Universidad de Valencia

Facultad de Ciencias Biológicas

Usher syndrome: molecular analysis of the USH2 genes and development of a next-generation sequencing platform

Gema García García

TESIS DOCTORAL

2013

Programa de doctorado en Biotecnología

Directores de tesis:

Dr. José Mª Millán Salvador

Dra. Elena Aller Mañas

El Dr. José María Millán Salvador, Titulado Superior de la Unidad de Genética y Diagnóstico Prenatal del Hospital La Fe, y la Dra. Elena Aller Mañas, Investigadora Postdoctoral del Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER) en el Instituto de Investigación Sanitaria La Fe,

CERTIFICAN: Que la memoria titulada "Usher syndrome: molecular analysis of the USH2 genes and development of a next-generation sequencing platform", ha sido realizada bajo su codirección en el Grupo de Enfermedades Neurosensoriales del Instituto de Investigación Sanitaria La Fe, por Gema García García, Licenciada en Ciencias Biológicas.

Para que conste, en cumplimiento de la legislación vigente, firman el presente certificado en Valencia a 03 de Mayo de 2013.

Fdo. José Mª Millán Salvador

Fdo. Elena Aller Mañas

Fdo. Gema García García

“A person who never made a mistake, never tried anything new”

Albert Einstein

“Life is not easy for any of us. But what of that? We must have perseverance and above all confidence in ourselves. We must believe that we are gifted for something and that this thing must be attained.”

Marie Curie

Agradecimientos

Son muchas las personas que de alguna forma han compartido conmigo estos últimos cuatro años, pero en primer lugar me gustaría agradecer a mis directores de tesis Chema y Elena. A **Chema**, por darme la oportunidad de dar mis primeros pasos en la investigación en su grupo, apoyarme y aceptar mis decisiones, por el respeto que tiene a su trabajo y pensar siempre en lo mejor para los demás. A **Elena**, quien tuvo la paciencia necesaria para formarme en mis inicios y siempre ha estado ahí cuando lo he necesitado, ha sido un placer trabajar contigo. Agradecer también al resto del Grupo de Enfermedades Neurosensoriales. No podría haber elegir un grupo mejor para realizar mi tesis y ha sido enriquecedor compartir laboratorio con todos vosotros. **Teresa**, gracias por escucharme e intentar ayudarme con mis dudas y toma de decisiones, resaltar que has sido un apoyo esencial para mí en estos últimos meses. **Rafa**, gracias por tus "no pasa nada", tu buen humor y porque siempre se puede contar contigo. **Regina**, siempre con una sonrisa y dispuesta a tenderte una mano, es un placer estar rodeada de gente como tú. **Majo**, que decir… a mi lado desde el primer día de Universidad hasta el último de la tesis, vacaciones en Jávea, cenas, charlas… en fin, gracias por estar siempre que lo he necesitado. **Loli**, gracias por el apoyo, las risas y por tus anécdotas, se te echa de menos cuando no estás en el laboratorio. **Cristina,** trabajadora incansable pero siempre pendiente del resto, seguro que conseguirás lo que te propongas. Recordar también a todas las personas que han pasado por el laboratorio, en especial a **Elena Grau**, **Ana María** y **Lucía,** gracias por vuestra simpatía y vuestros consejos. Sin olvidar a **Vicente Arnau**, porque siempre he guardado un buen recuerdo y sé que se puede contar contigo.

A mon groupe Usher français. Merci **Anne-Françoise**, pour m'avoir ouvert les portes de ton groupe, pour me permettre « d'apprendre de toi » jour après jour, pour tes conseils, tu seras toujours ma troisième directrice de thèse. Et j'attends ta phrase en espagnol.... **Christel**, je ne te remercierai jamais assez de l'opportunité que tu m'as donnée de travailler avec toi. Merci pour tes conseils, d'avoir pris soin de moi et pour tous les bons moments! **David**, mon Sauveur en informatique, merci pour ta disponibilité, toujours prêt à m'aider, ta bonne humeur et aussi pour ton honnêteté dont on a toujours besoin. **Lise**, **Val**

et **Susana**, les meilleures techniciennes du monde, merci de m'avoir aidé et toujours avec le sourire. **Thomas**, merci pour ton aide, ta patience et ton soutien tout au long de mon stage. Je garde en mémoire toutes nos conversations et nos bons moments, cela a été un plaisir de travailler avec toi. Et je ne peux pas oublier: **Jessica,** travailleuse, aimable, sympathique…enfin une fille exceptionnelle qui m'a toujours écouté et soutenu quand j'en avais besoin... autour d'une bière ou d'un bon film!". **Anne**, ta vitalité et ta passion pour ton travail sont admirables, cela a été un plaisir de partager le bureau, les apéros et les rires avec toi et surtout merci pour ton soutien continu. **Victoria**, merci pour tes conseils et d'essayer de me comprendre, tout est beaucoup plus facile avec des collègues comme toi! **Aliya**, j'ai adoré nos longues conversations à plus de 7h du soir, toujours gentille, tu m'as aidé depuis mon premier jour au labo, vraiment merci. **Marie Champeyrache,** pour apporter toujours la bonne humeur et de la lumière au labo, merci de te préoccuper de moi. **Julie**, toujours prête à écouter, aider, et aussi à faire un apéro! Je n'ai pas la place pour vous remercier tous un par un, donc je donne un grand merci à tout l'ensemble du **Laboratoire de Génétique Moléculaire**, dirigé par le professeur **Mireille Claustres**, je suis ravie d'avoir passé presque 2 ans avec vous.

A tous les amis que j'ai fait à Montpellier, qui m'ont fait sentir comme chez moi: **Elvira, Julio, Iulia, Teresita, Aroa, Lucille, Laura, Josefina…** Mention spéciale à ma soeur et à mon frère italiens, **Erica et Alessandro**, avec qui j'ai partagé de longues conversations, autant autour d'un diner qu'à 4h du matin dans le hall de la résidence! Beaucoup de rires et de bons moments, cette expérience n'aurait pas été la même sans vous.

A todos mis compañeros de Universidad y de Máster, **Sergio**, **Esther**, **Jorge, Diego, Elena, Susana**…, por compartir trabajos, apuntes, comidas, risas y tantos buenos momentos. Especialmente agradecer a mis 6 perlas. **Ester, Ana, Majo** y **Pilar**, aún recuerdo el primer día de clase en el que os conocí, los descansos en la cafetería, las risas en las clases de física...y desde ahí inseparables! **Teresa y Clara**, a las que conocí algo más tarde, pero con las que enseguida congenié, porque cada uno de los instantes vividos con vosotras dos ha sido especial y porque aun nos queda mucho por compartir! Simplemente gracias a las 6 por ser como sois y estar siempre a mi lado.

A todos mis amigos de Castellón, tanto a los que conozco de toda la vida, como a los que he ido conociendo en diferentes épocas de mi vida. En especial, a **Eva** (quien me soporta desde los tres años!!) y a **Blanca**, gracias a las dos por los miles de buenos recuerdos, por contar siempre conmigo, incluso a pesar de la distacia, y espero que siga así durante mucho tiempo.. **Pili**, inmejorable compañera de piso, de universidad y amiga. A **Amalia** y **Fanny** por todos esos buenos momentos vividos desde el instituto.

A todos mis amigos de la **Iglesuela del Cid**, porque subir al pueblo y estar con vosotros siempre me ha hecho feliz y me ha servido para desconectar. **Silvia, Patricia, María José, Esther, Mari, Rocío, Alba, Virginia y Marta,** imposible olvidar todos esos veranos, viajes, fiestas, anécdotas… gracias por vuestro apoyo, por no tener en cuenta mis ausencias e intentar entender mi trabajo "sobre las ballenas".

Agradecer a toda mi familia. A mis **abuelos** por mimarme y darme todo su cariño. A todos mis **tíos** y **primos** y las personas que sin ser familia directa, considero parte de ella, gracias por vuestro apoyo incondicional.

A vosotros**, papá y mamá**, por la fuerza y el coraje que siempre habéis tenido, y porque soy consciente del esfuerzo y sacrificio que habéis hecho para que nosotros pudiésemos elegir libremente nuestro camino. Siempre habéis aceptado mis decisiones, apoyado en todo momento y confiado en mí, incluso más que yo en mí misma. Sois y siempre seréis un ejemplo para mí, y gracias a vosotros he podido llegado hasta aquí. **Enrique** y **Ángel**, no podría haber tenido unos hermanos mejores, sé que siempre puedo contar con vosotros y no os podéis imaginar lo orgullosa que estoy de teneros a mi lado. Los dos tenéis un corazón enorme y os merecéis toda la felicidad del mundo. A los cuatro, y consciente de que pocas veces os lo digo, os quiero mucho.

Citando a la película *Into the wild*: *"la felicidad sólo tiene sentido cuando es compartida"*, ¡gracias a todos por compartirla conmigo!

Abbreviations

AAV	Adeno-Associated Virus
ANK	Ankyrin
AON	Antisense oligonucleotide
APEX	Arrayed primer extension
CalX-B	Calcium echanger β
CC	Connecting cilium
CD	Carboxy-terminal cytoplasmic domain
CEN	Central region
Array-CGH	Array comparative genomic hybridization
CIB2	Calcium and integrin binding protein 2
CNTF	Ciliary neurotrophic factor
DFNA	Nonsyndromic deafness autosomal dominant
DFNB	Nonsyndromic deafness autosomal recessive
EAR/EPTP	Epilepsy associated repeat
EC	Extracellular cadherin repeats
EGF-Lam	Laminin EGF(epidermal growth factor)-like
ERG	Electroretinography
FERM	4.1, ezrin, radixin, moesin
FN3	Fibronectin type III
GPR98	G-protein coupled receptor 98
GPS	Glutathione S-transferase
IHC	Inner hair cell
IQ	Isoleucine-glutamine
IS	Inner segment
LamG	Laminin G-like
LamNT	Laminin N-terminal
LCA	Leber congenital amaurosis
MyTH4	Myosin tail homology 4
NGS	Next generation sequencing
OHC	Outer hair cell
OS	Outer segment
PBM	PDZ binding motif
PDZ	PSD95, discs large, ZO-1
PDZD7	PDZ domain containing protein 7

PST	Proline-Serine-Threonine rich
RCPG	G protein coupled receptor
RP	Retinitis pigmentosa
RPE	Retinal pigment epithelium
RPE65	Retinal pigment epithelium 65
SAM	Sterile alpha motif
SH3	Src homology 3
SP	Signal peptide
SSCP	Single strand conformation polymorphism
TM	Trasnmembrane
TRIDs	Translational read-through-inducing drugs
USH	Usher syndrome
USH1	Usher syndrome type 1
USH2	Usher syndrome type 2
USH3	Usher syndrome type 3
VLGR1	Very large G-coupled receptor
WES	Whole exome sequencing
WGS	Whole genome sequencing

List of tables and figures

Figures

Tables

Appendix

Table of contents

Capítulo I: Introducción general

CAPÍTULO I

Introducción General

1. Historia del síndrome de Usher

El síndrome de Usher (USH) es una enfermedad hereditaria autosómica recesiva, caracterizada por la asociación de retinosis pigmentaria (RP), hipoacusia neurosensorial y, en ocasiones, alteración de la función vestibular. La combinación de sordera y RP fue descrita por primera vez por el oftalmólogo alemán Albrecht von Graefe en 1858 (GRÄFE 1858). Sin embargo, fue el oftalmólogo escocés Charles Howard Usher quien apuntó la naturaleza hereditaria de esta enfermedad y finalmente le dio nombre en 1935, 77 años después de su primera descripción (USHER 1935).

En 1922, antes de que el síndrome tuviese su nombre definitivo, la genetista Julia Bell observó que existían diferencias en cuanto al grado de sordera entre varios pacientes, indicando ya la heterogeneidad clínica de la enfermedad (BELL 1922). Ella dividió los pacientes en dos grupos, uno con sordera profunda y otro con pérdida auditiva de moderada a severa. Estos hechos sirvieron de base para la posterior clasificación del síndrome de Usher descrita por Davenport y Omenn en 1977 (DAVENPORT S 1977). En 1994, Smith y colegas clasificaron finalmente el síndrome en 3 tipos clínicos diferentes en función de la progresión, gravedad y edad de aparición de los diferentes síntomas (SMITH *et al.* 1994).

Sin embargo, no fue hasta 1995, casi 140 años después de la primera descripción del síndrome de Usher, cuando se identificó el primer gen responsable de la enfermedad (*MYO7A*) (WEIL *et al.* 1995). A pesar de los numerosos avances que se han producido en el estudio de este trastorno, aún queda mucho que esclarecer sobre la genética y fisiopatología del síndrome de Usher.

2. La visión

La visión es uno de los sentidos más importantes, nos permite apreciar el mundo que nos rodea y también comunicarnos. Los ojos son, junto al cerebro, los órganos fundamentales para la visión. Está constituido por el globo ocular y otros órganos anexos. El globo ocular está formado por tres capas concéntricas (Figura 1):

- La capa más externa denominada túnica fibrosa está formada por la esclerótica en la parte posterior, formada por tejido conectivo y que tiene una función protectora y mantiene junto con el humor vítreo la forma esférica del ojo, y la córnea en la parte anterior, tejido transparente y avascular con función principalmente óptica. El limbo esclerocorneal separa la esclera de la córnea.
- La capa intermedia se denomina túnica vascular y se caracteriza por estar muy vascularizada. Está dividida en dos partes: la anterior o úvea, formada por el iris y el cuerpo ciliar y la posterior o coroides.
- La capa interna, ó túnica sensorial del ojo es la retina.

La luz entra al ojo a través de la córnea y es refractada por el cristalino. Posteriormente atraviesa el humor vítreo y es proyectada a la retina, donde los fotorreceptores convierten la luz en señales eléctricas que son transmitidas vía nervio óptico a la región de la corteza cerebral encargada de la visión.

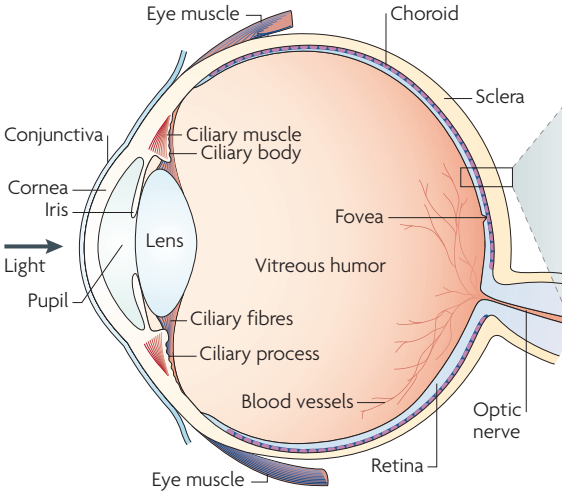


Figura 1. Anatomía del ojo humano. Imagen adaptada de (WRIGHT *et al.* 2010).

2.1 Desarrollo y estructura de la retina

La retina es una parte del sistema nervioso central que deriva del tubo neural. La retina se forma durante el desarrollo embrionario a partir de dos vesículas ópticas que nacen directamente del tubo neural. Estas dos vesículas se van aproximando y su interacción con el ectodermo superficial va haciendo que éste se vaya diferenciando (para formar la córnea y el cristalino) y que las vesículas se vayan transformando en cúpulas ópticas. Inicialmente cada una de las dos capas de la cúpula óptica está formada por una sola capa de células. A lo largo del desarrollo, la capa más externa se diferencia para formar el epitelio pigmentario, mientras que las células de la capa más interna se dividen repetidamente para dar lugar a la retina (Figura 2).

El desarrollo de la retina comienza en estadios muy tempranos y se van formando nuevas capas, que surgen como consecuencia de las sucesivas divisiones y migraciones celulares. De esta forma la retina se desarrolla desde dentro hacia fuera y tiene un aspecto estratificado en el que van apareciendo los diferentes tipos de células (Figura 2).

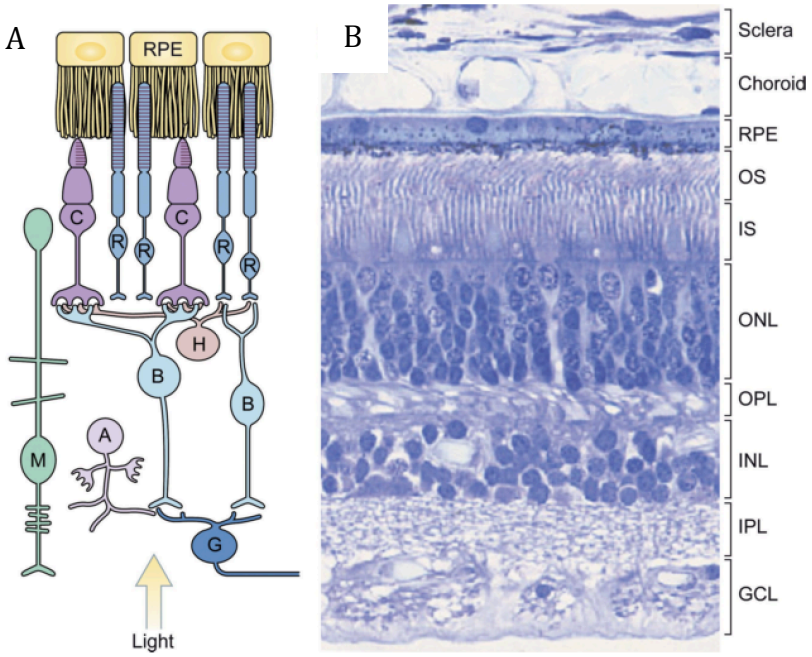


Figura 2. Estructura de la retina humana. A) Representación de los distintos tipos celulares. R, bastones; C, conos; B, célula bipolar; H, célula horizontal; A, célula amacrina; G, células ganglionares; M, célula de Müller. B) Sección de la retina humana. La retina está formada por una serie de capas. El núcleo de los fotorreceptores constituye la capa nuclear externa (ONL). El núcleo de las células bipolares, amacrinas, horizontales y las células gliales de Müller se encuentran en la capa nuclear interna (INL). Las células ganglionares forman la capa celular ganglionar (GCL). La

capa plexiforme externa (OPL) contiene las terminales sinápticas de los fotorreceptores, células horizontales y bipolares. En la capa plexiforme interna (IPL) se encuentran las terminales de las células bipolares, amacrinas y ganglionares. RPE (epitelio pigmentario retiniano), OS (segmento externo), IS (segmento interno). Imagen adaptada de (SUNG and CHUANG 2010).

Los fotorreceptores son las células sensibles a la luz y más abundantes de la retina. Hay dos tipos de fotorreceptores, conos y bastones. Los bastones son más numerosos, alrededor de 120 millones. Sin embargo, no son sensibles al color. Son los responsables de nuestra visión nocturna (visión escotópica) y en blanco y negro y, su fotopigmento es la rodopsina. Los conos son los fotorreceptores responsables de la visión diurna (visón fotópica) y de la percepción de los colores. Hay unos 6-7 millones de conos que pueden ser divididos en "rojos" (64%), verdes (34%) y azules (2%), dependiendo de la fotopsina (pigmento fotosensible) que contengan. Los conos están implicados en la agudeza visual.

Respecto a la densidad de conos y bastones en la retina, ésta muestra una gran cantidad de conos (principalmente verdes y rojos) en la fóvea, zona de la retina de mayor agudeza visual. Los conos "azules" tienen una mayor sensibilidad y se encuentran principalmente en la mácula, fuera de la fóvea. Los bastones están ausentes en la fóvea, a medida que nos vamos alejando de ella la densidad de bastones va aumentando y extendiéndose por la retina, siendo así los responsables de la visión periférica.

Ambos tipos de fotorreceptores tienen la misma estructura, constan de una región externa (de forma cónica en los conos y alargada en los bastones) y otra interna (Figura 3). El segmento externo está compuesto de discos membranosos que contienen el fotopigmento y se encuentra inmerso en el epitelio pigmentario de la retina, mientras que el segmento interno contiene los orgánulos y el núcleo de la célula y presenta terminales sinápticos que contactan con las células bipolares y horizontales.

Ambos segmentos están unidos por el cilio conector, un cilio modificado que tiene un papel fundamental en la organización y función de los fotorreceptores (Figura 3). Las membranas del segmento externo se renuevan continuamente, discos nuevos se forman en la base mientras que en la zona distal se fagocitan por las células del RPE. Este reemplazamiento requiere la síntesis de nuevas proteínas en el compartimento metabólicamente activo, el segmento interno, seguido de su transporte intracelular a través del cilio conector hacia la base del segmento externo. Además, hay proteínas presentes en el cilio conector que participan en la morfogénesis y mantenimiento de los discos del segmento externo.

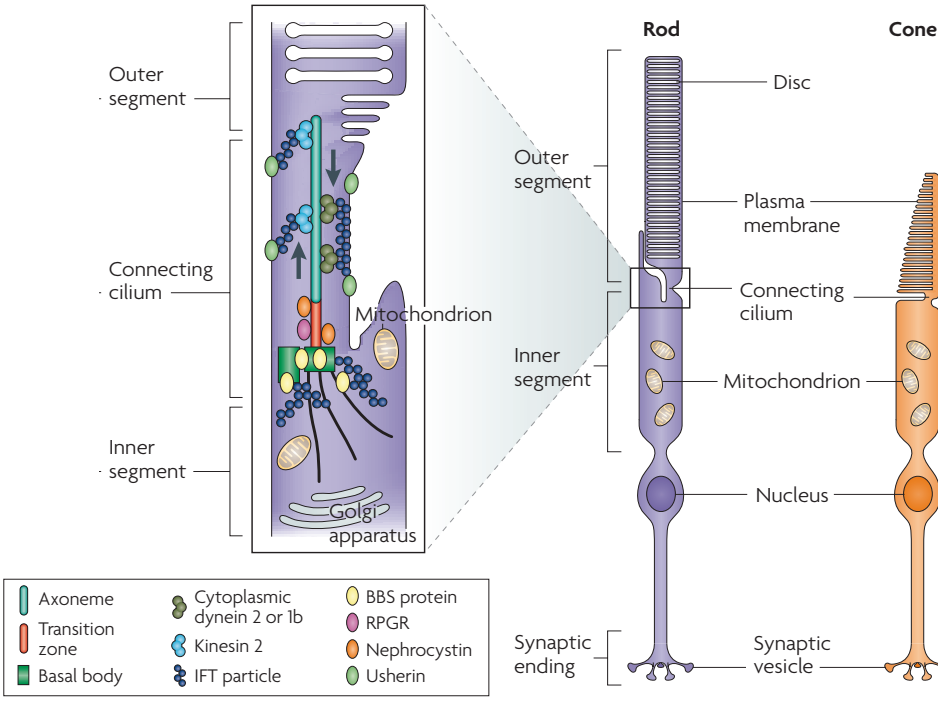


Figura 3. Estructura de los fotorreceptores. En la parte derecha de la imagen están representados los dos tipos de fotorreceptores: un bastón y un cono. A la izquierda de la figura se muestra en detalle la estructura del cilio conector. El cilio conector está formado por un eje principal, el axonema, compuesto de 9 microtúbulos dobles dispuestos en círculo. El cilio está anclado al cuerpo basal (basal body), que organiza el ensamblaje de los microtúbulos. El segmento interno contiene la mayor parte de la maquinaria encargada del metabolismo y del tráfico, mientras que el segmento externo contiene el aparato de fototransducción. Hay un flujo elevado de proteínas desde el segmento interno al externo, a través de mecanismos de transporte llevados a cabo por kinesinas motoras (hacia la región superior del axonema) o por dineínas (hacía el cuerpo basal). BBS, síndrome de Bardet–Biedl; RPGR, retinitis pigmentosa GTPase regulator. Imagen adaptada de (WRIGHT *et al.* 2010).

La absorción de la luz por el fotopigmento del segmento externo de los fotorreceptores inicia una cascada de reacciones bioquímicas, denominada fototransducción, que cambia el potencial de membrana y la cantidad de neurotransmisor liberado en la sinapsis del receptor con las células bipolares, las cuales transmitirán el impulso nervioso a las células ganglionares, cuyos axones formaran el nervio óptico llevando la información al resto del sistema nervioso central.

2.2 Retinosis Pigmentaria

La retinosis pigmentaria pertenece al grupo de retinopatías pigmentarias, nombre que engloba todas las distrofias de la retina con pérdida de fotorreceptores y depósitos de pigmento.

La prevalencia de la RP no sindrómica está estimada en 1 caso por cada 3.000-4.000 personas. Hay casi 1,5 millones de afectos en todo el mundo, siendo la forma más frecuente de distrofia retiniana hereditaria.

Su patrón de herencia es variable, pudiendo ser autosómica recesiva, autosómica dominante, ligada al cromosoma X, de herencia mitocondrial o de herencia digénica. La RP es normalmente no sindrómica, pero también hay también muchas formas sindrómicas, siendo la más frecuente el síndrome de Usher.

Como la RP típica afecta principalmente a los bastones, el síntoma más frecuente al inicio de la enfermedad es la ceguera nocturna (nictalopía) y pérdida de la visión periférica, ambas debidas a la muerte de los bastones en las dos primeras décadas de la vida, a lo que le sigue un progresivo empeoramiento de la visión a causa de la pérdida de los fotorreceptores conos.

A continuación se describen los principales signos clínicos, que sirven como criterio para al diagnóstico:

- **Nictalopía**: el primer síntoma de la RP es generalmente la ceguera nocturna. Los pacientes tienen problemas en la adaptación a la oscuridad y presentan dificultades en ambientes de poca luminosidad.

- **Disminución del campo visual**: generalmente, aparece después de la ceguera nocturna y es otro de los síntomas característicos de esta enfermedad. Al verse afectados en primer lugar los bastones, el campo visual periférico se ve disminuido (conocido como "visión en túnel"). Ésta pérdida es normalmente simétrica y avanza más rápido en el exterior y de una forma más lenta hacia el campo central. La agudeza visual está preservada hasta etapas más avanzadas.

- **Fotofobia**: una sensibilidad anormal a la luz es un signo frecuente, sobre todo en estados avanzados de la enfermedad.

- **Fotopsias**: los pacientes refieren la visión de puntos luminosos, chispas o como rayos luminosos en el campo intermedio entre la región central y periférica.

- **Problemas en la visión en color**: la visión en color permanece intacta hasta que los conos empiezan a estar afectados, sobre todo dentro del rango de color azul-amarillo.

- **Alteración del fondo de ojo**: el aspecto del fondo de ojo varía dependiendo del grado de degeneración de la retina (Figura 4). Suelen aparecer depósitos pigmentarios, inicialmente en la periferia de la retina, atenuación de los vasos sanguíneos, palidez papilar y varios grados de atrofia retinal.

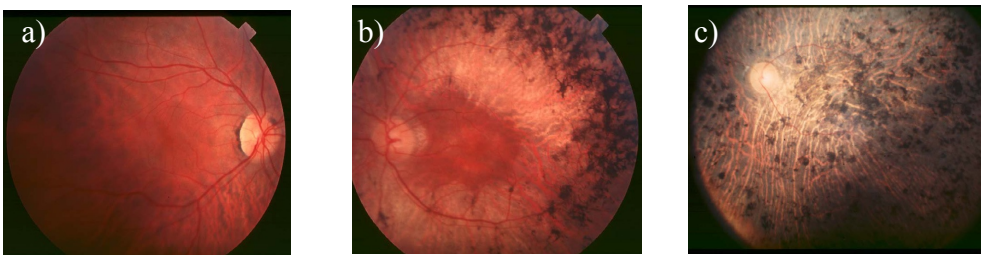


Figura 4. Imágenes de fondo de ojo de pacientes con RP en diferentes etapas de la enfermedad. a) Estado temprano de la RP. b) Estadio medio: presencia de depósitos pigmentarios, atenuación de los vasos. c) Fase avanzada de la RP: los depósitos están presentes en toda la retina, vasos sanguíneos muy finos y palidez papilar. (Imagen adaptada de (HAMEL 2006)).

- **Alteración de la función visual**: la función visual se determina mediante el electroretinograma (ERG). El ERG es un examen importante para el diagnóstico de la RP, pone en evidencia el estado funcional de los fotorreceptores y permite diferenciar la respuesta de los bastones de la de los conos y conocer así el grado de progresión de la enfermedad.

3. La audición

La audición nos permite interactuar con el entorno y es uno de los procesos fisiológicos más importantes que permiten el aprendizaje.

El oído humano es el encargado de convertir las ondas sonoras en señales eléctricas que se transmiten por el nervio auditivo hasta el cerebro, donde el sonido es interpretado. También es esencial para el equilibrio u orientación espacial. Se divide en tres partes: el oído externo, el oído medio y el oído interno (Figura 5a):

- **Oído externo**: constituido por el pabellón auditivo y el conducto auditivo externo (este conducto termina en el tímpano). El pabellón tiene como función localizar espacialmente el sonido, mientras que el conducto auditivo protege las estructuras más internas y conduce la vibración sonora hasta la membrana del tímpano.

- **Oído medio** (o caja timpánica): cavidad entre el tímpano y el oído interno encargada de la transmisión mecánica de la señal sonora. Esta cavidad timpánica consta de una cadena de tres huesecillos (el martillo, el yunque y el estribo) unidos entre sí de forma articulada, que actúan amplificando las vibraciones desde el tímpano hasta el oído interno.

- **Oído interno**: en el oído interno existen dos órganos, el auditivo o coclear (ubicado en la cóclea) y el órgano del equilibrio o vestibular.

3.1 Desarrollo y estructura del oído interno

El oído interno se desarrolla a partir de la placa ótica, fina región del ectodermo craneal que se invagina y cierra para llegar a formar una vesícula ótica, dónde ocurre un temprano desarrollo y diferenciación del epitelio y estructuras neuronales. La morfogénesis de las células mecanosensoriales ocurre en etapas más tardías del desarrollo del oído interno.

La detección del sonido empieza cuando las ondas alcanzan el oído externo y son transmitidas por el conducto auditivo externo hasta llegar al tímpano, donde provoca vibraciones. Estas vibraciones generan movimientos oscilantes de los huesecillos del oído medio y transmiten la señal hasta el oído interno (cóclea) donde, las vibraciones mecánicas son transformadas en impulsos nerviosos, que serán conducidos hasta la corteza cerebral para ser interpretados.

La cóclea es un órgano en forma de espiral y está dividida en tres regiones: la escala vestibular, la escala timpánica y la escala media. La escala media, está llena de endolinfa, caracterizada por tener una baja concentración en Na^+ y alta concentración de K^+, y contiene el órgano de Corti. El órgano de Corti es el epitelio sensorial del aparato auditivo y donde se produce la transformación del estímulo mecánico en impulsos

eléctricos. Este órgano descansa sobre la membrana basilar y está cubierto por la membrana tectorial (Figura 5b).

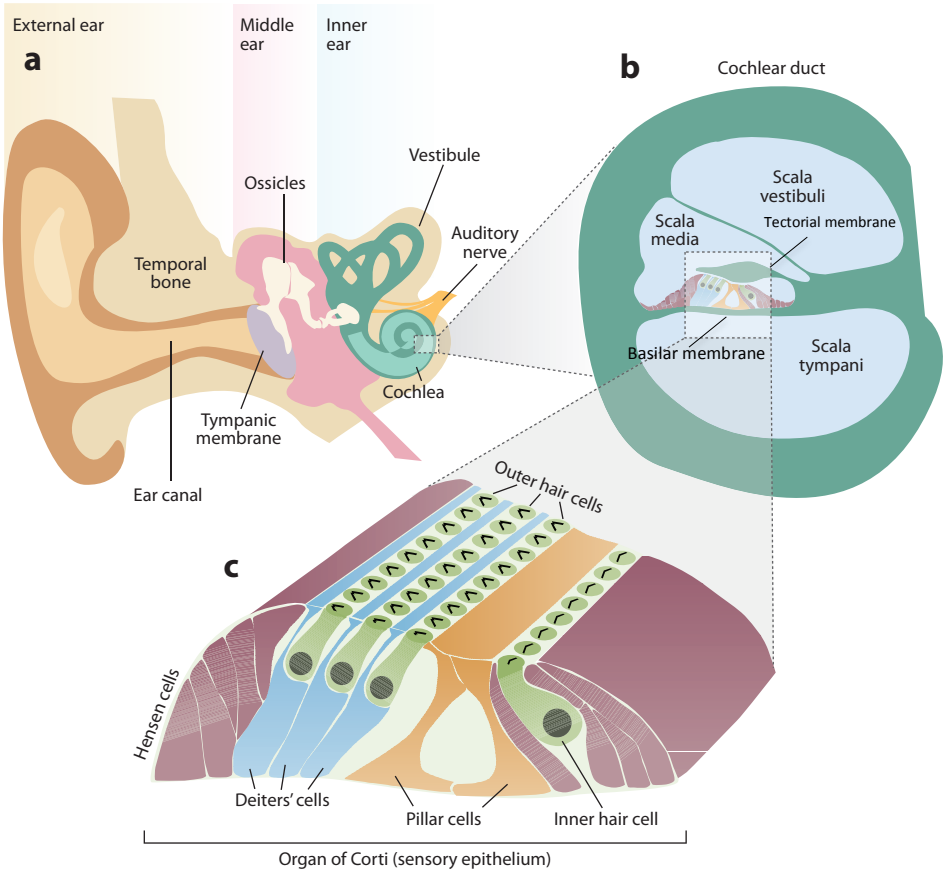


Figura 5. Esquema del oído humano. A) El oído está dividido en tres partes, el oído externo, medio e interno. B) Sección del conducto coclear que muestra los diferentes compartimentos llenos de fluido del oído interno. C) El órgano de Corti reside en la escala media, con las células sensoriales (células ciliadas) rodeadas de las células de sostén (que incluyen las células Hensen, Deiters y Pillar). Imagen adaptada de (DROR and AVRAHAM 2009).

En el órgano de Corti se distinguen dos tipos de células sensoriales especializadas, llamadas células ciliadas, que presentan una disposición altamente organizada: una fila interior compuesta por las "células ciliadas internas" (IHCs) y tres filas exteriores formadas por las "células ciliadas externas" (OHCs) (Figura 5c y 6).

Ambas convierten estímulos mecánicos en variaciones del potencial intracelular. Sin embargo, solo las IHCs parecen transmitir información sobre el ambiente acústico al

cerebro, mientras que las OHCs reciben abundantes inervaciones eferentes, pero pocas aferentes, pareciendo así que su tarea principal es potenciar el estímulo por un *feedback* electromecánico. Además, las OHCs muestran una movilidad única dada por voltaje e independiente de ATP, que proporcionaría un *feedback* positivo para amplificar las vibraciones inducidas por el sonido del órgano de Corti (FROLENKOV *et al*. 2004; FETTIPLACE and HACKNEY 2006).

Tanto las IHCs como las OHCs contienen en su parte apical microvellosidades especializadas, ricas en actina, llamadas estereocilios (Figura 6a). Para su correcto funcionamiento estos estereocilios se encuentran ordenados en filas escalonadas y conectados lateralmente entre sí por una serie de uniones que estabilizan su estructura y les permiten actuar como unidad funcional. En la figura 6b se pueden observar los diferentes tipos de uniones: uniones apicales (*tip links*), uniones superiores horizontales (*horizontal top links*), uniones laterales intermedias (*side links*) y uniones basales (*ankle links*).

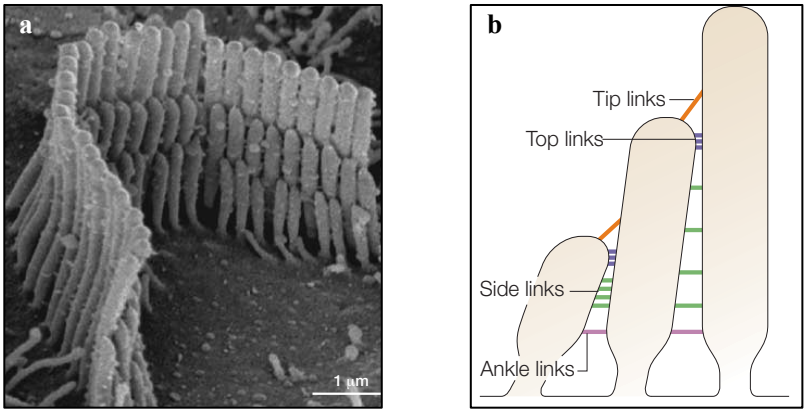


Figura 6. Representación de los estereocilios presentes en las células ciliadas. A) Imagen a microscopía electrónica de la especializada organización en escala de los esterocilios. B) Los cuatro tipos de uniones conocidas entre esterocilios adyacentes. Imágenes adaptadas de (FROLENKOV *et al.* 2004).

Las ondas generadas por el sonido hacen vibrar la membrana basilar, y excitan las células ciliadas por la desviación de sus estereocilios al rozar con la membrana tectorial, Ésto provoca la apertura de los canales iónicos de las células ciliadas, permitiendo el flujo de potasio de la endolinfa al interior de las células, iniciándose la despolarización.

Estas células mecanosensoriales pueden degenerar por el envejecimiento, exposición al ruido, drogas ototóxicas (por ejemplo antibióticos aminoglucósidos) o por la herencia de mutaciones en genes que son esenciales para la audición.

3.2 Hipoacusia

La hipoacusia es un defecto sensorial (disminución del nivel de audición) que en su forma más severa afecta a más de 28 y 22.5 millones de americanos y europeos, respectivamente. Estas cifras son todavía mucho más elevadas en regiones del Este o India, donde la consanguinidad contribuye al mayor número de casos de sordera hereditaria (DROR and AVRAHAM 2009). La incidencia de hipoacusias se estima en 1 de cada 1000 recién nacidos, con un 60% de los casos debidos a causas genéticas. Dentro de los casos genéticos, hay un 30% de hipoacusia sindrómica, con casi 400 formas de sordera asociada con otras anomalías clínicas (por ejemplo el síndrome de Usher), y un 70% donde la pérdida auditiva es un síntoma aislado (BROWNSTEIN *et al.* 2011). Para la hipoacusia no sindrómica se han descrito todos los patrones mendelianos de herencia: dominante, recesiva, ligada al X y mitocondrial.

Hay varios criterios que son importantes para el diagnóstico y clasificación de la pérdida auditiva:

1. Localización del daño:

    - Hipoacusia conductiva: se ve afectada la transmisión de la vibración sonora y en la mayoría de los casos es producida por lesiones en el oído externo o medio.
    - Hipoacusia neurosensorial: el defecto tiene lugar en el nervio auditivo o en la cóclea.
    - Hipoacusia mixta: en este tipo los problemas de transmisión y neurosensoriales ocurren simultáneamente.

2. Etiología:

    - Hipoacusia adquirida: debida a infecciones o toma de fármacos.
    - Hipoacusia genética: en este caso la pérdida auditiva tiene una base genética.

3. Edad de aparición:

    - Prelingual: la pérdida auditiva aparece con anterioridad a la adquisición del lenguaje, y por tanto el niño es incapaz de aprender a hablar.
    - Postlingual: la pérdida auditiva aparece con posterioridad a la adquisición del lenguaje.

4. Grado de la pérdida auditiva:

    La hipoacusia también se puede clasificar según el grado de la pérdida. La tabla 1 muestra la clasificación según la ASHA (*América Speech and Hearing Association*). Los números representan el rango de pérdida en decibelios (dB).

Tabla 1. Diferentes grados de hipoacusia según la ASHA.

Grado Hipoacusia	Rango Pérdida dB
Normal	–10 a 15
Ligera	16 a 25
Leve	26 a 40
Moderada	41 a 55
Moderada-Severa	56 a 70
Severa	71 a 90
Profunda	>91

Es muy importante diagnosticar y tratar la hipoacusia lo más temprano posible, por el impacto que tiene en el aprendizaje y desarrollo, y en general en la calidad de vida de las personas.

Como hemos comentado anteriormente, la mitad de los casos de pérdida auditiva tienen un origen genético. Sin embargo, debido a la gran heterogeneidad genética (69 genes descritos hasta la fecha) el diagnóstico genético queda limitado al *screening* de las mutaciones más frecuentes en cada población.

4. El equilibrio

El equilibrio es la capacidad de nuestro organismo para mantener la postura corporal respecto a la gravedad, pudiendo percibir de esta forma el entorno y conocer la posición del cuerpo en el espacio.

El aparato vestibular, situado en el oído interno, es el órgano del equilibrio. Gracias a la información aportada por el sistema vestibular, más la visual, propioceptiva (propiocepción: sentido que informa sobre la posición de los músculos) y el cerebelo, logramos nuestra coordinación postural y control motor. Está compuesto por 5 estructuras (Figura 7):

- **Tres canales semicirculares** responsables de las aceleraciones angulares (rotaciones de la cabeza). Hay tres conductos semicirculares, cada uno orientado a 90º respecto al otro y al final de cada conducto hay una dilatación llamada "ampolla". En la base de la "ampolla" hay una estructura conocida como "cresta", donde se encuentran un grupo de células ciliadas inervadas. Estas células ciliadas tienen en su zona apical una serie de estereocilios correctamente orientados, que se encuentran embebidos en una cúpula gelatinosa. Cuando la cabeza se mueve, el fluido en uno o más canales se desplaza empujando y curvando contra la cúpula los estereocilios de las células ciliadas. El sistema de transducción es similar al del sistema auditivo, el estímulo mecánico originado por el desplazamiento de los cilios causa la apertura de los canales K^+ y la generación de impulsos nerviosos, que llevan la información hasta el cerebro.

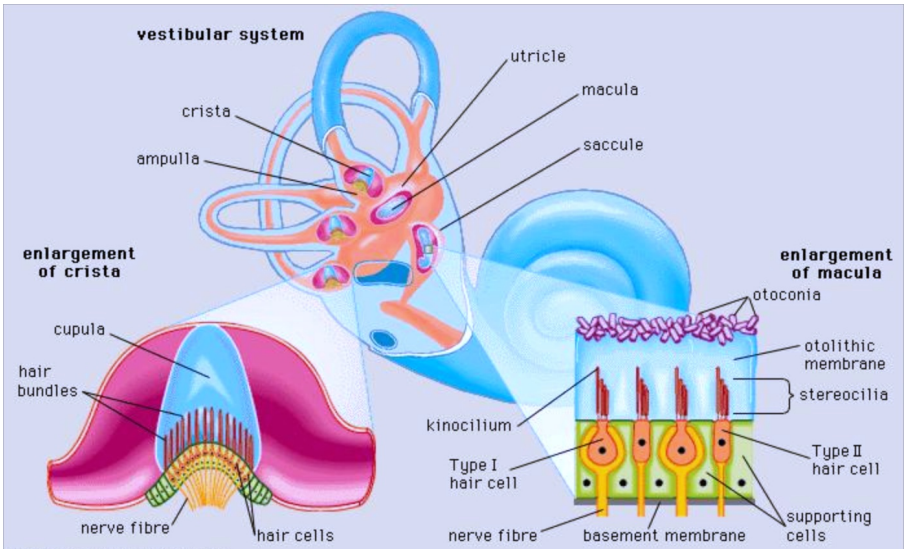


Figura 7. Estructura del sistema vestibular. Imagen adaptada de www.Britannica.com

- **Dos órganos** responsables de las aceleraciones lineales, "sáculo" y "utrículo". Estos órganos están situados en las paredes del oído interno, entre los conductos semicirculares y la cóclea, y están situados a 90º uno del otro. En estas estructuras se encuentra un órgano receptor conocido como "mácula". La mácula está integrada por células sensoriales ciliadas, cubiertas por una membrana horizontal. En la superficie de esta membrana gelatinosa se encuentran unos pequeños cristales de carbonato cálcico, llamados "otolitos" (otoconia). Con cualquier posición de la cabeza, la gravedad doblará los cilios de las células ciliadas, debido al peso de los otoconias. Este curvamiento de los cilios, es lo que desencadena la cascada de mecanotransducción y la información es enviada y procesada.

5. Tipos clínicos y evaluación del síndrome de Usher

Hay más de 50 síndromes que afectan al sistema visual y auditivo simultáneamente. Sin embargo, el síndrome de Usher está considerado como la causa más frecuente de sordo-ceguera de causa genética en humanos, siendo responsable de más del 50% de los individuos sordo-ciegos (SAIHAN *et al.* 2009). También está implicado en un 5% de todas las sorderas congénitas y un 18% de RP no sindrómica. La prevalencia de esta enfermedad varía según el estudio y está estimada en un rango de 3.2 a 6.2 por cada 100.000 (MILLAN *et al.* 2011). En España se estima en un 4.2/100.000 nacidos vivos (ESPINOS and MILLAN 1998).

Históricamente el síndrome de Usher se ha divido en tres tipos clínicos, todos manifestando hipoacusia neurosensorial asociada a una pérdida visual, debida a la RP. Además en algunos casos aparecen problemas vestibulares (Tabla 2). Los tres tipos son principalmente clasificados según la gravedad y progresión de la hipoacusia y la presencia o ausencia de alteración vestibular (DAVENPORT S 1977; SMITH *et al.* 1994).

Tabla 2. Clasificación clínica del síndrome de Usher.

	Síndrome Usher tipo 1	**Síndrome Usher tipo 2**	**Síndrome Usher tipo 3**
Pérdida Auditiva	Severa a profunda	Moderada a severa	Variable
	Congénita	Congénita	Postlingual
	Estable	Estable	Progresiva
Inicio RP	Antes de la pubertad	Después de la pubertad	Variable
Función Vestibular	Arreflexia vestibular	Normal	Variable

Sin embargo, hay algunos pacientes en los que las manifestaciones clínicas no encajan exactamente con las características de los tres grupos definidos, clasificándose como síndrome de Usher atípico.

5.1 Síndrome de Usher tipo 1

El síndrome de Usher tipo 1 (USH1) es la forma más grave de la enfermedad y está caracterizado por hipoacusia neurosensorial congénita y estable, de carácter severo a profundo, con una audición residual solo en frecuencias bajas. Típicamente, estos pacientes

pierden los restos de audición a lo largo de su primer año de vida, de forma que no serán capaces de desarrollar el habla, a menos que sean intervenidos con un implante coclear.

Además, el USH1 se caracteriza por una hiporreflexia vestibular. La afectación del aparato vestibular se refleja ya en el primer año de vida, presentando éstos un retraso en el desarrollo motor. La mayoría de ellos no pueden sentarse sin apoyo hasta los 6 meses de edad y no son capaces de andar independientemente hasta después de los 18 meses (SMITH *et al.* 1994).

Como ya hemos comentado, la pérdida visual en pacientes con síndrome de Usher está causada por RP, siendo el primer signo la falta de adaptación a la oscuridad, seguida de una distrofia retiniana que afecta inicialmente a los bastones, resultando en una "visión en túnel". En los pacientes USH1 los primeros signos de ceguera nocturna aparecen en la primera década de vida y la RP suele ser diagnosticada a mediados de la segunda. Sin embargo, existen pacientes con un USH atípico, que presentan una RP lentamente progresiva y con un diagnóstico más tardío (TSILOU *et al.* 2002; PENNINGS *et al.* 2004). Anomalías en la retina pueden ser ya detectadas alrededor de los 2-3 años mediante una ERG, incluso antes de que aparezcan los primeros signos de la enfermedad. Más tarde se desarrollarán anomalías en el fondo de ojo, con depósitos de pigmentos y estrechamiento de los vasos sanguíneos. Un estudio de Fishman *et al*., mostró que las lesiones en la fóvea son más prevalentes en USH1 comparado con USH2 (FISHMAN *et al.* 1995).

5.2 Síndrome de Usher tipo 2

El síndrome de Usher tipo 2 (USH2) es el tipo clínico más frecuente y es menos grave que el USH1. El grado de hipoacusia incrementa de moderada en frecuencias bajas, a severa en frecuencias más elevadas, tendiendo a permanecer estable. De esta forma los audiogramas tienen un perfil característico en pendiente descendiente, conocido como "*down sloping*"(ABADIE *et al.* 2012). Aunque la pérdida auditiva es congénita, ésta puede ser detectada en una etapa más tardía, cuando los niños empiezan a presentar problemas en el desarrollo del lenguaje. No obstante, dado el grado de hipoacusia que padecen, son capaces de desarrollar el habla y poder comunicarse. Además, el grado de audición puede mejorar notablemente mediante el uso de audiófonos.

La función vestibular es normal, por lo que los niños no sufren alteraciones en su desarrollo motor.

Los síntomas de la RP se manifiestan habitualmente de forma más tardía en los USH2, generalmente en la segunda década de vida, es decir unos 6-8 años más tarde en comparación a los USH1. Sin embargo, la edad de inicio de la RP puede variar enormemente dentro de los USH1 y USH2, pudiendo haber solapamientos entre ambos tipos (HOPE *et al.* 1997; IANNACCONE *et al.* 2004; PENNINGS *et al.* 2004). Se ha sugerido que un diagnóstico más temprano de la RP en los casos USH1, comparado a los USH2, es debido que éstos padecen una sordera profunda combinada con una alteración de la función vestibular, por lo que los pacientes USH1 dependen más de su visión para su orientación.

El rango y grado de la pérdida visual es muy variable inter e intrafamiliarmente, pero generalmente tiende a progresar de forma más lenta que los USH1.

5.3 Síndrome de Usher tipo 3

Los pacientes con síndrome de Usher tipo 3 (USH3), sufren una hipoacusia neurosensorial postlingual y progresiva no lineal, pudiendo aparecer desde la primera a la tercera década de vida. En etapas iniciales el grado de pérdida auditiva es similar al USH2, con un mayor impacto en las frecuencias elevadas. La progresión es muy variable, pero en la mayor parte de los casos se llega a una sordera profunda. Los niveles de audición permanecen generalmente normales durante la etapa de adquisición del lenguaje, de forma que estos niños no tienen problemas para desarrollar el habla.

La presencia de alteración en la función vestibular es variable, encontrándose en un 50% de los pacientes USH3. Raramente se observa una ausencia completa de la función vestibular, y no se sabe con certeza si la alteración progresa o no (PAKARINEN *et al.* 1995; HOPE *et al.* 1997)

El inicio de los síntomas de la RP es variable, aunque suele aparecer entre los 10-15 años. La progresión visual no difiere de los USH1 y USH2.

6. Prevalencia del síndrome de Usher

Las estimaciones de la prevalencia del síndrome de Usher están basadas en estudios de unos pocos países o regiones. En general, se acepta una estimación entorno a un 3.2-6.2 de cada 100.000 (MILLAN *et al.* 2011), con una distribución de los diferentes tipos (USH1, USH2 y USH3) variable dependiendo de la población de estudio. En la siguiente tabla (Tabla 3) se observan los datos de prevalencia obtenidos en diferentes poblaciones.

Tabla 3. Datos de prevalencia del síndrome de Usher en diferentes poblaciones.

País	Prevalencia	Trabajo
Alemania	6.2/100.000	(SPANDAU and ROHRSCHNEIDER 2002)
Colombia	3.2/100.000	(TAMAYO *et al.* 1991)
Dinamarca	5/100.000	(ROSENBERG *et al.* 1997)
España	4.2/100.000	(ESPINOS *et al.* 1998)
Finlandia	3.5/100.000	(VASTINSALO *et al.* 2012)
Inglaterra	6.2/100.000	(HOPE *et al.* 1997)
Noruega	3.6/100.000	(GRONDAHL and MJOEN 1986)
USA	4.4/100.000	(BOUGHMAN *et al.* 1983)

En España 1/3 de los pacientes con síndrome de Usher corresponden al tipo 1 y 2/3 son USH2, siendo muy bajo el número de casos con USH3. En el estudio realizado por Rosenberg *et al*. sobre población danesa se observó la siguiente distribución: 1.5/100,000 para USH1, 2.2/100,000 para USH2 y 0.1/100,000 para USH3, correspondiendo a un ratio 2:3 entre USH1 y USH2 (ROSENBERG *et al.* 1997). Sin embargo, esta distribución observada en la mayoría de poblaciones, puede diferir. En la población de Birmingham (UK) un 33% de los pacientes afectados de Usher fueron clasificados como USH1, 47% como USH2 y el porcentaje de USH3 llegaba hasta el 20%. Lo mismo ocurre en Finlandia y en población judía ashkenazi, donde la distribución es aún más diferente, ya que USH3, el cual generalmente es el menos predominante en el resto de poblaciones, es el más prevalente, llegando a representar un 40% de todos los casos, debido al efecto de una mutación fundadora en el gen implicado en cada una de estas poblaciones (JOENSUU *et al.* 2001; NESS *et al.* 2003).

7. Genes y proteínas implicados en el síndrome de Usher

Acabamos de ver que el síndrome de Usher es heterogéneo clínicamente, pero también lo es a nivel genético, ya que cada tipo clínico puede a su vez dividirse en subtipos genéticos. Hasta el momento se han descrito 9 loci para el USH1 (habiéndose identificado 6 de los genes responsables), 3 loci para el USH2 (con los tres genes localizados) y 2 loci para el USH3, aunque solo se conoce uno de los genes implicados (Tabla 4).

Tabla 4. Loci, genes y proteínas identificados en el síndrome de Usher.

Tipo Clínico	Locus	Localización	Gen	Proteína
USH1	USH1B	11q13.5	*MYO7A*	miosina VIIa
	USH1C	11p15.1	*USH1C*	harmonina
	USH1D	10q22.1	*CDH23*	cadherina 23
	USH1E	21q21	*desconocido*	
	USH1F	10q21.1	*PCDH15*	protocadherina 15
	USH1G	17q25.1	*USH1G*	SANS
	USH1H	15q22-23	*desconocido*	
	USH1J	15q24	*CIB2*	calcium and integrin binding protein 2
	USH1K	10p11.21-q21.1	*desconocido*	
USH2	USH2A	1q41	*USH2A*	usherina
	USH2C	5q14.3	*GPR98*	GPR98
	USH2D	9q32	*DFNB31*	whirlina
USH3	USH3A	3q25.1	*USH3A*	clarina 1
	USH3B	20q	*desconocido*	

Inicialmente, se describieron dos loci adicionales, USH1A y USH2B. Sin embargo, en los últimos años se ha descartado su implicación en el síndrome de Usher (GERBER *et al.* 2006; KREMER *et al.* 2006).

Además, en 2010, Ebermann *et al*. describieron la implicación del gen *PDZD7* en el síndrome de Usher, bien en un modelo de digenismo junto al gen *GPR98* o como modificador, agravando el fenotipo retiniano en pacientes con mutaciones en *USH2A* (EBERMANN *et al.* 2010).

7.1 Genes y proteínas implicadas en USH1

Hasta la fecha, se han descrito 9 loci para el síndrome de Usher tipo 1, aislándose los genes responsables de la enfermedad en 6 de estas localizaciones. La mayoría de estos genes también están implicados en sorderas no asociadas a RP, tanto dominantes (DFNA) como recesivas (DFNB).

7.1.1 USH1B / *MYO7A*

Fue en 1992 cuando Kimberling *et al*. describieron el locus USH1B en el cromosoma 11q13.5, mostrando que varias familias presentaban ligamiento a esta región (KIMBERLING *et al*. 1992). Posteriormente, en 1995, el gen responsable de este subtipo, *MYO7A*, fue identificado (WEIL *et al.* 1995). Este gen también se ha asociado a hipoacusia no sindrómica tanto de carácter autosómico dominante (DFNA11) (LIU *et al.* 1997a), como autosómico recesivo (DFNB2) (LIU *et al.* 1997b; RIAZUDDIN *et al.* 2008).

El gen *MYO7A* está formado por 49 exones, 48 de ellos codificantes, abarcando 87Kb de ADN genómico. Aunque se ha descrito la existencia de *splicing* alternativo (SAHLY *et al.* 1997), hasta la fecha, solo se ha identificado una isoforma proteica que consiste en 2215 aminoácidos (254 kDa).

Esta proteína tiene una estructura característica formada por un dominio motor altamente conservado y situado en el extremo N-terminal, seguido de una región denominada cuello y finalmente la cola (Figura 8a) (KELLEY *et al.* 1997). Las miosinas son proteínas motoras capaces de desplazarse a través de los filamentos de actina. La energía requerida para este movimiento se genera por la hidrólisis de ATP. El dominio de la cabeza contiene sitios de unión a la actina y al ATP. La región del cuello está formada por 5 motivos IQ (isoleucina-glutamina), implicados en la unión a calmodulina. La cola consiste en un dominio coiled-coil, probablemente implicado en la homodimerización, seguido por una repetición en tándem de MyTH4 (*myosin tail homology 4*) y FERM (*4.1, ezrin, radixin, moesin*). Este tándem MyTH4-FERM forma un complejo de interacción con otros dominios altamente conservados (WU *et al.* 2011). Las dos repeticiones están separadas por un dominio SH3 (*src homology 3*), que interactúa con dominios proteicos ricos en prolina.

Muchos estudios han focalizado su atención sobre el rol de la miosina VIIa en el oído interno y la retina. En la cóclea la proteína se localiza en los estereocilios de las células ciliadas, donde está implicada en su morfogénesis y en la mecanotransducción. Además su actividad motora contribuye a procesos como la diferenciación y organización de los estereocilios. En la retina, la miosina VIIa se expresa en el RPE y los fotorreceptores. Está implicada en la distribución y migración de melanosomas y fagosomas en el RPE y en la regulación del transporte de opsina hasta el segmento externo de los fotorreceptores (HEISSLER and MANSTEIN 2012).

7.1.2 USH1C / *USH1C*

En 1992, Smith *et al*. identificaron el locus USH1C en el cromosoma 11p14.3, gracias al análisis de ligamiento en una familia Acadiana de Louisiana (SMITH *et al.* 1992). Este locus solapaba con el locus descrito para una hipoacusia autosómica recesiva, DFNB48, sugiriendo que mutaciones en un gen podrían explicar ambas enfermedades (JAIN *et al.* 1998). Años más tarde se identificó el gen responsable por dos equipos distintos al mismo tiempo (BITNER-GLINDZICZ *et al.* 2000; VERPY *et al.* 2000). Este gen, que codifica para la proteína harmonina, consta de 28 exones, 20 de ellos constitutivos y 8 con s*plicing* alternativo, por lo que se han descrito diferentes transcritos. Las distintas isoformas podrían clasificarse en tres clases (A-C) (Figura 8b). Las isoformas de la clase A contienen 3 dominios PDZ (*PSD95, discs large, ZO-1*) y uno *coiled-coil*. Ambos tipos de dominios están implicados en las interacciones proteína-proteína. La clase B, es la isoforma más larga y, contiene un dominio *coiled-coil* adicional, además de una región PST (*Proline-Serine-Threonine rich region*) capaz de unirse a dominios como SH3 o la proteína reguladora de la actina (profilina) y un dominio PBM (*PDZ Binding Motif)*, terminal de unión a PDZ. Y por último, las isoformas tipo C tienen solo 2 dominios PDZ y un *coiled-coil*.

Tanto en los fotorreceptores de la retina, como en los estereocilios de las células ciliadas del oído interno se han encontrado diferentes isoformas de la harmonina. Por estas razones la proteína harmonina parece desempeñar un papel importante en el ensamblaje de las proteínas Usher en ambos tipos celulares.

7.1.3 USH1D / *CDH23*

Este locus se identificó en el cromosoma 10, mediante estudios de cartografiado de homocigosidad en una familia pakistaní (WAYNE *et al.* 1996). Esta región solapaba con la descrita para sordera autosómica recesiva no síndrómica, DFNB12 (CHAIB *et al.* 1996). El gen se describió por dos grupos al mismo tiempo en 2001, al detectarse mutaciones tanto en familias USH1 como en familias con DFNB12 (BOLZ *et al.* 2001; BORK *et al.* 2001). El gen *CDH23* está formado por 69 exones codificantes, para los que se han descrito varios *splicing* alternativos.

La isoforma más larga (isoforma A), comprende 3354 aminoácidos, que predicen una proteína con una región extracelular formada por 27 repeticiones EC (*extracellular cadherin repeats*), un dominio transmembrana y una región intracelular (Figura 8c). Los dominios EC poseen el motivo de unión a Ca^{+2}, característico de las cadherinas, y están implicados en la interacción dependiente de Ca^{+2} entre cadherinas, que median en la adhesión célula-célula. La región intracelular contiene un PBM en el extremo C-terminal que parece interactuar con el segundo dominio PDZ de la harmonina (BOEDA *et al.* 2002; SIEMENS *et al.* 2002; ADATO *et al.* 2005a). A diferencia de la isoforma A, la B solo tiene 7 de las 27 repeticiones EC y la C solo consiste en la región intracelular, sin los dominios EC ni transmembrana.

En el oído interno, la proteína cadherina 23 se ha localizado en las células ciliadas, donde desempeña un papel fundamental en las uniones entre estereocilios adyacentes y también en el mecanismo de mecanotransducción (KAZMIERCZAK *et al.* 2007). En la retina, también se ha localizado en la región del cilio conector y la región sináptica (REINERS *et al.* 2003).

7.1.4 USH1E

El locus USH1E fue el quinto locus USH1 identificado, en 1997, por análisis de ligamiento en una familia de Marruecos que presentaba consanguinidad (CHAIB *et al.* 1997). La región de ligamiento en el cromosoma 21q21 contiene 48 genes conocidos, pero hasta la fecha no se han detectado mutaciones en ninguno.

7.1.5 USH1F / *PCDH15*

Este locus se identificó en la región del cromosoma 10q21-q22 gracias a estudios de homocigosidad en varias familias que presentaban consanguinidad. Posteriormente, se identificó el gen responsable y se encontraron mutaciones en el gen *PCDH15* en las familias previamente descritas (AHMED *et al.* 2001; ALAGRAMAM *et al.* 2001b). Dos años más tarde se describieron mutaciones en este mismo gen, pero en pacientes con sordera autosómica recesiva no sindrómica (AHMED *et al.* 2003).

Este gen presenta *splicing* alternativo e inicialmente constaba de 33 exones. Posteriormente, se caracterizaron 6 exones adicionales y las diferentes isoformas se clasificaron en 4 clases (Ahmed *et al*., 2008). Las principales clases de isoformas CD1, CD2 y CD3, están compuestas por 11 dominios EC, una región transmembrana y un motivo CD (*carboxy-terminal cytoplasmic domain*). Esta región intracelular es la que difiere entre las tres isoformas. Las isoformas de la clase SI, isoforma secretada encontrada en células de linfoma humano NK/T, solo tienen 10 dominios EC y carecen de la región que contiene el dominio transmembrana y citoplasmático (ROUGET-QUERMALET *et al.* 2006).

La protocadherina 15 se ha localizado en las células ciliadas del oído interno y en los fotorreceptores de la retina. En los fotorreceptores ha sido detectada en la región sináptica y en la membrana y discos del segmento externo (AHMED *et al.* 2003). En el oído interno, la protocadherina-15 (como cadherina-23) está presente en las células ciliadas en desarrollo, concretamente en las uniones extracelulares que conectan los estereocilios y el cinetocilio y regula su morfogénesis. En células maduras esta proteína forma parte, junto con la cadherina-23, de las *tips links*, que regulan los canales de mecanotransducción (KAZMIERCZAK *et al.* 2007).

7.1.6 USH1G / *USH1G*

Los estudios en una familia palestina permitieron la identificación del locus USH1G en el cromosoma 17 (MUSTAPHA *et al.* 2002b), en una región de unos 23cM que solapaba con dos loci descritos para sorderas autosómicas dominantes no sindrómicas, DFNA20 y DFNA26 (MORELL *et al.* 2000; DEWAN *et al.* 2003; KEMPERMAN *et al.* 2004). Posteriormente, se detectaron mutaciones en el gen *ACTG1*, para todas las familias con sordera dominante mencionadas anteriormente (VAN WIJK *et al.* 2003; ZHU *et al.* 2003; RENDTORFF *et al.* 2006; LIU *et al.* 2008). Un cartografiado más fino en dos familias consanguíneas excluyó el gen *ACTG1* como responsable de USH1G y posteriores estudios mutacionales en pacientes USH1 permitieron la detección de mutaciones en el gen *USH1G*, que codifica para la proteína SANS (WEIL *et al.* 2003).

Esta proteína de 461 amino ácidos está compuesta por 3 repeticiones *ankirina* (ANK), una región central (CEN), un dominio SAM (*sterile alpha motif*) y un dominio PBM. Es conocida como una proteína "de andamiaje", capaz de unir a otras proteínas para formar complejos. Los dominios ANK están implicados en la interacción proteína-proteína, la región CEN parece interaccionar con la proteína miosina VIIa, el dominio SAM está implicado en la transducción de señal célula-célula y también en dimerización y, por último el motivo PBM es capaz de unirse a los dominios PDZ de la harmonina y whirlina.

En el oído interno SANS está presente en las células ciliadas externas y vestibulares, también en el cinetocilio y sinapsis de las células ciliadas (ADATO *et al.* 2005a). En los fotorreceptores se localizan en el segmento interno, especialmente concentrada en el cuerpo basal ciliar, cilio conector y en la región sináptica (MAERKER *et al.* 2008).

7.1.7 USH1H

Este locus se identificó en el 2008 en dos familias consanguíneas de origen pakistaní, en el cromosoma 15q22-q23, región que abarca unos 4.9cM. Este locus solapa con el locus descrito para un tipo de sordera autosómica recesiva no sindrómica, DFNB48. Se han estudiado varios genes candidatos presentes en dicha región, pero por el momento no se han encontrado mutaciones y el gen implicado sigue sin identificar (AHMED *et al.* 2009).

