



## TESIS DE DOCTORADO

# **NeoSeq: una herramienta de diagnóstico genético rápido para recién nacidos críticamente enfermos con sospecha de enfermedad genética**

María José de Castro López

ESCOLA DE DOUTORAMENTO

PROGRAMA DE DOUTORAMENTO EN INVESTIGACIÓN CLÍNICA EN MEDICINA

SANTIAGO DE COMPOSTELA

2021





## DECLARACIÓN DEL AUTOR/A DE LA TESIS

D./Dña. **María José de Castro López**

Título de la tesis: **NeoSeq: una herramienta de diagnóstico genético rápido para recién nacidos críticamente enfermos con sospecha de enfermedad genética**

Presento mi tesis, siguiendo el procedimiento adecuado al Reglamento y declaro que:

- 1) La tesis abarca los resultados de la elaboración de mi trabajo.
- 2) De ser el caso, en la tesis se hace referencia a las colaboraciones que tuvo este trabajo.
- 3) Confirmo que la tesis no incurre en ningún tipo de plagio de otros autores ni de trabajos presentados por mí para la obtención de otros títulos.

Y me comprometo a presentar el Compromiso Documental de Supervisión en el caso que el original no esté depositado en la Escuela.

En **Santiago de Compostela, 6 de junio de 2021.**

*Fdo. María José de Castro López*





## AUTORIZACIÓN DEL DIRECTOR DE LA TESIS

**NeoSeq: una herramienta de diagnóstico genético rápido pararecién nacidos críticamente enfermos con sospecha de enfermedad genética**

D<sup>a</sup>. María Luz Couce Pico

INFORMA:

Que la presente tesis, se corresponde con el trabajo realizado por D<sup>a</sup>. María José de Castro López, bajo mi dirección/tutorización, y autorizo su presentación, considerando que reúne los requisitos exigidos en el Reglamento de Estudios de Doctorado de la USC, y que como director de esta no incurre en las causas de abstención establecidas en la Ley 40/2015.

De acuerdo con lo indicado en el Reglamento de Estudios de Doctorado, declara también que la presente tesis doctoral es idónea para ser defendida en base a la modalidad de Monográfica con reproducción de publicaciones, en los que la participación del doctorando/a fue decisiva para su elaboración y las publicaciones se ajustan al Plan de Investigación.

En Santiago de Compostela, 24 de mayo de 2021

Fdo. María Luz Couce Pico





## AUTORIZACIÓN DEL DIRECTOR/TUTOR DE LA TESIS

**NeoSeq: una herramienta de diagnóstico genético rápido pararecién nacidos críticamente enfermos con sospecha de enfermedad genética**

D<sup>a</sup>. María Rosaura Leis Trabazo

INFORMA:

Que la presente tesis, se corresponde con el trabajo realizado por D<sup>a</sup>. María José de Castro López, bajo mi dirección/tutorización, y autorizo su presentación, considerando que reúne los requisitos exigidos en el Reglamento de Estudios de Doctorado de la USC, y que como director de esta no incurre en las causas de abstención establecidas en la Ley 40/2015.

De acuerdo con lo indicado en el Reglamento de Estudios de Doctorado, declara también que la presente tesis doctoral es idónea para ser defendida en base a la modalidad de Monográfica con reproducción de publicaciones, en los que la participación del doctorando/a fue decisiva para su elaboración y las publicaciones se ajustan al Plan de Investigación.

En Santiago de Compostela, 24 de mayo de 2021

Fdo. María Rosaura Leis Trabazo





## DECLARACIÓN DE CONFLICTO DE INTERESES

**NeoSeq: una herramienta de diagnóstico genético rápido para recién nacidos críticamente enfermos con sospecha de enfermedad genética**

Dña. **María José de Castro López,**

Declaro no tener ningún conflicto de intereses  
en relación con la tesis doctoral.

**En Santiago de Compostela, 6 de junio de 2021**



## **AGRADECIMIENTOS**

A todas las personas buenas,  
se hayan cruzado o no en mi camino.

A mi tío José, fue bueno y no tenía miedo.  
Hasta que nos volvamos a abrazar.





*To be nobody but yourself in a world which is doing its best, night and day, to make you everybody else - means to fight the hardest battle which any human being can fight; and never stop fighting.*

e.e. Cummings

*Si cumples tu cometido en el presente según la recta razón, con cuidado, persistencia y serenidad, sin distraerte con nada extraño; si conservas en su pureza al dios que te anima, como si tuvieras que restituirlo en cualquier instante, si, apegado a estos mismos principios, no te ves atormentado por el deseo o por el temor; si te limitas a actuar conforme a la naturaleza de tu ser y a decir sencillamente la verdad en todos tus discursos y en todas tus palabras, vivirás feliz. Y nadie puede impedirte que te conduzcas de este modo.*

Marco Aurelio

*One night I was already asleep and Richard (Feynman) got permission to wake me up, and I can still remember him at my bed saying: "Joan, I want to show you something wonderful". And I got out of bed, and we walked down the street to a golf course. And he said: "Look up". And there, in the sky, was an aurora. And he told me that it was an aurora and that nobody knew how they were formed. And I've spent my scientific career studying that. And it was nice of him to give me the aurora, and to know that I would think it was wonderful.*

Joan Feynman

- 1. Nosotros queremos cantar el amor al peligro, el hábito de la energía y de la temeridad.*
- 2. El coraje, la audacia y la rebeldía serán elementos esenciales de nuestra poesía.*
- 3. Nuestra pintura y arte resalta el movimiento agresivo, el insomnio febril, la carrera, el salto mortal, la bofetada y el puñetazo.*
- 4. Afirmamos que el esplendor del mundo se ha enriquecido con una belleza nueva: la belleza de la velocidad.*
- 5. Queremos alabar al hombre que tiene el volante, cuya lanza ideal atraviesa la Tierra, lanzada ella misma por el circuito de su órbita.*

Filippo Tommaso Marinetti



## RESUMO

Os recém nados con trastornos xenéticos e malformacións conxénitas non só comprenden unha proporción sustancial das admisións nas unidades de coidados intensivos neonatais, senón que contribúen de forma significativa á morbilidade nos primeiros meses de vida, estimándose responsables de arredor dun 30-50% das mortes neonatais e infantís. O diagnóstico destas patoloxías é un dos desafíos máis complexos aos que se enfrontan os especialistas en neonatoloxía, dado que se caracterizan por unha ampla diversidade de síntomas e signos pouco específicos que varían non só dunha enfermidade a outra, senón tamén dun paciente a outro inda padecendo a mesma enfermidade. Estas dificultades inevitablemente conducen ao atraso no diagnóstico e tratamento dos neonatos críticamente enfermos, poboación especialmente vulnerable pola súa fragilidade.

As novas técnicas de secuenciación xenómica móstranse moi prometedoras no campo da neonatoloxía, aumentando a taxa de diagnóstico e reducindo o tempo ata ó mesmo. O avance destas tecnoloxías permite a análise simultánea de centos a miles de xenes en moi pouco tempo e a baixo custo, ofrecendo varias estratexias para favorecer o diagnóstico dos neonatos con sospeita de enfermidade xenética, entre as que se inclúen os paneis xénicos e secuenciación de exoma clínico (CES), exoma completo (WES) e xenoma completo (WGS). Porén, varios obstáculos dificultan a incorporación desta tecnoloxía á práctica clínica habitual, incluíndo a necesidade de recursos humanos, técnicos e económicos moi específicos con distinto nivel de complexidade.

Neste proxecto deseñouse un panel *ad hoc* chamado NeoSeq para recén nados críticamente enfermos con sospeita de padecer unha enfermidade xenética, constituído por 1870 xenes implicados en malformacións conxénitas así como trastornos neurolóxicos e metabólicos de inicio precoz. O obxectivo principal deste estudio foi avaliar prospectivamente a utilidade do panel NeoSeq no diagnóstico molecular de neonatos críticamente enfermos con sospeita de enfermidade xenética, sendo os obxectivos secundarios comparar o rendemento diagnóstico con outras ferramentas de secuenciación masiva, aumentar o coñecemento das bases moleculares da enfermidade nesta poboación de pacientes e avaliar a utilidade da escala Phenomizer a través do fenotipo descrito mediante termos HPO. O fin último deste estudio é poder implementar o panel NeoSeq no diagnóstico das enfermidades xenéticas en neonatos críticamente enfermos ingresados en UCIN.

Para levalo a cabo aplicouse un enfoque en trío incluíndo tanto o paciente como a ambos proxenitores. Este panel solicitóuselle a pacientes críticamente enfermos de menos de dous meses de idade ingresados en unidades de coidados intensivos de tres hospitais terciarios de España e que cumprían os seguintes criterios de inclusión e exclusión: (i) Malformacións conxénitas non relacionadas cun síndrome xenético clinicamente identificable, (ii) descompensación metabólica non asociada con parámetros bioquímicos e / ou cribado neonatal que indicase una sospeita de enfermidade metabólica hereditaria, (iii) epilepsia ou enfermidades do neurodesenvolvemento de probable orixe xenética. Excluíronse do estudio ós neonatos con indicios claros dun síndrome específico que puidieran diagnosticarse mediante análise dirixida de xenes coñecidos ou variantes estruturais. Tódolos probandos foron fenotipados empregando termos HPO extraídos manualmente dos rexistros de saúde electrónicos. Posteriormente xerouse unha lista de posibles diagnósticos diferenciais baseada nos

achados clínicos empregando a aplicación Phenomizer ([compbio.charite.de/phenomizer/](http://compbio.charite.de/phenomizer/)).

Finalmente, reclutáronse 33 dun total de 51 neonatos, nos que se estableceu un diagnóstico xenético para 12 (36.4%) doentes cun tempo medio de resposta de 7.5 días. Ademais, realizouse unha análise dos estudos prospectivos previos de secuenciación masiva con diferentes abordaxes diagnósticas: secuenciación de xenoma completo, exoma completo e exoma clínico, en recém nados ingresados en coidados intensivos con sospeita de trastornos xenéticos, o que arroxou unha taxa diagnóstica media comparable (37.2%) pero cunha mediana de tempo diagnóstica maior (30.2 días) que a obtida con NeoSeq. A taxa diagnóstica do panel está fortemente asociado coa especificidade dos achados fenotípicos. Así, o panel alcanzou una maior rendemento naqueles pacientes con sospeita de enfermidade metabólica y enfermidade neurolóxica (43% e 35%, respectivamente) que o obtido no grupo de anomalías conxénitas, cun 23%. Ó avaliar o impacto do diagnóstico xenético na evolución clínica, achouse que o establecemento dun diagnóstico molecular afectou directamente ó manexo médico en 11 dos 12 pacientes, resultando en cambios na medicación ou dieta (9 pacientes), inicio de novas valoracións por subespecialistas específicos que non se consideraran previamente (5 pacientes) ou retirada do tratamento de coidados intensivos con reorientación cara coidados paliativos (5 pacientes). A análise dos termos HPO mediante a ferramenta Phenomizer correlacionou o fenotipo co diagnóstico molecular obtido de forma estadísticamente significativa ( $p < 0.05$ ) nun 41.6% dos pacientes.

O noso estudio demostra que o panel de xenes NeoSeq dirixido por fenotipos e específico para enfermidades xenéticas en neonatos críticamente enfermos é unha alternativa asequible a secuenciación de exoma clínico, exoma completo e xenoma completo ofrecendo unha

eficacia diagnóstica comparable cun tempo de resposta máis rápido, o que apoiaría a súa aplicación como proba de primeiro nivel nas unidades de coidados intensivos neonatais. Ademais, o impacto do diagnóstico xenético alcanzado con NeoSeq demostra a utilidade clínica desta ferramenta en neonatos críticamente enfermos. Por outra parte, a abordaxe en trios permite obter o diagnóstico definitivo de maneira máis rápida, ó facer innecesarios os posteriores estudos de segregación das variantes identificadas nos pacientes índice.

En resumo, as tecnoloxías de secuenciación masiva e especificamente o panel NeoSeq teñen un enorme potencial como test de primeira liña para o diagnóstico de neonatos críticamente enfermos con sospeita de padecer unha enfermidade xenética. Mentres que as taxas de diagnóstico máis altas deberían agardarse da secuenciación de WGS que de WES, e de WES que de CES ou paneis específicos, encontramos que ésta non se correlaciona co tamaño da porción do genoma analizado xa que NeoSeq proporcionou unha taxa diagnóstica similar ás obtidas cos enfoques WGS y WES, a pesar de interrogar 10 veces (WES) e 500 veces (WGS) menos pares de bases. A implementación clínica destas probas permite ademais un asesoramento xenético adecuado para o paciente e familia, así como adecuar a actitude terapéutica e o seguimento avanzando na liña da medicina personalizada aplicada a unha poboación de doentes especialmente vulnerable.

**Palabras chave:** Tecnoloxías de Secuenciación Masiva, Paneis Xenéticos, Enfermidades Raras, Neonatos, Unidades de Coidados Intensivos. Diagnóstico Xenético.

## RESUMEN

Los recién nacidos con trastornos genéticos y malformaciones congénitas no solo comprenden una proporción sustancial de los ingresos en las unidades de cuidados intensivos neonatales, sino que también contribuyen significativamente a la morbilidad y mortalidad en los primeros meses de vida, estimándose que son responsables de alrededor del 30-50% de las muertes neonatales e infantiles. El diagnóstico de estas patologías es uno de los retos más complejos que afrontan los neonatólogos, ya que se caracterizan por una amplia variedad de síntomas y signos inespecíficos que varían no solo de una enfermedad a otra, sino también de un paciente a otro. Estas dificultades conducen inevitablemente a retrasos en el diagnóstico y tratamiento de los recién nacidos críticamente enfermos, una población especialmente vulnerable debido a su fragilidad.

Las nuevas técnicas de secuenciación genómica son herramientas muy prometedoras en el campo de la neonatología, ya que aumentan la tasa de diagnóstico y reducen el tiempo hasta el mismo. El avance de estas tecnologías permite el análisis simultáneo de cientos a miles de genes en muy poco tiempo y a bajo costo, ofreciendo varias estrategias para facilitar el diagnóstico de neonatos con sospecha de enfermedad genética, incluyendo la secuenciación de paneles de genes, del exoma clínico (CES), exoma completo (WES) y genoma completo (WGS). Sin embargo, varios obstáculos dificultan la incorporación de esta tecnología a la práctica clínica habitual, incluyendo la necesidad de recursos humanos, técnicos y económicos específicos con distintos niveles de complejidad.

En este proyecto se diseñó un panel *ad hoc* llamado NeoSeq para recién nacidos críticamente enfermos con sospecha de padecer una enfermedad genética el cual consta de 1.870 genes implicados en malformaciones congénitas, así como trastornos neurológicos y metabólicos de inicio temprano. El objetivo principal de este estudio fue evaluar prospectivamente la utilidad diagnóstica del panel NeoSeq en el diagnóstico molecular de neonatos críticamente enfermos con sospecha de enfermedad genética, siendo los objetivos secundarios comparar su rendimiento con otras herramientas de secuenciación masiva, aumentar el conocimiento de las bases moleculares de la enfermedad en esta población de pacientes y evaluar la utilidad de la escala Phenomizer a través del fenotipo descrito utilizando términos HPO. El fin último de este estudio es poder implementar el panel NeoSeq en el diagnóstico de enfermedades genéticas en neonatos críticamente enfermos ingresados en UCIN.

Para llevarlo a cabo se aplicó un enfoque en trío que incluía tanto al caso índice como a ambos padres. Este panel se aplicó a los pacientes críticos menores de dos meses ingresados en unidades de cuidados intensivos de tres hospitales terciarios de España que cumplían los siguientes criterios de inclusión y exclusión: (i) Malformaciones congénitas no relacionadas con un síndrome genético clínicamente identificable, (ii) descompensación metabólica no asociada a parámetros bioquímicos y / o detección neonatal que indicase una sospecha de enfermedad metabólica hereditaria, (iii) epilepsia o enfermedad del neurodesarrollo de probable origen genético. Se excluyó del estudio a los recién nacidos con hallazgos claros de un síndrome específico que podría diagnosticarse mediante un análisis dirigido de genes conocidos o variantes estructurales. Todas las muestras fueron fenotipadas utilizando términos HPO extraídos manualmente de los registros de salud electrónicos. Posteriormente se generó una lista de posibles diagnósticos

diferenciales en base a los hallazgos clínicos obtenidos utilizando la aplicación Phenomizer ([compbio.charite.de/phenomizer/](http://compbio.charite.de/phenomizer/)).

