata inoltre effettuata la segregazione parentale, che ha permesso di definire la tipologia di trasmissione delle variazioni individuate e l'appartenenza delle stesse a differenti alleli: la prima variante è risultata essere di origine paterna, la seconda di origine materna.

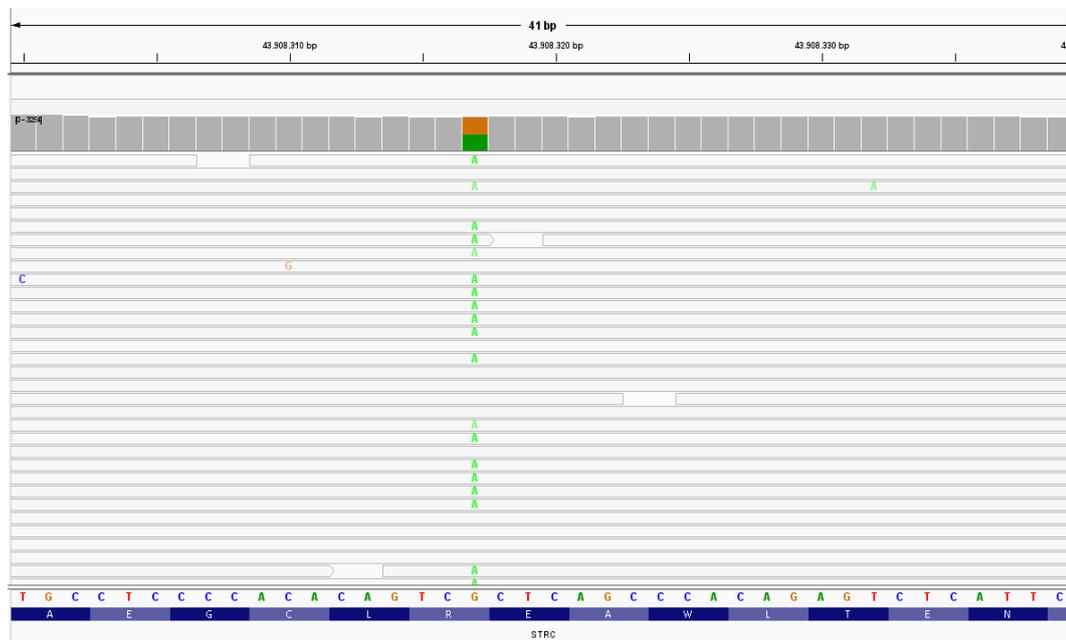


Figura 63. Risultato del sequenziamento NGS effettuato con protocollo di arricchimento Nextera eseguito nel paziente 17. Si può osservare la presenza in eterozigosi della sostituzione puntiforme di origine paterna in eterozigosi.

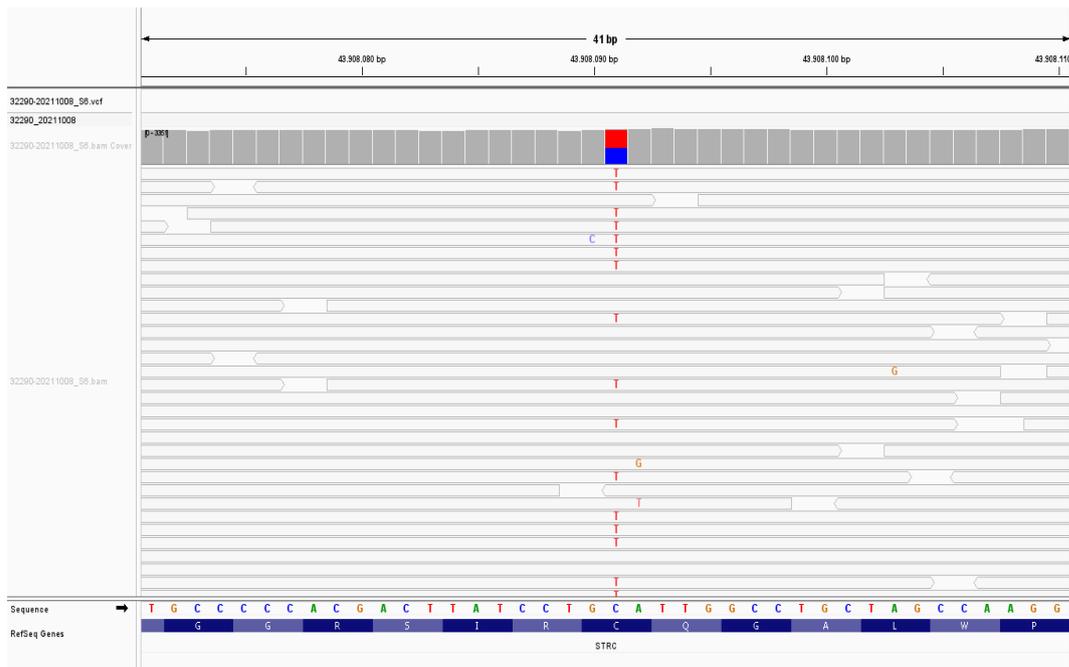


Figura 64. Risultato del sequenziamento NGS effettuato con protocollo di arricchimento Nextera eseguito nel paziente 17. Si può osservare la presenza in eterozigosi della sostituzione puntiforme di origine materna in eterozigosi.

Non è stato possibile risalire ad informazioni cliniche ed anamnestiche approfondite a riguardo, tuttavia sappiamo che l'ipoacusia è caratterizzata da grado lieve, e che l'esordio del quadro clinico è avvenuto in età pre-linguale.

6.3.18 PAZIENTI 18A E 18B

La casistica numero 18 comprende due pazienti, identificati come 18A e 18B, fratelli caratterizzati da genotipo analogo. 18A è un soggetto di sesso femminile, mentre 18B è un soggetto di sesso maschile; essi sono stati sottoposti ad analisi genetica rispettivamente a 18 e 14 anni, dopo aver eseguito una consulenza genetica esternamente ad AOPD.

Essi sono stati sottoposti alle seguenti tecniche di indagine: NGS di pannello multigenico per geni associati ad ipoacusia, ricerca bioinformatica di CNV con software *Agilent SureCall* a partire dai dati ottenuti con NGS, LR-PCR associata a NGS con protocollo di arricchimento *Nextera* specifica per gene *STRC*.

Esse hanno permesso di individuare le seguenti variazioni:

- NM_153700.2: c.1583C>T p.(Pro528Leu). Essa consiste in una sostituzione nucleotidica puntiforme di una citosina con una timina in posizione 1583, che determina nella proteina codificata dal gene *STRC* una sostituzione aminoacidica di una prolina con una leucina in posizione 528.

In accordo con quanto stabilito secondo i criteri di classificazione ACMG, si tratta di una variante di significato incerto (VUS).

- NM_153700.2: c.1002G>A p.(=). Essa consiste in una sostituzione nucleotidica puntiforme di una guanina con un'adenina in posizione 1002: si tratta di una variante sinonima in quanto non determina nella proteina codificata dal gene *STRC* un'alterazione diretta della sequenza aminoacidica.