7.1.8 USH1J */ CIB2*

Este locus se ha descrito recientemente por el mismo equipo que cartografió previamente el locus USH1H (RIAZUDDIN *et al.* 2012). Este grupo encontró mutaciones en el gen *CIB2* en familias pakistanís con DFNB48, 54 de ellas presentaban una misma mutación en homocigosis y otras 2 familias una mutación que segregaba con la sordera. Este gen está situado distalmente a la región crítica delimitada para el locus USH1H, y estudios de *CIB2* en las familias en las que se cartografió el locus USH1H resultaron

negativos. Sin embargo, en una familia con USH1 se detectó una mutación en homocigosis en el gen *CIB2*. Por tanto se ha caracterizado un nuevo locus responsable de síndrome de Usher tipo 1J y sordera no sindrómica DFNB48.

El gen está compuesto por 6 exones y codifica para tres isoformas distintas (A-C). La proteína pertenece a la familias de las *calcium and integrin binding protein*, y contiene 3 o 4 dominios *EF-hand* que cambian de conformación por la unión a Ca^{++} y mediarían en la señal intracelular de Ca^{++}. Además es capaz de interaccionar con la miosina VIIa y whirlina, por lo que entraría a integrar la red proteica formada por las proteínas Usher.

La proteína tiene una elevada expresión en oído interno y retina de mamífero. En adultos se ha localizado en IHCs, OHCs, citoplasma de las células sostén y en los estereocilios, donde podría estar implicada en la señal de Ca^{++} que regula la mecanotransducción. En retina se ha localizado tanto en el segmento externo como interno y en el RPE, y se ha visto que es importante para el mantenimiento y función de los fotorreceptores y para la homeostasis del Ca^{++} (RIAZUDDIN *et al.* 2012).

7.1.9 USH1K

El análisis de haplotipos en dos familias consanguíneas pakistanís revelaron este nuevo locus USH1K en el cromosoma 10p11.21-q21.1, que abarca une región de 11.74cM (JAWOREK *et al.* 2012). El gen *PCDH15*, responsable de USH1F y DFNB23, se encuentra en esta región, pero no se han encontrado mutaciones en ninguna de las dos familias. En este intervalo de 20.20Mb, se encuentran 85 genes conocidos, 53 de los cuales se expresan en el oído interno. Cuatro de ellos (*GJD4, FZD4, RET* y *LRRC18*) se eligieron como principales candidatos, pero no se encontraron mutaciones.

Además, la región descrita solapa con el locus DFNB33, implicado en sordera autosómica recesiva no sindrómica, por lo que se sugiere que ambos trastornos podrían estar causados por mutaciones alélicas.

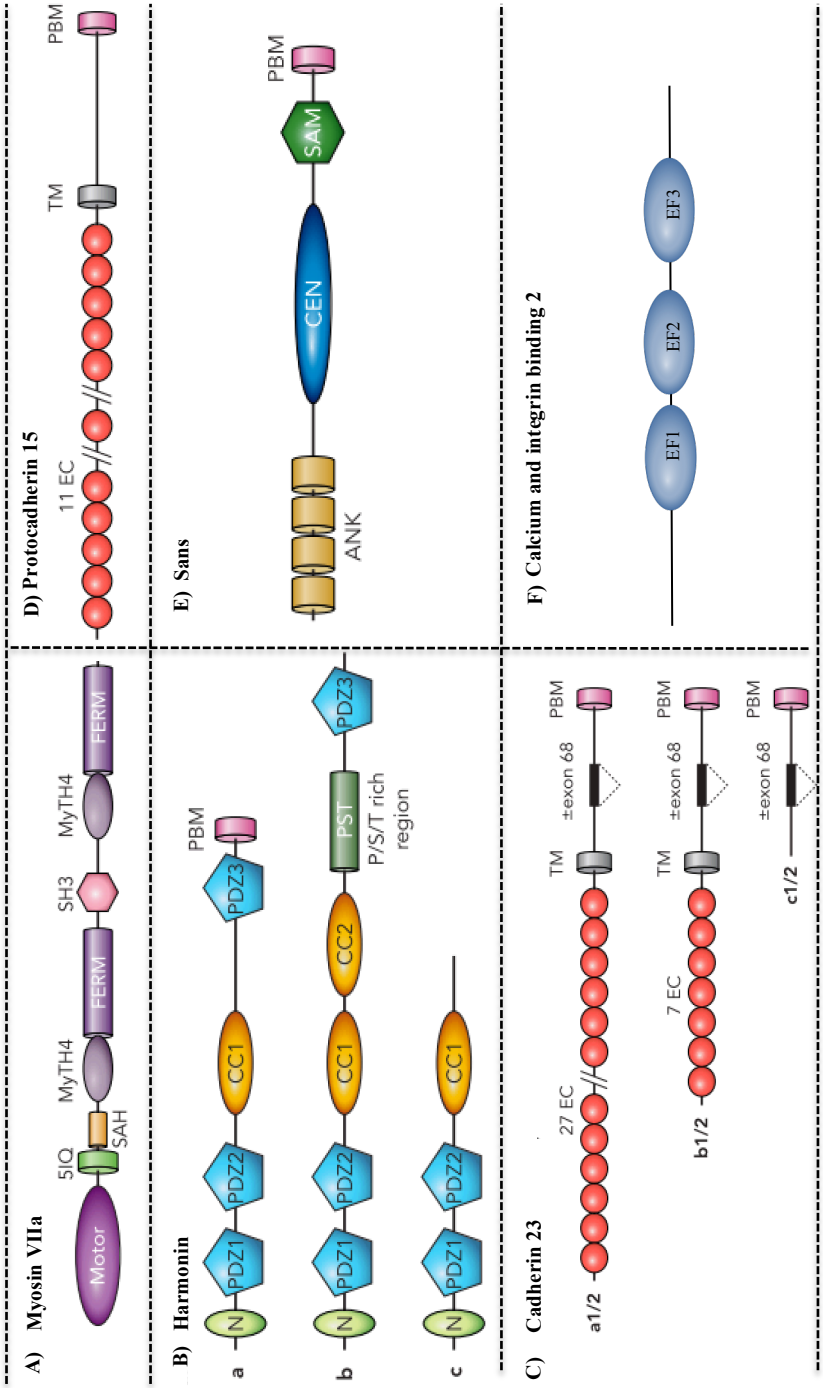


Figura 8. Representación de las proteínas codificadas por los genes USH1. (N) NH_3 terminal-domain, (PDZ) PDZ-domain, (CC) coiled-coil domain, (PST) P/S/T rich domain, (PBM) PDZ-binding motif, (TM) transmembrane domain, (EC) extracellular cadherin repeat, (ANK) ankyrin repeat, (CEN) central domain, (SAM) sterile alpha domain, (SH3) Src homology 3 domain, (Calx) calcium echanger, (MyTH4-FERM) MyTH4-FERM tandem. (IO) IO-motif. (Motor) Motor domain. (EF) EF-hand domain. Imáaenes adaptadas de (PAN and ZHANG 2012).

7.1.10 Implicación de los genes USH1 en el síndrome de Usher

Recientemente, se han publicado varios estudios que nos permiten comparar la implicación de los genes USH1 con los obtenidos en el seno de nuestro laboratorio a lo largo de los años. Le Quesne *et al*. y Bonnet *et al*. presentaron sendos trabajos en 2012 en los que estudiaban 47 (población inglesa) y 27 (población francesa) pacientes respectivamente, afectos de síndrome de Usher tipo I (BONNET and EL-AMRAOUI 2012; LE QUESNE STABEJ *et al.* 2012). Anteriormente, Roux *et al*. publicaron un estudio con un número bastante superior de pacientes USH1 franceses, un total de 92 (ROUX *et al.* 2011). Los resultados obtenidos en los diferentes trabajos se encuentran representados en la Figura 9.

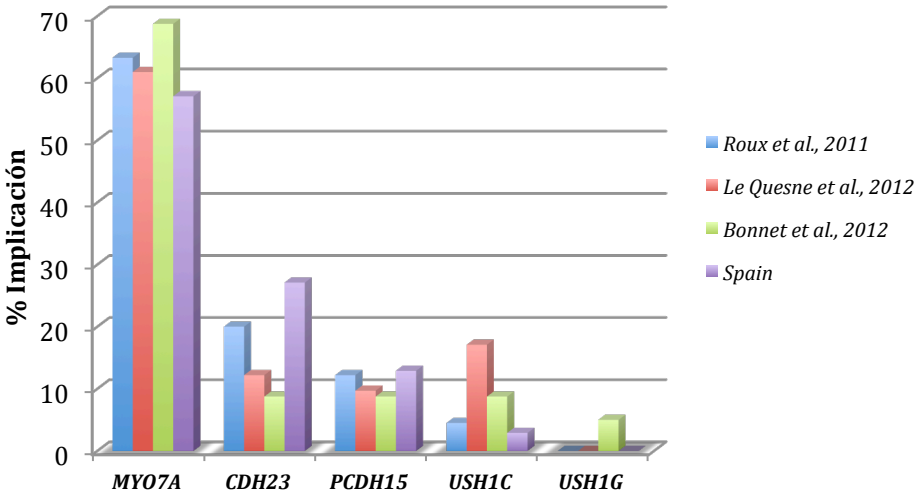


Figura 9. Porcentaje de implicación de los genes USH1 en diferentes estudios. Esta gráfica solo tiene en cuenta los pacientes para los que un genotipo patogénico ha sido descrito.

En todos los estudios el gen *MYO7A* resulta tener una mayor implicación en pacientes con USH1, resultando ser el responsable de alrededor de un 60% de los casos. Con un porcentaje de implicación bastante menor siguen los genes *CDH23* y *PCDH15*, excepto en el caso de la población inglesa donde el segundo gen de mayor implicación es *USH1C*, cuyo porcentaje en el resto de estudios es mucho menor. Esto se puede explicar debido a la presencia de una mutación de *splicing* bastante frecuente en esa población, c.496+1G>A, y nunca encontrada en población francesa y española hasta el momento. Varias excepciones como ésta han sido previamente descritas, causadas por el efecto de mutaciones fundadoras. Dos claros ejemplos son la mutación c.216G>A (*USH1C*) en población de origen Acadiano o c.733C>T (p.R254X; *PCDH15*) en familias de origen Ashkenazi, ambas elevando la implicación de dichos genes a un porcentaje mucho más

elevado al detectado en otras poblaciones. Solo en el estudio de Bonnet *et al*. se han descrito mutaciones en el gen *USH1G*, indicando la baja implicación de este gen.

7.2 Genes y proteínas implicadas en USH2

Hasta la fecha se han descrito tres locus para USH2, y se han identificado los tres genes responsables. Algunos de ellos están implicados en sorderas o RP no sindrómicas.

7.2.1 USH2A / *USH2A*

El locus USH2A fue el primer locus descrito para el síndrome de Usher, y fue identificado por Kimberling *et al*. en el año 1990 (KIMBERLING *et al.* 1990). Estudios de ligamiento en familias con USH2 le permitieron detectar ligamiento al cromosoma 1q y, cinco años más tarde este intervalo fue afinado localizándose en 1q41 (KIMBERLING *et al.* 1995). En 1998, se identificó el gen responsable, *USH2A*, y se describieron las primeras mutaciones en pacientes USH2 (EUDY *et al.* 1998). En 2000, se detectó la mutación p.C759F en el gen *USH2A*, en pacientes con RP pero sin pérdida auditiva, describiendo por primera vez la implicación de este gen en RP no sindrómica (RIVOLTA *et al.* 2000).

La estructura inicial de este gen constaba de 21 exones (Figura 10a), que codificaba para una proteína (usherina) de 1546 aminoácidos (isoforma A). Esta proteína está compuesta por un péptido señal (SP) seguido de un dominio similar a laminina G (LamG), un dominio laminina N-terminal (LamNT), 10 laminina tipo EGF-like (EGF Lam) y 4 dominios fibronectina tipo III (FN3) (EUDY *et al.* 1998). Sin embargo, la tasa de mutaciones detectada en pacientes USH2 era mucho menor de la esperada, solo se detectaron aproximadamente un 67% de las esperadas. Estos resultados hicieron revisar la estructura del gen y fue en 2004 cuando van Wijk *et al*. identificaron 51 exones adicionales en el extremos 3', así como la existencia de un splicing alternativo. Además, se encontraron mutaciones en estos nuevos exones en pacientes USH2 (VAN WIJK *et al.* 2004). De esta forma, el transcrito más largo, conocido como isoforma B, consta de 72 exones (el primero de ellos no codificante) y codifica para una proteína de 5202 aminoácidos. Además de los dominios descritos previamente, esta isoforma contiene dos dominios LamG, 28 FN3, una región transmembrana (TM) y un dominio intracelular con un motivo de unión a PDZ (PBM) en el extremo C-terminal (VAN WIJK *et al.* 2004). En 2005, se identificó un exón 71 alternativo, específico del oído interno, que añadiría 24 aminoácidos al dominio intracitoplasmático de la proteína (ADATO *et al.* 2005b).

En 2002, Bhattacharya y su equipo localizaron esta proteína en la membrana basal de retina y oído interno. Esta localización y su interacción con el colágeno IV y la fibronectina, pronto apuntaron a un importante papel estructural de la usherina en estos tejidos (BHATTACHARYA *et al.* 2004; BHATTACHARYA and COSGROVE 2005). Anticuerpos contra la isoforma B han localizado a la proteína en los *ankle links* de los estereocilios, región apical de las IHCs y células soporte, así como en la sinapsis de las células ciliadas (ADATO *et al.* 2005b). En la retina se ha observado en la región periciliar de los

fotorreceptores, cilio conector, cuerpo basal y región sináptica (VAN WIJK *et al.* 2006).

7.2.2 USH2C / *GPR98*

En 2000, diez años después de la descripción del primer locus implicado en USH, se localizó el locus USH2C en el cromosoma 5q14.3-q21 en pacientes con USH2 (PIEKE-DAHL *et al.* 2000). Cuatro años más tarde, (WESTON *et al.* 2004) identificaron mutaciones en el gen *VLGR1* (Very Large G-coupled receptor), demostrando su implicación en USH2C. Actualmente, este gen también es conocido como *GPR98* (G-protein coupled receptor 98).

El gen *GPR98* consta de 90 exones, abarcando 600Kb. Se conocen tres isoformas generadas por un splicing alternativo (MCMILLAN *et al.* 2002). La isoforma más larga (isoforma B) codifica para una de las proteínas más grandes descritas, formada por 6306 aminoácidos (Figura 10b). Como el resto de proteínas transmembrana, consta principalmente de tres regiones, una extracelular, una transmembrana y una intracelular. La parte extracelular está formada por un péptido señal, seguido de 35 dominios CalX-B (*calcium echanger β*) (homologos a los dominios específicos de proteínas que regulan los intercambios Na^{+}/Ca^{2+}), interrumpidos por un dominio LamG/TspN/PTX y 7 EAR/EPTP (*epilepsy associated repeat*). Después, la región C-terminal está compuesta por un dominio GPS (*Glutathione S-transferase*), 7 regiones trasnmembrana (TM) y, al igual que *USH2A*, un dominio de unión a PDZ. Esta región transmembrana en 7 hélices hidrofóbicas originó su clasificación en el grupo de los receptores acoplados a proteínas G (RCPG). La isoforma A, solo contiene 6 dominios CalX-B, la región transmembrana 7TM y el dominio C-terminal PBM, mientras que la isoforma C contiene un dominio LamG, más 15 motivos Calx-B pero carece de la región transmembrana y la región intracelular (WESTON *et al.* 2004).

En el oído interno esta proteína es esencial en la base de los estrereocilios, concretamente en los *ankle links* (MCGEE *et al.* 2006; YAGI *et al.* 2007). En la retina, está presente en las uniones entre la región periciliar y cilio conector de las células fotorreceptoras. Además en ambos sistemas está localizada en las regiones sinápticas (MAERKER *et al.* 2008).

7.2.3 USH2D / *DFNB31*

Este locus fue descrito inicialmente en una familia jordana con hipoacusia congénita, en el cromosoma 9q32-34 (MUSTAPHA *et al.* 2002a). Un año más tarde, se identificaron mutaciones en el gen *DFNB31* en pacientes con sordera autosómica recesiva no sindrómica (MBURU *et al.* 2003). Sin embargo, no fue hasta años más tarde, y asumiendo que las proteínas que interaccionan con las proteínas Usher, quizá fuesen responsables de algún tipo de síndrome de Usher, cuando Ebermann *et al*., describieron su implicación en familias con USH2 (EBERMANN *et al.* 2007).

Este gen consta de 12 exones y codifica para la proteína whirlina. Se han descrito dos isoformas (Figura 10c). La isoforma más larga (isoforma A), consta de los 12 exones que codifican para una proteína de 907 amino ácidos. Está compuesta por 3 dominios PDZ, una región rica en prolina y un dominio PBM, mediante el que puede formar homómeros. La isoforma corta (isoforma B), codificada por lo exones 6-12, solo contiene la región rica en prolina, el tercer dominio PDZ y el motivo PBM.

La proteína whirlina es esencial, tanto en la retina como en el oído interno. En el oído interno tiene un papel importante en la elongación de los estereocilios, en la dinámica de los filamentos de actina y el desarrollo de las células ciliadas de la cóclea. Además, se ha demostrado que es la proteína principal de anclaje de la red proteica Usher encontrada en las *ankle links* de las IHCs (BELYANTSEVA *et al.* 2005; MICHALSKI *et al.* 2007; MANOR *et al.* 2011). En la retina, la whirlina es clave en el interactoma de la región periciliar de las células fotorreceptoras (MAERKER *et al.* 2008).

7.2.4 PDZD7 / *PDZD7*

El gen *PDZD7* está localizado en el cromosoma 10 y se han descrito varias isoformas. El transcrito más largo está codificado por 18 exones y da lugar a una proteína de 1033 aminoácidos, PDZ *domain containing protein* 7 (PDZD7), con 3 dominios PDZ y un PBM (Figura 10d).

La implicación de este gen en el síndrome de Usher se describió en 2010 (EBERMANN *et al.* 2010). En un paciente con una sola mutación de codón de parada prematuro en el gen *GPR98*, se encontró otra mutación en el gen *PDZD7*, cada una de las mutaciones heredadas de un parental, sugiriéndose así un modelo de digenismo. Además, se propuso al gen *PDZD7* como posible modificador, agravando el fenotipo retiniano. Se describieron dos hermanas USH2 con una mutación en homocigosis en el gen *USH2A*, pero con fenotipo diferentes. En una de las hermanas la RP era más grave y los síntomas clínicos fueron más precoces, y justo en ella se detectó un codón de parada prematuro en el gen *PDZD7*. Además, se utilizó *zebrafish* como modelo haciendo un knockdown de *PDZD7* para la comprobación de dichas hipótesis.

Recientemente, se ha demostrado que la proteína PDZD7 colocaliza con usherina y GPR98 en los *ankle links* de los estereocilios de las células ciliadas, por lo que entraría a formar parte de la red proteica conformada por las proteínas Usher (GRATI *et al.* 2012).

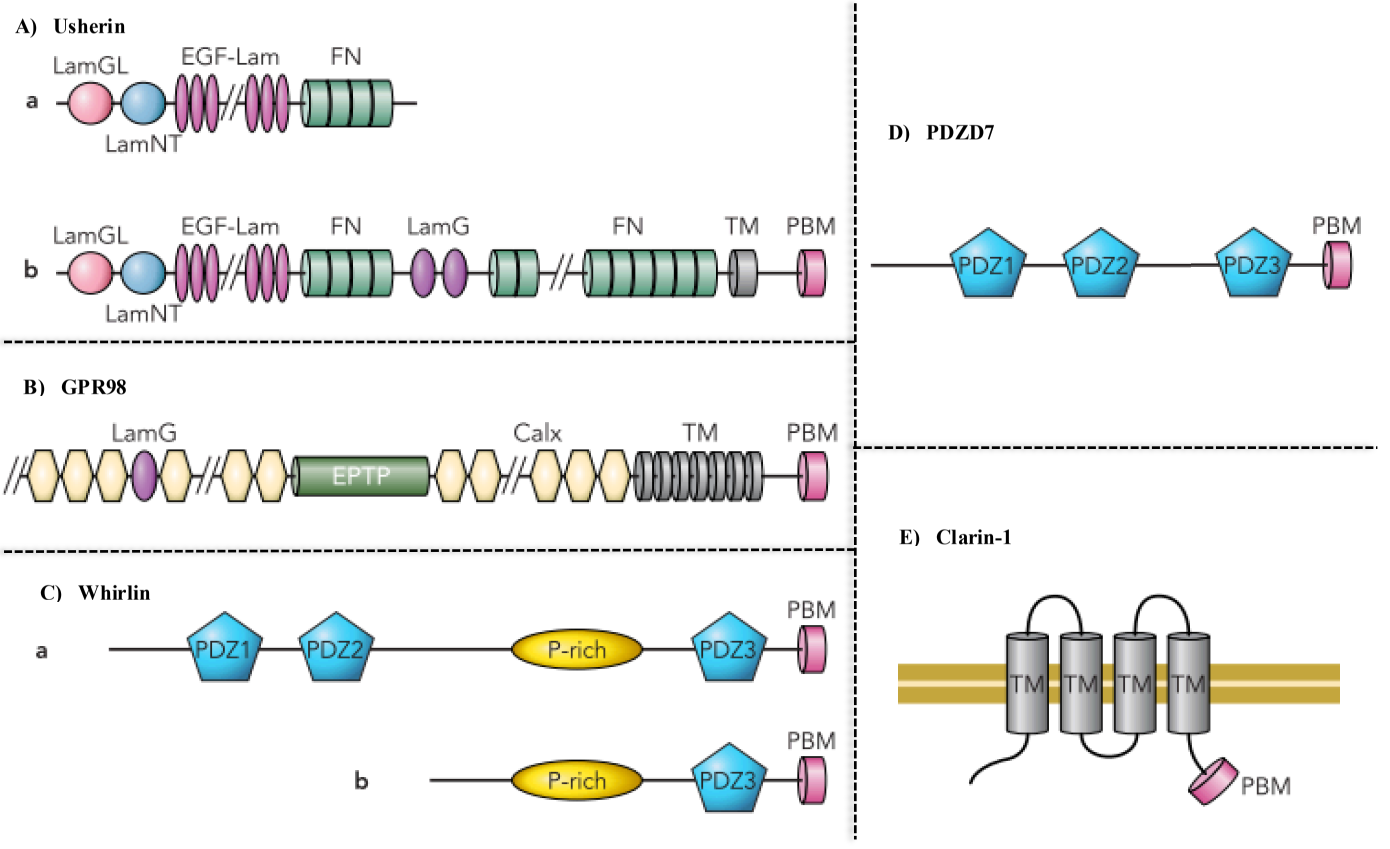


Figura 10. Representación de las proteínas USH2 y USH3. (TM) Transmembrane domain, (PBM) PDZ-binding motive, (LamG) Laminin G domain, (LamGL) Laminin G-like domain, (LamNT) N-terminal laminin domain, (FN) Fibronectin domain, (EGF-Lam) Laminin type EGF-like modules, (P-rich) Proline rich- domain, (EPTP) EPTP repeats. Imagen adaptada de (PAN and ZHANG 2012).

7.2.4 Implicación de los genes USH2 en el síndrome de Usher

Como para los genes USH1, son varios los estudios que recogen datos sobre la implicación de los genes USH2 en diferentes poblaciones (Figura 11). Dos estudios importantes son los realizados por (BESNARD *et al.* 2012) y (LE QUESNE STABEJ *et al.* 2012) en población francesa e inglesa, respectivamente. En ambos trabajos se analizó un gran número de pacientes y los resultados fueron similares, obteniendo una mayor implicación del gen *USH2A* (cerca de un 80%), seguido por *GPR98*. Sin embargo, en población inglesa no se identificaron mutaciones en el gen *DFNB31,* aunque su implicación en población francesa también fue muy baja. Otro estudió sobre los genes USH2 también en población francesa, fue realizado por (BONNET *et al.* 2011). En este caso los resultados fueron diferentes, la implicación estimada de los genes *GPR98* y *DFNB31* fue bastante superior a la encontrada en los estudios comentados anteriormente. Sin embargo, esta diferencia puede ser debida a la pequeña cohorte sobre la que se realizó el estudio, apenas 21 pacientes, que podría producir un sesgo en los porcentajes.

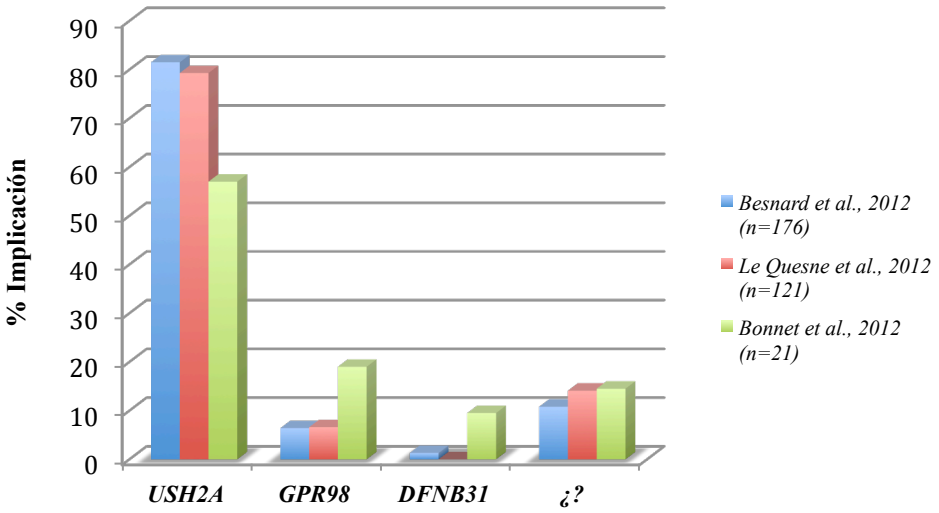


Figura 11. Representación de la implicación de cada uno de los 3 genes USH2 descritos. "¿?" = hace referencia al porcentaje de casos en los que no se han detectado mutaciones en ninguno de los tres genes USH2. "n"= número de pacientes incluidos en el estudio. El gen *PDZD7* solo fue estudiado en el trabajo de (BESNARD *et al.* 2012), y no se detectaron mutaciones, por ello este gen no se ha incluido en la gráfica.

En población española, nuestro grupo ha realizado estudios sobre el *screening* de los 21 primeros exones del gen *USH2A*, y posteriormente de sus 52 exones adicionales (NAJERA *et al.* 2002; ALLER *et al.* 2004; ALLER *et al.* 2006). Sin embargo, no se ha realizado un análisis exhaustivo de todo el gen *USH2A* en una gran cohorte, ni tampoco se han estudiado los otros dos genes USH2 (*GPR98* y *DFNB31*). Este apartado será ampliado en la sección de Resultados y discusión.

7.3 Genes y proteínas implicadas en USH3

Hasta la fecha solo un gen, *USH3A* (también conocido como *CLRN1*), se ha descrito como responsable de USH3. Petit *et al*., en 2001 sugirieron un segundo locus, USH3B, en 20q, pero no se ha podido identificar el gen responsable (PETIT 2001).

7.3.1 USH3A / *USH3A*

El locus USH3A fue localizado en el cromosoma 3 en el año 1995, gracias a estudios de ligamiento en familias finlandesas (SANKILA *et al.* 1995). Años más tarde, en 2001, se identificó el único gen reconocido como responsable de USH3 (JOENSUU *et al.* 2001).

Se han descrito al menos 11 isoformas distintas (VASTINSALO *et al.* 2011). El transcrito más largo, está codificado por 2359 nucleótidos, repartidos en 3 exones que codifican para una proteína de 232 aminoácidos, la clarina-1. La clarina-1 forma parte de la familia de proteínas con 4 motivos transmembrana (4TM). Además de esta región transmembrana, posee un dominio PBM en el extremo C-terminal (Figura 10e).

Esta proteína se expresa en las células ciliadas del oído interno y en la retina, y aunque su función no se conoce muy bien, su expresión espaciotemporal descrita en las células ciliadas de ratón, parece indicar una implicación en la maduración sináptica (GENG *et al.* 2009; ZALLOCCHI *et al.* 2009). La homología estructural y de secuencia de la clarina-1 con la proteína sináptica stargazin (proteína que desempeña un papel importante en la formación y mantenimiento de la sinapsis en el cerebelo), sugieren un rol de la clarina-1 en la sinapsis de las células neurosensoriales de la cóclea y la retina (ADATO *et al.* 2002).

El síndrome de Usher tipo 3 es el menos frecuente de los tres tipos clínicos, excepto en judíos Ashkenazi y población finlandesa, donde su implicación representa hasta un 40% de todos los casos Usher (MILLAN *et al.* 2011). Esto es debido a la presencia de mutaciones fundadoras en ambas poblaciones.

8. El interactoma "Usher"

La publicación de estudios sobre la localización e interacciones de las proteínas codificadas por los genes Usher, ha puesto en evidencia la existencia de una red proteica constituida por estas proteínas, conocida como "interactoma Usher". Esta red la encontraríamos tanto en la retina como en el oído interno, donde desempeña un papel esencial para el correcto funcionamiento de ambos sistemas.

El eje central de este interactoma está conformado por proteínas con un dominio PDZ, las homólogas harmonina y whirlina. En la mayoría de los casos uno o más de sus dominios PDZ interaccionan con los motivos de unión a PDZ clase I C-terminal (PBM) u otros dominios de unión a PDZ internos de las proteínas a las que se unen. Todas las proteínas USH1 y 2 interaccionan con ambas proteínas. Además, en estas interacciones proteína-proteína, muchas de las proteínas de esta red (SANS, harmonina, whirlina y myosina VIIa) son capaces de formas homodímeros (KREMER *et al.* 2006).

8.1 Proteínas Usher en el oído interno

En el oído interno desempeñan un papel fundamental. Gracias a los fenotipos observados en los diferentes modelos animales mutantes para los genes Usher, se ha demostrado que estas proteínas son esenciales para el desarrollo y cohesión de los estereocilios en las células ciliadas de cóclea y vestíbulo. Durante el crecimiento y desarrollo, y también en la etapa adulta, los estereocilios mantienen su cohesión mediante uniones inter-estereociliares y también estereocilio-cinetocilio. El cinetocilio es un cilio verdadero esencial para la correcta orientación de los estereocilios durante el desarrollo que desparecerá en las células maduras. Estas uniones van evolucionando a medida que se avanza en el desarrollo (Figura 12). En las primeras etapas encontramos los *transient lateral links*, que disminuyen a medida que aparecen los *ankle links* (son *transient basal links*) en la base de los estereocilios. Estos últimos van desapareciendo para formarse los *horizontal top conector*, que ya son mantenidos en el adulto. Otro tipo de uniones importantes son los *tip links*, que conectan la región superior de un estereocilio con el estereocilio vecino más alto. Las proteínas Usher son fundamentales para la formación y mantenimiento de estas uniones (KREMER *et al.* 2006; MULLER 2008).

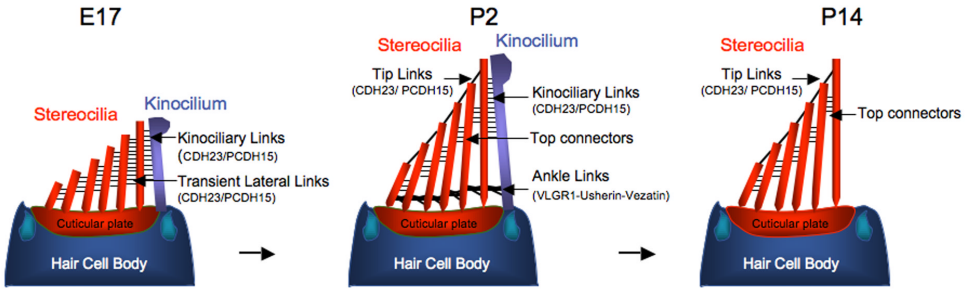


Figura 12. Desarrollo de los estereocilios en las células ciliadas de la cóclea. (E) Embrión, (P) Postnatal. Imagen modificada de (MÜLLER *et al*. 2009).

Principalmente nos encontramos con dos interactomas que participarían en el desarrollo y mantenimiento de las células ciliadas (Figura 13).

Uno estaría compuesto principalmente por las proteínas USH2, es decir por las proteínas transmembrana usherina y GPR98, que en asociación con la whirlina constituyen los *ankle links*, donde también parece estar implicada la miosina VIIa (MICHALSKI *et al.* 2007). Además recientemente se ha publicado un estudio en el que se observa que PDZD7 colocaliza con las proteínas usherina, GPR98 y whirlina en esos *ankle links* y que además los dominios citoplasmáticos de usherina y GPR98 serían capaces de interaccionar con whirlina y PDZD7. Esto incluiría a PDZD7 en el interactoma Usher, formando parte de estos *ankle links* (GRATI *et al.* 2012).

Por otro lado tendríamos el interactoma formado por las proteínas USH1, donde harmonina y SANS formarían un eje central (Figura 13). Las cadherinas, protocadherina 15 y cadherina 23, además de estar presentes en los *transient lateral links* forman parte de los *tips links*. En 2007, Kazmierczak *et al*., describió que cada tip link está formado por homodímeros de cadherina 23 que interaccionan en *trans* con homodímeros de protocadherina 15, donde cadherina 23 está en la parte superior del *tip link* y protocadherina 15 en la parte inferior. Además, se piensa que en estas regiones de uniones cadherina 23 - protocadherina 15 hay proteínas como la harmonina necesarias para el anclaje de estos *links* al citoplasma. También se ha propuesto la presencia de varias miosinas (miosina Ic, miosina IIIa, miosina VIIa, y miosina XVa) que participarían en la adaptación de la mecanotransducción, en la dinámica de la actina de los estereocilios, localización de las *lateral links* y en el transporte (KAZMIERCZAK *et al.* 2007; MULLER 2008; SAKAGUCHI *et al.* 2009). Además se ha descrito que mutaciones en proteínas como cadherina 23, protocadherina 15 o miosina VIIa afectan a la formación de los *tips links* y el proceso de mecanotransducción, confirmando su importancia en esos procesos (ALAGRAMAM *et al.* 2011).

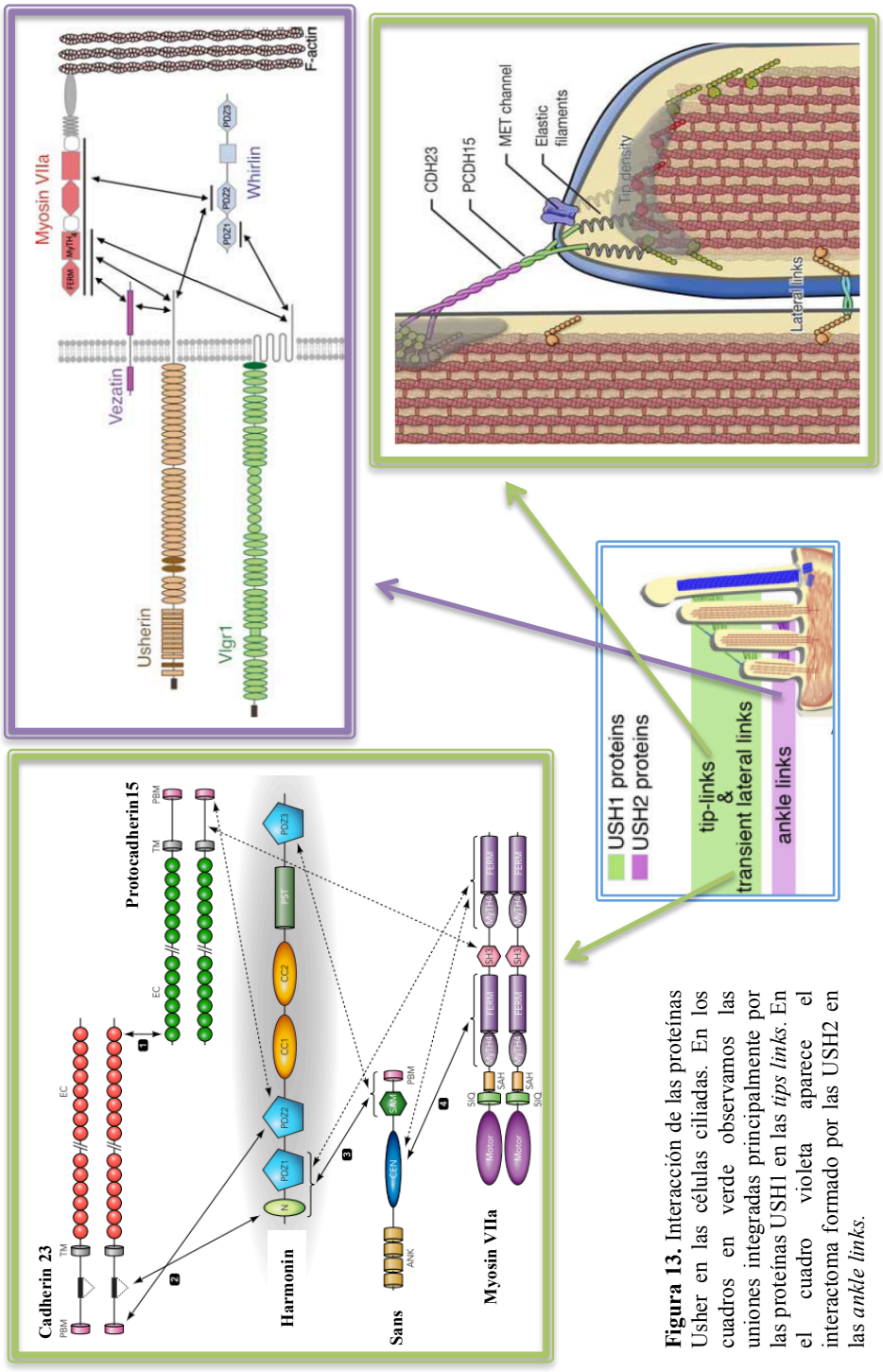


Figura 13. Interacción de las proteínas Usher en las células ciliadas. En los cuadros en verde observamos las uniones integradas principalmente por las proteínas USH1 en las *tips links*. En el cuadro violeta aparece el interactoma formado por las USH2 en las *ankle links*.

8.2 Proteínas Usher en la retina

En cuanto a la retina, las proteínas Usher se han localizado en la terminal sináptica de los fotorreceptores, en la región ciliar entre el segmento interno y externo, más concretamente en el cilio conector y en los procesos *calyceal*.

En 2008, Maerker *et al.* describen la presencia de una red proteíca USH1 y USH2 en la región periciliar de la zona apical del segmento interno, y la adyacente al cilio conector en los fotorreceptores (Figura 14A) (MAERKER *et al.* 2008). Esta red está organizada por las proteínas de andamiaje whirlina y SANS. Se observa una asociación de SANS con los microtúbulos, lo que sugiere una unión de esta red al citoesqueleto de microtúbulos y también una posible función en el transporte intracelular. La asociación vista en los *ankle links* de los estereocilios de GPR98 y usherina con whirlina parece ser similar en esta región periciliar. Hay una interacción directa de la región citoplasmática de estas dos proteínas transmembrana con whirlina y SANS, conectando así membranas adyacentes, y parece posible que GPR98 y usherina formen homodímeros o heterodímeros por interacción de sus largos ectodominios, esenciales para la adhesión celular. En otras palabras, hacen de puente entre el cilio conector y la región periciliar para mantener la estructura de soporte y también transporte de vesículas por la asociación de SANS a microtúbulos. Además, la proteína miosina VIIa es capaz de interaccionar con las proteínas presentes en esta red y participaría en el transporte de componentes de membrana a través del filamento de actina. Esta proteína motora también se une directamente a las regiones citoplasmáticas de usherina y GPR98, por lo que podría tener también una función de apoyo de esta red a la membrana periciliar.

Además recientemente, se ha publicado un trabajo en el que se estudian las proteínas Usher en fotorrecpetores de macaco (SAHLY *et al.* 2012). Se ha observado la co-localización de todas las proteínas USH1 en la región entre el segmento externo e interno, en el proceso *calyceal* y en la región basolateral del segmento externo de bastones y conos (Figura 14B). Con estos resultados se propone la participación de una red formada por las proteínas USH1, esencial para el mantenimiento de los fotorreceptores. Además, en fotorreceptores de ratones no se detectó esta red y éstos carecen de este proceso *calyceal*. Estos hechos podrían explicar por qué los roedores no mimetizan la degeneración retiniana observada en pacientes con síndrome de Usher.

El papel de las proteínas USH en RPE y sinapsis no es tan conocido, pero se podría postular que algunas de estas proteínas podrían tener una función sinaptognénica y ayudarían a transportar las moléculas de desecho producidas por la renovación de los discos del segmento externo de los fotorreceptores.

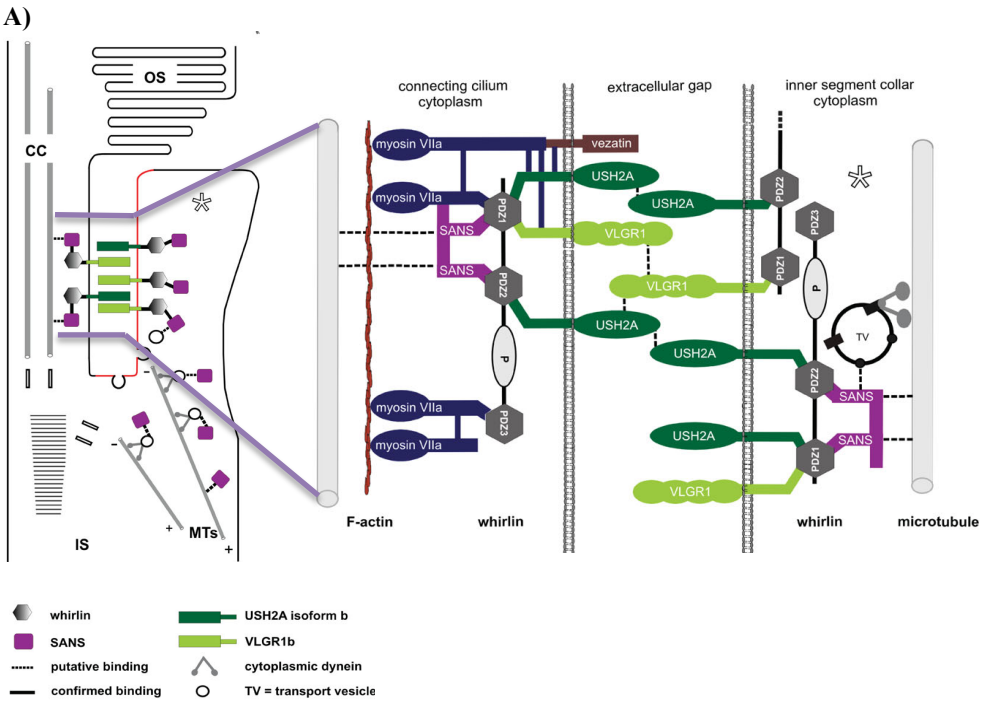


Figura 14. Localización de las proteínas Usher en los fotoreceptores. A) Representación de la red proteica integrada por las proteínas Usher. CC=cilio conector (connecting cilium). OS: segmento externo (outer segment). IS: segmento interno (inner segment). MTs: transporte intracelular asociado a microtúbulos (microtubule-associated intracellular transport). ✱ Indica cuello del segmento interno (inner segment collar). B) Diferencias en la arquitectura de la región más apical del segmento interno de los fotorreceptores de ratón y macaco. En el cono de macaco los procesos *calyceal* no están representados para poder observar la distribución de las proteínas USH1, presentes en la unión entre el segmento externo e interno.

9. Modelos animales

El desarrollo de modelos animales es esencial para dilucidar las bases moleculares y fisiológicas implicadas en el síndrome de Usher. Una de las características principales a la hora de la elección de un modelo animal, aparte de su fácil manejo, es que mimeticen el fenotipo observado en humanos, sobre todo de cara al desarrollo de posibles terapias. Dos de los animales más utilizados son el ratón y el pez *Danio rerio (zebrafish)*. En la tabla 5 se presentan los diferentes modelos desarrollados para cada uno de los genes Usher.

Tabla 5. Modelos animales desarrollados para cada uno de los genes Usher**.**

	Modelo Animal			
Gen	***Mus Musculus***	**Referencias**	***Danio rerio***	**Referencias**
MYO7A	Shaker1	(GIBSON *et al.* 1995; MBURU *et al.* 1997)	mariner	(ERNEST *et al.* 2000)
USH1C	Deaf circler	(JOHNSON *et al.* 2003)	Ush1c	(PHILLIPS *et al.* 2011)
CDH23	Waltzer	(DI PALMA *et al.* 2001)	sputnik	(SOLLNER *et al.* 2004)
PCDH15	Ames waltzer	(ALAGRAMAM *et al.* 2001a)	orbiter	(SEILER *et al.* 2005)
USH1G	Jackson shaker	(KIKKAWA *et al.* 2003)	--	--
CIB2	--	--	morpholino knockdown	(RIAZUDDIN *et al.* 2012)
USH2A	Knockout	(LIU *et al.* 2007)	morpholino knockdown	(EBERMANN *et al.* 2010)
GPR98	Vlgr1 del7TM	(MCMILLAN and WHITE 2004; MCGEE *et al.* 2006)	morpholino knockdown	(EBERMANN *et al.* 2010)
DFNB31	Whirler	(LANE 1963; HOLME *et al.* 2002; MBURU *et al.* 2003)	--	--
CLRN1	Knockout	(GELLER *et al.* 2009)	--	--

Muchos de los modelos de ratón para el síndrome de Usher se identificaron gracias a su comportamiento, movimientos en círculo y movimientos bruscos de cabeza, debido a una disfunción vestibular. Estas características dieron nombres a varios de ellos.

Se han descrito modelos de ratón para todos los genes Usher, excepto para *CIB2* que se ha identificado recientemente. De esta forma los ratones, *shaker1* (sh1), *deaf circler* (dfcr), *waltzer* (v), *ames waltzer* (av) y *Jackson shaker* (js) tienen mutaciones en genes homólogos a los genes humanos *MYO7A*, *USH1C*, *CDH23*, *PCDH15*, *USH1G1*, respectivamente (Tabla 5). El ratón *knockout* para *USH2A* (LIU *et al.* 2007) no exhibe este comportamiento, consistente con la función vestibular normal que presentan los pacientes USH2. El modelo para *DFNB31*, *whirler*, gira en círculos (como su nombre indica) y mueve la cabeza bruscamente (LANE 1963)

A pesar del éxito de la manifestación en estos ratones de la sordera y problemas del equilibrio, no poseían el fenotipo mutante en la retina. Ninguno de los modelos USH1 en ratón presentaban degeneración retiniana. El ratón mutante expresando una proteína GPR98 carente de su dominio transmembrana y citoplasmático (Vlgr/del7TM), también presentaba una histología de la retina normal. El único modelo que ha presentado una clara degeneración retiniana es el *Knockout* para *USH2A*, estos ratones presentan una morfología y número de fotorreceptores normal hasta los 10 meses, pero a los 20, la mitad de ellos se han perdido (LIU *et al.* 2007).

Aunque la mayoría de los modelos no presentan una clara degeneración de la retina, sí que se ha descrito para alguno de ellos una reducción en las amplitudes de los electroretinogramas, indicando un defecto en la respuesta de los fotorreceptores, concretamente para algunos alelos de *shaker1*, *waltzer*, *ames waltzer*, *Gpr98*-mutante (LIBBY and STEEL 2001; HAYWOOD-WATSON *et al.* 2006). Como ya se ha comentado en un apartado anterior, Shaly *et al*., describieron recientemente en fotorreceptores de macaco una red integrada por las proteínas USH1 en la región limitante entre el segmento interno y externo no detectada en roedores(SAHLY *et al.* 2012). Además, estos últimos no poseen los procesos *calyceal*, conservados en primates, pudiendo ser una explicación del por qué en ratones no se observa una degeneración de los fotorreceptores.

También se han generado modelos utilizando el pez *zebrafish,* con genes mutantes o *knockdown* ortólogos para los genes Usher. Uno de ellos es *mariner*, que expresa una miosina VIIa mutante (similar a *shaker1* en ratón) y exhibe un comportamiento en círculos y defectos funcionales y morfológicos en las células ciliadas del oído interno (ERNEST *et al.* 2000). *Zebrafish* tiene dos ortólogos para el gen *PCDH15*, siendo *orbiter* mutante para la *Pcdh15a*. *Sputnik* expresa una Cadherina 23 mutante, con una mecanotransducción reducida o nula.

También se están desarrollando mutantes mediante el uso de morfolinos, como es el caso de *PDZD7, USH2A, GPR98* y *CIB2* (EBERMANN *et al.* 2010; RIAZUDDIN *et al.* 2012). En *zebrafish* se identificaron dos ortólogos para *PDZD7*, *pdzd7a* y *pdzd7b*. Ambos se expresan en las células mecanosensoriales del oído interno y células de la retina, pero *pdzd7a* mostraba una mayor expresión, por lo que se eligió para el estudio. El *knockdown pdzd7a* presentaba síntomas similares a los encontrados en otros modelos descritos para el USH, movimientos en círculos y desorganización de los estereocilios. Las proteínas Gpr98 y Ush2a en este modelo se localizan en el cilio conector y sinapsis de las células fotorreceptoras. Cuando se generó el *knockdown pdzd7*, se observó una disminución de la

proteína Gpr98 en esas regiones. El mutante *pdzd7/gpr98* presentaba el típico fenotipo USH (movimiento en círculos y degeneración de la retina), apoyando la hipótesis de una herencia digénica de mutaciones en ambos genes.

En el *knockdown* de *zebrafish* para *CIB2* se observó que hasta el 80% de los peces no respondían a estímulos acústicos y que el número de neuromastos con células ciliadas en la línea lateral era menor. Para este gen también se desarrollaron mutantes usando como modelo *Drosophila,* que contiene un gen que codifica para una proteína similar a CIB2 y se expresa en varios tejidos, incluido el ojo adulto. Se generó un *knockdown* a través de RNA de interferencia y se observó una mayor degeneración de los fotorreceptores y una respuesta menor a los estímulos, posiblemente debido a un incremento de Ca^{++} intracelular (RIAZUDDIN *et al.* 2012).

10. Diagnóstico molecular del síndrome de Usher

La gran heterogeneidad genética y alélica del síndrome de Usher, junto al gran tamaño de alguno de los genes implicados, dificulta enormemente su diagnóstico molecular. A continuación se va a describir brevemente la evolución de las tecnologías aplicadas al diagnóstico molecular del USH.

10.1 Secuenciación Sanger

Tradicionalmente el estudio se ha realizado analizando gen a gen, empezando por el de mayor implicación dentro de cada tipo clínico. Varias técnicas se han ido utilizando a lo largo de los años para la identificación de las mutaciones responsables de la enfermedad. Inicialmente estos estudios se realizaban mediante la combinación de SSCPs (polimorfismo conformacional de cadena sencilla) y Sanger. Los SSCPs permitían la detección de variaciones de secuencia en función del diferente plegamiento de las cadenas simples de ADN al ser analizadas en un gel de poliacrilamida y si se observaba un patrón electroforético anómalo se procedía a su secuenciación automática mediante Sanger. Sin embargo, su menor sensibilidad en comparación a la secuenciación directa, hicieron que rápidamente se pasara a la secuenciación completa de los genes. Se han realizado muchos análisis mutacionales de los genes Usher, y en diferentes tipos de poblaciones, gracias a los que se han descrito cientos de variaciones y mutaciones. Aunque tenga una elevada sensibilidad, debido al tamaño de genes como *USH2A*, *GPR98* o *MYO7A,* la presencia de casos atípicos, que presentan mutaciones en genes típicamente responsables de otro tipo clínico, y de pacientes sin clasificar debido a la falta de datos clínicos, hacen que la secuenciación de los genes Usher en cascada tenga un elevado coste de tiempo y dinero.

10.2 Microarrays de genotipado

Los microarrays de genotipado tienen como finalidad facilitar, en la medida de lo posible, los estudios genéticos. Consisten en múltiples sondas de ADN unidas a una superficie sólida sobre la que se hibridarán muestras de ADN de los pacientes y permitirá la identificación de variaciones genéticas o mutaciones.

En el año 2007, se aplicó ésta tecnología para el estudio del síndrome de Usher (CREMERS *et al.* 2007). Éste microarray, comercializado por la empresa Asper Biotech (Tartu, Estonia), está basado en la técnica APEX (Arrayed Primer Extension) e inicialmente era capaz de detectar 298 mutaciones previamente identificadas en ocho de los genes Usher. En 2010, Jaijo *et al*. evaluaron una nueva versión del microchip que contenía 429 mutaciones conocidas en 183 familias con síndrome de Usher (JAIJO *et al.* 2010). Identificaron 32 mutaciones en los pacientes estudiados, pero también cambios de dudosa patogeneicidad, así como, errores en la detección de mutaciones o en su nomenclatura.

Aunque el microarray solo detecta mutaciones previamente descritas y son necesarias continuas actualizaciones (la última versión incluye 651 mutaciones), debido a su bajo coste, esta técnica puede ser útil para poder priorizar el gen a estudiar.

10.3 Secuenciación de nueva generación

Hace más de 30 años, en 1977 concretamente, se produjo un gran avance para el mundo de la genética, Allan Maxam y Walter Gilbert, y Frederick Sanger desarrollaron simultáneamente métodos para la determinación de la secuencia de nucleótidos del ADN. Con las subsecuentes mejoras el método de secuenciación Sanger ha sido usado de forma rutinaria en genética.

El método de Sanger fue el usado en la secuenciación del primer genoma humano completado en el año 2003, gracias al conocido Proyecto del Genoma humano (HGP), un trabajo llevado a cabo por varios centros en cooperación y que supuso 13 años de esfuerzo, y cerca de 2.700 millones de dólares.

Poco tiempo después, en el 2005, se desarrolló la primera plataforma de secuenciación de nueva generación (NGS: *next generation sequencing*), lo que sería el inicio de una nueva era. Por comparación al coste que supuso el primer genoma secuenciado, tanto en tiempo como en dinero, en 2008 se produjo la secuenciación de un genoma humano en tan solo 5 meses y 1.5 millones, gracias a las tecnologías de nueva generación que han emergido en estos últimos años (WHEELER *et al.* 2008). En agosto de ese mismo año se anunciaba el reto de poder secuenciar en un futuro, un genoma humano por unos 1000 dólares.

Desde aquella primera plataforma, se han producido muchos avances tanto en la química utilizada como en la detección de la señal generada. Estas plataformas tienen la capacidad de secuenciar millones de bases, a un reducido coste por nucleótido, y también en poco tiempo. Tal cantidad de información plantea retos a la hora de su almacenamiento, procesado y manejo de los datos, alineamiento o ensamblaje, anotación, interpretación y una posible translación de estos datos a la clínica, que poco a poco se van solventando.

Son múltiples las aplicaciones de estas nuevas plataformas de secuenciación:

- Secuenciación de genomas (WGS: *whole genome sequencing*).
- Secuenciación de exomas (WES: *whole exome sequencing*).
- Secuenciación de determinadas regiones (*targeted resequencing*).
- Estudios de trancriptomas para perfiles celulares, de tejido u órganos.
- Identificación de marcas epigenéticas.

Actualmente son varias las plataformas disponibles y todas se basan en una primera amplificación del ADN seguida de la secuenciación. Estos son conocidos como los secuenciadores de segunda generación, pero ya se están desarrollando los llamados de tercera generación, en los que esta primera amplificación ya no sería necesaria.

Los secuenciadores que acabamos de comentar tienen una gran capacidad, por lo que posteriormente se desarrollaron los secuenciadores de bancada, conocidos como *benchtop*. Son pequeños sistemas NGS, menos caros, con protocolos más sencillos y

también un análisis de datos más simple. Principalmente hay tres plataformas que han desarrollado estos secuenciadores de bancada y en la tabla 6 podemos observar las principales características de cada una de ellas.

Tabla 6. Comparación de las tres plataformas con secuenciadores *benchtop* disponibles.

Plataforma	**MiSeq**	**454 GS Junior**	**Ion Torrent PGM**
Compañía	Illumina	Roche	Life Technology
Método Amplificación	Amplificación por puentes	PCR en emulsión	PCR en emulsión
Método Secuenciación	Secuenciación por síntesis utilizando terminadores marcados	Secuenciación por síntesis. Pirosecuenciación	Secuenciación por síntesis. Semiconducción de iones
Longitud Lectura (bases)	2x150	400	35-400
Capacidad/run	4.5Gb	40Mb	10Mb-1Gb
Tiempo/run	24h	8h	4h
Precio/run ($)	750	1100	625

Los tres secuenciadores difieren principalmente en el tipo de secuenciación utilizado. A través de diferentes trabajos se ha visto que el MiSeq es el menos propenso a errores de lectura, problema principal de los otras dos plataformas que presentan problemas en regiones con homopolímeros (LOMAN *et al.* 2012) debido a su sistema de detección de la señal, variaciones en la intensidad lumínica en el caso de GS Junior o cambios en el Ph en el Ion Torrent. También vemos diferencias en cuanto a la capacidad de cada aparato y el tamaño de lectura generado.

Una de las posibles aplicaciones que hemos comentado anteriormente es la secuenciación de pequeñas regiones del genoma (*targeted resequencing*), y para ello una de las estrategias más utilizadas es la combinación de la secuenciación masiva asociada a la captura de secuencias. El objetivo es que parte del genoma sea enriquecida usando diferentes estrategias, bien por amplificación directa de pequeños o grandes amplicones o la captura de esas regiones de interés mediante sondas marcadas. Después de este enriquecimiento de ciertas zonas, procederemos a su secuenciación.

Hay múltiples estudios que han utilizado en los últimos años esta estrategia, sobre todo para el estudio de enfermedades genéticamente heterogéneas, en el que un gran número de genes están implicados. Dos claros ejemplos son el de Neveling *et al*. donde se

analizan 52 genes implicados en RP (NEVELING *et al.* 2012) o los de Browstein *et al.* y De Keulenaer *et al*. para el estudio de hipoacusias hereditarias (BROWSTEIN *et al.* 2011) y (DE KEULENAER *et al.* 2012).

Creemos que este enfoque puede resultar útil para el estudio genético del síndrome de Usher, clínica y genéticamente heterogéneo. Licastro *et al*. utilizaron dos estrategias diferentes, amplificación de largos amplicones y captura de todo el exoma, seguidas de secuenciación masiva, para el estudio de los genes Usher (LICASTRO *et al.* 2012). Sin embargo, mediante el estudio de todo el exoma la cobertura no fue ni homogénea, ni suficiente, sugiriendo que no es la mejor estrategia para el estudio molecular del síndrome de Usher en el marco de una aplicación en diagnóstico. Por eso, éste es uno de los objetivos de esta tesis, desarrollar e implementar una plataforma de secuenciación de nueva generación adecuada para el estudio de los genes Usher.

11. Terapias y tratamientos para el síndrome de Usher

En los últimos años se están produciendo avances prometedores en el campo de las terapias para el síndrome de Usher. Principalmente dos órganos sensoriales se ven afectados en este transtorno: el oído interno y la retina. Hasta el momento los diferentes estudios se centran en el tratamiento de cada órgano por separado.

11.1 Tratamiento para la hipoacusia

Principalmente dos estrategias (audífonos e implantes cocleares) se están siguiendo dependiendo de la gravedad de la hipoacusia.

- Audífonos: el audífono es un aparato electrónico diseñado para amplificar la señal acústica de forma que los sonidos sean audibles para el paciente y se suele usar en pacientes en los que la sordera es de leve a moderada, como son los USH2.

- Implantes cocleares: en pacientes que sufren una sordera profunda desde el nacimiento (USH1) o en aquellos en los que la sordera es progresiva y derivan en un grado de hipoacusia grave-profundo, se están utilizando con éxito los implantes cocleares. Se trata de un pequeño dispositivo electrónico que se implanta de forma quirúrgica. Generalmente constan de dos partes, una que se introduce en el hueso que rodea el oído (hueso temporal) y consta de un estimulador-receptor que recibe, descodifica y envía una señal al cerebro y, una segunda parte externa que consta de un micrófono/receptor, que recibe el sonido, lo convierte en señal eléctrica y lo envía a la parte interna del implante coclear.

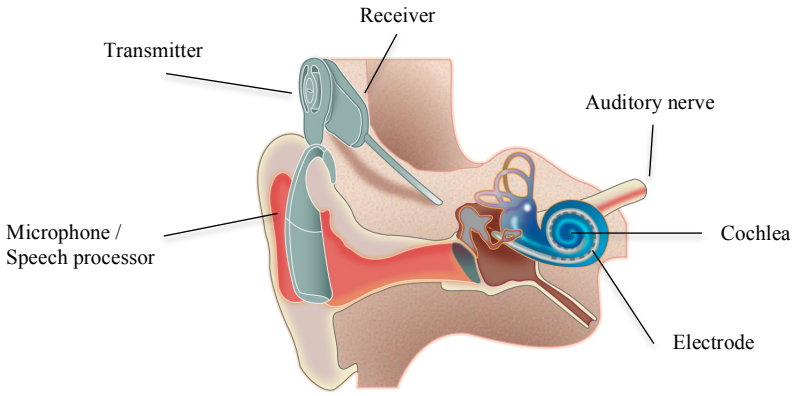


Figura 14. Esquema de un implante coclear.

- Terapia génica: hasta el momento, se han realizado pocos estudios para el uso de la terapia génica en el oído. Este tipo de trabajos siempre se han centrado más en la terapia para la RP debido a las características del ojo y también porque con los audífonos e implantes se mejora bastante la audición. Recientemente se ha publicado un estudio en el que aplicaban esta tecnología para un tipo de sordera profunda. Se trata de la deleción del gen *VGLUT3*, en la que la ausencia de la proteína VGLUT3 en las células ciliadas internas produce una sordera profunda. Mediante la introducción de una copia *wild-type* mediada por vectores AAV (vectores derivados de virus adenoasociados) en un ratón sordo deficiente para ese gen, se vio que se recuperaba de la hipoacusia (MARTIN and RAPHAEL 2012).

11.2 Perspectivas terapéuticas para la RP

La terapia ocular está favorecida por el hecho de que la retina es un órgano muy accesible y se pueden realizar inyecciones locales fácilmente, además de por su privilegiado sistema inmune y las propias barreras internas del ojo que limitan la difusión sistémica de los vectores virales que puedan ser utilizados en este tipo de estrategias. Todo esto hace que se hayan desarrollado más ensayos terapéuticos que para el oído.

- Implantes artificiales: están basados en la estimulación eléctrica de la red neuronal de la retina que todavía permanece funcional, mediante la implantación de microchips o electrodos. En 2011, se publicó un estudio en el que los autores desarrollaron un *array* de micro-electrodos subretinianos que permite a los pacientes reconocer y localizar objetos. Sin embargo, aún hay que mejorar la resolución espacial de estos implantes y sobretodo su biocompatibilidad y estabilidad a largo plazo (ZRENNER *et al.* 2011).

- Optogenética: es una alternativa a los implantes eléctricos que acabamos de ver. Se parecen en el sentido de que la luz es capturada por fotosensores artificiales. Sin embargo, la optogenética se trata de la expresión ectópica de moléculas que permitan cambiar el potencial de membrana a las células. La inyección de vectores AAV con la información necesaria para la expresión de determinados canales o bombas sensibles a la luz en determinadas células que todavía permanecen en esa red de la retina, va a permitir restaurar la fotosensibilidad. Busskamp *et al*., tuvieron éxito usando esta estrategia para restaurar la sensibilidad a la luz en modelos de ratón con RP y reactivar la sensibilidad a la luz en fotorreceptores de retinas humanas *ex-vivo* (BUSSKAMP *et al.* 2012). Por el momento solo se trata de ensayos preclínicos.

- Terapia génica: la terapia génica es una de las estrategias más simples y atractivas conceptualmente. Sin embargo tiene algunos problemas como que el tipo celular en el que se exprese tiene que estar vivo o que el tamaño del gen a incorporar exceda la capacidad de los vectores disponibles para el uso en humano. Este tipo de terapia está en progreso para pacientes con Amaurosis congénita de Leber (LCA), producida por defectos en el gen *RPE65* (retinal pigment epithelium 65), mediante la inyección subretiniana de un

vector AAV con el gen funcional (BARKER *et al.* 2009; SIMONELLI *et al.* 2010). Tras el éxito de los ensayos con ratones y perros, hay tres ensayos clínicos en progreso y parece detectarse una restauración de la visión. Esto da confianza para el desarrollo de terapias génicas en otras distrofias retinianas, demostrando la seguridad de los vectores AAV. En 2011, se publicó un estudio en el que se reemplazaba la proteína whirlina mediante el uso de vectores AAV en un ratón deficiente para esta proteína y ha sido capaz de restaurar el complejo proteico USH2 en la región periciliar de los fotorreceptores (ZOU *et al.* 2011). Una estrategia similar utilizó (HASHIMOTO *et al.* 2007) para el reemplazamiento de otro gen implicado en Usher, *MYO7A* en ratones deficientes para dicho gen. En este caso se basó en el uso de lentivirus debido al gran tamaño del gen. Recientemente, se han iniciado los ensayos clínicos de fase I/II de terapia génica para pacientes con USH1B. Este tratamiento, conocido como UshStat®, permitirá el transporte de copias correctas del gen *MYO7A* a células de la retina afectadas mediante el uso de la tecnología Oxford's LentiVector®, que permite el transporte de genes grandes.