Finalmente, se reclutaron 33 de un total de 51 neonatos, en los que se estableció un diagnóstico genético en 12 (36,4%) pacientes con un tiempo medio de respuesta de 7,5 días. Además, se realizó un análisis de estudios previos prospectivos que empleaban diferentes herramientas de secuenciación masiva: secuenciación completa del genoma, exoma completo y exoma clínico en recién nacidos de cuidados intensivos con sospecha de trastornos genéticos, que arrojó una tasa de diagnóstico promedio comparable (37,2%) pero con una mediana de tiempo superior (30,2 días) que la obtenida con NeoSeq. El rendimiento del panel NeoSeq está fuertemente asociado con la especificidad de los hallazgos fenotípicos. Así, el panel logró una mayor tasa de diagnóstico en aquellos pacientes con sospecha de enfermedad metabólica y neurológica (43% y 35%, respectivamente) que en el grupo de anomalías congénitas, con un 23%. Al evaluar el impacto del diagnóstico genético en la evolución clínica, se encontró que el establecimiento de un diagnóstico molecular afectó directamente el manejo médico en 11 de los 12 pacientes, resultando en cambios en la medicación o dieta (9 pacientes), inicio de nuevas evaluaciones por subespecialistas específicos no consideradas previamente (5 pacientes) o retirada del tratamiento de cuidados intensivos con reorientación hacia los cuidados paliativos (5 pacientes). El análisis de los términos HPO mediante la herramienta Phenomizer correlacionó el fenotipo con el diagnóstico molecular obtenido de forma estadísticamente significativa ( $p < 0,05$ ) en el 41,6% de los pacientes.

Nuestro estudio demuestra que el panel de genes NeoSeq específico para fenotipo en recién nacidos críticamente enfermos es una alternativa asequible a la secuenciación del exoma clínico, del exoma completo y del genoma completo que ofrece una eficacia

diagnóstica comparable pero con un tiempo de respuesta más rápido, lo que respaldaría su aplicación como prueba de primer nivel en neonatos ingresados en unidades de cuidados intensivos. Además, el impacto del diagnóstico genético logrado con NeoSeq demuestra la utilidad clínica de esta herramienta en neonatos críticamente enfermos. Por otro lado, el abordaje trío permite obtener el diagnóstico definitivo de manera más rápida, al hacer innecesarios los estudios posteriores de segregación de las variantes identificadas en los pacientes índice.

En resumen, las tecnologías de secuenciación masiva y específicamente el panel NeoSeq tienen un enorme potencial como prueba de primera línea para el diagnóstico de neonatos críticamente enfermos con sospecha de padecer una enfermedad genética. Si bien se deberían esperar tasas de diagnóstico más altas de WGS que de WES, y de WES que CES o paneles específicos, descubrimos que esto no se correlacionó con el tamaño de la porción del genoma analizado ya que NeoSeq proporcionó una tasa de diagnóstico similar a las obtenidas con los enfoques WGS y WES, a pesar de interrogar 10 veces (WES) y 500 veces (WGS) menos pares de bases. La implementación clínica de estas pruebas también permite un adecuado asesoramiento genético para el paciente y su familia, así como adecuar la actitud terapéutica y el seguimiento avanzando en la línea de la medicina personalizada aplicada a una población de pacientes especialmente vulnerable.

**Palabras clave:** Tecnologías de Secuenciación Masiva, Paneles Genéticos, Enfermedades Raras, Recién Nacidos, Unidades de Cuidados Intensivos. Diagnóstico genético.

## ABSTRACT

Newborns with genetic disorders and congenital malformations not only comprise a substantial proportion of admissions to neonatal intensive care units, but also contribute significantly to morbidity and mortality in the first months of life, being estimated to be responsible for about 30-50% of neonatal and infant deaths. The diagnosis of these pathologies is one of the most complex challenges faced by neonatologists, as they are characterized by a wide variety of unspecific symptoms and signs that vary not only from one disease to another, but also from one patient to another while still suffering the same disease. These difficulties inevitably lead to delays in the diagnosis and treatment of critically ill neonates, a population especially vulnerable to their fragility.

New genomic sequencing techniques are proving very promising in the field of neonatology, increasing the rate of diagnosis and reducing the time to the same. The advancement of these technologies allows the simultaneous analysis of hundreds to thousands of genes in a very short time and at low cost, offering several strategies to promote the diagnosis of neonates with suspected genetic disease, including gene panels and clinical exome sequencing, complete exome and genome. However, several obstacles hinder the incorporation of this technology into routine clinical practice, including the need for specific human, technical, and economic resources with varying levels of complexity.

In this project a specific gene panel called NeoSeq was designed for critically ill newborns suspected of suffering from a genetic

disease consisting of 1870 genes involved in congenital malformations as well as early-onset neurological and metabolic disorders (before 2 months of age). The main objective of this study was to prospectively evaluate the diagnostic utility of the NeoSeq panel in the molecular diagnosis of critically ill neonates with suspected genetic disease, with secondary objectives of acquiring diagnostic performance with other massive sequencing tools, increasing knowledge of the molecular basis of disease in this patient population and to evaluate the usefulness of the Phenomizer scale through the phenotype described using HPO terms. The final objective of this study is to be able to implement the NeoSeq panel in the diagnosis of genetic diseases in critically ill neonates and to increase the knowledge of the molecular basis of genetic diseases in neonates admitted to the NICU.

A trio approach was applied including both the patient and both parents. This panel was applied to all critically ill patients under two months of age admitted to intensive care units of three tertiary hospitals in Spain who met the following inclusion and exclusion criteria: (i) Congenital malformations that are not related to a genetic syndrome clinically identifiable, (ii) metabolic decompensation not associated with biochemical parameters and / or neonatal screening indicating a suspicion of inherited metabolic disease, (iii) epilepsy or neurodevelopmental disease of probable genetic origin. Neonates with clear indications of a specific syndrome that could be diagnosed by targeted analysis of known genes or structural variants were excluded from the study. All samples were phenotyped using HPO terms manually extracted from electronic health records. A list of possible differential diagnoses was subsequently generated based on the clinical findings obtained using the Phenomizer application ([compbio.charite.de/phenomizer/](http://compbio.charite.de/phenomizer/)).

Finally 33 of a total of 51 neonates were recruited, in which a genetic diagnosis rate was established for 12 (36.4%) patients with a

mean response time of 7.5 days. In addition, an analysis of previous prospective mass sequencing studies with different diagnostic approaches was conducted: complete genome sequencing, complete exome and clinical exome in intensive care newborns with suspected genetic disorders, which yielded a comparable average diagnostic rate (37.2%) but with a higher median diagnostic time (30.2 days) than obtained with NeoSeq. Panel diagnostic performance is strongly associated with the specificity of phenotypic findings. Thus, the panel achieved a higher rate of diagnosis in those patients with suspected metabolic disease and neurological disease (43% and 35%, respectively), when compared with the diagnostic yield obtained in the group of congenital anomalies, with 23%. When assessing the impact of genetic diagnosis on clinical evolution, it was found that the establishment of a molecular diagnosis directly affected medical management in 11 of the 12 patients, resulting in changes in medication or diet (9 patients), initiation of new assessments by specific subspecialists. were previously considered (5 patients) or withdrawn from intensive care treatment with reorientation toward palliative care (5 patients). Analysis of HPO terms using the Phenomizer tool statistically significantly correlated the phenotype with the obtained molecular diagnosis ( $p < 0.05$ ) in 41.6% of patients.

Our study demonstrates that the phenotype-specific, gene-specific NeoSeq gene panel in critically ill neonates is an affordable alternative offering comparable diagnostic efficacy to complete exome sequencing with a faster response time, which would support its application as first-level test in neonatal intensive care units. In addition, the impact of the genetic diagnosis achieved with NeoSeq demonstrates the clinical utility of this tool in critically ill neonates. On the other hand, the trio approach allows to obtain the definitive diagnosis more quickly, by making unnecessary the subsequent studies of segregation of the variants identified in the index patients.

In summary, mass sequencing technologies and specifically the NeoSeq panel have enormous potential as a first-line test for the diagnosis of critically ill neonates suspected of suffering from a genetic disease. While higher diagnostic rates should be expected from WGS than WES, and from WES than CES or specific panels, we found that this did not correlate with the portion size of the genome analyzed as NeoSeq provided a diagnostic rate similar to those obtained with the approaches. WGS and WES, despite interrogating 10 times (WES) and 500 times (WGS) fewer base pairs. The clinical implementation of these tests also allows adequate genetic counseling for the patient and family, as well as adapting the therapeutic attitude and follow-up advancing in the line of personalized medicine applied to a particularly vulnerable patient population.

**Keywords:** Mass Sequencing Technologies, Genetic Panels, Rare Diseases, Newborns, Intensive Care Units. Genetic diagnosis.

## ABREVIATURAS

ACMG	Colegio Americano de Medicina Genómica
ADN	Ácido desoxirribonucleico
ARN	Ácido ribonucleico
ARN-seq	Secuenciación de ARN
B	Benigna
BFNE	Epilepsia Neonatal Familiar Benigna
BWA	Burrows-Wheeler Aligners
CES	Secuenciación de exoma clínico
CMA	Microarray Cromosómico
CNV	Variantes en el Número de Copias
CONDEL	CONsensus DELeiteriousness
CoQ10	Coenzima Q10
CPS1	Carbamoil fosfato sintetasa 1
DE	Desviación Estándar
DECIPHER	Base de datos de variantes genómicas y fenotipo en humanos mediante recursos Ensembl
EDTA	Ácido etilendiaminotetraacético
EEG	Electroencefalograma
ExAC	Consortio de Agregación de Exomas
FISH	Hibridación Fluorescente In Situ
GnomAD	Base de datos de Agregación de Genomas

GERP	Genomic Evolutionary Rate Profiling
HANAC	Angiopatía hereditaria con nefropatía, aneurismas y calambres musculares
HMG	Base de Datos de Mutaciones Genéticas Humanas
HPO	Ontología de Fenotipos Humanos
HSF	Human Splicing Finding
LB	Probable benigna
LCCS	Síndrome de Contractura Congénita Letal
LP	Probablemente patógena
MAF	Minor allele frequency
NGS	Secuenciación de nueva generación
NEE	Encefalopatía Epiléptica Precoz
OEIS	Onfalocela, extrofia vesical y rectal, ano imperforado y defectos espinales
OMIM	Online Mendelian Inheritance in Man
OTGS	Secuenciación Optimizada de Genoma en Tríos
OXPPOS	Sistema de Fosforilación Oxidativa Mitocondrial
P	Patógena
PCHW	Neuropatía desmielinizante periférica, dismielinización central, síndrome de Waardenburg y enfermedad de Hirschsprung
PCR	Reacción en cadena de la polimerasa
PhenIX	Interpretación Fenotípica de Exomas
SMA	Atrofia muscular espinal
SNP	Polimorfismos de un solo nucleótido
SNVs	Variante Individuales de Nucleótidos
RTA	Análisis en tiempo real

rWES	Secuenciación de Exoma Rápido
rWGS	Secuenciación de Genoma Rápido
UCIN	Unidad de Cuidados Intensivos Neonatales
UCIP	Unidad de Cuidados Intensivos Pediátricos
urWGS	Secuenciación de Genoma Ultrarrápido
VACTERL	Defectos vertebrales, atresia anal, defectos cardíacos, fístula tráqueo-esofágica, anomalías renales y anomalías de las extremidades
VATER	Defectos vertebrales, atresia anal, fístula tráqueo-esofágica, anomalías renales y anomalías de las extremidades
VCF	Variant calling format
VUS	Variante de significado incierto
WES	Secuenciación de exoma completo
WGS	Secuenciación de genoma completo



# ÍNDICE

RESUMO .....	15
RESUMEN .....	19
ABSTRACT .....	23
ABREVIATURAS .....	27
ÍNDICE DE TABLAS .....	35
ÍNDICE DE FIGURAS .....	37
1. INTRODUCCIÓN.....	39
1.1. Enfermedades raras .....	39
1.2. Retos para el diagnóstico de las enfermedades raras .....	40
1.3. Enfermedades genéticas en neonatos críticamente enfermos.....	42
1.4 Expresión clínica de las enfermedades genéticas en neonatos críticamente enfermos .....	43
1.4.1 Rasgos dismórficos.....	43
1.4.2 Cluster de signos y síntomas .....	46
1.4.3 Deterioro neurocognitivo .....	47
1.5 Retos para el diagnóstico de las enfermedades genéticas en neonatos críticamente enfermos .....	48
1.6 Técnicas de diagnóstico genético en pediatría .....	49
1.6.1 Diagnóstico genético clásico .....	49
1.6.2 Diagnóstico genético en la era de la secuenciación masiva ...	51
1.6.3 Pruebas de diagnóstico genético basados en tecnologías de secuenciación masiva .....	53

1.6.4 Enfoque en un solo probando o en tríos .....	55
1.7 Integración e interpretación de los datos .....	55
1.7.1 Llamada y anotación de variante .....	58
1.7.2 Filtrado de variantes .....	59
1.7.3 Asignación de patogenicidad .....	60
1.8. Consideraciones para la aplicación clínica .....	62
1.8.1 Human Phenotype Ontology .....	63
1.8.2 Experiencia clínica .....	66
1.8.3 Selección del paciente candidato a estudio genético .....	67
1.8.4 Elección del estudio genético .....	69
1.8.5 Consideraciones éticas y legales .....	70
1.9. Estudios piloto prospectivos que han aplicado técnicas de secuenciación masiva a neonatos críticamente enfermos .....	74
2. OBJETIVOS .....	87
2.1. Objetivo principal: .....	87
2.2 Objetivos secundarios: .....	87
2.3. Objetivo Final: .....	87
3. MATERIAL Y MÉTODOS .....	89
3.1. Diseño del estudio .....	89
3.2. Pacientes .....	89
3.3 Variables del estudio .....	90
3.4 Diseño del panel NeoSeq .....	91
3.5 Secuenciación NGS .....	91
3.5.1 Recepción de muestras y extracción de ADN .....	93
3.5.2 Librerías de ADN con captura de secuencia .....	94
3.5.3 Dilución de muestras, agrupación y preparación de bibliotecas para secuenciación .....	96
3.5.4 Secuenciación en la plataforma MiSeq (Illumina) .....	96
3.6 Análisis bioinformático .....	96
3.6.1 Análisis primario .....	97
3.6.1.1 Análisis de imágenes y llamadas de base .....	97
3.6.1.2 Control de calidad .....	98

3.6.2	Análisis secundario .....	99
3.6.2.1	Alineamiento de secuencias .....	99
3.6.2.2	Llamada de variante .....	100
3.6.3	Análisis terciario .....	102
3.6.3.1	Identificación y anotación de las variantes.....	102
3.6.3.1	Filtrado y priorización de las variantes .....	104
3.6.4	Informe de variantes.....	106
3.7.	Valoración clínica del resultado genético considerado positivo .....	107
4.	RESULTADOS .....	109
4.1.	Características de los pacientes a estudio.....	109
4.2.	Estudios genéticos. Tiempo de respuesta y tasa diagnóstica del panel NeoSeq.....	110
4.2.1	Pacientes con variantes definitivamente/probablemente causales .....	111
4.2.2	Pacientes con variantes no definitivamente/probablemente causales.....	134
4.2.3	Pacientes con enfermedad genética no confirmada mediante el panel NeoSeq.....	144
4.2.	Impacto del diagnóstico genético sobre el manejo clínico .....	145
5.	DISCUSIÓN.....	149
6.	CONCLUSIONES.....	171
7.	PERSPECTIVAS FUTURAS .....	173
8.	ANEXO I: Genes incluidos en el panel NeoSeq .....	175
9.	ANEXO II: Publicaciones .....	187
10.	ANEXO III: Permiso de reproducción de la Figura 1 .....	203
11.	REFERENCIAS .....	205