Non e' tuttavia escludibile che la variazione possa alterare il meccanismo di *splicing* del trascritto, per quanto alcuni software predittivi non supportino tale ipotesi.

In accordo con quanto stabilito secondo i criteri di classificazione ACMG, si tratta di una variante di significato incerto (VUS).

Nei pazienti in esame non è stata eseguita la segregazione parentale, pertanto non è possibile determinare le modalità di trasmissione ereditaria delle variazioni individuate e l'appartenenza delle stesse a differenti alleli.

Non è stato possibile risalire ad informazioni cliniche ed anamnestiche approfondite a riguardo, tuttavia sappiamo che l'ipoacusia è caratterizzata da grado severo.

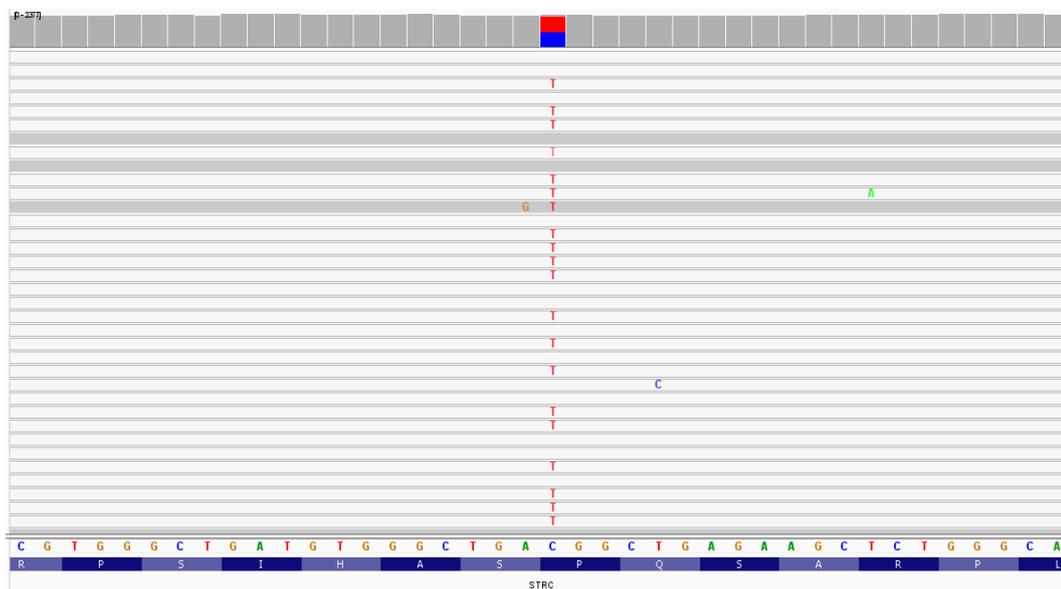


Figura 65. Risultato del sequenziamento NGS effettuato con protocollo di arricchimento Nextera eseguito nel paziente 18. Si può osservare la presenza in eterozigosi della sostituzione puntiforme missenso.

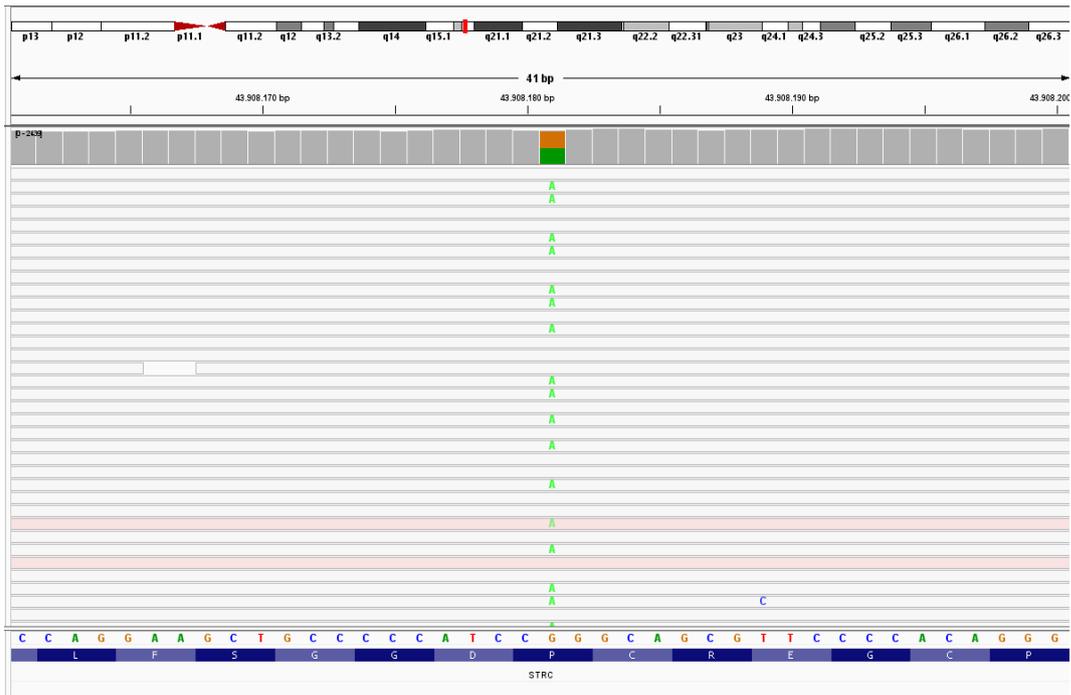


Figura 66. Risultato del sequenziamento NGS effettuato con protocollo di arricchimento Nextera eseguito nel paziente 18. Si può osservare la presenza in eterozigosi della sostituzione puntiforme sinonima.

7. DISCUSSIONE E CONCLUSIONI

7.1 FENOTIPO AUDIOLOGICO E GENOTIPO

Sulla base dei risultati a disposizione riguardo le manifestazioni fenotipiche dei soggetti indagati, emerge un quadro complessivo che va a confermare quanto già noto in letteratura riguardo l'ipoacusia correlata a mutazione di *STRC*. (21)

La perdita di udito è infatti risultata sempre bilaterale, di grado principalmente variabile tra lieve e moderato, ed esordio principalmente prelinguale.

La severità del quadro clinico è stata documentata in 12 soggetti: essa è risultata lieve in 4 di essi (33%), lieve-moderata in 3 (25%), moderata in 2 (17%), severa in 3 (25%). In accordo coi dati disponibili in letteratura, nessun caso è caratterizzato da ipoacusia di grado profondo e solo una porzione di casi pari al 25% è caratterizzata da ipoacusia di grado severo; la maggior parte degli affetti presenta un deficit audiologico variabile da lieve a moderato.

Ciò trova conferma anche nella tipologia di trattamento effettuata. Tra i casi in cui essa è stata documentata con certezza, infatti, nessuno è stato sottoposto ad impianto cocleare, dispositivo di scelta nei pazienti presentanti ipoacusia di grado profondo; tutti i pazienti trattati utilizzavano invece protesi acustiche.