- TRIDs (translational read-through-inducing drugs): es un tipo de terapia basada en solucionar el problema que se presenta cuando hay una mutación que genera un codón de parada prematuro con la subsecuente producción de una proteína truncada. La clave es inyectar una serie de moléculas que hacen que durante la traducción la maquinaria reconozca el codón de parada de la mutación como un triplete que codifica para un amino ácido, resultando en una proteína entera a partir de un ARN mensajero mutante. Estudios de este tipo se han realizado para el gen *USH1C* y la mutación p.R31X, probando dos tipos de moléculas: antibióticos aminoglucósidos (NB54) o aminoglucósidos sintéticos como el PTC124 con buena biocompatibilidad (GOLDMANN *et al.* 2012).

- AONs (oligonucleótidos antisentido): la presencia de mutaciones puntuales en regiones intrónicas profundas, que producen la activación de un sitio críptico de *splicing*, produciendo un ARN mensajero aberrante había sido previamente descrita en otras enfermedades (VETRINI *et al.* 2006; RINCON *et al.* 2007). Sin embargo, fue en el 2012 cuando Vaché *et al*. publicaron la primera mutación de este tipo en el síndrome de Usher (VACHE *et al.* 2012). Esto abre las puertas al uso de oligonucleótidos antisentido diseñados contra la mutación, que evitarían la inclusión de ese pseudoexón mediante el bloqueo de los sitios de reconocimiento de *splicing* (Figura 15). De esta forma se podría recuperar el *splicing* normal y poder obtener una proteína funcional, como ya ha sido demostrado en otras patologías (PEREZ *et al.* 2010). Y además, este mismo año se ha publicado un estudio en el que consiguen rescatar la hipoacusia y disfunción vestibular en un modelo de ratón con la mutación c.216G>A en el gen *USH1C,* demostrando la eficacia de los AONs para corregir *splicings* aberrantes (LENTZ *et al.* 2013).

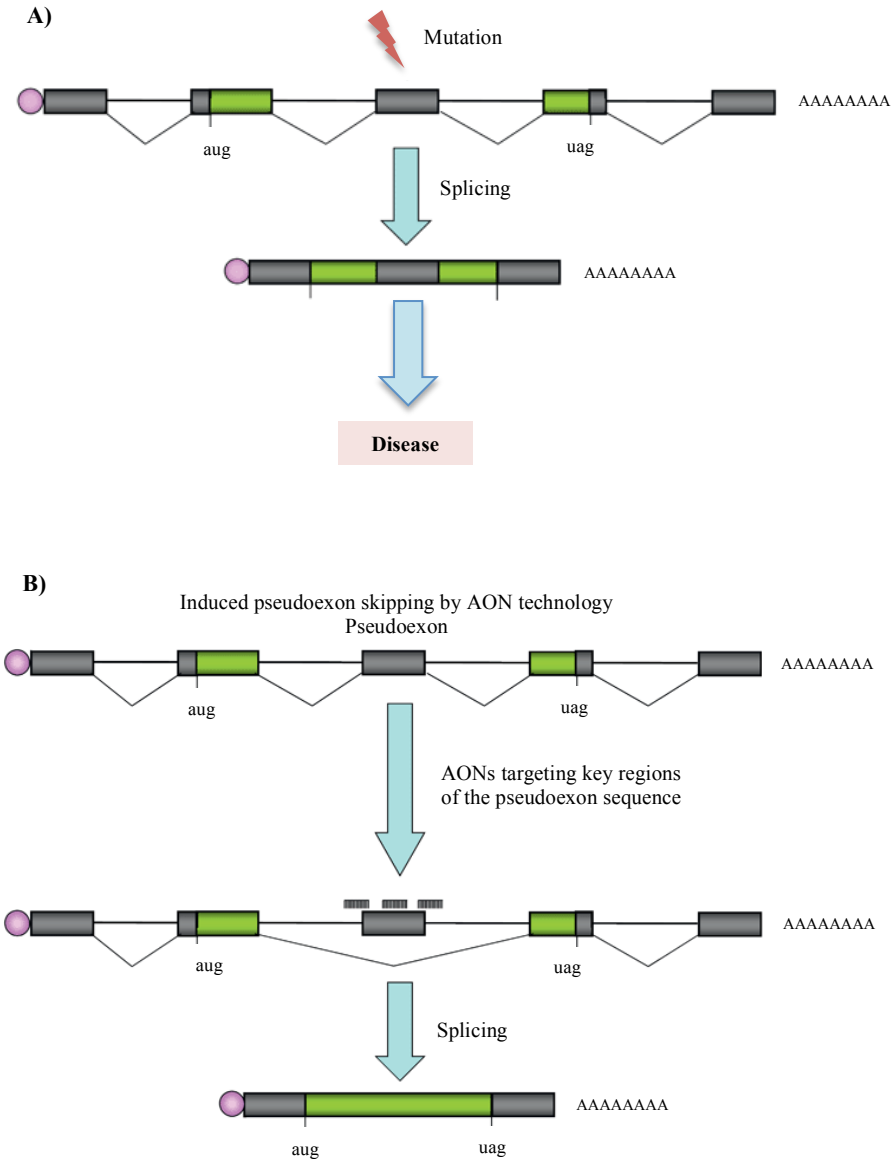


Figura 15. Representación del uso de AONs. A) En la imagen se observa la inclusión de un pseudoexón debido a la presencia de una mutación en un intrón. B) Mediante el uso de AONs se consigue la exclusión del pseudoexón. En verde aparecen los exones y en gris los intrones.

- Células encapsuladas: estas capsulas permiten la administración de drogas en el ojo. Uno de los factores que se ha administrado de esta forma es el CNTF (*ciliary neurotrophic factor*) y ha mostrado proteger a la retina de la degeneración en 13 modelos animales distintos (TAO 2006). En 2006 ya se empezaron los primeros ensayos clínicos en pacientes con RP (SIEVING *et al.* 2006) y el transporte de este factor mediante células encapsuladas también se ha probado en pacientes con degeneración macular asociada a la edad (ZHANG *et al.* 2011), observándose un enlentecimiento de la pérdida visual. Un año más tarde, Kauper *et al*., presentaron un estudio en el que analizaban la farmacocinética del CNTF inyectado intraocularmente mediante células encapsuladas en pacientes con RP obteniendo resultados positivos (KAUPER *et al.* 2012).

- Trasplante de células progenitoras: una de las estrategias más recientes es la diferenciación de células madre de adultos o de embrión y su posterior trasplante en la retina. Lu *et al*., trasplantaron células progenitoras derivadas del cerebro en el espacio subretinal en un animal modelo para *USH2A (LU et al. 2010)*. Recientemente (PEARSON *et al.* 2012) demostraron que células precursoras de fotorreceptores obtenidas de retinas en un determinado estado de desarrollo pueden ser trasplantadas y, esas células son capaces de generar sinapsis funcionales y contribuir a la función visual.

CHAPTER II

Hypothesis and objectives

1. Hypothesis

1.

a) Three genes have been identified for Usher syndrome type 2: *USH2A*, *GPR98* and *DFNB31*. *USH2A* is the most frequent causative gene in most populations. *GPR98* and *DFNB31* are responsible for a minor number of cases. The screening of the three USH2 genes will allow us to know the mutational spectrum present in our cohort, to determine the involvement of each gene in the Spanish population and to establish comparisons with other studies.

b) *CLRN1* is the only gene identified to be responsible for Usher syndrome type 3 and also responsible for cases diagnosed of USH1 and USH2. The screening of this gene in patients with USH2 and USH3 will identify pathological mutations.

2.

a) Usher syndrome is a heterogeneous disease. In the last years, NGS has emerged as a powerful tool for identifying variants in this type of disorders. The development of a NGS platform is an appropriate strategy for the study of all the Usher genes in a cost-effective way.

2. Objectives

1.

a) To determine the involvement of each USH2 gene (*USH2A*, *GPR98* and *DFNB31)* in a cohort of Spanish Usher syndrome patients.

b) To determine the role of *CLRN1* mutations among clinically USH2 and USH3 patients.

2.

a) To implement a NGS platform for the molecular genetic diagnosis of Usher syndrome.

CHAPTER III

Mutational screening of the USH2 and USH3 genes

1. Context

Among the three clinical types of Usher syndrome, USH2 is the most common form, accounting for approximately two thirds of all Usher syndrome patients in the Spanish population. As stated in the introduction, several studies have been published about the involvement of Usher genes in different populations. To date, an exhaustive molecular analysis of the three causative USH2 genes has not been performed in our country. In this chapter we will present the results of mutational screenings of the *USH2A*, *GPR98* and *DFNB31* genes. The study was not limited to the search of point mutations; CGH-array analysis, computational predictions and functional studies were carried out when necessary.

Initially, a total of 88 patients with Usher syndrome were enrolled in the study. The work started with the molecular analysis of the *USH2A* gene, since it has been reported to be the main causative gene for USH2 patients. Subsequently, the two other USH2 genes, *GPR98* and *DFNB31,* were analysed when patients showed to be negative for *USH2A* and whenever DNA from these patients was available.

Vaché *et al*. reported the first deep intronic mutation causing activation of a pseudoexon in *USH2A* involved in Usher syndrome (VACHE *et al.* 2012). We collaborated in this work, in which we studied the involvement of the new intronic deleterious variant in the Spanish USH patients.

USH3 is thought to be a minor type of USH in our population. Mutations in *CLRN1* were initially responsible for USH3, but mutations in this gene have also been reported in patients with clinical features similar to USH1 or USH2. We report here the results of the molecular analysis of *CLRN1* in a cohort of 19 USH patients.

2. Article:

Mutational screening of the *USH2A* gene in Spanish USH patients reveals 23 novel pathogenic mutations

Gema Garcia-Garcia[1], Maria J Aparisi[1], Teresa Jaijo[1,2], Regina Rodrigo[1], Ana M Leon[2], Almudena Avila-Fernandez[2,3], Fiona Blanco-Kelly[3], Sara Bernal[2,4], Rafael Navarro[5], Manuel Diaz-Llopis[6], Montserrat Baiget[2,4], Carmen Ayuso[2,3], Jose M Millan[1,2,7], Elena Aller[1,2]

1. Grupo de Investigación en Enfermedades Neurosensoriales. Instituto de Investigación Sanitaria IIS-La Fe, Valencia, Spain.
2. CIBER de Enfermedades Raras (CIBERER), Valencia, Spain.
3. Servicio de Genética, Fundación Jiménez Díaz, Madrid, Spain.
4. Servei de Genètica, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain.
5. Instituto de Microcirugía Ocular, Barcelona, Spain.
6. Servicio de Oftalmología, Hospital Universitario La Fe, Valencia, Spain.
7. Unidad de Genética y Diagnóstico Prenatal, Hospital Universitario La Fe, Valencia, Spain.

Orphanet J Rare Dis. 2011 Oct 17;6:65.

ABSTRACT

Background: Usher Syndrome type II (USH2) is an autosomal recessive disorder, characterized by moderate to severe hearing impairment and retinitis pigmentosa (RP). Among the three genes implicated, mutations in the *USH2A* gene account for 74-90% of the USH2 cases.
Methods: To identify the genetic cause of the disease and determine the frequency of *USH2A* mutations in a cohort of 88 unrelated USH Spanish patients, we carried out a mutation screening of the 72 coding exons of this gene by direct sequencing. Moreover, we performed functional minigene studies for those changes that were predicted to affect splicing.
Results: As a result, a total of 144 DNA sequence variants were identified. Based upon previous studies, allele frequencies, segregation analysis, bioinformatics' predictions and *in vitro* experiments, 37 variants (23 of them novel) were classified as pathogenic mutations.
Conclusions: This report provide a wide spectrum of *USH2A* mutations and clinical features, including atypical Usher syndrome phenotypes resembling Usher syndrome type I. Considering only the patients clearly diagnosed with Usher syndrome type II, and results obtained in this and previous studies, we can state that mutations in *USH2A* are responsible for 76.1% of USH2 disease in patients of Spanish origin.

Keywords: Usher Syndrome, *USH2A*, Mutations, Sequence Variants.

Disease name and definition:

Usher syndrome: Usher syndrome (USH) is an autosomal recessive disease characterized by hearing loss, retinitis pigmentosa (RP), and, in some cases, vestibular dysfunction. It is clinically and genetically heterogeneous and is the most common cause underlying deafness and blindness of genetic origin. Clinically, USH is divided into three types. Usher type I (USH1) is the most severe form and is characterized by severe to profound congenital deafness, vestibular areflexia, and prepubertal onset of progressive RP. Type II (USH2) displays moderate to severe hearing loss, absence of vestibular dysfunction, and later onset of retinal degeneration. Type III (USH3) shows progressive postlingual hearing loss, variable onset of RP, and variable vestibular response. To date, five USH1 genes have been identified: *MYO7A* (USH1B), *CDH23* (USH1D), *PCDH15* (USH1F), *USH1C* (USH1C), and *USH1G* (USH1G). Three genes are involved in USH2, namely, *USH2A* (USH2A), *GPR98* (USH2C), and *DFNB31* (USH2D). USH3 is rare except in certain populations, and the gene responsible for this type is *USH3A*.

BACKGROUND

Usher syndrome (USH) is an autosomal recessive disease characterized by the association of hearing loss and visual impairment due to retinitis pigmentosa (RP), with or without vestibular dysfunction. It is the most frequent cause of concurrent deafness and blindness of genetic origin and its general prevalence ranges from 3.3 to 6.4 per 100.000 live births [1]. In Spain, the estimation is 4.2/100.000 [2].

USH is clinically and genetically heterogeneous. Three clinical forms are distinguished: USH1, USH2 and USH3 and nine genes have been identified responsible so far. Five causative genes have been reported for USH1: *MYO7A*, *USH1C*, *CDH23*, *PCDH15* and *USH1G*. Three genes for USH2: *USH2A*, *GPR98* and *DFNB31*. Meanwhile only one gene has been described for USH3: *USH3A* [3,4].

USH2 appears to be the most common clinical form of the disorder, accounting for more than 50% of all Usher cases [5,6]. Among the three genes described for USH2, *USH2A* is the most commonly mutated gene. It is responsible for approximately 74-90% of USH2 cases [2,7]. Mutations in *USH2A,* are also responsible for atypical Usher syndrome and recessive non-syndromic RP [8,9]. The *USH2A* gene, located on chromosome 1q41 [10], was initially described as comprising 21 exons, encoding a protein of 1546 amino acids [11,12]. However, in 2004, van Wijk et al. (2004) identified 51 additional exons at the 3´ end of *USH2A* [13]. The longest transcript consists of 72 exons, encoding a protein of 5202 amino acids. In addition, Adato et al. (2005), identified an alternative spliced exon 71 in mouse transcripts, expressed in the inner ear and well conserved in vertebrates [14]. The long *isoform b* is characterized by containing a transmembrane region, followed by an intracellular domain with a PDZ-binding motif, which interacts with the PDZ domain of harmonin and whirlin, integrating USH2A into the USH protein network [14,15].

Initially, studies of the *USH2A* gene covered just exons 2-21 [11,16-18,8,12,19,20,9,21,22], but only 45-63% of the expected *USH2A* mutations were identified. Nevertheless, since the discovery of the long isoform, several mutational analyses of all 72 exons have been carried out in diverse populations [23-30]. As a result, many novel pathogenic mutations have been identified, including splicing mutations at non-canonical positions of splice sites [31]. However, the majority of these changes only appear in a few cases, with the exception of the common ancestral p.Glu767fs mutation, located at exon 13, which is the most prevalent *USH2A* mutation in several populations [16].

In the present study we have performed an exhaustive mutational screening of the long *isoform b* of *USH2A* to identify new patients with mutations in this gene, and to detect the second mutation in patients with one previously detected *USH2A* mutation. Some cases had previously been studied for exon 13 or for the 21 first *USH2A* exons [19,9], or analyzed using the genotyping microarray for Usher syndrome (Asper Biotech, Tartu, Estonia; [32]). Furthermore, we have used *in silico* and *in vitro* analysis to evaluate the functional consequences on gene expression and protein function of several nucleotide changes.

MATERIAL AND METHODS

Subjects

Eighty-eight (88) unrelated Spanish patients diagnosed of Usher syndrome were included in this study. They were recruited from the *Federación de Afectados de Retinosis Pigmentaria de España* (FARPE) and also from the Ophthalmology and ENT Services of several Spanish Hospitals as part of a large-scale study on the genetics of Usher syndrome in the Spanish population.

On the basis of their clinical history and ophthalmologic, audiological, neurophysiological and vestibular tests, 58 of these families were clinically classified as USH2 while 11 displayed atypical Usher syndrome. Detailed clinical data could not be obtained for 19 patients and these remained as non classified (USHNC).

Previously, 40 of these 88 patients were studied for exon 13, while 24 were analysed for the first 21 exons of *USH2A* and 42 were analyzed with the genotyping microarray for Usher syndrome (Asper Biotech, Tartu, Estonia). At that time, the version of the array detected 429 previously described mutations in eight of the nine genes reported for the disease. As a result of these previous analyses, eighteen of them were found to carry one mutated allele, but the second mutation could not be detected. These mutations have been included in the statistical summaries presented herein. These patients were subjected to mutation screening of the exons that had not been analyzed. In the remaining patients, we carried out the study of exons 2-72 (including the alternatively spliced exon 71).

When DNA samples from patients' relatives were available, we carried out a segregation analysis. One hundred unrelated individuals of Spanish origin without hearing

loss or RP family history were screened as controls to evaluate the frequency of the mutations found in the patient sample.

Mutation analysis

Genomic DNA from patients and controls was extracted from peripheral blood samples following standard protocols. The coding exons and flanking intronic sequences of *USH2A* were amplified by PCR using primers and conditions described by Aller et al. (2004; 2006) [9,23]. The amplified DNA fragments were analysed by direct sequencing using the Big Dye Terminator v.3.1 kit *(*Applied Byosistems, Carlsbad, CA*)*, and purified sequencing reactions were analysed in an ABI PRISM 3730 DNA analyzer *(*Applied Byosistems, Carlsbad, CA). The obtained sequences were compared with the consensus sequence NM_206933.2. The +1 position corresponds to A in the ATG translation initiation codon.

Predictions of the pathogenic effect of missense variations

To predict whether a rare missense variant is deleterious, we used the combined results of three different computer algorithms:

- Sort Intolerant From Tolerant (SIFT) (available at http://sift.jcvi.org) uses sequence homology to predict whether a change is tolerated or deleterious.
- The polymorphism phenotyping program, PolyPhen (available at http://genetics.bwh.harvard.edu/pph/) uses sequence conservation, structure and SWISS-PROT annotation to characterize an amino acid substitution as benign, possibly deleterious or probably deleterious.
- Pmut (available at http://mmb2.pcb.ub.es:8080/PMut/) provides prediction by neural networks, which use internal databases, secondary structure prediction and sequence conservation. This program provides a binary prediction of “neutral” or “pathologic”.

Splice-site Prediction programs

Intronic, isocoding and missense changes were analyzed using the programs NNSPLICE (http://fruitfly.org:9005/seq_tools/splice.html), Human Splicing Finder (HSF) version 2.4 (http://www.umd.be/HSF/) and NetGene2 (http://www.cbs.dtu.dk/services/NetGene2/) in order to predict whether those changes could be affecting, creating or eliminating donor/acceptor splice sites.

Minigene constructions and expression

Minigene constructs were generated, using the exon trapping expression vector pSPL3. For each mutation, the exon and intronic flanking sequences were amplified from the patient’s DNA, using the High Fidelity Phusion polymerase (Finnzymes, Espoo, Finland). Amplicons were inserted between the XhoI/NheI and XhoI/BamHI restriction sites for the variants p.E2496E and p.V382M, respectively, using T4 DNA ligase (Invitrogen Corporation, Carlsbad, CA). The p.V382M mutation was generated by site-directed mutagenesis. All vectors were confirmed by direct sequencing. The minigene constructs were transfected into COS-7 cells as described before [33]. RNA extraction and

RT-PCR analysis was perfomed as previously described [34,31]. Missplicing percentages were measured using the Alpha Imager 2200 (version 3.1.2) software (AlphaInnotech Corporation, San Francisco, CA, USA).

RESULTS

The molecular analysis of the *USH2A* gene in 88 unrelated USH Spanish patients revealed 37 different pathogenic mutations. Among these, a total of 23 mutations were novel (See Tables 1, 2 and 3). At least one pathogenic mutation was found in 43 out of 88 unrelated patients (48.9%). Thirty-three patients were classified as USH2, five as USHA (atypical Usher syndrome) and five as USHNC (Usher syndrome non classified). In 25 out of these 43 cases the two causative mutations were detected (58.1%), five patients were homozygous and 20 compound heterozygous. Detailed clinical manifestations of these 25 patients and 3 additional patients with one pathogenic and one probably pathogenic mutation (UV3; likely to be pathogenic but cannot formally be proven) are summarized in Table 4.

In this study, a total of 144 variants were detected: 25 were truncating mutations and five were splice-site mutations (located at the conserved AG/GT dinucleotides of the splice site). The pathogenic effect of these variants is clear. But, in addition, 48 missense, 20 silent and 46 intronic variants were identified. According to previous studies, allele frequencies, segregation analysis, bioinformatics' predictions and *in vitro* experiments, the missense, silent and intronic changes were classified into 4 different categories: pathogenic, possibly pathogenic (UV3), possibly non-pathogenic (possibly neutral, UV2) and non-pathogenic (neutral). (See Tables 1, 2, 3 and 5).

Table 1. *USH2A* truncating and *splice-site* mutations

Nucleotide change	Exon	Predicted effect	Predicted pathology	No. of alleles	References
Nonsense mutations					
c.820C>T	5	**p.R274X**	+	2	Present study
c.1518T>A	8	p.Y506X	+	1	Bernal *et al*., 2005
c.3883C>T	18	p.R1295X	+	1	Dreyer *et al*., 2000
c.4474G>T	21	p.E1492X	+	2	Bernal *et al*., 2005
c.4645C>T	22	p.R1549X	+	1	Baux *et al*., 2007
c.7854G>C	41	**p.W2618X**	+	1	Present study
c.9753T>A	50	**p.C3251X**	+	2	Present study
c.10102C>T	51	**p.Q3368X**[a]	+	1	Present study

Table 1 *USH2A* truncating and *splice-site* mutations (continued)

Nucleotide change	Exon	Predicted effect	Predicted pathology	No. of alleles	References
c.10759C>T	55	p.Q3587X	+	2	Baux *et al*., 2007
c.11146C>T	57	**p.Q3716X**	+	1	Present study
c.14175G>A	65	**P.W4725X**	+	1	Present study
Deletions and insertions					
c.918_919insGCTG	6	**p.S307AfsX17**	+	1	Present study
c.1214delA	7	p.N405IfsX3	+	5	Bernal *et al*., 2005
c.1629_1645del	10	**p.F543LfsX2**	+	1	Present study
c.2299delG	13	p.E767SfsX21	+	8	Eudy *et al.,* 1998
c.5278delG	26	**p.D1760MfsX10**[a]	+	1	Present study
c.5540_5541dup	27	**p.N1848X**	+	1	Present study
c.6319_6324delinsTAAA	32	**p.V2107X**	+	1	Present study
c.8890dupT	45	**p.W2964LfsX89**	+	1	Present study
c.8954delG	45	**p.G2985AfsX3**	+	1	Present study
c.9261delT	47	**p.E3088KfsX9**	+	1	Present study
c.10272_10273dupTT	52	p.C3425FfsX4	+	1	Aller *et al.,* 2006
c.11566delA	60	**p.S3856VfsX28**	+	1	Present study
c.12093delC	62	**p.Y4031X**	+	1	Present study
c.13140delA	63	**p.V4381CfsX10**	+	1	Present study
***Splice-site* mutations**					
c.1328+1G>T	IVS7	Ex7 splice defect	+	1	Present study
c.1841-2A>G	IVS10	Ex11 splice defect	+	2	Bernal *et al*., 2003
c.11548+2T>G	IVS59	Ex59 splice defect	+	1	Present study
c.12067-2A>G	IVS61	Ex62 splice defect	+	4	Auslender *et al.,* 2008
c.15053-1G>A	IVS69	Ex70 splice defect	+	1	Present study

+: Denotes pathogenic mutations. **No. of alleles:** Number of alleles identified in patients. [a]These two mutations are allelic. Novel pathogenic mutations described in this study are in bold.

Missense variants

Fourty-eight missense variants were identified (See Table 2). Twenty-nine were considered as non-pathogenic because all of them were already described as non-pathogenic in other studies [https://grenada.lumc.nl/LOVD2/Usher_montpellier, 35].

Nine nucleotide changes were classified as possibly non-pathologic (UV2). p.G713R, p.S841Y, p.S2196T, p.S2639P, p.G4692R and p.K5026E have already been reported in other works and categorized as possibly non-deleterious or of unknown pathogenecity [https://grenada.lumc.nl/LOVD2/Usher_montpellier, 35]. Meanwhile, p.N2377S, p.N2394K and p.E4921K have not been described previously, so they were analyzed with the three sequence analysis programs (SIFT, Polyphen and PMUT). None of those changes was predicted to be clearly deleterious (See Table 6).

Table 2 Missense changes in *USH2A*

Nucleotide change	Exon	Amino acid change	Predicted pathology	No. of alleles	Ref.
c.373G>A	2	p.A125T	-	81	Dreyer *et al*., 2000
c.130G>A	2	**p.G44R**	+	1	Present Study
c.688G>A	4	p.V230M	-	3	Dreyer *et al*., 2000
c.908G>A	6	p.R303H[e]	UV3	1	Yan *et al*., 2009
c.1144G>A	7	**p.V382M[b]**	UV3	1	Present Study
c.1434G>C	8	p.E478D	-	4	Seyedahmadi *et al*., 2004
c.1663C>G	10	p.L555V	-	2	Bernal *et al*., 2003
c.1931A>T	11	p.D644V	-	6	Weston *et al*., 2000
c.2137G>C	12	P.G713R	UV2	2	Dreyer *et al*., 2000
c.2276G>T	13	p.C759F	+	3	Dreyer *et al*., 2000
c.2522C>A	13	p.S841Y	UV2	1	Jaijo *et al*., 2009
c.4457G>A	21	p.R1486K	-	75	Dreyer *et al*., 2000
c.4714C>T	22	p.L1572F	-	7	Dreyer *et al*., 2008
c.4994T>C	25	p.I1665T	-	31	Kaiserman *et al*., 2007
c.5975A>G	30	p.Y1992C[e]	UV3	1	McGee et al., 2010
c.6317T>C	32	p.I2106T	-	117	Aller *et al*., 2006
c.6506T>C	34	p.I2169T	-	85	Aller *et al*., 2006
c.6587G>C	34	p.S2196T	UV2	2	Jaijo *et al*., 2009
c.6713A>C	35	p.E2238A	-	1	Aller *et al*., 2006
c.6875G>A	36	p.R2292H	-	2	Dreyer *et al*., 2008

Table 2 Missense changes in *USH2A* (continued)

Nucleotide change	Exon	Amino acid change	Predicted pathology	No. of alleles	Ref.
c.7130A>G	38	p.N2377S	UV2	1	Present Study
c.7182C>A	38	p.N2394K	UV2	1	Present Study
c.7506G>A	40	p.P2502P	-	11	Baux *et al* ., 2008
c.7685T>C	41	p.V2562A	-	1	Dreyer *et al*., 2008
c.7915T>C	41	p.S2639P	UV2	3	McGee et al., 2010
c.8624G>A	43	p.R2875Q	-	9	Aller *et al*., 2006
c.8656C>T	43	p.L2886F	-	9	Aller *et al*., 2006
c.9262G>A	47	p.E3088K	-	1	Dreyer *et al*., 2008
c.9296A>G	47	p.N3099S	-	12	Aller *et al*., 2006
c.9343A>G	47	p.T3115A	-	9	Dreyer *et al*., 2008
c.9430G>A	48	p.D3144N	-	8	Aller *et al*., 2006
c.9595A>G	49	p.N3199D	-	9	Baux *et al*., 2007
c.9799T>C	50	p.C3267R	+	5	Aller *et al*., 2006
c.10073G>A	51	p.C3358Y	+	1	McGee *et al*., 2010
c.10232A>C	52	p.E3411A	-	94	Aller *et al*., 2006
c.10636G>A	54	**p.G3546R**	+	4	Present Study
c.11504C>T	59	p.T3835I	-	30	Present Study
c.11602A>G	60	p.M3868V	-	34	Aller *et al*., 2006
c.11677C>A	60	p.P3893T	-	2	Dreyer *et al*., 2008
c.11680A>G	60	**p.N3894D**	UV3	1	Present Study
c.12343C>T	63	p.R4115C	-	2	van Wijk *et al*., 2004
c.14074G>A	64	p.G4692R[c]	UV2	1	McGee *et al*., 2010
c.14453C>T	66	p.P4818L	+	1	Aller *et al*., 2006
c.14513G>A	66	p.G4838E	-	1	McGee *et al*., 2010
c.14543G>A	66	p.R4848Q	-	1	McGee *et al*., 2010
c.14761G>A	67	p.E4921K	UV2	1	Present Study
c.15076A>G	70	p.K5026E	UV2	1	McGee *et al*., 2010
c.15091C>T	70	p.R5031W	-	2	Dreyer *et al*., 2008

+: Denotes pathogenic mutations; **UV3**: Probably pathogenic mutations; **UV2**: probably non-pathogenic mutations; **-**: neutral variants. **No. of alleles:** Number of alleles identified in

patients.[b]These variant may alter normal splicing. [c]Patient with this change also has two other clearly pathogenic mutations in *USH2A*. [e] p.R303H and p.Y1992C were initially described as being pathogenic mutations by Yan *et al*., 2009 (ref. 29) and McGee et al., 2010 (ref. 30) respectively, but we have classified them as UV3 in accordance with the specific locus database for Usher syndrome: https://grenada.lumc.nl/LOVD2/Usher_montpellier. Novel pathogenic mutations and novel probably pathogenic mutations (UV3) described in this study are in bold.

Four missense variants were classified as possibly-pathogenic (UV3). The variants p.R303H and p.Y1992C were described previously [https://grenada.lumc.nl/LOVD2/Usher_montpellier, 35]. The novel change p.N3894D was not found in 200 control alleles and the segregation analysis proved that it co-segregates with the disease. However, only one program considered it as clearly pathogenic (See Table 6). The new p.V382M change, which affects the first base of exon 7, was not found in control samples and it was predicted to slightly affect splicing. The minigene assays only revealed a mild increase of the transcript excluding exon 7 (Fig. 1, band d) when the variant was present, in comparison to the wild-type sequence.

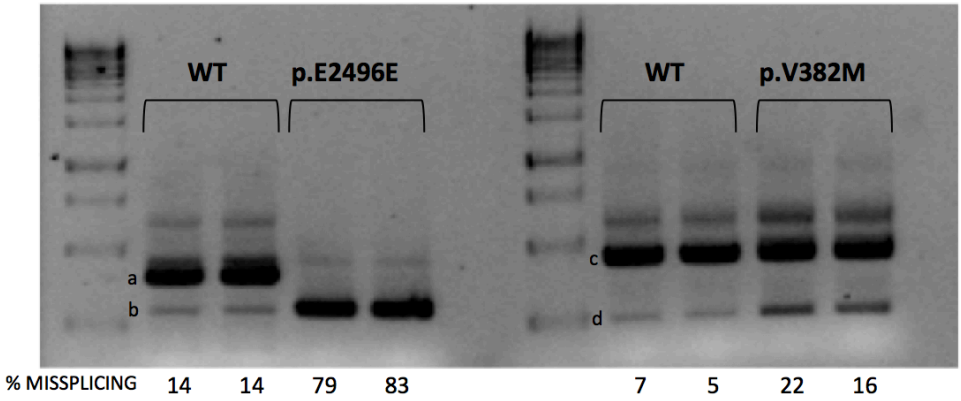


Fig. 1 In vitro splicing assays for p.E2496E and p.V382M mutations. Gel electrophoresis shows the different splicing processes for WT minigene and mutants constructions. COS-7 cells transfection experiments were performed in duplicate. Numbers at the bottom of gels indicate the proportion (%) of misspliced transcripts compared to the full-length transcript. For the p.E2496E mutation, an evident increase of band b (corresponding to the aberrant transcript that only contains 37pb of the exon 40) can be observed with regard to the WT minigene expression product. For the p.V382M variant a small increase of the exon 7 skipping in the mutant minigene expression is observed (band d).

Finally, six missense mutations were considered as pathogenic. p.C759F and p.C3267R were already described by others authors as damaging [https://grenada.lumc.nl/LOVD2/Usher_montpellier, 35]. p.C3358Y and p.P4818L were classified by McGee *et al*. (2010) [30] and the LOVD-USH Database as likely-pathogenic (UV3). However, in the present study, p.C3358Y was detected in a patient together with another nucleotide change (p.C3267R) and the segregation analysis confirmed that the

mutations were not in the same allele. p.P4818L was detected in a patient together with two other mutations that directly or indirectly cause a truncated protein (p.Q3368X and c.5278delG). The segregation analysis confirmed that the deletion and the nonsense mutation were in *cis* and the missense variant was in the other allele and cosegregated with the disease. The segregation analyses support the damaging effect of p.C3358Y and p.P4818L, so we have considered them as pathogenic. p.G44R and p.G3546R were novel; none of them was detected in 200 control alleles. The variant p.G44R was detected in a single patient and p.G3546R in three cases (one homozygous and two compound heterozygous cases); the results of the three computational analyses classified them as pathogenic (See Table 6).

Table 6 Results from the three different analysis programs used to predict the pathogenicity of novel missense changes.

	SIFT	**PolyPhen**	**PMUT**
p.G44R	Affect (Score 0.01)	Probably damaging	Pathogenic (NN output 0.5113)
p.N2377S	Tolerated (Score 0.44)	Possibly damaging	Neutral (NN output 0.2372)
p.N2394K	Tolerated (Score 0.23)	Possibly damaging	Neutral (NN output 0.5113)
p.G3546R	Affect (Score 0.01)	Probably damaging	Pathogenic (NN output 0.1799)
p.N3894D	Tolerated (Score 0.05)	Probably damaging	Neutral (NN output 0.0522)
p.E4921K	Tolerated (Score 0.85)	Benign	Neutral (NN output 0.3130)

SIFT: SIFT Score ranges from 0 to 1. The amino acid substitution is predicted to be damaging if the score is < 0.05, and tolerated if the score is 0/> 0.05.

PolyPhen: "Probably damaging" (it is believed most likely to affect protein function or structure), "Possibly damaging" (it is believed to affect protein function or structure), "Benign" (most likely lacking any phenotypic effect).

PMUT: NN output is the original output from the neural network for this mutation and its parameters. If that output is bigger than 0.5 it is predicted as pathogenic, otherwise as neutral.

Silent variants

We also identified 20 silent variants, 17 were previously described as neutral [36, https://grenada.lumc.nl/LOVD2/Usher_montpellier, 35] and three were novel (See Table 3). Only the variant p.E2496E was categorized as pathogenic. It was not found in 200 control alleles and the segregation analysis confirmed that this mutation co-segregates with the disease. We detected this variant in *trans* in a patient who also had a premature stop codon. According to *in silico* analyses, it was predicted to create a *de novo* donor splice site

(data not shown). The splicing alteration was confirmed using hybrid minigenes. The mutant construct generated a transcript that lacked the last 106 nucleotides of exon 40 (Fig. 1, band b). This loss of nucleotides creates a new open reading frame, leading to a premature stop codon five amino acids downstream.

Table 3. Silent variants in *USH2A*

Nucleotide change	Exon	Amino acid position	Predicted pathology	No. of alleles	Ref.
c.504A>G	3	p.T168T	-	75	Baux *et al*., 2008
c.1179A>G	7	p.Q393Q	-	1	Aller, 2008
c.1419C>T	8	p.T473T	-	36	Dreyer *et al*., 2000
c.2109T>C	12	p.D703D	-	7	Weston *et al*., 2000
c.2256T>C	13	p.H752H	-	3	Dreyer *et al*., 2008
c.3945T>C	18	p.N1315N	UV2	3	Present Study
c.4371G>A	20	p.S1457S	-	1	Dreyer *et al*., 2000
c.5031C>A	25	p.G1671G	-	22	Aller *et al*., 2006
c.5751C>T	28	p.Y1917Y	UV2	1	McGee *et al*., 2010
c.7488A>G	40	**p.E2496E**[b]	+	1	Present Study
c.7506G>A	40	p.P2502P	-	11	McGee *et al*., 2010
c.11736G>A	61	p.E3912E	-	2	Dreyer *et al*., 2008
c.11907A>T	61	p.P3969P	-	2	Dreyer *et al*., 2008
c.11946A>T	61	p.L3982L	-	29	Dreyer *et al*., 2008
c.12093C>T	62	p.Y4031Y	-	1	Dreyer *et al*., 2008
c.12612A>G	63	p.T4204T	-	129	Dreyer *et al.*, 2008
c.12666A>G	63	p.T4222T	-	58	Dreyer *et al*., 2008
c.13191G>A	63	p.E4397E	-	23	Dreyer *et al*., 2008
c.14481C>T	66	p.A4827A	-	1	McGee *et al*., 2010
c.14664G>A	67	p.T4888T	UV2	1	Present Study

+: Denotes pathogenic mutations; **UV2**: probably non- pathogenic mutations; **-** : neutral variants.
No. of alleles: Number of alleles identified in patients.
[b]These variant may alter normal splicing.
[c]Patient with this change also has two other clearly pathogenic mutations in *USH2A*.
Novel pathogenic mutations described in this study are in bold.

Intronic variants

Fourty-six intronic variants located at non-canonical positions of splice sites, of which 20 are novel, were detected in the *USH2A* gene sequence. According to computational analysis, most of these novel variants were classified as possibly non-pathologic (UV2). (See Table 4)

Table 4. Novel intronic variants

Nucleotide change	Intron	Predicted pathology	No. of alleles
c.1328+52T>C	IVS7	-	3
c.1841-61G>A	IVS10	-	11
c.4627+32G>T	IVS21	UV2	1
c.6485+18C>T	IVS33	UV2	1
c.6486-54T>C	IVS33	UV2	2
c.6486-43T>A	IVS33	UV2	1
c.6657+29C>A	IVS34	UV2	1
c.8681+18A>G	IVS43	UV2	1
c.8681+53T>G	IVS43	UV2	1
c.8681+118A>G	IVS43	UV2	1
c.9056-52G>T	IVS45	UV	1
c.9372-50A>G	IVS47	UV2	1
c.9740-59G>A	IVS49	UV2	1
c.9958+128A>G	IVS50	-	4
c.10388-123T>C	IVS52	UV2	1
c.13812-78A>G	IVS63	UV2	1
c.14134-53T>C	IVS64	UV2	1
c.14343+36G>C	IVS65	UV2	1
c.15298-35T>A	IVS70	UV2	1
c.15298-1153G>A(g.798209G>A) [d]	IVS71	UV2	1

UV2: probably non- pathogenic mutations; - : neutral variants. No. of alleles: Number of alleles identified in patients.

[d]The intronic variant g.798209G>A affects the last nucleotide of the cochlea specific exon 71. g.DNA numbering starts at nucleotide position 1 in Human Refseq: NG_009497.1 GI:222352133, which represents the minus (-) strand of *USH2A*.

DISCUSSION

In the present study, we have performed a wide mutational screening of the *USH2A* gene in 88 unrelated Spanish patients diagnosed with Usher syndrome. This analysis has led us to identify a total of 37 different pathogenic mutations, 23 of which had not been previously described: six nonsense, eleven deletions/insertions, two missense, three splice-site mutations and one isocoding variant. At least one mutation was identified in 43 cases and the two responsible mutations were detected in 25 patients (five homozygous and 20 compound heterozygous cases).

The genotype-phenotype correlation for those patients bearing two mutations is illustrated in Table 5. Most cases presented with classical USH2 clinical features. But, interestingly, in one patient (RP-259), the sensorineural hearing loss was profound, RP started at the age of 6 years and he also had vestibular dysfunction (clinical findings typical for USH1). In another intriguing case, phenotype manifestations started at the age of 50 years (RP-1703). We cannot discard the possibility that additional changes in *USH2A* or in other USH genes, present in these patients, have some modifying effect on the phenotype [37,38].

It is complicated to predict the consequences of missense, silent and intronic changes, in order to discriminate neutral variants from those with a pathogenic effect. We have used a number of bioinformatics' tools to predict the damaging effect of these variants. However, we must bear in mind that these results are only computing predictions and additional studies are necessary to confirm the effect of those changes not clearly classified. In this sense, i*n vitro* analyses for two variants located at non canonical splice sites which were predicted to affect the splicing (p.E2496E and p.V382M) showed that p.E2496E creates a *de novo* donor splice site stronger than the wild type site that leads to the loss of the last 106 nucleotides of exon 40. Thus, we have considered it as pathogenic. On the other hand, the presence of p.V382M revealed a mild increase of the transcript excluding exon 7 (Fig. 1, band d) when the variant was present, but still, the normal transcript has a stronger expression. For this reason, this change has been classified as UV3.

The majority of mutations were found once or twice. Only the c.2299delG mutation was identified in more than 5 alleles. However, the cohort in our study is biased, because those patients in whom two mutations were detected in previous analyses (study of exon 13, exons 2-21 or microarray analyses) were not included in this work. Actually, the allele frequency of the c.2299delG mutation in the Spanish population is 15%, which is lower, in any case, than in other populations [39].

Table 4 Genotype-phenotype correlations of USH patients with both mutations found in this study

Patient	Mutations	Year of Birth	Diagnosis	Age	Sensorineural Hearing Loss	Vestibular Function	Onset of Night Blindness	Onset of Visual Field Loss	Visual Field	Visual Acuity	Eye Fundus	ERG	Cataracts
RP1310*	**c.12067-2A>G / c.12067-2A>G**		**USHNC**										
RP1274*	p.E1492X / p.E1492X	1969	USH2		Moderate-severe and stable	Normal[A]					1		
RP1633	p.E1492X / p.E1492X	1962	USH2	25	Since infancy	Normal[A]	25	25	Concentric loss	0,2 / 0,4	1	No response	BE
RP1607	p.G3546R / p.G3546R	1925	USH2	33	Mild-moderate	Normal[A]	16	25	Marked concentric loss	0,6 / 0,6	1	Moderate alteration	BE
RP1599*	c.1214delA / c.1214delA	1980	USH2		Moderate, since infancy	Normal[A]	20	30	Moderate concentric loss	0,6 / 0,7	1	Moderate alteration	BE
RP259	**c.1214delA / p.C3267R**	**1974**	**USHA**	**19**	**Profound since 6 years**	**Vestibular Dysfunction[A]**	**6**		**Concentric loss (19 years)**	**0,35 / 0,35 (19 years)**	**1**	**No response**	**No (19 years)**
RP1349	c.2299delG / c.1214delA	1954	USH2		Moderate-severe	Normal[B]	15	20	Marked concentric loss	0,4 / 0,3 (30 years)	2		
RP1493	c.2299delG / c.8890dupT	1973	USH2	30	Congenital, moderate and stable	Normal[B]	29	29	Concentric loss (at 31 years)	Normal (31 years)	1	No response (31 years)	No (31 years)
RP1632	c.2299delG / c.8954delG	1961	USH2		Since infancy	Normal[B]	23	25		<0,1BE	2	No response	BE
RP1715*	c.2299delG / c.1629_1645del	1986	USH2	22	Moderate and stable since 6 years	Normal[B]	16	20	Slight concentric loss (at 25 years)	Normal (25 years)	1	No response	No

Table 4 Genotype-phenotype correlations of USH patients with both mutations found in this study (continued)

Patient	Mutations	Year of Birth	Diagnosis	Age	Sensorineural Hearing Loss	Vestibular Function	Onset of Night Blindness	Onset of Visual Field Loss	Visual Field	Visual Acuity	Eye Fundus	ERG	Cataracts
RP1775	**c.2299delG / p.R303H[#]**	**1961**	**USHA**	**30**	**Moderate since 7 years and progressive**		**23**	**22**					
RP1618	p.C3267R / c.6319_6324delinsTAAA	1964	USH2		Severe-profound	Normal[A]	30	30	Marked concentric loss	0,1 / 0,1	2	No response	BE
RP1442*	p.C3267R / c.12093delC	1962	USH2	20	Congenital, moderate and stable	Normal[B]	15	20	Concentric loss, 5º (at 43 years)	0,5 / 0,2 (43 years)	2		LE (37 years) RE (43 years)
RP1703*	**p.C3267R / p.C3358Y**	**1936**	**USHA**	**50**	**Since 64 years**		**50**	**55**	**Reduced (67 years)**		**1**		**Yes**
RP1759	p.C3267R / p.Y1992C[#]	1945	USH2	62	Congenital, moderate and stable	Normal[B]		62				Abnormal response	Yes (62 years)
RP1625*	c.12067-2A>G / p.R274X	1976	USH2	25	Since infancy	Normal[A]	19	24	Concentric loss	0,1 / 0,1	2	No response	BE
RP1631	c.1841-2A>G / p.R274X	1976	USH2		Since infancy	Normal[A]	28	28	Concentric loss	0,7 / 0,6	1		No
RP951	c.13140delA / p.C3251X	1969	USH2	25	Congenital, severe and stable	Normal[B]	25	18	Reduced (23 years)		2	No response (30 years)	No (30 years)
RP1558	p.R1549X / c.1328+1G>T	1933	USH2	43	Severe and progressive since 20 years	Normal[B]	Before puberty		-10ºC	0,007 / 0.03	2		

Table 4 Genotype-phenotype correlations of USH patients with both mutations found in this study (contiued)

Patient	Mutations	Year of Birth	Diagnosis	Age	Sensorineural Hearing Loss	Vestibular Function	Onset of Night Blindness	Onset of Visual Field Loss	Visual Field	Visual Acuity	Eye Fundus	ERG	Cataracts
RP1539*	p.R1295X / p.N3894D[#]	1987	USH2	20	Congenital, moderate and stable	Normal[B]	20	17	Concentric loss (20 years)		2		
RP1172*	c.10272_10273dupTT / p.W2618X	1964	USH2	23	Moderate and stable since 7 years	Normal[B]	23	18	Concentric loss (38 years)			No response (39 years)	
RP1641*	**p.C759F / p.W4725X**	**1967**	**USHA**		**Moderate and progressive**	**Central vestibular pathology[A]**	**18**	**22**	**Marked concentric loss**	**0,1 / 0,2**	**2**	**No response**	**BE**
RP1667*	p.C759F / c.11548+2T>G	1954	USH2	15	Mild and slightly progressive	Normal[B]	15	8					RE (30 years)
RP690M*	p.P4818L / p.Q3368X + c.5278delG	1972	USH2	22	Congenital, moderate and stable	Normal[A]	22	8		0,3 / 0,3 (32 years)	1	No response (31 years)	BE (25 years)
RP532	p.Y506X / p.Q3587X	1968	USH2		Profound	Normal[A]	15	20	Marked concentric loss	0,1 / 0,1	2	No response	BE
RP946*	**p.Q3587X / p.E2496E**	**1982**	**USHNC**	**17**									
RP1613	c.9260delT / p.G44R	1971	USH2	30	Severe	Normal[A]	20	20	Marked concentric loss	0,1 / 0,6	2	Abnormal response	BE
RP1615	c.11566delA / c.15053-1G>A	1975	USH2	27	Moderate since infancy	Normal[A]	24	25	Marked concentric loss	0,5 / 0,5	2	No response	BE

*Parental origin of the mutations was determined; [#]Patients with possibly pathogenic mutations (UV3). Age (of diagnosis), onset of night blindness and visual field loss are expressed in years. [A]: results of clinical examinations; [B]: self reported symptoms. Eye fundus 1: Bone spicules deposits, attenuation of vessels and waxy pollar of the optic nerve head. Eye fundus 2: 1 + macular affectation. ERG: Electroretinography; BE: Both Eyes; LE: Left Eye; RE: Right Eye. Patients clinically classified as USHA or USHNC are highlighted in bold.

Figure 2 shows the distribution of all the pathogenic mutations detected in the present study, along the different domains of the USH2A protein. Mutations are located evenly throughout the protein and no "hot spots" were observed. Interestingly, there are two domains in which mutations are not detected: the transmembrane and intracellular domains. None of the studies performed in *USH2A* have detected mutations in the intracytoplasmatic region, involved in the interaction of the USH2A protein with harmonin and whirlin.

There are more than 160 pathogenic variants described in previous studies. Noteworthy, 23 mutations reported in this work are novel. If we compare the mutations detected in this study with those found in other Caucasian populations, only five mutations are common with the studies of Baux *et al*. (2007) [25] and McGee *et al*. (2010) [30] and three with the work of Dreyer *et al*. (2008) [28]. However, we can also find similarities with other populations, like non-Ashkenazi Jews. The splice-site variant c.12062-2A>G was detected in three patients, in homozygous state in one of them. This mutation was initially described by Auslender *et al*. (2008) [26], as one of the most *USH2A* prevalent mutations in non-Ashkenazi Jews. Later, it was also detected in the American population [30]. We do not know the origin of our three patients, but it is tempting to speculate that they are descendant of those Sephardic Jews that were expelled from Spain in 1492 [40].

We did not find any mutation in 45 families while in 18 the second mutation remained unidentified. The number of detected pathogenic variants is probably underestimated, because there may be mutations in regions which have not been analyzed (introns, 3' and 5' untranslated regions (UTRs), promoter region, distant enhancers…) or large insertions, deletions and rearrangements that cannot be detected with the conventional PCR techniques. Moreover, some of these patients may have mutations in other genes like *GPR98,* which seems to be responsible for approximately 3-6% of USH2 cases [41,42] or *DFNB31*, although the studies indicate a minor role of *DFNB31* in USH2 [43,44]. Furthermore, USH1 genes may be responsible for phenotypically USH2 patients. Jaijo *et al*. (2010) [32] found two mutations in *CDH23* in two patients diagnosed as USH2 and a high phenotypic heterogeneity due to *CDH23* variants has been reported [45,46].

In this report, we have detected at least one mutation in 48.9% (43/88) of total patients. Considering only the patients clearly diagnosed with Usher syndrome type II, the mutation detection ratio raises to 56.9% (33/58). This detection rate is lower than expected because, as it has been mentioned before, the patient sample included in this study is biased. Thus, if we take into account all our USH2 patients studied so far (including results from previous studies [19,9,23,32, unpublished data] and the present work), our database includes 102 typical USH2 patients with at least one mutation detected in the *USH2A* gene and 32 typical USH2 patients who have been studied for all exons of this gene and no mutation was found (Table 7). Thereby, our mutation detection rate rises considerably to 76.1% (102/134), making our percentage similar to those obtained by Baux *et al*. (2007) [25] and Dreyer *et al*. (2008) [28].

Table 7. Comparison between results obtained from the sample included in the present report with the global results for our total series.

	PRESENT REPORT SAMPLE				**TOTAL SERIES[#]**			
	0 MUT	1 MUT	2 MUT	1 + 2 MUT[*]	0 MUT	1 MUT	2 MUT	1 + 2 MUT[*]
USH2	43.1% (25/58)	22.4% (13/58)	34.5% (20/58)	56.9% (33/58)	23.9% (32/134)	16.4% (22/134)	59.7% (80/134)	76.1% (102/134)
USHA	54.5% (6/11)	18.2% (2/11)	27.3% (3/11)	45.5% (5/11)	33.3% (8/24)	25% (6/24)	41.6% (10/24)	66.7% (16/24)
USHNC	73.7% (14/19)	15.8% (3/19)	10.5% (2/19)	26.3% (5/19)	64% (16/25)	12% (3/25)	24% (6/25)	36% (9/25)
TOTAL	51.1% (45/88)	20.5% (18/88)	28.4% (25/88)	48.9% (43/88)	30.6% (56/183)	16.9% (31/183)	52.5% (96/183)	69.4% (127/183)

We consider as a mutation (MUT) only those clearly pathogenic. If we also consider UV3, the total figures in the present report sample are 51.1% (45/88); 17% (15/88); 31.9% (28/88) and 48.9% (43/88) for 0 MUT, 1 MUT, 2 MUT and 1+2 MUT respectively. * Percentage of patients with at least one mutation in *USH2A*: patients with only one mutation + patients with 2 mutations (1MUT + 2MUT). [#] TOTAL SERIES represents the global results obtained from this, together with previous studies [19,9,23,32, unpublished data and present work].

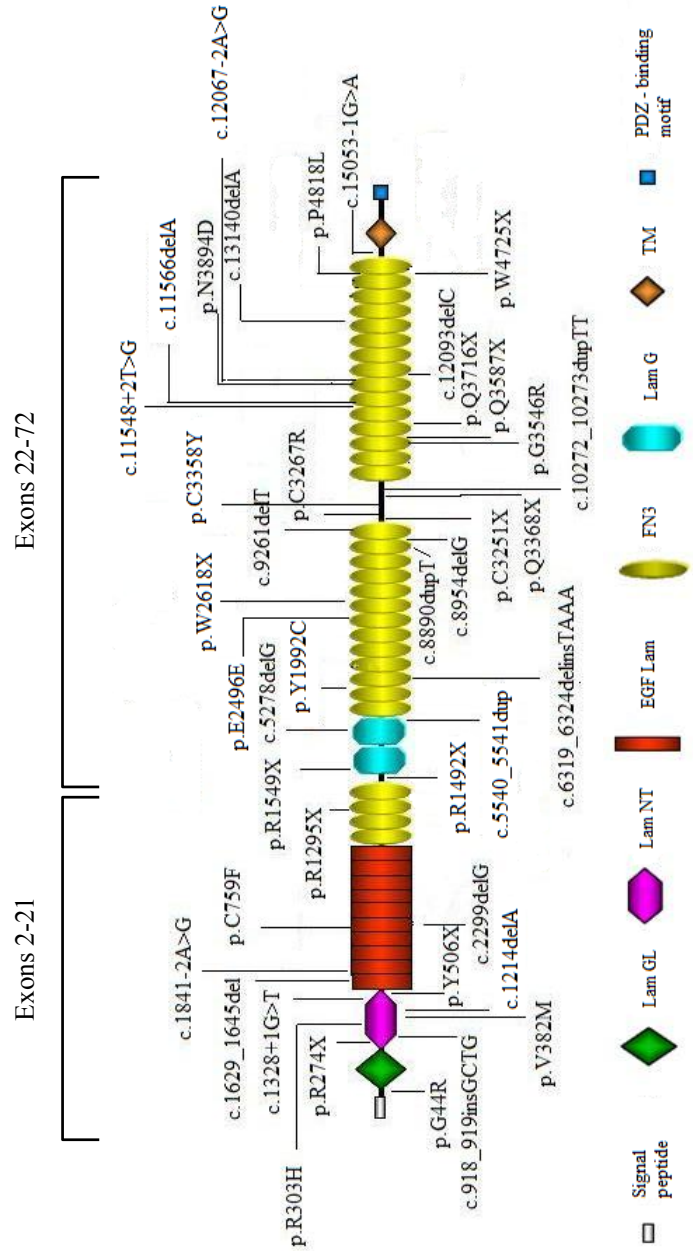


Figure. 2 Schematic illustration of the pathogenic and possibly pathogenic (UV3) mutations identified in this study along the USH2A protein domain.

Competing interests

There are no competing interests.

Authors' contributions

GG carried out the mutational screening of the *USH2A* gene in Usher syndrome patients, participated in the minigene constructions and expression and drafted the manuscript. MA carried out the mutational screening of the *USH2A* gene in controls and participated in the minigene constructions and expression. TJ carried out the Splice-site predictions and drafted the manuscript. RR carried out bioinformatics' predictions of the pathogenic effect of missense variations. AL participated in the mutational screening of the *USH2A* gene in Usher syndrome patients. AA and SB participated in patients' and controls' genomic DNA extraction. FB participated in clinical anamnestic data collection. RN, MB and CA participated in patient recruitment, pedigree and clinical anamnestic data collection. MD participated in the clinical diagnosis of patients. JM participated in patient recruitment, pedigree and clinical anamnestic data collection and drafted the manuscript. EA coordinated and supervised the study and drafted the manuscript. All authors read and approved the manuscript.

Acknowledgements

Authors are grateful to the patients participating in the study and to their family members, and also to the FARPE for their help and co-operation. The exon trapping expression vector pSPL3 was kindly provided by S. Tuffery-Giraud and I. Bottillo. Funding: This work was supported by PI07/0558, PI08/90311 and PI10/01825, from the Spanish Ministry of Science and Innovation and GVPRE/2008/024 from the Conselleria de Sanitat de la Comunitat Valenciana. CIBERER is an initiative of the Institute of Health Carlos III from the Spanish Ministry of Science and Innovation. Gema García-García and María José Aparisi are recipient of a fellowship from the Spanish Ministry of Education (REF: AP2008-02760 and AP2009-3344, respectively). Regina Rodrigo has a Contrato-Investigador SNS Miguel Servet (CP09/118) from Instituto de Salud Carlos III, Ministerio de Ciencia e Innovación. Patient consent: obtained; Ethics approval: This study was conducted according to the tenets of Helsinki declaration and with the approval of the Hospital La Fe ethics committee.

REFERENCES

1. Keats BJ, Corey DP: The usher syndromes. Am J Med Genet 1999, 89:158-66. Review.
2. Espinós C, Millán JM, Beneyto M, Nájera C: Epidemiology of Usher syndrome in Valencia and Spain. Community Genet 1998, 1:223-8.
3. Saihan Z, Webster AR, Luxon L, Bitner-Glindzicz M: Update on Usher syndrome. Curr Opin Neurol 2009, 22:19-27.
4. Millán JM, Aller E, Jaijo T, Blanco-Kelly F, Gimenez-Pardo A, Ayuso C: An update on the genetics of usher syndrome. J Ophthalmol 2011, 2011:417217.

5. Rosenberg, T; Haim, M; Hauch, AM; Parving, A: The prevalence of Usher syndrome and other retinal dystrophy-hearing impairment associations. Clin. Genet 1997, 51:314-21.
6. Spandau UH, Rohrschneider K: Prevalence and geographical distribution of Usher syndrome in Germany. Graefes Arch Clin Exp Ophthalmol 2002, 240:495-498.
7. Pieke-Dahl S, van Aarem A, Dobin A, Cremers CW, Kimberling WJ: Genetic heterogeneity of Usher syndrome type II in a Dutch population. J Med Genet 1996, 33:753-7.
8. Rivolta C, Sweklo EA, Berson EL, Dryja TP: Missense mutation in the USH2A gene: association with recessive retinitis pigmentosa without hearing loss. Am J Hum Genet 2000, 66:1975-1978.
9. Aller E, Najera C, Millan JM, Oltra JS, Perez-Garrigues H, Vilela C, Navea A, Beneyto M: Genetic analysis of 2299delG and C759F mutations (USH2A) in patients with visual and/or auditory impairments. Eur J Hum Genet 2004, 12:407-10.
10. Kimberling WJ, Weston MD, Möller C, van Aarem, Cremers CWRJ, Sumegi J, Ing PS, Connolly C, Martini A, Milani M, Tamayo ML, Bernal J, Greenberg J, Ayuso C: Gene mapping of Usher syndrome type IIa: localization of the gene to a 2.1 –cM segment on chromosome 1q41. Am J Human Genet 1995, 56:216-223.
11. Eudy JD, Weston MD, Yao S, Hoover DM, Rehm HL, Ma-Edmonds M, Yan D, Ahmad I, Cheng JJ, Ayuso C, Cremers C, Davenport S, Moller C, Talmadge CB, Beisel KW, Tamayo M, Morton CC, Swaroop A, Kimberling WJ, Sumegi J: Mutation of a gene encoding a protein with extracellular matrix motifs in Usher syndrome type IIa. Science 1998, 280:1753-7.
12. Weston MD, Eudy JD, Fujita S, Yao S, Usami S, Cremers C, Greenberg J, Ramesar R, Martini A, Moller C, Smith RJ, Sumegi J, Kimberling WJ: Genomic structure and identification of novel mutations in usherin, the gene responsible for Usher syndrome type IIa. Am J Hum Genet 2000, 66:1199-210.
13. van Wijk E, Pennings RJ, te Brinke H, Claassen A, Yntema HG, Hoefsloot LH, Cremers FP, Cremers CW, Kremer H: Identification of 51 novel exons of the Usher syndrome type 2A (USH2A) gene that encode multiple conserved functional domains and that are mutated in patients with Usher syndrome type II. Am J Hum Genet 2004, 74:738-44.
14. Adato A, Lefevre G, Delprat B, Michel V, Michalski N, Chardenoux S, Weil D, El-Amraoui A, Petit C: Usherin, the defective protein in Usher syndrome type IIA, is likely to be a component of interstereocilia ankle links in the inner ear sensory cells. Hum Mol Genet 2005, 14:3921-32.
15. van Wijk E, van der Zwaag B, Peters T, Zimmermann U, Te Brinke H, Kersten FF, Märker T, Aller E, Hoefsloot LH, Cremers CW, Cremers FP, Wolfrum U, Knipper M, Roepman R, Kremer H: The DFNB31 gene product whirlin connects to the Usher

protein network in the cochlea and retina by direct association with USH2A and VLGR1. Hum Mol Genet. 2006, 15(5):751-65.
16. Liu XZ, Hope C, Liang CY, Zou JM, Xu LR, Cole T, Mueller RF, Bundey S, Nance W, Brown SD: A mutation (2314delG) in the Usher syndrome type IIA gene: high prevalence and phenotypic variation. Am J Hum Genet 1999, 64:1221-1225.
17. Adato A, Weston MD, Berry A, Kimberling WJ, Bonne-Tamir: A Three novel mutations and twelve polymorphisms identified in the USH2A gene in Israeli USH2 families. Hum Mutat 2000, 15:388.
18. Dreyer B, Tranebjaerg L, Rosenberg T, Weston M.D, Kimberling WJ, Nilssen: Identification of novel USH2A mutations: implications for the structure of USH2A protein. Eur J Hum Genet 2000, 8:500-506.
19. Najera C, Beneyto M, Blanca J, Aller E, Fontcuberta A, Millan JM, Ayuso C Mutations in myosin VIIA (MYO7A) and usherin (USH2A) in Spanish patients with Usher syndrome types I and II, respectively. Hum Mutat 2002, 20:76-77.
20. Bernal S, Ayuso C, Antiñolo G, Gimenez A, Borrego S, Trujillo MJ, Marcos I, Calaf M, Del Rio E, Baiget M: Mutations in USH2A in Spanish patients with autosomal recessive retinitis pigmentosa: high prevalence and phenotypic variation. J Med Genet 2003, 40:e8.
21. Pennings RJ, Te Brinke H, Weston MD, Claassen A, Orten DJ, Weekamp H, Van Aarem A, Huygen PL, Deutman AF, Hoefsloot LH, Cremers FP, Cremers CW, Kimberling WJ, Kremer H: USH2A mutational analysis in 70 Dutch families with Usher syndrome type II. Hum Mutat. 2004, 24:185.
22. Maubaret C, Griffoin J-M, Arnaud B, Hamel CP: Novel mutations in MYO7A and USH2A in Usher syndrome. Ophthalm Genet. (2005) 26:25-29.
23. Aller E, Jaijo T, Beneyto M, Nájera C, Oltra S, Ayuso C, Baiget M, Carballo M, Antiñolo G, Valverde D, Moreno F, Vilela C, Collado D, Pérez-Garrigues H, Návea A, Millán JM: Identification of 14 novel mutations in the long isoform of USH2A in Spanish patients with Usher syndrome type II. J Med Genet 2006, 43:e55.
24. Cremers FP, Kimberling WJ, Külm M, de Brouwer AP, van Wijk E, te Brinke H, Cremers CW, Hoefsloot LH, Banfi S, Simonelli F, Fleischhauer JC, Berger W, Kelley PM, Haralambous E, Bitner-Glindzicz M, Webster AR, Saihan Z, De Baere E, Leroy BP, Silvestri G, McKay GJ, Koenekoop RK, Millan JM, Rosenberg T, Joensuu T, Sankila EM, Weil D, Weston MD, Wissinger B, Kremer H: Development of a genotyping microarray for Usher syndrome. J Med Genet 2007, 44:153-160.
25. Baux D, Larrieu L, Blanchet C, Hamel C, Ben Salah S, Vielle A, Gilbert-Dussardier B, Holder M, Calvas P, Philip N, Edery P, Bonneau D, Claustress M, Malcolm S, Roux AF: Molecular and In Silico Analyses of the Full-lenth Isoform of Usherin Identify New Pathogenic Alleles in Usher Type II Patients. Hum Mut 2007, 28:781-9.
26. Auslender N, Bandah D, Rizel L, Behar DM, Shohat M, Banin E, Allon-Shalev S, Sharony R, Sharon D, Ben-Yosef T: Four USH2A founder mutations underlie the

majority of Usher syndrome type 2 cases among non-Ashkenazi Jews. Genet Test 2008, 12: 289–94.
27. Dai H, Zhang X, Zhao X, Deng T, Dong B, Wang J, Li Y: Identification of five novel mutations in the long isoform of the USH2A gene in Chinese families with Usher syndrome type II. Mol Vis 2008, 4:2067-75.
28. Dreyer B, Brox V, Tranebjaerg L, Rosenberg T, Sadeghi AM, Möller C, Nilssen O: Spectrum of USH2A mutations in Scandinavian patients with Usher syndrome type II. Hum Mutat 2008, 29(3):451.
29. Yan D, Ouyang X, Patterson DM, Du LL, Jacobson SG, Liu XZ: Mutation analysis in the long isoform of USH2A in American patients with Usher Syndrome type II. J Hum Genet 2009, 54(12):732-8.
30. McGee TL, Seyedahmadi BJ, Sweeney MO, Dryja TP, Berson EL: Novel mutations in the long isoform of the USH2A gene in patients with Usher syndrome type II or non-syndromic retinitis pigmentosa. J Med Genet 2010, 47(7):499-506.
31. Le Guédard-Méreuze S, Vaché C, Baux D, Faugère V, Larrieu L, Abadie C, Janecke A, Claustres M, Roux AF, Tuffery-Giraud S: Ex vivo splicing assays of mutations at noncanonical positions of splice sites in Usher genes. Hum Mutat 2010, 31(3):347-55.
32. Jaijo T, Aller E, García-García G, Aparisi MJ, Bernal S, Avila-Fernández A, Barragán I, Baiget M, Ayuso C, Antiñolo G, Díaz-Llopis M, Külm M, Beneyto M, Nájera C, Millán JM: Microarray-based mutation analysis of 183 Spanish families with Usher syndrome. Invest Ophthalmol Vis Sci. 2010, 51(3):1311-7.
33. Jaijo T, Aller E, Aparisi MJ, García-García G, Hernan I, Gamundi MJ, Nájera C, Carballo M, Millán JM: Functional analysis of splicing mutations in MYO7A and USH2A genes. Clin Genet 2011, 79(3):282-8.
34. Bottillo I, De Luca A, Schirinzi A, Guida V, Torrente I, Calvieri S, Gervasini C, Larizza L, Pizzuti A, Dallapiccola B: Functional analysis of splicing mutations in exon 7 of NF1 gene. BMC Med Genet 2007, 12;8:4.
35. Baux D, Faugère V, Larrieu L, Le Guédard-Méreuze S, Hamroun D, Béroud C, Malcolm S, Claustres M, Roux AF: UMD-USHbases: a comprehensive set of databases to record and analyse pathogenic mutations and unclassified variants in seven Usher syndrome causing genes. Hum Mutat 2008, 29(8);E76-E87.
36. Aller, E: Molecular bases of Usher syndrome. Identifying mutations in known genes. Functional studies searching for candidate genes. PhD thesis. Valencia University, Genetic Department; 2007.
37. de Pontual L, Pelet A, Clement-Ziza M, Trochet D, Antonarakis SE, Attie-Bitach T, Beales PL, Blouin JL, Dastot-Le Moal F, Dollfus H, Goossens M, Katsanis N, Touraine R, Feingold J, Munnich A, Lyonnet S, Amiel J: Epistatic interactions with a common hypomorphic RET allele in syndromic Hirschsprung disease. Hum Mut 2007, 28:790-6.