## ÍNDICE DE TABLAS

<b>Tabla 1.</b> Indicaciones para la realización de WES, WGS o paneles génicos en pediatría.....	69
<b>Tabla 2.</b> Variantes identificadas en el caso 1 .....	112
<b>Tabla 3.</b> Análisis in silico de la patogenicidad en el caso 1 .....	112
<b>Tabla 4.</b> Variantes identificadas en el caso 2 .....	114
<b>Tabla 5.</b> Análisis in silico de la patogenicidad en el caso 2 .....	114
<b>Tabla 6.</b> Variantes identificadas en el caso 3 .....	116
<b>Tabla 7.</b> Análisis in silico de la patogenicidad en el caso 3 .....	116
<b>Tabla 8.</b> Variantes identificadas en el caso 4 .....	117
<b>Tabla 9.</b> Análisis in silico de la patogenicidad en el caso 4 .....	117
<b>Tabla 10.</b> Variantes identificadas en el caso 5 .....	119
<b>Tabla 11.</b> Análisis in silico de la patogenicidad en el caso 5 .....	119
<b>Tabla 12.</b> Variantes identificadas en el caso 6 .....	120
<b>Tabla 13.</b> Análisis in silico de la patogenicidad en el caso 6 .....	120
<b>Tabla 14.</b> Variantes identificadas en el caso 7 .....	122
<b>Tabla 15.</b> Análisis in silico de la patogenicidad en el caso 7 .....	122
<b>Tabla 16.</b> Variantes identificadas en el caso 8 .....	124
<b>Tabla 17.</b> Análisis in silico de la patogenicidad en el caso 8 .....	124
<b>Tabla 18.</b> Variantes identificadas en el caso 9 .....	126

<b>Tabla 19.</b> Análisis in silico de la patogenicidad en el caso 9.....	126
<b>Tabla 20.</b> Variantes identificadas en el caso 10.....	128
<b>Tabla 21.</b> Análisis in silico de la patogenicidad en el caso 10.....	129
<b>Tabla 22.</b> Variantes identificadas en el caso 11.....	131
<b>Tabla 23.</b> Análisis in silico de la patogenicidad en el caso 11.....	131
<b>Tabla 24.</b> Variantes identificadas en el caso 12.....	133
<b>Tabla 25.</b> Análisis in silico de la patogenicidad en el caso 12.....	133
<b>Tabla 26.</b> Características clínicas y demográficas de los pacientes incluidos en el estudio. ....	138
<b>Tabla 27.</b> Hallazgos demográficos, clínicos y genéticos de los pacientes con variantes causales/probables. ....	141
<b>Tabla 28.</b> Hallazgos demográficos, clínicos y genéticos de los pacientes con variantes no definitivamente/no probablemente causales. ....	143
<b>Tabla 29.</b> Impacto del diagnóstico molecular en el manejo médico.....	148
<b>Tabla 30.</b> Estudios prospectivos del rendimiento diagnóstico de WES, WGS y CES en recién nacidos con sospecha de trastorno genético CES:secuenciación de exoma clínico.....	151

## ÍNDICE DE FIGURAS

<b>Figuras 1A y 1B.</b> Características dismórficas a tener en cuenta en la exploración física de un recién nacido. ....	44
<b>Figura 2.</b> Estrategias de secuenciación genética. ....	49
<b>Figura 3.</b> Integración de flujos de trabajo clínicos y de laboratorio para optimizar el diagnóstico de enfermedades raras mediante secuenciación de nueva generación. ....	57
<b>Figura 4.</b> Criterios de información de resultados en el proyecto BabySeq. ....	73
<b>Figura 5.</b> Flujo de trabajo para la secuenciación y análisis de diferentes regiones del genoma humano. ....	92
<b>Figura 6.</b> Comparación del panel NeoSeq con otras estrategias diagnósticas de secuenciación genética. ....	152
<b>Figura 7.</b> Retos y aproximaciones diagnósticas en pacientes con sospecha de enfermedad genética sin confirmación mediante NGS. ....	166



# 1. INTRODUCCIÓN

## 1.1. ENFERMEDADES RARAS

Las enfermedades raras, llamadas también “huérfanas”, “minoritarias” o “poco frecuentes” se definen, en función de su prevalencia, como aquellas patologías que afectan a un número bajo de personas en una población. El límite a partir del cual se considera minoritaria una enfermedad es variable, en la Unión Europea se denomina como enfermedad rara aquella en la que el número de sujetos afectados es menor de 5:10.000, mientras que en los EEUU incluye aquellas condiciones que padecen menos de 200.000 personas en el país, correspondiendo a una prevalencia aproximada de 60:100.000(1). Frecuentemente se trata de patologías crónicas y debilitantes, asociadas a discapacidad y a una disminución en la esperanza de vida, de las que se deriva un alto impacto socio económico (2).

La mayoría (60-75%) de las enfermedades raras afectan a los niños y muchos son trastornos multisistémicos graves con una amplia variedad de fenotipos. Tomadas en conjunto, las enfermedades raras son responsables del 35% de muertes en el primer año de vida y una importante causa de ingresos hospitalarios pediátricos (3); de hecho hasta un tercio de los niños que nacen con una enfermedad rara fallecerán antes de alcanzar los cinco años de edad. Se estima que en total existen 7000 enfermedades raras, de las cuales un 80% tienen una causa genética (4).

Debido a que estas enfermedades pueden ser difíciles de reconocer basándose únicamente en las características clínicas, el uso de pruebas genéticas en la población pediátrica es crítico para facilitar su diagnóstico y tratamiento (5).

La identificación de la causa molecular precisa (genotipo) que explican las características clínicas de la enfermedad (fenotipo) es la piedra angular de la asistencia de estos pacientes ya que desbloquea el acceso a una gran cantidad de información sobre el manejo adecuado, la terapia indicada y el pronóstico. Además, permite la determinación precisa del riesgo existente y futuro en otros miembros de la familia (6) y también facilita el acceso específico a grupos de apoyo, lo que reduce el aislamiento de las familias afectadas por un trastorno poco común. A reducir este aislamiento también se suman las asociaciones de afectados y familias. En este sentido, EURORDIS, una alianza no gubernamental de organizaciones de pacientes en Europa, representa a 929 organizaciones de pacientes con enfermedades raras en 72 países, siendo la voz de 30 millones de personas afectadas por enfermedades raras en toda Europa (7).

## **1.2 RETOS PARA EL DIAGNÓSTICO DE LAS ENFERMEDADES RARAS**

Encontrar un diagnóstico para cada individuo afecto sigue siendo un desafío considerable debido a la variabilidad genética y fenotípica asociada con estas enfermedades, así como a nuestro conocimiento incompleto y en constante evolución tanto de su historia natural como de sus bases moleculares (8, 9).

Superar este reto constituye uno de los objetivos principales del Consorcio Internacional de Investigación de Enfermedades Raras para el periodo 2017-2027 “permitiendo que todas las personas con una enfermedad rara reciban un diagnóstico preciso, cuidado y terapia disponible en el plazo de un año tras requerir asistencia médica”(10).

Las variantes genéticas causantes de enfermedad son diversas y varían en tamaño desde la sustitución, eliminación o duplicación de un único par de bases a variantes estructurales y variaciones de número de copia de un cromosoma completo (aneuploidía) o genoma (como

en el mosaicismo diploide / triploide). La enorme heterogeneidad, tanto genética como clínica, asociada a estas enfermedades es uno de los factores que explican la dificultad de su diagnóstico. Además estas enfermedades no sólo son muy raras o ultrararas (<1 : 50.000) a nivel individual sino que también es frecuente que presenten una gran variabilidad en su penetrancia y en la expresividad de sus características fenotípicas dentro de cada una de las entidades patológicas (5).

Una proporción de esta variabilidad asociada a las enfermedades genéticas se atribuye a la heterogeneidad del locus y la heterogeneidad alélica, que puede ser extrema; por ejemplo, la clínica característica de "discapacidad intelectual" puede ser causada por variantes monoalélicas o bialélicas en uno o más de 700 genes diferentes (11, 12). Otras fuentes importantes de variabilidad incluyen variantes genéticas en uno o más loci (modificadores) y factores ambientales, que actualmente son difíciles de identificar y cuantificar (13).

Por otro lado, hay que tener presente que no todos los trastornos raros son de origen genético y actualmente sigue habiendo una contribución pequeña pero importante de la exposición a teratógenos (por ejemplo, síndrome de alcohol fetal (14), toxoplasmosis congénita (15) o virus Zika (16)). Todos estos factores explican cómo, hasta hace poco, la mayoría de los pacientes con enfermedades raras, especialmente aquellos con trastornos del desarrollo, habían permanecido sin diagnosticar.

Debido a estas dificultades, el proceso para llegar a un diagnóstico se ha llegado a definir como "odisea diagnóstica"(17). Algunos pacientes viven durante meses, años o en muchos casos, toda su vida, con una enfermedad "sin nombre". Obtener un diagnóstico puede ser un proceso largo y difícil; una encuesta de EURORDIS sobre las ocho enfermedades raras más prevalentes en Europa mostró que el 25% de los pacientes esperaban más de cinco años hasta conocer la causa de su enfermedad, y

durante ese tiempo el 40% recibían un diagnóstico incorrecto lo cual supone la instauración de tratamientos inadecuados que pueden ser yatrogénicos. Además, otra importante consecuencia de no tener un diagnóstico es la angustia y el estrés que sienten los pacientes y sus familias ante la incertidumbre, que se combinan con sentimientos de aislamiento y exclusión. Es frecuente que durante este proceso el paciente sea derivado a diferentes especialistas y se le realicen numerosas pruebas complementarias, muchas de ellas invasivas.

### **1.3. ENFERMEDADES GENÉTICAS EN NEONATOS CRÍTICAMENTE ENFERMOS**

Los recién nacidos ingresados en una unidad de cuidados intensivos (UCIN) pueden padecer una enfermedad genética rara, siendo este grupo de enfermedades una causa importante de morbilidad y mortalidad tanto en el periodo neonatal como en la primera infancia (18). El término “trastornos genéticos” es amplio y en este contexto se refiere tanto a anomalías cromosómicas (19), como los síndromes de aneuploidía o trastornos por delección o duplicación cromosómica (20), como a enfermedades mendelianas (21), incluyendo los errores congénitos del metabolismo (22). Por otra parte, las malformaciones congénitas, que afectan aproximadamente al 2% de los nacidos vivos, son responsables de una gran proporción de los ingresos en una UCIN y son la causa principal de mortalidad infantil en los países desarrollados (23), aunque la etiología subyacente de estas malformaciones puede no ser genética, como en la embriopatía diabética o en la exposición intraútero a teratógenos. En resumen, los neonatos con trastornos genéticos y malformaciones congénitas que sugieren un posible síndrome genético no solo comprenden una proporción sustancial de las admisiones a la UCIN, sino que contribuyen de manera importante a la mortalidad neonatal, habiéndose estimado que dichos trastornos son responsables de alrededor del 30-50% de las muertes neonatales e infantiles (24, 25).

El diagnóstico genético tiene un impacto inmediato en el cuidado y manejo de estos pacientes. Utilizando los recursos diagnósticos actuales, en demasiadas ocasiones no es posible conocer la causa molecular de la enfermedad o si se conoce es de forma tardía. Es por ello que la Neonatología se ha convertido en una de las áreas en la que más interés ha despertado el uso de las nuevas herramientas de secuenciación genómica.

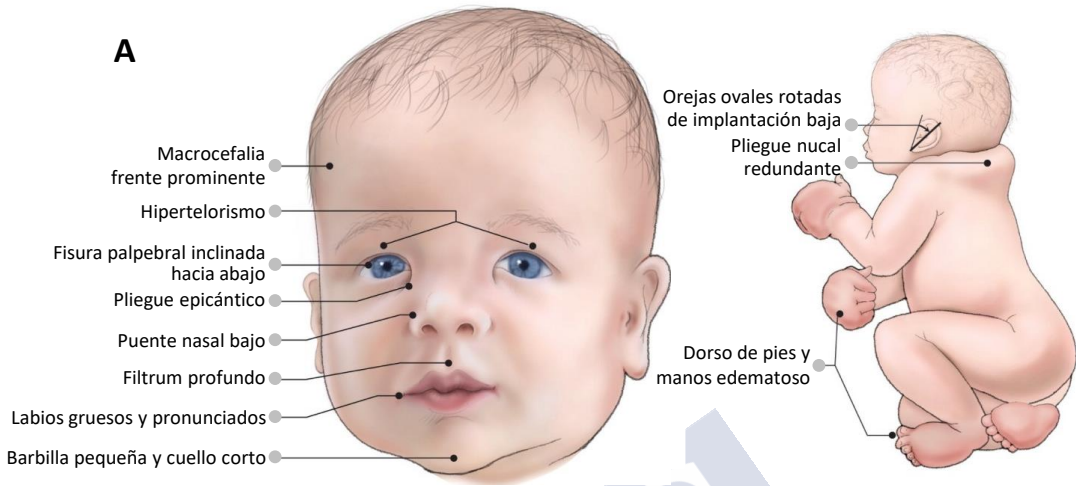
#### **1.4 EXPRESIÓN CLÍNICA DE LAS ENFERMEDADES GENÉTICAS EN NEONATOS CRÍTICAMENTE ENFERMOS**

Los hallazgos en el examen físico y la historia clínica que sugieren la presencia de una enfermedad genética pueden ayudar en el diagnóstico y tratamiento de los pacientes afectados, así como la derivación a las diferentes subespecialistas que precisen para su tratamiento y seguimiento, ya que la mayoría de estas patologías son multisistémicas. Entre los datos de alerta que pueden orientar hacia una condición genética se incluyen los rasgos dismórficos; las anomalías múltiples; el deterioro neurocognitivo inexplicable; y los antecedentes familiares que sugieran una enfermedad hereditaria (26).

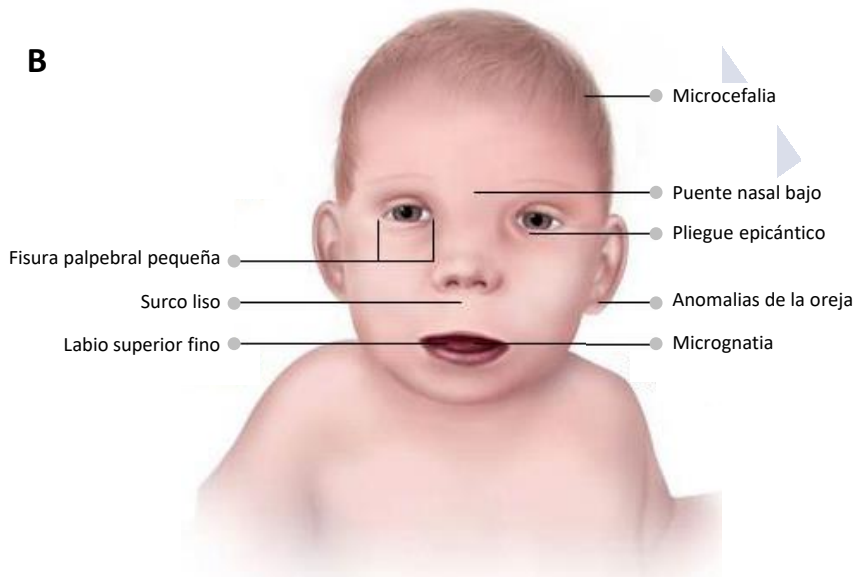
##### **1.4.1 Rasgos dismórficos**

Los rasgos dismórficos se deben a una perturbación en el desarrollo humano, la cual puede ser un efecto directo de una mutación genética o puede involucrar indirectamente un factor exógeno, como en el caso de la exposición gestacional a un teratógeno. Las Figuras 1A y 1B son ejemplos de características dismórficas que se pueden encontrar en el examen físico de un recién nacido, la primera en relación a síndrome de Noonan y la segunda a alcohol fetal.

**A**



**B**



**Figuras 1A y 1B. Características dismórficas a tener en cuenta en la exploración física de un recién nacido.** En la figura 1A se representan los hallazgos faciales y corporales de un neonato con síndrome de Noonan. En la figura 1B se muestran los hallazgos faciales en un recién nacido con síndrome de alcohol fetal. Figura 1A adaptada de Am Fam Physician. 2014 Jan 1;89(1):37-43 ([www.aafp.org/afp/2014/0101/p37.html](http://www.aafp.org/afp/2014/0101/p37.html)); y la figura 1B proveniente del Centers for Disease Control and Prevention. ANEXO III: Permiso de reproducción de la Figura 1.

Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the Centers for Disease Control and

prevention or the Department of Health and Human Services

Además del peso, la longitud y la circunferencia de la cabeza, ejemplos de otras medidas importantes incluyen la distancia interpupilar (la distancia entre el centro de las pupilas) si se sospecha un defecto en la línea media (como en el caso de holoprosencefalia, la malformación más común del prosencéfalo humano), o el tamaño de las extremidades en el caso de una displasia esquelética. Estos valores deben interpretarse en el contexto del crecimiento longitudinal del paciente, así como de los antecedentes familiares (27).