I dati relativi alle modalità di esordio del deficit sono stati invece reperiti in 13 soggetti: tra questi, 7 (54%) hanno presentato un'insorgenza di tipo prelinguale, i restanti 6 (46%) di tipo postlinguale. Questi dati, per quanto concordi con quelli disponibili in letteratura riguardo la maggior frequenza dell'insorgenza prelinguale, non risultano fortemente dirimenti e suggeriscono che le due modalità d'esordio siano pressoché epidemiologicamente sovrapponibili. È importante evidenziare in tal senso l'eventualità che ci sia stato un *bias* nella classificazione di alcune tra le ipoacusie definite ad esordio postlinguale: è infatti possibile che parte di questi casi fosse affetta da un deficit molto a lieve prelinguale rimasto tuttavia misconosciuto. La ricerca delle varianti mutazionali in ciascuno dei pazienti facenti parte del campione oggetto di studio ha permesso l'identificazione di 36 varianti alleliche: 16 (44%) erano delezioni multiesoniche, di cui 13 (81%) interessavano sia i geni *STRC* che *CATSPER2* mentre 3 (19%) interessavano solo *STRC*; 3 (8%) erano

verosimili conversioni geniche, di cui una confermata e le altre in fase di accertamento; le restanti 17 (48%) erano piccole varianti introniche o esoniche.

L'elevata frequenza di CNV a carico della regione interessante il gene *STRC* va a confermare quanto già osservato in precedenza da vari altri studiosi del settore (22).

Le ragioni di tale tendenza sono state caratterizzate anche dal punto di vista teorico e risiedono nella particolare struttura del locus genico 15q15.3, ospitante il gene: come ampiamente trattato nel paragrafo 3.2.3, la regione genomica che contiene il gene *STRC* risulta pressochè completamente duplicata, vista la presenza a valle di uno pseudogene *pSTRC* caratterizzato da elevatissima omologia di sequenza. L'omologia di sequenza tra il gene e lo pseudogene favorisce quindi l'insorgenza di riarrangiamenti strutturali più o meno complessi, che possono essere copy number variations (CNV) quali duplicazioni e delezioni, originatisi principalmente in seguito a fenomeni di ricombinazione omologa non allelica; oppure conversioni geniche, che si sviluppano a causa del trasferimento di informazioni genetiche dal gene allo pseudogene, o viceversa, durante il processo di meiosi o di riparazione del DNA.

Nel caso specifico di *STRC*, invero, in accordo sia con quanto osservato nello studio oggetto dell'elaborato sia con le informazioni presenti in letteratura (23), sembrerebbe che le CNV siano in massima parte delezioni, mentre le duplicazioni risultano di gran lunga più rare. Le ragioni di tale osservazione non sono ancora del tutto chiare: è possibile che ciò sia causa del fatto che esse non determinino l'insorgenza di un fenotipo patologico nei soggetti che le presentano.

Un ulteriore spunto di discussione che emerge dalla casistica considerata nello studio consiste nell'analisi dell'estensione delle delezioni individuate; nella maggior parte dei casi (81%) risulta essere interessato anche il gene *CATSPER2*, posto a monte di *STRC*, mentre solo il 19% interessa quest'ultimo in maniera esclusiva.

Mentre possiamo assumere che le delezioni multiesoniche interessanti sia *STRC* che *CATSPER2* abbiano un'estensione che ricopre l'interezza dei geni, risulta molto complicato, alla luce dei limiti propri delle tecniche analitiche ad oggi disponibili, definire l'estensione delle delezioni multiesoniche che apparentemente interessano il solo gene *STRC*. È inoltre verosimile che alcune delezioni di minori dimensioni non riescano ad essere individuate tramite le metodiche d'esame utilizzate nello studio e nella pratica clinica. La tecnica maggiormente accurata impiegata di

routine per individuare le CNV è infatti costituita dall'MLPA: essa tuttavia è in grado di esaminare esclusivamente determinati esoni del gene, in quanto le sonde in essa utilizzate sono strutturate in modo tale da riconoscere alcune tra le poche aree di non omologia tra gene e pseudogene, impedendo di conseguenza un'analisi esaustiva dell'intera sequenza genica, e determinando eventualmente l'impossibilità di individuare determinate delezioni che riguardano esoni caratterizzati da pressochè completa omologia di sequenza con lo pseudogene. Queste criticità possono tuttavia essere superate utilizzando tecniche di sequenziamento di terza generazione come *Oxford Nanopore*.

Oltre alle CNV, sono stati individuati nella coorte oggetto di studio tre riarrangiamenti genici complessi, di cui uno è stato caratterizzato molecularmente come conversione genica grazie all'impiego del sequenziamento di terza generazione, mentre gli altri due sono ancora in fase di conferma; si ritiene comunque che anche questi siano conversioni geniche.

Il caso confermato consiste nel primo esempio di caratterizzazione molecolare di una conversione a carico del gene *STRC*: ciò potrebbe sembrare controintuitivo, dal momento in cui in un campione di dimensioni relativamente ridotte si presume che le conversioni geniche siano ben 3 (8%). La motivazione per cui prima d'ora nessun altro era riuscito a caratterizzare molecularmente una conversione genica di *STRC* non risiede infatti nella rarità della variazione, che, come anticipato in precedenza, tende a formarsi con relativa facilità a causa delle caratteristiche intrinseche della regione genomica d'interesse; bensì nei limiti che presentano le metodologie d'indagine tradizionali che vengono utilizzate nella *routine* laboratoristica, che permettono solamente di ipotizzare in maniera speculativa la presenza di conversioni geniche di *STRC*, senza però poter dare conferme concrete: per poter far ciò risulta necessario l'impiego di tecniche di sequenziamento di terza generazione.

L'analisi bioinformatica delle CNV e l'MLPA, infatti, non riescono ad identificare eventuali conversioni geniche, interpretandole tutt'al più come delezioni, se non addirittura non segnalando alcun sbilanciamento a causa dell'elevata omologia di sequenza tra la regione di provenienza del materiale genetico (pseudogene) e quella di arrivo (gene): sta pertanto all'esperienza ed alla preparazione del professionista incaricato dell'interpretazione dei risultati delle analisi effettuate sollevare un

dubbio riguardante l'effettiva veridicità del responso restituito e, se necessario e possibile, richiedere un sequenziamento di terza generazione.

Sulla base di quanto discusso, è possibile ipotizzare che molte delle varianti che sono state classificate in letteratura come delezioni multiesoniche isolate a carico del gene *STRC* costituiscano in realtà delle conversioni geniche misconosciute.

Nel complesso, 17 delle 36 varianti alleliche individuate nei pazienti studiati (48%) sono risultate essere piccole mutazioni interessanti singoli nucleotidi o un numero ridotto di nucleotidi all'interno dello stesso esone o dello stesso introne. Queste varianti sono state quindi sottoclassificate sulla base della loro localizzazione in introniche (12%) ed esoniche (88%); sulla base del loro effetto sulla proteina tradotta in missenso (70%), troncanti (24%) e sinonime (6%). Al netto di tutto ciò, è necessario considerare come alcune varianti (12%), classificate come sinonime o missenso, possano, vista la loro localizzazione in prossimità di una giunzione introne-esone, verosimilmente essere coinvolte in alterazioni del meccanismo di *splicing*, determinando pertanto effetti ben più impattanti nella proteina tradotta.