38. Khanna H, Davis EE, Murga-Zamalloa CA, Estrada-Cuzcano A, Lopez I, den Hollander AI, Zonneveld MN, Othman MI, Waseem N, Chakarova CF, Maubaret C, Diaz-Font A, Macdonald I, Muzny DM, Wheeler DA, Morgan M, Lewis LR, Logan CV, Tan PL, Beer MA, Inglehearn CF, Lewis RA, Jacobson SG, Bergmann C, Beales PL, Attié-Bitach T, Johnson CA, Otto EA, Bhattacharya SS, Hildebrandt F, Gibbs RA, Koenekoop RK, Swaroop A, Katsanis N: A common allele in RPGRIP1L is a modifier of retinal degeneration in ciliopathies. Nat Genet 2009, 41(6):739-45.
39. Aller E, Larrieu L, Jaijo T, Baux D, Espinós C, González-Candelas F, Nájera C, Palau F, Claustres M, Roux AF, Millán JM. The USH2A c.2299delG mutation: dating its common origin in a Southern European population. Eur J Hum Genet 2010, 18(7):788-93.
40. Millán JM, Aller E, Jaijo T, Blanco-Kelly F, Corton M, Ayuso C: Molecular Epidemiology of Usher syndrome. Satpal Ahuja (ed) Usher syndrome: Pathogenesis, Diagnosis and Therapy. Nova Publishers, Hauppauge NY [Epub ahead of print]. 2011
41. Weston MD, Luijendijk MW, Humphrey KD, Möller C, Kimberling WJ: Mutations in the VLGR1 gene implicate G-protein signaling in the pathogenesis of Usher syndrome type II. Am J Hum Genet 2004, 74:357-66. Erratum in: Am J Hum Genet. 2004;74:1080.
42. Ebermann I, Wiesen MH, Zrenner E, Lopez I, Pigeon R, Kohl S, Löwenheim H, Koenekoop RK, Bolz HJ: GPR98 mutations cause Usher syndrome type 2 in males. J Med Genet 2009, 46(4):277-80.
43. Ebermann I, Scholl HP, Charbel Issa P, Becirovic E, Lamprecht J, Jurklies B, Millán JM, Aller E, Mitter D, Bolz H: A novel gene for Usher syndrome type 2: mutations in the long isoform of whirlin are associated with retinitis pigmentosa and sensorineural hearing loss. Hum Genet 2007, 121:203-11.
44. Aller E, Jaijo T, van Wijk E, Ebermann I, Kersten F, García-García G, Voesenek K, Aparisi MJ, Hoefsloot L, Cremers C, Díaz-Llopis M, Pennings R, Bolz HJ, Kremer H, Millán JM: Sequence variants of the DFNB31 gene among Usher syndrome patients of diverse origin. Mol Vis 2010, 23;16:495-500.
45. Bolz H, von Brederlow B, Ramírez A, Bryda EC, Kutsche K, Nothwang HG, Seeliger M, del C-Salcedó Cabrera M, Vila MC, Molina OP, Gal A, Kubisch C: Mutation of CDH23, encoding a new member of the cadherin gene family, causes Usher syndrome type 1D. Nat Genet 2001, 27:108-12.
46. Astuto LM, Bork JM, Weston MD, Askew JW, Fields RR, Orten DJ, Ohliger SJ, Riazuddin S, Morell RJ, Khan S, Riazuddin S, Kremer H, van Hauwe P, Moller CG, Cremers CW, Ayuso C, Heckenlively JR, Rohrschneider K, Spandau U, Greenberg J, Ramesar R, Reardon W, Bitoun P, Millan J, Legge R, Friedman TB, Kimberling WJ: CDH23 mutation and phenotype heterogeneity: a profile of 107 diverse families with Usher syndrome and nonsyndromic deafness. Am J Hum Genet 2002, 71(2):262-75.

3. Article:

The contribution of *GPR98* and *DFNB31* genes to a Spanish Usher syndrome type 2 cohort

Gema García-García [1,2], Thomas Besnard [1,3,4], David Baux [3], Christel Vaché [3], Elena Aller [2], Sue Malcolm [5], Mireille Claustres [1,3,4], Jose M. Millan [2], Anne-Françoise Roux [1,3]#

1. INSERM, U827, Montpellier, F-34000, France
2. Grupo de Investigación en Enfermedades Neurosensoriales, Instituto de Investigación Sanitaria IIS-La Fe and Centro de Investigación Biomédica en Red de Enfermedades Raras, Valencia, Spain
3. Centre Hospitalier Universitaire, Montpellier, Laboratoire de Génétique Moléculaire, Montpellier, F-34000, France
4. Univ, Montpellier I, Montpellier, F-34000, France
5. Clinical and Molecular Genetics, Institute of Child Health, University College London, London, United Kingdom

Molecular Vision. 2013 Feb 19:367-73

ABSTRACT

Background: Usher syndrome type 2 (USH2) is an autosomal recessive disease characterised by moderate to severe hearing loss and retinitis pigmentosa. To date, three disease-causing genes have been identified, *USH2A*, *GPR98* and *DFNB31* of which *USH2A* is clearly the major contributor. The aim of this work was to determine the contribution of *GPR98* and *DFNB31* genes in a Spanish cohort of *USH2A* negative patients using exhaustive molecular analysis, including sequencing, dosage and splicing analysis.
Methods: Linkage analysis was performed to prioritise the gene to study, followed by sequencing of exons and intron-exon boundaries of the selected gene, *GPR98* (90 exons) or *DFNB31* (12 exons). Functional splicing analyses and CGH-array to detect large rearrangements were performed when appropriate.
Results: We confirmed that mutations in *GPR98* contribute a significant but minor role to Usher type 2. In a group of patients referred for molecular diagnosis, 43 had been found to be positive for *USH2A* mutations, 19 remaining without *USH2A* alterations were screened and seven different mutations were identified in the *GPR98* gene in 7 patients (five in the homozygous state), of which 6 were novel. All detected mutations result in a truncated protein; deleterious missense mutations were not found. No pathological mutations were identified in *DFNB31* gene.
Conclusions: In Spain, *USH2A* and *GPR98* are responsible for 94.8 and 5.2 % of USH2 mutated cases, respectively. *DFNB31* plays a minor role in the Spanish population. There was a group of patients in whom no mutation could be found. These findings confirm the importance of including at least *GPR98* analysis for comprehensive USH2 molecular diagnosis.
Key Words: Usher syndrome, *GPR98*, *VLGR1*, *DFNB31*, molecular analysis

BACKGROUND

Usher syndrome (USH, OMIM 276900, OMIM 276905, OMIM 605472) is recessive inherited disease characterized by sensorineural hearing loss (HL), visual loss due to retinitis pigmentosa (RP) and, in some cases, vestibular dysfunction. The syndrome is the most common cause of combined visual and hearing loss, accounting for more than 50% of adult cases with deaf-blindness [1]. Prevalence estimates have ranged from 3.2-6.2/100.000 with a recent study indicating that USH prevalence could be much higher at up to 1/6000 [2].

Patients with USH are classified into three clinical subtypes (USH1, USH2, USH3), based on the severity and progression of hearing impairment and presence or absence of vestibular dysfunction [3, 4]. USH2, the subject of this study, is the most common type and is characterized by moderate to severe congenital HL and normal vestibular function [5]. Usually RP develops during the second decade.

Three USH2 genes are known, *USH2A*, *GPR98* (also known as *VLGR1*) and *DFNB31*. The long isoforms of *USH2A* (*USH2Ab*) and *GPR98 (VLGR1b)* encode two transmembrane proteins, usherin and G protein-coupled receptor 98, respectively, that contain large extracellular domains. *DFNB31* encodes the post synaptic density protein (PSD95), Drosophila disc large tumor suppressor (Dlg1), and zonula occludens-1 protein (ZO-1; PDZ) domain-containing scaffold protein, whirlin. These three USH2 proteins are part of the Usher protein complex, in which USH1 and USH2 proteins are assembled in a multiprotein scaffold with a major function in the cochlea hair cells as well as in the photoreceptor cells [6-8].

Among the three known genes responsible for USH2, results from large European cohorts [9-12] have shown that *USH2A* is by far the most frequently involved gene and accounts for at least 75% of USH2 cases. Molecular analyses of the *GPR98* and *DFNB31* genes remained scarce until recently because of their minor involvement and, logistically, because of the high number of exons (n=90) to screen in *GPR98*. The most thorough study of *GPR98* and *DFNB31* was performed by Besnard *et al*., who reported 17 mutations in *GPR98* and 2 in *DFNB31* equivalent to an involvement in USH2 of 6.4% and 1.3% for *GPR98* and *DFNB31*, respectively [12]. Two other analyses found a contribution of 6% or 19 % for *GPR98* and 0% or 9.5 % for *DFNB31,* respectively [11,13].

Recently, another gene, *PDZD7,* was shown to contribute to USH2 as a modifier of the retinal phenotype on a *USH2A* background or in digenic inheritance with *GPR98* [14]. We have previously studied the *USH2A* gene in a Spanish cohort, which accounts for 76.1% of USH2 patients [10], leaving a significant percentage of unresolved cases. We present in this work findings of the exhaustive mutational screening of *GPR98* and *DFNB31* performed in this *USH2A* negative cohort.

MATERIAL AND METHODS

Subjects

The patients were recruited from the Federación de Afectados de Retinosis Pigmentaria de España (FARPE) and also from the Ophthalmology and ENT Services of several Spanish Hospitals as part of a large-scale study on the genetics of Usher syndrome in the Spanish population. The 19 *USH2A* negative patients genotyped in this work were previously studied in [10] and are divided as follows: 12 patients were classified as USH2, 5 displayed atypical Usher syndrome and in 2 cases, detailed clinical data could not be obtained. The subjects had been classified on the basis of their clinical history and ophthalmologic, audiometric and vestibular tests.

Molecular analyses

Haplotypes and sequencing analyses of *GPR98* and *DFNB31* were performed as already described [12]. Nomenclature of the variants follows the Human Genome Variation Society (HGVS9 recommendations. A laboratory-designed comparative genomic hybridization (CGH) microarray chip (12×135 k), which includes all Usher genes and their 5′ and 3′ regions, was used to detect large genomic rearrangements

In silico analyses

The potential effects on splicing of any sequence variation were analysed with the Human Splicing Finder (HSF) tool. The multi-step analysis described in Baux *et al.*, [16] and Roux *et al*., [15] was used to classify the variants. In particular, USMA was used to predict the impact of the missense variants on protein structure.

Pathogenicity grades

The classification system for Unknown Variants is the same as that used in USHbases and is as follow: UV1: variant certainly neutral; UV2: variant likely neutral; UV3: variant likely pathogenic; UV4: variant certainly pathogenic. This classification is in line with the guidelines published by the clinical and molecular genetics society (http://www.cmgs.org/BPGs/Best_Practice_Guidelines.htm).

Briefly variants were classified based on the following critera: previously published, allele frequencies, whether they are in cis or trans to deleterious mutations/UVs, predictions from bioinformatics as to whether the change is in a conserved region and whether it is likely to alter protein structure. The two latter are considered as the main criteria.

Minigene construction and expression

In vitro analyses were performed to evaluate the functional consequence at RNA level of variant c.14368C>T. We used a minigene construct based on the expression vector pSPL3 [17], generated by Besnard *et al.*, [12] which included the wild-type exon 70 and surrounding sequences of *GPR98*. The c.14368C>T variant was generated by site-directed mutagenesis (Quick change II, Stratagene, La Jolla, CA). The minigene construct was transiently transfected into ARPE-19 cells (ATCC, CRL-2502TM) during 24h. Briefly, 70-80% confluence cells plated in six well plates were transfected with the

FuGENE6 Transfection Reagent (Roche Diagnosis, Indianapolis,IN) according to the manufacturer's instructions. Reverse transcriptase reactions were carried out with the Superscript II Reverse Transcriptase (Invitrogen, Cergy-Pontoise, France) on total RNA extracted from cells with the Nucleospin RNAII kit (Macherey-Nagel, Hoerdt, France). Polymerase chain reactions were performed using vector-specific primers (5'-CATCCTGGTCAGCTGGACG-3'; 5'- GTAGGTCAGGGTGGTCACGA-3') and amplification products were analysed as previously described [18].

GenBank numbers

GenBank reference sequences: *GPR98*: NM_32119.3; *DFNB31*: NM_015404.2. The +1 position corresponds to A in the ATG translation initial codon.

URLs

HGVS http://www.hgvs.org/rec.html
1000 genomes: http://www.1000genomes.org/
Exome Variant Server: http://evs.gs.washington.edu/EVS/
USHbases : https://grenada.lumc.nl/LOVD2/Usher_montpellier/
USMA : https://neuro-2.iurc.montp.inserm.fr/USMA/
HSF: http://www.umd.be/HSF/

RESULTS

Haplotype analyses

Haplotype analyses were performed as a first step at the USH2C (*GPR98*) and USH2D (*DFNB31*) loci because consanguinity was documented in some families (n= 3) or because several sibs were available. Homozygosity was revealed at the USH2C locus, in five families (RP1188, RP153, RP1157, RP952, RP1068, Table 1). Subsequent sequencing of the *GPR98* gene identified a homozygous mutation in all the cases (see below). Haplotype analyses excluded both USH2C and USH2D loci in one family, RP98. This family did not undergo subsequent investigation in this study.

Mutational analysis

GPR98

Sequencing of the 90 coding exons of *GPR98* revealed, in seven of the 18 patients (who had not been excluded by haplotyping for the USH2C locus) seven different mutations, of which 6 were novel (Table 2). Five patients were homozygotes, one patient was a compound heterozygote and one patient carried only one identified mutation (RP1634) (Table 1). Five of the patients with *GPR98* mutations were diagnosed as Usher syndrome type 2, and two patients could not be classified because of lack of clinical data (table 1).

Table 1. Genotype of the patients bearing mutations in *GPR98*. All listed mutations are predicted to lead to a truncated protein due to the apparition of a PTC either directly, or because of a frameshift.

Patient	Mutations	Diagnosis	Year of birth	Reported Consanguinity
RP1068	c.17368_17369delinsTTAT / c.17368_17369delinsTTAT	USH2	1966	Yes
RP1157	c.18261delA / c.18261delA	USH2	1953	Yes
RP1188	c.17204+4_17204+7del / c.17204+4_17204+7del	USH2		No
RP153	c.6932_6939dup / c.6932_6939dup	USH2	1958	No
RP952	c.12528-1G>T / c.12528-1G>T	USH2		Yes
RP1590	c.10301delT / c.12528-1G>T	USH		
RP1634	c.17386C>T / -	USH		

Five of the seven mutations predict a premature termination codon (PTC), leading to a truncated protein. These variants include a small duplication, two small deletions, a deletion/insertion and a nonsense mutation (Table 2). All were classified as *a priori* deleterious.

The other two pathological mutations affect splice sites, altering the correct splicing mechanisms and they were classified as UV4 (unknown variant probably pathologic). The variant c.12528-1G>T (intron 61), was detected in two families: in a homozygous state in RP952 and *in trans* to c.10301delT in RP1590 (table 1). The second splicing variant detected in a homozygous state in RP1068, is a deletion of four nucleotides (c.17204+4_17204+7del) previously described by Besnard *et al*., [12] that abolishes the +4/+5 positions of the 5' siplice site (SS), and results in the exon skipping of exon 79.

Seventy-five non-deleterious variants recorded in USHBases by our group or others were detected. Nineteen additional variants were identified (Appendix 1)), eight of them absent from any of the databases (dbSNP, 1000 genomes, Exome Variant Server). All were classified as neutral, UV1 or UV2 based on allele frequency, bioinformatic predictions or, in the case of c.14368C>T, *in vitro* experiments.

Table 2. Pathogenic mutations identified in *GPR98*

Exon	cDNA	Protein	Splice effect predicted	Reference
31	c.6932_6939dup	p.(Glu2314fs)	No	Novel
49	c.10301delT	p.(Leu3434fs)	No	Novel
IVS61	c.12528-1G>T	p.(?)	Yes	Novel
IVS79	c.17204+4_17204+7del	p.(?)	Yes	[12]
80	c.17368_17369delinsTTAT	p.(Ser5790fs)	No	Novel
80	c.17386C>T	p.(Gln5796*)	No	Novel
86	c.18261delA	p.(Gln6088fs)	No	Novel

In silico analysis of the c.14368C>T variant predicted an increase in the strength of a cryptic donor splice site recognition (score of 52.41 to 79.25 for HSF and -6.61 to 1.13 for Maximum Entropy software (MaxEnt)). An *in vitro* splicing assay was performed to test for a splicing alteration. No altered splicing was detected. Results clearly show that c.14368C>T, identified in a single patient (RP1059) did not alter proper splicing of exon 70 *in vitro*.

Analysis of the 90 *GPR98* exons was completed by CGH-array analysis for the three patients carrying a single deleterious or newly identified missense variant: patient RP1634 heterozygous for the pathological mutation c.17386C>T (p.(Gln5796*) (Table 1)) and patients RP1059 and RP1611, heterozygous for the missense alterations c.14368C>T and c.8585A>G, respectively (Appendix 1). Deletions or duplications were not detected in any of these patients, supporting the non-pathogenicity of the two missense variants, which remained as UV1 or UV2.

Audiograms for two of the genotyped patients are shown in Figure 1. They are characterized by moderate to severe hearing loss with a down-sloping configuration. This is similar to that observed by Abadie *et al*., for patients with *GPR98* mutations [5]. In both patients, tone loss is slightly stronger at high frequencies confirming the tendency for *GPR98* mutated patients to present with a more severe hearing loss than those mutated in *USH2A* [5] .

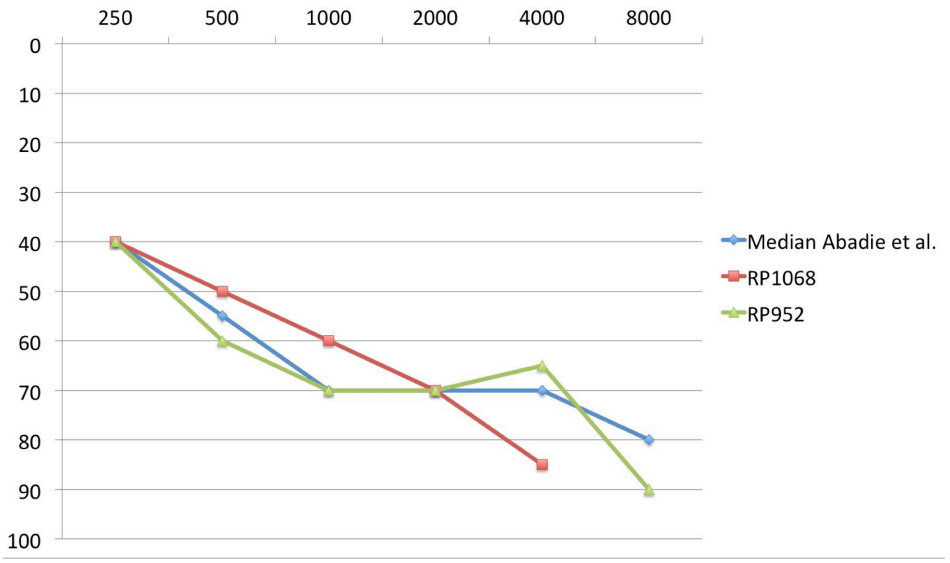


Figure 1. Audiograms from two patients mutated in *GPR98*. Age and sex corrected audiograms (ISO 7029) of RP952 and RP1068 mutated in *GPR98* are represented along with the median audiogram of *GPR98* mutated patients from Abadie *et al*., study [5]. Hearing loss (0-100) is in dB and frequencies (500-8000) are in Hz.

DFNB31

Mutational analysis of the *DFNB31* gene was performed in 13 patients for whom no homozygosity could be detected at either locus. Fourteen different variants were identified. All but one were previously recorded in the public databases with frequencies suggestive of a benign interpretation. A single novel isocoding variant c.2112G>T, localized in exon 8, was identified in family RP1600, but a deleterious effect on splicing was not predicted by *in silico* analyses and no other clear mutation could be detected in this patient. This variant was not reported in 1000 genomes or in the Exome Variant Server but was considered non-pathogenic.

DISCUSSION

Seven deleterious mutations were identified in *GPR98*. They include small insertions and deletions, point mutations and splicing alterations; all of them predicting premature termination codons. Six of them are new. This study raises the total number of established pathogenic mutations to 40. It confirms that the mutational spectrum of *GPR98* differs from that of *USH2A* in that no missense causative mutations were identified here. Mutations were spread throughout the whole gene, mainly localised in the terminal end [12].

In silico analyses of the new potential splice site mutation c.12528-1G>T predict that this mutation abolishes the wild type 3'SS of exon 62 (reducing the scores from 79.61 to 50.67 and 7.31 to -1.27, for HSF and MaxEnt matrices, respectively). The expected effects could be either a skip of exon 62 or the use of a cryptic site localised 11 nucleotides downstream the wild type acceptor site (increased strength from 79.26 to 82.42 and -2.6 to 2.8, for HSF and MaxEnt, respectively). In both hypotheses, this variation results in the disruption of the coding phase (deletion of 139 or 11 nucleotides) and therefore should be clearly considered as pathogenic.

We observed in this cohort a high number of homozygous cases, five of the seven patients positive for *GPR98,* as expected for rare mutations. Only three of these were reported to be consanguineous which confirms the helpfulness of carrying out preliminary haplotype analysis.

Our overall results are very similar to those obtained in UK [11] and in France [12]. In more than 100 USH2 patients studied in each study, both groups found an involvement of about 80 % for *USH2A* and 6 % for *GPR98* with mutations in *DFNB31* being absent or negligible. In a much smaller sample (21) Bonnet *et al*., reported a lower contribution of *USH2A (5*7%) with 4 patients and 2 patients for *GPR98* and *DFNB31* respectively [13].

Recently, Vaché *et al.*, described the first example in Usher syndrome of a deep intronic mutation causing activation of a pseudoexon, through analyses of RNA from nasal cells in a patient with only one mutation detected in the *USH2A* gene [19]. The same type of mutations could arise in *GPR98,* as patient RP1634 carries a single mutation, absence of additional genomic rearrangements has been tested and no mutation could be identified in PDZD7 (not shown). Interestingly, several patients with a single *GPR98* mutation have been identified in other studies [11, 12]. Limitations of molecular studies to the sequencing of all the coding exons and their boundaries together with CGH-array leave some unresolved cases that require further studies at RNA level. A mutation in the 5' or 3' untranslated region cannot be excluded.

Twelve USH2 patients remain with no mutation in either of the USH2 genes. These patients will undergo next generation sequencing (NGS) applied to "Usher-exome" (i.e. targeted exome of the Usher genes) as this approach is becoming available. Several patients have been reported who have presented with a clinical subtype of Usher in which the mutated gene is usually responsible for a different subtype. Several examples have been identified [11, 13, 20, 21]. Although they remain rare, they represent a real pitfall in terms of molecular diagnosis using conventional approaches such a Sanger sequencing focusing on cascade sequencing of the different genes.

Patient	Sensorineural Hearing Loss	Vestibular Function	Onset of Night Blindness	Onset of Visual Field Loss	Visual Field	Visual Acuity	Eye Fundus	ERG	Cataracts
RP1068	moderate since 7 years	Normal	30		Marked Concentric Loss	0.25 / 0.25 (at 34 years)	1		BE
RP1157	moderate-severe	?	11		Marked Concentric Loss	0.16 / 0.16 (2010)	1	Abolish (37 years)	BE
RP1188	moderate since 1 year	Normal	17	16		0.5 / 1 (2002)	1		No
RP153	moderate-severe	Normal	16		Marked Concentric Loss		1		BE
RP952	moderate since 4 years	Normal	25	35	Moderate Concentric Loss (at 45 years)	0.7 / 0.7 (at 45 years)	1	Abolish	BE

Table 3. Clinical description of the patients bearing mutations in *GPR98*.
Onset of night blindness and visual field loss are expressed in years. Eye fundus 1: Bone spicules deposits, attenuation of vessels and pale optic nerve. ERG: Electroretinography. BE: Both Eyes. No clinical description available for patients RP1590 and RP1634.

Appendix 1. Novel variants detected in *GPR98* classified as non-pathologic.

Exon/ Intron	Variant	Protein	Classification	Number Alleles[1]	Splicing Prediction	dbSNP/MAF	1000 genomes (global frequency)	EVS (all alleles)	Main Contributor to classification
5'UTR	c.-96-50G>A	-	UV1	1	No	rs73176098/0.003	0.003	-	a
1	c.30C>T	p.(=)	UV1	1	No	No	No	No	a
IVS3	c.358-149A>G	-	Neutral	1	No	rs16868853/ 0.012	0.063	-	b
IVS3	c.358-199C>G	-	UV1	1	No	rs73181609/ 0.015	0.063	-	b
IVS4	c.453+23A>C	-	Neutral	1	No	rs16868854/ 0.014	0.063	-	b
IVS5	c.558+103_IVS5+123delinsCTCCAGG	-	Neutral	7	No	No	No	-	c, d
IVS11	c.2240+73delT	-	Neutral	3	No	No	No	-	d
19	c.3509A>G	p.(Tyr1170Cys)	UV1	1	No	rs188772875/ 0.001	0.001	0.001	a
23	c.4939A>G	p.(Ile1647Val)	UV1	3	No	rs72782753/ 0.004	0.005	0.005	d
36	c.8291C>T	p.(Ser2764Leu)	Neutral	2	No	rs16869016/ 0.066	0.082	0.046	b, d
IVS38	c.8730+21dupT	-	Neutral	9	No	rs60522638/ 0.016	NAF	-	c
38	c.8585A>G	p.(Tyr2862Cys)	UV2	1	No	No	No	No	a, d
51	c.10577T>C	p.(Met3526Thr)	UV1	1	No	rs41311343/ 0.015	0.041	0.022	b
IVS51	c.10769+71C>A	-	UV1	1	No	rs57252870/ 0.015	0.027	-	b
IVS58	c.12121-31dupT	-	Neutral	9	No	No	No	-	c
70	c.14368C>T	p.(Leu4790Phe)	UV1[2]	1	Yes	No	No	No	a*
IVS73	c.14972+77C>T	-	UV2	1	No	No	No	-	a
IVS87	c.18433-18_18433-17insTC	-	Neutral	9	No	rs35858094 /NAF	NAF	-	c
3'UTR	c.*281del	-	UV1	2	No	No	No	-	d

[1] The total number of screened alleles is 36 for each amplicon; EVS Exome Variant Server. [2] Classification was performed after minigene analysis; -: not applicable; NAF: No Available Frequency. a No predicted effect (using splicing predictors and bioinformatics tools for missense variants) b Allele Frequency (public databases). c Allele frequency (patients). d patient genotype (can refer either to a genotype already containing two deleterious mutations, or to a genotype with no other identified pathogenic allele). *conservation study was not convincing; splicing prediction was unclear, however minigene assay did not show any effect

Acknowledegments
This work was supported in part by le Ministère de la Recherche "PHRC National 2004, PROM 7802". GG-G is a recipient of a fellowship from the Spanish Ministry of Education (REF: AP2008-02760). TB is a recipient of a UNADEV Foundation fellowship.

REFERENCES

1. Boughman JA, Vernon M, Shaver KA. Usher syndrome: definition and estimate of prevalence from two high-risk populations. J Chronic Dis 1983; 36(8):595-603.
2. Kimberling WJ, Hildebrand MS, Shearer AE, Jensen ML, Halder JA, Trzupek K, Cohn ES, Weleber RG, Stone EM, Smith RJ. Frequency of Usher syndrome in two pediatric populations: Implications for genetic screening of deaf and hard of hearing children. Genet Med 2010; 12(8):512-6.
3. Saihan Z, Webster AR, Luxon L, Bitner-Glindzicz M. Update on Usher syndrome. Curr Opin Neurol 2009; 22(1):19-27.
4. Millan JM, Aller E, Jaijo T, Blanco-Kelly F, Gimenez-Pardo A, Ayuso C. An update on the genetics of usher syndrome. J Ophthalmol 2011; 2011:417217.
5. Abadie C, Blanchet C, Baux D, Larrieu L, Besnard T, Ravel P, Biboulet R, Hamel C, Malcolm S, Mondain M, Claustres M, Roux AF. Audiological findings in 100 USH2 patients. Clinical genetics 2011.
6. van Wijk E, van der Zwaag B, Peters T, Zimmermann U, Te Brinke H, Kersten FF, Marker T, Aller E, Hoefsloot LH, Cremers CW, Cremers FP, Wolfrum U, Knipper M, Roepman R, Kremer H. The DFNB31 gene product whirlin connects to the Usher protein network in the cochlea and retina by direct association with USH2A and VLGR1. Hum Mol Genet 2006; 15(5):751-65.
7. Michalski N, Michel V, Bahloul A, Lefevre G, Barral J, Yagi H, Chardenoux S, Weil D, Martin P, Hardelin JP, Sato M, Petit C. Molecular characterization of the ankle-link complex in cochlear hair cells and its role in the hair bundle functioning. J Neurosci 2007; 27(24):6478-88.
8. Maerker T, van Wijk E, Overlack N, Kersten FF, McGee J, Goldmann T, Sehn E, Roepman R, Walsh EJ, Kremer H, Wolfrum U. A novel Usher protein network at the periciliary reloading point between molecular transport machineries in vertebrate photoreceptor cells. Hum Mol Genet 2008; 17(1):71-86.
9. Dreyer B, Brox V, Tranebjaerg L, Rosenberg T, Sadeghi AM, Moller C, Nilssen O. Spectrum of USH2A mutations in Scandinavian patients with Usher syndrome type II. Hum Mutat 2008; 29(3):451.
10. Garcia-Garcia G, Aparisi MJ, Jaijo T, Rodrigo R, Leon AM, Avila-Fernandez A, Blanco-Kelly F, Bernal S, Navarro R, Diaz-Llopis M, Baiget M, Ayuso C, Millan JM, Aller E. Mutational screening of the USH2A gene in Spanish USH patients reveals 23 novel pathogenic mutations. Orphanet J Rare Dis 2011; 6:65.

11. Le Quesne Stabej P, Saihan Z, Rangesh N, Steele-Stallard HB, Ambrose J, Coffey A, Emmerson J, Haralambous E, Hughes Y, Steel KP, Luxon LM, Webster AR, Bitner-Glindzicz M. Comprehensive sequence analysis of nine Usher syndrome genes in the UK National Collaborative Usher Study. Journal of medical genetics 2012; 49(1):27-36.
12. Besnard T, Vaché C, Baux D, Larrieu L, Abadie C, Blanchet C, Odent S, Blanchet P, Calvas P, Hamel C, Dollfus H, Lina-Granade G, Lespinasse J, David A, Isidor B, Morin G, Malcolm S, Tuffery-Giraud S, Claustres M, Roux AF. Non-USH2A Mutations in USH2 Patients. Human mutation 2012; 33(3):504-10.
13. Bonnet C, Grati M, Marlin S, Levilliers J, Hardelin JP, Parodi M, Niasme-Grare M, Zelenika D, Delepine M, Feldmann D, Jonard L, El-Amraoui A, Weil D, Delobel B, Vincent C, Dollfus H, Eliot MM, David A, Calais C, Vigneron J, Montaut-Verient B, Bonneau D, Dubin J, Thauvin C, Duvillard A, Francannet C, Mom T, Lacombe D, Duriez F, Drouin-Garraud V, Thuillier-Obstoy MF, Sigaudy S, Frances AM, Collignon P, Challe G, Couderc R, Lathrop M, Sahel JA, Weissenbach J, Petit C, Denoyelle F. Complete exon sequencing of all known Usher syndrome genes greatly improves molecular diagnosis. Orphanet J Rare Dis 2011; 6(1):21.
14. Ebermann I, Phillips JB, Liebau MC, Koenekoop RK, Schermer B, Lopez I, Schafer E, Roux AF, Dafinger C, Bernd A, Zrenner E, Claustres M, Blanco B, Nurnberg G, Nurnberg P, Ruland R, Westerfield M, Benzing T, Bolz HJ. PDZD7 is a modifier of retinal disease and a contributor to digenic Usher syndrome. J Clin Invest 2010; 120(6):1812-23.
15. Roux AF, Faugere V, Vache C, Baux D, Besnard T, Leonard S, Blanchet C, Hamel C, Mondain M, Gilbert-Dussardier B, Edery P, Lacombe D, Bonneau D, Holder-Espinasse M, Ambrosetti U, Journel H, David A, Lina-Granade G, Malcolm S, Claustres M. Four year follow-up of diagnostic service in USH1 patients. Invest Ophthalmol Vis Sci 2011; 52(7):4063-71.
16. Baux D, Larrieu L, Blanchet C, Hamel C, Ben Salah S, Vielle A, Gilbert-Dussardier B, Holder M, Calvas P, Philip N, Edery P, Bonneau D, Claustres M, Malcolm S, Roux AF. Molecular and in silico analyses of the full-length isoform of usherin identify new pathogenic alleles in Usher type II patients. Hum Mutat 2007; 28(8):781-9.
17. Bottillo I, De Luca A, Schirinzi A, Guida V, Torrente I, Calvieri S, Gervasini C, Larizza L, Pizzuti A, Dallapiccola B. Functional analysis of splicing mutations in exon 7 of NF1 gene. BMC Med Genet 2007; 8:4.
18. Le Guédard-Mereuze S, Vaché C, Baux D, Faugère V, Larrieu L, Abadie C, Janecke A, Claustres M, Roux AF, Tuffery-Giraud S. Ex vivo splicing assays of mutations at non-canonical positions of splice sites in USHER genes. Hum Mutat 2010; 31:347-55.
19. Vaché C, Besnard T, le Berre P, Garcia-Garcia G, Baux D, Larrieu L, Abadie C, Blanchet C, Bolz HJ, Millan J, Hamel C, Malcolm S, Claustres M, Roux AF. Usher syndrome type 2 caused by activation of an USH2A pseudoexon: Implications for diagnosis and therapy. Human mutation 2012; 33(1):104-8.

20. Astuto LM, Weston MD, Carney CA, Hoover DM, Cremers CW, Wagenaar M, Moller C, Smith RJ, Pieke-Dahl S, Greenberg J, Ramesar R, Jacobson SG, Ayuso C, Heckenlively JR, Tamayo M, Gorin MB, Reardon W, Kimberling WJ. Genetic heterogeneity of Usher syndrome: analysis of 151 families with Usher type I. Am J Hum Genet 2000; 67(6):1569-74.
21. Kalay E, de Brouwer AP, Caylan R, Nabuurs SB, Wollnik B, Karaguzel A, Heister JG, Erdol H, Cremers FP, Cremers CW, Brunner HG, Kremer H. A novel D458V mutation in the SANS PDZ binding motif causes atypical Usher syndrome. J Mol Med 2005; 83(12):1025-32.

4. Article:

Two novel disease-causing mutations in the *CLRN1* gene in USH3 patients.

Gema García-García,[1] María J. Aparisi,[1] Regina Rodrigo,[1] María D. Sequedo,[1] Carmen Espinós,[1,2] Jordi Rosell,[3] José L. Olea,[4] M. Paz Mendívil,[5] María A Ramos-Arroyo,[6] Carmen Ayuso,[2,7] Teresa Jaijo,[1,2] Elena Aller,[1,2] José M. Millán[1,2,8]

(The first two authors contributed equally to this work)

1.Grupo de Investigación en Enfermedades Neurosensoriales. Instituto de Investigación Sanitaria IIS-La Fe, Valencia, Spain

2. Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Valencia, Spain

3. Servicio de Genética. Hospital Universitario Son Espases. Palma de Mallorca, Spain

4. Servicio de Oftalmología. Hospital Universitario Son Espases. Palma de Mallorca, Spain

5. Servicio de Oftalmología. Hospital de Basurto. Basurto, Spain

6. Servicio de Genética. Hospital Virgen del Camino, Pamplona, Spain

7. Servicio de Genética, Instituto de Investigacion Sanitaria-Fundacion Jimenez Diaz (IIS-FJD). Madrid, Spain

8. Unidad de Genética. Hospital Universitario La Fe. Valencia, Spain

Molecular Vision. 2012 Dec;18:3070-8.

ABSTRACT

Purpose: To identify the genetic defect in Spanish families with Usher syndrome (USH) and probable involvement of the *CLRN1* gene.
Methods: DNA samples of the affected members of our cohort of USH families were tested using an USH genotyping array, and/or genotyped with polymorphic markers specific for the USH3A locus. Based on these previous analyses and clinical findings, *CLRN1* was directly sequenced in 17 patients susceptible of carrying mutations in this gene.
Results: Microarray analysis revealed the previously reported mutation p.Y63X in two unrelated patients, one of them homozygous for the mutation. After *CLRN1* sequencing, we found two novel mutations, p.R207X and p.I168N. Both novel mutations segregated with the phenotype.
Conclusions: To date, 18 mutations in *CLRN1* have been reported. In this work, we report two novel mutations and a third one previously identified in the Spanish USH sample. The prevalence of *CLRN1* among our USH sample is low.

INTRODUCTION

Usher syndrome (USH) is an autosomal recessive disease characterized by the association of hearing loss and visual impairment due to retinitis pigmentosa (RP), with or without vestibular dysfunction. It is the most frequent cause of concurrent deafness and blindness of genetic origin and its general prevalence ranges from 3.3 to 6.4 per 100,000 live births. However, recent studies indicate that the prevalence might be as high as 1 per 6,000 [1, 2].

USH is clinically and genetically heterogeneous. Three clinical forms are distinguished: USH1, USH2 and USH3 and nine genes have been identified responsible so far [3,4]. Five causative genes have been reported for USH1: *MYO7A*, *USH1C*, *CDH23*, *PCDH15* and *USH1G*. Three genes for USH2: *USH2A*, *GPR98* and *DFNB31* and one gene has been described for USH3: *CLRN1*. There is growing evidence suggesting that at least eight of these proteins (all except clarin 1) form a network, which is critical for the development and maintenance of the sensorineural cells in the inner ear and the retina [5-7].

The *CLRN1* gene is complex. At least 11 splice variants have been reported [8]. The main splice variant is composed of three exons that code for a 232 amino acid protein, clarin 1. Clarin 1 is a four transmembrane protein expressed in the hair cells of the organ of Corti and in the neural retina. [9, 10].

The clarin 1 protein is thought to be expressed in mouse cochlea transiently from E18 to postnatal day P6 in basal parts of the hair cells, whereas in apical parts (stereocilia) the clarin 1 expression is lost already at P1. In adult mouse retina clarin 1 localizes to inner segments, connecting cilia and ribbon synapses. The function of clarin 1 remains unknown; however, the spatiotemporal expression pattern of clarin 1 in hair cells implicates protein involvement in synaptic maturation [11, 12]. Structural and sequence homology with the synaptic protein stargazin suggest a role for clarin 1 in the plasma membranes surrounding ribbon synapses of the inner ear and retina transport [9].

USH3 is characterized by progressive hearing loss, retinitis pigmentosa and variable vestibular dysfunction, being the progressiveness of hearing loss the distinctive feature with USH1 and USH2 [13]. Although the *CLRN1* gene was initially described as responsible for USH3 cases, recent studies have demonstrated that mutations in *CLRN1* can also course with Usher clinical forms similar to USH1 and USH2 or even isolated RP [14-16].

Usher syndrome type III is rare except among Finnish and Ashkenazi Jews. In fact, most of studies show that USH3 represents less than 5% in the majority of populations. In Finland, at least 40% of Usher patients display type III whereas USH1 accounts for 34% of cases and USH2 represents only a 12% [17]. Furthermore, most of patients carry one of the two named “Finnish mutations”. The Finnish founder mutation, Finmajor, is a nonsense mutation c.528T>G at codon 176 (p.Y176X), and is responsible for

94% of the USH3 patients studied [18]. The Finminor mutation is c.359T>A (p.M120K), which is responsible for virtually the rest of *CLRN1* alleles.
Among Ashkenazi Jews the single mutation c.144T>C (p.N48K) is responsible for approximately 40% of USH cases, that means virtually the 100% of USH3 cases in the Ashkenazi population [9]. Haplotype analysis suggests a founder effect and an ancestral origin for these three mutations.

In this study, we sequenced the coding region of *CLRN1* in 17 unrelated USH patients according to our algorithm for USH3 diagnosis (figure 1). We found the mutation previously described p.Y63X [9] and two novel pathogenic mutations, p.R207X and p.I168N.

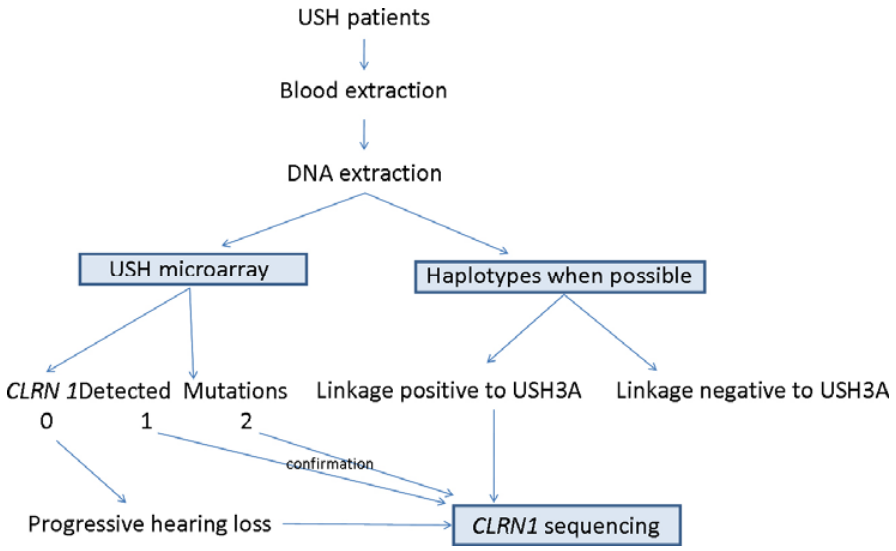


Figure 1. Algorithm for the study of USH3 in our series. When the family has more than one affected member or there is consanguinity, we haplotype the family to discard or not discard linkage to *CLRN1*. In parallel, we send DNA for USH microarray genotyping. If haplotypes are compatible with USH3A linkage or the USH micro-array detects only one mutation, we perform direct sequencing of *CLRN1*. All the mutations detected with the USH microarray were confirmed with direct sequencing.

MATERIAL AND METHODS

Subjects

Spanish subjects with Usher syndrome were mainly recruited from the Federación de Asociaciones de Afectados de Retinosis Pigmentaria del Estado Español (FARPE) and also from the Ophthalmology and ENT Services of several Spanish Hospitals as part of a large study into the genetics of Usher syndrome in Spain. An informed consent approved

by the Ethic Committee of the Hospital La Fe was obtained for all these patients and this study followed the tenets of the Declaration of Helsinki.

Subjects were classified as USH on the basis of their clinical history and underwent mutation screening of *CLRN1* as part of the algorithm for molecular diagnostic of USH carried out in our laboratory which includes the previous analysis with Asper Biotech Usher genotyping microarray and linkage analysis and haplotype compatibility when possible. DNA from 50 individuals without family history of hearing impairment or visual alterations were screened as healthy controls to evaluate the frequency of the mutations found in the patient sample.

Clinical Evaluation

Auditory function was assessed by otoscopy, standard pure tone audiometry (250-800 Hz) and tympanometry. In many cases the localization of hearing loss in the cochlea was confirmed by otoacoustic emissions (OAE) and auditory brainstem responses. Ophthalmological evaluations comprised funduscopy, visual acutity test, visual field test and electroretinography (ERG). Vestibular evaluation was performed by bithermal-caloric test, rotatory test, computerized dynamic posturography and Romberg test.

Microarray analysis

Genomic DNA was extractd from peripheral blood collected in EDTA tubes using an automated DNA extractor (MagNA Pure compact Instrument, Roche Applied Science, GmbH). Five micrograms DNA were sent to Asper Biotech (Tartu, Estonia; www.asperbio.com) to be analyzed by the USH microarray. This microarray detects 612 mutations from the nine USH genes.

The results obtained from the microarray were confirmed by direct sequencing. All exons where a mutation was identified were amplified by PCR. Amplicons were directly sequenced with dye termination chemistry (Prism Big Dye Terminator ver.1.1, Applied Biosystems, Inc. [ABI]Foster City, CA) and purified sequencing reactions were resolved in a sequencer (Prism 3130xl; ABI). To ascertain parental origin, segregation analysis was performed in those cases in which at least two pathologic variants were identified.

Construction of haplotypes

Haplotypes were constructed using the intragenic single nucleotide polymorphism (SNP) c.57A>T rs3796242 and two flanking extragenic markers D3S3661 and D3S1279 located upstream and downstream of *CLRN1* respectively.

Mutation screening

Exons 0, 2 and 3 (corresponding to the main isoform; NM_174878.2) and exon 1 (expressed in isoform NM_052995), including intron-exon boundaries of the *CLRN1* gene, were amplified using primers previously described by Adato et al. [9] with some modifications and standard polymerase chain reaction (PCR) conditions. Primers used for amplification of the four *CLRN1* exons are shown in Table 1. PCR products were sequenced using manufacturer's recommendations (Applied Biosystems). The sequences

obtained were compared with the consensus genomic sequence NG_009168.1 using the BLAST program (http://ncbi.nlm.nih.gov/blast/Blast.cgi). In those cases where mutations were detected, we carried out segregation analysis. For the construction of family trees we used Cyrillic version 2.02 software.

Predictions of the pathogenic effect of missense variations

To predict whether a rare missense variant is deleterious, we used the combined results of three different computer algorithms:

- Sort Intolerant From Tolerant (SIFT) (available at http://sift.jcvi.org) uses sequence homology to predict whether a change is tolerated or deleterious.

- The polymorphism phenotyping program, PolyPhen (available at http://genetics.bwh.harvard.edu/pph/) uses sequence conservation, structure and SWISS-PROT annotation to characterize an amino acid substitution as probably damaging, possibly damaging, benign or unknown.

- Pmut (available at http://mmb2.pcb.ub.es:8080/PMut/) provides prediction by neural networks, which use internal databases, secondary structure prediction and sequence conservation. This program provides a binary prediction of "neutral" or "pathologic".

Table 1. Primers used for amplification of the coding sequence of the *CLRN1* gene

Exon	Primers	SEQUENCES (5'-3')	Size (pb)	Primers described by
0	0-D	TCCCATTGCTCACAAAGGTCTTGTTTTTG	380	Present study
	0-R	ATTCCTCGCAACACTGGGAA		Present study
1	1-D	TCACTATCTGAAACTATCTTGTTGT	910	9
	1-R	AAGCCCCTGAACTTTATAGG		9
2	2-D	TCAGAAGGATTTTAGTGATGTTTGA	358	9
	2-R	AGACGGTCTTTTTGACATATTGAAAAGCACA		Present study
3	3AB-D	ATGTCAATGGGGATGATGGT	981	9
	3AB-D-N	TTTACACATTGACCCTCTTCC	946	Present study
	3AB-R	CAGGCTGTAACTCGAACTCC		Present study
	3B-D	GTAGCTGCAGATCTAATGTAC	242	Present study
	3B-R	GTCAAGCAATTTCCCACCAG		Present study
	3BC-D	AAGTATACTCTTAGGCCAGGC	944	Present study
	3BC-R	CCTTTGTGGCTAGACTGAATT		Present study

RESULTS

Mutation Analysis

Previous studies allowed to identify two *CLRN1* mutations responsible for USH in two Spanish families respectively. The mutation c.189C>A that led to a substitution of a tyrosin by a stop codon resulting in a truncated protein (p.Y63X) was found by Adato *et al*. in 2002 [9] in homozygous state in a family diagnosed as USH1. The second mutation (p.C40G) was homozygously found in a possible consanguineous Spanish USH3 family by Aller *et al*. in 2004 [21]. Both mutations were incorporated onto the Usher genotyping microchip from Asper Ophthamics (Tartu, Estonia; www.asperbio.com). Further microchip analyses identified the p.Y63X mutation in homozygous state in the patient of another consanguineous family (FRP-44) and in heterozygous state in a patient from a second family (FRP-529).

The screening of the *CLRN1* gene in the affected female of FRP-529 revealed a novel nonsense mutation in exon 3, c.619C>T, that led to a truncated protein in codon 207, p.R207X. Figure 2A shows the electropherogram of the sequence showing the c.619C>T mutation. The segregation analysis revealed that the mother was the carrier of the novel mutation p.R207X whereas the father was carried of p.Y63X. The healthy sib was carrier of p.Y63X.

Haplotype analysis in the family FRP360 was compatible with disease-association with the USH3A locus. After *CLRN1* mutation screening in this family, the c.619C>T (p.R207X) mutation was detected heterozygously in both affected males together with the missense mutation c.503T>A (p.I168N). Figure 2B shows the electropherogram corresponding to the c.503T>A mutation.

A) p.R207X c.619C>T

B) p.I168N c.503T>A

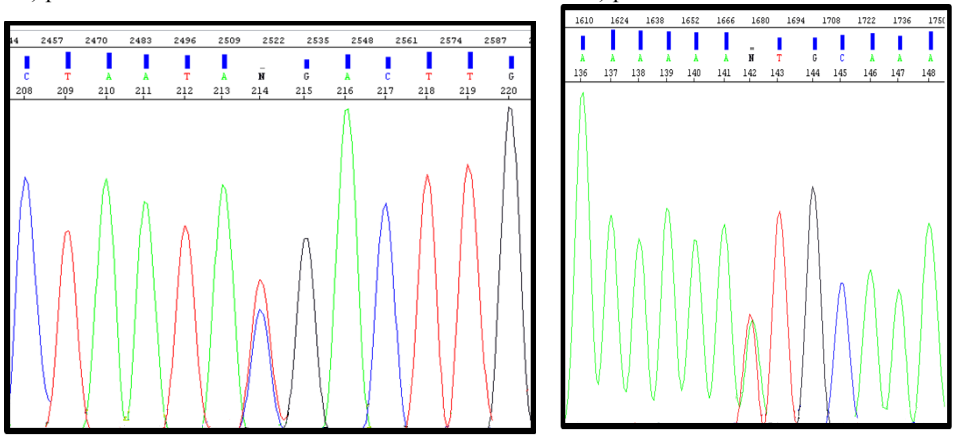


Figure 2. Electropherograms showing the novel *CLRN1* mutations identified in this study. **A**: c.619C>T mutation (p.R207X). **B**: c.503T>A mutation (p.I168N).

No other pathological mutations were found in *CLRN1* in any of the additional 15 USH samples. Table 2 shows the mutations in *CLRN1* found in the Spanish USH population detected in this and previous studies.

In silico prediction

We performed computational analysis with the programs SIFT, PolyPhen, and Pmut to infer the pathologic effect of the variant p.I168N. These programs generated a pathogenic nature prediction for the change. The SIFT program predicted that the substitution of a isoleucine for an asparagine in codon 168 (p.I168N) of the protein would affect its function (score of 0.00; normalized probabilities less than 0.05 are predicted to be deleterious; those greater than or equal to 0.05 are predicted to be tolerated). The program Pmut predicted that change p.I168N was pathologic with a reliability of 0.9485. Finally, the program PolyPhen predicted that the amino acid substitution was probably damaging the protein.

Table 2. Mutations identified in the CLRN1 gene in the Spanish USH population.

Nucleotide change	Exon	Predicted effect	No. of alleles	No. of families	References
c.118C>T	0	p.C40G	2	1	(23)
c.189C>A	0	p.Y63X	9	3	(9)
c.503T>A	**3**	**p.I168N**	1	1	Present Study
c.619C>T	**3**	**p.R207X**	2	2	Present Study

The novel mutations detected in this study are marked in bold. The reference sequence for nomenclature is NM_174878.2. Column 4 indicates the number of mutated alleles in the total Spanish USH series. Column 5 indicates the number of families in which the mutation was detected.

Clinical description

Patients included in this work displayed a wide range of ophthalmological and audiological symptoms even among patients from the same family.

The patient RP614 from family FRP-44 carried the p.Y63X homozygously. She was the daughter of a consanguineous marriage. Her hearing loss was prelingual and severe. The patient used audiphones until the age of 33 when she underwent a cochlear implantation. Hearing loss appeared to be progressive. The onset of visual phenotype occurred at 14 years old but it progressed rapidly to a severe RP. No data about vestibular function was available.

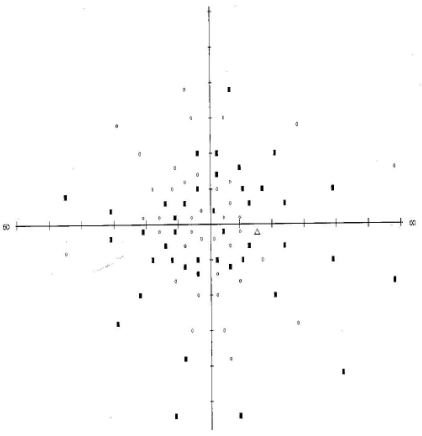
The patient RP1850 from FRP-259 carried the p.Y63X mutation together with the novel p.R207X. She was a 19 years old female who presented with a bilateral severe progressive sensorineural hearing loss corrected with audiphones. Currently, she is awaiting for cochlear implantation. She showed a delay in gait development and a vestibular

hyporeflexia. The onset of her RP was at nine years old including night blindness and peripheral visual loss. Fundus ophthalmoscopy showed pigmentary anomalies typical of RP since young, attenuated arteriolar vessels and increased bright of the internal limitant membrane. The visual field showed an anular scotoma reduced by 15º. Her best-corrected visual acuity was 0.4 in both eyes but she refers a rapid progression of visual loss in the last year. The eye fundus and visual field of the patient is shown in figure 3.

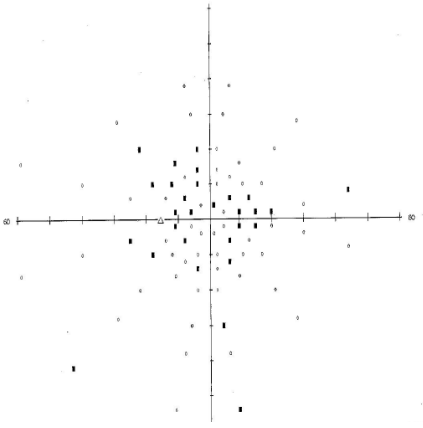
The patients from FRP360 were compound heterozygotes for p.R207X and p.I168N. They displayed very discordant phenotypes. Patient RP-1520 is a 52 years old male that started with night visual alterations at six years old, but he was not diagnosed until 15. Visual acuity was 0.05 in both eyes, he presented with pseudophakya. The eye fundus was typical of RP and the visual field was reduced by 5º around the fixation point in both eyes. He had a normal speech acquisition and normal motor milestones (he started gait at nine months old). At 13 years old, he presented with a progressive hearing loss, without tinnitus nor balance dysfunction. He needed audiphones at 42 years old and currently he shows a moderate bilateral sensorineural hearing loss ranging 79-80 dB and normal vestibular function.

Patient RP-1521 is a 49 years old male who presented with severe hearing loss since born. He showed a delay in speech acquisition, starting to speak at 36 months and requiring deaf school education. He refers vestibular dysfunction at 10 years old. The first ophthalmological symptoms were detected at the age of 8 years old as night blindness and tunnel vision. The eye fundus was typical of RP with a reduction of 10º in OD and 20º in OI. The visual acuity was 0.2 (OD) and 0.3 (OI). A brief clinical description of patients with mutations in *CLRN1* is shown in table 3.

3A) Visual Field Right Eye



Visual Field Left Eye



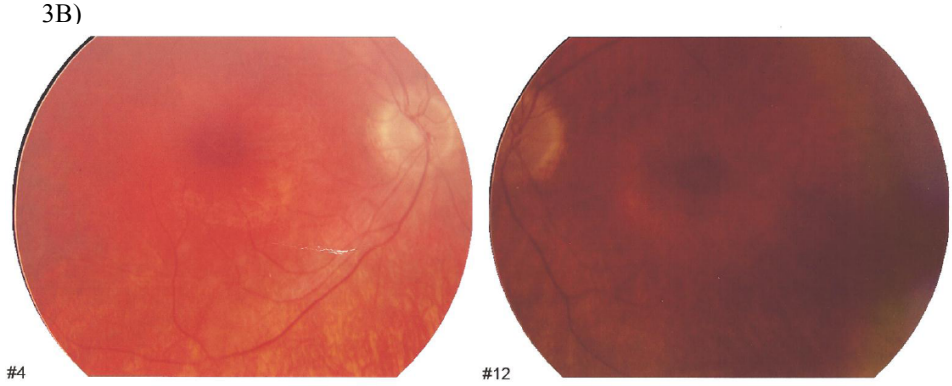


Figure 3. Ophthalmological features of the patient compound heterozygote for p.Y63X and p.R207X. **A**: Visual field, right eye and visual field, left eye. **B**: Eye fundus photography.

DISCUSSION

A total of 18 *CLRN1* mutations have been documented in USH or RP patients around the world (Human Gene Mutation Database; www.hgmd.cf.ac.uk). With the exception of the mutations p.Y176X and p.M120K found in Finnish population [18] and p.N48K found in Ashkenazi Jews [9], the rest of them were mostly found in a single family.

Besides of these three common mutations, seven indels, one nonsense and seven missense mutations have been reported to date: a deletion c.459_461delATT [18], a gross deletion of 23 bp (c.187_209del23bp) that led to a frameshift at codon 63 and a stop codon 25 amino acids forward, a p.Y63X stop mutation [9], a c.165delC deletion and a small indel mutation c.145_152delCAGGinsTGTCCAAT [10], a c.149-152delCAGG [16], a c.181delA [22] and a c.502dupA insertion [23], the c.161T>C (p.L54P) [24], the c.449T>C (p.L150P) [10], the c.118T>G (p.C40G) [21], the c.313T>C (p.S105P) [16] and the c.368C>A (p.A123D) [25], Finally, Khan et al. [26] found two homozygous missense mutations, c.92C>T (p.P31L) and c.461T>G (p.L154W) in two consanguineous Pakistani families with isolated RP and no audiological symptoms.

Two of the mutations found in this study are truncating, the p.Y63X mutation that has been found in two unrelated families being the carriers of Catalonian origin and p.R207X, found in heterozygous state in two unrelated families being the carriers of Basque origin. The other mutation, p.I168N is a missense and was found in a family together with the p.R207X mutation.

The physiological consequences of p.R207X and p.I168N mutations are unknown. Despite *CLRN1* was identified in 2002, little is known about the protein, its function and

domains. Clarin 1 is a membrane protein having four transmembrane domains that belongs to the family of vertebrate tetraspanins. The p.I168N mutation is located in the second extracellular loop. It affects an isoleucine residue, hydrophobic and aliphatic which is substituted by the polar hydrophilic asparagine. To date, only two mutations have been described in the second extracellular loop and they lead to truncated proteins. It is conceivable that p.I168N impairs clarin 1 in some way, maybe by altering the trafficking of the protein from endoplasmic reticulum to the plasma membrane as it has been observed for other missense mutations [26]. The mutation p.R207X leads to a clarin 1 that lacks the C-terminal domain. This C-terminal has a TMV signature and could serve as a protein binding motif (PBM) [9].

In the total Spanish series, the mutation p.Y63X was observed in 9 alleles of affected patients from 3 different families. The p.R207X mutation was detected in 3 alleles of affected patients from 2 families. A common haplotype was seen for p.R207X but not for p.Y63X (figure 4).

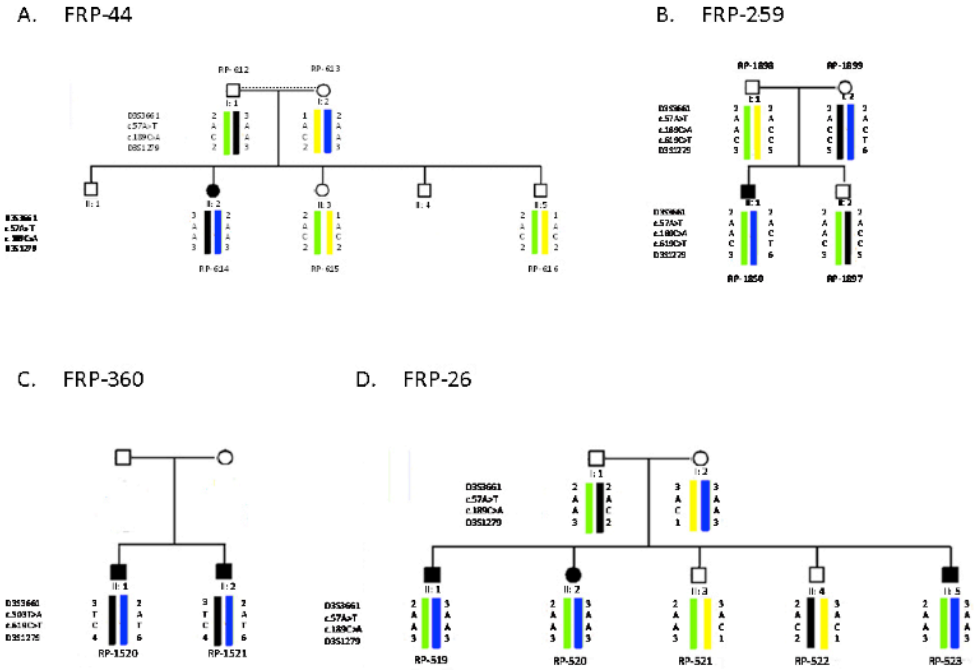


Figure 4. Pedigrees and haplotypes from four Spanish families with *CLRN1* mutations. **A**: Patient homozygote for c.189C>A (p.Y63X). **B**: Patient compound heterozygote for c.189C>A (p.Y63X) and c.619C>T (p.R207X). **C**: Patients compound heterozygotes for c.619C>T (p.Y63X) and c.503T>A (p.I168N). **D**: Patients homozygotes for c.189C>A (p.Y63X) initially described in [9].

The classical ophthalmological features of USH3 were defined by Joensuu [13] as night-blindness, mid peripheral visual field defects and slowly progressive tunnel function vision at a mean age of 30 years. Fundus appearance was typical of RP with thin vessels and granular pigmentation in the midperiphery retina. The progression of RP resulted in a severe visual handicap at 20-30 years. The diagnosis of RP is made usually at the mean age of 17 years. Later on, Plantiga et al. [27] found a visual acuity less than 0.05, severely impaired at mean age 37 years old, and tubular vision without peripheral islands at 30 years old. The visual deterioration rate was also similar to USH2A and USH1B, maybe with a more rapid progression in USH3A than USH2A, especially concerning visual field loss. Retinal degeneration pattern and structural changes in USH3A were similar to those in USH2A [28].