La ausencia o presencia de ciertas características en los miembros de la familia puede constituir una pista de que una característica física sea patológica o simplemente una variante familiar o étnica. Por ejemplo, los pliegues epicánticos internos (pequeños pliegues de piel sobre los ojos a nivel medial) pueden aparecer en personas con síndrome de Down y también se describen en más de otros 50 síndromes, incluido el síndrome de Noonan, el síndrome de Rubinstein-Taybi y el síndrome de Smith-Lemli-Opitz. Sin embargo, dichos pliegues también son un hallazgo normal en muchas personas de ascendencia asiática o nativa americana (28).

Recientemente, se ha hecho un esfuerzo para codificar descriptores físicos, que pueden ayudar a comunicar de forma estandarizada los hallazgos presentes en un paciente así como a establecer los potenciales diagnósticos diferenciales.

Debe tenerse en cuenta que la evidencia de una malformación obvia no debería limitar la evaluación completa, porque los hallazgos adicionales y más sutiles a menudo van a resultar de gran importancia en el diagnóstico diferencial. Por ejemplo, es posible que se observe

inmediatamente que un recién nacido tiene un paladar hendido, lo cual, como hallazgo aislado, tiene un extenso diagnóstico diferencial. Sin embargo, un paciente que además del paladar hendido presente hipocalcemia, un defecto del tabique ventricular, y sutiles características físicas, como una raíz nasal ancha, hélices de las orejas superpuestas y dedos estrechos y largos de manos y pies, probablemente estará afecto de un síndrome de delección 22q11.2 (anteriormente llamado síndrome de DiGeorge), que tiene una prevalencia de aproximadamente uno de cada 4.000 nacidos vivos (29).

#### 1.4.2 Cluster de signos y síntomas

Muchas condiciones genéticas se manifiestan a través de una combinación de características clínicas, que incluyen la apariencia física del paciente, anomalías de laboratorio y aspectos relativos a la familia. Por ejemplo, las anomalías del dedo pulgar de un recién nacido que también presenta anemia aplásica pueden sugerir anemia de Fanconi y la confirmación de este diagnóstico sería importante al permitir instaurar las medidas de seguimiento que permitan identificar precozmente las potenciales complicaciones a largo plazo (30). Otro ejemplo sería la presencia de rabdomiomas cardíacos y nódulos subependimarios en un paciente cuya madre o padre presenta fibromas periungueales y manchas hipocrómicas, lo cual nos orienta hacia una esclerosis tuberosa (31).

Se han propuesto varias hipótesis para explicar la presencia de un grupo de hallazgos en un paciente con un síndrome genético. Una de ellas es la pleiotropía, en la que una mutación en un solo gen ejerce múltiples efectos en diferentes órganos y sistemas. Por ejemplo, en el síndrome de Holt-Oram, a veces llamado síndrome corazón mano las variantes patógenas en el gen *TBX5* causan malformaciones congénitas tanto del corazón como de las extremidades (32). Otra posible explicación para la presencia de un grupo de hallazgos es que

el paciente tiene un síndrome genético contiguo de modo que se afectan todos los genes en los intervalos eliminados o duplicados dando lugar a un cuadro clínico complejo. El síndrome de deleción 22q11.2 es un ejemplo bien conocido de un síndrome genético contiguo (29).

### 1.4.3 Deterioro neurocognitivo

En muchas condiciones genéticas, el deterioro neurocognitivo puede ser el primer y más evidente signo de enfermedad. Las enfermedades genéticas pueden producir un afectación en el neurodesarrollo mediante diferentes mecanismos, los cuales incluyen tanto las malformaciones estructurales del sistema nervioso central, como la señalización aberrante que implica genes que desempeñan importantes funciones neurológicas así como la acumulación de sustancias tóxicas o el déficit energético de los errores innatos del metabolismo (33).

Estos últimos pueden causar una patología neurológica debido a una cantidad inadecuada de un sustrato necesario para el correcto funcionamiento del sistema nervioso o a la acumulación de un metabolito tóxico para el mismo. En este sentido la fenilcetonuria, la galactosemia, la aciduria metilmalónica y la enfermedad de la orina con jarabe de arce son ejemplos de afecciones causadas por errores innatos del metabolismo incluidos en el cribado neonatal, lo que permite su diagnóstico precoz. Sin embargo, muchas otras enfermedades metabólicas no están incluidas en dicho cribado, como es el caso de las enfermedades peroxisomales y lisosomales. Las manifestaciones de deterioro neurocognitivo en recién nacidos pueden incluir la hipotonía, las convulsiones y la encefalopatía. Desde el punto de vista bioquímico, la hiperamoniemia, la acidosis metabólica, la hipoglucemia y la hiperlactacidemia pueden sugerir un error congénito del metabolismo (34). Asimismo, en la evaluación del

retraso del desarrollo y el deterioro neurocognitivo, debe incluirse la realización de pruebas de neuroimagen ante la sospecha de un proceso estructural, degenerativo o metabólico.

Además, es fundamental recoger en la historia médica el grado y el tipo de disfunción; así como el momento en el que se observó por primera vez y cómo el deterioro ha cambiado con el tiempo (incluyendo si ha habido regresión, o pérdida de hitos del desarrollo); y si hubiese algún desencadenante reconocido.

Los errores innatos del metabolismo merecen una atención especial, porque las intervenciones médicas de forma precoz en estos trastornos son especialmente importantes para prevenir las secuelas neurológicas. Por ejemplo, puede ser necesario eliminar un tipo de alimento fuente que no se puede metabolizar adecuadamente. Estas intervenciones pueden influir en el grado de deterioro cognitivo, y pueden ser una cuestión de vida o muerte.

### **1.5 RETOS PARA EL DIAGNÓSTICO DE LAS ENFERMEDADES GENÉTICAS EN NEONATOS CRÍTICAMENTE ENFERMOS**




Si bien se sospecha una enfermedad genética en más del 50% de los niños que ingresan en la UCIN y mueren durante el primer año de vida, el diagnóstico se confirma solo en el 20-30% de los casos, a menudo post-mortem(35). Existen múltiples factores que dificultan el diagnóstico genético en los recién nacidos, incluida la heterogeneidad genética (a día 28 de febrero de 2021 existen 5615 enfermedades genéticas con causa molecular conocida en la base de datos OMIM) (36), la heterogeneidad clínica (por ejemplo, la aparición de formas frustradas de fenotipos clásicos) (37) y la comorbilidad relacionada con la fragilidad de los recién nacidos (38). Asimismo, la enfermedad tiende a progresar más rápidamente en el periodo neonatal que en otras etapas de la vida.

Por todos estos motivos, los pacientes de la UCIN podrían beneficiarse enormemente de un diagnóstico genético rápido, ya que éste mejoraría el manejo clínico del neonato críticamente enfermo, bien con tratamientos específicos o bien reorientando la atención hacia cuidados paliativos y la adecuación del esfuerzo terapéutico, lo cual es de particular importancia en la UCIN.

## 1.6 TÉCNICAS DE DIAGNÓSTICO GENÉTICO EN PEDIATRÍA

### 1.6.1 Diagnóstico genético clásico

Durante la mayor parte de su historia la genética clínica se ha basado en dos tipos de pruebas: análisis molecular de alta resolución de un solo gen y pruebas citogenéticas de todo el genoma de baja resolución (Figura 2) (39).

	Microscopía óptica	Cariotipo de bandas G	Microarray	Secuenciación del exoma completo	Secuenciación de genoma completo
Apariencia				<pre>CGGATGATTACCCGTT G.....GCTC TAGCTAGCTATA....</pre>	<pre>CGGATGATTACCCGTT GATATAGCTCTCGCTC GCTCTAGCTAGCTATA GCCTATGGGTGGGGGC</pre>
Resolución	Cromosoma completo	5-10 Mb	50-100 kb	1 bp	1 bp
Número de loci sondadas	N/A	-500	-0.05-2 millones	-50 millones	3 mil millones
Variantes detectadas	Aneuploidía, poliploidía	Variantes >5 Mb	Variantes en el número de copias	Regiones de codificación	Mayoría de variantes
Variantes por persona	0 o 1	0 o 1	10-100s	-20,000	4-5 millón
Rendimiento diagnóstico	Baja	—————→			Alta
Hallazgos incidentales	Baja	—————→			Alta

**Figura 2. Estrategias de secuenciación genética.** En la figura se compara la apariencia, la resolución, el número de locus analizados, las variantes por persona, la tasa diagnóstica y los hallazgos incidentales de las diferentes técnicas de diagnóstico genético. A medida que aumenta la resolución de la prueba, el número de variantes detectadas también aumenta, lo cual mejora el rendimiento

diagnóstico pero también incrementa sustancialmente la probabilidad de detectar hallazgos incidentales y variantes de importancia clínica incierta. Figura adaptada de Wright CF, et al. *Nat Rev Genet* 2018;19(5):253-268 (5).

En las pruebas moleculares de un solo gen, se selecciona un gen particular para la secuenciación o genotipado mediante la técnica de Sanger de acuerdo con la presentación clínica del paciente. Esta técnica fue desarrollada en 1977 y permite establecer la secuencia de tramos cortos de ADN, pudiendo identificar variantes de nucleótido individuales (SNVs), pequeñas inserciones y pequeñas deleciones. La probabilidad de que la prueba conduzca al diagnóstico depende de identificar correctamente la causa genética subyacente de la afección y seleccionar la prueba correcta. Por ejemplo, la mayoría de los trastornos genéticos "clásicos", como la fibrosis quística (causada por variantes en el gen *CFTR*) (40) o distrofia muscular de Duchenne (causada por variantes en el gen *DMD*) (41), se diagnostican utilizando este enfoque. Por el contrario, una herramienta que estudie el genoma completo a baja resolución, como un cariotipo basado en microscopía de banda G (resolución ~ 5-7Mb), se puede utilizar para diagnosticar trisomías comunes y aneuploidías segmentarias junto con algunos trastornos del desarrollo aparentemente esporádicos que son causados por grandes desequilibrios únicos de material cromosómico (42). Sin embargo, debe tenerse en cuenta que el rendimiento diagnóstico del cariotipo de banda G es limitado porque no suele detectar anomalías por debajo de las 5-10 megabases.

A principios de la década de 1990, se desarrolló la hibridación in situ mediante inmunofluorescencia (FISH). Esta técnica detecta anomalías genéticas por debajo del umbral del cariotipo de banda G, facilitando la identificación de eventos submicroscópicos (por ejemplo, deleciones adyacentes a los telómeros) implicados en enfermedad. Sin embargo, esta herramienta diagnóstica se ve limitada a la evaluación de las regiones cromosómicas que pueden ser objeto de sondas y las anomalías por debajo del rango de 50–300 kb no se pueden

detectar(43). El cariotipo de bandas G y el FISH han sido reemplazados en gran medida por el microarray cromosómico (CMA), que permite la evaluación simultánea de todos los cromosomas para los desequilibrios del número de copias y, en algunos casos, la isodisomía uniparental. El uso de CMA ha revelado que las anomalías citogenéticas submicroscópicas contribuyen significativamente a los defectos congénitos y trastornos neurológicos neonatales. El CMA suele ser una prueba de primer nivel en la evaluación de los niños que presentan enfermedades con sospecha de base genética, como las anomalías congénitas múltiples y el retraso del desarrollo, obteniéndose una tasa de diagnóstico de entre el 15% y el 20%, dependiendo de la indicación clínica(44). Existen diferencias entre las plataformas CMA que afectan al rango de anomalías detectadas. Las matrices basadas en conjuntos estándar de polimorfismos de un solo nucleótido (SNP) pueden tener una resolución algo menor para genes particulares en comparación con la hibridación genómica comparativa, pero son superiores para la detección de mosaicos de variaciones de número de copia (CNV), así como anomalías neutrales de copia clínicamente importantes e isodisomía. El CMA no puede detectar translocaciones equilibradas e inversiones(45).

Sin embargo, su sensibilidad es baja en comparación con las técnicas de secuenciación masiva y solo ~ 10% de los pacientes con una enfermedad pediátrica rara puede diagnosticarse utilizando estos enfoques(42).

### **1.6.2 Diagnóstico genético en la era de la secuenciación masiva**

Las tecnologías de secuenciación de nueva generación (NGS) (Figura 2) permiten la secuenciación simultánea de una gran cantidad de material genético, lo cual ha revolucionado la genética clínica al analizar múltiples genes o incluso exomas o genomas completos. Este enfoque de alta resolución combina los beneficios de las pruebas

moleculares y las pruebas citogenéticas de todo el genoma(46). Por ejemplo, el síndrome de Bardet-Biedl puede ser causado por variantes en más de 20 genes con presentaciones clínicas indistinguibles (47). Del mismo modo, hay muchas causas de deficiencia auditiva neurosensorial congénita que son clínicamente indistinguibles, pero genéticamente diferentes. En ambos casos, la secuenciación en paralelo de todos los genes que pueden estar potencialmente implicados mejora y acelera el proceso de diagnóstico. Las tecnologías de secuenciación masiva son particularmente útiles para diagnosticar trastornos del desarrollo que podrían ser causados por SNVs o pequeñas inserciones y/o deleciones (indels) en cualquiera de los miles de genes en el genoma donde probar cada gen candidato individualmente ya no es un enfoque factible.

Esta estrategia se ejemplifica mediante la aplicación de NGS al diagnóstico genético de encefalopatía epiléptica infantil, una entidad clínica heterogénea que puede ser causada por variantes en docenas de genes, donde la aplicación de las técnicas de NGS permitió la identificación de 31 genes nuevos entre 2012 y 2015. La sensibilidad de las tecnologías NGS para detectar SNVs e indels es extremadamente alta (48) y, aunque la secuenciación de Sanger todavía se considera la técnica de referencia para la confirmación de este tipo de diagnóstico, en algunos casos las tecnologías NGS pueden resultar superiores, permitiendo, por ejemplo, detectar las variantes en mosaico que sólo están presentes en un subconjunto de células. Los datos de NGS también se pueden utilizar para detectar CNVs y otras variantes estructurales, aunque su validez es inferior a la de los microarrays. No obstante, los paneles genéticos optimizados secuenciado a alta profundidad utilizando NGS han demostrado ser útiles para detectar pequeñas eliminaciones de exones (<10 kb) y tanto la secuenciación de exoma completo (WES) y genoma completo (WGS) se han utilizado con éxito para detectar grandes CNVs (>500 kb)(49).

Las técnicas de secuenciación masiva de lectura corta no detectan de forma fiable las expansiones de repetición de tripletes, que son la causa molecular de enfermedades como el síndrome de X frágil, la distrofia miotónica congénita y la ataxia de Friedreich (50). También presentan limitaciones en la detección de pequeños CNV, como eliminaciones de exones o duplicaciones, que son una de las principales causas de enfermedad en patologías como la distrofia muscular de Duchenne y atrofia muscular espinal (SMA) tipo 1. En zonas de baja cobertura pueden no detectarse variantes presentes en mosaico en el paciente o considerar como de novo variantes heredadas de progenitores que la presentan en mosaico somático y no está afectos (51). Por lo tanto, es crucial considerar los parámetros analíticos de cada ensayo para determinar si el gen o genes relevantes están convenientemente cubiertos y si es el adecuado para detectar el tipo de variante típicamente asociada con enfermedad.

### **1.6.3 Pruebas de diagnóstico genético basados en tecnologías de secuenciación masiva**

Las herramientas de NGS se pueden aplicar de diferentes maneras para el diagnóstico clínico, dependiendo del número y el tipo de regiones que se deseen estudiar. Los enfoques dirigidos incluyen la secuenciación de genes individuales y paneles de genes asociados a una enfermedad o fenotipo clínico concreto; así como la secuenciación de todos los exones de los genes actualmente conocidos como causantes de enfermedad, lo que también se conoce como el exoma clínico (CES)(52).

Por el contrario, la secuenciación de los exones de los 20.000 genes codificantes de proteínas incluidos en el exoma y los genomas completos son esencialmente pruebas no dirigidas (53). Todos los enfoques diagnósticos basados en técnicas de secuenciación masiva generan grandes cantidades de datos, pero la diferencia de escala entre

las pruebas puede ser muy amplia, desde unos pocos cientos de pares de bases de ADN para un gen pequeño, donde solo un cierto número de variantes serán detectadas, a 3 mil millones de pares de bases para un genoma completo, que como promedio contiene de 4 a 5 millones de variantes por persona.