Chiaramente, mentre il ruolo patogenetico delle varianti strutturali, interessanti porzioni considerevoli del gene se non l'intero gene, risulta facilmente deducibile poiché queste impediscono completamente la sintesi di stereocilina, la valutazione del grado di patogenicità delle varianti intraesoniche ed introniche rappresenta un processo molto più articolato, che talvolta richiede l'utilizzo di complessi *software* bioinformatici.

A tale scopo, sono disponibili delle linee guida, stilate dall'*American College of Medical Genetics* (ACMG) (71), che, sulla base di molteplici parametri quali la frequenza allelica, i dati presenti in letteratura, la conservazione evolutiva delle sequenze e il tipo di mutazione, permettono al professionista, incaricato dell'interpretazione dei risultati delle analisi effettuate, di classificare le varianti individuate sulla base del loro grado di patogenicità in: varianti benigne, varianti verosimilmente benigne, varianti di significato incerto, varianti verosimilmente patogenetiche e varianti patogenetiche.

Le varianti troncanti (varianti non senso e varianti frameshift) e le varianti di *splicing* che interessano i nucleotidi ± 1 e ± 2 della giunzione introne-esone, esitando nella maggior parte dei casi nella produzione di una proteina tronca non funzionale,

vengono agevolmente classificate come patogenetiche nella gran maggioranza dei casi.

Per quanto riguarda, invece, varianti missenso, varianti sinonime e varianti di *splicing* che non interessano i siti canonici di *splicing* sovracitati, il processo di definizione del grado di patogenicità risulta molto più difficoltoso ed esita frequentemente nella rilevazione di varianti di significato incerto (VUS), poiché non è chiaro il loro effetto sulla funzionalità della proteina e/o non è presente in letteratura una casistica sufficientemente ampia.

Tra le varianti non strutturali individuate nei pazienti oggetto di studio, è stato possibile comunque classificare il 47% delle stesse come patogenetiche, il 18% come verosimilmente patogenetiche, e solo il 35% come varianti di significato incerto, raggiungendo pertanto un risultato soddisfacente, alla luce delle difficoltà analitiche discusse.

Diversi pazienti, d'altro canto, presentavano ulteriori VUS, che non sono state però tenute in considerazione ai fini dello studio: la distinzione tra quali VUS reputare accettabili e quali no è stata effettuata sulla base della loro probabilità di incidere nella patogenesi del quadro clinico calcolata dai *software* predittivi ed interpretativi utilizzati, e sulla base della loro posizione all'interno del gene, ritenuta più o meno prona ad esitare in un'alterazione del processo di *splicing* e/o in alterazioni della funzionalità della proteina mutante.

7.2 PROTOCOLLO DI ANALISI

In funzione delle ben note criticità associate all'indagine genetica del locus 15q15.3, le metodiche impiegate al fine di individuare variazioni a carico del gene *STRC* nei soggetti facenti parte della coorte oggetto di studio sono state molteplici. In ciascuno dei pazienti è stato eseguito il sequenziamento NGS di un pannello di geni associati ad ipoacusia non sindromica, effettuato con metodiche di arricchimento tradizionali (protocolli *SureSelect* o *Haloplex*), e, a partire dai dati ottenuti col sequenziamento NGS, l'analisi bioinformatica delle CNV, svolta impiegando il software dedicato *SureCall*.

Sono stati quindi eseguiti approfondimenti specifici in base alle caratteristiche delle eventuali varianti suggerite dalle indagini di primo livello: tra questi, MLPA, LR-PCR specifica per il gene *STRC* associata a sequenziamento Sanger o

sequenziamento NGS con protocollo di arricchimento *Nextera*, sequenziamento di terza generazione Oxford Nanopore.

Grazie all'utilizzo delle sovraccitate tecnologie, si è riusciti a caratterizzare con relativa precisione le varianti mutazionali presenti in ognuno dei pazienti studiati: considerando quanto sia difficoltosa l'indagine del gene *STRC*, si tratta di un risultato assolutamente soddisfacente, che ci permette di consigliare il protocollo di analisi impiegato e di riproporlo in futuro.

7.2.1 SEQUENZIAMENTO NGS E ANALISI DELLE CNV

I pazienti appartenenti alla coorte di studio sono stati in prima battuta sottoposti a sequenziamento NGS, effettuato con protocollo di arricchimento *SureSelect* o *Haloplex*, di un pannello multigenico correlato ad ipoacusie non sindromiche.

Alla luce delle peculiarità della regione genomica d'interesse, caratterizzata dalla presenza a valle di uno pseudogene (*pSTRC*) ad elevatissima omologia di sequenza (99%), l'impiego di tecniche di primo livello quali il sequenziamento di seconda generazione non risulta sufficiente ad effettuare una diagnosi nella maggior parte dei casi, in quanto non in grado di discernere se le varianti individuate siano effettivamente di pertinenza genica piuttosto che pseudogenica, a meno che le stesse non si presentino proprio in corrispondenza delle ridotte porzioni di non omologia. Il sequenziamento NGS determina, infatti, la produzione di *reads* di dimensione pari a circa 150 paia di basi, lunghezza non adatta per un allineamento univoco al gene.

Al fine di superare le criticità legate al discernimento tra varianti geniche e pseudogeniche, è stato possibile amplificare selettivamente il gene mediante un'unica reazione di *Long Range PCR*, evitando quindi la "contaminazione" del campione da parte di sequenze dello pseudogene. È stato quindi effettuato il sequenziamento dell'intero amplificato mediante NGS con protocollo di arricchimento *Nextera* (caratterizzato da maggior profondità di lettura, che arriva fino a 10000 *reads*) oppure della sola regione di interesse contenente la variante da confermare (in tal caso è necessaria una ulteriore amplificazione con PCR) mediante metodo Sanger. Questa procedura ha permesso di confermare con maggiore precisione e accuratezza le piccole varianti esoniche e introniche che erano state suggerite dal sequenziamento NGS del pannello multigenico.

A partire dai dati grezzi ottenuti con il sequenziamento NGS è stata quindi eseguita, in tutti i pazienti, l'analisi bioinformatica delle CNV, al fine di rilevare l'eventuale presenza di riarrangiamenti strutturali maggiori a carico di *STRC*.

Tale metodica, che sfrutta l'impiego del software *SureCall*, permette di arginare uno dei principali limiti del sequenziamento di seconda generazione, restituendo una stima della presenza di CNV sulla base di dati generati per l'individuazione di piccole variazioni di sequenza. Il software bioinformatico è infatti in grado di evidenziare, sulla base di valutazioni statistiche legate alla profondità di lettura del sequenziamento effettuato nel caso in studio rispetto a quello in uno o più controlli, la presenza di variazioni quantitative quali delezioni o duplicazioni esoniche o multiesoniche.

Si tratta di una stima approssimativa, non sufficiente alla formulazione di una diagnosi certa, e che richiede pertanto analisi di conferma mediante altre metodiche, come l'analisi MLPA: nel contesto dello studio, l'esecuzione di MLPA ha comunque restituito risultati pressochè sovrapponibili all'analisi bioinformatica, indicando l'elevata accuratezza di quest'ultima come indagine di screening in pazienti con ipoacusia genetica.