In the patients reported here the mean age of diagnosis of RP was 11.5 years. Nevertheless, patient RP1521 from family FRP360 was diagnosed at the time he noticed his first symptoms (eight years old), probably, because the parents and the ophthalmologist were aware of the disease of his brother RP1520. Visual acuity was poor except for the younger patient, RP1850 who is 19 years old but she refers a considerable loss of vision in the last year. All the patients displayed nigh blindness and visual field loss as the first symptoms and the eye fundus was typical of RP. Cataracts were present in two patients (RP614 and RP1520), RP1850 has not developed cataracts at the age of 19 as well as RP1521 at the age of 50 years old while is elder brother had cataracts in both eyes.

Hearing loss was prelingual, severe and progressive and vestibular dysfunction was present except in patient RP1520 who had a postlingual moderate deafness with normal vestibular reflexes. These findings and especially the clinical features of the two affected sibs of FRP360 illustrates the impressive wide spectrum of sensorineural hearing impairment (SNHI) in type and degree and the high degree of intersubject and intrafamiliar variability due to *CLRN1* mutations as previously reported [14]

The variable phenotype may cause USH3 to be under-diagnosed and mutations in *CLRN1* might be more prevalent than previously thought. Nevertheless, the screening of *CLRN1* indicates that this gene is responsible for a low percentage of cases in our Spanish USH cohort.

Acknowledgements

We acknowledge the Usher patients, their families and FARPE (Federación de Afectados por Retinosis Pigmentaria de España) for their collaboration. This work has been supported by PI08/90311 and PI10/01825, from the Spanish Ministry of Economy and Competitiveness and AP112/11 from the Conselleria de Sanidad de la Comunidad Valenciana. Gema García-García and María José Aparisi are recipient of a fellowship from the Spanish Ministry of Education (REF: AP2008-02760 and AP2009-3344 respectively). Regina Rodrigo and Carmen Espinós have a researcher-contract SNS Miguel Servet (CP09/118 and CP08/53 respectively) from the Institute of Health Carlos III. CIBERER is an initiative of the Institute of Health Carlos III, Ministry of Economy and Competitiveness.

Annexe Table 1. Clinical description of the patients in whom mutations in the *CLRN1* gene were found

Patient	Mutations	Sex/Year of Birth	Diagnosis	Age of diagnosis	Sensorineural Hearing Loss	Vestibular Function	Onset of Night Blindness	Onset of Visual Field Loss	Visual Field	Visual Acuity	Eye Fundus	ERG	Cataracts
FRP44	PY63X/ p.Y63X	F/1977	USH3	14	Prelingual severe progressive	Not available	14	14	3	0,1/ 0,1	1	Not available	BE
FRP259	PY63X/ p.R207X	F/1992	USH3	9	Prelingual, severe Progressive	Vestibular Dysfunction[A]	9	9	3	0,4 / 0,4	1	No response	No
FRP360 (II-1)	p.R207X/ p.I168N	M/ 1959	USH3	15	Postlingual moderate progressive	Normal	6	7	3	0.05/ 0.05	1	No response	BE
FRP360 (II-2)	p.R207X/ p.I168N	M/ 1961	USH3	8	Prelingual severe progressive	Vestibular Dysfunction	8	8	3	0.2/0.3	1	No response	No

Eye fundus: 1. Peripheral bone spicules, attenuated vessels, pale optic nerve · Visual field: normal: 0/ peripheral constriction or diffuse scotoma: 1/ anular scotoma or tubular field with islotes: 2/ Tubular central: 3/ islots or absolute scotoma: 4

REFERENCES

1. Kimberling WJ, Möller C. Clinical and molecular genetics of Usher syndrome. J Am Acad Audiol. 1995; 6:63-72.
2. Kimberling WJ, Hildebrand MS, Shearer AE, Jensen ML, Halder JA, Cohn ES, Weleber RG, Stone EM, Smith RJH. Frequency of Usher syndrome in two pediatric populations: implications for genetic screening of deaf and hard of hearing children. Genet Med. 2010; 12:512-16.
3. Saihan Z, Webster AR, Luxon L, Bitner-Glindzicz M. Update on Usher syndrome. Curr Opin Neurol 2009; 22:19-27.
4. Millán JM, Aller E, Jaijo T, Blanco-Kelly F, Gimenez-Pardo A, Ayuso C. An update on the genetics of usher syndrome. J Ophthalmol 2011; 2011:417217.
5. Reiners J, Nagel-Wolfrum K, Jürgens K, Märker T, Wolfrum U. Molecular basis of human Usher syndrome: deciphering the meshes of the Usher protein network provides insights into the pathomechanisms of the Usher disease. Exp Eye Res. 2006; 83:97-119.
6. Kremer H, van Wijk E, Märker T, Wolfrum U, Roepman R. Usher syndrome: molecular links of pathogenesis, proteins and pathways. Hum Mol Genet. 2006; 15 Spec No 2:R262-70.
7. Maerker T, van Wijk E, Overlack N, Kersten FF, McGee J, Goldmann T, Sehn E, Roepman R, Walsh EJ, Kremer H, Wolfrum U. A novel Usher protein network at the periciliary reloading point between molecular transport machineries in vertebrate photoreceptor cells. Hum Mol Genet. 2008; 17:71-86.
8. Västinsalo H, Jalkanen R, Dinculescu A, Isosomppi J, Geller S, Flannery JG, Hauswirth WW, Sankila EM. Alternative splice variants of the USH3A gene Clarin 1 (CLRN1). Eur J Hum Genet. 2011; 19:30-5.
9. Adato A, Vreugde S, Joensuu T, Avidan N, Hamalainen R, Belenkiy O, Olender T, Bonne-Tamir B, Ben-Asher E, Espinos C, Millán JM, Lehesjoki AE, Flannery JG, Avraham KB, Pietrokovski S, Sankila EM, Beckmann JS, Lancet. USH3A transcripts encode clarin-1, a four-transmembrane-domain protein with a possible role in sensory synapses. Eur J Hum Genet. 2002; 10:339-50.
10. Fields RR, Zhou G, Huang D, Davis JR, Möller C, Jacobson SG, Kimberling WJ, Sumegi J. Usher syndrome type III: revised genomic structure of the USH3 gene and identification of novel mutations. Am J Hum Genet. 2002; 71:607-17.
11. Zallocchi M, Meehan DT, Delimont D, Askew C, Garige S, Gratton MA, Rothermund-Franklin CA, Cosgrove D. Localization and expression of clarin-1, the Clrn1 gene product, in auditory hair cells and photoreceptors. Hear Res. 2009; 255:109-20.
12. Geng R, Geller SF, Hayashi T, Ray CA, Reh TA, Bermingham-McDonogh O, Jones SM, Wright CG, Melki S, Imanishi Y, Palczewski K, Alagramam KN, Flannery JG. Usher syndrome IIIA gene clarin-1 is essential for hair cell function and associated neural activation. Hum Mol Genet. 2009; 18:2748-60.

13. Joensuu T. Positional cloning of the Usher syndrome type III gene (USH3). Doctoral Thesis. Department of Medical Genetics, Faculty of Medicine. University of Helsinki. 2002
14. Pennings RJ, Fields RR, Huygen PL, Deutman AF, Kimberling WJ, Cremers CW. Usher syndrome type III can mimic other types of Usher syndrome. Ann Otol Rhinol Laryngol. 2003; 112:525-30.
15. Aller E, Jaijo T, Beneyto M, Nájera C, Oltra S, Ayuso C, Baiget M, Carballo M, Antiñolo G, Valverde D, Moreno F, Vilela C, Collado D, Pérez-Garrigues H, Návea A, Millán JM: Identification of 14 novel mutations in the long isoform of USH2A in Spanish patients with Usher syndrome type II. *J Med Genet* 2006, 43:e55.
16. Sadeghi M, Cohn ES, Kimberling WJ, Tranebjaerg L, Möller C. Audiological and vestibular features in affected subjects with USH3: a genotype/phenotype correlation. Int J Audiol. 2005; 44:307-16.
17. Vastinsalo H, Isosomppi J, Aittakorpi A, Sankila EM. Two Finnish USH1B patients with three novel mutations in myosin VIIA. Mol Vis. 2006; 12:1093-7.
18. Joensuu T, Hämäläinen R, Yuan B, Johnson C, Tegelberg S, Gasparini P, Zelante L, Pirvola U, Pakarinen L, Lehesjoki AE, de la Chapelle A, Sankila EM. Mutations in a novel gene with transmembrane domains underlie Usher syndrome type 3. Am J Hum Genet. 2001; 69:673-84.
19. Cremers FP, Kimberling WJ, Külm M, de Brouwer AP, van Wijk E, te Brinke H, Cremers CW, Hoefsloot LH, Banfi S, Simonelli F, Fleischhauer JC, Berger W, Kelley PM, Haralambous E, Bitner-Glindzicz M, Webster AR, Saihan Z, De Baere E, Leroy BP, Silvestri G, McKay GJ, Koenekoop RK, Millan JM, Rosenberg T, Joensuu T, Sankila EM, Weil D, Weston MD, Wissinger B, Kremer H. Development of a genotyping microarray for Usher syndrome. J Med Genet. 2007; 44:153-60.
20. Kurg A, Tõnisson N, Georgiou I, Shumaker J, Tollett J, Metspalu A. Arrayed primer extension: solid-phase four-color DNA resequencing and mutation detection technology. Genet Test. 2000; 4:1-7.
21. Aller E, Jaijo T, Oltra S, Alió J, Galán F, Nájera C, Beneyto M, Millán JM. Mutation screening of USH3 gene (clarin-1) in Spanish patients with Usher syndrome: low prevalence and phenotypic variability. Clin Genet. 2004; 66:525-9.
22. Dreyer B, Brox V, Tranebjaerg L, Rosenberg T, Sadeghi AM, Möller C, Nilssen. Spectrum of USH2A mutations in Scandinavian patients with Usher syndrome type II. Hum Mutat. 2008; 29:451.
23. Ebermann I, Wilke R, Lauhoff T, Lübben D, Zrenner E, Bolz HJ. Two truncating USH3A mutations, including one novel, in a German family with Usher syndrome. Mol Vis. 2007a; 13:1539-47.
24. Isosomppi J, Västinsalo H, Geller SF, Heon E, Flannery JG, Sankila EM. Disease-causing mutations in the CLRN1 gene alter normal CLRN1 protein trafficking to the plasma membrane. Mol Vis. 2009; 15:1806-18.

25. Ebermann I, Lopez I, Bitner-Glindzicz M, Brown C, Koenekoop RK, Bolz HJ. Deafblindness in French Canadians from Quebec: a predominant founder mutation in the USH1C gene provides the first genetic link with the Acadian population. Genome Biol. 2007b; 8:R47.
26. Khan MI, Kersten FF, Azam M, Collin RW, Hussain A, Shah ST, Keunen JE, Kremer H, Cremers FP, Qamar R, den Hollander AI. CLRN1 mutations cause nonsyndromic retinitis pigmentosa. Ophthalmology. 2011; 118:1444-8.
27. Plantinga RF, Pennings RJ, Huygen PL, Sankila EM, Tuppurainen K, Kleemola L, Cremers CW, Deutman AF. Visual impairment in Finnish Usher syndrome type III. Acta Ophthalmol Scand. 2006; 84:36-41.
28. Herrera W, Aleman TS, Cideciyan AV, Roman AJ, Banin E, Ben-Yosef T, Gardner LM, Sumaroka A, Windsor EA, Schwartz SB, Stone EM, Liu XZ, Kimberling WJ, Jacobson SG. Retinal disease in Usher syndrome III caused by mutations in the clarin-1 gene. Invest Ophthalmol Vis Sci. 2008; 49:2651-60

5. Additional Studies

5.1 Large rearrangements in the *USH2A* gene

Large deletions or duplications have been described to be responsible for Usher syndrome (LE GUEDARD *et al.* 2007; ALLER *et al.* 2010a; ROUX *et al.* 2011). Hence, it is necessary to perform specific analyses to detect this type of mutations. In our cohort several patients remained without mutations or some others presented only one clear pathological mutation after mutational screening of the *USH2A* gene. These patients might carry large rearrangements, which are not detected by routinely PCR techniques. Hence, the presence of this type of mutations within these patients was studied whenever DNA remained available after our preliminary studies.

Eleven subjects previously studied for the *USH2A* mutations were included, 10 with one pathologic mutation. The 11th patient, RP1622, showed no point mutations, but all polymorphisms detected were in homozygous state and exon 14 was impossible to amplify, strongly suggesting a possible deletion in this region.

Patients were analysed using a customized array-CGH (array comparative genomic hybridization) chip (12x135k) developed in the laboratory of Anne-Françoise Roux, which includes all the Usher syndrome genes currently known (with the exception of *CIB2*) (ROUX *et al.* 2011). The design of the array includes thousands of probes covering the genes and 10.000 nucleotides of 5' and 3' UTRs (untranslated regions). The technique was performed in a high-resolution microarray platform following manufacturer's instructions (Nimblegen, Roche Diagnostics).

Next step consisted in the identification of deletion breakpoints. One pair of primers was designed to specifically amplify patient's genomic DNA in the region where deletion breakpoints were indicated by the array-CGH analysis. The amplified PCR products were directly sequenced. The genomic *USH2A* reference sequence used for nomenclature was NG_009497.1.

We found two large deletions in our samples (Figure 16) (data not published). We detected a large deletion encompassing exons 5 to 9 in a patient (RP1638) in whom we previously identified the mutation c.5540_5541dup (Figure 16A). We also detected a homozygous deletion of exon 14 in patient RP1622 (Figure 16B), in whom we did not identify any mutation after *USH2A* screening.

To determine the exact position of deletions breakpoints the primers were located upstream and downstream of last deleted probe in each extreme of deletion, indicated by CGH-array. For patient RP1638, the forward primer was designed in intron 4 and reverse primer in intron 9. For patient RP1622, forward primer was located in intron 13 and reverse primer in intron 14.

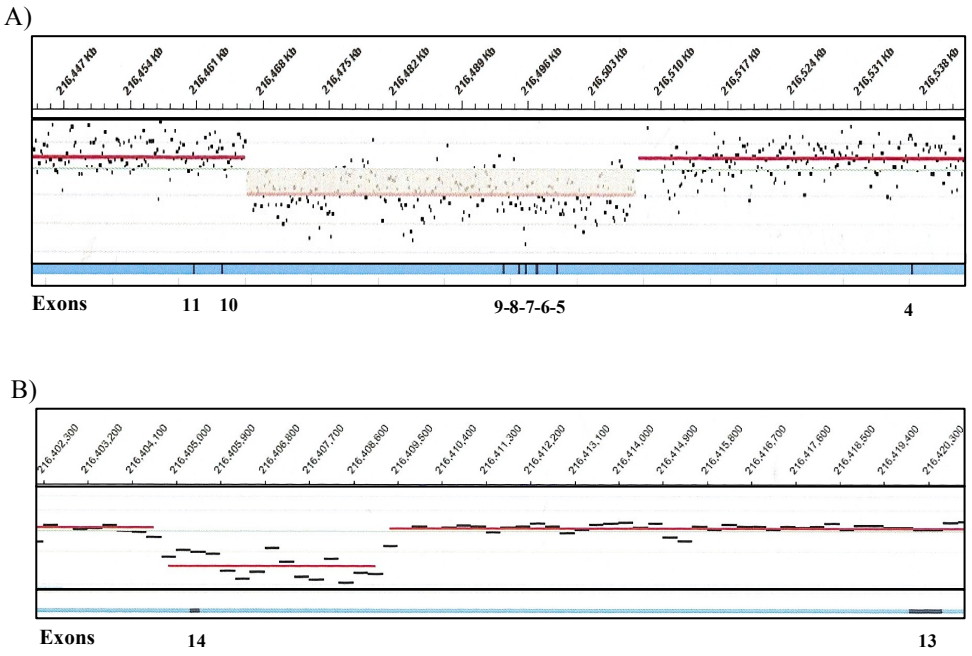


Figure 16. Representation of deletions detected in *USH2A* using array-CGH analysis. A) Deletion of exons 5-9 in subject RP1638. B) Deletion of exon 14 in patient RP1622. Blue lines at the bottom of each diagram show intronic DNA. Black boxes within these blue lines represent exons. Exon numbers are shown below each black box. Black coloured points/lines represent the probes used in this study. Red line represents the ratio of the mean of probes joined to the sample of the patient respect to the healthy control.

In figure 17, we show the sequences surrounding the deletion breakpoint. In patient RP1638, the exact breakpoints were located in position g.79625 (intron 4) and g.135433 (intron 9) and the deletion spanned more than 55Kb. Exons 5 to 9 were deleted, a total of 848 coding nucleotides. In patient RP1622, the breakpoints were positioned in g.192139 (intron 13) and g.192161 (intron 14), which shows that there is a deletion of more than 5Kb in this region. This deletion produces the loss of 172 coding nucleotides corresponding to exon 14.

Both of them are out-of-frame deletions and would lead to premature stops codons, with the subsequently truncated protein.

The four intervals surrounding the breakpoints were analysed through Repeat Masker (http:// www.repeatmasker.org/). This analysis showed the presence of repetitive elements such as short interspersed elements (SINE), long interspersed elements (LINE), long terminal repeats (LTR) or simple repeats, were found in all cases. An accurate analysis of the breakpoint regions showed the dinucleotide TA (in patient RP1638) and trinucleotide

ATA (in patient RP1622) immediately adjacent to the deletion breakpoints. Both are oligonucleotides over-represented in GRaBD (Gross Rearrangement Breakpoint Database) in deletion breakpoint junction (ABEYSINGHE *et al*. 2003). Furthermore, surrounding intronic regions of breakpoints detected in our patients are AT-rich, which is a typical feature of deletion breakpoints (ABEYSINGHE *et al*. 2003). Our results are in agreement with the previous findings in studies of deletions/duplications of other USH genes, as P*CDH15* (LE GUEDARD *et al.* 2007; ALLER *et al.* 2010a), or of the dystrophin gene in DMD (Duchenne Distrophy Muscular) (GUALANDI *et al.* 2006).

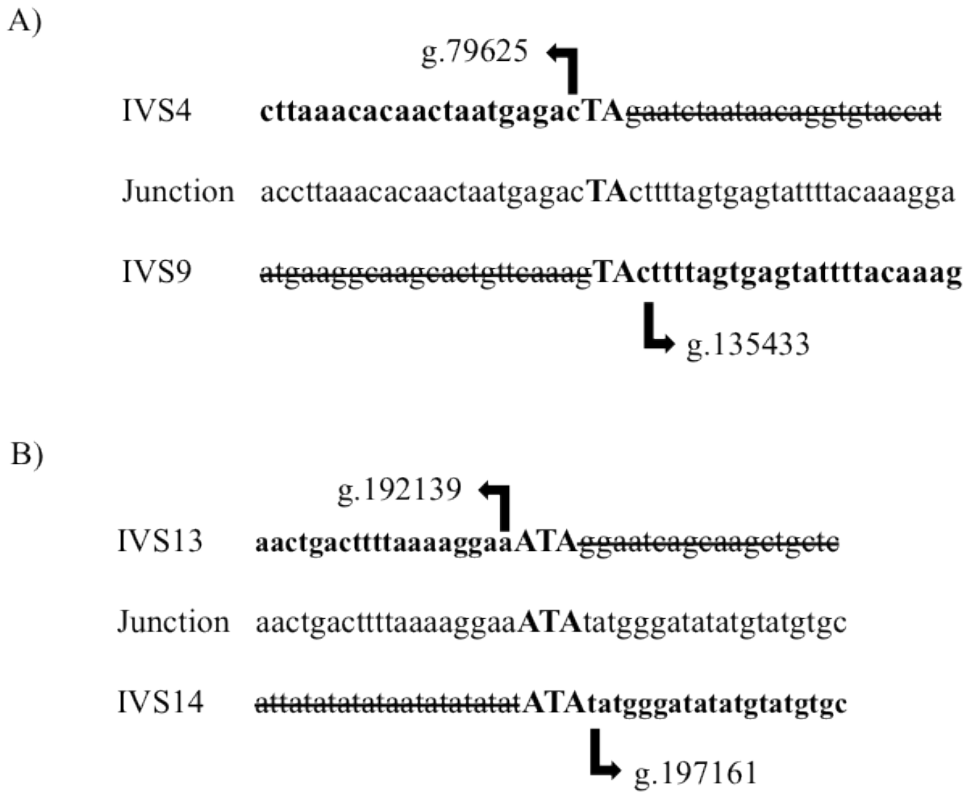


Figure 17. Schematic localization of the deletion breakpoints within the *USH2A* gene. The junction fragments identified in patients RP1638 (A) and RP1622 (B) are aligned with the corresponding wild-type sequences spanning the 5' and 3' breakpoints. The deleted sequences are crossed out. The nucleotides located directly adjacent to the breakpoints are shown in bold.

We have identified large deletions in 2 out of the 11 patients analysed. Because of the large size of the *USH2A* gene and the presence of large intronic regions (only 19% of the gene sequence is coding) is not surprisingly that large rearrangements are detected.

5.2 Deep intronic *USH2A* mutation

The presence of deep intronic variants, which modify the normal splicing, had already been described in genes related to other diseases (TUFFERY-GIRAUD *et al.* 2003; SPENA *et al.* 2007; PEREZ *et al.* 2010), but not for Usher syndrome. It was in 2011 when Vaché *et al*. reported the first case in Usher syndrome with a deep intronic mutation causing an activation of a pseudoexon (VACHE *et al.* 2012). The article is attached in Annexes.

The work was carried out in a large family with three members affected of Usher syndrome type 2. Haplotype analysis, mutational screening for *USH2A* gene and array-CGH for large rearrangements were performed for the three subjects, but only one mutation could be identified. However, haplotype analysis showed a common haplotype for the three patients, suggesting they should share a second mutation not detected using the conventional technics. Taking together these results, the authors proposed that the second mutational event could occur in *USH2A* deep intronic sequences.

To test this hypothesis we decided to explore the splicing status of the *USH2A* gene within the Usher patients. Obtaining RNA of *USH2A* from retinas or inner ear from patients is challenging for obvious reasons. However, it has been described that nasal cells expresses detectable levels of *USH2A* mRNA, and these cells are easily accesible, without harsh invasive techniques. Hence, we extracted total RNA from these cells from one of the three affected subjects and we analysed the mRNA of *USH2A*. A total of 14 overlapping RT-PCRs were necessary to study the entire *USH2A* sequence. The same study was performed in 12 controls.

In the fragment encompassing the exons 39-44, an additional and larger fragment was detected which was not observed in healthy controls. Direct sequencing of this fragment confirmed the existence of two populations of transcripts, that we called normal and aberrant transcripts. The aberrant transcript included an insertion of 152 nucleotides between the exons 40 and 41 (Figure 18). The inclusion of this out-of-frame sequence produced a direct stop codon in exon 41, generating a truncated protein of 2671 amino acids. This protein would lack of transmembrane and intracellular domains and could not be anchored to the membrane. The next step was to ascertain the origin of this additional sequence. To do so we performed Blast analysis, which determined that this sequence belongs to an intronic region localised at -2996 base pairs from exon 41 (Figure 18).

Then, *in silico* analysis of these 152 nucleotides and the surrounding region were performed using HSF (Human Splicing Finder). It revealed that this fragment had a strong 3' splice site, but a weak 5' splice site, due to the presence of an adenine in position +1. So, it would be possible that the inclusion of these 152 nucleotides extra was due to the presence of a mutation that reinforced the splicing site in 5' and produced the activation of the pseudoexon.

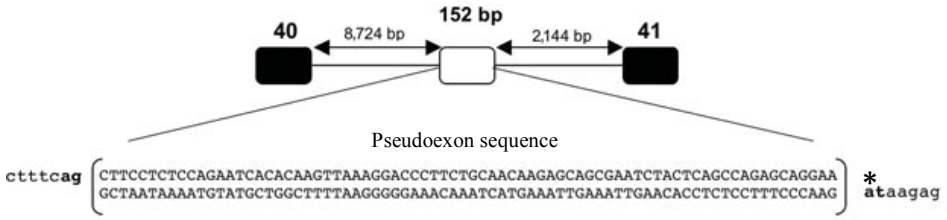


Figure 18. Representation of the pseudoexon sequence at DNA level. Black boxes correspond to exons and white boxe to pseudoexon. *Nucleotide in position +1, "a" in wild-type sequence and "g" in mutated sequence. Adapted from (VACHE *et al.* 2012).

Subsequently, to confirm the hypothesis, the DNA sequence of this potential pseudoexon and flanking regions were amplified. A nucleotide change (A>G) in position +1 of the inserted sequence was found in the patients but not in the 12 controls. This change increased hugely the strength of the 5' splice site from 2,39 to 10,57 according to the Maximum Entropy (MaxEnt) software, necessary for the pseudoexon activation. The mutation was not described in dbSNP database, not detected in 338 alleles and it segregated with the disease in the family. All these results supported the idea it was the second mutation responsible of the disease. To reinforce this hypothesis, *in vitro* analyses were performed using the pSPL3 vector trapping and HELA cells. The results showed to be nicely coherent with the data previously obtained.

To study the prevalence of this new variation, we screened it in 20 French USH2 patients having 0 or 1 *USH2A* mutations. This analysis revealed 4 additional cases with the variation. To know if the variation was also so frequent in others populations, we performed the analysis in eighteen Spanish patients with only one mutation detected in *USH2A* and this intronic mutation was present in 4 patients. The study was extended to 16 subjects that remained without mutations after screening of the *USH2A* gene, but the result was negative. The variation was also studied in additional 90 controls of Spanish origin. In none of the cases the variation was detected.

In summary, the variation was detected in heterozygous state in a total of 8 additional patients carrying only a single mutation previously detected in *USH2A*.

This study highlights the importance of the analyses at RNA level to detect variations that otherwise would be missed with the routinely screenings. Moreover, the novel mutation detected is quite frequent in French and Spanish population and it would be interesting to include this mutation in future *USH2A* screenings. Furthermore, this type of intronic mutations has already been corrected using antisense oligonucleotides (AONs), opening the possibility to set up a new therapy for Usher syndrome.

CHAPTER IV

Application of next generation sequencing to molecular diagnosis of USH

1. Context

The high genetic heterogeneity of Usher syndrome makes the molecular diagnosis of this disease very complex. Patients with the same USH clinical type are indistinguishable independently of which is the gene responsible and, sometimes, there are patients with atypical clinical features or patients that remain without classification due to lack of clinical data. Additionally, if in the most of cases clinical diagnosis is in line with the causative gene, in some others, patients carry mutations in a gene responsible for a different clinical subtype.

We have used Sanger sequencing, coupled with array-CGH to identify mutations in USH patients. However, comprehensive sequencing of the many genes involved in the disease according to their implication is impractical with traditional Sanger methods.

In the last years, next generation sequencing (NGS) has emerged as a new tool in the clinical laboratory for the rapid and cost-effective detection of mutations in genetically heterogeneous disorders. Several sequencing platforms (Illumina, Solid, 454 Roche, Ion Torrent…) and enrichment methods (PCR amplification, Long-PCR amplification, "in-solution capture", microdroplet-based PCR amplification) have been described and used in different studies. We have chosen the combination of "in solution" targeted capture (Nimblegen, Roche) of Usher genes and additional genes, with massively parallel sequencing, using GS Junior platform (454, Roche).

A total of 71 patients divided in three groups were analysed with this approach. Forty-seven patients previously studied for at least one Usher gene formed the control group. In this group, we included the patients who remained with 0 or 1 mutation after mutational screening of the USH2 genes in previous studies (GARCIA-GARCIA *et al*. 2011; GARCIA-GARCIA *et al*. 2013). The second group was defined by 13 new Usher patients and the third one comprised 11 patients with non-syndromic hearing loss.

2. Article:

Usher syndrome: experience of targeted exome sequencing as a clinical test

Thomas Besnard *[1,2], **Gema García-García** *[1,3], David Baux *[4], Christel Vaché [4], Valérie Faugère [4], Lise Larrieu [4], Susana Léonard [4], Jose M. Millan [3], Sue Malcolm [5], Mireille Claustres [1,2,4], Anne-Françoise Roux [1,4]

***The authors contributed equally to this work**

1. Inserm, U827, Montpellier, F-34000, France

2. Univ, Montpellier I, Montpellier, F-34000, France

3. Grupo de Investigación en Enfermedades Neurosensoriales, Instituto de Investigación Sanitaria IIS-La Fe and CIBERER, Valencia, Spain

4. CHU Montpellier, Laboratoire de Génétique Moléculaire, Montpellier, F-34000, France

5. Clinical and Molecular Genetics, Institute of Child Health, University College London, London, United Kingdom

Molecular Genetics & Genomic Medicine. Submitted. **2013**

ABSTRACT

We show that massively parallel targeted sequencing (NGS) of 19 genes provides a new and reliable strategy for molecular diagnosis of Usher syndrome and non syndromic deafness, particularly appropriate for these disorders characterized by a high clinical and genetic heterogeneity and a complex structure of several of the genes involved. A series of 71 patients including Usher patients previously screened by Sanger sequencing plus newly referred patients was studied. Ninety-eight percent of the variants previously identified by Sanger sequencing were found by NGS. NGS proved to be efficient as it offers exhaustivity that is laborious to reach with Sanger sequencing. Among the 13 newly referred Usher patients, both mutations in the same gene were identified in 77% of cases (10 patients) and one candidate pathogenic variant in two additional patients. This work can be considered as pilot for implementing NGS for genetic heterogeneous diseases in clinical service.

Key Words: Usher syndrome, NSHL, Next generation sequencing, Variant prioritization, Bioinformatics

INTRODUCTION

Usher syndrome (USH) is an autosomal recessive disorder with a prevalence of at least 5/100,000 characterized by the association of sensorineural hearing loss (HL) and visual impairment due to retinitis pigmentosa (RP). Some patients will also exhibit vestibular areflexia (VA). USH is the most common form of deaf-blindness (Saihan, et al., 2009). Three clinical subtypes (USH1, USH2 and USH3) are distinguished depending on the severity and progression of HL and presence or absence of VA and this distinction is generally used to guide molecular diagnosis. USH1 is the most severe form with congenital profound HL and VA. USH2 appears to be the most common clinical form of the disorder, accounting for over a half of all USH cases and is characterized by congenital moderate-to-severe HL, with normal vestibular function. In USH3, the HL is progressive and leads to variable vestibular function. USH3 is rare except in some populations with recognized founder effects where it can responsible for more than 40% of the Finnish and Jewish Ashkenazi USH cases (Saihan, et al., 2009).

This clinical heterogeneity is accompanied by high genetic heterogeneity. To date a minimum of eleven genes responsible for the disease are known. Five USH1 genes have been extensively studied: *MYO7A* (USH1B), *CDH23* (USH1D), *PCDH15* (USH1F), *USH1C* (USH1C) and *USH1G* (USH1G), with *MYO7A* being the most prevalent (Le Quesne Stabej, et al., 2012; Roux, et al., 2011). A sixth gene, *CIB2* (USH1J), has been very recently reported in a single family (Riazuddin, et al., 2012). Among the three identified USH2 genes, *USH2A* (USH2A), *GPR98* (USH2C) and *DFNB31* (USH2D), *USH2A* has been shown to be responsible for 70-80% of the USH2 cases (Besnard, et al., 2012; Le Quesne Stabej, et al., 2012). Until recently, *CLRN1* (USH3A) was the only gene known responsible for USH3 (Joensuu, et al., 2001), but the *HARS* gene was recently proposed as a novel USH3 gene (Puffenberger, et al., 2012). In addition to those, a twelfth gene, *PDZD7*, was shown to contribute to USH2 as a modifier of the retinal phenotype on a *USH2A* background or in digenic inheritance with *GPR98* (Ebermann, et al., 2010). Multiple isoforms have been described for most of these genes. Among them, *USH1C*, *PCDH15*, *USH2A* isoforms have been well characterised (Ahmed, et al., 2008; Bitner-Glindzicz, et al., 2000; van Wijk, et al., 2004; Verpy, et al., 2000).

Mutations in seven of these genes (*MYO7A*, *USH1C*, *CDH23, PCDH15, DFNB31* and *CIB2*) can also cause non-syndromic hearing loss (NSHL) and mutations in *USH2A* and *CLRN1* give rise to isolated autosomal recessive RP (see Retinal and hearing impairment genetic mutation database, which includes USHbases and other NSHL genes: https://grenada.lumc.nl/LOVD2/Usher_montpellier/). Recently, a mutation in the short isoform of *USH1C* has been shown to be associated with RP and late-onset deafness (Khateb, et al., 2012).

Molecular genetic diagnosis for Usher syndrome evolved from the scanning of restricted portions of USH genes (Adato, et al., 1997) to extensive direct sequencing (Aller, et al., 2006; Baux, et al., 2007; Besnard, et al., 2012; Bonnet, et al., 2011; Dreyer, et al., 2008; Garcia-Garcia, et al., 2011; Le Quesne Stabej, et al., 2012; Roux, et al., 2006; Roux, et al., 2011). Because of the genetic heterogeneity, prioritization of the genes to be

sequenced was facilitated by linkage analysis (Roux, et al., 2006; Roux, et al., 2011). Due to the large size of most Usher genes (totalizing more than 350 exons), Sanger sequencing of genes one-by-one remains expensive and time-consuming. Furthermore, large rearrangements have been described in *MYO7A, CDH23, GPR98, USH2A* and, particularly, in *PCDH15,* and their detection requires array-CGH studies (see USHbases). Taken together, these strategies allow a reliable diagnosis for Usher patients with a mutation detection rate of about 90% for USH1 and USH2 patients (Besnard, et al., 2012; Roux, et al., 2011). A genotyping microarray commercially available (Cremers, et al., 2006) allows rapid epidemiological screening of hundreds of previously identified variations in nine USH genes (Vozzi, et al., 2011). However its application in clinical diagnosis is hampered by a very low detection rate as most USH-causing DNA alterations are private or restricted to one or two families (see USHBases).

NGS technology has recently demonstrated its capacity to detect DNA variants in sensorineural disorders known to be genetically heterogeneous (Brownstein, et al., 2011; Neveling, et al., 2012; Redin, et al., 2012), and a targeted NGS protocol assayed on 9 samples provided technical performances compatible with potential use as a diagnostic platform applied to hearing loss (HL) (Shearer, et al., 2010). A recent study applied to Usher syndrome compared two different enrichment methods and reported a higher efficiency in mutation detection by using a Long-Range PCR targeted approach as compared with whole exome capture (Licastro, et al., 2012).

Here, we have designed a NGS-based workflow using a solution-based capture method, that we applied to an important cohort of patients (n = 71) with the aim to rigorously evaluate the feasibility of NGS for the screening of Usher genes in a clinical diagnostic setting. Forty-seven Usher patients previously partially positive or negative for USH genes analyzed by standard Sanger sequencing, were used as a test cohort to establish criteria and thresholds for accurate generation and filtration of the data, as well as prioritization and annotation of the variants, and calculation of analytical sensitivity. The validated protocol was then applied to 13 newly referred Usher patients and to 11 NSHL patients.

MATERIAL AND METHODS

Patients

A total of 71 subjects (21 Spanish and 50 French), classified on the basis of their clinical history and ophthalmologic, audiometric and vestibular tests, were enrolled in this study. Audiograms from patients presenting with NSHL were collected and confirmed profound HL. The local Ethics Committee approved molecular analyses and consent to genetic testing was obtained from adult probands or parents in the case of minors. DNA was extracted from blood samples and quality and quantity assessed by using standard techniques.

Test Sample

Among the 47 patients included in this group, 7 were classified as USH1, 34 as USH2 and 2 as USH3 (Table S1). Four of them could not be classified because of lack of clinical data. All patients had previously been studied for at least one Usher gene by Sanger sequencing (Table S1) that led to the identification of one putative causative mutation in 22 of them. The remaining 25 patients had no identified pathogenic mutation in any of the genes screened by Sanger sequencing.

Diagnostic Sample

Usher Group

Thirteen patients were included in this group. No molecular study had been performed prior NGS. Five of them were considered to be USH1, 7 to be USH2 and one to be USH3.

NSHL group

Eleven patients presenting with NSHL were selected. A genetic origin of deafness was suspected based on the absence of any environmental or infectious cause, or presence of familial cases or documented consanguinity. All had been previously screened for mutations at the DFNB1 locus and one of them was *GJB2* c.35delG heterozygote.

Sequence capture and sequencing

A custom solution-based sequence capture manufactured by Roche Nimblegen (SeqCap EZ Choice Library) included a total of 634 exons (and 100 bp of the flanking intronic regions) from 19 genes (9 known Usher genes, 2 candidate Usher genes (*PDZD7* and *VEZT*), 7 NSHL genes and *CHM* (REP-1) gene), and the 5' and 3' untranslated regions. All annotated transcripts were included. The design also included the intronic *USH2A* region encompassing the pseudoexon recently described (Vaché, et al., 2012). The entire custom design spanned 326 kb. The final capture size was 364 kb covered by more than 32,000 different biotinylated probes. By merging the overlapping regions, the design encompassed 535 different regions.

Sequence Capture was performed according to the User's Guide "NimbleGen SeqCap EZ Library LR" (Version 2.0, November 2011). DNA libraries were prepared following the instructions from the manufacturer (GS FLX Titanium Rapid Library Preparation Method Manual, January 2010). Genomic DNA (500 ng) was sheared by fragmentation (with a majority of fragments between 400-700 pb). Fragments were end repaired, A-tailed and ligated to the adapters. Small fragments were removed using Agencourt AMPure XP beads (Beckman Coulter, Agencourt, Beverly, MA, USA). The libraries were amplified for 12 cycles by pre-captured ligation-mediated polymerase chain reaction (pre-capture LM-PCR) with primers specific for the adaptors. The amplified libraries were then hybridized to the designed biotinylated probes for 66-72 hours at 47°C.

The biotinylated probes-DNA hybrids were purified with streptavidin-conjugated magnetic beads and washed. Finally, the captured DNA fragments were eluted/recovered and amplified for 15 cycles (post-capture LM-PCR). The final concentration of each captured library was calculated with a Qubit fluorometer and diluted at 10^7 molecules/µl. Emulsion PCRs were performed according to the manufacturer's instructions (emPCR Amplification Method Manual Lib-L GS Junior Titanium Series, May 2010, Rev. April 2011). The ratio 1 molecule per bead was chosen as input to perform the emPCRs. Sequencing of each library was carried out on Roche GS Junior sequencer according to the manufacturer's protocol (Sequencing Method Manual GS Junior Titanium Series, May 2010, Rev. June 2010).

Bioinformatics pipeline and prioritization

Assembly, coverage and variant calling

Sequence reads were mapped against the human chromosomes reference (hg19) using the GS Reference Mapper software (Roche, version 2.6 and 2.7). Average depth of coverage (aDOC) for each region was calculated by dividing the sum of the DOC per base within the specific regions by the total region size (in base pairs).

$$\text{aDOC} = \frac{\sum_{i=1}^{n} \text{DOC}}{\text{Region size}}$$

The percentage of *on target* was defined as a ratio between the number of bases aligned in targeted regions and the number of base pairs mapped in total.

$$\text{On target \%} = \frac{\text{bp aligned on target}}{\text{total mapped bp}} \times 100$$

Artefact variants were removed following these criteria:

i) Variants detected in less than 20% of total reads.
ii) *Indels* with a coverage >20 reads but with a disequilibrium between number of forward and reverse sequences (Fwd or Rev <10%).
iii) *Indels* located at distance from exons (more than +/-20 intronic flanking nucleotides).

The remaining variants were annotated adding gene name, known polymorphism from dbSNP131, localisation in gene, cDNA and protein nomenclature, using either Annovar (Wang, et al., 2010) or Mutalyzer (Wildeman, et al., 2008).

We developed a in-house software called "GS data online treatment" (GSdot), available at https://neuro-2.iurc.montp.inserm.fr/454/ to automate the calculations and filters described above. It uses as input files initially generated by GS Reference mapper. More details on how to use the software and the different steps can be obtained from the website.

Prioritization of variants and determination of pathogenicity

After automatic filtering performed by GSdot, all the annotated variant files generated (one per patient) were merged into a single one for manual prioritization of the variants. Prioritization consisted in keeping any known pathogenic mutation and any variant of unknown clinical significance (VUCS) identified fewer than 5 times on DNAs from the Test Sample, and localised in exons or within the 20 bp intronic regions.

The pathogenicity of the candidate variants was investigated examining the frequency in patients or controls from our internal database, the locus specific database LOVD-USHbases, the deafness variation database (http://www.deafnessvariationdatabase.com/), the exome variant server database (EVS, http://evs.gs.washington.edu/EVS/) or the genome variant database (1000 genomes, http://www.1000genomes.org/, and dbSNP, http://www.ncbi.nlm.nih.gov/snp). Specific *in silico* studies were conducted following our multistep analysis described in Roux 2011 (Roux, et al., 2011), using the Usher Syndrome Missense Analysis software (USMA, https://neuro-2.iurc.montp.inserm.fr/USMA/) and Human Splicing Finder (HSF (Desmet, et al., 2009), http://www.umd.be/HSF/), which includes two distinct algorithms, namely HSF and MaxEnt ((Yeo and Burge, 2004)). In order to assess the impact of missense candidates in non-Usher genes, ortholog and domain alignments were studied from UCSC (http://genome.ucsc.edu/) and Prosite (http://prosite.expasy.org/), respectively.

All likely pathogenic variants were confirmed by Sanger sequencing, and familial segregation analyses were performed whenever possible. The latter contribute to classification of the VUCSs as already described (Baux, et al., 2013; Roux, et al., 2011).

Ex vivo splicing assay

DNA from U1157 was used as template in a PCR amplification including exons 62 to 65 of *CDH23* with the High Fidelity Phusion Polymerase (Finnzymes, Espoo, Finland). Amplicons were inserted in the pSPL3 exon-trapping vector between the *NotI* and *XhoI* restrictions sites and the constructs were transfected in a human retinal pigment epithelial cells line (ARPE-19) as previously described (Guédard-Méreuze, et al., 2009). Forty-eight hours after transfection, RNA was extracted with the Nucleospin RNAII kit (Macherey-Nagel, Hoerdt, France). RT-PCR and splicing alterations analyses were carried out as described before (Le Guédard-Mereuze, et al., 2010).

RESULTS

Raw data quality

Data obtained from the Test Sample were used to evaluate the quality of raw data. The number of reads per run was estimated on average to 129,783 of which 98,149 were mapped with a mean length of 431 bp. The average amount of mappable sequence data was 53 Mb. Eighty percent of these sequences overlapped the targeted region and 52% of data were mapped on target (Figure 1).

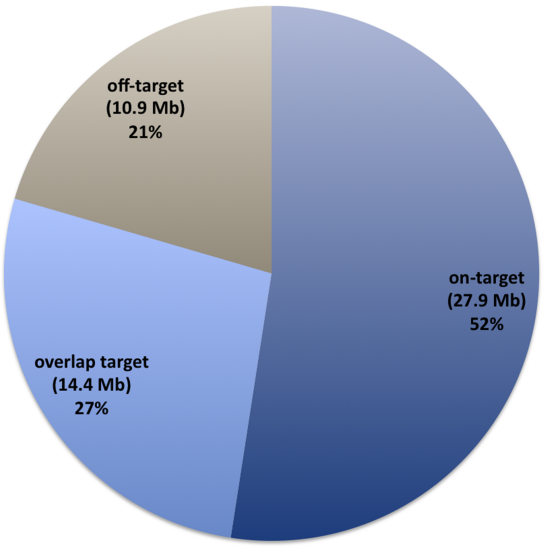


Figure 1: Distribution of mapped bases. The percentages refer to reads mapped to the targeted region (on-target), sequences partially mapped to the targeted region (overlap-target) or reads completely aligned out of region (off-target).

The overall depth of coverage was estimated to 77X across the whole design, ranging from 55.8X (*USH1G*) to 106.9X (*TECTA*) (Figure 2). Among the 634 targeted regions, only 37 (5.8%) were covered with less than 40X (22 regions between 20-40X) (Figure S1). Eighteen of these correspond to high GC-content (> 60%).

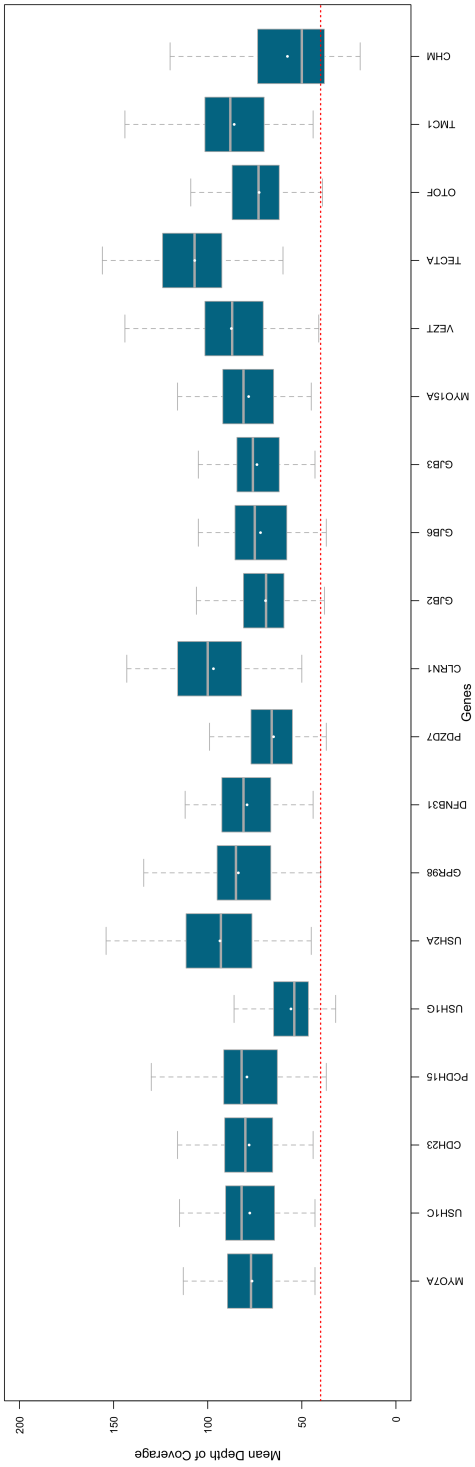


Figure 2: Average coverage for targeted genes. Red line represents a Depth of Coverage (DOC) of 40 reads, the defined limit for proper validation. The horizontal grey line corresponds to the median. Mean values are denoted by a white dot.

Filtering/Prioritization/Classification of variants

In the Test Sample, a mean of 4,674 putative variants were identified per patient, however this was reduced to 8 candidate variants per patient when the analysis pipeline was applied as described above (Figure 3). First, filtering was performed to eliminate artefacts from raw data. This task has been automated in a dedicated publicly available tool named GSDot. Then, the cohort data were used to mask likely non-pathogenic variants, i.e., when variants were present in more than 4 patients or were distant of more than 20 bp of exon boundaries. The 8 remaining Usher variants, representing 0.17% of the original pool of candidate variants, underwent specific analysis for scoring as detailed below.

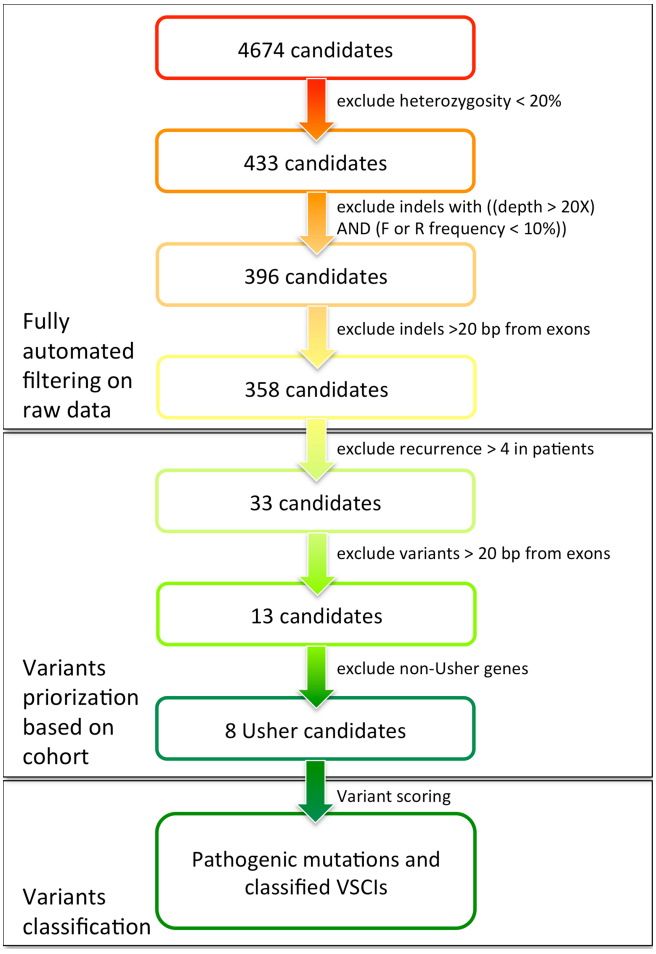


Figure 3: Pipeline designed for efficient filtering, prioritization and classification of variants. Indicated figures are related to the average of the control group. F, forward strand sequences. R, reverse strand sequences. VUCS=Variants of Unkown Clinical Significance.

Sensitivity of the strategy

To assess the analytical sensitivity of this approach, we checked whether 687 variants from 24 patients previously detected by Sanger sequencing were also detected with NGS. These variants were widespread throughout the nine Usher genes known at the time of the study plus *VEZT* (Figure 4). All these variations were located in exons or within the 20 bp adjacent intronic sequences, which is in line with the filters applied to NGS data. The variant detection rate of our NGS protocol was 98% (674/687). Among the 13 false negative variants, 6 were lying in homopolymeric regions, 4 could be visually detected but were misaligned and therefore not considered by the variant calling software provided by Roche, and 3 were localized in poorly covered regions (<40 X).

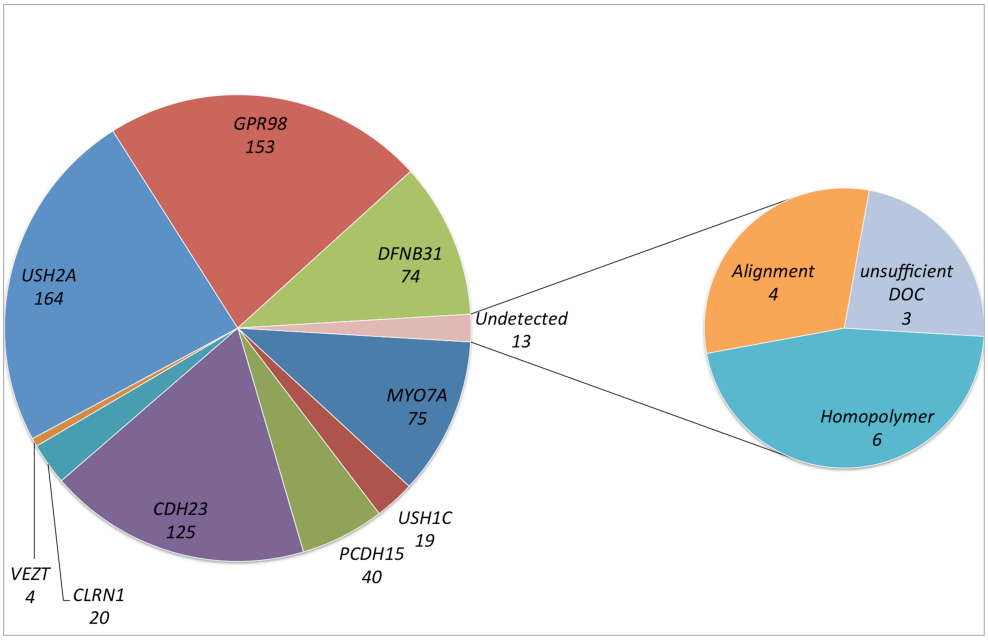


Figure 4: Analytical sensibility of targeted capture approach. The graphic shows the number of variations previously identified in each gene using Sanger sequencing that were also detected by next-generation sequencing. Undetected variants are detailed.

Identification of previously undetected variations in the Control group

In addition to the concordant variants (NGS *vs* Sanger) found in 24 patients, additional pathogenic mutations could be detected in 12 of the 47 patients of the Test Sample (Table 1).

Table 1: Mutations previously identified in Sanger sequencing and the additional detected by NGS in 12 patients from the Test Sample.

Patient	USH type	Gene	Mutations identified by Sanger	Additional mutations identified by NGS
U654	II	*USH2A*	c.5528C>T – p.(Pro1843Leu)	c.11864G>A – p.(Trp3955*) °
RP1616	II	*USH2A*		c.11864G>A – p.(Trp3955*) ° c.11864G>A – p.(Trp3955*) °
RP98	II	*USH2A*		c.14803C>T – p.(Arg4935*) °
RP1578	II	*USH2A*	c.2299delG – p.(Glu767fs) °	c.13811+2T>G – p.(?)
U996	Undef	*USH2A*	NA	c.2299delG – p.(Glu767fs) ° c.2176T>C – p.(Cys726Arg)
U286	II	*USH2A*	c.2299delG – p.(Glu767fs) °	
		GPR98	NA	c.10458G>A – p.(Trp3486*)
U277	I	*GPR98*	NA	c.13536_13537delTC - p.(Pro4513fs) c.13536_13537delTC - p.(Pro4513fs)
RP1024	I	*MYO7A*	NA	c.3610C>A – p.(Pro1204Thr) c.3764delA – p.(Lys1255fs) °
U585	Undef	*MYO7A*	NA	c.2283-1G>T – p.(?) °
U1080	I	*MYO7A*	c.5803C>A – p.(Leu1935Met)	
		USH1C	NA	c.311G>A – p.(Gly104Asp) c.226C>T – p.(Gln76*)
RP1604	II	*CDH23*	NA	c.7221C>A – p.(Tyr2407*) c.7221C>A – p.(Tyr2407*)
RP1611	II	*CLRN1*	NA	c.67G>T – p.(Gly23*) c.67G>T – p.(Gly23*)

NA= gene not analysed by Sanger sequencing. Undef: data were not accurate enough to clearly discriminate a subtype. All missense were classified as likely pathogenic (UV3), with the contribution of segregation analysis, low frequencies in public databases and *in silico* predictions. ° Already described variant. References for published alterations, as well as dbSNP identifiers are all included in USHbases.

In patients RP98, RP1578, U654, and RP1616, one mutation had been missed by Sanger sequencing. The c.11864G>A mutation had not been detected in U654 and RP1616 because the sequencing primer was masking the variant. Two other mutations c.14803C>T (RP98) and c.13811+2T>G (RP1578) were not detected due to errors in reading the Sanger sequences. In 8 patients, mutations were identified in genes that had not been previously sequenced. Patients U277 and U286, clinically classified as USH2 were found to have truncating mutations in *GPR98*. Patient U1080, diagnosed as USH1 and previously found to carry a rare *MYO7A* missense, was also harboring mutations in *USH1C*. Patients RP1604 and RP1611 diagnosed as USH2, and patient RP1024 classified as USH3, were found mutated in genes (*CDH23, CLRN1* and *MYO7A,* respectively) usually implicated in a different clinical subtype. Patients U996 and U585 could not be classified in any of the subtypes based on the available clinical data. Two *USH2A* mutations were identified in U996, establishing a diagnosis of USH2, but only one *MYO7A* alteration was found in U585.

In 14 patients carrying a single mutation identified by Sanger analysis, no additional mutation could be detected by NGS.

Usher group

Thirteen USH patients (clinically classified as 5 USH1, 7 USH2 and 1 USH3) with no preliminary haplotyping or Sanger sequencing analysis underwent Usher-exome screening by NGS. The validated filtering and prioritization strategy was applied and selected 80 USH variants candidates, which represent an average of 6 variants of interest per patient. Among those, some were already described as non-pathogenic in our internal database or in USHbases and were therefore eliminated. The remaining variants are shown in Table 2. We then applied our multi-step analysis to classify these 49 variants (Roux, et al., 2011).

Table 2: Classification of 49 variations identified in the Usher group.

Patient	USH type	Gene	Variant	Class	Main contributor
U1067	I	*MYO15A*	c.10242C>T – p.(=)	UV1	b,e
U1084	I	*MYO7A*	c.6220C>T – p.(Pro2074Ser)	UV2	d
		USH1C	c.307C>T – p.(Arg103Cys)	UV3	d
		USH2A	c.14662A>T – p.(Thr4888Ser)	UV1	d
		GPR98	c.2727C>A – p.(=)	UV2	e
			c.9366A>G – p.(=)	UV2	e
		PDZD7	c.1916C>G – p.(Ala639Gly)	UV1	d
U1093	II	*USH1C*	c.2585C>T – p.(Pro862Leu)	UV1	d
			c.101A>G – p.(His34Arg)	UV2	b
		GPR98	c.8992_8994delinsAAGTTCC – p.(Ala2998fs)	Pathogenic	a
			c.8992_8994delinsAAGTTCC – p.(Ala2998fs)		
		MYO15A	c.7468G>A – p.(Ala2490Thr)	UV1	b
U1120	I	*CDH23*	c.9904G>A – p.(Glu3302Lys)	UV3	d
U1141	II	*USH2A*	c.1055C>T – p.(Thr352Ile)	UV4	c,d °
			c.949C>A – p.(?)	UV4	c °
U1148	II	*USH2A*	c.14475_14484del – p.(Ala4827fs)	Pathogenic	g,a
			c.14766delG – p.(Trp4922*)	Pathogenic	g,a
		PDZD7	c.2144C>T – p.(Pro715Leu)	UV1	b
		MYO15A	c.54G>A – p.(=)	UV1	b,e
			c.7908C>T – p.(=)	UV1	b,e
U1157	I	*CDH23*	c.9278+5G>C – p.(?)	UV4	f
			c.9278+5G>C – p.(?)		
U1163	I	*CDH23*	c.5015_5016delAT – p.(Tyr1672fs)	Pathogenic	a
			c.6829+1G>A – p.(?)	UV4	e
		PDZD7	c.1011C>T – p.(=)	Neutral	b,e
		MYO15A	c.1385G>A – p.(Gly462Asp)	UV1	b
			c.5133+15A>G – p.(=)	UV1	b,e
U1167	II	*USH2A*	c.1256G>T – p.(Cys419Phe)	UV4	c,d °
			c.2882delA – p.(His961fs)	Pathogenic	a
		GPR98	c.10820T>C – p.(Val3607Ala)	UV3	d

		CDH23	c.4891G>A – p.(Ala1631Thr)	UV1	d
U1178	II	*GPR98*	c.2864C>A – p.(Ser955*)	Pathogenic	a
			c.2864C>A – p.(Ser955*)		
			c.17756-4A>G – p.(=)	UV2	e
			c.17756-4A>G – p.(=)		
U1170	I	*MYO7A*	c.472G>A – p.(Gly158Arg)	UV3	d
			c.5502G>A – p.(Trp1834*)	Pathogenic	a
		CDH23	c.9319G>T – p.(Gly3107Trp)	UV2	h [$]
		CLRN1	c.472+4C>T – p.(=)	Neutral	b
		VEZT	c.1396G>A – p.(Glu466Lys)	UV2	h
		MYO15A	c.3413A>G – p.(Gln1138Arg)	Neutral	b
			c.4655+11G>A	UV1	b
U1171	II	*GPR98*	c.7129C>T – p.(Arg2377*)	Pathogenic	a
			c.13536_13537delTC – p.(Pro4513fs)	Pathogenic	a
		MYO7A	c.1134C>T – p.(=)	UV2	e
U1185	II	*USH2A*	c.653T>A – p.(Val218Glu)	UV3	c,d °
			c.13010C>T – p.(Thr4337Met)	UV3	d °
		DFNB31	c.1204-17A>G – p.(=)	UV2	e
		USH1C	c.496+1C>T – p.(?)	Pathogenic	a °
		VEZT	c.1831+4A>G – p.(=)	UV2	e [$]
		MYO15A	c.9478C>T – p.(Leu3160Phe)	UV2	b,d
			c.10182G>A – p.(=)	UV1	b,e

[$] Variant could alter splicing ° Previously described variant. References for published DNA variations, as well as dbSNP identifiers are all included in USHbases. Contributors to classification: a, protein translation predicts a PTC; b, allele frequency (public databases or control samples analysed by our laboratory); c, allele frequency (patients); d, *in silico* predictions (missense variants); e, *in silico* predictions (splicing); f, minigene analysis; g, segregation analysis; h, patient genotype.

NGS successfully identified the pathogenic genotype in 10 out of 13 patients (77%). Patients U1157 and U1163 were found with *CDH23* alterations. Because U1157 carried a newly described variant in position +5 of exon 63, a minigene analysis was performed to assess the impact of the substitution on the splicing process. In this context the c.9278+5G>C variation leads to a premature stop codon either by a retention of intron 63, a deletion of the last fifteen nucleotides of exon 63 (use of a cryptic donor splice site) or a total skipping of this exon (Figure 5). A patient USH1 (U1170) carried mutations in *MYO7A*, a truncating mutation and a newly described missense, p.(Gly158Arg). Among the three patients carrying pathogenic mutations in *GPR98*, two (U1093 and U1178) were homozygotes for truncating mutations and the other (U1171) was compound heterozygous for two additional truncating mutations. Four patients (U1084, U1141, U1167, U1185) were *USH2A* compound heterozygotes; three truncating mutations (p.(Ala4827fs), p.(Trp4922*), p.(His961fs)) are newly described whereas the five others, well known (see USHbases), include 4 missense (p.(Thr352Ile), p.(Cys419Phe), p.(Val218Glu), p.(Thr4337Met)) and a splicing alteration (c.949C>A).

In two additional patients, NGS detected a single rare candidate UV3 variant, *USH1C* p.(Arg103Cys) in U1084 and *CDH23* p.(Glu3302Lys) in U1120. All *USH1C* and

CDH23 exons were further sequenced by Sanger in U1084 and in U1120 respectively to avoid any missed mutation in an homopolymeric region.

Finally, in only one patient, U1067 clinically diagnosed as USH3, no candidate pathogenic alteration could be identified.

In addition to the pathogenic and UV3 variants identified, several rare variants were also detected among the thirteen patients. Most of them were classified as non-pathogenic. Interestingly, U1185 carried, in addition to an *USH2A* pathogenic genotype, the c.496+1G>T variant in *USH1C*.

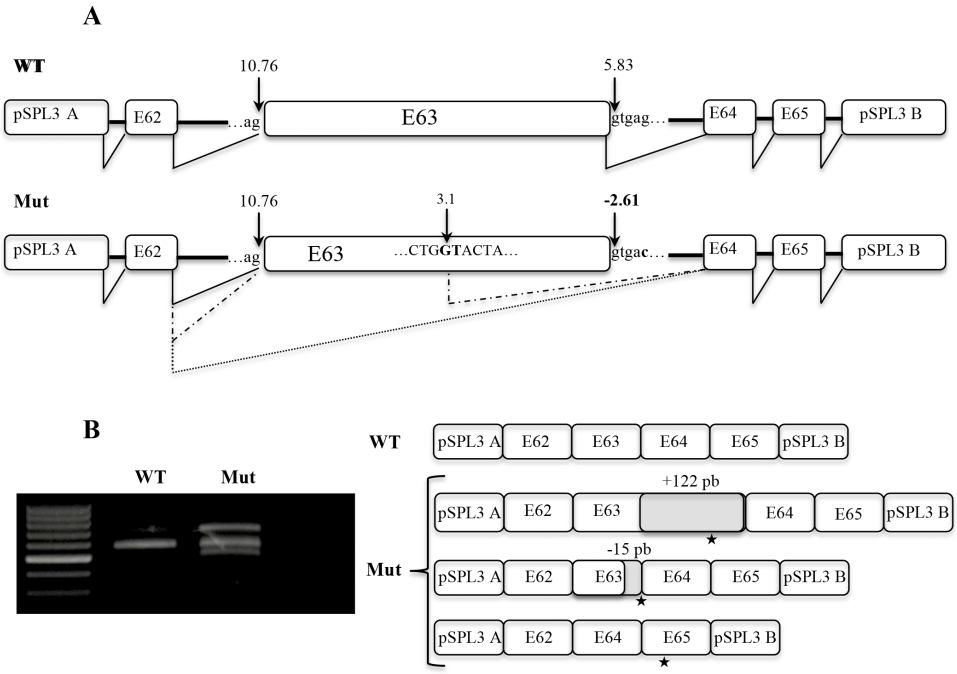


Figure 5: Ex vivo analysis of *CDH23* c.9278+5G>C. (A) Schematic wild-type (WT) and mutated (Mut) pSPL3 minigenes including exons 62 to 65 of *CDH23*. Exons are represented by boxes and introns by horizontal lines. The strengths of splice sites are mentioned by the MaxEnt scores. (B) Agarose gel and schematic sequences of transcripts obtained by RT-PCR from WT and Mut minigenes. Black stars indicate premature termination codons (PTC).

A single rare candidate variant (UV3) was identified in two additional patients: U1084 and U1120. *USH1C* p.(Arg103Cys) in U1084 and *CDH23* p.(Glu3302Lys). The whole *USH1C* and *CDH23* exons were further sequenced by Sanger in U1084 an U1120 respectively, to avoid any missed mutation in an homopolymeric region.

Remains one patient, U1067, clinically diagnosed as USH3, in whom no candidate pathogenic alteration could be identified.

In addition to the pathogenic and UV3 variants identified, several rare variants were also detected among the thirteen patients. Most of them were classified as non-pathogenic. Interestingly, U1185 carried, in addition to an *USH2A* pathogenic genotype, the c.496+1G>T variant in *USH1C*.