A medida que se secuencian más genes en paralelo y se obtienen más datos la sensibilidad de la prueba aumenta, pero la especificidad disminuye y se incrementan los desafíos logísticos y éticos. En particular, el problema de cómo lidiar con hallazgos secundarios no solicitados es complejo debido a la enorme cantidad de variantes en cada genoma y la dificultad en su interpretación. Cuando se aplica un enfoque de panel de genes virtual a analizar datos de WES o WGS, se debe lograr un equilibrio entre limitar el panel a un conjunto muy específico y conservador de genes que están directamente relacionados con el fenotipo particular del paciente (por ejemplo, limitar las pruebas a los genes relacionados con la epilepsia en lugar de incluir todos los genes relacionados con trastornos del neurodesarrollo), lo que aumentará la especificidad de la prueba, y realizar una prueba no dirigida, como WES o WGS, que potencialmente será capaz de identificar cualquier defecto genético en el paciente a costa de disminuir la especificidad e incrementar la dificultad de la interpretación de los resultados.

El panel es clínicamente atractivo porque reduce el ruido de hallazgos no relacionados y hallazgos incidentales, pero puede resultar en errores diagnósticos debidos sobre todo a falsos negativos. Por otro lado, el número de variantes genéticas presentes en cada individuo, en muchos casos privadas, puede conducir a la sobreinterpretación de la relación de estas variantes con el fenotipo del paciente, lo que conduce a un sobrediagnóstico, que será más probable al aumentar la escala de la prueba. Todavía está en debate cuál es la mejor aproximación para ajustar el análisis genómico con el fin de maximizar la identificación

de verdaderas variantes de diagnóstico y minimizar la identificación de variantes de confusión.

#### **1.6.4 Enfoque en un solo probando o en tríos**

Una excelente estrategia para gestionar la riqueza de variantes genéticas identificadas por CES, WES o WGS es utilizar la aproximación en tríos, lo cual consiste en analizar las muestras del niño y de ambos padres biológicos. Esto permite filtrar variantes raras benignas privadas, identificar variantes de novo que están presentes sólo en el niño y confirmar la herencia.

Para familias en las que ninguno de los padres se ve afectado por el mismo trastorno que el hijo, la secuenciación de tríos ofrece una reducción de aproximadamente diez veces en el número de variantes candidatas, así como un aumento del 50% en el rendimiento diagnóstico aumentando sustancialmente la velocidad y probabilidad de llegar a un diagnóstico preciso (54).

Este aumento del poder analítico debe sopesarse con el coste adicional de la secuenciación parental, aunque dicho coste se puede compensar en parte por la eliminación de la necesidad de realizar estudios de segregación de variantes candidatas que es necesario en enfoques de un solo probando (55, 56).

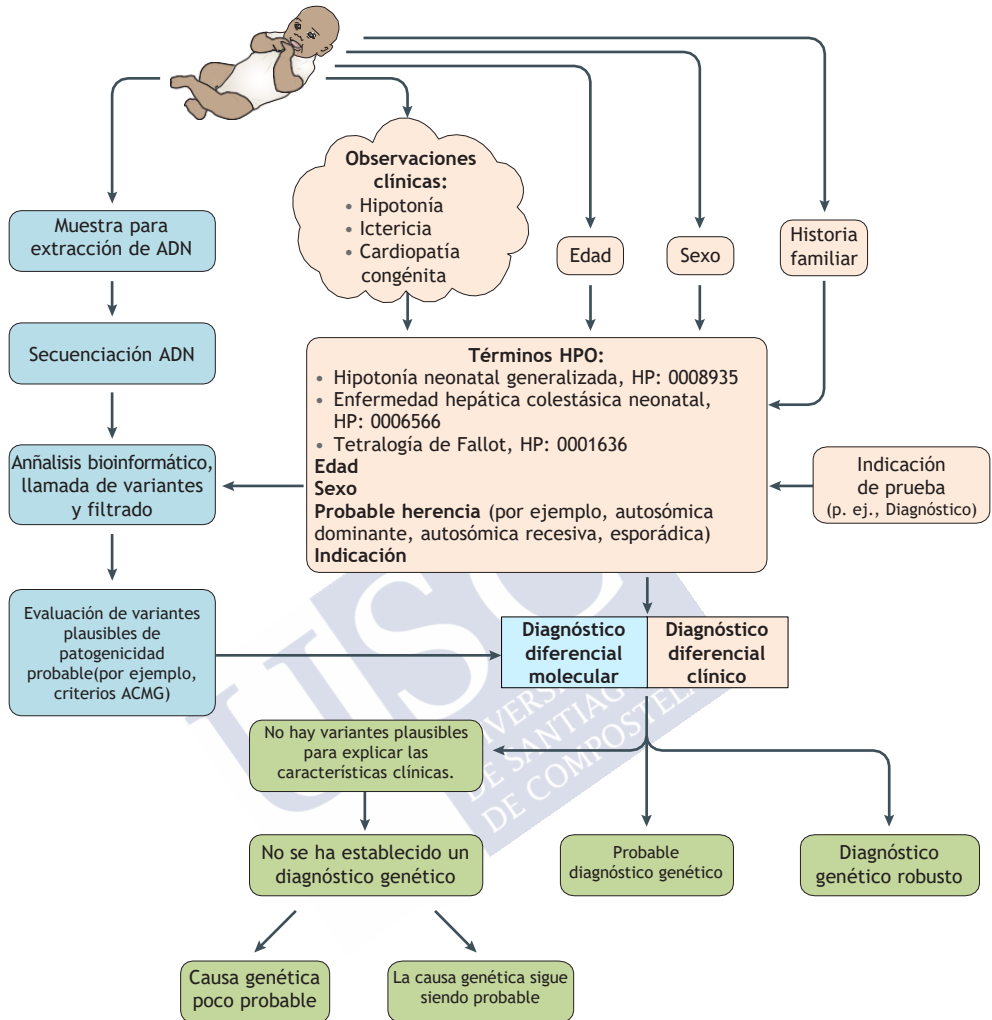
### **1.7 INTEGRACIÓN E INTERPRETACIÓN DE LOS DATOS**

Durante la última década, se ha vuelto cada vez más evidente que un genoma individual no puede interpretarse de forma aislada. Se requiere conocimiento de la variación de fondo en una población para filtrar variantes comunes. Idealmente, esta información se obtendría de forma aleatoria en individuos muestreados que no han sido seleccionados por cualquier característica clínica o social específica, para lo cual se necesitan miles de estos genomas. La clasificación

óptima de las variantes también requiere información clínica detallada y completa sobre el probando, en lugar de una simple etiqueta de diagnóstico. La escala de datos WES y WGS necesita la automatización del filtrado de variantes para centrarse en los intervalos genómicos y / o variantes relevantes con relación a la clínica.

Por otra parte, los flujos de trabajo analíticos para el diagnóstico de enfermedades raras que implican el procesamiento de datos de secuenciación crudos, la identificación de variantes patogénicas y la integración de datos clínicos para lograr un diagnóstico genético robusto son complejos y no están completamente estandarizados (Figura 3) (57, 58).





**Figura 3. Integración de flujos de trabajo clínicos y de laboratorio para optimizar el diagnóstico de enfermedades raras mediante secuenciación de nueva generación.** La implementación efectiva y segura de un servicio de genómica pediátrica requiere un diseño diligente para optimizar la experiencia de los médicos, bioinformáticos y científicos en varias partes del flujo de trabajo. Cada paso del proceso de filtrado de variantes merece una evaluación cuidadosa para reducir el número de variantes un número suficientemente bajo que permita la revisión de las mismas sin eliminar variantes valiosas para el diagnóstico. La evaluación clínica del paciente se lleva a cabo tanto al principio como al final de todo el proceso diagnóstico. Figura adaptada de Wright CF, et al. Nat Rev Genet 2018;19(5):253-268 (5)

### 1.7.1 Llamada y anotación de variante

Los procesos bioinformáticos de llamada y anotación de variantes por los que pasan los datos crudos generados tras la secuenciación del ADN, implican una serie de pasos claves en la posterior interpretación de los resultados.

La llamada y la anotación de variantes son procesos muy complejos, y la reproducibilidad para el análisis repetido de la misma muestra de WGS utilizando la misma canalización es generalmente inferior al 100% (59). La llamada de variante se usa para identificar diferencias alélicas en la composición base de un individuo comparando cada posición representada en los datos con la posición equivalente en un genoma humano de referencia. La sensibilidad depende de la calidad de los datos de prueba y de los datos de referencia, así como de los algoritmos empleados.

La presencia de datos de baja calidad o cobertura pueden resultar en la pérdida de variantes importantes para el diagnóstico. La variante de llamada, particularmente de pequeños indels, es difícil en secuencias repetitivas, lo que puede conducir a la omisión de diagnósticos a pesar de que la variante esté presente en los datos brutos (60).

Los datos de variante de llamada se almacenan en archivos de formato “*variant calling*” (VCF) estandarizados, que posteriormente se modifican para el análisis clínico incluyendo la anotación de cada variante (61). La anotación de las variantes implica la generación de una serie de parámetros que son cruciales para la interpretación clínica de dicha variante, especialmente la frecuencia alélica y la predicción de las consecuencias de la variante.

Todas las anotaciones, pero los datos frecuencia alélica en particular, varían con los datos de partida y el software utilizado en el momento de la anotación. Como resultado, los archivos VCF pueden variar sustancialmente entre proveedores y entre versiones a pesar de tener una estructura estandarizada (62).

### 1.7.2 Filtrado de variantes

Las decisiones sobre qué variantes incluir o excluir en la evaluación diagnóstica o en los análisis posteriores dependen en gran medida de los recursos utilizados y la fecha en la que se accedió a los mismos. Las variantes se suelen excluir si no tienen predicciones o consecuencias funcionales conocidas, si se considera que son demasiado comunes para ser la causa de una enfermedad en particular o si no encajan con el patrón observado de herencia para la enfermedad(63).

Si se aplica un enfoque en trío, se debe realizar una cuidadosa evaluación de los padres para asegurarse de que no tienen un fenotipo subclínico leve relevante para el trastorno en su hijo; de lo contrario, se establecerá un patrón de herencia incorrecta confundiendo el análisis. Filtrar variantes sin función conocida requiere una predicción precisa del efecto de la variante genética por software como por ejemplo Ensembl VEP71(64, 65) o SNPeff (66). Sin embargo, estas predicciones pueden no reflejar la consecuencia biológica real de la variante. Además, la misma variante genómica puede tener diferentes consecuencias a nivel pronóstico debido a la presencia de transcritos alternativos, ya que la falta de un método estandarizado y sencillo de selección para determinar el transcrito de la mayoría de los genes es una fuente significativa de variación entre los laboratorios de diagnóstico. La identificación de variantes comunes requiere acceso a recursos de población a gran escala, como ExAC (67) y gnomAD (68). A día de 28 de febrero de 2021, estas bases de datos contenían variantes de más de 125.748 exomas y 76.156 genomas de una serie de cohortes de estudio para adultos y niños. No obstante, muchos grupos siguen estando subrepresentados (69).

Las variantes genéticas que causan enfermedades graves en la infancia rara vez están presentes en este tipo de bases de datos, al contrario que lo que sucede con variantes asociadas con enfermedades

de debut en la edad adulta, ya que los factores relacionados con el estilo de vida y las exposiciones ambientales a largo plazo no contribuyen de forma significativa a la variabilidad de expresión de las enfermedades de aparición temprana.

En este sentido cuando el análisis se aplica a población pediátrica, puede ser conveniente establecer unos umbrales de frecuencia alélica muy bajos que permitan seleccionar sólo variantes muy raras. Sin embargo, la etnia del individuo que se está secuenciando y su grado de representación en conjuntos de datos de referencia puede afectar la eficacia del filtrado de variantes, y debe tenerse en cuenta que sigue habiendo un sesgo europeo en el análisis genómico (70).

### **1.7.3 Asignación de patogenicidad**

Un examen clínico exhaustivo sigue siendo crucial para interpretar las variantes candidatas identificadas por secuenciación genómica y para garantizar las buenas prácticas clínicas (Figura 2). Asignar patogenicidad a una variante rara identificada mediante el uso de un algoritmo bioinformático y establecer un diagnóstico genético puede ser un proceso extremadamente desafiante.

Las variantes en genes previamente asociados con una enfermedad en cuestión a menudo se priorizan utilizando paneles genéticos virtuales (71), así mismo la capacidad de identificar a las personas que no tienen una de las causas conocidas de una enfermedad a menudo facilita el descubrimiento de nuevos determinantes de enfermedades (72).

Desafortunadamente las asociaciones espúreas entre un gen y una enfermedad son relativamente frecuentes en la literatura, lo que resulta en enormes diferencias entre los paneles génicos y una falta de consenso claro con respecto al nivel de evidencia requerido para que un gen se incluya en un panel. Además, las bases de datos de variantes

patogénicas, como la *Human Gene Mutation Database* (HGMD) (73), contienen numerosos errores, ya que contienen variantes benignas a las que en algún momento se les asignó erróneamente patogenicidad (74, 75).

Es esperable que esta situación mejore a medida que aumente el número de genomas secuenciados, pero la importancia de estas grandes bases de datos que comprenden colecciones de genomas tanto de individuos sanos como de individuos con enfermedades raras, como gnomAD y ClinVar (76), no debe ser sobreestimada. Grandes cohortes pediátricas como *Deciphering Developmental Disorders Study* en el Reino Unido han facilitado el desarrollo de métodos robustos para establecer la patogenicidad únicamente sobre la base de datos genómicos, si bien estos estudios también cuentan con la ayuda de datos fenotípicos categóricos y cuantitativos (77). De hecho, los enfoques categóricos utilizan ontologías estructuradas, como el Human Phenotype Ontology (HPO) (78), para capturar características clínicas importantes, mientras que los datos cuantitativos en genética pediátrica incluyen el crecimiento y los hitos del desarrollo.

La publicación de las guías Association for Molecular Pathology (ACMG–AMP) por parte del American College of Medical Genetics and Genomics (ACMG) (79, 80) fue uno de los principales hitos hacia el establecimiento de un marco común para la clasificación de variantes.

Por otra parte, plataformas como DECIPHER (81, 82) permiten describir el fenotipo de forma estandarizada por parte de los clínicos y su integración con el genotipo como un registro a nivel de paciente dentro de un único recurso, lo que permite la valoración holística de datos relevantes. DECIPHER muestra información de fenotipos, datos de imágenes, información de genotipos (incluyendo tipo de variante (como SNV o CNV), consecuencia y modo de herencia) y la estructura de la proteína resultante en 2D y 3D cuando está disponible,

así como la referencia y los recursos poblacionales. Además, se complementa con una herramienta de evaluación para facilitar la toma de decisiones basadas en la integración de la experiencia clínica con el fin de lograr un diagnóstico genético sólido.

El uso sistemático de datos categóricos y fenotípicos cuantitativos para la clasificación fenotípica de cada niño con un trastorno genético raro o ultrararo también facilita los enfoques estadísticos para identificar patrones característicos de cada trastorno. Esto permitirá una evaluación de la concordancia entre el fenotipo observado en el individuo analizado y el fenotipo esperado asociado con cada uno de los genotipos causales plausibles que haya sido identificado mediante el análisis genético.

### **1.8. CONSIDERACIONES PARA LA APLICACIÓN CLÍNICA**

La experiencia para evaluar clínicamente la importancia diagnóstica de las variantes genómicas identificadas y comunicar los resultados significativos a la familia y a otros profesionales de la salud es un requisito previo esencial para la introducción de las técnicas NGS en un sistema sanitario (83). La búsqueda de un diagnóstico, que puede prolongarse a lo largo de años, a menudo se resuelve cuando la nueva tecnología (como la NGS) y los nuevos conocimientos (como el descubrimiento de nuevas enfermedades y genes) se combinan. El uso adecuado de estas herramientas tiene el potencial de acortar drásticamente la "odisea diagnóstica" para pacientes con trastornos pediátricos raros y ultra-raros (84, 85). Además, puede evitar la necesidad de realizar investigaciones más invasivas y costosas, como la biopsia de tejidos.