L'utilizzo di software bioinformatici, inoltre, pur non essendo in grado di rilevare direttamente riarrangiamenti complessi come le conversioni geniche, poiché le interpreta tutt'al più come CNV, permette al professionista incaricato della refertazione di formulare delle ipotesi riguardo alla loro presenza: esaminando scrupolosamente gli sbilanciamenti individuati a carico sia del gene che dello pseudogene, e osservando l'eventuale reciprocità tra le stesse, è infatti possibile postulare, in maniera speculativa e da confermare con metodiche più adeguate, l'avvenimento di fenomeni di conversione genica.

I 13 pazienti in cui è stata rilevata mediante analisi bioinformatica la possibile presenza di CNV a carico del gene *STRC* sono sempre stati approfonditi con metodica MLPA, eccetto che in tre casi in cui ciò non è stato possibile per ragioni logistiche.

Si tratta di una tecnica caratterizzata da elevatissima accuratezza e precisione nella rilevazione di CNV, tanto da essere generalmente in grado di individuare alterazioni a carico di singoli esoni.

Ciononostante, l'impiego dell'analisi MLPA nel contesto del locus 15q15.3 presenta comunque diversi limiti correlati alle caratteristiche intrinseche della regione. Al fine di individuare correttamente le varianti geniche, senza incorrere in errori legati all'erroneo coinvolgimento dello pseudogene, le sonde del *kit* specifico per la regione genomica d'interesse vengono disegnate appositamente per appaiarsi a sequenze di pertinenza esclusiva di *STRC*, sfruttando l'1% di non omologia con *pSTRC*. Ciò si traduce nella possibilità di analizzare solamente alcune porzioni geniche, corrispondenti agli esoni 19, 20, 23, 24, 25 e 28, tralasciando eventuali CNV a carico delle porzioni restanti, che dovranno essere indagate tramite metodiche alternative.

Così come l'analisi bioinformatica con software *SureCall*, anche l'MLPA non è stata concepita per rilevare riarrangiamenti strutturali complessi quali le conversioni geniche, ma, ancora una volta, un'attenta interpretazione delle varianti individuate può fare emergere informazioni che indirizzino l'ipotesi diagnostica in tal senso.

7.2.2 SEQUENZIAMENTO DI TERZA GENERAZIONE OXFORD NANOPORE

I 3 soggetti in cui, tramite l'impiego delle metodiche appena discusse, è stata ipotizzata la presenza di conversioni a carico del gene *STRC* sono stati sottoposti a sequenziamento di terza generazione *Oxford Nanopore*: esso ha permesso di confermare la conversione genica in uno dei pazienti, mentre i restanti due sono ancora in fase di accertamento dei risultati ottenuti.

La peculiarità di questa tipologia di sequenziamento, che permette di riuscire ad individuare in maniera precisa varianti altrimenti non caratterizzabili in *STRC*, risiede nella sua capacità di produrre *reads* che possono raggiungere la lunghezza di decine di migliaia di basi: è pertanto possibile raccogliere l'intera sequenza del gene in un'unica *read*, garantendo di conseguenza un'agevole discriminazione delle varianti geniche da quelle pseudogeniche.

Risulta quindi possibile valutare varianti strutturali complesse come le conversioni geniche, riuscendo inoltre ad individuare i punti di rottura del riarrangiamento cromosomico.

Nonostante gli indiscutibili pregi del sequenziamento di terza generazione, che hanno permesso di caratterizzare per la prima volta i punti di rottura di una

conversione genica a carico di *STRC*, questa metodologia d'indagine presenta ancora diverse criticità, su cui sarà necessario lavorare in futuro al fine di permetterne l'utilizzo in contesti di *routine* diagnostica.

Si tratta infatti di una tecnologia dispendiosa in termini di tempo e denaro, sia dal punto di vista operativo che per quanto riguarda le strumentazioni necessarie ad interpretarne i risultati. Sono richiesti *tool* bioinformatici molto complessi, il cui utilizzo, a differenza dei *software* impiegati nel sequenziamento NGS, richiede il coinvolgimento di computer potenti, nonché il possesso di competenze tecniche altamente specialistiche: in particolare è necessario, al fine dell'interpretazione dei dati generati dal sequenziamento *Oxford Nanopore*, l'apporto professionale di un esperto bioinformatico.

Un'ulteriore criticità del sequenziamento di terza generazione di *STRC*, individuata nel corso dello studio, sta nella necessità di eseguire un arricchimento del materiale genomico da analizzare. Nonostante teoricamente uno dei principali vantaggi del *Nanopore* sia infatti costituito dalla possibilità di sequenziare DNA nativo (*Adaptive sampling*), di fatto la profondità di lettura che si è riusciti ad ottenere a partire da quest'ultimo, pari a circa 3 *reads* per il gene e 8 per lo pseudogene, non è risultata sufficiente, ed è stato necessario eseguire un arricchimento della regione genomica d'interesse tramite *Long Range PCR*. Questo accorgimento, tuttavia, fa sì che eventuali errori verificatisi durante il processo di amplificazione si ripercuotano sul risultato finale dell'analisi, inficiandone la precisione e l'accuratezza.

Infine, il sequenziamento *Oxford Nanopore* rappresenta una metodica soggetta ad elevato tasso d'errore, che arriva fino al 15%, e riguarda principalmente le varianti esoniche di dimensioni minori, quali inserzioni, delezioni e, in minor parte, sostituzioni. Per questo motivo è consigliato modificare alcune impostazioni al fine di mascherare le indels inferiori a 10 paia di basi e i mismatch che presentano una VAF (*Variant Allele Frequency*) inferiore al 25%: in questo modo si ottiene un dato pulito da eventuali errori di sequenziamento.



Figura 67. Dato grezzo ottenuto dal sequenziamento di STRC con Oxford Nanopore. Si può vedere come le reads siano ricche di errori tipo inserzioni e delezioni.

7.3 RISVOLTI CLINICI DELLO STUDIO

L'analisi mutazionale condotta nella coorte oggetto di studio tramite metodiche di sequenziamento di seconda e terza generazione ha permesso l'individuazione delle varianti che, verosimilmente, determinano il quadro di ipoacusia neurosensoriale di ciascun paziente. La possibilità di caratterizzare puntualmente le mutazioni patogenetiche presenti nei soggetti affetti, anche in caso queste interessino regioni genomiche tipicamente ostili all'indagine di *routine*, presenta importanti risvolti nella pratica clinica.

Per quanto il conoscenza delle esatte varianti causative del quadro fenotipico non influenzi particolarmente le strategie applicate in termini di terapia e di sorveglianza, ciò si rivela invece fondamentale in ottica preventiva.

Una volta formulata la diagnosi eziologica, infatti, il medico, nel contesto della fase post-test della consulenza genetica, comunica al paziente e ai familiari le modalità di trasmissione ereditaria dell'ipoacusia che interessa il probando, fornendo agli stessi le informazioni necessarie alla comprensione del concetto di rischio di ricorrenza, ovvero la probabilità che il quadro patologico si ripresenti a carico della discendenza. Dopo aver quantificato il rischio di ricorrenza per ciascun membro della famiglia, grazie anche allo svolgimento della segregazione parentale, lo

specialista illustra le modalità di prevenzione del rischio di cui possono eventualmente usufruire gli interessati.