NSHL group

We assigned unambiguous disease causing mutations in only 1/11 cases (S91) although a number of potentially disease causing changes were present as heterozygotes in a number of cases. Among the 11 probands (Table 3) presenting with NSHL, seven were from consanguineous families (S4, S11, S25, S28, S91, S634, S789). A first step analysis focused on the search for homozygous variants in the different genes. Homozygosity was found for the USH1B (*MYO7A)* locus for S28 but only non-pathogenic variants were identified. Two unknown missense variants were identified in *PCDH15* in a heterozygous state but did not segregate with the disease. S91 showed homozygosity at the *TMC1* locus for a previously described deleterious mutation c.1165C>T was identified (see Retinal and hearing impairment genetic mutation database), and segregation analysis of parental DNA confirmed the homozygous status of the mutation.

In the cases not known or suspected to be consanguineous a number of single changes were observed. S334 was found with a heterozygous known pathogenic mutation in *CDH23,* c.6050-9G>A, but no other pathogenic variant could be identified in this gene (confirmed by Sanger) nor in any other genes. ArrayCGH analysis did not reveal any large rearrangement. S565 was already known to be a c.35delG heterozygote in GJB2 but negative for the common del*GJB6*(D13S1830) deletion. Interestingly, the c.35delG mutation was not detected with GS Reference Mapper as the deleted guanine is located in a homopolymeric sequence of 6 consecutive G. Two rare variants in *TECTA* were identified but none was classified as a likely pathogenic variant. S660 was found to be a carrier of *OTOF* c.1392+1G>T mutation. This mutation did not segregate with the HL in the family As for the two previous patients, no other pathogenic mutation could be identified and OTOF was excluded as responsible by linkage analysis.

Patient S789 was found heterozygous for the synonymous sequence change c.687C>T in exon 7 of *MYO7A*. *In silico* analysis suggests the creation of a donor splice site that would remove the last fifty nucleotides of exon 7 and create a PTC in exon 8. Because two other siblings presenting with HL were also available, haplotype analysis was performed and excluded the USH1B locus, i.e. *MYO7A* as a possible candidate gene.

Patients S4, S25, S634, S385 carry rare variants in different genes at heterozygous state (Table 3). Subsequent classification considered all the variants as non-pathogenic. S11 did not carry any candidate pathogenic or rare variants.

Table 3: List of candidate variants detected in NSHL group.

Patient	Documented consanguinity	Gene	Variant	Class	Main contributor
S4	Yes	*GJB3*	c.94C>T – p.(Arg32Trp)	UV1	b °
		PCDH15	c.1205G>C – p.(Gly402Ala)	UV2	d
		USH2A	c.8575C>T – p.(Arg2859Cys)	UV1	d
S11	Yes	-	-	-	-
S25	Yes	*MYO7A*	c.849+7C>G – p.(=)	UV2	e
S28	Yes	*PCDH15*	c.875C>G – p.(Pro292Arg)	UV3	b,d,h
			c.55T>A – p.(Ser19Thr)	UV1	d,h
S91	Yes	*TMC1*	c.1165C>T – p.(Arg389*)	Pathogenic	a °
			c.1165C>T – p.(Arg389*)		
S634	Yes	*USH2A*	c.14904C>T – p.(=)	UV2	e
S789	Yes	*MYO7A*	c.687C>T – p.(=)	UV3	e
S334	No	*CDH23*	c.6050-9G>A – p.(?)	Pathogenic	c °
S385	No	*MYO15A*	c.1656C>T – p.(=)	UV2	e
		OTOF	c.1966C>T – p.(Arg656Trp)	UV1	d
S565	No	*GJB2*	c.35delG – p.(Gly12fs)	Pathogenic	a °
		TECTA	c.3239A>T – p.(Asp1080Val)	UV2	d
			c.3456G>A – p.(=)	UV2	e
		MYO15A	c.806C>T – p.(Pro269Leu)	UV1	d
		PDZD7	c.1452G>T – p.(=)	UV2	e
S660	No	*OTOF*	c.1392+1G>T – p.(?)	Pathogenic	a
		CDH23	c.3995T>G – p.(Ile1332Ser)	UV3	d

° Previously described variant. References for published DNA variations as well as dbSNP identifiers are all included in USHbases. Contributors to classification: a, protein translation predicts a PTC; b, allele frequency (public databases or control samples analysed by our laboratory); c, allele frequency (patients); d, *in silico* predictions (missense variants); e, *in silico* predictions (splicing); f, minigene analysis; g, segregation analysis; h, patient genotype. No candidate pathogenic alterations have been identified for patient S11.

DISCUSSION

NGS in clinical services

As a reference laboratory, we have developed over the last 6 years a comprehensive approach that allows a mutation detection rate of more than 90 % of the cases for USH1 (Roux, et al., 2006; Roux, et al., 2011) and USH2 (Baux, et al., 2007; Besnard, et al., 2012). This includes screening for large rearrangements (Le Guédard, et al., 2007; Roux, et al., 2011) and the analysis of USH transcripts from nasal epithelial cells (Vaché, et al., 2010; Vaché, et al., 2012) as well as the development of a multistep analysis to interpret the variants of unknown clinical significance (Baux, et al., 2013). Because Sanger sequencing is time consuming, we included a preliminary linkage analysis at the USH1 loci that prioritizes the gene to be sequenced in 44 % of the cases. In our cohort, *MYO7A* is the most prevalent gene responsible for more than 60 % of the USH1 cases. *USH2A* accounts for 80 % of Usher type 2 cases (Besnard, et al., 2012) and is usually screened as a first step unless siblings are available or consanguinity is present, in which

cases, haplotype analyses are performed. In the present study, we have evaluated the Usher-exome coupled with a benchtop NGS machine. Sequencing in parallel all candidate genes has clear advantages as it can resolve not only atypical/undefined USH cases but also cases with a misclassified clinical subtype. But we have also noticed that despite a sensitivity of 98 %, failure to identify pathogenic mutations, particularly one of the founder mutations, was real and was inherent to technology limitations and the difficulty of sequencing short repeats. Taking all this together, we have worked out a "decision-making diagram" as represented on Figure S3. If clinical criteria are compatible with USH1, 2 or 3, we would still recommend performing haplotype analyses and/or Sanger sequencing the two major genes (i.e *MYO7A* for USH1 and *USH2A* for USH2); if negative, then NGS should be performed. For the atypical/undefined cases, NGS should be performed first.

Raw data quality and technical issues

Using the Roche GS junior 454 sequencer, we were able to generate an average of 40 Mb of pertinent nucleotides per run (Figure 1). The deficit of 10 Mb per run due to off-target sequences is inevitable to obtain a reasonable coverage for the regions of interest. While we observed that 96% of our target regions were covered by at least 40 reads (Figure 2, Figure S1), and that 98% of variants previously identified by Sanger sequencing were correctly found by NGS (Figure 4), we identified some weaknesses in the base-calling and alignment system used (namely 454 base-caller and GS Reference Mapper which are provided by Roche). Misalignment was noted not only in homopolymeric regions but also in some neighbouring regions (for example, the *USH2A* c.2299delG mutation lies in the vicinity of a stretch of 6 A and could not be detected at first when using the default parameters, see below). This is very important as such homopolymers are frequent causes of mutation due to slippage. Base-calling and alignment, two crucial steps, must be improved in the future. Base-calling is defined as the analysis of the sensor data to predict the individual bases (Ledergerber and Dessimoz, 2011). In the case of 454 pyrosequencing, this relies on the quantification of the emitted light during a single nucleotide flow (Margulies, et al., 2005). Until recently, only two base-callers were available to analyse 454 data (Datta, et al., 2010), the native 454 base-caller and Pyrobayes (Quinlan, et al., 2008). Pyrobayes has been reported to be more accurate than the built-in 454 base-caller in substitution error rate for example, but it does not improve errors due to homopolymeric regions. To address this point, a new method called HPcal has been claimed to reduce homopolymeric length errors by 35% (Beuf, et al., 2012).

The quality of the generated data could also be improved by modifying the alignment method. Alignment of hundreds of thousands reads on a reference genome is a huge task, and therefore dedicated software needs to proceed in two steps: a fast mapping of the reads is first performed on the genome to identify candidates regions and is followed by a fine alignment of the same reads on those regions. The latter is realized using a classical algorithm such as Smith-Watermann (Smith and Waterman, 1981), but the first mapping on candidates regions is achieved by different methods. Several software packages are available. The most efficient ones, such as BWA (Li and Durbin, 2009) or Bowtie (Langmead, et al., 2009), are based on the Burrows-Wheeler transform approach. One of the advantages of 454 sequencing is to generate long reads (average of 431 bp in this study), but this reduces the number of optimized alignment and base-calling methods

that are available. Among those, AGILE (AliGnIng Long rEads) seems to be promising in terms of accuracy, memory usage and speed (Misra, et al., 2011).

Filtering, Prioritization and Classification of variants

We divided the filtering, prioritization and classification of our NGS data into three distinct stages (Figure 3). Analysis of the copious amounts of data generated raises problems at two levels, firstly in checking the validity of the reads (steps 1 to 3) and, secondly, in determining its relevance to disease causing changes. All the settings were done on the Test Sample, i.e. focusing on the Usher genes mainly. The first step aims to eliminate the maximum of false positives without removing from the dataset the true positives. This step has been automated with GSdot software. The generic software GS Reference Mapper generates two lists of possible DNA variations, one containing all signals, and a second supposedly to include only the true variants. We first followed the recommendations of the manufacturer and worked with the ”cleaned file” (High Confidence Variations file) but realized that the most frequent mutation *USH2A* c.2299delG, responsible for 10-45% of *USH2A* pathogenic alleles in Europe (Aller, et al., 2010; Dreyer, et al., 2008; Le Quesne Stabej, et al., 2012), was systematically excluded from the second list. This pinpoints a drawback of the software, which did not detect a neighbouring alteration of an homopolymer, confirmed for other variants localized, not within, but next to homopolymeric regions. We therefore chose to use as input for our custom software GSdot the complete list of aligned DNA alterations, which explains the relatively high number of candidate DNA variants at the very beginning of the workflow (mean 4674), the majority of which, 92.3% (i.e 4316/4674) are removed/excluded by the first filter.

The second stage consisted in the selection of candidate pathogenic variants. This step remains to be automated. We chose not to depend on external software such as pathogenic predictors or external databases. This strategy has proven to be effective as only 8 variants per patient remained after this prioritization. At this step, every variant was confirmed by Sanger sequencing, which revealed that a few false positives were still present, and this suggests that more stringent filters can be applied. External software and databases were however fully integrated in the third step, which focused on the classification of the Usher candidate variants. Our multistep strategy of classification was applied to any putative splicing alteration as well as to variants liable to impact the protein structure, and proved to be very efficient (Baux, et al., 2013; Roux, et al., 2011). It should be noted that this was possible thanks to our in house experience of molecular alterations of the Usher genes in USH patients.

The same strategy was applied to the NSHL group and was quite laborious and time consuming in part because of our novice experience of the NSHL genes, and also because we did not have access to a tool similar to USMA (Usher Syndrome Missense Analysis) which contains and correlates all the published information. This software could be extended to a restricted number of NSHL genes. However, for larger numbers of genes, it is probably necessary to apply classical filters for missense based on the accumulation of predictive results (Neveling, et al., 2012; Perrault, et al., 2012; Redin, et al., 2012; Thompson, et al., 2013).

The massive amounts of data generated by NGS represent both a challenge for analysis and a powerful tool to classify sequence variations. The numbers of variants identified in 19 genes by Sanger sequencing (2475) or NGS (16851) in the 47 patients of the Test Sample are displayed on Figure S2. These data provide a unique resource in terms of distribution of variants identified in patients which will greatly facilitate future diagnoses.

Data from the Test Sample illustrate the limitations of Sanger sequencing in terms of efficiency in identifying pathogenic genotypes in the following situations: i) high genetic heterogeneity, particularly if the most prevalent genes (i.e. *MYO7A* and *USH2A*) are not involved, ii) large genes to be sequenced, and iii) patients with atypical Usher syndrome. In these cases, Sanger sequencing was time and money consuming that limited the completeness of the service in routine diagnostics labs. We have indeed improved USH diagnosis for 12/47 patients mainly because we increased exhaustivity.

Sensitivity of the strategy

We could estimate a sensitivity of 98% for our NGS approach as it successfully detected 674 variants among the 687 identified with Sanger sequencing. Among the 13 false negative variants, 10 were not detected because of misalignments and/or lying in homopolymeric regions, two major drawbacks that were previously discussed. The 3 remaining were lying in poorly covered regions within the *DFNB31* region, which requires an optimization of the design. However, this test only validates the high capacity of the 454 sequencing to identify nucleotide substitutions (670 among the 674 identified variations). This strong bias is due to the nature of the genotypes: the large majority of variants included in genotypes are neutral polymorphisms; as we mainly focused on exonic sequences, substitutions are over-represented, comprising 97.5% of all variants (670/687). Among the 496 USH patients genotyped to date in our laboratory, Sanger sequencing identified a total of 13,399 variations, of which 93% are DNA substitutions. This supports the fact that 454 sequencing would be able to identify the vast majority of our genotypes. However, if we exclude neutral polymorphisms, the proportion of substitutions for this series decreases to 68.9% (1073/1557). As a direct consequence, one of the 13 missed variants was a pathogenic mutation (c.6116_6119delGCTG in *USH2A)*, due to base-calling default and misalignment. This confirms the urgency to improve these methods and in the meantime, Sanger sequencing should be performed systematically for patients in whom one pathogenic variant is identified by NGS.

Among the 60 USH patients analysed with the “Usher-exome” designed in this study (47 from the *Test Sample* (Table 1, Supp Mat S1) and 13 from the Usher group), the pathogenic genotype (2 mutations) has been identified in 22 patients. No causative mutation could be found in 19 patients in any candidate USH genes or NSHL genes also included in the design. Deep intronic mutations (such as c.7595-2144A>G recently identified in *USH2A* by mRNA studies (Vaché, et al., 2012)), may be involved in rare cases. However the most likely explanation for negative cases is that the patients are not "classical" USH patients, i.e. carrying mutations in "USH genes". Only Whole Exome Sequencing (WES) or Whole Genome Sequencing (WGS) are likely to bring some answers and could possibly redefine the clinical diagnosis for some cases. Among 19 patients

carrying a single mutation, 10 were heterozygous for an *USH2A* pathogenic allele. The carrier frequency of an *USH2A* mutation is estimated to be of 1/70 in UK (95% CI=1/333 - 1/40) (Le Quesne Stabej, et al., 2012) and in our cohort (95% CI=1/111- 1/53) (unpublished results), therefore these patients cannot all be random carriers and it is most likely that the second mutation has not yet been detected. For two patients, U838 carrying *MYO7A* p.(Cys31*), a common mutation in Scandinavian populations (Janecke, et al., 1999), and U585 carrying c.2283-1G>T, a common mutation in North African populations, (unpublished results), the clinical description did not allow classification into a particular subgroup so that the question still remains as to whether they are random carriers.

Among the genotyped USH patients, two could be put forward as potential oligogenic or digenic cases. Patient U1185 (presenting with typical USH2 clinical signs) carries 3 pathogenic mutations: two in *USH2A* and the *USH1C* c.496+1G>T splice mutation frequently identified in the USH population originating from UK (Le Quesne Stabej, et al., 2012). Patient U286 carries the *USH2A* c.2299delG and the *GPR98* p.(Trp3486*) truncating mutation. Again digenic mechanism could be postulated, although this patient could as well be a random carrier of the frequent c.2299delG mutation with two pathogenic *GPR98* mutations. We have already described a patient as a random c.2299delG carrier associated with a *CDH23* linked USH1 syndrome (Roux, et al., 2011).

Incidental findings are a direct consequence of an exhaustive screening with NGS. Although the number of genes screened with this approach is targeted, it already pinpoints the presence of additional mutations, which probably just reflect the carrier rate frequency in the general population. In Usher syndrome, the carrier frequency of one USH gene mutation could be estimated to 1/42 (95% CI= [1/90 -1/27], considering a MAF for c.2299delG of 0.009 (EVS and (Baux, et al., 2007)) with this mutation representing 7.9% of the pathogenic USH alleles in our series.

NSHL

The pilot study was conducted on 11 NSHL patients for 5 genes responsible for NSHL: *OTOF, TMC1, TECTA, MYO15A, GJB3* plus *GJB2* and *GJB6* (Table S2) in addition to Usher genes. All the identified variants underwent the filtering and prioritization previously defined. Two mutations were detected in only one patient, S91, with Spanish and Algerian origins, homozygous for the *TMC1* p.(Arg389*) mutation. *TMC1* is the sixth most common cause of recessive HL worldwide (Hilgert, et al., 2009), and the most prevalent in Iran, Turkey, Israel and Jews of Moroccan origin (Brownstein, et al., 2011). Single mutations were found in two further cases but their significance is unclear. S334 is a carrier of the *CDH23* c.6050-9G>A mutation. This splicing mutation represents 9.7 % (47/483) of the *CDH23* pathogenic alleles linked to USH1 (see USHbases). A single non-syndromic deafness case has been reported in a young patient (Kimberling, et al., 2010), however ophthalmologic follow-up might have discovered an ocular phenotype. Because a second mutation has not been identified, S334 could be a random carrier of the general population. S660 was identified with a single *CDH23* variant p.(Ile1332Ser) (UV3) and an *OTOF* mutation, c.1392+1G>T. Linkage analysis at the USH1D/DFNB12 locus excluded the *CDH23* involvement, as well as for *OTOF*.

The most common mutation responsible for NSHL *GJB2* c.35delG was not detected as it lies in a homopolymeric region. Luckily, the screening of mutations at the DFNB1 locus (*GJB2* mutations and del(*GJB6*-D13S1830) that accounts for up to 50 % of the NSHL cases (Green, et al., 1999; Snoeckx, et al., 2005) is easy to set up, cost effective, and should definitely been done before considering NGS. For the "non-DFNB1" cases, NGS on targeted sequences covering more than 50 NSHL genes has already proven to be powerful using 454 GS FLX or Illumina GA (Shearer, et al., 2010) (Brownstein, et al., 2011) and seems to be the most relevant choice. Because our knowledge with the mutational spectrum of NSHL genes is still limited, interpretation of the variants was much more laborious and time consuming, and not as straightforward than for the Usher variants.

In conclusion, this pilot study of NGS applied to the molecular diagnosis of an heterogeneous disorder emphasizes the need of special expertise of the genes analyzed by NGS for correct interpretation of variants in a clinical context. In-house databases cumulating patients' data, as well as public available databases such as the Deafness variation database will be of great help to develop efficient NSHL diagnosis.

Acknowledgments

This work was supported in part by le Ministère de la Recherche "PHRC National 2004, PROM 7802". GG-G is a recipient of a fellowship from the Spanish Ministry of Education (REF: AP2008-02760). TB is a recipient of a UNADEV Foundation fellowship.

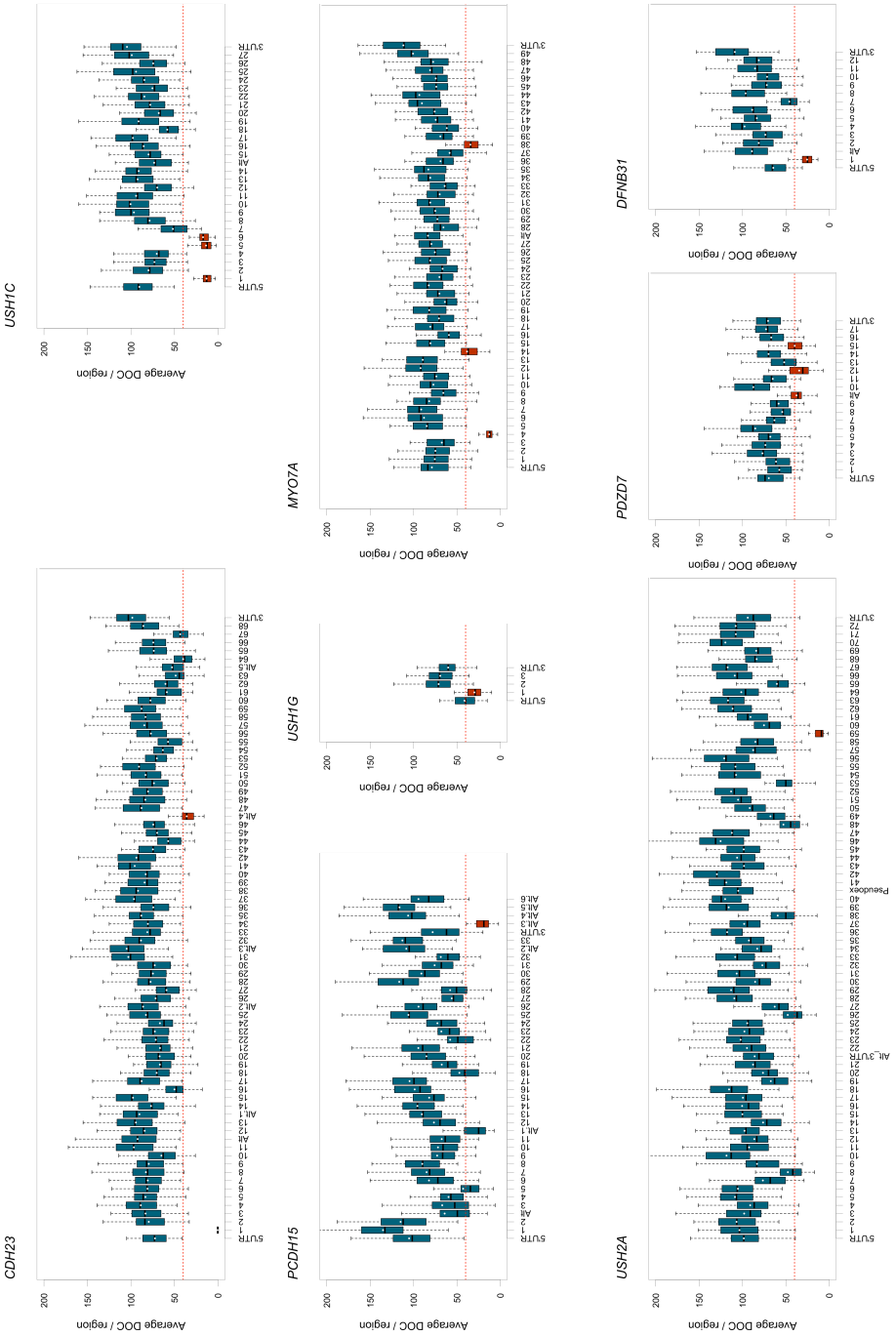
Supplementary Material

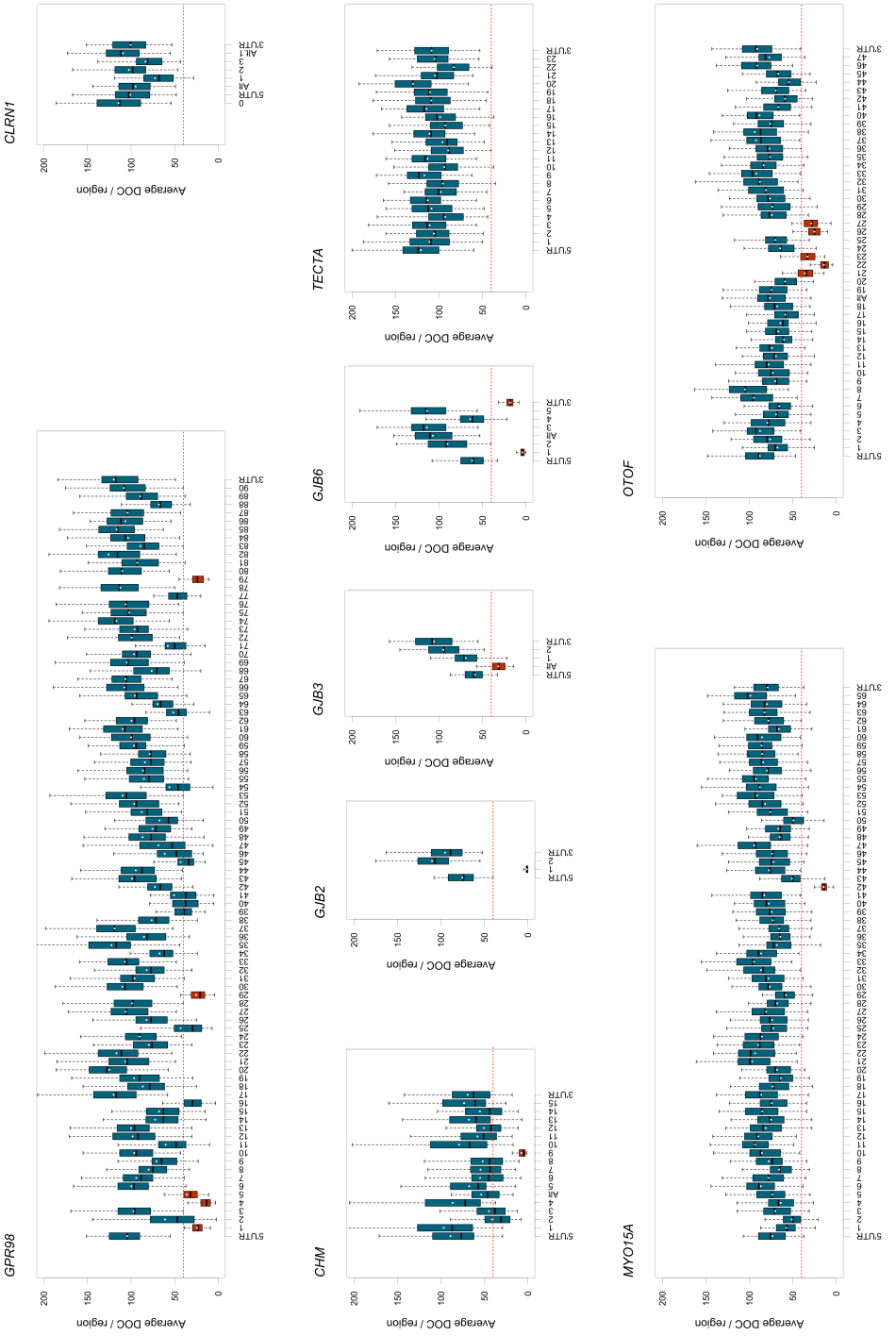
PATIENT	USH Type	*MYO7A*	*USH1C*	*CDH23*	*PCDH15*	*USH1G*	*USH2A*	*GPR98*	*DFNB31*	*PDZD7*	*CLRN1*	Results before NGS	aCGH
U810	I	X	X	X	X							NM_033056.3(*PCDH15*):c.1441-1G>A	X
U331	I	X	X	X	X	X			X	X	X	NM_033056.3(*PCDH15*):del E18-26	X
U1080	I	X										NM_000260.3(*MYO7A*):p.(Leu1935Met)	
U321	I	X		X	X		X				X	-	
U503	I			X	X						X	-	
U584	I	X	X	X	X						X	-	
U740	I	X	X	X	X		X				X	-	
U283	II						X			X		NM_206933.2(*USH2A*):p.(Gly2039fs)	
U466	II	X	X	X	X		X	X		X	X	NM_206933.2(*USH2A*):p.(Cys759Phe)	X
U286	II						X				X	NM_206933.2(*USH2A*):p.(Glu767fs)	
RP1360	II						X					NM_206933.2(*USH2A*):p.(Cys759Phe)	X
RP1485	II						X					NM_206933.2(*USH2A*):p.(Glu404fs)	X
RP1578	II						X					NM_206933.2(*USH2A*):p.(Glu767fs)	X
RP1617	II						X					NM_206933.2(*USH2A*):p.(His308fs)	X
RP1635	II						X					NM_206933.2(*USH2A*):p.(Cys759Phe)	X
RP963	II						X					NM_206933.2(*USH2A*):p.(Cys759Phe)	X
U329	II						X		X		X	NM_001195263.1(*PDZD7*):p.(Arg56fs)	
U787	II						X	X	X	X		NM_032119.3(*GPR98*):p.(Arg5688*)	X
RP1634	II						X	X	X			NM_032119.3(*GPR98*):p.(Gln5796*)	X
U496	II			X			X	X	X	X	X	NM_206933.2(*USH2A*):p.(Arg1777Trp)	X
U654	II						X	X	X	X	X	NM_206933.2(*USH2A*):p.(Pro1843Leu)	X
U932	II						X					NM_206933.2(*USH2A*):p.(Cys717Gly)	X
RP1600	II						X	X	X			NM_206933.2(*USH2A*):p.(Val382Met)	
U391	II						X	X	X	X	X	NM_032119.3(*GPR98*):p.(Gly2045Arg)	X
U444	II		X	X	X				X		X	-	X
U583	II						X	X	X	X	X	-	
U670	II						X	X	X	X	X	-	
U767	II	X		X			X	X	X	X	X	-	
U436	II						X	X	X	X	X	-	
U277	II						X		X	X	X	-	
RP98	II						X					-	
RP1059	II						X	X	X			-	
RP1278	II						X	X	X			-	
RP1417	II						X		X			-	
RP1448	II						X	X	X			-	
RP1604	II						X	X	X			-	
RP1606	II						X	X	X			-	
RP1611	II						X	X	X			-	
RP1616	II						X	X	X			-	
RP1636	II						X	X	X			-	
RP659	II						X					-	
RP1024	III						X	X	X			-	
RP1612	III						X	X	X			-	
U838	Undef	X		X			X	X	X	X	X	NM_000260.3(*MYO7A*):p.(Cys31*)	X
U461	Undef	X	X	X	X		X		X		X	NM_022124.5(*CDH23*):p.(Asp1130Tyr)	X
U585	Undef						X	X	X	X	X	-	
U996	Undef										X	-	

Table S1: List of 47 Usher patients included in control group. The genes previously studied using Sanger sequencing or aCGH for each subject are marked with a cross and the identified putative mutations are displayed (all the mutations were detected in heterozygous state).

Chr	gene	RefSeq AccNo (nucleotide)	RefSeq AccNo (protein)	exons	size (bp)	coding size (bp)	Amino acids	Alternative RefSeq (nucleotide)	Alternative RefSeq (protein)	Additional exon(s)	size (bp)
1	*GJB3*	NM_001005752.1 variant 2	NP_001005752.1 isoform a	2	1777	813	270	NM_024009.2 variant 1	NP_076872.1 isoform a	1	591
1	*USH2A*	NM_206933.2 variant 2	NP_996816.2 isoform b	72	18883	15609	5202	NM_007123.5 variant 1	NP_009054.5 isoform a	1	1534
2	*OTOF*	NM_194248.2 variant 1	NP_919224.1 isoform a	47	7171	5994	1997	NM_194322.2 variant 3	NP_919303.1 isoform c	1	395
3	*CLRN1*	NM_001195794.1 variant 5	NP_001182723.1 isoform d	4	2398	738	245	NM_052995.2 variant 4	NP_443721.1 isoform c	2	1129
5	*GPR98*	NM_032119.3 variant 1	NP_115495.3 isoform 1	90	19333	18921	6307				
10	*PCDH15*	NM_033056.3 variant C	NP_149045.3 isoform CD1-4	33	7021	5868	1955	NM_001142771.1 variant K	NP_001136243.1 isoform CD3-1	1	4359
								NM_001142769.1 variant I	NP_001136241.1 isoform CD2-1	3	990
								NM_001142770.1 variant J	NP_001136242.1 isoform CD2-2	1	74
								NM_001142773.1 variant H	NP_001136245.1 isoform CD1-10	1	6
								NM_001142763.1 variant A	NP_001136235.1 isoform CD1-1	1	16
10	*CDH23*	NM_022124.5 variant 1	NP_071407.4 isoform 1	69	11134	10065	3354	NM_001171932.1 variant 5	NP_001165403.1 isoform 5	1	402
								NM_052836.3 variant 2	NP_443068.1 isoform 2	1	1005
								NM_001171931.1 variant 4	NP_001165402.1 isoform 4	1	612
								NM_001171930.1 variant 3	NP_001165401.1 isoform 3	1	739
								NM_001171933.1 variant 6	NP_001165404.1 isoform 6	1	310
								NM_001171935.1 variant 8	NP_001165406.1 isoform 8	1	481
11	*USH1C*	NM_153676.3 variant b3	NP_710142.1 isoform b3	27	3246	2700	899	NM_005709.3 variant 1	NP_005700.2 isoform a	1	75
11	*MYO7A*	NM_000260.3 variant 1	NP_000251.3 isoform 1	49	7465	6648	2215	NM_001127179.2 variant 3	NP_001120651.2 isoform 3	1	410
17	*USH1G*	NM_173477.2	NP_775748.2	3	3561	1386	461				
9	*DFNB31*	NM_015404.3 variant 1	NP_619636.2 isoform 1	12	4079	2724	907	NM_001083885.2 variant 2	NP_001077354.2 isoform 2	1	110
10	*PDZD7*	NM_001195263.1 variant 1	NP_001182192.1 isoform 1	17	4164	3102	1033	NM_024895.4 variant 2	NP_079171.1 isoform 2	1	285
9	*TMC1*	NM_138691.2	NP_619636.2	24	3201	2283	760				
11	*TECTA*	NM_005422.2	NP_005413.2	23	6468	6468	2155				
12	*VEZT*	NM_017599.3 variant 1	NP_060069.3 isoform 1	12	4580	2340	779	NR_038242.1 variant 2	non coding	1	74
13	*GJB2*	NM_004004.5	NP_003995.2	2	2347	681	226				
13	*GJB6*	NM_001110219.2 variant 1	NP_001103689.1 isoform 1	5	2178	786	261	NM_006783.4 variant 3	NP_006774.2 isoform1	1	369
17	*MYO15A*	NM_016239.3	NP_057323.3	65	11876	10593	3530				
X	*CHM*	NM_000390.2 variant 1	NP_000381.1 isoform a	15	5442	1962	653	NM_001145414.1 variant 2	NP_001138886.1 isoform b	1	2493

Table S2: Table of 19 human genes targeted for the next-generation sequencing. Additional exon(s) refers to the number of exons differing from the main isoform.


USH1C
MYO7A
DFNB31
CDH23
USH1G
PDZD7
PCDH15
USH2A
Average DOC / region
5UTR
3UTR


CLRN1
GPR98
CHM
GJB2
GJB3
GJB6
TECTA
MYO15A
OTOF
Average DOC / region
5'UTR
3'UTR

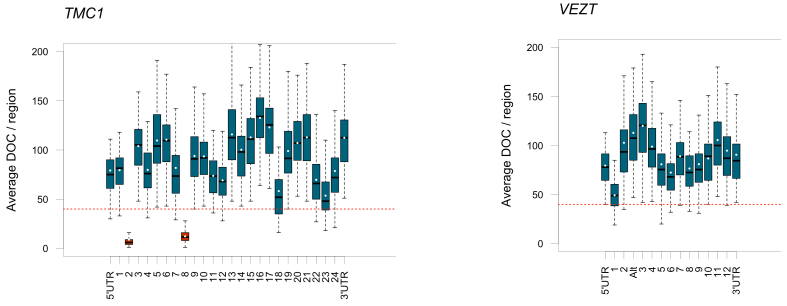


Figure S1: Box-plots of average DOC for targeted regions. Detailed coverage for the 634 regions are showed gene by gene. Boxes are showed in red when mean DOC in the region is lower than 40X. Red line represents a DOC of 40 reads, the minimum limit for a proper validation. The horizontal grey line corresponds to the median. The mean value is marked with a white point.

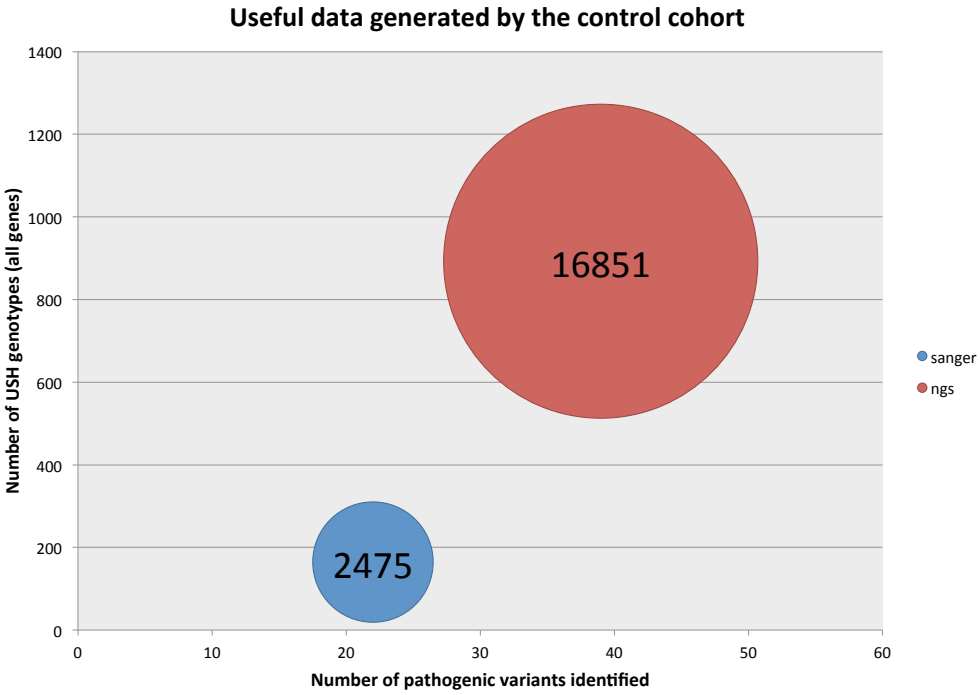


Figure S2: Comparison between useful data generated by NGS and Sanger sequencing in Usher control cohort. Numbers in the sphere correspond to the sum of variants identified in each approach.

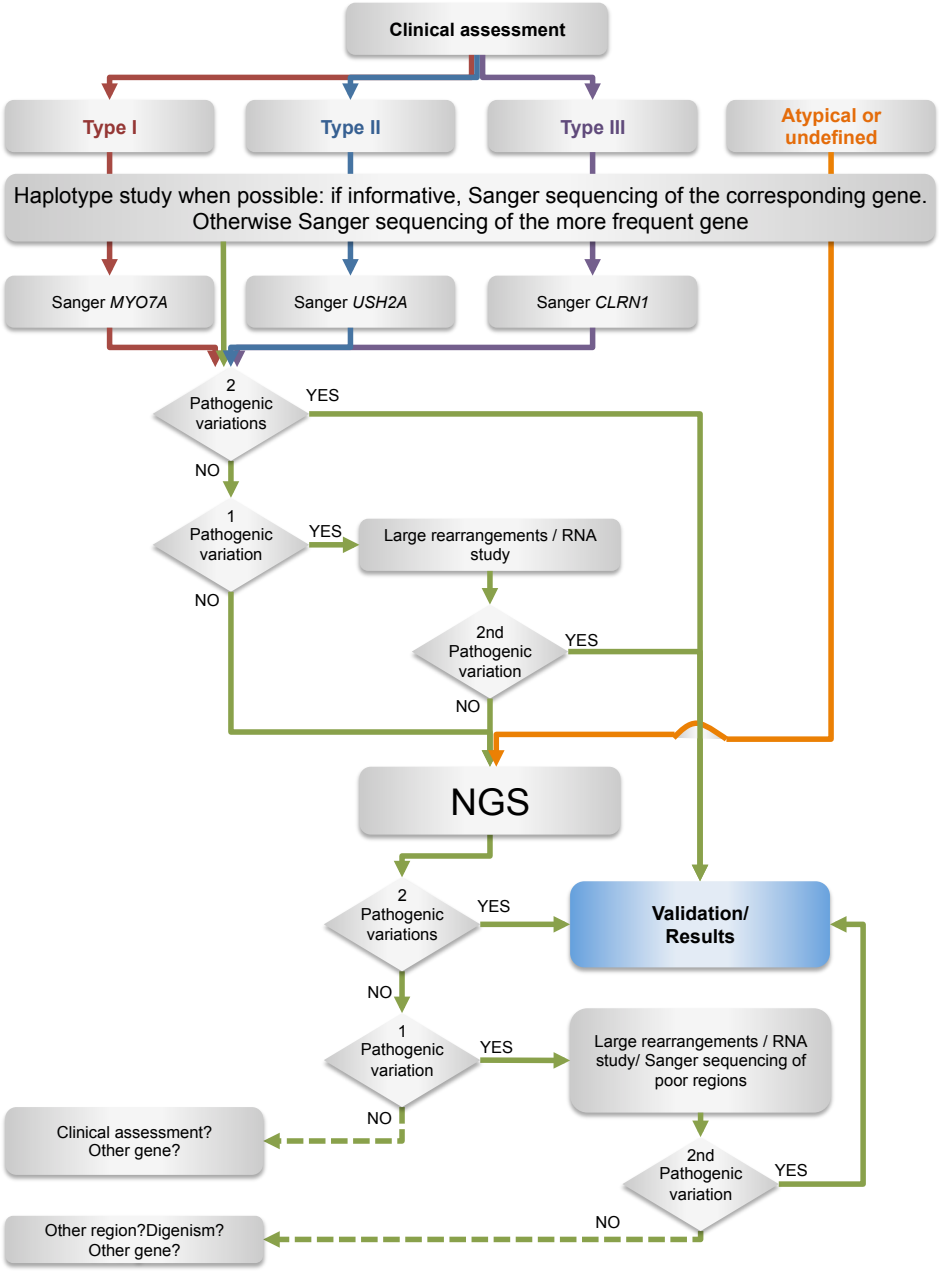


Figure S3: Inclusion of NGS in the decision-making diagram for molecular diagnosis of USH patients.

REFERENCES

Adato A, Weil D, Kalinski H, Pel-Or Y, Ayadi H, Petit C, Korostishevsky M, Bonne-Tamir B. 1997. Mutation profile of all 49 exons of the human myosin VIIA gene, and haplotype analysis, in Usher 1B families from diverse origins. Am J Hum Genet 61:813-21.

Ahmed ZM, Riazuddin S, Aye S, Ali RA, Venselaar H, Anwar S, Belyantseva PP, Qasim M, Riazuddin S, Friedman TB. 2008. Gene structure and mutant alleles of PCDH15: nonsyndromic deafness DFNB23 and type 1 Usher syndrome. Hum Genet 124:215-23.

Aller E, Jaijo T, Beneyto M, Najera C, Oltra S, Ayuso C, Baiget M, Carballo M, Antinolo G, Valverde D and others. 2006. Identification of 14 novel mutations in the long isoform of USH2A in Spanish patients with Usher syndrome type II. J Med Genet 43:e55.

Aller E, Larrieu L, Jaijo T, Baux D, Espinos C, Gonzalez-Candelas F, Najera C, Palau F, Claustres M, Roux AF and others. 2010. The USH2A c.2299delG mutation: dating its common origin in a Southern European population. Eur J Hum Genet 18:788-793.

Baux D, Larrieu L, Blanchet C, Hamel C, Ben Salah S, Vielle A, Gilbert-Dussardier B, Holder M, Calvas P, Philip N and others. 2007. Molecular and in silico analyses of the full-length isoform of usherin identify new pathogenic alleles in Usher type II patients. Hum Mutat 28:781-789.

Baux D, Vaché C, Malcolm S, Claustres M, Roux A-F. 2013. Interpretation of variants of uncertain clinical significance (VUCS): the paradigm of Usher syndrome In: Riazuddin S, Ahmed ZM, editors. Inner Ear Development and Hearing Loss. New York: Nova Science Publishers, Inc. p 159-174

Besnard T, Vache C, Baux D, Larrieu L, Abadie C, Blanchet C, Odent S, Blanchet P, Calvas P, Hamel C and others. 2012. Non-USH2A Mutations in USH2 Patients. Human mutation 33:504-510.

Beuf KD, Schrijver JD, Thas O, Criekinge WV, Irizarry RA, Clement L. 2012. Improved base-calling and quality scores for 454 sequencing based on a Hurdle Poisson model. BMC Bioinformatics 13:303.

Bitner-Glindzicz M, Lindley KJ, Rutland P, Blaydon D, Smith VV, Milla PJ, Hussain K, Furth-Lavi J, Cosgrove KE, Shepherd RM and others. 2000. A recessive contiguous gene deletion causing infantile hyperinsulinism, enteropathy and deafness identifies the usher type 1C gene. Nat Genet 26:56-60.

Bonnet C, Grati M, Marlin S, Levilliers J, Hardelin JP, Parodi M, Niasme-Grare M, Zelenika D, Delepine M, Feldmann D and others. 2011. Complete exon sequencing of all known Usher syndrome genes greatly improves molecular diagnosis. Orphanet J Rare Dis 6:21.

Brownstein Z, Friedman LM, Shahin H, Oron-Karni V, Kol N, Abu Rayyan A, Parzefall T, Lev D, Shalev S, Frydman M and others. 2011. Targeted genomic capture and massively parallel sequencing to identify genes for hereditary hearing loss in Middle Eastern families. Genome biology 12:R89.

Cremers FP, Kimberling WJ, Kulm M, de Brouwer A, van Wijk E, Te Brinke H, Cremers CW, Hoefsloot LH, Banfi S, Simonelli F and others. 2006. Development of a genotyping Microarray for usher syndrome. J Med Genet

Datta S, Kim S, Chakraborty S, Gill RS. 2010. Statistical Analyses of Next Generation Sequence Data: A Partial Overview. J Proteomics Bioinform 3:183-190.

Desmet FO, Hamroun D, Lalande M, Collod-Beroud G, Claustres M, Beroud C. 2009. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. Nucleic Acids Res 37:e67.

Dreyer B, Brox V, Tranebjaerg L, Rosenberg T, Sadeghi AM, Moller C, Nilssen O. 2008. Spectrum of USH2A mutations in Scandinavian patients with Usher syndrome type II. Hum Mutat 29:451.

Ebermann I, Phillips JB, Liebau MC, Koenekoop RK, Schermer B, Lopez I, Schafer E, Roux AF, Dafinger C, Bernd A and others. 2010. PDZD7 is a modifier of retinal disease and a contributor to digenic Usher syndrome. J Clin Invest 120:1812-23.

Garcia-Garcia G, Aparisi MJ, Jaijo T, Rodrigo R, Leon AM, Avila-Fernandez A, Blanco-Kelly F, Bernal S, Navarro R, Diaz-Llopis M and others. 2011. Mutational screening of the USH2A gene in Spanish USH patients reveals 23 novel pathogenic mutations. Orphanet J Rare Dis 6:65.

Green GE, Scott DA, McDonald JM, Woodworth GG, Sheffield VC, Smith RJ. 1999. Carrier rates in the midwestern United States for GJB2 mutations causing inherited deafness. Jama 281:2211-6.

Guédard-Méreuze SL, Vaché C, Molinari N, Vaudaine J, Claustres M, Roux AF, Tuffery-Giraud S. 2009. Sequence contexts that determine the pathogenicity of base substitutions at position +3 of donor splice-sites. Hum Mutat 30:1329-39.

Hilgert N, Smith RJ, Van Camp G. 2009. Forty-six genes causing nonsyndromic hearing impairment: Which ones should be analyzed in DNA diagnostics? Mutat Res 681:189-196.

Janecke AR, Meins M, Sadeghi M, Grundmann K, Apfelstedt-Sylla E, Zrenner E, Rosenberg T, Gal A. 1999. Twelve novel myosin VIIA mutations in 34 patients with Usher syndrome type I: confirmation of genetic heterogeneity. Hum Mutat 13:133-40.

Joensuu T, Hamalainen R, Yuan B, Johnson C, Tegelberg S, Gasparini P, Zelante L, Pirvola U, Pakarinen L, Lehesjoki AE and others. 2001. Mutations in a novel gene with transmembrane domains underlie Usher syndrome type 3. Am J Hum Genet 69:673-84.

Khateb S, Zelinger L, Ben-Yosef T, Merin S, Crystal-Shalit O, Gross M, Banin E, Sharon D. 2012. Exome Sequencing Identifies a Founder Frameshift Mutation in an Alternative Exon of USH1C as the Cause of Autosomal Recessive Retinitis Pigmentosa with Late-Onset Hearing Loss. PLoS One 7:e51566.

Kimberling WJ, Hildebrand MS, Shearer AE, Jensen ML, Halder JA, Trzupek K, Cohn ES, Weleber RG, Stone EM, Smith RJ. 2010. Frequency of Usher syndrome in two pediatric populations: Implications for genetic screening of deaf and hard of hearing children. Genet Med 12:512-6.

Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome biology 10:R25.

Le Guédard S, Faugère V, Malcolm S, Claustres M, Roux AF. 2007. Large genomic rearrangements within the PCDH15 gene are a significant cause of USH1F syndrome. Mol Vis 13:102-7.

Le Guédard-Mereuze S, Vaché C, Baux D, Faugère V, Larrieu L, Abadie C, Janecke A, Claustres M, Roux AF, Tuffery-Giraud S. 2010. Ex vivo splicing assays of mutations at non-canonical positions of splice sites in USHER genes. Hum Mutat 31:347-355.

Le Quesne Stabej P, Saihan Z, Rangesh N, Steele-Stallard HB, Ambrose J, Coffey A, Emmerson J, Haralambous E, Hughes Y, Steel KP and others. 2012. Comprehensive sequence analysis of nine Usher syndrome genes in the UK National Collaborative Usher Study. Journal of medical genetics 49:27-36.

Ledergerber C, Dessimoz C. 2011. Base-calling for next-generation sequencing platforms. Briefings in bioinformatics 12:489-97.

Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754-60.

Licastro D, Mutarelli M, Peluso I, Neveling K, Wieskamp N, Rispoli R, Vozzi D, Athanasakis E, D'Eustacchio A, Pizzo M and others. 2012. Molecular diagnosis of Usher syndrome: application of two different next generation sequencing-based procedures. PLoS One 7:e43799.

Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen Z and others. 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature 437:376-80.

Misra S, Agrawal A, Liao WK, Choudhary A. 2011. Anatomy of a hash-based long read sequence mapping algorithm for next generation DNA sequencing. Bioinformatics 27:189-95.

Neveling K, Collin RW, Gilissen C, van Huet RA, Visser L, Kwint MP, Gijsen SJ, Zonneveld MN, Wieskamp N, de Ligt J and others. 2012. Next-generation genetic testing for retinitis pigmentosa. Human mutation 33:963-72.

Perrault I, Hanein S, Zanlonghi X, Serre V, Nicouleau M, Defoort-Delhemmes S, Delphin N, Fares-Taie L, Gerber S, Xerri O and others. 2012. Mutations in NMNAT1 cause Leber congenital amaurosis with early-onset severe macular and optic atrophy. Nature genetics 44:975-7.

Puffenberger EG, Jinks RN, Sougnez C, Cibulskis K, Willert RA, Achilly NP, Cassidy RP, Fiorentini CJ, Heiken KF, Lawrence JJ and others. 2012. Genetic mapping and exome sequencing identify variants associated with five novel diseases. PLoS One 7:e28936.

Quinlan AR, Stewart DA, Stromberg MP, Marth GT. 2008. Pyrobayes: an improved base caller for SNP discovery in pyrosequences. Nat Methods 5:179-81.

Redin C, Le Gras S, Mhamdi O, Geoffroy V, Stoetzel C, Vincent MC, Chiurazzi P, Lacombe D, Ouertani I, Petit F and others. 2012. Targeted high-throughput sequencing for diagnosis of genetically heterogeneous diseases: efficient mutation detection in Bardet-Biedl and Alstrom syndromes. Journal of Medical Genetics 49:502-12.

Riazuddin S, Belyantseva IA, Giese AP, Lee K, Indzhykulian AA, Nandamuri SP, Yousaf R, Sinha GP, Lee S, Terrell D and others. 2012. Alterations of the CIB2 calcium-

and integrin-binding protein cause Usher syndrome type 1J and nonsyndromic deafness DFNB48. Nature genetics 44:1265-71.

Roux AF, Faugere V, Le Guedard S, Pallares-Ruiz N, Vielle A, Chambert S, Marlin S, Hamel C, Gilbert B, Malcolm S and others. 2006. Survey of the frequency of USH1 gene mutations in a cohort of Usher patients shows the importance of cadherin 23 and protocadherin 15 genes and establishes a detection rate of above 90%. J Med Genet 43:763-768.

Roux AF, Faugere V, Vache C, Baux D, Besnard T, Leonard S, Blanchet C, Hamel C, Mondain M, Gilbert-Dussardier B and others. 2011. Four year follow-up of diagnostic service in USH1 patients. Invest Ophthalmol Vis Sci 52:4063-4071.

Saihan Z, Webster AR, Luxon L, Bitner-Glindzicz M. 2009. Update on Usher syndrome. Curr Opin Neurol 22:19-27.

Shearer AE, DeLuca AP, Hildebrand MS, Taylor KR, Gurrola J, 2nd, Scherer S, Scheetz TE, Smith RJ. 2010. Comprehensive genetic testing for hereditary hearing loss using massively parallel sequencing. Proceedings of the National Academy of Sciences of the United States of America 107:21104-9.

Smith TF, Waterman MS. 1981. Identification of common molecular subsequences. J Mol Biol 147:195-7.

Snoeckx RL, Huygen PL, Feldmann D, Marlin S, Denoyelle F, Waligora J, Mueller-Malesinska M, Pollak A, Ploski R, Murgia A and others. 2005. GJB2 Mutations and Degree of Hearing Loss: A Multicenter Study. Am J Hum Genet 77:945-57.

Thompson BA, Greenblatt MS, Vallee MP, Herkert JC, Tessereau C, Young EL, Adzhubey IA, Li B, Bell R, Feng B and others. 2013. Calibration of multiple in silico tools for predicting pathogenicity of mismatch repair gene missense substitutions. Human mutation 34:255-65.

Vaché C, Besnard T, Blanchet C, Baux D, Larrieu L, Faugère V, Mondain M, Hamel C, Malcolm S, Claustres M and others. 2010. Nasal epithelial cells are a reliable source to study splicing variants in Usher syndrome. Hum Mutat 31:734-41.

Vaché C, Besnard T, le Berre P, Garcia-Garcia G, Baux D, Larrieu L, Abadie C, Blanchet C, Bolz HJ, Millan J and others. 2012. Usher syndrome type 2 caused by activation of an USH2A pseudoexon: Implications for diagnosis and therapy. Human mutation 33:104-8.

van Wijk E, Pennings RJ, te Brinke H, Claassen A, Yntema HG, Hoefsloot LH, Cremers FP, Cremers CW, Kremer H. 2004. Identification of 51 novel exons of the Usher syndrome type 2A (USH2A) gene that encode multiple conserved functional domains and that are mutated in patients with Usher syndrome type II. Am J Hum Genet 74:738-44.

Verpy E, Leibovici M, Zwaenepoel I, Liu XZ, Gal A, Salem N, Mansour A, Blanchard S, Kobayashi I, Keats BJ and others. 2000. A defect in harmonin, a PDZ domain-containing protein expressed in the inner ear sensory hair cells, underlies Usher syndrome type 1C. Nat Genet 26:51-5.

Vozzi D, Aaspollu A, Athanasakis E, Berto A, Fabretto A, Licastro D, Kulm M, Testa F, Trevisi P, Vahter M and others. 2011. Molecular epidemiology of Usher syndrome in Italy. Molecular Vision 17:1662-8.

Wang K, Li M, Hakonarson H. 2010. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic acids research 38:e164.

Wildeman M, van Ophuizen E, den Dunnen JT, Taschner PE. 2008. Improving sequence variant descriptions in mutation databases and literature using the Mutalyzer sequence variation nomenclature checker. Human mutation 29:6-13.

Yeo G, Burge CB. 2004. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. J Comput Biol 11:377-94.

CHAPTER V

Results and discussion

1. Exhaustive analysis of the USH2 genes in the Spanish population

Usher syndrome displays a high genetic and allelic heterogeneity. Chapter III describes the molecular study of the three genes involved in Usher syndrome type 2 (*USH2A*, *GPR98* and *DFNB31)* and we show the high variability of mutations that can be found in Usher genes.

1.1 Classification of variants

1.1.1 *USH2A* screening

Eighty-eight patients were included in the analysis of the *USH2A* gene. At least one mutation was identified in 43 out of 88 patients (48,9%). In 25 of these 43 subjects the two causative mutations were found, of which five cases were homozygotes for the mutation and in 20 patients the pathological variants were detected in heterozygous state.

In order to classify the variants several criteria were followed: previous studies, allele frequency, segregation analysis, *in silico* predictions and *in vitro* analysis. They were classified as pathogenic (+), probably pathogenic (UV3), probably non-pathogenic (UV2) and non-pathogenic (-). In figure 19 we show the classification of the 144 identified variants in these four categories.

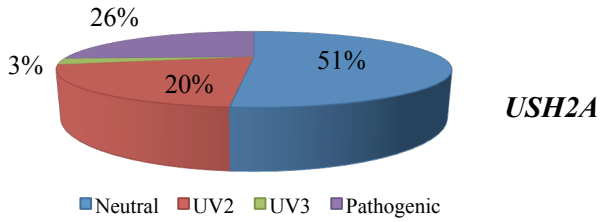


Figure 19. Percentage of variations on the basis of their pathogenicity.

As a result of our analysis, 37 different *USH2A* clearly pathological mutations were detected, 23 of them novel including: six nonsense, three splice-site, 11 small deletions/insertions, two missense and one isocoding mutation. Moreover, 107 additional variants were identified, including: 42 missense, 19 isocoding and 46 located in intronic regions. Four out of 107 were classified as UV3, 29 as UV2 and 74 as non-pathogenic.

In two variants (p.V382M and p.E2496E) *in vitro* analysis was needed to confirm a possible pathogenic effect predicted by the bioinformatics tools. The variant p.V382M affected the first base of exon seven. However, the results of minigene analysis were not conclusive and then we included this change in the UV3 group. For the novel isocoding

variant, p.E2496E, the predicted splicing effect was confirmed and hence, it was classified as pathogenic.

1.1.2 *GPR98* and *DFNB31* screening

Several patients remained without mutations after *USH2A* screening, so we decided to follow up with the analysis of the other USH2 genes in patients for whom DNA was still available. Nineteen patients negative for *USH2A* analysis were included in the study of *GPR98* and *DFNB31*. In five patients, haplotype analysis revealed homozygosity at the USH2C locus. *GPR98* was directly studied in those five cases. The remaining patients were first screened for *DFNB31* mutations, due to the huge size of *GPR98*. The USH2C and USH2D loci were excluded in one patient by haplotype analysis, so it was discarded of the study.

Mutations in *GPR98* were found in seven out of 18 analysed subjects, of whom five were classified as USH2 and two as USHNC. In five patients mutations were detected in homozygous state (as predicted by haplotype analysis). A total of seven different mutations were identified, six of them previously unreported. Five out of seven pathogenic variants led to a premature termination codon, including one duplication, two deletions, one insertion/deletion and one nonsense mutation. The other two mutations affected the splice-sites, one of them novel, and both were classified as UV4.

Seventy-five additional variants were classified as non-deleterious and nineteen were considered as UV1 or UV2 on the basis of the allele frequency, bioinformatics predictions, *in vitro* analysis or presence of mutations in *cis*. In figure 20 the percentage of mutations included in each group is showed. For *GPR98* and *DFNB31* two novel categories, UV1 (variant certainly neutral) and UV4 (variant certainly pathogenic) were added, respect to the *USH2A* classification.

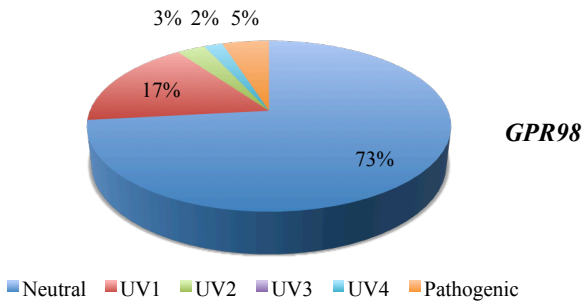


Figure 20. Classification of 101 variations detected in *GPR98* in our cohort.

Concerning the *DFNB31* gene, no pathogenic mutations were detected in the 13 patients analysed.

1.1.3 Additional analysis

We also looked for large rearrangements using array-CGH, to perform a complete analysis of the genes. Two novel large deletions were identified in *USH2A*, one of them in a patient with a previously detected mutation and another one in homozygous state. Large deletions or duplications were not detected in the *GPR98* gene.

Sequencing analysis of the intronic region, where a deep intronic mutation was recently reported in *USH2A* (VACHE et al. 2012), led us to detect the second mutation in four additional Spanish patients who remained with only one pathogenic variation after *USH2A* screening.

The results of the array-CGH analysis and the finding of the intronic mutation, led us to detect the second mutation in a total of five additional patients that remained with an incomplete genotype after the *USH2A* screening.

1.2 Genotype-phenotype correlation

Whenever possible, clinical data of patients carrying mutations in any of the genes were gathered. We managed to obtain clinical data from 25 patients with both pathogenic mutations in *USH2A* and from three patients with one pathogenic plus one UV3 in *USH2A*. Most of patients showed a phenotype characteristic of USH2. However, there were two interesting cases, RP259 and RP1703. The first one displayed a profound hearing loss, the RP started at six years old and the vestibular function was altered; all features typical from Usher syndrome type 1. USH1 genes were not studied in this patient. Maybe the subject carried the mutations in an USH1 gene, additionally of *USH2A* mutations, responsible of the phenotype. The second case was the patient RP1703, for whom the first symptoms of the disease were noticed at 50 years old, later than usual. We cannot discard the presence of variations in other genes that could be involved in the phenotype modification.

Concerning *GPR98*, clinical data could be obtained from five out of seven subjects carrying mutations, and the clinical characteristics were compatible with a typical Usher syndrome type 2. In two patients bearing mutations in *GPR98,* audiograms were analysed. Both of them presented a down-slopping configuration with a tone loss slightly stronger in high frequencies. The results confirm the trend observed in the study of Abadie *et al*. where patients mutated in *GPR98* showed a more severe hearing loss than those mutated in *USH2A* (ABADIE *et al*. 2012).

1.3 Recurrent mutations

Several studies have been performed about the USH2 genes and a high number of mutations have already been reported. More than 200 mutations are known in *USH2A* and more than 40 in *GPR98*. Despite of all previous studies, the mutational screening of the genes involved in Usher syndrome type 2 in our Spanish cohort revealed 23 novel

pathological mutations in *USH2A*, and six in *GPR98*. A considerable number of mutations were detected in only one or few families, showing that most of the mutations responsible for Usher syndrome are private. When we compare the mutational spectrum found in different populations, we can observe that the number of recurrent mutations is rare. Only 6 *USH2A* and one *GPR98* pathogenic mutations detected in our study were reported in other populations. However, a prevalent mutation located in exon 13 of *USH2A*, designated as c.2299delG, is frequently found in the European and US patients. The allele frequency distribution of c.2299delG varies geographically in Europe. This mutation accounts for 47.5% of *USH2A* alleles in Denmark and for 36% in Scandinavia, whereas an allelic frequency of 31% was found in the Netherlands, 16–36 % in the United Kingdom and 15% in France (ALLER *et al.* 2010b). In the Spanish population the allelic frequency was estimated around 15%. To date, in our cohort, after c.2299delG, the most frequent mutation is p.C3267R, representing a 5,6% of mutated alleles. Contrarily to *USH2A*, a prevalent mutation has been not identified in *GPR98*.

In our cohort, only five out of 43 cases were homozygous for *USH2A*. However, we detected a high number of homozygous cases in *GPR98*, five out of seven with mutations in that gene. Consanguinity was only reported in three out of five patients with homozygous mutations. It confirms the advantage of carrying out preliminary haplotype analysis; it can be a useful tool to prioritize the screening, especially in Usher syndrome, where huge genes are involved.

1.4 Distribution of mutations

A representation of the detected mutations along the usherin protein is shown in García-García *et al*. and we can observe that these mutations are homogeneously distributed along all the protein and that hot spots are not present (GARCIA-GARCIA *et al.* 2011). As in previous studies, no mutations have been identified in the transmembrane domain or the intracellular region of usherin. A similar figure for GPR98 protein, with the distribution of mutations detected in our cohort and in other previous studies can be found in Appendix (Figure S1). Like in usherin, *a priori* there are not domains with a clear accumulation of mutations in GPR98 protein. However, after a careful analysis, we can observe that a 45% of the reported mutations are located in the carboxy-terminal end. In our patients, no mutations have been detected in the intracellular regions. Besnard *et al*. found two frameshifts in the transmembrane domain of GPR98 that would predict a protein lacking of the PDZ domain, essential for the interaction with others proteins (BESNARD *et al.* 2012).

1.5 Mutational spectrum

The mutational spectrum of *USH2A* involves all type of mutations: splicing mutations, missense, nonsense, small and large rearrangements. Surprisingly, in our cohort pathogenic missense variations have not been detected in *GPR98*. The seven mutations identified here produced directly or indirectly a premature stop codon, with the subsequent

truncated protein. This difference respect to the *USH2A* gene could be due to the lower number of patients analysed for the G*PR98* gene (18), but if we take into account all known *GPR98* mutations, the missense mutations represent only a 15,6% (8/51) of the total. The percentage is considerably lower than the one found in *USH2A*. In our study, missense *USH2A* mutations represented a 27% (10/37) of the total mutations and if we gather together all the mutations reported in USH-database, the percentage rises to 32% (108/338) (Figure 21). According to the results, *USH2A* would be more prone to generate missense variations compared to *GPR98*. Alternatively, we could hypothesise that *GPR98* would tolerate better the nucleotide changes. The noticeable difference between these transmembrane proteins could be explained by the amino acid organization, as proposed by Besnard *et al*. Nearly 88% sequence of usherin protein belongs to an identified domain, meanwhile over 40% of *GPR98* protein has an unknown function, which maybe tolerates better the codon changes (BESNARD *et al.* 2012).