Sin embargo, y especialmente con pacientes muy jóvenes, las investigaciones que pueden consolidar un diagnóstico clínico deben continuar en paralelo con las pruebas genéticas de secuenciación

masiva, ya que los resultados pueden ser útiles para establecer la validez de los diagnósticos sugeridos por variantes candidatas. Investigaciones de seguimiento (como los estudios bioquímicos o enzimáticos o las pruebas de imagen) pueden ser necesarias para confirmar si las características fenotípicas normalmente asociados al gen candidato están presentes y si el diagnóstico sugerido es clínicamente plausible.

Cuanto menos específica sea la presentación (por ejemplo, la hipotonía) más difícil resultará establecer la relevancia diagnóstica de las variantes candidatas identificadas. Con una cuidadosa selección del paciente en una UCIN, ha sido posible establecer un diagnóstico genético para neonatos gravemente enfermos en tan solo 26 horas (86), pero este período de tiempo sigue siendo inalcanzable en la mayoría de los servicios de salud debido a la falta de recursos tecnológicos, así como a los costes sobreañadidos. Sin embargo, estos resultados demuestran que el diagnóstico genético rápido es factible y proporciona un fuerte impulso para mejorar la aplicabilidad de estas herramientas en la práctica clínica habitual, aspirando a aumentar su velocidad de respuesta pero manteniendo un equilibrio con los recursos necesarios, tanto humanos como económicos, para analizar los resultados obtenidos (87).

### **1.8.1 Human Phenotype Ontology**

La plataforma HPO proporciona recursos bioinformáticos integrales para el análisis de enfermedades y fenotipos humanos, ofreciendo un puente computacional entre la biología del genoma y la medicina clínica (88). HPO se publicó inicialmente en 2008 con el objetivo de permitir la integración de la información fenotípica en todos los campos científicos y bases de datos. Desde entonces, el proyecto ha crecido en términos de cobertura, alcance y sofisticación (89), y se ha convertido en un componente central de la Iniciativa

Monarch, permitiendo el análisis computacional de especies cruzadas (90).

La descripción de la variación fenotípica se ha convertido en un tema central para la investigación traslacional y la medicina genómica, y las descripciones 'computables' de la enfermedad humana utilizando perfiles fenotípicos HPO (también conocidos como 'anotaciones') se han convertido en un elemento clave en una serie de algoritmos que se utilizan para respaldar el descubrimiento y el diagnóstico genómico (91).

HPO está organizada como subontologías independientes que cubren diferentes categorías. La categoría más grande es la anomalía fenotípica. La subontología del modo de herencia permite definir modelos de enfermedad según los modos de herencia mendeliano o no mendeliano. De manera similar, la subontología Mortalidad / Envejecimiento permite anotar la edad de muerte típicamente asociada con una enfermedad u observada en un individuo específico. Finalmente, la subontología del modificador clínico está diseñada para proporcionar términos para caracterizar y especificar las anomalías fenotípicas definidas en la subontología de anomalías fenotípicas, con respecto a la gravedad, lateralidad, edad de inicio y otros aspectos.

La principal aplicación de dominio de HPO, hasta la fecha, ha sido sobre trastornos raros, y proporciona un gran corpus de perfiles de anotación de enfermedad-HPO utilizando OMIM, Orphanet y DECIPHER para entidades patológicas. Con los avances recientes en la medicina personalizada, es cada vez más importante proporcionar una base computacional para el análisis de genomas impulsado por el fenotipo y otras investigaciones traslacionales en otros campos de la medicina.

El uso de términos HPO para pacientes con epilepsia es un desafío. A diferencia de muchos otros trastornos genéticos, las

características fenotípicas de los pacientes con epilepsia son dinámicas y las características específicas, como la semiología de una convulsión compleja, a menudo son difíciles de incluir por completo en las ontologías fenotípicas sistemáticas. Por ejemplo, un paciente con convulsiones febriles simples puede tener convulsiones febriles autolimitadas, puede tener convulsiones febriles recurrentes después de los seis años o puede desarrollar la epilepsia intratable relacionada con la fiebre del síndrome de Dravet con el tiempo. Las tres entidades son distintas, pero dependiendo de la edad de los pacientes, pueden codificarse de forma idéntica en el HPO si no se utilizan modificadores que codifican la edad del paciente. El dilema de representar completamente los fenotipos neurológicos dinámicos enfatiza la necesidad del uso continuo de modificadores de HPO para lograr la dimensionalidad en los datos del fenotipo.

HPO se emplea para incorporar datos clínicos en el análisis de un panel de secuenciación de diagnóstico de NGS con casi todos los genes asociados con la enfermedad mendeliana conocidos; el algoritmo, interpretación fenotípica de exomas (PhenIX) contribuyó a una tasa de diagnóstico del 28% en los niños en los que los exámenes exhaustivos previos no habían revelado un diagnóstico (92). El uso de HPO para generar paneles de genes individualizados impulsados por fenotipos para el diagnóstico condujo a un aumento en el rendimiento diagnóstico(93). El Thrombo Genomics Consortium informó que la priorización computacional de variantes raras candidatas identificadas en pacientes con trastornos hemorrágicos, trombóticos o plaquetarios utilizando fenotipos codificados por HPO asignó las puntuaciones más altas a variantes patogénicas o probables patogénicas en el 85% de los casos (94), lo que demuestra que los algoritmos basados en HPO pueden hacer que las reuniones de diagnóstico multidisciplinarias sean más eficientes.

### 1.8.2 Experiencia clínica

Aunque para algunas patologías (como por ejemplo la acondroplasia) es posible establecer un diagnóstico molecular definitivo a partir del genotipo o un juicio clínico cierto a partir del fenotipo, el diagnóstico genético de la mayoría de las enfermedades raras requiere de la integración y correlación de los datos genómicos (genotipo) con las características clínicas (fenotipo) del paciente. La profundidad del fenotipo necesaria para hacerlo con precisión aún no está establecida para la mayoría de las condiciones, y se recomienda la evaluación especializada por un médico con experiencia en enfermedades raras para asegurar que un diagnóstico molecular de enfermedad genética en un paciente pediátrico está bien fundamentado. Del mismo modo, tanto si el diagnóstico molecular representa una explicación completa para la presentación clínica del paciente, como si éste presenta un fenotipo poco específico, también es apropiada una evaluación por un experto.

Hacer un diagnóstico genético de por vida en un paciente joven con un trastorno grave es un paso crucial para permitir un tratamiento y seguimiento adecuado para ese paciente y su familia. Dependiendo de las variantes genéticas, este proceso a menudo requiere revisión de publicaciones primarias recientes y evaluación clínica adicional para corroborar o refutar un diagnóstico, así como estudios de segregación para determinar si el diagnóstico molecular sugerido encaja con el patrón de herencia observado en la familia. Además, para pacientes con trastornos raros o ultra raros, es necesario dedicar tiempo para investigar adecuadamente la disponibilidad de estudios de investigación y ensayos de tratamiento. Este proceso es lento, pero esencial para la práctica segura en medicina genómica; el tiempo y la experiencia necesarios para realizarlo con garantías puede no ser fácil de acomodar en una consulta pediátrica sin el apoyo de un equipo multidisciplinar.

Asimismo, los pacientes con un informe inicial negativo de una prueba genética de secuenciación masiva pero en los que se sospecha fuertemente una base genética para la enfermedad también se beneficiarán de la evaluación de expertos en genética clínica. Por otra parte, a medida que el uso del WGS o el WES se generaliza el reanálisis clínicamente orientado de los datos existentes centrados en el paciente mejora el rendimiento diagnóstico mediante la identificación de variantes que se hayan podido pasar por alto en el análisis estándar inicial (95).

### **1.8.3 Selección del paciente candidato a estudio genético**

La identificación de los pacientes adecuados para someterse a un estudio genético ha evolucionado sustancialmente en los últimos 25 años, coincidiendo con la capacidad de diagnosticar de una manera cada vez más eficaz y económica un número creciente de trastornos. El ACMG publicó una guía de indicaciones para los casos en los que debe considerarse el uso de los estudios genéticos, incluyendo aquellos pacientes con una alta probabilidad de presentar una causa monogénica de enfermedad que no es fácilmente identificable (96). Ampliando esta idea, Bick et al. (42) sugieren algunas características generales adicionales de los trastornos que requerirían un abordaje mediante técnicas de secuenciación masiva:

- Datos del fenotipo o de la historia familiar que implican fuertemente una etiología genética, pero el fenotipo no se corresponde con un trastorno para el que exista una prueba genética dirigida a un gen específico.
- Diagnóstico clínico de un trastorno que se sabe que es causado por múltiples genes (heterogeneidad extensa del locus).

- Características clínicas que son insuficientes por sí mismas para hacer un diagnóstico clínico, pero que se sabe que están asociadas con múltiples trastornos genéticos.
- Un diagnóstico clínico de un trastorno genético conocido en el que un solo gen u otras pruebas dirigidas ha sido negativo.
- El paciente tiene un curso clínico atípico para la enfermedad considerada (por ejemplo, gravedad o duración inesperada, reacción idiosincrásica a los medicamentos).
- Inicio temprano de un trastorno que normalmente se observa en la edad adulta.
- Alteraciones clínicas o analíticas raras y específicas.
- Combinaciones atípicas o complejas de anomalías clínicas o signos y síntomas adicionales no explicados por un diagnóstico.

Por otra parte, Wright et al.(5) proponen las siguientes indicaciones para identificar pacientes pediátricos que se puedan beneficiar de un estudio genético mediante técnicas de secuenciación masiva (Tabla 1):

**Tabla 1.** Indicaciones para la realización de WES, WGS o paneles génicos en pediatría

1. Trastorno del neurodesarrollo incluyendo retraso psicomotor y/o dificultad de aprendizaje, encefalopatía epiléptica o parálisis cerebral grave.
2. Anomalías congénitas incluyendo múltiples anomalías congénitas (dos o más) o una sola anomalía congénita mayor junto a retraso de neurodesarrollo, crecimiento anormal, características dismórficas o comportamiento inusual.
3. Datos somatométricos anormales (talla/longitud, peso, perímetro de cráneo): dos o más parámetros superior o inferiores a  $>3$  DE o un parámetro superior o inferior a  $>4$  DE de media (excepto para obesidad aislada, donde se debe considerar  $<4.5$  DE y una fuerte sospecha de etiología genética).
4. Rasgos dismórficos.
5. Fenotipo comportamental anormal en conjunto con alguna de las anteriores o fenotipo comportamental extremo sugestivo de base genética.
6. Trastorno de magnitud considerable en el que se sospecha una base genética con alguna de las siguientes características: varios miembros de la familia afectados; otro miembro de la familia afecto con un fenotipo raro, distintivo y consistente; o caso único asociado con un fenotipo particular grave.

#### 1.8.4 Elección del estudio genético

La capacidad para descifrar el código genético ha mejorado notablemente en los últimos 10 años debido a múltiples circunstancias, las cuales incluyen costes exponencialmente decrecientes, una mayor precisión de las técnicas de secuenciación y un rápido desarrollo de bibliotecas de genes relacionados con diferentes patologías(97, 98). La elección de un test concreto de secuenciación masiva está influenciada por diferentes factores, entre los que se encuentran la especificidad de la clínica y la disponibilidad de recursos, tanto materiales como humanos. Los paneles de genes dirigidos tienen un coste inferior que los exomas y los genomas y proporcionan una profundidad de cobertura excelente, pero algunos genes de interés pueden no estar incluidos (99). Secuenciar un exoma

es más barato que secuenciar un genoma pero la cobertura de algunos exones puede ser subóptima, especialmente en regiones ricas en guanina y citosina (100). El análisis de genoma completo proporciona la mejor cobertura, pero también la más cara y la más laboriosa (101).

El tiempo de respuesta de cada una de estas pruebas también es variable, oscilando entre semanas y meses en función de la complejidad de su análisis, siendo más rápido en el caso de los paneles de genes y más largo para el exoma y el genoma. La aplicación de estas pruebas en unidades de cuidados intensivos neonatales y pediátricos tiene un notable atractivo. El reciente desarrollo de protocolos rápidos de secuenciación de exomas (rES) y genomas (rGS) en un tiempo aproximado de 1 a 2 semanas ha permitido la aplicación exitosa de estas pruebas en cuidados intensivos neonatales y pediátricos identificando números inesperados de lactantes con enfermedades agudas de origen genético y cambiando positivamente su atención médica (87, 102, 103). A pesar de estos éxitos, los rES y rGS siguen siendo costosos y no están fácilmente disponibles para la mayoría de las unidades de cuidados intensivos que atienden bebés y niños. Examinar un panel de genes con función conocida y las asociaciones de enfermedades establecidas simplificarían la interpretación de los datos pudiendo proporcionar una alternativa más simple a rES y rGS, aún con tasas de diagnóstico comparables (104).

### **1.8.5 Consideraciones éticas y legales**

Hay importantes implicaciones éticas asociadas con la aplicación clínica de las tecnologías de secuenciación masiva, particularmente en niños. La secuenciación de exomas y genomas identifica con frecuencia hallazgos incidentales o secundarios: variantes de importancia potencial para el niño o la familia que no están relacionados con las enfermedades para las que se realiza la prueba.

La presentación de conclusiones incidentales es controvertida y ha dado lugar a políticas de recomendaciones en muchos casos contradictorias (105, 106). Algunos grupos sugieren no informar variantes patogénicas que no impliquen una intervención terapéutica. Otros recomiendan ofrecer los hallazgos patogénicos tanto en trastornos tratables como intratables, así como el estado portador de enfermedades recesivas (107). Esta discusión es particularmente relevante en la población pediátrica ya que estos pacientes no se consideran legalmente competentes cuando se examinan, pero adquirirán competencia a medida que crecen (108). El hecho de que muchos adultos opten por no someterse a pruebas genéticas cuando se les ofrecen (109) genera dudas con respecto a la protección de la autonomía y la privacidad futuras (110).

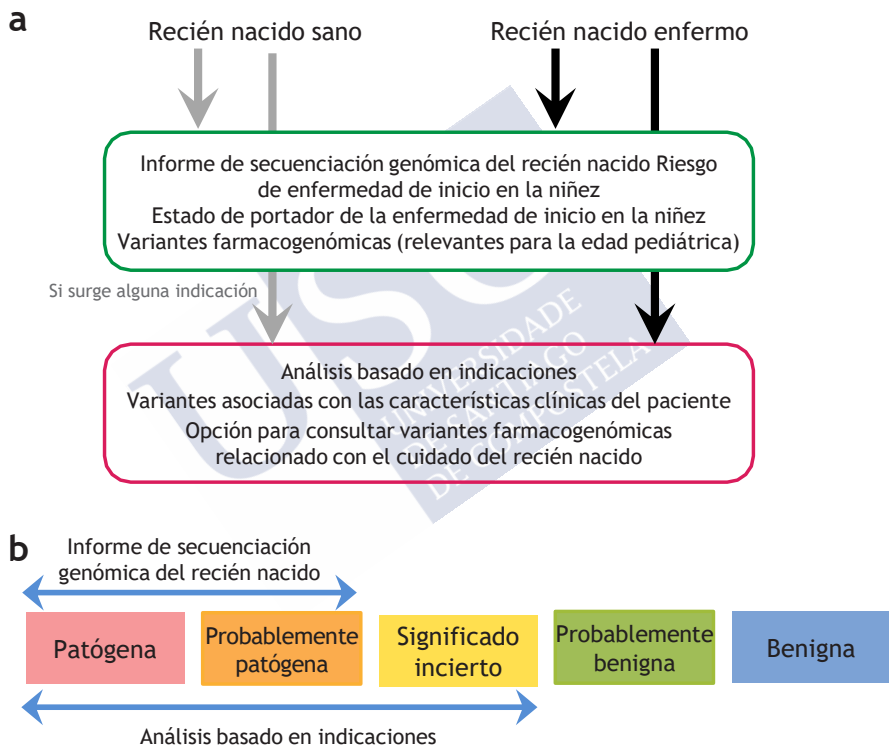
Es por ello que la gran cantidad de variantes que se obtienen tras el análisis de exomas y genomas representa un desafío significativo para su uso en la práctica clínica habitual.