7.4 CONCLUSIONI

Lo studio effettuato ha permesso la caratterizzazione fenotipica e genotipica di una coorte di 19 pazienti affetti da ipoacusia neurosensoriale causata da alterazioni a carico del gene *STRC*: sia il quadro fenotipico che quello genotipico sono risultati tendenzialmente in linea con le informazioni disponibili in letteratura, permettendo quindi di consolidare le conoscenze acquisite riguardo la patologia d'interesse. In particolare, sono state individuate 36 varianti geniche, comprendenti varianti esoniche, varianti introniche, delezioni multiesoniche e conversioni geniche; si tratta inoltre del primo esempio presente in letteratura di caratterizzazione di una conversione genica a carico di *STRC*.

Alla luce delle criticità correlate all'indagine del *locus* 15q15.3, contenente il gene *STRC*, caratterizzato dalla presenza a valle di uno pseudogene *pSTRC* ad elevatissima omologia di sequenza (99%), i risultati ottenuti sono stati reputati soddisfacenti, e permettono di raccomandare il protocollo d'analisi impiegato e di riproporlo in futuro. Va comunque sottolineato come si sia reso necessario l'utilizzo di metodologie d'analisi avanzate quali il sequenziamento di terza generazione *Oxford Nanopore*, rivelatosi fondamentale nell'indagine delle conversioni geniche. Non è ancora possibile introdurre tale analisi nel protocollo diagnostico di *routine*, ma è necessario considerare l'impiego dello stesso nel caso gli esami di primo livello non si rivelino sufficienti alla formulazione di una corretta diagnosi.

8. BIBLIOGRAFIA

1. *International classification of impairments, disabilities and handicaps (ICIDH): a manual of classification relating to the consequences of diseases*. Organization, World Health. 1980, Vol. https://apps.who.int/iris/bitstream/handle/10665/41003/9241541261_eng.pdf?sequence=1
2. Organization, World Health. Deafness and hearing loss. https://www.who.int/health-topics/hearing-loss#tab=tab_1. 2023.
3. (NIDCD), National Institute on Deafness and Other Communication Disorders. Presbycusis Health Statistics. . <http://www.nidcd.nih.gov/health/hearing/presbycusis.htm>. 2011.
4. Robertson CM, Howarth TM, Bork DL, Dinu IA. Permanent bilateral sensory and neural hearing loss of children after neonatal intensive care because of extreme prematurity: a thirty-year study. *Pediatrics*. 2009, Vol. 123, pp. 797-807.
5. Doyle LW, Keir E, Kitchen WH, Ford GW, Rickards AL, Kelly EA. Audiologic assessment of extremely low birth weight infants: a preliminary report. *Pediatrics*. 1992, Vol. 90, pp. 744-749.
6. Morales AC, Azuara BN, Gallo TJ, Gonzalez AA, Rama QJ. Sensorineural hearing loss in cerebral palsy patients. *Acta Otorrinolaringol Esp*. 2006, Vol. 57, 7, pp. 300-302.
7. Reid SM, Modak MB, Berkowitz RG Reddihough DS. A population-based study and systematic review of hearing loss in children with cerebral palsy cohort. *Dev Medicine & Child Neurology*. 2011, Vol. 53, pp. 1038-1045.
8. Yoshinaga-Itano C, Coulter D, Thomson V. Developmental outcomes of children with hearing loss born in Colorado hospitals with and without universal newborn hearing screening programs. *Seminars in Neonatology*. 2001, Vol. 6, pp. 521-529.
9. Alshuaib WB, Al-Kandari JM, Hasan SM. Classification of hearing loss. [book auth.] Fayez Bahmad Jr. *Update on hearing loss*. s.l. : InTech, 2015.
10. Cunningham LL, Tucci DL. Hearing Loss in Adults. *New England Journal of Medicine*. 2017, Vol. 377, 25, pp. 2465-2473.
11. Angeli S, Lin X, Liu XZ. Genetics of hearing and deafness. *Anat Rec (Hoboken)*. 2012, Vol. 295, 11, pp. 1812-1829.
12. Yoshinaga-Itano C, Sedey AL, Coulter DK, Mehl AL. Language of early- and later-identified children with hearing loss. *Pediatrics*. 1998, Vol. 102, pp. 1161-1171.
13. Sheffield AM, Smith RJH. he Epidemiology of Deafness. *Cold Spring Harb Perspect Med*. 2019, Vol. 9, 9.
14. *Genetic Hearing Loss Overview*. Shearer E, Hildebrand MS, Smith RJH. s.l. : GeneReviews, 1999 (last update 2023).
15. homepage, Hereditary hearing loss. Hereditary hearing loss. <https://hereditaryhearingloss.org/>. 2021, pp. 131-142.

16. Chan DK, Chang KW. GJB2-Associated Hearing Loss: Systematic Review of Worldwide Prevalence. *Laryngoscope*. 2014, Vol. 124, pp. 34-53.
17. Francey LJ, Conlin, LK, Kadesch HE, Clark D, Berrodin, D, Sun Y, et al. Genome-wide SNP genotyping identifies the stereocilin (STRC) gene as a major contributor to pediatric bilateral sensorineural hearing impairment. *American Journal of Medical Genetics*. 2012, Vol. 158A, 2, pp. 298-308.
18. Verpy E, Masmoudi S, Zwaenepoel I, Leibovici M, Hutchin TP, DelCastillo 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. *Nature Genetics*. 2001, 29, pp. 345-349.
19. Elisabeth Verpy, Michel Leibovici, Nicolas Michalski, Richard J. Goodyear, Carine Houdon, Dominique Weil, Guy P. Richardson and Christine Petit. Stereocilin connects outer-hair-cell stereocilia to one another and to the tectorial membrane. *Journal of comparative neurology*. 2011, Vol. 2, 519, pp. 194-210.
20. Kennedy HJ, Crawford AC, Fettiplace R. Force generation by mammalian hair bundles supports a role in cochlear amplification. *Nature*. 2005, Vol. 7028, 433, pp. 880-883.
21. Shuang Han, Dejun Zhang, Yingyuan Guo, Zeming Fu and Guofang Guan. Prevalence and Characteristics of STRC Gene Mutations (DFNB16): A Systematic Review and Meta-Analysis. *Frontiers in genetics*. 2021, Vol. 12.
22. 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. Detection and Confirmation of Deafness-Causing Copy Number Variations in the STRC Gene by Massively Parallel Sequencing and Comparative Genomic Hybridization. *Annals of Otolaryngology, Rhinology and Laryngology*. 2016, Vol. 125, 11, pp. 918-923.
23. Shearer AE, Kolbe DL, Azaiez H, Sloan CM, Frees KL, Weaver AE, Clark ET, Nishimura CJ, Black-Ziegelbein EA, Smith RJ. Copy number variants are a common cause of non-syndromic hearing loss. *Genome Medicine*. 2014, Vol. 6, 5.
24. Feng Zhang, Wenli Gu, Matthew E. Hurles, and James R. Lupski. Copy Number Variation in Human Health, Disease, and Evolution. *The Annual Review of Genomics and Human Genetics*. 2009, Vol. 10, 451.
25. Gu W, Zhang F, Lupski JR. Mechanisms for human genomic rearrangements. *Pathogenetics*. 2008, Vol. 1, 1.
26. Matthew M Parks, Charles E Lawrence & Benjamin J Raphael. Detecting non-allelic homologous recombination from high-throughput sequencing data. *Genome Biology*. 2015, Vol. 16, 72.
27. Howard H. Y. Chang, Nicholas R. Pannunzio, Noritaka Adachi & Michael R. Lieber. Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nature Reviews Molecular Cell Biology* volume. 2017, 18, pp. 495-506.
28. Feng Zhang, Mehrdad Khajavi, Anne M Connolly, Charles F Towne, Sat Dev Batish, James R Lupski. The DNA replication FoSTeS/MMBIR mechanism can generate genomic, genic and exonic complex rearrangements in humans. *Nature Genetics*. 2009, Vol. 41, 7, pp. 849-853.