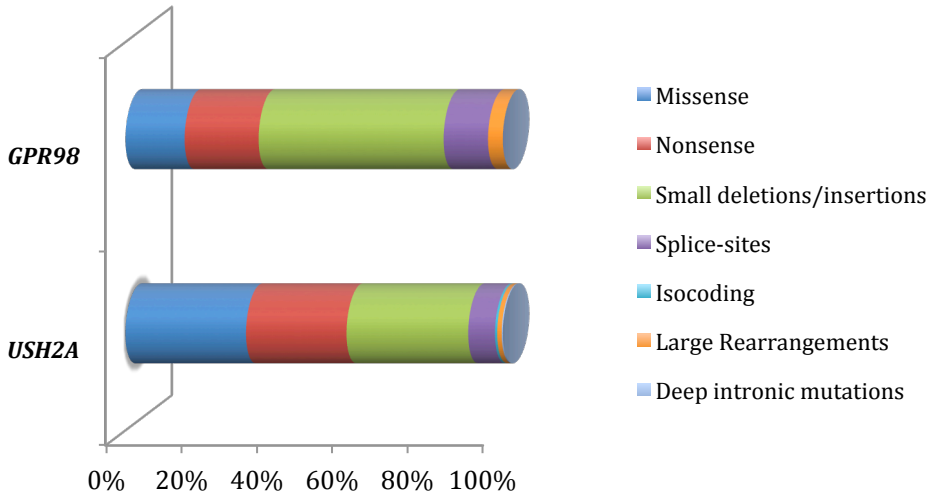


Figure 21. Distribution of the different types of mutations detected in the *USH2A* and *GPR98* genes. Fifty-one and 338 pathogenic mutations have been reported in *GPR98* and *USH2A,* respectively. Pathogenic, UV4 and UV3 are included. Data obtained from: Retinal and hearing impairment genetic mutation database (https://grenada.lumc.nl/LOVD2/Usher_montpellier).

In addition, the first example of a deep intronic mutation in Usher syndrome was detected in the *USH2A* gene. This variant was found in five French patients and in four Spanish subjects, allowing us to complete the molecular diagnosis of these subjects. In this regard, it is tempting to speculate that the same type of mutation may be found in the future in other Usher genes. As an example, RP1634 presented only one mutation in *GPR98*, large rearrangements were discarded and *PDZD7* screening resulted negative. It could be possible that the second mutation responsible for the disease was located in a deep intronic region.

1.6 Involvement of USH2 genes in the Spanish population

To minimize the bias in the estimation of the mutation detection ratio in *USH2A*, due to the nature of our initial cohort of patients, we took into account all the *USH2A* screenings carried out by our research group on our own USH2 cohort. A total of 134 patients classified as USH2 have been studied for the *USH2A* gene, identifying mutations in 102 cases. The estimated ratio was 76.1%, quite similar to those previously obtained in others studies (BAUX *et al.* 2007; DREYER *et al.* 2008).

Pathogenic mutations in *GPR98* were found in seven out of 19 negative *USH2A* patients (36.8%). If we include all the USH2 patients analysed by our group, the percentage decrease to 5.2%. Mutations in *DFBN31* were not detected in our Spanish cohort, indicating the low involvement of this gene in USH2 pathogenesis.

When we compare our global results with those from other studies, we can see that the percentages found in the Spanish population are similar to those showed in Besnard *et al*. or Le Quesne *et al*, in French and English population, respectively (Figure 22) (BESNARD *et al.* 2012; LE QUESNE STABEJ *et al.* 2012). As said in the introduction, Bonnet *et al*. showed a higher contribution for *GPR98* and *WHRN* than the others studies, but probably it is due to the low number of patients analysed (n=21) (BONNET *et al.* 2011).

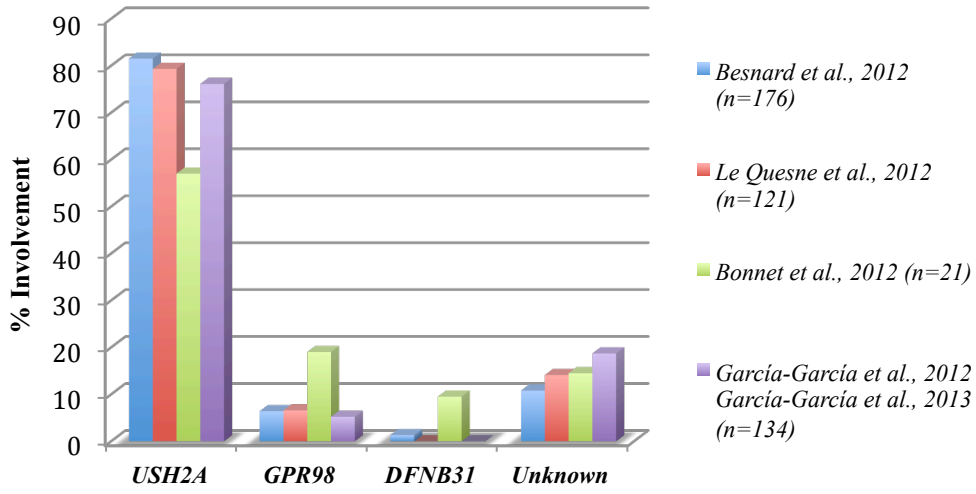


Figure 22. Involvement of the USH2 genes in different studies. N= number of patients included in the study.

1.7 Future prospects

We have performed an exhaustive molecular analysis of USH2 genes in our cohort of Spanish patients. However, there are some patients who remain genetically undiagnosed. Some of them could bear mutations located in regions not analysed, as deep intronic regions or 5' and 3' UTR, which highlights the need to perform further analyses at DNA and RNA level.

Nevertheless, it would be interesting to review in detail the clinical data of these patients, to confirm they are really affected of Usher syndrome. Eisenberg *et al*. reported in 2012 the study of a consanguineous family with two affected members displaying clinical features compatible with Usher syndrome type 3 (EISENBERGER *et al.* 2012). As result of analysing the homozygous regions detected using a SNP array mapping by next generation sequencing, a mutation was detected in *ABHD12* gene, known to be involved in PHARC (polyneuropathy, hearing loss, ataxia, retinitis pigmentosa, and early-onset cataract). Further neurological exams revealed the patient had ataxia, concluding that the patient was affected from a variant of PHARC, rather than USH3. It confirms the importance of an accurate clinical diagnosis and the helpfulness of massively parallel sequencing to improve the molecular diagnosis.

Patients displaying characteristics of a clinical Usher syndrome subtype, but with mutations detected in genes usually involved in other subtype have been previously reported. Thus, we cannot discard some of our unresolved patients carry mutations in some of the USH1 genes. In order to confirm this hypothesis, these patients were posteriorly included in the cohort of patients that was analysed using the next generation sequencing platform developed for all the Usher syndrome genes described in chapter IV.

2. *CLRN1* gene analysis

Molecular analysis of the USH3 gene was carried out in 17 USH patients. The study of the *CLRN1* gene was performed according to the algorithm established in our laboratory.

2.1 Mutational spectrum

As a result of analysis using the Asper Biotech Usher genotyping microarray, the mutation p.Y63X was detected in homozygous state in a consanguineous family (FRP44) and in heterozygous state in a second family (FRP529). In order to find the second mutation responsible for the disease, all the coding region and flanking sequences were sequenced. A novel nonsense mutation was detected in exon 3, p.R207X. Haplotype analysis was performed and the variant segregated with the disease.

In a third family (FRP360), haplotype analysis was compatible with an involvement of the *USH3A* locus. Sequencing resulted in the identification of two mutations, p.R207X and the novel missense p.I168N. A pathologic effect was predicted for this novel variant according to the results of computational tools.

To date, a total of 18 mutations had been reported in the *CLRN1* gene in patients with Usher syndrome: eight missense, three nonsense and seven small deletions/insertions. In this thesis we have reported two additional mutations, the premature stop codon p.R207X and the aminoacid change p.I168N.

Only four different pathogenic mutations have been found in the *CLRN1* gene in the Spanish population: p.C40G, p.Y63X, p.I168N and p.R207X. Most of them are found in only one or two families. The nonsense mutation p.Y63X, which is the most frequent, was detected in three families and in a total of nine alleles.

2.2 Involvement of USH3

Usher syndrome type 3 is considered to be the rarest USH type worldwide; however, among Ashkenazi Jews and in the Finnish population, up to 40% of the USH cases have the USH3 subtype as result of multiple founder effects (MILLAN *et al.* 2011). In the rest of populations this type of USH accounts for only 1-6%.

In our series, 23 families classified as USH3 were studied and only seven families carried mutations in *CLRN1*. Therefore several families initially classified as USH3 carried mutations in *USH2A*. Similar findings were observed in the study of Bonnet *et al*. where in six USH3 patients, pathogenic variations were identified in *USH2A* (BONNET *et al*. 2011). In these cases it would be interesting to make a revision of the clinical features to confirm they are really USH3. Again, it is worth to point out the importance and need to have an

accurate clinical diagnosis and make a follow up of the patients in order to check the progression of the symptoms, to allow the correct classification and molecular analysis.

2.3 Future prospects

In the three families with mutations in *CLRN1* reported here, the index case was previously diagnosed as USH3 on the basis of their clinical characteristics. However, some studies have demonstrated that mutations in *CLRN1* can also course with clinical forms similar to USH1 or USH2. Therefore, due to the high phenotypic variability and the small size of the gene, it would be interesting to include the analysis of the *CLRN1* gene in all USH patients, independently of the clinical subtype.

3. Development of a new NGS platform for Usher genes

For the sake of easy molecular diagnosis of Usher syndrome we decided to test the reliability of the new techniques of massively parallel sequencing in this disease. In Chapter IV we have described the development and implementation of a new NGS platform to study the Usher genes, candidate Usher genes and several genes involved in sensorineural hearing loss. Additional details of the thecnique are available in Appendix (Figure S2 and Figure S3).

3.1 Raw data

Forty-seven patients as a control group were used to obtain overall results about quality of raw data. On average, around 130.000 reads (410 bp of mean) per sequencing run passed the filters, corresponding to more than 53 million of nucleotides.

3.1.1 Major filters

The rejected reads did not pass the follow quality filters:

- Key Pass Filter: this filter discards reads that do not start with a valid 4-base 'key' sequence corresponding to either a library read (GACT for Rapid Library). The sequencing key is a known sequence of four nucleotides located immediately downstream from the sequencing primer and, therefore, the first to be sequenced in each well.

- Dot Filter: this filter discards reads that are too short or have too many poor incorporations or interruptions.

- Mixed Filter: this filter discards reads with too many nucleotide incorporations, possibly occurring from a bead carrying two or more DNA fragments attached, a well containing more than one DNA bead or signal contamination from a neighbouring well.

- Sequences are too short or the quality is not good enough.

3.1.2 Alignment of sequences

Around a 20% of analysable reads were non-specifically captured and aligned off the target regions. The *on target* was estimated in 52%. These data are shown in figure 23:

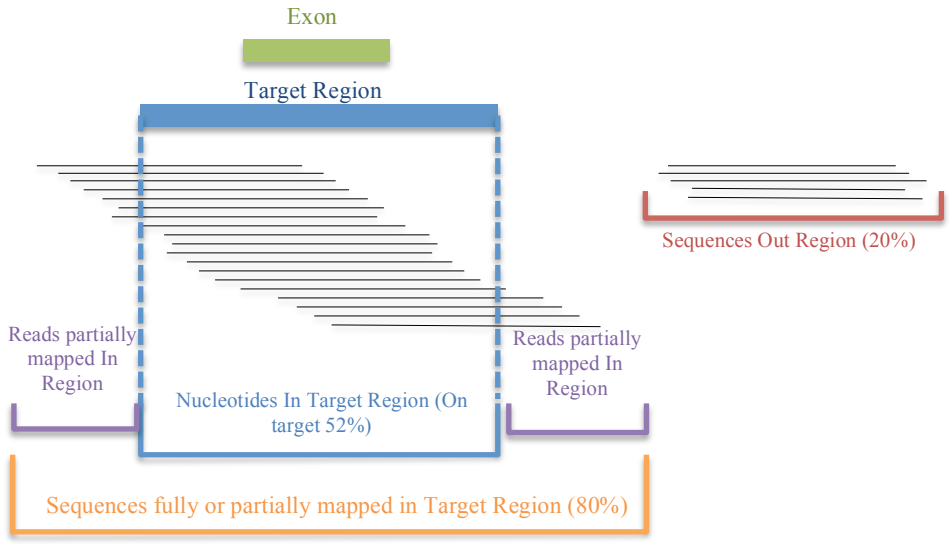


Figure 23. Representation of mapping reads outcome of the sequencing. The exonic region is represented in green, and the region designed to be captured in blue. Twenty per cent of reads were mapped outside of desired regions (in red). In reads partially mapped inside of the region designed for the capture there were nucleotides bordering the region (in purple), but for *on target* estimation only nucleotides exactly inside of the region (in blue) were accounted.

3.2 Coverage depth

The mean coverage depth was estimated in 77X but the coverage was not homogeneous along all the regions. A minimum value of 40X is assumed for an accurate validation and we can see in figure 24 that most of regions overcome this limit. Only 37 out of 634 regions were poorly covered.

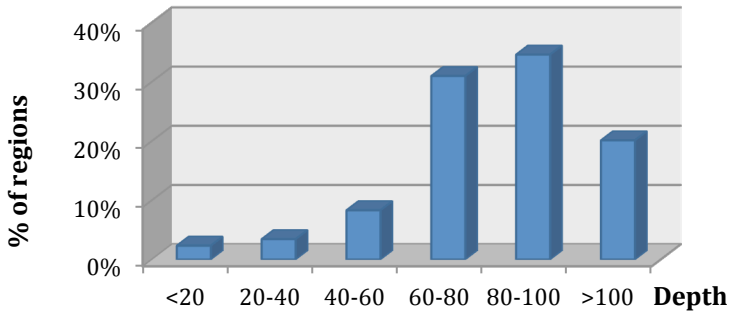


Figure 24. Global results of mean coverage depth obtained in the 635 regions. Around 85% of regions were covered more than 60X and only 2% less than 20X.

3.2.1 GC content

The regions with a weak coverage were common in all the patients, suggesting this uneven average could be due to the nature of the regions. It has been reported that GC rich sequences are difficult to capture (SULONEN *et al*. 2011). Eighteen regions with a poor coverage had GC percentages higher than 60% compared to the average GC content of 41% for the 50 best covered exons.

In figure 25 an average of coverage respect to the GC content is showed and it is evident that regions with more than 70% GC were really bad covered.

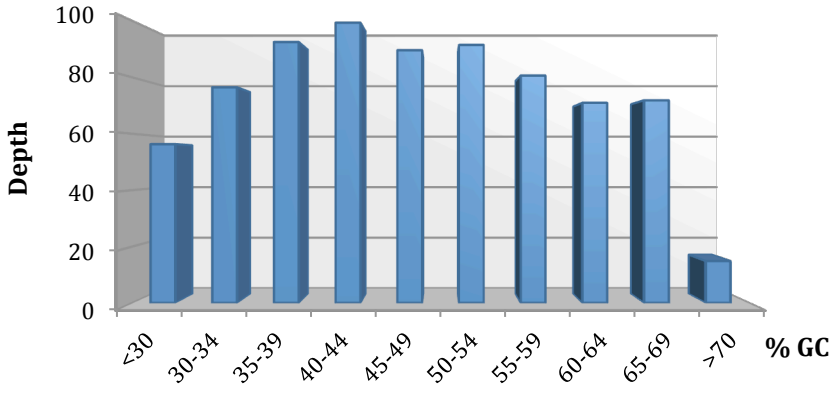


Figure 25. Mean of depth of regions depending on their GC percentage.

3.2.2 Repetitive elements

Most of these regions presented also repetitive elements that could be also responsible for the poor capture efficiency in these cases. As an example, exon 59 of *USH2A* was difficult to sequence by Sanger and also by NGS. This region had not an unusual GC content (36%), but there was a (TA)n simple repetitive element overlapping with the targeted region. This type of repetitions are spread genome wide, so it could be generating a competition in the hybridization, with the subsequently poor coverage in targeted region for exon 59.

Although in some regions the problem could be due to the complexity of the sequence, there are regions with apparently normal sequence and low coverage so, a thoroughly revision of the design will be essential to improve capture efficiency.

3.3 Sanger *vs* NGS

A key point was to know the sensitivity of the technique versus Sanger sequencing, used traditionally. Comparing 687 variants previously detected by Sanger, only a 2% were not detected using NGS. Three variants were not detected due to a low coverage. However, the other variants were mainly missed due to the presence of homopolymeric regions and/or alignment problems. In this way, it would be recommendable to use additional softwares to confirm that mutations are not missed due to these problems. In this way, we also tested the SeqNext software (JSI Medical Systems). In most cases, this software led us to identify the mutations not detected previously with GS Reference Mapper, but it was less specific, reporting a high number of variants, and possibly more false positive. To deal with these problems, new base-caller and alignment softwares to analyse 454 data are being developed. The native 454 base-calling has difficulties to distinguish the number of incorporated nucleotides in homopolymeric stretches, consequently this system has a relatively high error rate for calling insertions and deletions (indels) in homopolymers or in their vecinity. Recently, Beuf *et al*. have proposed an alternative method called HPcall which integrates more quality information in the base-calling and the variant detection seems to be more accurate (BEUF *et al.* 2012). The same problem occurs with the alignment, alternative methods for long reads have been developed to improve the Reference Mapper provided by Roche.

3.4 Developing GSDot

Initially, to analyse the variants we used a file generated for Reference Mapper, which theoretically contained the true variants. However, the most frequent mutation, c.2299delG, in the *USH2A* gene was not included in this file. Probably it was discarded due to the presence of a homopolymeric stretch surround the deletion (Figure 26).

```
>chr1   216420437      216420437      C       -       USH2A

Reference                        GT-CCTTT-GGCTTC-TTTTTT--G-CACTCACACTG-CCCAGAGTGAGGATTGCAGAATT
                                                                *

HHI1QPW01BGTAI          459+ GT-CCTTT-GGCTTC-TTTTT---G-CACT-ACACTG-CCCAGAGT          495

HHI1QPW01AM60F           74- GT-CCTTT-GGCTTC-TTTTTT--G-CACT-ACACTG-CCCAGAGTGAGGATTGCAGAATT 22

HHI1QPW01AXGWP           73- GT-CCTTT-GGCTTC-TTTTT---G-CACT-ACACTG-CCCAGAGTGAGGATTGCAGAATT 22
                                                                *

HHI1QPW01BWHUB           62- GT-CCTTT-GGCTTC-TTTTTT--G-CACTCACACTG-CCCAGAGTGAGGATTGCAGAATT 9
HHI1QPW01BKA1S           84- GT-CCTTT-GGCTTC-TTTTTT--G-CACTCACACTG-CCCAGAGTGAGGATTGCAGAATT 31
HHI1QPW01A7HCS          379+ GT-CCTTT-GGCTTC-TTTTT---G-CACTCACACTG-CCCAGAGTGAGGATTGCAGAATT 431
```

Figure 26. Alignment of sequences in a patient with the pathogenic mutation c.2299delG. Different sequences in the negative strand, with or without the deletion, are represented. *Denotes the position with the variation. Deleted nucleotide is in bold. The homopolymeric region in the vicinity of the mutation is highlighted in grey.

This fact prompted us to develop our own filters to eliminate artefacts using a customized software, called GSDot. The data management is an important and laborious task. We have tried to establish and apply appropriate successive filters to make the posterior analysis of variants easier. After GSDot, a manual prioritization of the remaining variations was performed. In these two steps the initial variant number (4674) was drastically reduced to 8, which were classified following several criteria.

3.5 Additional mutations in control group

The study led us to identify additional mutations in 12 out of the 47 patients included in control group. In four patients mutations have been missed in Sanger sequencing. It was due to the presence of a polymorphism in the primer sequencing or an error in sequences reading. In the remaining eight patients, the mutations were detected in genes not previously studied by Sanger. Among these patients, there were three patients included in Chapter III, who remained without mutations after USH2 genes screening. After NGS sequencing, we detected mutations in *MYO7A*, *CDH23* and *CLRN1,* genes not frequently mutated in clinical type USH2. An advantage of whole Usher exome study is its ability to detect mutations in genes typically involved in other clinical type. It facilitates the molecular diagnosis of atypical patients or non-clinical classified patients.

3.6 Detection ratio in new USH patients

When our workflow and bioinformatics pipeline were properly defined on the basis of the results obtained in the control group, we proceeded with the analysis of 13 new Usher patients not studied before. The mutation detection ratio was estimated in 85% with two mutations identified in 10 out of 13 patients. Although the number of patients included in this cohort is low, the results are promising.

3.7 Non-syndromic hearing loss cohort

Finally, 11 subjects with non-syndromic hearing loss were studied following the same strategy. Mutations were identified in *TMC1*, *CDH23*, *OTOF* and *MYO7A* in four unrelated patients. One patient was carrier of the mutation c.35delG (*GJB2*) but GS Reference Mapper did not detect this small deletion, maybe because it was located in a stretch of six G. *GJB2* is a small gene responsible for around 50% of hereditary cases with ARNSHL (autosomal recessive nonsyndromic hearing loss). In this way, the best approach would be to analyse this gene using Sanger sequencing, and when the result is negative to continue with the study of the rest of genes using NGS. Massively parallel sequencing has already been tested in hereditary hearing loss in several studies in an efficient way (SHEARER *et al.* 2010; BROWNSTEIN *et al.* 2011; DE KEULENAER *et al.* 2012) and in all of them *GJB2* was previously analysed by direct sequencing.

3.8 Digenism or oligogenism

Sequencing of all Usher genes is a great test to assay the presence of digenic or oligogenic inheritance in Usher syndrome. In our cohort we have found two patients with pathogenic mutations in several genes. For different reasons (see manuscript in Chapter IV) we do not suggest a digenism or oligogenism in these subjects. However, it has already been reported by Ebermann *et al*. and Zheng *et al*. and we cannot discard the possibility of finding other similar cases (ZHENG *et al.* 2005; EBERMANN *et al.* 2010).

3.9 Future prospects

After sequencing of the Usher exome, 18 patients have still remained without mutations and 17 patients with only one mutation identified. It is remarkable that 10 out of 17 patients with one pathogenic variant bore the mutation in the *USH2A* gene, and in four cases the variant was p.C759F. Regardless these patients could carry mutation in regions non included in the design (as deep intronic regions), the number of patients with an incompletely genotype is quite high. Hence, it would be also possible that some subjects were solely carriers of a mutation in an USH gene while the real causative gene is another one.

Recently, Licastro *et al*. used Long-PCR and WES (Whole exome sequencing) for molecular diagnostic of Usher syndrome (LICASTRO *et al.* 2012). However, the coverage obtained was not homogeneous and satisfactory; using WES only 50% of regions were covered more than 25X, resulting inappropriate in a molecular diagnosis context. In this study, we have assayed a new approach and the combined targeted capture and NGS appears to be an accurate strategy to study the Usher genes. Although slight improvements will be needed, the results are promising and a future transfer to diagnostics will be possible.

4. Global perspectives and future directions

The studies described in this thesis provide a strong basis toward a better understanding of molecular mechanisms involved in the Usher syndrome. Molecular diagnosis of USH patients is challenging and the reported mutational spectrum confirms the need of including alternative approaches to identify mutations not detected by conventional PCR techniques (for example large rearrangements). Moreover, the deep intronic mutation recently described in *USH2A* shows the requirement of analysing the transcripts of Usher genes to detect this type of mutations in other genes and, in this way, to improve the molecular diagnosis of patients who remain with an incomplete genotype.

The elucidation of the mutations responsible for the disease is needed to develop targeted therapies for Usher syndrome, such as gene replacement or suppression of nonsense mutations by aminoglycosides or similar compounds. Moreover, the finding of a pseudoexon-activating mutation opens the way toward a new type of therapy using antisense oligonucleotides.

In most cases, NGS is an effective way to study all the Usher genes. However, there are some patients in whom mutations could not be identified. For these subjects, after an in-depth phenotyping, it would be interesting to perform a WES. WES would lead us to identify mutations in genes involved in other disorders with overlapping phenotype to Usher syndrome, or to identify novel genes responsible for USH.

CHAPTER VI

Conclusions

Conclusions

1. Mutations in the *USH2A* gene are responsible for the 76.1% of USH2 patients in the Spanish population. After c.2299delG, p.C3267R is the most frequent mutation. Hot spots are not observed and most of the mutations are private.

2. *GPR98* is the second most prevalent USH2 gene, with an involvement of 5.2% in USH2 patients of Spanish origin. There are not recurrent mutations and missense mutations are less frequent than in *USH2A*. Mutations are distributed along the entire gene, with a slightly accumulation in the last exons.

3. The involvement of *DFNB31* in the Spanish population is very low. To date, no pathogenic mutation in *DFNB31* has been identified in our cohort.

4. Minigenes are an accurate approach to test the effect of the variants in the splicing mechanisms. *In silico* predictions and functional analyses are needed to determine the pathogenicity of the variants in a properly way.

5. Large rearrangements in the USH genes are responsible for Usher syndrome. Array-CGH analyses have led us to identify two large deletions and the exact breakpoints. Techniques which allow detecting large rearrangements should be included in molecular diagnosis in a routinely way.

6. The deep intronic mutation c.7595-2144A>G produces the activation of a pseudoexon in the *USH2A* gene and causes USH2. This mutation is quite frequent; it has been detected in 9 patients with USH2. Further analyses of transcripts in Usher genes will allow us to detect additional mutations in profound intronic regions. This mutation offers the possibility of a new therapeutic approach using AONs.

7. USH3 is the less frequent type of USH in our population. Mutations in *CLRN1* are found only in a 2.7% of all USH patients.

8. Targeted capture of the Usher exome, coupled to massively parallel sequencing offers a better cost-effective and faster molecular diagnostic assay compared to the conventional gene-by-gene Sanger sequencing approach.

9. The Usher panel facilitates the molecular diagnosis of patients with an atypical phenotype or without classification due to lack of clinical data. This approach let us to identify mutations in genes typically involved in other clinical subtype and to detect possible cases of digenic or oligogenic inheritance.

10. Successful identification of pathogenic mutations demonstrates that we have developed an accurate experimental and bioinformatics pipeline. Gene targeted sequencing is a highly effective tool for genetic studies of heterogeneous disorders.

CHAPTER VII

General bibliography

ABADIE, C., C. BLANCHET, D. BAUX, L. LARRIEU, T. BESNARD *et al.*, 2012 Audiological findings in 100 USH2 patients. Clin Genet **82:** 433-438.

ABEYSINGHE, S. S., N. CHUZHANOVA, M. KRAWCZAK, E. V. BALL and D. N. COOPER, 2003 Translocation and gross deletion breakpoints in human inherited disease and cancer I: Nucleotide composition and recombination-associated motifs. Hum Mutat **22:** 229-244.

ADATO, A., S. VREUGDE, T. JOENSUU, N. AVIDAN, R. HAMALAINEN *et al.*, 2002 USH3A transcripts encode clarin-1, a four-transmembrane-domain protein with a possible role in sensory synapses. Eur J Hum Genet **10:** 339-350.

ADATO, A., V. MICHEL, Y. KIKKAWA, J. REINERS, K. N. ALAGRAMAM *et al.*, 2005a Interactions in the network of Usher syndrome type 1 proteins. Hum Mol Genet **14:** 347-356.

ADATO, A., G. LEFEVRE, B. DELPRAT, V. MICHEL, N. MICHALSKI *et al.*, 2005b Usherin, the defective protein in Usher syndrome type IIA, is likely to be a component of interstereocilia ankle links in the inner ear sensory cells. Hum Mol Genet **14:** 3921-3932.

AHMED, Z. M., S. RIAZUDDIN, S. L. BERNSTEIN, Z. AHMED, S. KHAN *et al.*, 2001 Mutations of the protocadherin gene PCDH15 cause Usher syndrome type 1F. Am J Hum Genet **69:** 25-34.

AHMED, Z. M., S. RIAZUDDIN, J. AHMAD, S. L. BERNSTEIN, Y. GUO *et al.*, 2003 PCDH15 is expressed in the neurosensory epithelium of the eye and ear and mutant alleles are responsible for both USH1F and DFNB23. Hum Mol Genet **12:** 3215-3223.

AHMED, Z. M., S. RIAZUDDIN, S. N. KHAN, P. L. FRIEDMAN and T. B. FRIEDMAN, 2009 USH1H, a novel locus for type I Usher syndrome, maps to chromosome 15q22-23. Clin Genet **75:** 86-91.

ALAGRAMAM, K. N., C. L. MURCIA, H. Y. KWON, K. S. PAWLOWSKI, C. G. WRIGHT *et al.*, 2001a The mouse Ames waltzer hearing-loss mutant is caused by mutation of Pcdh15, a novel protocadherin gene. Nat Genet **27:** 99-102.

ALAGRAMAM, K. N., H. YUAN, M. H. KUEHN, C. L. MURCIA, S. WAYNE *et al.*, 2001b Mutations in the novel protocadherin PCDH15 cause Usher syndrome type 1F. Hum Mol Genet **10:** 1709-1718.

ALAGRAMAM, K. N., R. J. GOODYEAR, R. GENG, D. N. FURNESS, A. F. VAN AKEN *et al.*, 2011 Mutations in protocadherin 15 and cadherin 23 affect tip links and mechanotransduction in mammalian sensory hair cells. PLoS One **6:** e19183.

ALLER, E., C. NAJERA, J. M. MILLAN, J. S. OLTRA, H. PEREZ-GARRIGUES *et al.*, 2004 Genetic analysis of 2299delG and C759F mutations (USH2A) in patients with visual and/or auditory impairments. Eur J Hum Genet **12:** 407-410.

ALLER, E., T. JAIJO, M. BENEYTO, C. NAJERA, S. OLTRA *et al.*, 2006 Identification of 14 novel mutations in the long isoform of USH2A in Spanish patients with Usher syndrome type II. J Med Genet **43:** e55.

ALLER, E., T. JAIJO, G. GARCIA-GARCIA, M. J. APARISI, D. BLESA *et al.*, 2010a Identification of large rearrangements of the PCDH15 gene by combined MLPA and a CGH: large duplications are responsible for Usher syndrome. Invest Ophthalmol Vis Sci **51:** 5480-5485.

ALLER, E., L. LARRIEU, T. JAIJO, D. BAUX, C. ESPINOS *et al.*, 2010b The USH2A c.2299delG mutation: dating its common origin in a Southern European population. Eur J Hum Genet **18:** 788-793.

AVILA-FERNANDEZ, A., D. CANTALAPIEDRA, E. ALLER, E. VALLESPIN, J. AGUIRRE-LAMBAN *et al.*, 2010 Mutation analysis of 272 Spanish families affected by autosomal recessive retinitis pigmentosa using a genotyping microarray. Mol Vis **16:** 2550-2558.

BARKER, S. E., C. A. BRODERICK, S. J. ROBBIE, Y. DURAN, M. NATKUNARAJAH *et al.*, 2009 Subretinal delivery of adeno-associated virus serotype 2 results in minimal immune responses that allow repeat vector administration in immunocompetent mice. J Gene Med **11:** 486-497.

BAUX, D., L. LARRIEU, C. BLANCHET, C. HAMEL, S. BEN SALAH *et al.*, 2007 Molecular and in silico analyses of the full-length isoform of usherin identify new pathogenic alleles in Usher type II patients. Hum Mutat **28:** 781-789.

BELYANTSEVA, I. A., E. T. BOGER, S. NAZ, G. I. FROLENKOV, J. R. SELLERS *et al.*, 2005 Myosin-XVa is required for tip localization of whirlin and differential elongation of hair-cell stereocilia. Nat Cell Biol **7:** 148-156.

BELL, J., 1922 *Retinitis pigmentosa and allied diseases*. Cambridge University Press, London.

BESNARD, T., C. VACHE, D. BAUX, L. LARRIEU, C. ABADIE *et al.*, 2012 Non-USH2A mutations in USH2 patients. Hum Mutat **33:** 504-510.

BEUF, K. D., J. D. SCHRIJVER, O. THAS, W. V. CRIEKINGE, R. A. IRIZARRY *et al.*, 2012 Improved base-calling and quality scores for 454 sequencing based on a Hurdle Poisson model. BMC Bioinformatics **13:** 303.

BHATTACHARYA, G., R. KALLURI, D. J. ORTEN, W. J. KIMBERLING and D. COSGROVE, 2004 A domain-specific usherin/collagen IV interaction may be required for stable integration into the basement membrane superstructure. J Cell Sci **117:** 233-242.

BHATTACHARYA, G., and D. COSGROVE, 2005 Evidence for functional importance of usherin/fibronectin interactions in retinal basement membranes. Biochemistry **44:** 11518-11524.

BITNER-GLINDZICZ, M., K. J. LINDLEY, P. RUTLAND, D. BLAYDON, V. V. SMITH *et al.*, 2000 A recessive contiguous gene deletion causing infantile hyperinsulinism, enteropathy and deafness identifies the Usher type 1C gene. Nat Genet **26:** 56-60.

BOEDA, B., A. EL-AMRAOUI, A. BAHLOUL, R. GOODYEAR, L. DAVIET *et al.*, 2002 Myosin VIIa, harmonin and cadherin 23, three Usher I gene products that cooperate to shape the sensory hair cell bundle. EMBO J **21:** 6689-6699.

BOLZ, H., B. VON BREDERLOW, A. RAMIREZ, E. C. BRYDA, K. KUTSCHE *et al.*, 2001 Mutation of CDH23, encoding a new member of the cadherin gene family, causes Usher syndrome type 1D. Nat Genet **27:** 108-112.

BONNET, C., M. GRATI, S. MARLIN, J. LEVILLIERS, J. P. HARDELIN *et al.*, 2011 Complete exon sequencing of all known Usher syndrome genes greatly improves molecular diagnosis. Orphanet J Rare Dis **6:** 21.

BONNET, C., and A. EL-AMRAOUI, 2012 Usher syndrome (sensorineural deafness and retinitis pigmentosa): pathogenesis, molecular diagnosis and therapeutic approaches. Curr Opin Neurol **25:** 42-49.

BORK, J. M., L. M. PETERS, S. RIAZUDDIN, S. L. BERNSTEIN, Z. M. AHMED *et al.*, 2001 Usher syndrome 1D and nonsyndromic autosomal recessive deafness DFNB12 are caused by allelic mutations of the novel cadherin-like gene CDH23. Am J Hum Genet **68:** 26-37.

BOUGHMAN, J. A., M. VERNON and K. A. SHAVER, 1983 Usher syndrome: definition and estimate of prevalence from two high-risk populations. J Chronic Dis **36:** 595-603.

BROWNSTEIN, Z., L. M. FRIEDMAN, H. SHAHIN, V. ORON-KARNI, N. KOL *et al.*, 2011 Targeted genomic capture and massively parallel sequencing to identify genes for hereditary hearing loss in Middle Eastern families. Genome Biol **12:** R89.

BUSSKAMP, V., S. PICAUD, J. A. SAHEL and B. ROSKA, 2012 Optogenetic therapy for retinitis pigmentosa. Gene Ther **19:** 169-175.

CREMERS, F. P., W. J. KIMBERLING, M. KULM, A. P. DE BROUWER, E. VAN WIJK *et al.*, 2007 Development of a genotyping microarray for Usher syndrome. J Med Genet **44:** 153-160.

CHAIB, H., C. PLACE, N. SALEM, C. DODE, S. CHARDENOUX *et al.*, 1996 Mapping of DFNB12, a gene for a non-syndromal autosomal recessive deafness, to chromosome 10q21-22. Hum Mol Genet **5:** 1061-1064.

CHAIB, H., J. KAPLAN, S. GERBER, C. VINCENT, H. AYADI *et al.*, 1997 A newly identified locus for Usher syndrome type I, USH1E, maps to chromosome 21q21. Hum Mol Genet **6:** 27-31.

DAVENPORT S, O. G., 1977 The heterogeneity of Usher's syndrome., pp. Proceedings of the 5th International Conference of Birth Defects.

DE KEULENAER, S., J. HELLEMANS, S. LEFEVER, J. P. RENARD, J. DE SCHRIJVER *et al.*, 2012 Molecular diagnostics for congenital hearing loss including 15 deafness genes using a next generation sequencing platform. BMC Med Genomics **5:** 17.

DEWAN, A. T., A. R. PARRADO and S. M. LEAL, 2003 A second kindred linked to DFNA20 (17q25.3) reduces the genetic interval. Clin Genet **63:** 39-45.

DI PALMA, F., R. H. HOLME, E. C. BRYDA, I. A. BELYANTSEVA, R. PELLEGRINO *et al.*, 2001 Mutations in Cdh23, encoding a new type of cadherin, cause stereocilia disorganization in waltzer, the mouse model for Usher syndrome type 1D. Nat Genet **27:** 103-107.

DREYER, B., V. BROX, L. TRANEBJAERG, T. ROSENBERG, A. M. SADEGHI *et al.*, 2008 Spectrum of USH2A mutations in Scandinavian patients with Usher syndrome type II. Hum Mutat **29:** 451.

DROR, A. A., and K. B. AVRAHAM, 2009 Hearing loss: mechanisms revealed by genetics and cell biology. Annu Rev Genet **43:** 411-437.

EBERMANN, I., H. P. SCHOLL, P. CHARBEL ISSA, E. BECIROVIC, J. LAMPRECHT *et al.*, 2007 A novel gene for Usher syndrome type 2: mutations in the long isoform of whirlin are associated with retinitis pigmentosa and sensorineural hearing loss. Hum Genet **121:** 203-211.

EBERMANN, I., J. B. PHILLIPS, M. C. LIEBAU, R. K. KOENEKOOP, B. SCHERMER *et al.*, 2010 PDZD7 is a modifier of retinal disease and a contributor to digenic Usher syndrome. J Clin Invest **120:** 1812-1823.

EISENBERGER, T., R. SLIM, A. MANSOUR, M. NAUCK, G. NURNBERG *et al.*, 2012 Targeted next-generation sequencing identifies a homozygous nonsense mutation in ABHD12, the gene underlying PHARC, in a family clinically diagnosed with Usher syndrome type 3. Orphanet J Rare Dis **7:** 59.

ERNEST, S., G. J. RAUCH, P. HAFFTER, R. GEISLER, C. PETIT *et al.*, 2000 Mariner is defective in myosin VIIA: a zebrafish model for human hereditary deafness. Hum Mol Genet **9:** 2189-2196.

ESPINOS, C., J. M. MILLAN, M. BENEYTO and C. NAJERA, 1998 Epidemiology of Usher syndrome in Valencia and Spain. Community Genet **1:** 223-228.

ESPINOS, C., and J. M. MILLAN, 1998 [Genetic research on Usher's syndrome in Spain]. Med Clin (Barc) **110:** 340-341.

EUDY, J. D., M. D. WESTON, S. YAO, D. M. HOOVER, H. L. REHM *et al.*, 1998 Mutation of a gene encoding a protein with extracellular matrix motifs in Usher syndrome type IIa. Science **280:** 1753-1757.

FETTIPLACE, R., and C. M. HACKNEY, 2006 The sensory and motor roles of auditory hair cells. Nat Rev Neurosci **7:** 19-29.

FISHMAN, G. A., R. J. ANDERSON, B. L. LAM and D. J. DERLACKI, 1995 Prevalence of foveal lesions in type 1 and type 2 Usher's syndrome. Arch Ophthalmol **113:** 770-773.

FROLENKOV, G. I., I. A. BELYANTSEVA, T. B. FRIEDMAN and A. J. GRIFFITH, 2004 Genetic insights into the morphogenesis of inner ear hair cells. Nat Rev Genet **5:** 489-498.

GARCIA-GARCIA, G., M. J. APARISI, T. JAIJO, R. RODRIGO, A. M. LEON *et al.*, 2011 Mutational screening of the USH2A gene in Spanish USH patients reveals 23 novel pathogenic mutations. Orphanet J Rare Dis **6:** 65.

GARCIA-GARCIA, G., T. BESNARD, D. BAUX, C. VACHE, E. ALLER *et al.*, 2013 The contribution of GPR98 and DFNB31 genes to a Spanish Usher syndrome type 2 cohort. Mol Vis **19:** 367-373.

GELLER, S. F., K. I. GUERIN, M. VISEL, A. PHAM, E. S. LEE *et al.*, 2009 CLRN1 is nonessential in the mouse retina but is required for cochlear hair cell development. PLoS Genet **5:** e1000607.

GENG, R., S. F. GELLER, T. HAYASHI, C. A. RAY, T. A. REH *et al.*, 2009 Usher syndrome IIIA gene clarin-1 is essential for hair cell function and associated neural activation. Hum Mol Genet **18:** 2748-2760.

GERBER, S., D. BONNEAU, B. GILBERT, A. MUNNICH, J. L. DUFIER *et al.*, 2006 USH1A: chronicle of a slow death. Am J Hum Genet **78:** 357-359.

GIBSON, F., J. WALSH, P. MBURU, A. VARELA, K. A. BROWN *et al.*, 1995 A type VII myosin encoded by the mouse deafness gene shaker-1. Nature **374:** 62-64.

GOLDMANN, T., N. OVERLACK, F. MOLLER, V. BELAKHOV, M. VAN WYK *et al.*, 2012 A comparative evaluation of NB30, NB54 and PTC124 in translational read-through efficacy for treatment of an USH1C nonsense mutation. EMBO Mol Med **4:** 1186-1199.

GRÄFE, V., 1858 Exceptionelles Verhalten des Gesichtsfeldes bei Pigmententartung der Netzhaut. Archiv für Ophthalmologie **4:** 250–253.

GRATI, M., J. B. SHIN, M. D. WESTON, J. GREEN, M. A. BHAT *et al.*, 2012 Localization of PDZD7 to the stereocilia ankle-link associates this scaffolding protein with the Usher syndrome protein network. J Neurosci **32:** 14288-14293.

GRONDAHL, J., and S. MJOEN, 1986 Usher syndrome in four Norwegian counties. Clin Genet **30:** 14-28.

GUALANDI, F., P. RIMESSI, C. TRABANELLI, P. SPITALI, M. NERI *et al.*, 2006 Intronic breakpoint definition and transcription analysis in DMD/BMD patients with deletion/duplication at the 5' mutation hot spot of the dystrophin gene. Gene **370:** 26-33.

HAMEL, C., 2006 Retinitis pigmentosa. Orphanet J Rare Dis **1:** 40.

HASHIMOTO, T., D. GIBBS, C. LILLO, S. M. AZARIAN, E. LEGACKI *et al.*, 2007 Lentiviral gene replacement therapy of retinas in a mouse model for Usher syndrome type 1B. Gene Ther **14:** 584-594.

HAYWOOD-WATSON, R. J., 2ND, Z. M. AHMED, S. KJELLSTROM, R. A. BUSH, Y. TAKADA *et al.*, 2006 Ames Waltzer deaf mice have reduced electroretinogram amplitudes and complex alternative splicing of Pcdh15 transcripts. Invest Ophthalmol Vis Sci **47:** 3074-3084.

HEISSLER, S. M., and D. J. MANSTEIN, 2012 Functional characterization of the human myosin-7a motor domain. Cell Mol Life Sci **69:** 299-311.

HOLME, R. H., B. W. KIERNAN, S. D. BROWN and K. P. STEEL, 2002 Elongation of hair cell stereocilia is defective in the mouse mutant whirler. J Comp Neurol **450:** 94-102.

HOPE, C. I., S. BUNDEY, D. PROOPS and A. R. FIELDER, 1997 Usher syndrome in the city of Birmingham--prevalence and clinical classification. Br J Ophthalmol **81:** 46-53.

IANNACCONE, A., S. B. KRITCHEVSKY, M. L. CICCARELLI, S. A. TEDESCO, C. MACALUSO *et al.*, 2004 Kinetics of visual field loss in Usher syndrome Type II. Invest Ophthalmol Vis Sci **45:** 784-792.

JAIJO, T., E. ALLER, G. GARCIA-GARCIA, M. J. APARISI, S. BERNAL *et al.*, 2010 Microarray-based mutation analysis of 183 Spanish families with Usher syndrome. Invest Ophthalmol Vis Sci **51:** 1311-1317.

JAIN, P. K., A. K. LALWANI, X. C. LI, T. L. SINGLETON, T. N. SMITH *et al.*, 1998 A gene for recessive nonsyndromic sensorineural deafness (DFNB18) maps to the chromosomal region 11p14-p15.1 containing the Usher syndrome type 1C gene. Genomics **50:** 290-292.

JAWOREK, T. J., R. BHATTI, N. LATIEF, S. N. KHAN, S. RIAZUDDIN *et al.*, 2012 USH1K, a novel locus for type I Usher syndrome, maps to chromosome 10p11.21-q21.1. J Hum Genet **57:** 633-637.

JOENSUU, T., R. HAMALAINEN, B. YUAN, C. JOHNSON, S. TEGELBERG *et al.*, 2001 Mutations in a novel gene with transmembrane domains underlie Usher syndrome type 3. Am J Hum Genet **69:** 673-684.

JOHNSON, K. R., L. H. GAGNON, L. S. WEBB, L. L. PETERS, N. L. HAWES *et al.*, 2003 Mouse models of USH1C and DFNB18: phenotypic and molecular analyses of two new spontaneous mutations of the Ush1c gene. Hum Mol Genet **12:** 3075-3086.

KAUPER, K., C. MCGOVERN, S. SHERMAN, P. HEATHERTON, R. RAPOZA *et al.*, 2012 Two-year intraocular delivery of ciliary neurotrophic factor by encapsulated cell technology implants in patients with chronic retinal degenerative diseases. Invest Ophthalmol Vis Sci **53:** 7484-7491.

KAZMIERCZAK, P., H. SAKAGUCHI, J. TOKITA, E. M. WILSON-KUBALEK, R. A. MILLIGAN *et al.*, 2007 Cadherin 23 and protocadherin 15 interact to form tip-link filaments in sensory hair cells. Nature **449:** 87-91.

KELLEY, P. M., M. D. WESTON, Z. Y. CHEN, D. J. ORTEN, T. HASSON *et al.*, 1997 The genomic structure of the gene defective in Usher syndrome type Ib (MYO7A). Genomics **40:** 73-79.

KEMPERMAN, M. H., E. M. DE LEENHEER, P. L. HUYGEN, E. VAN WIJK, G. VAN DUIJNHOVEN *et al.*, 2004 A Dutch family with hearing loss linked to the DFNA20/26 locus: longitudinal analysis of hearing impairment. Arch Otolaryngol Head Neck Surg **130:** 281-288.

KIKKAWA, Y., H. SHITARA, S. WAKANA, Y. KOHARA, T. TAKADA *et al.*, 2003 Mutations in a new scaffold protein Sans cause deafness in Jackson shaker mice. Hum Mol Genet **12:** 453-461.

KIMBERLING, W. J., M. D. WESTON, C. MOLLER, S. L. DAVENPORT, Y. Y. SHUGART *et al.*, 1990 Localization of Usher syndrome type II to chromosome 1q. Genomics **7:** 245-249.

Kimberling, W. J., C. G. Moller, S. Davenport, I. A. Priluck, P. H. Beighton *et al.*, 1992 Linkage of Usher syndrome type I gene (USH1B) to the long arm of chromosome 11. Genomics **14:** 988-994.

KIMBERLING, W. J., M. D. WESTON, C. MOLLER, A. VAN AAREM, C. W. CREMERS *et al.*, 1995 Gene mapping of Usher syndrome type IIa: localization of the gene to a 2.1-cM segment on chromosome 1q41. Am J Hum Genet **56:** 216-223.

KREMER, H., E. VAN WIJK, T. MARKER, U. WOLFRUM and R. ROEPMAN, 2006 Usher syndrome: molecular links of pathogenesis, proteins and pathways. Hum Mol Genet **15 Spec No 2:** R262-270.

LANE, P. W., 1963 Whirler Mice: A Recessive Behavior Mutation in Linkage Group Viii. J Hered **54:** 263-266.

LE GUEDARD, S., V. FAUGERE, S. MALCOLM, M. CLAUSTRES and A. F. ROUX, 2007 Large genomic rearrangements within the PCDH15 gene are a significant cause of USH1F syndrome. Mol Vis **13:** 102-107.

LE QUESNE STABEJ, P., Z. SAIHAN, N. RANGESH, H. B. STEELE-STALLARD, J. AMBROSE *et al.*, 2012 Comprehensive sequence analysis of nine Usher syndrome genes in the UK National Collaborative Usher Study. J Med Genet **49:** 27-36.

LENTZ, J. J., F. M. JODELKA, A. J. HINRICH, K. E. MCCAFFREY, H. E. FARRIS *et al.*, 2013 Rescue of hearing and vestibular function by antisense oligonucleotides in a mouse model of human deafness. Nat Med.

LIBBY, R. T., and K. P. STEEL, 2001 Electroretinographic anomalies in mice with mutations in Myo7a, the gene involved in human Usher syndrome type 1B. Invest Ophthalmol Vis Sci **42:** 770-778.

LICASTRO, D., M. MUTARELLI, I. PELUSO, K. NEVELING, N. WIESKAMP *et al.*, 2012 Molecular diagnosis of Usher syndrome: application of two different next generation sequencing-based procedures. PLoS One **7:** e43799.

LIU, P., H. LI, X. REN, H. MAO, Q. ZHU *et al.*, 2008 Novel ACTG1 mutation causing autosomal dominant non-syndromic hearing impairment in a Chinese family. J Genet Genomics **35:** 553-558.

LIU, X., O. V. BULGAKOV, K. N. DARROW, B. PAWLYK, M. ADAMIAN *et al.*, 2007 Usherin is required for maintenance of retinal photoreceptors and normal development of cochlear hair cells. Proc Natl Acad Sci U S A **104:** 4413-4418.

LIU, X. Z., J. WALSH, Y. TAMAGAWA, K. KITAMURA, M. NISHIZAWA *et al.*, 1997a Autosomal dominant non-syndromic deafness caused by a mutation in the myosin VIIA gene. Nat Genet **17:** 268-269.

LIU, X. Z., J. WALSH, P. MBURU, J. KENDRICK-JONES, M. J. COPE *et al.*, 1997b Mutations in the myosin VIIA gene cause non-syndromic recessive deafness. Nat Genet **16:** 188-190.

LOMAN, N. J., R. V. MISRA, T. J. DALLMAN, C. CONSTANTINIDOU, S. E. GHARBIA *et al.*, 2012 Performance comparison of benchtop high-throughput sequencing platforms. Nat Biotechnol **30:** 434-439.

LU, B., S. WANG, P. J. FRANCIS, T. LI, D. M. GAMM *et al.*, 2010 Cell transplantation to arrest early changes in an ush2a animal model. Invest Ophthalmol Vis Sci **51:** 2269-2276.

MAERKER, T., E. VAN WIJK, N. OVERLACK, F. F. KERSTEN, J. MCGEE *et al.*, 2008 A novel Usher protein network at the periciliary reloading point between molecular transport machineries in vertebrate photoreceptor cells. Hum Mol Genet **17:** 71-86.

MANOR, U., A. DISANZA, M. GRATI, L. ANDRADE, H. LIN *et al.*, 2011 Regulation of stereocilia length by myosin XVa and whirlin depends on the actin-regulatory protein Eps8. Curr Biol **21:** 167-172.

MARTIN, D. M., and Y. RAPHAEL, 2012 Have you heard? Viral-mediated gene therapy restores hearing. Neuron **75:** 188-190.

MBURU, P., X. Z. LIU, J. WALSH, D. SAW, JR., M. J. COPE *et al.*, 1997 Mutation analysis of the mouse myosin VIIA deafness gene. Genes Funct **1:** 191-203.

MBURU, P., M. MUSTAPHA, A. VARELA, D. WEIL, A. EL-AMRAOUI *et al.*, 2003 Defects in whirlin, a PDZ domain molecule involved in stereocilia elongation, cause deafness in the whirler mouse and families with DFNB31. Nat Genet **34:** 421-428.

MCGEE, J., R. J. GOODYEAR, D. R. MCMILLAN, E. A. STAUFFER, J. R. HOLT *et al.*, 2006 The very large G-protein-coupled receptor VLGR1: a component of the ankle link complex required for the normal development of auditory hair bundles. J Neurosci **26:** 6543-6553.

MCMILLAN, D. R., K. M. KAYES-WANDOVER, J. A. RICHARDSON and P. C. WHITE, 2002 Very large G protein-coupled receptor-1, the largest known cell surface protein, is highly expressed in the developing central nervous system. J Biol Chem **277:** 785-792.

MCMILLAN, D. R., and P. C. WHITE, 2004 Loss of the transmembrane and cytoplasmic domains of the very large G-protein-coupled receptor-1 (VLGR1 or Mass1) causes audiogenic seizures in mice. Mol Cell Neurosci **26:** 322-329.

MICHALSKI, N., V. MICHEL, A. BAHLOUL, G. LEFEVRE, J. BARRAL *et al.*, 2007 Molecular characterization of the ankle-link complex in cochlear hair cells and its role in the hair bundle functioning. J Neurosci **27:** 6478-6488.

MILLAN, J. M., E. ALLER, T. JAIJO, F. BLANCO-KELLY, A. GIMENEZ-PARDO *et al.*, 2011 An update on the genetics of usher syndrome. J Ophthalmol **2011:** 417217.

MORELL, R. J., K. H. FRIDERICI, S. WEI, J. L. ELFENBEIN, T. B. FRIEDMAN *et al.*, 2000 A new locus for late-onset, progressive, hereditary hearing loss DFNA20 maps to 17q25. Genomics **63:** 1-6.

MULLER, U., 2008 Cadherins and mechanotransduction by hair cells. Curr Opin Cell Biol **20:** 557-566.

MUSTAPHA, M., E. CHOUERY, S. CHARDENOUX, M. NABOULSI, J. PARONNAUD *et al.*, 2002a DFNB31, a recessive form of sensorineural hearing loss, maps to chromosome 9q32-34. Eur J Hum Genet **10:** 210-212.

MUSTAPHA, M., E. CHOUERY, D. TORCHARD-PAGNEZ, S. NOUAILLE, A. KHRAIS *et al.*, 2002b A novel locus for Usher syndrome type I, USH1G, maps to chromosome 17q24-25. Hum Genet **110:** 348-350.

NAJERA, C., M. BENEYTO, J. BLANCA, E. ALLER, A. FONTCUBERTA *et al.*, 2002 Mutations in myosin VIIA (MYO7A) and usherin (USH2A) in Spanish patients with Usher syndrome types I and II, respectively. Hum Mutat **20:** 76-77.

NESS, S. L., T. BEN-YOSEF, A. BAR-LEV, A. C. MADEO, C. C. BREWER *et al.*, 2003 Genetic homogeneity and phenotypic variability among Ashkenazi Jews with Usher syndrome type III. J Med Genet **40:** 767-772.

NEVELING, K., R. W. COLLIN, C. GILISSEN, R. A. VAN HUET, L. VISSER *et al.*, 2012 Next-generation genetic testing for retinitis pigmentosa. Hum Mutat **33:** 963-972.

PAKARINEN, L., K. TUPPURAINEN, P. LAIPPALA, M. MANTYJARVI and H. PUHAKKA, 1995 The ophthalmological course of Usher syndrome type III. Int Ophthalmol **19:** 307-311.

PAN, L., and M. ZHANG, 2012 Structures of usher syndrome 1 proteins and their complexes. Physiology (Bethesda) **27:** 25-42.

PEARSON, R. A., A. C. BARBER, M. RIZZI, C. HIPPERT, T. XUE *et al.*, 2012 Restoration of vision after transplantation of photoreceptors. Nature **485:** 99-103.

PENNINGS, R. J., P. L. HUYGEN, D. J. ORTEN, M. WAGENAAR, A. VAN AAREM *et al.*, 2004 Evaluation of visual impairment in Usher syndrome 1b and Usher syndrome 2a. Acta Ophthalmol Scand **82:** 131-139.

PEREZ, B., L. RODRIGUEZ-PASCAU, L. VILAGELIU, D. GRINBERG, M. UGARTE *et al.*, 2010 Present and future of antisense therapy for splicing modulation in inherited metabolic disease. J Inherit Metab Dis **33:** 397-403.

PETIT, C., 2001 Usher syndrome: from genetics to pathogenesis. Annu Rev Genomics Hum Genet **2:** 271-297.

PHILLIPS, J. B., B. BLANCO-SANCHEZ, J. J. LENTZ, A. TALLAFUSS, K. KHANOBDEE *et al.*, 2011 Harmonin (Ush1c) is required in zebrafish Muller glial cells for photoreceptor synaptic development and function. Dis Model Mech **4:** 786-800.

PIEKE-DAHL, S., C. G. MOLLER, P. M. KELLEY, L. M. ASTUTO, C. W. CREMERS *et al.*, 2000 Genetic heterogeneity of Usher syndrome type II: localisation to chromosome 5q. J Med Genet **37:** 256-262.

REINERS, J., B. REIDEL, A. EL-AMRAOUI, B. BOEDA, I. HUBER *et al.*, 2003 Differential distribution of harmonin isoforms and their possible role in Usher-1 protein complexes in mammalian photoreceptor cells. Invest Ophthalmol Vis Sci **44:** 5006-5015.

RENDTORFF, N. D., M. ZHU, T. FAGERHEIM, T. L. ANTAL, M. JONES *et al.*, 2006 A novel missense mutation in ACTG1 causes dominant deafness in a Norwegian DFNA20/26 family, but ACTG1 mutations are not frequent among families with hereditary hearing impairment. Eur J Hum Genet **14:** 1097-1105.

RIAZUDDIN, S., S. NAZLI, Z. M. AHMED, Y. YANG, F. ZULFIQAR *et al.*, 2008 Mutation spectrum of MYO7A and evaluation of a novel nonsyndromic deafness DFNB2 allele with residual function. Hum Mutat **29:** 502-511.

RIAZUDDIN, S., I. A. BELYANTSEVA, A. P. GIESE, K. LEE, A. A. INDZHYKULIAN *et al.*, 2012 Alterations of the CIB2 calcium- and integrin-binding protein cause Usher syndrome type 1J and nonsyndromic deafness DFNB48. Nat Genet **44:** 1265-1271.

RINCON, A., C. AGUADO, L. R. DESVIAT, R. SANCHEZ-ALCUDIA, M. UGARTE *et al.*, 2007 Propionic and methylmalonic acidemia: antisense therapeutics for intronic variations causing aberrantly spliced messenger RNA. Am J Hum Genet **81:** 1262-1270.

RIVOLTA, C., E. A. SWEKLO, E. L. BERSON and T. P. DRYJA, 2000 Missense mutation in the USH2A gene: association with recessive retinitis pigmentosa without hearing loss. Am J Hum Genet **66:** 1975-1978.

ROSENBERG, T., M. HAIM, A. M. HAUCH and A. PARVING, 1997 The prevalence of Usher syndrome and other retinal dystrophy-hearing impairment associations. Clin Genet **51:** 314-321.

ROUGET-QUERMALET, V., J. GIUSTINIANI, A. MARIE-CARDINE, G. BEAUD, F. BESNARD *et al.*, 2006 Protocadherin 15 (PCDH15): a new secreted isoform and a potential marker for NK/T cell lymphomas. Oncogene **25:** 2807-2811.

ROUX, A. F., V. FAUGERE, C. VACHE, D. BAUX, T. BESNARD *et al.*, 2011 Four-year follow-up of diagnostic service in USH1 patients. Invest Ophthalmol Vis Sci **52:** 4063-4071.

SAHLY, I., A. EL-AMRAOUI, M. ABITBOL, C. PETIT and J. L. DUFIER, 1997 Expression of myosin VIIA during mouse embryogenesis. Anat Embryol (Berl) **196:** 159-170.

SAHLY, I., E. DUFOUR, C. SCHIETROMA, V. MICHEL, A. BAHLOUL *et al.*, 2012 Localization of Usher 1 proteins to the photoreceptor calyceal processes, which are absent from mice. J Cell Biol **199:** 381-399.

SAIHAN, Z., A. R. WEBSTER, L. LUXON and M. BITNER-GLINDZICZ, 2009 Update on Usher syndrome. Curr Opin Neurol **22:** 19-27.

SAKAGUCHI, H., J. TOKITA, U. MULLER and B. KACHAR, 2009 Tip links in hair cells: molecular composition and role in hearing loss. Curr Opin Otolaryngol Head Neck Surg **17:** 388-393.

SANKILA, E. M., L. PAKARINEN, H. KAARIAINEN, K. AITTOMAKI, S. KARJALAINEN *et al.*, 1995 Assignment of an Usher syndrome type III (USH3) gene to chromosome 3q. Hum Mol Genet **4:** 93-98.

SEILER, C., K. C. FINGER-BAIER, O. RINNER, Y. V. MAKHANKOV, H. SCHWARZ *et al.*, 2005 Duplicated genes with split functions: independent roles of protocadherin15 orthologues in zebrafish hearing and vision. Development **132:** 615-623.

SHEARER, A. E., A. P. DELUCA, M. S. HILDEBRAND, K. R. TAYLOR, J. GURROLA, 2ND *et al.*, 2010 Comprehensive genetic testing for hereditary hearing loss using massively parallel sequencing. Proc Natl Acad Sci U S A **107:** 21104-21109.

SIEMENS, J., P. KAZMIERCZAK, A. REYNOLDS, M. STICKER, A. LITTLEWOOD-EVANS *et al.*, 2002 The Usher syndrome proteins cadherin 23 and harmonin form a complex by means of PDZ-domain interactions. Proc Natl Acad Sci U S A **99:** 14946-14951.

SIEVING, P. A., R. C. CARUSO, W. TAO, H. R. COLEMAN, D. J. THOMPSON *et al.*, 2006 Ciliary neurotrophic factor (CNTF) for human retinal degeneration: phase I trial of CNTF delivered by encapsulated cell intraocular implants. Proc Natl Acad Sci U S A **103:** 3896-3901.

SIMONELLI, F., A. M. MAGUIRE, F. TESTA, E. A. PIERCE, F. MINGOZZI *et al.*, 2010 Gene therapy for Leber's congenital amaurosis is safe and effective through 1.5 years after vector administration. Mol Ther **18:** 643-650.

SMITH, R. J., E. C. LEE, W. J. KIMBERLING, S. P. DAIGER, M. Z. PELIAS *et al.*, 1992 Localization of two genes for Usher syndrome type I to chromosome 11. Genomics **14:** 995-1002.

SMITH, R. J., C. I. BERLIN, J. F. HEJTMANCIK, B. J. KEATS, W. J. KIMBERLING *et al.*, 1994 Clinical diagnosis of the Usher syndromes. Usher Syndrome Consortium. Am J Med Genet **50:** 32-38.

SOLLNER, C., G. J. RAUCH, J. SIEMENS, R. GEISLER, S. C. SCHUSTER *et al.*, 2004 Mutations in cadherin 23 affect tip links in zebrafish sensory hair cells. Nature **428:** 955-959.

SPANDAU, U. H., and K. ROHRSCHNEIDER, 2002 Prevalence and geographical distribution of Usher syndrome in Germany. Graefes Arch Clin Exp Ophthalmol **240:** 495-498.

SPENA, S., R. ASSELTA, M. PLATE, G. CASTAMAN, S. DUGA *et al.*, 2007 Pseudo-exon activation caused by a deep-intronic mutation in the fibrinogen gamma-chain gene as a novel mechanism for congenital afibrinogenaemia. Br J Haematol **139:** 128-132.

SULONEN, A. M., P. ELLONEN, H. ALMUSA, M. LEPISTO, S. ELDFORS *et al.*, 2011 Comparison of solution-based exome capture methods for next generation sequencing. Genome Biol **12:** R94.

SUNG, C. H., and J. Z. CHUANG, 2010 The cell biology of vision. J Cell Biol **190:** 953-963.

TAMAYO, M. L., J. E. BERNAL, G. E. TAMAYO, J. L. FRIAS, G. ALVIRA *et al.*, 1991 Usher syndrome: results of a screening program in Colombia. Clin Genet **40:** 304-311.

TAO, W., 2006 Application of encapsulated cell technology for retinal degenerative diseases. Expert Opin Biol Ther **6:** 717-726.

TSILOU, E. T., B. I. RUBIN, R. C. CARUSO, G. F. REED, A. PIKUS *et al.*, 2002 Usher syndrome clinical types I and II: could ocular symptoms and signs differentiate between the two types? Acta Ophthalmol Scand **80:** 196-201.

TUFFERY-GIRAUD, S., C. SAQUET, S. CHAMBERT and M. CLAUSTRES, 2003 Pseudoexon activation in the DMD gene as a novel mechanism for Becker muscular dystrophy. Hum Mutat **21:** 608-614.

USHER, C., 1935 Bowman lecture: On a few hereditary eye affection. Trans Ophthlamol Soc UK **55:** 164.

VACHE, C., T. BESNARD, P. LE BERRE, G. GARCIA-GARCIA, D. BAUX *et al.*, 2012 Usher syndrome type 2 caused by activation of an USH2A pseudoexon: implications for diagnosis and therapy. Hum Mutat **33:** 104-108.