Los laboratorios suelen utilizar herramientas bioinformáticas para reducir el número de variantes que requieren evaluación manual, las cuales se basan en un conjunto de criterios de filtración como la frecuencia alélica, el impacto previsto sobre la proteína y la información disponible en bases de datos sobre la patogenicidad de las variantes. Sin embargo, una parte importante del proceso de interpretación -determinar si el gen afectado por la variante está fuertemente asociado con la enfermedad y cumple con los criterios para ser informado como causa de la enfermedad- implica revisar manualmente la validez del papel de cada gen en la enfermedad teniendo en cuenta otros atributos como la penetrancia y la edad de comienzo.

En este sentido, las herramientas bioinformáticas que integran elementos fenotípicos en la selección de variantes son una opción muy atractiva para generar una lista de genes y variantes candidatos y así disminuir tanto el tiempo de análisis como los hallazgos incidentales.

Para dar respuesta a estas preguntas sobre las tecnologías de secuenciación masiva aplicadas específicamente a población neonatal, es interesante reseñar el proyecto BabySeq (111). El proyecto BabySeq es un ensayo controlado y aleatorizado diseñado para explorar los resultados médicos, conductuales y económicos asociados con el uso de la secuenciación masiva, en concreto del análisis del genoma, en recién nacidos (112). Se seleccionaron dos cohortes de neonatos: recién nacidos sanos de la maternidad y recién nacidos enfermos de las unidades de cuidados intensivos neonatales y pediátricos. Con el fin de facilitar la interpretación de resultados, los investigadores seleccionaron 1.514 genes para analizar el papel de cada uno de ellos en la enfermedad, la edad de aparición, la penetrancia y el modo de herencia basada en un conjunto de criterios determinados por un grupo interdisciplinar de genetistas clínicos y moleculares, pediatras, neonatólogos y expertos en bioética. Usando esta información determinaron un conjunto de criterios para el retorno de resultados que se incorporaron al protocolo del proyecto. Los criterios se desarrollaron para maximizar el beneficio mientras se minimiza la incertidumbre de reportar trastornos con baja penetrancia, aparición tardía o evidencia subóptima de asociación, limitándose a cuatro grupos de resultados: (i) inicio en la niñez (inicio más temprano informado antes de los 18 años), involucrando genes con evidencia sólida para causar trastornos de aparición en la niñez altamente penetrantes; (ii) genes con evidencia moderada y / o penetrancia moderada asociada con condiciones para las cuales la acción durante la infancia puede prevenir un resultado devastador más adelante en la vida; (iii) genes con fuertes asociaciones farmacogenómicas (clase 1 y genes 2A en la base de datos PharmGKB (<https://www.pharmgkb.org>)) que son relevantes para la población pediátrica, incluyendo *RYRI* asociado con hipertermia maligna, *TPMT* asociado con toxicidad por tiopurina, y *G6PD* asociado con anemia hemolítica debido a la deficiencia de glucosa-6 fosfato deshidrogenasa; y (iv) estado de portador de cualquier gen que cumpla con estos criterios (Figura 4).

En el caso de los neonatos sanos se decidió informar sólo de las variantes patogénicas o muy probablemente patogénicas, mientras que en el caso de los neonatos críticamente enfermos todas las variantes con evidencia para apoyar una contribución clara o posible a la indicación del recién nacido deberían ser informadas. Finalmente, y en base a estos criterios, se filtraron para informar 954 genes entre los más de 20000 incluidos en la secuenciación del genoma(109).



**Figura 4. Criterios de información de resultados en el proyecto BabySeq.** Todos los recién nacidos del grupo sano reciben un informe de secuenciación genómica de recién nacidos que devuelve el riesgo y el estado portador de la enfermedad de inicio infantil y variantes farmacogenómicas que pueden ser relevantes para la población pediátrica. Además los recién nacidos enfermos reciben un análisis basado en indicaciones que devuelve todas las variantes con evidencia para causar o contribuir a la enfermedad del bebé, con una opción para consultar variantes farmacogenómicas relacionadas con el cuidado del bebé. Figura adaptada de Ceyhan-Birsoy O, et al. Genet Med 2017;19(7):809-818 (113).

### 1.9. ESTUDIOS PILOTO PROSPECTIVOS QUE HAN APLICADO TÉCNICAS DE SECUENCIACIÓN MASIVA A NEONATOS CRÍTICAMENTE ENFERMOS

Las enfermedades raras a menudo se presentan en los primeros días y semanas de vida, pudiendo requerir un manejo complejo en el contexto de una UCIN. Las investigaciones genéticas o metabólicas tradicionales son costosas y, a menudo, no logran llegar a un diagnóstico final cuando no se sospecha un síndrome reconocible. Por este motivo, hasta la fecha se han llevado a cabo proyectos piloto evaluando la viabilidad de diferentes técnicas de secuenciación masiva como herramienta para mejorar el diagnóstico de enfermedades raras en recién nacidos ingresados en la UCIN. En este sentido, debe tenerse en consideración que se trata de estudios realizados en el ámbito de la investigación y que se han basado en diferentes aproximaciones: secuenciación de CES, WES y WGS.

A continuación se detallan las características de cada uno de ellos, agrupándolos según el tipo de herramienta de secuenciación masiva empleada:

- a) *Secuenciación de CES*. Este tipo de aproximación se basa en la secuenciación de las regiones codificantes de un grupo de genes seleccionados. Los estudios piloto que han aplicado esta aproximación a pacientes neonatales críticamente enfermos incluyen los siguientes:
  - Daoud, et al (102) valoraron su utilización basada en el empleo de un panel grande en tríos que incluye 4813 genes relacionados con enfermedad. Los recién nacidos incluidos debían tener al menos uno de los siguientes hallazgos: una o más malformaciones congénitas, características dismórficas, anomalías en los parámetros de crecimiento, deterioro neurológico (incluida encefalopatía, convulsiones o

hipotonía) o características que sugieran una afección metabólica. De los 20 recién nacidos estudiados, 8 recibieron un diagnóstico molecular (tasa de diagnóstico del 40%). Los diagnósticos fueron disgenesia tubular renal, síndrome de encefalopatía relacionada con SCN1A, miopatía miotubular, síndrome de deficiencia de FTO, displasia craneoectodérmica, síndrome miasténico congénito, síndrome de discapacidad intelectual autosómica dominante tipo 7 y síndrome de Denys-Drash. De estos pacientes dos tuvieron un cambio en el tratamiento médico debido al diagnóstico. Para el paciente con mutaciones en el gen de la enzima de conversión de la angiotensina, el diagnóstico molecular guió la evaluación posterior de la aldosterona y el inicio de la fludrocortisona, que finalmente resultó en una mejor función renal. En otro paciente, el diagnóstico de una variante de novo en *WT1* supuso un cambio en el manejo quirúrgico, debiéndose extirpar ambos riñones durante el trasplante de riñón, para reducir el riesgo sustancial de tumor de Wilms. El conocimiento del diagnóstico en los otros 6 casos brindó información beneficiosa a sus médicos sobre la historia natural y el pronóstico y permitió un asesoramiento preciso sobre el riesgo de recurrencia y el diagnóstico prenatal. El tiempo transcurrido hasta la devolución de los resultados osciló entre 2 y 58 semanas (media de 15,2 semanas), lo que refleja la disponibilidad limitada de la plataforma MiSeq para fines de investigación.

- Meng, et al (103) lo aplicaron a población pediátrica ingresada en unidades de cuidados intensivos, tanto UCIN como UCIP. Basado en el uso de un exoma clínico comercial en un solo probando (178 neonatos y lactantes

hasta 100 días de vida) o en tríos (100 neonatos y lactantes hasta 100 días de vida, de los cuales 63 eran críticamente enfermos –*critical trio CES*-). La evaluación de la necesidad de CES fue realizada por varios genetistas clínicos certificados por la junta del Texas Children's Hospital. De los 278 pacientes, 102 (36,7%) obtuvieron un diagnóstico molecular. El exoma trío crítico proporcionó diagnósticos moleculares significativamente más altos (50,8%) que el exoma probando (32,6%) y el exoma trío (32,4%). La mediana del tiempo de respuesta fue de 13,0 días para el *critical trio CES* menor que la del exoma probando (95,0 días) y el exoma trío (51,1 días). Aproximadamente el 38% de los individuos diagnosticados tenían una presentación infantil atípica o no reconocida de trastornos genéticos. Para los trastornos genéticos bien conocidos, como el síndrome de Kabuki causado por variante patógena en *KMT2D*, las características craneofaciales fueron atípicas o poco reconocidas en los cuatro bebés. Otro ejemplo de presentación atípica de trastornos mendelianos conocidos en recién nacidos incluyen el síndrome de megalencefalia-polimicrogiria-polidactilia-hidrocefalia relacionado con *AKT3* en un individuo con hipoglucemia, acidosis metabólica por hiperlactatemia y ventrículos laterales prominentes en el límite sin macrocefalia al nacer.

- Van Diemen, et al (114); estos autores secuenciaron el genoma completo pero en el análisis de las variantes, éstas se filtraron por CES. La enfermedad crítica se definió como insuficiencia cardiorrespiratoria que necesitaba asistencia respiratoria o disfunción orgánica grave (cerebro, corazón, pulmones, hígado o riñones). Los criterios de inclusión

fueron: edad <1 año y la presencia de una o más anomalías congénitas y / o síntomas neurológicos graves, como convulsiones intratables, que sugiriesen una causa genética de la enfermedad. Se incluyeron 23 niños en estado crítico menores de 12 meses ingresados en UCI en los que no se había podido realizar un diagnóstico rápido tras la evaluación clínica y las pruebas diagnósticas de rutina. El análisis dirigido de 3426 genes de enfermedades conocidas se realizó utilizando datos de secuenciación del genoma completo. Se obtuvo un diagnóstico genético en 7 pacientes (30%), con una mediana de tiempo de respuesta de 12 días. Se identificamos mutaciones heterocigotas compuestas en el gen *EPG5* (síndrome de Vici), *RMND1* (deficiencia combinada de fosforilación oxidativa-11) y el gen *EIF2B5* (sustancia blanca evanescente), y mutaciones homocigotas en el gen *KLHL41* (miopatía nemalínica), *GFER* (miopatía mitocondrial progresiva) y en *GLBI* (gangliosidosis GM1). Además, se detectó una microdelección 1p36.33p36.32 en un niño con miocardiopatía. En cuanto al impacto sobre el manejo clínico, estos diagnósticos reorientaron el tratamiento a cuidados paliativos en 5 de los 7 niños. Además, los diagnósticos específicos y el asesoramiento genético apropiado tuvieron un impacto en las decisiones que tomaron los padres para la futura descendencia.

- Brunelli, et al (104) se basaron en un panel diseñado por el mismo grupo (RapSeq, 4503 genes asociados a enfermedad) aplicado en tríos a neonatos con insuficiencia respiratoria y / o cardiovascular, encefalopatía, hipotonía profunda, malformaciones cerebrales complejas, alteraciones metabólicas graves o anomalías congénitas múltiples (sin un síndrome conocido), enfermedad inusualmente grave o

prolongada o insuficiencia multiorgánica. Se realizó un diagnóstico en 10/20 neonatos reclutados (50%); ocho de los cuales tenían variantes de novo (*ASXL1*, *CHD*, *FBN1*, *KMT2D*, *FANCB*, *FLNA*, *PAX3*). Los informes preliminares se generaron de media a los 9,6 días; informes finales después de la secuenciación de Sanger a los 16,3 días. En todos los casos positivos, el diagnóstico cambió la actitud terapéutica.

b) *Secuenciación de WES*. En este caso se busca secuenciar todas las regiones codificantes del genoma humano, que representa entre el 1% y el 1.5% del total. Los estudios piloto que han aplicado esta aproximación a pacientes neonatales críticamente enfermos incluyen los siguientes:

- Stark, et al (115, 116)-que lo aplicaron a población pediátrica ingresada en unidades de cuidados intensivos, tanto neonatal como pediátrica, en pacientes de 0 a 6 meses de edad. Los criterios de inclusión del estudio incluían la presencia de múltiples anomalías congénitas y características dismórficas u otras características fuertemente sugestivas de trastornos monogénicos, por ejemplo, afecciones neurometabólicas y displasias esqueléticas. El estudio pretendía estudiar la implementación de un protocolo basado en un test WES rápido (rWES) realizado únicamente al paciente índice. De 40 pacientes reclutados para el estudio, se alcanzó un diagnóstico genético en 21 (52.5%) con un tiempo medio hasta diagnóstico de 16 días. El manejo clínico cambió en 12 de los 21 pacientes diagnosticados (57%), incluyendo instaurar tratamiento para salvar la vida, evitar biopsias invasivas y la reorientación en algunos casos a cuidados

paliativos. En comparación con un estudio del mismo grupo basado también en el uso de WES en 80 pacientes pediátricos de 0 a 2 años (33 de los cuales estaban ingresados en la UCIN), pero un protocolo habitual (no rápido) (116), los resultados en cuanto a rendimiento diagnóstico y utilidad clínica fueron muy semejantes, pero difiriendo en el tiempo medio hasta diagnóstico (149 días frente a los 16 usando el rWES).

- Ceyhan-Birsoy, et al. (58) con su proyecto BabySeq, cuyos resultados han sido recientemente publicados es un ensayo clínico randomizado que pretende explorar el impacto médico, psicosocial y económico de la aplicación de la secuenciación genómica mediante WES en neonatos, tanto sanos, como ingresados en unidades de cuidados intensivos. 127 neonatos sanos y 32 enfermos fueron reclutados para este estudio. Con una selección restrictiva de las variantes informativas, variantes patogénicas o probablemente patogénicas en genes de alta penetrancia asociados a enfermedades clínicamente relevantes durante la infancia, o genes con evidencia o penetrancia moderada, pero asociadas a enfermedades muy graves, los autores identificaron variantes en 15 de los 159 neonatos (9.4%), 10 en sujetos sanos. Si bien el proyecto BabySeq ha ofrecido una gran cantidad de información en lo relacionado al manejo de la información genética obtenida, priorización de genes, información de portadores (117), percepción de los riesgos, beneficios y utilidad de la secuenciación genómica por parte de los padres de los pacientes y de los clínicos (118) e impacto psicosocial de la información genética sobre las familias (119), aspectos todos ellos a tener en cuenta a la hora de implementar un protocolo de esta naturaleza.

- Elliot, et al (120) con su estudio RAPIDOMICS. Se trata de un proyecto piloto de WES rápido que incluyó 25 recién nacidos con sospecha de trastornos genéticos ingresados en la UCIN del BC Women's Hospital. El tiempo medio hasta el diagnóstico preliminar fue de 7,2 días. Los criterios de selección incluían a neonatos de la UCIN con una o más de las siguientes presentaciones: convulsiones inexplicables, alteraciones metabólicas, anomalías neurológicas o disminución del nivel de conciencia, anomalías congénitas múltiples o alteración fisiológica de acuerdo con un trastorno genético (por ejemplo, retraso del crecimiento fetal). En 15 pacientes se detectaron variantes definitivamente o posiblemente causales del fenotipo de los recién nacidos. Estas variantes incluyeron 13 asociadas con herencia autosómica dominante en los genes *RAF1*, *KCNT1*, *CHD7*, *SMARCB1*, *EFTUD2*, *KCNQ2*, *SAMD9*, *COL4A1*, *PHOX2B*, *HRAS*, *ZMYND11*, *SLC12A5* y *KMT2D* y dos con herencia autosómica recesiva en los genes *AGK* y *BBS4*. Todas las variantes asociadas a condiciones autosómicas dominantes fueron de novo con la excepción de *PHOX2B*, donde se encontró en mosaico en un progenitor. Un total de 34 decisiones médicas asociadas tuvieron lugar en 15 de los 18 pacientes diagnosticados. La mayoría de estas decisiones se orientaron hacia la realización de pruebas complementarias para la detección de posibles complicaciones ocultas como deficiencias visuales o auditivas, alteraciones hormonales, malignidad, anomalías adicionales o inmunodeficiencia y el desarrollo de protocolos de seguimiento. El tiempo medio hasta el diagnóstico preliminar fue de 7,2 días.