29. A. P. Jason de Koning, Wanjun Gu, Todd A. Castoe, Mark A. Batzer, David D. Pollock. Repetitive Elements May Comprise Over Two-Thirds of the Human Genome. *PLoS Genetics*. 2011, Vol. 10, 1371.
30. Perry, G.H. The evolutionary significance of copy number variation in the human genome. *Cytogenetic and Genome Research*. 2009, Vol. 123, 1-4, pp. 283-287.
31. Jian-Min Chen, Claude Férec and David N. Cooper. Gene Conversion in Human Genetic Disease. *Genes (Basel)*. 2010, Vol. 1, 3, pp. 550-563.
32. Hastings, P.J. Conversion Mechanisms of Ectopic Gene. *Genes*. 2010, Vol. 1, 3, pp. 427-439.
33. Back D, Shehata-Dieler W, Vona B, Hofrichter MAH, Schroeder J, Haaf T, Rahne T, Hagen R, Schraven SP. Phenotypic Characterization of DFNB16- associated Hearing Loss. *Otology&Neurotology*. 2019, Vol. 40, 1, pp. 48-55.
34. Čada Z, Šafka Brožková D, Balatková Z, Plevová P, Rašková D, Laštůvková J, Černý R, Bandúrová V, Koucký V, Hrubá S, Komarc M, Jenčík J, Poisson Marková S, Plzák J, Kluh J, Seeman P. Moderate sensorineural hearing loss is typical for DFNB16 caused by various types of mutations affecting the STRC gene. *European Archives of Otorhinolaryngology*. 2019, Vol. 276, 12, pp. 3353-3358.
35. 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 JS. CATSPER2, a human autosomal nonsyndromic male infertility gene. *European Journal of Human Genetics*. 2003, Vol. 11, 7, pp. 497-502.
36. Zhang Y, Malekpour M, Al-Madani N, Kahrizi K, Zanganeh M, Lohr NJ, Mohseni M, Mojahedi F, Daneshi A, Najmabadi H, Smith RJ. Sensorineural deafness and male infertility: a contiguous gene deletion syndrome. *Journal of Medical Genetics*. 2007, Vol. 44, pp. 233-240.
37. C Jacq, J R Miller, G G Brownlee. A pseudogene structure in 5S DNA of *Xenopus laevis*. *Cell Press*. 1977, Vol. 12, 1, pp. 109-120.
38. Salmena, Leonardo. Pseudogenes: Four Decades of Discovery. *Methods on Molecular Biology*. 2021, Vol. 2324, pp. 3-18.
39. Robin-Lee Troskie, Geoffrey J Faulkner, Seth W Cheetham. Processed pseudogenes: A substrate for evolutionary innovation: Retrotransposition contributes to genome evolution by propagating pseudogene sequences with rich regulatory potential throughout the genome. *Bioessays*. 2021, Vol. 43, 11.
40. Mikita Suyama, Eoghan Harrington, Peer Bork, David Torrents. Identification and analysis of genes and pseudogenes within duplicated regions in the human and mouse genomes. *PLoS Computational Biology*. 2006, Vol. 2, 6.
41. Zhengdong D Zhang, Adam Frankish, Toby Hunt, Jennifer Harrow, Mark Gerstein. Identification and analysis of unitary pseudogenes: historic and contemporary gene losses in humans and other primates. *Genome Biology*. 2010, Vol. 11, 3.

42. Michael Lynch, Matthew S Ackerman, Jean-Francois Gout, Hongan Long, Way Sung, W Kelley Thomas, Patricia L Foster. Genetic drift, selection and the evolution of the mutation rate. *Nature reviews. Genetics*. 2016, Vol. 17, 11, pp. 704-714.
43. Chen X, Wan L, Wang W, Xi WJ, Yang AG, Wang T. Re-recognition of pseudogenes: From molecular to clinical applications. *Theranostics*. 2020, Vol. 10, 4, pp. 1479-1499.
44. Portin, Petter. The birth and development of the DNA theory of inheritance: sixty years since the discovery of the structure of DNA. *Journal of genetics*. 2014, Vol. 93, 1, pp. 293-302.
45. F Sanger, G M Air, B G Barrell, N L Brown, A R Coulson, C A Fiddes, C A Hutchison, P M Slocombe, M Smith. Nucleotide sequence of bacteriophage phi X174 DNA. *Nature*. 1977, Vol. 265, 5596, pp. 687-695.
46. Francis S Collins, Michael Morgan, Aristides Patrinos. The Human Genome Project: lessons from large-scale biology. *Science*. 2003, Vol. 300, 5617, pp. 286-290.
47. F Sanger, S Nicklen, A R Coulson. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America*. 1977, Vol. 74, 12, pp. 5463-5467.
48. J Zimmermann, H Voss, C Schwager, J Stegemann, W Ansorge. Automated Sanger dideoxy sequencing reaction protocol. *FEBS Letters*. 1988, Vol. 233, 2, pp. 432-436.
49. G, Karki. Sanger's method of gene sequencing. *Online Biology Notes*. 2017, Vols. <https://www.onlinebiologynotes.com/sangers-method-gene-sequencing/>.
50. Jun He, Yang Na, Jiyang Liu. Result of Sanger sequencing for newborn carriers of single heterozygous variants of GJB2 or SLC26A4 gene by genechip analysis. *Chinese Journal of Medical Genetics*. 2020, Vol. 37, 11, pp. 1213-1216.
51. Barton E Slatko, Andrew F Gardner, Frederick M Ausubel. Overview of Next-Generation Sequencing Technologies. *Current Protocols in Molecular Biology*. 2018, Vol. 122, 1.
52. Dale Muzzey, Eric A Evans, Caroline Lieber. Understanding the Basics of NGS: From Mechanism to Variant Calling. *Current genetic medicine reports*. 2015, Vol. 3, 4, pp. 158-165.
53. Steven R Head, H Kiyomi Komori, Sarah A LaMere, Thomas Whisenant, Filip Van Nieuwerburgh, Daniel R Salomon, Phillip Ordoukhanian. Library construction for next-generation sequencing: overviews and challenges. *BioTechniques*. 2014, Vol. 56, 2, pp. 51-54.
54. Gaetan De Waele, Jim Clauwaert, Gerben Menschaert, Willem Waegeman. CpG Transformer for imputation of single-cell methylomes. *Bioinformatics*. 2022, Vol. 38, 3, pp. 597-603.
55. Vinzenz May, Leonard Koch, Björn Fischer-Zirnsak, Denise Horn, Petra Gehle, Uwe Kornak, Dieter Beule, Manuel Holtgrewe. ClearCNV: CNV calling from NGS panel data in the presence of ambiguity and noise. *Bioinformatics*. 2022, Vol. 38, 16, pp. 3871-3876.