VAN WIJK, E., E. KRIEGER, M. H. KEMPERMAN, E. M. DE LEENHEER, P. L. HUYGEN *et al.*, 2003 A mutation in the gamma actin 1 (ACTG1) gene causes autosomal dominant hearing loss (DFNA20/26). J Med Genet **40:** 879-884.

VAN WIJK, E., R. J. PENNINGS, H. TE BRINKE, A. CLAASSEN, H. G. YNTEMA *et al.*, 2004 Identification of 51 novel exons of the Usher syndrome type 2A (USH2A) gene that encode multiple conserved functional domains and that are mutated in patients with Usher syndrome type II. Am J Hum Genet **74:** 738-744.

VAN WIJK, E., B. VAN DER ZWAAG, T. PETERS, U. ZIMMERMANN, H. TE BRINKE *et al.*, 2006 The DFNB31 gene product whirlin connects to the Usher protein network in the cochlea and retina by direct association with USH2A and VLGR1. Hum Mol Genet **15:** 751-765.

VASTINSALO, H., R. JALKANEN, A. DINCULESCU, J. ISOSOMPPI, S. GELLER *et al.*, 2011 Alternative splice variants of the USH3A gene Clarin 1 (CLRN1). Eur J Hum Genet **19:** 30-35.

VASTINSALO, H., R. JALKANEN, C. BERGMANN, C. NEUHAUS, L. KLEEMOLA *et al.*, 2012 Extended mutation spectrum of Usher syndrome in Finland. Acta Ophthalmol.

VERPY, E., M. LEIBOVICI, I. ZWAENEPOEL, X. Z. LIU, A. GAL *et al.*, 2000 A defect in harmonin, a PDZ domain-containing protein expressed in the inner ear sensory hair cells, underlies Usher syndrome type 1C. Nat Genet **26:** 51-55.

VETRINI, F., R. TAMMARO, S. BONDANZA, E. M. SURACE, A. AURICCHIO *et al.*, 2006 Aberrant splicing in the ocular albinism type 1 gene (OA1/GPR143) is corrected in vitro by morpholino antisense oligonucleotides. Hum Mutat **27:** 420-426.

WAYNE, S., V. M. DER KALOUSTIAN, M. SCHLOSS, R. POLOMENO, D. A. SCOTT *et al.*, 1996 Localization of the Usher syndrome type ID gene (Ush1D) to chromosome 10. Hum Mol Genet **5:** 1689-1692.

WEIL, D., S. BLANCHARD, J. KAPLAN, P. GUILFORD, F. GIBSON *et al.*, 1995 Defective myosin VIIA gene responsible for Usher syndrome type 1B. Nature **374:** 60-61.

WEIL, D., A. EL-AMRAOUI, S. MASMOUDI, M. MUSTAPHA, Y. KIKKAWA *et al.*, 2003 Usher syndrome type I G (USH1G) is caused by mutations in the gene encoding SANS, a protein that associates with the USH1C protein, harmonin. Hum Mol Genet **12:** 463-471.

WESTON, M. D., M. W. LUIJENDIJK, K. D. HUMPHREY, C. MOLLER and W. J. KIMBERLING, 2004 Mutations in the VLGR1 gene implicate G-protein signaling in the pathogenesis of Usher syndrome type II. Am J Hum Genet **74:** 357-366.

WHEELER, D. A., M. SRINIVASAN, M. EGHOLM, Y. SHEN, L. CHEN *et al.*, 2008 The complete genome of an individual by massively parallel DNA sequencing. Nature **452:** 872-876.

WRIGHT, A. F., C. F. CHAKAROVA, M. M. ABD EL-AZIZ and S. S. BHATTACHARYA, 2010 Photoreceptor degeneration: genetic and mechanistic dissection of a complex trait. Nat Rev Genet **11:** 273-284.

WU, L., L. PAN, Z. WEI and M. ZHANG, 2011 Structure of MyTH4-FERM domains in myosin VIIa tail bound to cargo. Science **331:** 757-760.

YAGI, H., H. TOKANO, M. MAEDA, T. TAKABAYASHI, T. NAGANO *et al.*, 2007 Vlgr1 is required for proper stereocilia maturation of cochlear hair cells. Genes Cells **12:** 235-250.

ZALLOCCHI, M., D. T. MEEHAN, D. DELIMONT, C. ASKEW, S. GARIGE *et al.*, 2009 Localization and expression of clarin-1, the Clrn1 gene product, in auditory hair cells and photoreceptors. Hear Res **255:** 109-120.

ZHANG, K., J. J. HOPKINS, J. S. HEIER, D. G. BIRCH, L. S. HALPERIN *et al.*, 2011 Ciliary neurotrophic factor delivered by encapsulated cell intraocular implants for treatment of geographic atrophy in age-related macular degeneration. Proc Natl Acad Sci U S A **108:** 6241-6245.

ZHENG, Q. Y., D. YAN, X. M. OUYANG, L. L. DU, H. YU *et al.*, 2005 Digenic inheritance of deafness caused by mutations in genes encoding cadherin 23 and protocadherin 15 in mice and humans. Hum Mol Genet **14:** 103-111.

ZHU, M., T. YANG, S. WEI, A. T. DEWAN, R. J. MORELL *et al.*, 2003 Mutations in the gamma-actin gene (ACTG1) are associated with dominant progressive deafness (DFNA20/26). Am J Hum Genet **73:** 1082-1091.

ZOU, J., L. LUO, Z. SHEN, V. A. CHIODO, B. K. AMBATI *et al.*, 2011 Whirlin replacement restores the formation of the USH2 protein complex in whirlin knockout photoreceptors. Invest Ophthalmol Vis Sci **52:** 2343-2351.

ZRENNER, E., K. U. BARTZ-SCHMIDT, H. BENAV, D. BESCH, A. BRUCKMANN *et al.*, 2011 Subretinal electronic chips allow blind patients to read letters and combine them to words. Proc Biol Sci **278:** 1489-1497.

CHAPTER VIII

Appendix

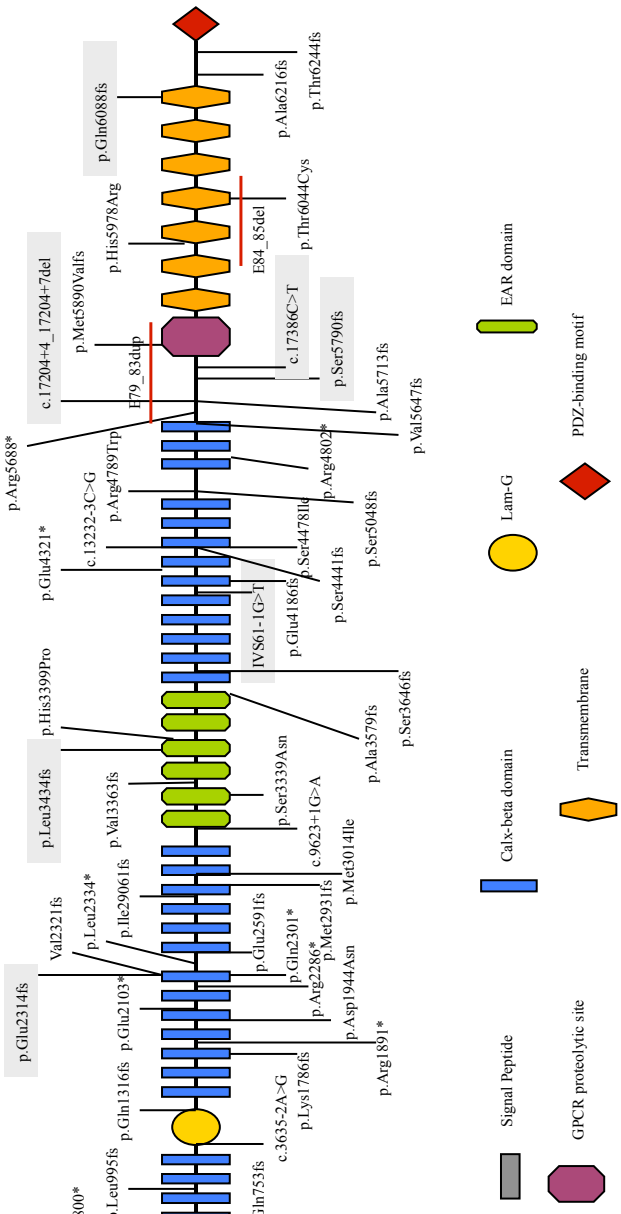


Figure S1. Representation of *GPR98* mutations reported in different studies. UV3, UV4 and pathogenic variants are included. Mutations identified in present work are highlighted in grey (García-García *et al.*, 2013).

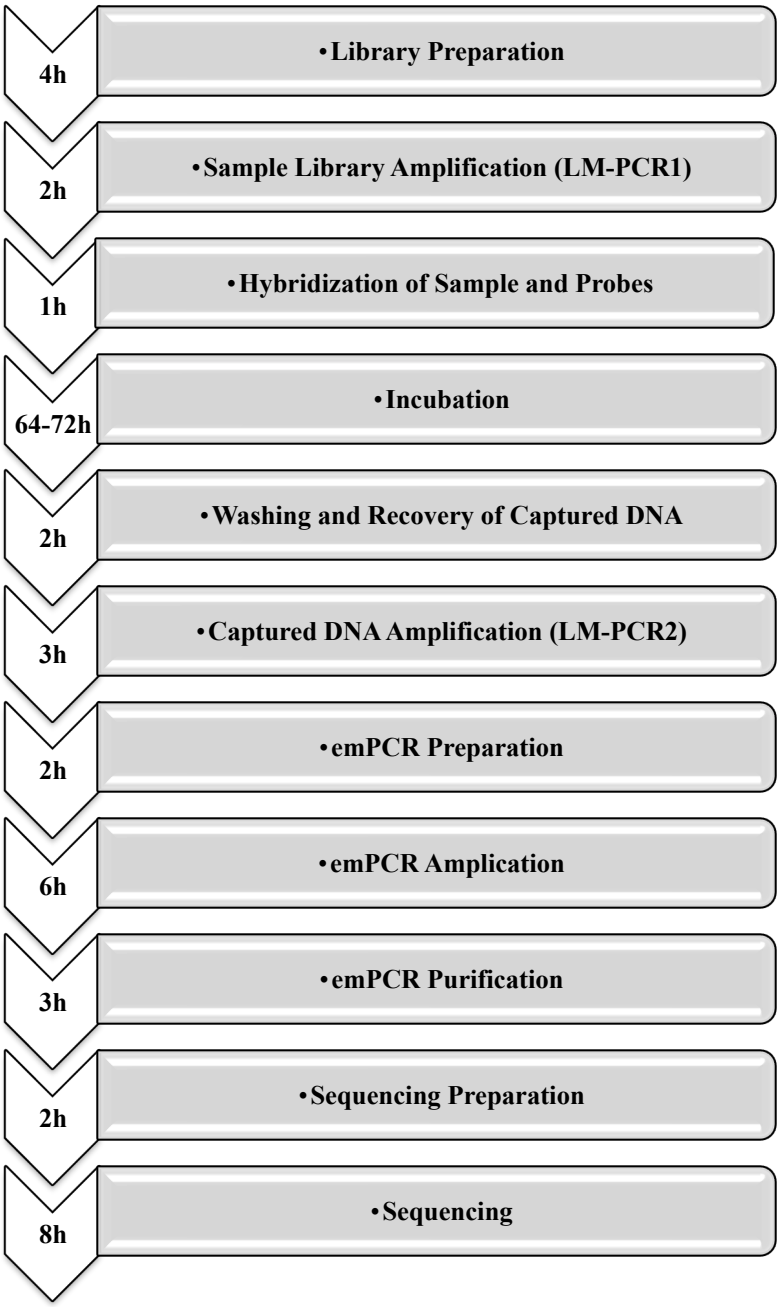
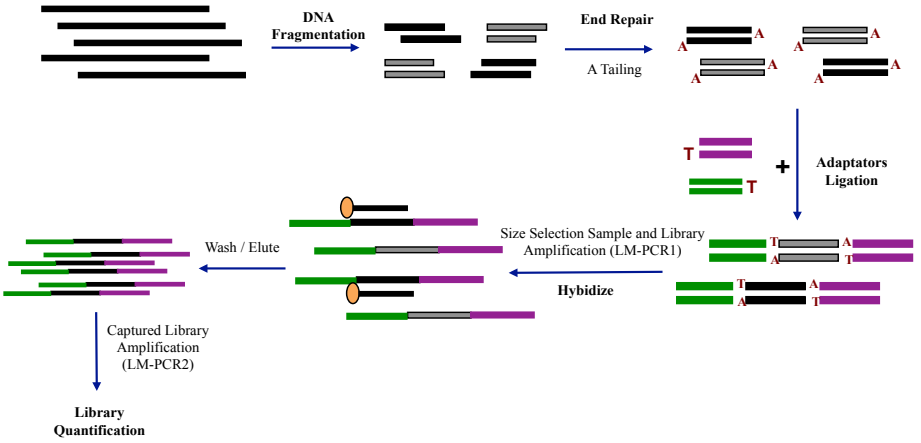
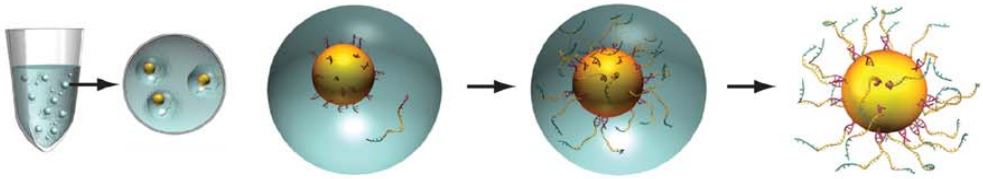


Figure S2. Detailed workflow with the approximate time needed in each step. LM-PCR (ligation-mediated PCR); emPCR (emulsion-based PCR).

A) Library Preparation



B) emPCR



Emulsify beads and PCR reagents in water-in-oil microeactors

Clonal Amplification

Break microreactors , enrich for DNA-positive beads

C) 454 Sequencing

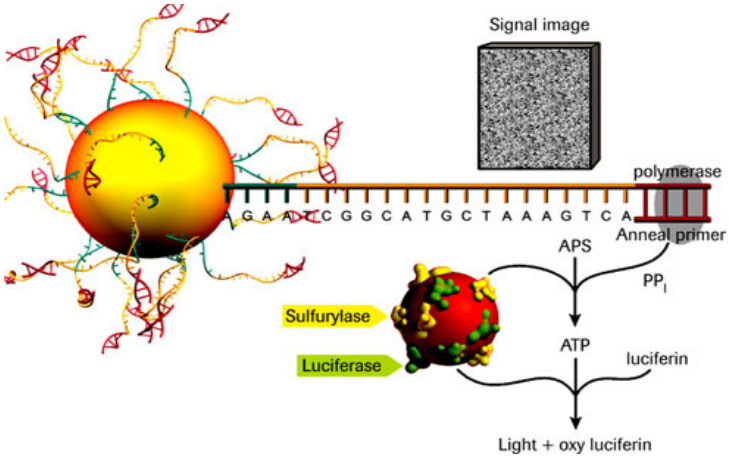


Figure S3. Representation of the three major steps. A) The first step is the preparation of the library to be sequenced. We start with 500ng of sample DNA that we fragment by nebulization, applying 2.1 bar of N_2 for 1 minute. Next, we perform the fragment End-Repair (a polish reaction to create blunt end for adaptor ligation and the A-tailing to allow for TA ligation of the library adaptor), to finally bind the adaptors by ligation. The adaptors contain sequences complementary to primers needed for the amplifications and the sequencing. Then we purify the samples to remove small fragments (<300 pb) using AMPure beads and the library is amplified. It is a LM-PCR pre-capture (LM-PCR1) and we use primers complementary to adaptors sequences. Our sample is now ready for the hybridization with the set of biotinylated oligonucleotides probes to perform the sequence capture. We incubate at 47ºC for 64-72h. After the hybridization period, we proceed with the washing and recovering of captured DNA. For this step we use magnetic beads recovered of streptavidin (Streptavidin Dynabeads), which bind to biotinylated probes, and these probes are specifically bind to desired regions. The captured library is now amplified, known as LM-PCR2 (same principle as LM-PCR1). Next step is the quantification of our library to use it in the emPCR (emulsion PCR). B) Our library will be clonally amplified by emPCR. The library is diluted to single-molecule concentration, denatured, and hybridized to individual beads containing sequences complementary to adapter oligonucleotides. The beads are compartmentalized into water-in-oil microvesicles, where millions of copies of each single fragment bound to the beads occur during emulsion PCR. After amplification, the emulsion is broken, and the beads containing clonally amplified template DNA are enriched, providing a library of sequencing templates. Last step before sequencing preparation is the annealing of sequencing primer. C) Now the sample is ready to be sequenced. The beads are subsequently deposited into individual wells of a picotiter plate where the pyrosequencing is carried out. Smaller enzyme-carrying beads are also loaded to generate sequencing reactors. Across sequencing cycles, four nucleotides are successively added to the reaction mixture in each well where the complementary strand of the DNA template is synthesized by the incorporation of proper nucleotides. During sequencing, the nucleotide incorporation releases an inorganic pyrophosphate (PPi). This PPi is converted to ATP by the ATP sulfurylase enzyme in presence of APS (adenosine 5' phosphosulfate). The ATP generated is used for the luciferase enzyme to the conversion of luciferin to oxyluciferin and the subsequently production of luminescence, which is detected by a charge-coupled device (CCD) camera. Unincorporated nucleotides are degraded by the apyrase before next addition of next nucleotides. The luminescence intensity produced is directly proportional to number of nucleotides incorporated in the sequence in each cycle.

Article:

Usher syndrome type 2 caused by activation of an *USH2A* pseudoexon: implications for diagnosis and therapy

Christel Vaché [(1)], Thomas Besnard [(1,2,3)], Pauline le Berre [(1)], **Gema García-García** [(2,4)], David Baux [(1)], Lise Larrieu [(1)], Caroline Abadie [(1)], Catherine Blanchet [(5)], Hanno Jörn Bolz [(6,7)], Jose Millan [(4)], Christian Hamel [(5)], Sue Malcolm [(8)], Mireille Claustres [(1,2,3)], Anne-Francoise Roux [(1,2)#]

1. CHU Montpellier, Laboratoire de Génétique Moléculaire, Montpellier, F-34000, France
2. Inserm, U827, Montpellier, F-34000, France
3. Univ, Montpellier I, Montpellier, F-34000, France
4. Grupo de Investigación en Enfermedades Neurosensoriales, Instituto de Investigación Sanitaria IIS-La Fe and CIBERER, Valencia, Spain
5. CHU Montpellier, Centre National de Référence maladies rares "Affections Sensorielles Génétiques", Montpellier, F-34000, France
6. Bioscientia Center for Human Genetics, Ingelheim, Germany
7. Institute for Human Genetics, University Hospital of Cologne, Cologne, Germany
8. Clinical and Molecular Genetics, Institute of Child Health, University College London, London, United Kingdom

Human Mutation. 2012 Jan 33; 104-8

ABSTRACT

USH2A sequencing in three affected members of a large family, referred for the recessive USH2 syndrome, identified a single pathogenic alteration in one of them and a different mutation in the two affected nieces. As the patients carried a common *USH2A* haplotype, they likely shared a mutation not found by standard sequencing techniques. Analysis of RNA from nasal cells in one affected individual identified an additional pseudoexon (PE) resulting from a deep intronic mutation. This was confirmed by minigene assay. This is the first example in Usher syndrome (USH) with a mutation causing activation of a PE. The finding of this alteration in eight other individuals of mixed European origin emphasizes the importance of including RNA analysis in a comprehensive diagnostic service. Finally, this mutation, which would not have been found by whole-exome sequencing, could offer, for the first time in USH, the possibility of therapeutic correction by antisense oligonucleotides (AONs).

Key words: Usher syndrome, pseudoexon, diagnosis, *USH2A*

Usher syndrome (USH) is an autosomal recessive disorder defined by the presence of sensorineural hearing loss together with retinitis pigmentosa (RP). Usher syndrome is the cause of more than 50% of cases of hereditary deaf-blindness and has a prevalence recently reestimated to be 1/6000 (Kimberling, et al., 2010). Three clinical types USH1, USH2 and USH3, are distinguished on the basis of severity of hearing loss and the presence or absence of vestibular dysfunction (Smith, et al., 1994).

USH type 1, the most severe form, presents with a severe-to-profound congenital hearing loss, vestibular dysfunction and prepubertal progressive RP. In USH type 2, the most frequent form, hearing loss is described as moderate-to-severe, vestibular function is normal and onset of RP is around puberty. USH type 3 is characterized by a progressive hearing loss, a variable vestibular dysfunction and RP.

This clinical heterogeneity is associated with high genetic heterogeneity and to date nine genes implicated in the disease have been identified: *MYO7A* (MIM# 276903), *USH1C* (MIM# 605242), *CDH23* (MIM# 605516), *PCDH15* (MIM# 605514), *USH1G* (MIM# 607696) for USH type 1, *USH2A* (MIM# 608400), *GPR98* (MIM# 602851), *DFNB31* (MIM# 607928) for type 2 and *CLRN1* (MIM# 606397) for type 3 (Online Mendelian Inheritance in Man (OMIM), http://www.omim.org/). In addition, the *PDZD7* gene has been described as a contributor to digenism and as a modifier of the retinal phenotype in Usher syndrome (Ebermann, et al., 2010).

In USH type 2, *USH2A* is the most commonly mutated gene. There is a recurrent c.2299delG mutation, for which a founder effect has been described (Aller, et al., 2010; Dreyer, et al., 2001) but a large proportion of published mutations are private (see USHbases https://grenada.lumc.nl/LOVD2/Usher_montpellier/USHbases.html; (Roux, et al., 2011)). They are spread throughout the 72 *USH2A* exons and their flanking intronic sequences and consist of nonsense and missense mutations, deletions, duplications, large rearrangements and splicing variants (USHbases and unpublished results). Classic molecular diagnostic analyses of *USH2A* are carried out by direct sequencing of the exons and their flanking intronic sequences, and by Quantitative Multiplex PCR of Short Fragments (QMPSF) or array comparative genomic hybridization (CGH) to detect large rearrangements. Although this diagnostic approach is, in most cases, sufficient to identify the causative mutations on both alleles, some clinically well defined USH2 patients still present with no or only a single heterozygote mutation in *USH2A*. Resolving these cases represents a major advance in complete molecular diagnosis and in understanding the mechanisms involved in the pathology. Presence of uncharacterized mutations in deep intronic sequences or in 5' and 3' untranslated regions of the *USH2A* gene have then to be considered.

This study concerns family U249 with three of the 13 members presenting with typical clinical signs of USH type 2 disorder (Fig.1).

Written informed consent to genetic testing was obtained from all participating subjects.

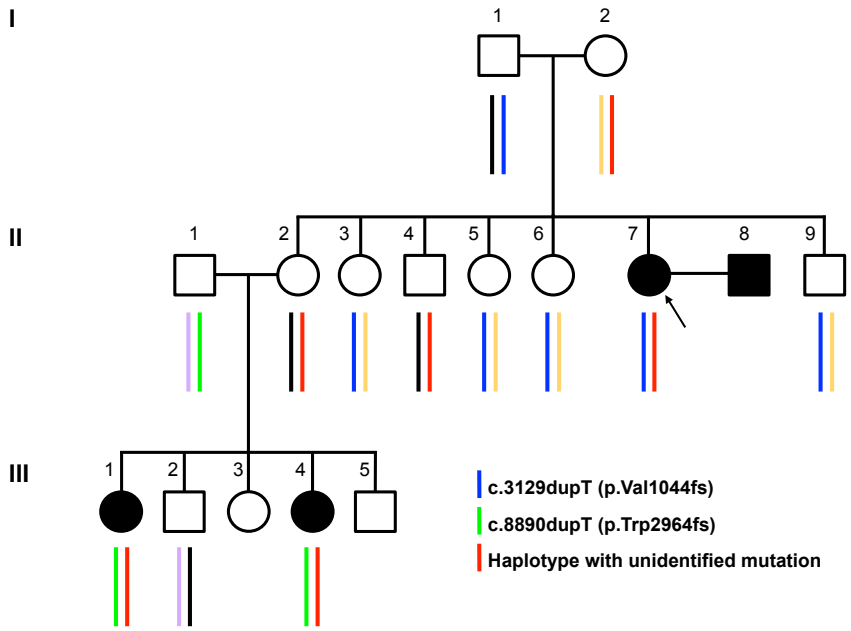


Figure 1. Pedigree of family U249. Black and white symbols indicated, respectively, affected and unaffected individuals. Patient II-7 is indicated with a black arrow. Colored vertical lines correspond to haplotypes defined by the Genethon microsatellites markers D1S474, AFM248TC1, D1S2827, D1S229, D1S227 and D1S2860. The haplotype with unidentified mutation (red) and the haplotypes carrying the identified mutations (blue and green) are indicated. The three normal haplotypes segregating in the family are identified (black, yellow and purple).

The nomenclature for the variations observed at DNA (cds) and RNA level follows the recommendations of the Human Genome Variation Society (http://www.hgvs.org) (Horaitis and Cotton, 2004) where nucleotide +1 is the A of the ATG translation initiation codon in the *USH2A* reference sequence NM_206933.2.

Molecular diagnostic investigations on *USH2A* consisting of haplotype analyses, direct sequencing of the 72 *USH2A* exons and their proximal intronic sequences, and screening for large rearrangements identified a c.8890dupT p.(Trp2964fs) mutation in patients III-1 and III-4 and a c.3129dupT p.(Val1044fs) alteration in patient II-7 transmitted by their respective fathers (Fig.1). As no other potential pathogenic *USH2A* variant was found associated with the common allele transmitted by I-2 we hypothesized that the second mutational event occurred in the *USH2A* deep intronic sequences.

To test this hypothesis, a molecular analysis at the RNA level was conducted in patient II-7. Total RNA was isolated from nasal cells of the patient and 12 controls as described previously (Vaché, et al., 2010) and used as a template in 14 overlapping RT-PCR reactions encompassing the total coding region of the *USH2A* gene. Primers (Supp. Table S1) were designed in order to include at least one heterozygous single nucleotide

polymorphism (SNP), previously identified in the patient's DNA, in each PCR reaction to verify that both alleles have been amplified.

Analysis of the 14 RT-PCR amplifications (size from 619 to 2,737 bp) revealed in patient II-7 an additional larger product obtained by PCR of exons 39-44 (Fig. 2A). This abnormal band pattern was not detected in any of the 12 controls that underwent similar amplifications between these exons. Direct sequencing of these products established that the 2 alleles were both amplified and confirmed the presence of two populations of transcripts in II-7: normal transcripts and aberrant transcripts consisting of an insertion of 152 bp at the junction of exons 40 and 41 (Fig.2A). This out-of-frame inserted sequence leads to a premature stop codon in exon 41 predicting, if translated, a truncated protein deleted of 2,671 amino acids. This protein would not be anchored to the membrane due to the absence of the C-terminal transmembrane domain required for the normal function of usherin.

To determine the origin of this extra sequence a Blast analysis (Basic Local Alignment Search Tool; http://blast.ncbi.nlm.nih.gov/Blast.cgi) was performed and revealed a perfect alignment with an intronic sequence located at -2,296 bp from exon 41 (Fig.2B). In consequence the nomenclature of this insertion is r.7594_7595ins7595-2296_7595-2143.

In silico analysis using the Human Splicing Finder (HSF) tool (http://www.umd.be/HSF/) (Desmet, et al., 2009) of the genomic region encompassing the intronic 152 bp detected in transcripts showed that the inserted sequence was surrounded by 3' and 5' splice sites (iss for inserted sequence splice site). When compared to the Maximum Entropy (MaxEnt (Yeo and Burge, 2004)) scores average for 3' and 5'constitutive *USH2A* splice sites (3'Ushss and 5'Ushss) the 3'iss can be considered a strong site (MaxEnt scores: 8.16±2.28 for 3'Ushss and 10.11 for the 3'iss) in contrast to the 5'iss that is probably a very weak site due to the presence of an A at position +1 (MaxEnt scores: 8.76±1.58 for 5'Ushss and 2.39 for the 5'iss).

In accordance with this result, and the fact that the inserted sequence was absent in 12 controls, we hypothesized that the aberrant splicing resulted from a pseudoexon (PE) activation due to a mutational event reinforcing the 5'iss.

To confirm this hypothesis, DNA studies were conducted in patient II-7 and the 12 controls previously analyzed at RNA level. Primers, Pe40F (5'-CAATTACACCCACGGAGAGC-3') and Pe40R (5'-TCCCAGAGAAAAGCAAGTACG-3'), were designed to amplify the potential PE and its surrounding genomic sequences. PCR reactions were performed under standard conditions (annealing temperature 60°C; 30 cycles).

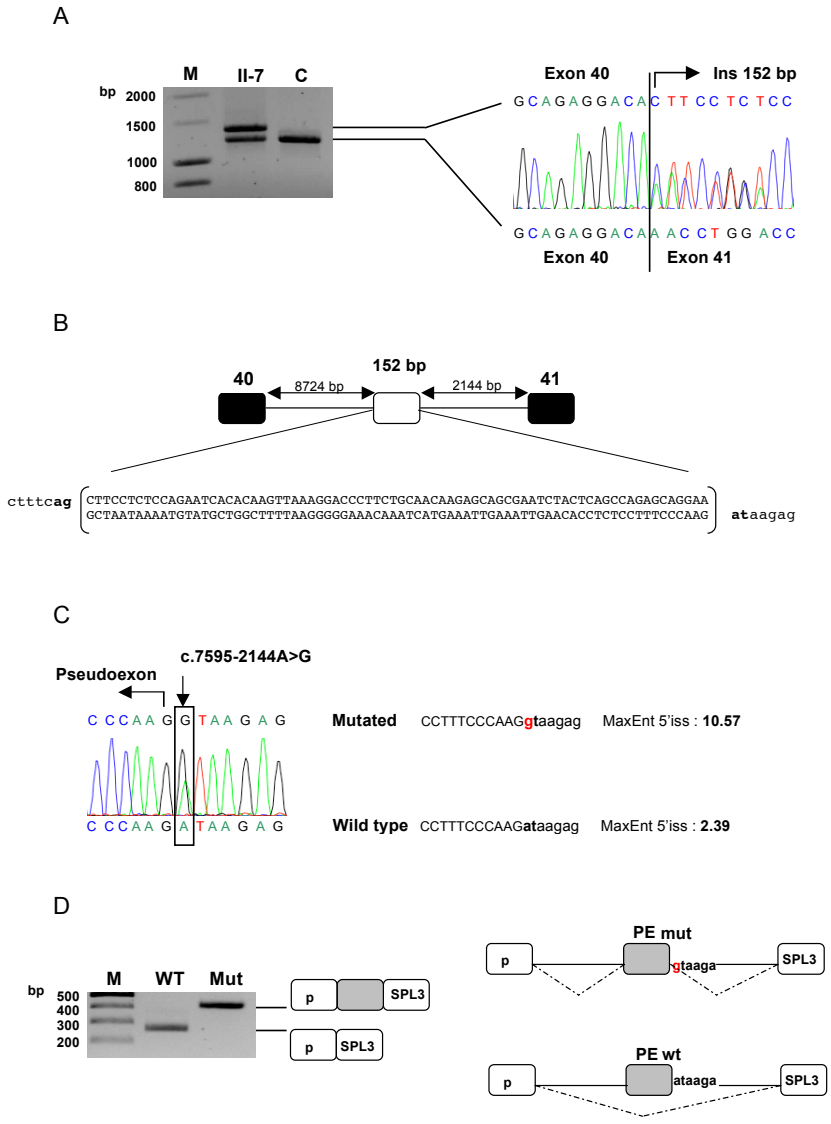


Figure 2. Molecular characterization of the pseudoexon (PE) inclusion in patient II-7. **A.**Total RNA was isolated from nasal cells of patient II-7 and controls with the Nucleo Spin RNA II isolation kit (Macherey-Nagel) and reverse transcribed using Superscript™ III reverse transcriptase (Invitrogen, Carlsbad, CA, USA)). Nested PCRs were carried out using primers localized in exons 39 and 44 in first rounds and 40 and 43 in second rounds (expected size of 1,208 bp) including the heterozygous c.8624G>A SNP located in exon 43 (not shown). Agarose gel is shown for one representative RT-PCR from patient II-7 and a control (C). An additional band of 1,360 bp is observed in II-7. Sequencing of the two products obtained for II-7, using the BigDye Terminator v1.1 cycle

sequencing kit of products, is presented for the exonic junctions. **B**.Schematic localization and sequence of the PE at DNA level. Exons are represented by black boxes, introns by horizontal lines and PE by a white box. Sequence in uppercase corresponds to PE and lowercase to the intronic flanking sequences. **C.** Identification of the c.7595-2144A>G mutation in the genomic sequence of patient II-7. The A>G substitution at position +1 of the PE is boxed and its effect on the donor splice site strength is indicated (MaxEnt score). **D.** *Ex vivo* splicing assay. DNA of patient II-7 was used as template to generate a fragment containing the PE encompassed by 208 bp of the 5' and 197 bp of 3' intronic flanking sequences using the Phusion™ High-Fidelity DNA polymerase (Finnzymes, Espoo, Finland). Wild type (WT) and mutated (Mut) pSPL3 minigenes were constructed and transfected into ARPE-19 and HeLa cells as previously described (Le Guédard-Mereuze, et al., 2010); 48 hr after transfection total RNA was extracted and RT-PCR were performed. Primers sequences and RT-PCR conditions are available upon request. As no difference in splicing effect was observed between the two cell lines only results from ARPE-19 cells are shown. Schemes of the amplified products are presented on the right of the agarose gel. White boxes represent pSPL3 exons and grey box the 152 bp PE.

In the patient's DNA, direct sequencing of PCR products (Fig.2C) revealed, a heterozygous base-change A>G located at position +1 of the inserted sequence (c.7595-2144A>G) increasing the 5'iss strength from 2.39 to 10.57 (MaxEnt scores). This increase (+342.26%) is the result of the creation of a high-quality donor splice site (MaxEnt scores: 8.76±1.58 for 5'Ushss). This variation was not observed in the 12 controls.

The c.7595-2144A>G variant was analyzed in the other U249 family members and segregated perfectly with the phenotype: subjects I-2, II-2 and II-4 were identified as carriers whereas III-1 and III-4 were compound heterozygotes c.[7595-2144A>G];[8890dupT].

The variant was not described in dbSNP and not detected in 338 tested alleles. All these results therefore strongly suggest that this variant was the second necessary and sufficient mutational event responsible for the Usher syndrome in U249.

To definitively confirm this hypothesis, we carried out a functional analysis for this pseudoexon (Fig.2D) on HeLa and human retinal pigment epithelial cells (ARPE-19) using the exon-trapping vector pSPL3 (Le Guédard-Mereuze, et al., 2010). Only transcripts obtained from the mutated c.7595-2144A>G construction included the PE.

These results, combined with those obtained on RNA from nasal epithelial cells, show for the first time the involvement of a deep intronic mutation in USH type 2.

To establish the prevalence of this mutation, direct sequencing of DNA was performed in two groups of our cohort. The first group was composed of 20 *USH2A* patients with no identified mutation or only a single pathogenic mutation or likely/probably pathogenic variant (UV3/UV4 according to our classification criteria (Roux, et al., 2011)). The second group contained 54 *USH2A* patients carrying a heterozygous mutation plus a variant described as UV3/UV4 or two variants classified as UV3 or UV4, tat is, both likely mutations had already been identified. Analyses of the 74 amplified sequences showed that the c.7595-2144A>G mutation was present in four patients of group one at heterozygous state (Supp. Table S2) and was not detected in group two.

An additional DNA study carried out in a Spanish cohort composed of 18 patients with incomplete USH2A genotypes identified the c.7595-2144A>G mutation in four cases

(Supp. Table S2), while an epidemiological study on 180 Spanish alleles was not able to detect it.

For these two European cohorts, when possible, we showed that this base change A>G always segregated with the disease /haplotypes in each family.

Comparison of all the haplotypes (built with intragenic neutral variants) carrying the intronic mutation concluded that c.7595-2144A>G was not linked to a specific haplotype. Nevertheless, four patients (two French and two Spanish, Supp. Table S2) carry the c.7595-2144A>G mutation and for three of them it is *in cis* to the rare variant c.6049+76A>T (rs55865063) (Aller, et al., 2006) that we have classified as nonpathogenic by *ex vivo* splicing assays (unpublished results). This association suggests a possible ancestral event.

Although mid-intronic mutations leading to PE activation have been already described in several genes (Dhir and Buratti, 2010) this mechanism is, as yet, an infrequently observed phenomenon mainly because it is indeed a rare event as splicing needs a set of determinants (donor and acceptor splice sites, polypyrimidine tract and lariat branch point) all in the correct position and, second, because their detection escapes both classic mutational analyses and next-generation exome sequencing. This study shows for the first time this type of pathogenic mechanism in USH2 syndrome and highlights the importance of analyses at both RNA and DNA levels.

In view of the frequency and the presence of the c.7595-2144A>G mutation in both a French and a Spanish population, we suggest that the DNA analysis of the region encompassing the PE should be integrated into *USH2A* molecular diagnostic services. Furthermore, similar studies for the eight USH genes expressed in nasal epithelial cells (Vaché, et al., 2010) could be carried out when a single mutation has been detected in an USH patient.

There has been considerable interest as to whether there might be digenic inheritance in USH in view of the large number of genes involved and the large number of polymorphisms in each of them. Indeed, *PDZD7* has been described as a contributor to digenism in USH2 (Ebermann, et al., 2010). We have found no evidence of digenic inheritance in either USH1 (Roux, et al., 2011) or in this study of USH2. If identification of only a single mutation is sometimes regarded as possible evidence for digenic inheritance, this study comprehensively shows that more exhaustive analysis, including of RNA, should be undertaken to identify a likely second mutation.

This PE identification in nine unrelated *USH2A* patients not only greatly improves molecular diagnosis, but opens a new therapeutic approach for USH. In this case, where all the native exons are present, induction of PE exclusion would result in restoring a complete wild type RNA with a functional protein and a corrected phenotype. In other diseases this therapeutic approach has already been developed using antisense oligonucleotide (AON) chemistry to rescue a normal splicing process and promising results have been reported (Perez, et al., 2010). Considering the advances in improving the stability and the safety of AON by the use of modified molecules and the progress in nanotechnologies and controlled drug delivery systems (Fattal and Barratt, 2009; Perez, et al., 2010) patients

carrying the c.7595-2144A>G mutation could, in the not too distant future, benefit from antisense-mediated exon skipping therapy. To our knowledge, this approach had not been yet considered for USH.

Supp.Table S1. Primers, product sizes and amplification conditions used for the 14 overlapping RT- PCR reactions. * already published by (van Wijk, et al., 2004)

N° PCR	Forward Primer (5'-3')	Reverse Primer (5'-3')	Expected product size (bp)	Annealing temperature (C°)
1	GCTCTGAGCTTCAGGTAACC	GATCTCCAGAGAAGACTTCC		58
	GTCTTCCTTCCAGCATTCC	GTGCCACTTGGTATAATCG	1189	56
2	CAATGGCGTGGAGAAGGATC	CCCAGTCTTATCACAGTTGC		58
	CTCTAAGTGGTTCAATTACAG	CAGGTGTCACACTGAAGTCC	1618	58
3	CCTTGTCAATGCAACAGCC	CAAACACACTGACCAGTCAGG		58
	CATTACAACATCTCTGTAGACC	GATTAACTGCACCAGTTGTATGG	1129	62
4	GAAGGGAGACAGTGCAATAAATG	CCTTCGTAGGAAACACATGG		60
	GTGTGTGCCTAATCGTCAAGG	GTGTCACAGAGTCTGAGCC	737	60
5	GCTTACTTACAGTTTACTCAGGG	GAGGATGGTATAACTTCGCG		60
	CACAACAGAGGATCAATACCC	CAGTACTACCATTCAGGATGG	1684	60
6	CCAAATCACTCTGGATGGG	CCAAGGAGCAAATCCGTAAGC		60
	GCCATCCTGAATGGTAGTAC	CACGATAGCTGAGTTCTGAGG	1984	58
7	CCTGGGATGAACCTGTTGTC*	CGATGGCTGTGTGGTTAAGG		60
	CTCTGCCAGTGCTGAATTTG*	TTCTCGGCTCGGTGTAAAAC*	2737	58
8	CAGCTCGTAATAACGCTCC	TTCTCGGCTCGGTGTAAAAC		58
	TTTCCAATCCTTCTGCATCG	TTCATAAAGCCACTGAGTTCC	1208	58
9	GAGCTGTTCTCTGATACTCC	GAATACTCAGTGACAACACC		58
	GTAGAAGTTACCACAAGACC	GTACTGTTGATGATGACAACC	959	58
10	GTAAACTCTCATGTCATTGGC	GTGCTGTTCATCTGGACAGC		60
	GTCCACAGCATCAACAGTGC	CTTCACAACAGCGATGTCCAG	619	62
11	GATCAGAGTGATGAGCTCTGC	AGCTGTTCCAGGCAGAAATC		60
	CCTGAATCTATCTGTGGAC	CACAGTGCAAATGTGGCTGG	756	56
12	GAGCCAGAAATGCTGTAATGG	GATGATGACCAACGGAGAAGG		62
	GCTCTGACAAGATTTCAACTGG	CATTGAGTAAGACATTGTACTCC	1290	64
13	CAAACACCTATGTCAACACC	CATCAAAGGTGCAATCTCAGG		58
	CACCAGAAGAAATCTATCCTCC	CACTTGATGCACATCCCAGG	2550	62
14	CCACCAGCATCACAACTCTG*	CCTCTATTAGGAAGGAACAC		58
	GTCAGCCCTCCAGATCTTTG*	CCTTGATGGCGTTCATCAGG	2482	58

Supp. Table S2. Overview of the *USH2A* patients carrying the c.7595-2144A>G mutation. The c.6049+76A>T variant (rs55865063) has not been identified among 174 control chromosomes. [a] in cis to c.7595-2144A>G

Origin	Family / Patients		2nd mutation	Originally referenced in	c.6049+76A>T (Aller, et al., 2006)	Segregation
France	U249	II-7	c.3129dupT p.(Val1044fs)	This study	no	in trans
		III-1, III-4	c.8890dupT p.(Trp2964fs)	This study		
	U414	II-3, II-4, II-5	c.3680delG p.(Gly1227fs)	(Baux, et al., 2007)	yes[a]	in trans
	U446	II-1	c.2168-1G>C	(Baux, et al., 2007)	no	in trans
	U753	II-3	c.920_923dup p.(His308fs)	(Dreyer, et al., 2000)	yes[a]	in trans
	U835	II-1	c.2299delG p.(Glu767fs)	(Liu, et al., 1999)	no	not performed
Spain	FRP389	II-1, II-2	c.10636G>A p.(Gly3546Arg) = UV3	This study	yes[a]	in trans
	FRP380	II-1	c.2299delG p.(Glu767fs)	(Liu, et al., 1999)	no	not performed
	FRP30	II-1, II-2	c.11146C>T p.(Gln3716X)	This Study	no	in trans
	FRP484	II-1	c.2276G>T p.(Cys759Phe)	(Dreyer, et al., 2000)	yes	not performed

Acknowledgments

We acknowledge our colleague Dr. Sylvie Tuffery-Giraud for fruitful adivce. Carsten Bergmann and Dr. Christine Neuhaus for referring patients. We are also very grateful to family U249 for their interest in molecular diagnosis and their participation to this study.

REFERENCES

Aller E, Jaijo T, Beneyto M, Najera C, Oltra S, Ayuso C, Baiget M, Carballo M, Antinolo G, Valverde D, Moreno F, Vilela C, Collado D, Perez-Garrigues H, Navea A, Millan JM. 2006. Identification of 14 novel mutations in the long isoform of USH2A in Spanish patients with Usher syndrome type II. J Med Genet 43:e55.

Aller E, Larrieu L, Jaijo T, Baux D, Espinos C, Gonzalez-Candelas F, Najera C, Palau F, Claustres M, Roux AF, Millan JM. 2010. The USH2A c.2299delG mutation: dating its common origin in a Southern European population. Eur J Hum Genet 18:788-793.

Baux D, Larrieu L, Blanchet C, Hamel C, Ben Salah S, Vielle A, Gilbert-Dussardier B, Holder M, Calvas P, Philip N, Edery P, Bonneau D, Claustres M, Malcolm S, Roux AF. 2007. Molecular and in silico analyses of the full-length isoform of usherin identify new pathogenic alleles in Usher type II patients. Hum Mutat 28:781-789.

Desmet FO, Hamroun D, Lalande M, Collod-Beroud G, Claustres M, Beroud C. 2009. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. Nucleic Acids Res 37:e67.

Dhir A, Buratti E. 2010. Alternative splicing: role of pseudoexons in human disease and potential therapeutic strategies. The FEBS journal 277:841-55.

Dreyer B, Tranebjaerg L, Brox V, Rosenberg T, Moller C, Beneyto M, Weston MD, Kimberling WJ, Cremers CW, Liu XZ, Nilssen O. 2001. A common ancestral origin of the frequent and widespread 2299delG USH2A mutation. Am J Hum Genet 69:228-34.

Dreyer B, Tranebjaerg L, Rosenberg T, Weston MD, Kimberling WJ, Nilssen O. 2000. Identification of novel USH2A mutations: implications for the structure of USH2A protein. Eur J Hum Genet 8:500-6.

Ebermann I, Phillips JB, Liebau MC, Koenekoop RK, Schermer B, Lopez I, Schafer E, Roux AF, Dafinger C, Bernd A, Zrenner E, Claustres M, Blanco B, Nurnberg G, Nurnberg P, Ruland R, Westerfield M, Benzing T, Bolz HJ. 2010. PDZD7 is a modifier of retinal disease and a contributor to digenic Usher syndrome. J Clin Invest 120:1812-23.

Fattal E, Barratt G. 2009. Nanotechnologies and controlled release systems for the delivery of antisense oligonucleotides and small interfering RNA. Br J Pharmacol 157:179-94.

Horaitis O, Cotton RG. 2004. The challenge of documenting mutation across the genome: the human genome variation society approach. Hum Mutat 23:447-52.

Kimberling WJ, Hildebrand MS, Shearer AE, Jensen ML, Halder JA, Trzupek K, Cohn ES, Weleber RG, Stone EM, Smith RJ. 2010. Frequency of Usher syndrome in two

pediatric populations: Implications for genetic screening of deaf and hard of hearing children. Genet Med 12:512-6.

Le Guédard-Mereuze S, Vaché C, Baux D, Faugère V, Larrieu L, Abadie C, Janecke A, Claustres M, Roux AF, Tuffery-Giraud S. 2010. Ex vivo splicing assays of mutations at non-canonical positions of splice sites in USHER genes. Hum Mutat 31:347-355.

Liu XZ, Hope C, Liang CY, Zou JM, Xu LR, Cole T, Mueller RF, Bundey S, Nance W, Steel KP, Brown SD. 1999. A mutation (2314delG) in the Usher syndrome type IIA gene: high prevalence and phenotypic variation. Am J Hum Genet 64:1221-5.

Perez B, Rodriguez-Pascau L, Vilageliu L, Grinberg D, Ugarte M, Desviat LR. 2010. Present and future of antisense therapy for splicing modulation in inherited metabolic disease. J Inherit Metab Dis 33:397-403.

Roux AF, Faugere V, Vaché C, Baux D, Besnard T, Leonard S, Blanchet C, Hamel C, Mondain M, Gilbert-Dussardier B, Edery P, Lacombe D, Bonneau D, Holder-Espinasse M, Ambrosetti U, Journel H, David A, Lina-Granade G, Malcolm S, Claustres M. 2011. Four year follow-up of diagnostic service in USH1 patients. Invest Ophthalmol Vis Sci 52:4063-4071.

Smith RJ, Berlin CI, Hejtmancik JF, Keats BJ, Kimberling WJ, Lewis RA, Moller CG, Pelias MZ, Tranebjaerg L. 1994. Clinical diagnosis of the Usher syndromes. Usher Syndrome Consortium. Am J Med Genet 50:32-8.

Vaché C, Besnard T, Blanchet C, Baux D, Larrieu L, Faugère V, Mondain M, Hamel C, Malcolm S, Claustres M, Roux AF. 2010. Nasal epithelial cells are a reliable source to study splicing variants in Usher syndrome. Hum Mutat 31:734-41.

van Wijk E, Pennings RJ, te Brinke H, Claassen A, Yntema HG, Hoefsloot LH, Cremers FP, Cremers CW, Kremer H. 2004. Identification of 51 novel exons of the Usher syndrome type 2A (USH2A) gene that encode multiple conserved functional domains and that are mutated in patients with Usher syndrome type II. Am J Hum Genet 74:738-44.

Yeo G, Burge CB. 2004. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. J Comput Biol 11:377-94.

References cited only in the sup. Material

(Baux, et al., 2007) (Dreyer, et al., 2000) (Liu, et al., 1999) (van Wijk, et al., 2004)

Resumen

Introducción

El síndrome de Usher (USH) es una enfermedad hereditaria autosómica recesiva, caracterizada por la asociación de hipoacusia neurosensorial, retinosis pigmentaria y, en ocasiones, alteración de la función vestibular. Clínicamente, el USH se puede clasificar en tres tipos (USH1, USH2 y USH3), principalmente en base a la gravedad y progresión de la hipoacusia y presencia o no de disfunción vestibular. El USH es heterogéneo tanto a nivel clínico como genético y, hasta la fecha, se han descrito 10 genes implicados en la enfermedad. Se conocen seis genes responsables del síndrome de tipo 1: *MYO7A*, *USH1C*, *CDH23*, *PCDH15*, *USH1G* y *CIB2*; tres genes responsables del tipo 2: *USH2A*, *GPR98* y *DFNB31*; y un gen responsable del tipo 3: *CLRN1*. Además, se ha propuesto la implicación de *PDZD7* en un modelo de digenismo junto con *GPR98* o como modificador, agravando el fenotipo de pacientes con mutaciones en *USH2A*.

Los genes USH codifican para proteínas que intervienen en la misma red de interacción proteica, conocida como *interactoma-Usher*. Este interactoma tiene un papel fundamental en las células del oído interno y la retina, siendo la harmonina y la whirlina a través de sus dominios PDZ, las proteínas centrales a las que se unen las demás.

El USH2 es la forma más común y tres son los genes responsables conocidos: *USH2A (72 exones)*, *GPR98 (90 exones)* y *DFNB31 (12 exones)*. *USH2A* es responsable de más del 70% de los casos USH2, teniendo una menor implicación los otros dos genes. Se han realizado rastreos mutacionales de los genes USH2 en otras poblaciones, pero hasta la realización de esta tesis, no se había llevado a cabo un análisis exhaustivo de estos tres genes en pacientes españoles con síndrome de Usher. Este trabajo nos permitirá conocer el espectro mutacional en nuestra cohorte, conocer la implicación de cada uno de los genes y poder compararlo con otras poblaciones. *CLRN1* es el único gen responsable del USH3 conocido. Sin embargo, se ha publicado que este gen también puede estar implicado en pacientes con USH1 o USH2. El estudio de este gen en pacientes USH2 y USH3 nos permitirá identificar nuevas mutaciones.

En la actualidad, los métodos para la determinación del defecto causante de la enfermedad pasan por la secuenciación exón a exón de los genes responsables. Sin embargo, debido a la elevada heterogeneidad genética y alélica del USH y al gran tamaño de la mayoría de los genes implicados, el diagnóstico del USH mediante las técnicas de secuenciación tradicionales resulta una tarea muy laboriosa. En los últimos años, las técnicas de secuenciación de nueva generación se han desarrollado considerablemente y son muchas sus posibles aplicaciones. La secuenciación masiva dirigida mediante el enriquecimiento de regiones genómicas especificas es una estrategia que permite maximizar la eficacia de esta tecnología, facilitando el análisis y manejo de los datos, resultando muy útil en el caso de enfermedades heterogéneas. El desarrollo de una plataforma de NGS es una estrategia apropiada para el estudio de los genes Usher.

Objetivos

1.
   a) Determinar la implicación de cada gen USH2 (*USH2A, GPR98, DFNB31*) en una cohorte de pacientes españoles con síndrome de Usher.

   b) *Screening* mutacional del gen *CLRN1* en una cohorte de 17 pacientes.

2.
   a) Desarrollo e implementación de una plataforma de secuenciación de nueva generación para el diagnóstico genético del USH.

Metodología

Análisis molecular de los genes USH2 y USH3

Para el estudio molecular de los genes USH2 (cohorte inicial de 88 pacientes) y *CLRN1* (17 pacientes) se amplificaron mediante PCR las secuencias exónicas junto con las regiones intrónicas flanqueantes y, posteriormente, se secuenciaron los productos amplificados. Como particularidad, para el estudio de los genes *GPR98* y *DFNB31* se aplicó el método SCAIP (Single Condition Amplification/Internal Primers). Esta metodología permite estudiar todos los exones de un gen en un mismo paciente mediante secuenciación directa en placa en unas mismas condiciones, por lo que los resultados del estudio genético son obtenidos en un corto espacio de tiempo.

El estudio no se limitó a la detección de mutaciones puntuales, las variantes detectadas fueron analizadas con varios programas bioinformáticos para predecir su posible repercusión sobre la estructura y función de la proteína o su posible impacto en el mecanismo de *splicing*. Cuando fue necesario, se procedió a la clonación de este tipo de cambios en minigenes de expresión especialmente diseñados para el análisis del efecto de estos cambios sobre el mecanismo de *splicing* y tratar de confirmar las predicciones informáticas y el carácter patológico de las variantes.

Para identificar posibles reordenamientos genómicos, no detectados por las técnicas tradicionales, en cualquiera de los genes responsables de síndrome de Usher se realizó un "high density oligonucleotide comparative genomic hybridization" (CGH-array).

Estudio del exoma-Usher mediante NGS

El estudio del exoma Usher se realizará combinando la captura de regiones (Nimblegen) con la secuenciación masiva en la plataforma GSJunior (Roche). Se estudiaron las regiones codificantes, secuencias intrónicas flanqueantes y regiones 5' y 3' UTR de los 11 genes descritos implicados hasta la fecha en el síndrome de Usher, más otros genes candidatos o implicados en sordera no sindrómica; un total de 19 genes. El

método elegido para el enriquecimiento de las regiones de interés es la captura en solución liquida (SeqCap EZ Choice Library, Roche NimbleGen). Una serie de sondas biotiniladas de unos 70 nucleótidos de longitud se diseñaron para poder capturar de forma efectiva todas las regiones deseadas.

Resultados

Análisis molecular de los genes USH2 y USH3

El estudio se inició con *USH2A,* y en los pacientes en los que no se encontraron mutaciones en este gen, se continuó con el análisis de *GPR98* y *DFNB31*. Treinta y siete (23 nuevas) mutaciones diferentes se detectaron en *USH2A* y 7 (5 nuevas) en *GPR98*. No se detectaron mutaciones en *DFNB31*. Las mutaciones detectadas fueron de naturaleza muy diversa (cambios de aminoácido, isocodificantes, codones de parada prematuro directos, pequeñas deleciones e inserciones) y en la mayoría de casos las mutaciones fueron privadas, demostrando la elevada heterogeneidad alélica de esta enfermedad. Además, se identificaron dos grandes deleciones en el gen *USH2A* gracias al uso de la técnica de CGH-array. En el desarrollo de este estudio se describió por primera vez en el USH, la presencia de una mutación intrónica profunda en el gen *USH2A* que produce la activación de un pseudoexón. Esta variante fue detectada en 4 pacientes españoles, completando y mejorando de esta forma el diagnóstico molecular del USH. Esta parte del estudio nos permitió estimar la prevalencia de los genes USH2 en población española. *USH2A* está implicado en un 76% de los casos USH2, seguido de *GPR98* con un 5,2%. *DFNB31* tiene una menor implicación en nuestra población.

El análisis del gen *USH3A* en una cohorte de 17 pacientes con USH nos ha permitido identificar 2 nuevas mutaciones y confirmar la baja implicación de este gen en pacientes españoles.

Estudio del exoma-Usher mediante NGS

Nosotros hemos elegido el método de captura de regiones (Nimblegen, Roche) seguido de secuenciación masiva mediante la plataforma GS Junior (454, Roche) para el estudio de los genes Usher y su posible aplicación al diagnóstico. Un total de 19 genes fueron incluidos en el diseño (genes Usher, candidatos y otros implicados en sordera). De media, se obtuvieron cerca de 53Mb de secuencias, con un *on target* del 52%. La profundidad de cobertura obtenida fue de 77X, superando los 40X necesarios. Para evaluar la sensibilidad y fiabilidad de la técnica, en una primera fase se estudiaron 47 pacientes que tenían analizados previamente algunos de los genes Usher. Esta primera cohorte de pacientes nos permitió establecer los criterios adecuados para los sucesivos filtros que se aplicaron a las variantes obtenidas. Esto se automatizó mediante el diseño de un propio programa bioinformático, llamado *GSdot*. Solo un 2% de las variaciones conocidas no fueron detectadas por NGS debido principalmente a la presencia de regiones con homopolímeros, problemas de alineamiento o una baja cobertura. También nos permitió

detectar mutaciones adicionales en 12 de los 47 pacientes, bien porque las mutaciones no se habían detectado por Sanger o el gen implicado no se había estudiado. Una vez comprobada la validez de la técnica, se estudiaron 13 nuevos pacientes con USH pero sin ningún gen analizado anteriormente. Se detectaron las dos mutaciones en 10 de los 13 casos y en otros dos pacientes se detectó una variante clasificada como UV3. Además, se estudiaron 11 pacientes con sordera no sindrómica, lo que permitió identificar 3 mutaciones en 3 de ellos. En varios pacientes solo se identificó una mutación, y aunque alguno de ellos pueda ser simplemente portador de una mutación en los genes Usher y las mutaciones responsables estén en otro gen, la mayoría de ellos deben tener una segunda mutación, posiblemente en regiones no analizadas con esta técnica (por ejemplo: regiones intrónicas). En otros pacientes, no se pudo detectar ninguna mutación. En estos casos, el estudio del exoma entero ayudará a encontrar el gen implicado. Aunque es necesario mejorar la técnica, los resultados obtenidos son muy prometedores para una futura traslación al diagnóstico. Esta estrategia facilitará el diagnóstico del USH de pacientes que no presentan mutaciones en los genes principalmente implicados o de aquellos casos que presentan mutaciones en genes típicamente responsables de otro subtipo clínico.

Conclusiones

1. Mutaciones en el gen *USH2A* son responsables de un 76.1% de los pacientes USH2 en población española. Después de la deleción c.2299delG, p.C3267R es la mutación más frecuente. La mayoría de mutaciones son privadas y no se observan regiones con acumulación de mutaciones.

2. *GPR98* es el segundo gen USH2 más prevalente, con una implicación del 5.2%. No hay mutaciones recurrentes y las mutaciones tipo *missense* son menos frecuentes que en *USH2A*. Las mutaciones se distribuyen a lo largo de todo el gen, con una ligera acumulación en los últimos exones.

3. La implicación de *DFNB31* en población española es muy baja. Hasta la fecha, no se han detectado mutaciones patológicas en este gen en nuestra cohorte.

4. Los minigenes son una buena estrategia para comprobar el efecto de las variantes en el mecanismo de *splicing*. Las predicciones *in silico* y los análisis funcionales son necesarios para determinar la posible patogenicidad de las variantes.

5. Los grandes reordenamientos en los genes USH son responsables de síndrome de Usher. La técnica de CGH-array nos ha permitido identificar dos grandes deleciones y sus puntos de rotura. Este tipo de estudios deberían ser incluidos en el diagnóstico molecular.

6. La mutación intrónica profunda c.7595-2144A>G produce la activación de un *pseudoexon* en el gen *USH2A*. Esta mutación es bastante frecuente, se ha detectado en 9 pacientes con USH2. Estudios adicionales de los transcritos de los genes USH nos

permitirán detectar otras mutaciones de este tipo. Esta mutación ofrece la posibilidad de una terapia mediante el uso de AONs (oligonucleótidos antisentido).

7. USH3 es el tipo de USH menos frecuente en nuestra población. Mutaciones en *CLRN1* se han identificado en un 2.7% de todos nuestros pacientes.

8. La captura del exoma Usher, seguido de la secuenciación masiva, ofrece un diagnóstico molecular más rápido y económico en comparación a la convencional secuenciación gen a gen mediante Sanger.

9. El panel de genes USH facilita el diagnóstico molecular de pacientes con un fenotipo atípico o sin datos clínicos disponibles. Este estudio nos permite identificar mutaciones en genes típicamente implicados en otro subtipo clínico y posibles casos de digenismo.

10. La identificación eficaz de mutaciones demuestra que hemos desarrollado una estrategia experimental y bioinformática apropiada. La secuenciación dirigida es una herramienta adecuada para el estudio de enfermedades heterogéneas.

List of publications

Besnard, T*., **G. García-García*,** D. Baux*, C. Vaché, V. Faugère, L. Larrieu, S. Léonard, J. M. Millan, S. Malcolm, M. Claustres, A. F. Roux. Usher syndrome: experience of targeted exome sequencing as a clinical test. Molecular Genetics & Genomic Medicine Submitted. 2013.

Garcia-Garcia, G., T. Besnard, D. Baux, C. Vache, E. Aller, S. Malcolm, M. Claustres, J. M. Millan and A. F. Roux. The contribution of GPR98 and DFNB31 genes to a Spanish Usher syndrome type 2 cohort. Mol Vis. 2013 19**:** 367-73.

Aparisi, M. J., **G. Garcia-Garcia**, E. Aller, M. D. Sequedo, C. Martinez-Fernandez de la Camara, R. Rodrigo, M. Armengot, J. Cortijo, J. Milara, M. Diaz-Llopis, T. Jaijo and J. M. Millan. Study of USH1 Splicing Variants through Minigenes and Transcript Analysis from Nasal Epithelial Cells. PLoS One. 2013 8**:** e57506.

Garcia-Garcia, G., M. J. Aparisi, R. Rodrigo, M. D. Sequedo, C. Espinos, J. Rosell, J. L. Olea, M. P. Mendivil, M. A. Ramos-Arroyo, C. Ayuso, T. Jaijo, E. Aller and J. M. Millan. Two novel disease-causing mutations in the CLRN1 gene in patients with Usher syndrome type 3. Mol Vis. 2012 18**:** 3070-8.

Vache, C., T. Besnard, P. le Berre, **G. Garcia-Garcia**, D. Baux, L. Larrieu, C. Abadie, C. Blanchet, H. J. Bolz, J. Millan, C. Hamel, S. Malcolm, M. Claustres and A. F. Roux. Usher syndrome type 2 caused by activation of an USH2A pseudoexon: implications for diagnosis and therapy. Hum Mutat. 2012 Jan; 33**:** 104-8.

Garcia-Garcia, G., M. J. Aparisi, T. Jaijo, R. Rodrigo, A. M. Leon, A. Avila-Fernandez, F. Blanco-Kelly, S. Bernal, R. Navarro, M. Diaz-Llopis, M. Baiget, C. Ayuso, J. M. Millan and E. Aller. Mutational screening of the USH2A gene in Spanish USH patients reveals 23 novel pathogenic mutations. Orphanet J Rare Dis. 2011 6**:** 65.

Jaijo, T., E. Aller, M. J. Aparisi, **G. Garcia-Garcia**, I. Hernan, M. J. Gamundi, C. Najera, M. Carballo and J. M. Millan. Functional analysis of splicing mutations in MYO7A and USH2A genes. Clin Genet. 2011 Mar; 79**:** 282-8.

Aparisi, M. J., **G. Garcia-Garcia**, T. Jaijo, R. Rodrigo, C. Graziano, M. Seri, T. Simsek, E. Simsek, S. Bernal, M. Baiget, H. Perez-Garrigues, E. Aller and J. M. Millan. Novel mutations in the USH1C gene in Usher syndrome patients. Mol Vis. 2010 16**:** 2948-54.

Aller, E., T. Jaijo, **G. Garcia-Garcia**, M. J. Aparisi, D. Blesa, M. Diaz-Llopis, C. Ayuso and J. M. Millan. Identification of large rearrangements of the PCDH15 gene by combined MLPA and a CGH: large duplications are responsible for Usher syndrome. Invest Ophthalmol Vis Sci. 2010b Nov; 51**:** 5480-5.

Jaijo, T., E. Aller, **G. Garcia-Garcia**, M. J. Aparisi, S. Bernal, A. Avila-Fernandez, I. Barragan, M. Baiget, C. Ayuso, G. Antinolo, M. Diaz-Llopis, M. Kulm, M. Beneyto, C. Najera and J. M. Millan. Microarray-based mutation analysis of 183 Spanish families with Usher syndrome. Invest Ophthalmol Vis Sci. 2010 Mar; 51**:** 1311-7.

Aller, E., T. Jaijo, E. van Wijk, I. Ebermann, F. Kersten, **G. Garcia-Garcia**, K. Voesenek, M. J. Aparisi, L. Hoefsloot, C. Cremers, M. Diaz-Llopis, R. Pennings, H. J. Bolz, H. Kremer and J. M. Millan. Sequence variants of the DFNB31 gene among Usher syndrome patients of diverse origin. Mol Vis. 2010a 16**:** 495-500.