56. Haiying Jia, Yunfei Guo, Weiwei Zhao, Kai Wang. Long-range PCR in next-generation sequencing: comparison of six enzymes and evaluation on the MiSeq sequencer. *Scientific Reports*. 2014, Vol. 4, 5737.
57. William Christie, Ron Yadin, Kristy Ip, Kevin W George. Highly Multiplexed, Semiautomated Nextera Next-Generation Sequencing (NGS) Library Preparation. *Methods in molecular biology*. 2020, Vol. 2205, pp. 91-104.
58. Lora J. H. Bean PhD, Birgit Funke PhD, Colleen M. Carlston PhD, Jennifer L. Gannon MD, Sibel Kantarci PhD, Bryan L. Krock PhD, Shulin Zhang MD, PhD & Pinar Bayrak-Toydemir MD, PhD. Diagnostic gene sequencing panels: from design to report—a technical standard of the American College of Medical Genetics and Genomics (ACMG). *Genetics in medicine*. 2019, Vol. 22, pp. 453-461.
59. Vyne van der Schoot, Lonneke Haer-Wigman, Ilse Feenstra, Femke Tammer, Anke J M Oerlemans, Martine P A van Koolwijk, Frans van Agt, Yvonne H J M Arens, Han G Brunner, Lisenka E L M Vissers, Helger G Yntema. Lessons learned from unsolicited findings in clinical exome sequencing of 16,482 individuals. *European journal of human genetics*. 2022, Vol. 30, 2, pp. 170-177.
60. Wang Y, Zhao Y, Bollas A, Wang Y, Au KF. Nanopore sequencing technology, bioinformatics and applications. *Nature biotechnology*. 2021, Vol. 39, 11, pp. 1348-1365.
61. Gilpatrick T, Lee I, Graham JE, Raimondeau E, Bowen R, Heron A, Downs B, Sukumar S, Sedlazeck FJ, Timp W. Targeted nanopore sequencing with Cas9-guided adapter ligation. *Nature biotechnology*. 2020, Vol. 38, 4, pp. 433-438.
62. Martin S, Heavens D, Lan Y, Horsfield S, Clark MD, Leggett RM. Nanopore adaptive sampling: a tool for enrichment of low abundance species in metagenomic samples. *Genome biology*. 2022, Vol. 23, 1.
63. Karst SM, Ziels RM, Kirkegaard RH, Sørensen EA, McDonald D, Zhu Q, Knight R, Albertsen M. High-accuracy long-read amplicon sequences using unique molecular identifiers with Nanopore or PacBio sequencing. *Nature methods*. 2021, Vol. 18, 2, pp. 165-169.
64. Chen, Y., Nie, F., Xie, SQ. et al. Efficient assembly of nanopore reads via highly accurate and intact error correction. *Nature Communications*. 2021, Vol. 12, 60.
65. Cheung SW, Bi W. Novel applications of array comparative genomic hybridization in molecular diagnostics. *Expert review of molecular diagnostic*. 2018, Vol. 18, 6, pp. 531-542.
66. Stuppia L, Antonucci I, Palka G, Gatta V. Use of the MLPA assay in the molecular diagnosis of gene copy number alterations in human genetic diseases. *Int J Mol Sci*. 2012, Vol. 13, 3, pp. 3245-3276.
67. Bunyan DJ, Skinner AC, Ashton EJ, Sillibourne J, Brown T, Collins AL, Cross NC, Harvey JF, Robinson DO. Simultaneous MLPA-based multiplex point mutation and deletion analysis of the dystrophin gene. *Mol Biotechnol*. 2007, Vol. 35, 2, pp. 135-140.
68. Fu X, Shi Y, Ma J, Zhang K, Wang G, Li G, Xiao L, Wang H. Advances of multiplex ligation-dependent probe amplification technology in molecular diagnostics. *Biotechniques*. 2022.

69. Klimontov VV, Koshechkin KA, Orlova NG, Sekacheva MI, Orlov YL. Medical Genetics, Genomics and Bioinformatics-2022. *Int J Mol Sci.* 2023, Vol. 24, 10.
70. Roca I, González-Castro L, Fernández H, Couce ML, Fernández-Marmiesse A. Free-access copy-number variant detection tools for targeted next-generation sequencing data. *Mutat Res Rev Mutat Res.* 2019, 779, pp. 114-125.
71. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, Voelkerding K, Rehm HL. ACMG Laboratory Quality Assurance Committee. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med.* 2015, Vol. 17, 5.
72. Riggs ER, Andersen EF, Cherry AM, Kantarci S, Kearney H, Patel A, Raca G, Ritter DI, South ST, Thorland EC, Pineda-Alvarez D, Aradhya S, Martin CL. Technical standards for the interpretation and reporting of constitutional copy-number variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics (ACMG) and the Clinical Genome Resource (ClinGen). *Genet Med.* 2020, Vol. 22, 2, pp. 245-257.
73. Mandelker D, Amr SS, Pugh T, Gowrisankar S, Shakhbatyan R, Duffy E, Bowser M, Harrison B, Lafferty K, Mahanta L, Rehm HL, Funke BH. Comprehensive diagnostic testing for stereocilin: an approach for analyzing medically important genes with high homology. *Mol Diagn.* 6, 2014, 16, pp. 639-647.
74. Society, Human Genome Variation. Sequence variant nomenclature. <https://varnomen.hgvs.org/>.
75. *Genetics Glossary: Copy Number Variations.* Institute, National Human Genome Research. Vols. <https://www.genome.gov/genetics-glossary/Copy-Number-Variation>.
76. Borg J, Georgitsi M, Aleporou-Marinou V, Kollia P, Patrinos GP. Genetic recombination as a major cause of mutagenesis in the human globin gene clusters. *Clin Biochem.* 2009, Vol. 42, 18, pp. 1839-1850.
77. Tamrakar, Dipesh. A Brief introduction on DNA sequencing. (sanger sequencing and Next generation sequencing) & generations of sequencing. *Slideshare.* 2019, Vols. <https://www.slideshare.net/DipeshTamrakar2/dna-sequencing-171111663>.
78. Lin B, Hui J, Mao H. Nanopore Technology and Its Applications in Gene Sequencing. *Biosensors (Basel).* 2021, Vol. 11, 7.
79. Karampetsou E, Morrogh D, Chitty L. Microarray Technology for the Diagnosis of Fetal Chromosomal Aberrations: Which Platform Should We Use? *J Clin Med.* 2014, Vol. 3, 2, pp. 663-678.