- Gubbels, et al. (121, 122), en su estudio llevado a cabo en la Unidad de Cuidados Intensivos del Boston Children's Hospital se incluyeron a pacientes menores de 6 meses de edad con hipotonía, convulsiones, un fenotipo metabólico complejo y / o malformaciones congénitas múltiples se inscribieron prospectivamente para la secuenciación rápida del exoma (<7 días) basada en trío. Se obtuvo un diagnóstico genético en 29 de 50 (58%) casos secuenciados. Veintisiete (54%) pacientes recibieron un diagnóstico molecular que involucraba genes de enfermedades conocidas; dos casos adicionales (4%) se resolvieron con variantes patogénicas encontradas en genes de nuevas enfermedades. En 24 de los casos resueltos, el diagnóstico tuvo impacto en el manejo del paciente y / o familiares, incluyendo el cambio a cuidados paliativos, cambios en la medicación, participación de especialidades adicionales y la consideración de nuevas terapias experimentales. Como era de esperar, dado el entorno neonatal, muchos de los diagnósticos logrados a través de este estudio representan casos que se encuentran entre los más jóvenes reportados en la literatura. Los pacientes diagnosticados también tendían a tener fenotipos más graves de lo que se informa normalmente, especialmente para los trastornos metabólicos y neuromusculares. Para varios genes, los casos diagnosticados también representaron expansiones de fenotipos conocidos. Por ejemplo, un paciente con deficiencia de HIBCH presentó episodios de apnea y convulsiones en una etapa temprana de la vida, en comparación con presentaciones iniciales más sutiles (por ejemplo, retraso del crecimiento a los 3 meses o más) en otros casos notificados (123). Los diagnósticos fueron difíciles de alcanzar en algunos casos a pesar de los

fenotipos sorprendentes. un paciente por ejemplo, presentaba acidosis metabólica persistente, miocardiopatía hipertrófica y parentesco consanguíneo, pero no se pudo encontrar un diagnóstico definitivo. Estos pacientes podrían beneficiarse de WGS, secuenciación de ARN(124) y / o un futuro reanálisis del exoma (125). Estas técnicas podrían ofrecer nuevas vías interesantes para futuras investigaciones y uso clínico.

Lunke, et al (126). Realizaron un estudio multicéntrico en 12 hospitales australianos para evaluar prospectivamente el rendimiento de una red multicéntrica para el diagnóstico genómico ultrarrápido en un sistema público de salud. La población de estudio incluyó a 108 pacientes con una mediana de edad de 28 días (rango, 0 días a 17 años) y el 57% provenían de unidades de cuidados intensivos neonatales, el 33% de unidades de cuidados intensivos pediátricos y el 9% de otras salas del hospital. El tiempo medio desde la recepción de la muestra hasta el informe de secuenciación ultrarrápida del exoma fue de 3,3 días. Se estableció un diagnóstico molecular en 55 pacientes (51%). Once diagnósticos (20%) resultaron del uso de los siguientes enfoques para aumentar el análisis de secuenciación del exoma estándar: análisis de secuenciación del genoma mitocondrial, análisis del número de copias basado en la secuenciación del exoma, uso de bases de datos internacionales para identificar nuevas asociaciones entre genes y enfermedades, y análisis adicional de fenotipado y ARN. En 42 de 55 pacientes (76%) con diagnóstico molecular y 6 de 53 pacientes (11%) sin diagnóstico molecular, se consideró que el resultado de la secuenciación ultrarrápida del exoma influyó en el manejo clínico.

c) *Secuenciación de genoma completo* (WGS). Esta aproximación pretende determinar la secuencia completa de ADN de todo el genoma del paciente. Los estudios piloto que han aplicado esta aproximación a pacientes ingresados en UCIN incluyen los siguientes:

- Petrikin, et al. (127) que llevaron a cabo el proyecto NSIGHT1, un ensayo que busca evaluar las ventajas del uso de una aproximación WGS frente a estudios genéticos tradicionales. Se reclutaron 65 pacientes ingresados en UCIN, de los que 32 fueron estudiados mediante WGS y 33 fueron asignados al grupo control. Los fenotipos fueron muy diversos y típicamente estaban presentes al nacimiento, las indicaciones más comunes para la inclusión en el estudio fueron anomalías congénitas y trastornos neurológicos. Menos lactantes de control tuvieron hallazgos cardiovasculares que los casos, lo que podría haber afectado la probabilidad de enfermedad genética. La tasa diagnóstica en el grupo WGS fue del 41%, frente al 24% del grupo control. El objetivo del estudio era conseguir un diagnóstico genético en 28 días, objetivo que se consiguió en un 31% del grupo WGS frente a un 3% del grupo control. El impacto clínico a corto plazo de los diagnósticos se evaluó mediante revisiones de historias clínicas y encuestas a los médicos responsables de los pacientes. 20 (31%) de los 65 lactantes incluidos (95% de los que recibieron diagnósticos) tuvieron un cambio consecuente en el manejo. Diez (15%) neonatos (48% de los que recibieron el diagnóstico) tuvieron un cambio en el manejo que no fue el asesoramiento genético o reproductivo de los padres.
- French, et al. (122) realizaron WGS en tríos una cohorte prospectiva de familias reclutadas en los servicios de

pediatría (incluyendo UCIN y UCIP) de hospitales de Cambridge. Se desarrolló una línea de investigación en colaboración con el Servicio Nacional de Salud para entregar los resultados moleculares dentro de las 2 a 3 semanas posteriores al reclutamiento. En la UCIN, los criterios de exclusión fueron los niños ingresados para la vigilancia posparto de corta estancia, la prematuridad sin características adicionales, los bebés con antecedentes prenatales o de parto claros que sugieran una causa no genética y aquellos bebés en los que ya se realizó un diagnóstico genético. Los bebés con anomalías congénitas, síntomas neurológicos que incluyen convulsiones, sospecha de enfermedad metabólica, enterocolitis necrotizante quirúrgica, retraso extremo del crecimiento intrauterino y enfermedad crítica inexplicable de probable etiología genética fueron elegibles para el estudio. En la UCIP, los criterios de inclusión fueron los mismos que para la UCIN, pero los casos de traumatismo claro, malignidad y bronquiolitis o infecciones del tracto respiratorio en un niño sano no fueron elegibles. Se realizó un análisis del genoma completo de 195 familias (567 muestras) y el 21% recibió un diagnóstico molecular de la condición genética subyacente en el niño. La mayoría de los pacientes fueron reclutados en la UCIN (54%) y la UCIP (31%) y veintiocho probandos (15%) fueron derivados de neurología pediátrica o genética clínica. La descripción fenotípica del niño fue un mal predictor del gen identificado en el 90% de los casos, lo que aboga por las pruebas de genes agnósticos en la UCIN / UCIP. El diagnóstico afectó el manejo clínico en más del 65% de los casos (83% en recién nacidos), incluida la modificación de tratamientos y vías de atención y / o informando las decisiones de cuidados paliativos. Un plazo

de 2 a 3 semanas fue suficiente para afectar la mayoría de las decisiones clínicas.

- Wang, et al (128), con su estudio que se realizó utilizando una prueba de secuenciación optimizada del genoma trío (OTGS) en 130 pacientes pediátricos pertenecientes a hospital infantil terciario en China. Se reclutaron pacientes de la unidad de cuidados intensivos pediátricos o neonatales (UCIP/ NICU) con características clínicas complejas y sospecha de un trastorno genético. Las pruebas de OTGS identificaron variantes patogénicas en 62 de 130 individuos, lo que resultó en una tasa de diagnóstico del 47,7%. El tiempo de respuesta varió de 72 a 120 h. De los 62 lactantes con diagnóstico, 48 (77,4%) tenían variantes patogénicas de un solo nucleótido (SNV), 12 (19,4%) tenían variaciones patogénicas del número de copias (CNV) o variantes de estructura (SV) y 2 (3,2%) tenían pequeñas deleciones en un alelo más variantes patogénicas en otro alelo de genes autosómicos recesivos. Las estrategias terapéuticas para el 48,4% (30/62) de los pacientes diagnosticados fueron modificadas e incluyeron trasplante, recomendaciones dietéticas o cambio de fármacos, que evitaron morbilidades y mejoraron el pronóstico.
- Finalmente, debemos mencionar el estudio realizado por Kingsmore, et al (129) que llevaron a cabo el estudio NSIGHT2, un ensayo controlado aleatorio de la eficacia de la secuenciación rápida del genoma completo o del exoma (rWGS o rWES, respectivamente) y ultrarápida del genoma completo (urWGS) en neonatos gravemente enfermos, comparando su rendimiento analítico y diagnóstico. De 1.248 niños hospitalizados, 578 (46%) tenían enfermedades de etiología desconocida. Se inscribieron 213 recién nacidos

(37% de los elegibles), 24 de los cuales (11%) estaban muy enfermos y recibieron secuenciación ultrarrápida del genoma completo (urWGS). Los neonatos restantes fueron asignados al azar, 95 a rWES y 94 a rWGS. Si bien el rendimiento analítico de rWGS fue superior a rWES, incluidas las variantes que pueden afectar la función de las proteínas y las variantes patogénicas / probables de ClinVar, el rendimiento diagnóstico de rWGS y rWES fue similar (18 diagnósticos en 94 lactantes [19%] versus 19 diagnósticos en 95 lactantes [20%], respectivamente), al igual que el tiempo hasta el resultado (mediana 11,0 versus 11,2 días, respectivamente). Sin embargo, la proporción diagnosticada por urWGS (11 de 24 [46%]) fue mayor que rWES / rWGS y el tiempo hasta el resultado fue menor (mediana 4,6 días), probablemente por la selección de neonatos más graves dentro de los críticamente enfermos, lo que aumenta la posibilidad de que padezcan una enfermedad genética.

## 2. OBJETIVOS

### 2.1. OBJETIVO PRINCIPAL:

- a) Evaluar la utilidad diagnóstica del panel NeoSeq en el diagnóstico molecular de neonatos críticamente enfermos con sospecha de enfermedad genética.

### 2.2 OBJETIVOS SECUNDARIOS:

- a) Comparar el rendimiento diagnóstico de NeoSeq con otros abordajes de secuenciación masiva aplicados a neonatos críticamente enfermos con sospecha de enfermedad genética.
- b) Aumentar el conocimiento de las bases moleculares de las enfermedades genéticas en neonatos críticamente enfermos.
- c) Evaluar la utilidad de escalas de valoración como Phenomizer para sospecha genética a través del fenotipo descrito mediante términos HPO

### 2.3. OBJETIVO FINAL:

Implementar el uso del panel NeoSeq en el diagnóstico de las enfermedades genéticas en neonatos críticamente enfermos.



## 3. MATERIAL Y MÉTODOS

La presente tesis ha sido parcialmente publicada en *Journal of Clinical Medicine*. 2020; 9(8):2362.

<https://doi.org/10.3390/jcm9082362>

Factor de impacto: 3.303

Categoría y cuartil JCR: 36/165 (Q1) en 'Medicine, General & Internal'

<https://www.mdpi.com/2077-0383/9/8/2362>

### 3.1. DISEÑO DEL ESTUDIO

Estudio prospectivo multicéntrico que incluye a pacientes consecutivos ingresados en UCIN de nivel IIIB / C pertenecientes al Hospital Clínico Universitario de Santiago de Compostela, el Hospital Regional de Málaga y el Hospital Clínico Universitario Miguel Servet durante un período de 24 meses.

### 3.2. PACIENTES

Los pacientes críticamente enfermos de menos de dos meses de edad que cumplieran con uno o más de los siguientes criterios se consideraron elegibles para la inclusión: (i) Malformaciones congénitas que no están claramente relacionadas con un síndrome genético clínicamente identificable, (ii) descompensación metabólica

no asociada con parámetros bioquímicos. y / o cribado neonatal que indique una sospecha de enfermedad metabólica hereditaria, (iii) epilepsia o enfermedades del neurodesarrollo de probable origen genético. Se excluyó del estudio a los neonatos con indicios claros de un síndrome específico que pudieran diagnosticarse mediante análisis dirigido de genes conocidos o variaciones estructurales.

En todos los casos, previo al análisis genético, se obtuvo el consentimiento informado de los tutores legales del paciente para realizar el estudio genético. Se recogió la información clínica relevante de la historia clínica para la interpretación de los resultados genéticos. Para su ejecución, este proyecto fue sometido a la aprobación del Comité Autónomo de Ética de la Investigación Clínica de Galicia (código 2018/366). Este estudio respeta los principios fundamentales establecidos en la Declaración de Helsinki de la Asociación Médica Mundial, en el Convenio del Consejo de Europa relacionado con los derechos humanos y la biomedicina y cumple los requisitos establecidos en la legislación española en el campo de la investigación biomédica, la protección de datos personales y bioética.

### **3.3 VARIABLES DEL ESTUDIO**

Para cada paciente se evaluaron las siguientes variables: antecedentes familiares, datos de consanguinidad, antecedentes obstétricos maternos, sexo, parámetros antropométricos, edad de inicio de los síntomas, síntomas y signos clínicos presentados, parámetros bioquímicos de las pruebas de laboratorio, variables de imagen y pruebas adicionales realizadas para identificar la enfermedad de base, tratamientos administrados (farmacológicos, nutricionales, de soporte respiratorio, diálisis u otras medidas invasivas), análisis genético del paciente y de sus padres.

Todos los probandos fueron fenotipados utilizando términos HPO extraídos manualmente de los registros de salud electrónicos. Posteriormente se generó una lista de posibles diagnósticos diferenciales basada en los hallazgos clínicos obtenidos utilizando la aplicación de diagnóstico clínico Phenomizer ([compbio.charite.de/phenomizer/](http://compbio.charite.de/phenomizer/)). Al obtener los resultados del estudio genético, se correlacionaron las variables clínicas y moleculares.

### 3.4 DISEÑO DEL PANEL NEOSEQ

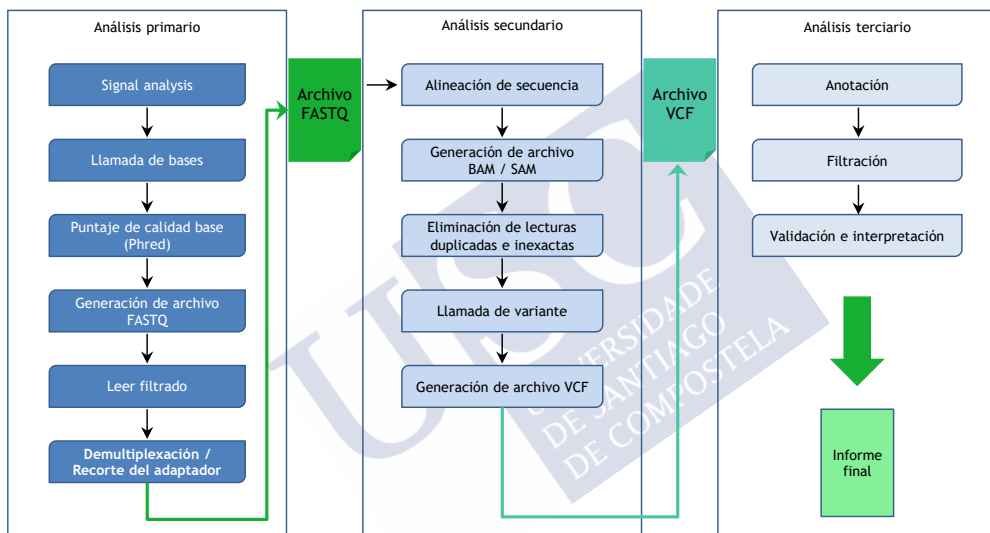
Se creó un panel de genes específico, llamado NeoSeq, que consta de 1.870 genes asociados con trastornos neurológicos o metabólicos y malformaciones congénitas de inicio temprano (durante los dos primeros meses de vida). El panel NeoSeq se diseñó utilizando la base de datos OMIM (<https://omim.org/>) como principal fuente de información y tras la revisión de la literatura científica. La cobertura promedio alcanzada es de 200-300X. El porcentaje medio de bases con cobertura <10X por muestra analizada es del 0,28% (28 de 10.000 pb se encuentran con cobertura <10X o insuficiente). Estos defectos de cobertura son causados por problemas intrínsecos al método de enriquecimiento utilizado. Sin embargo, este cálculo se realiza para lecturas de calidad superior a 20 (Q20 - Qscore, estima la probabilidad de un error en la identificación de una base de nucleótidos basada en un modelo de calidad, puntuación de calidad Phred). Para un Q20, la probabilidad de factor de error de base es de 1 en 100, lo que implica mayor confiabilidad de bases con cobertura insuficiente.

### 3.5 SECUENCIACIÓN NGS

El estudio genético se realiza con la creación de un pipeline de trabajo que combina métodos de secuenciación y análisis

bioinformático para la identificación y anotación de variantes con significación biológica relevante (Figura 5).

En esta sección se describe el proceso de secuenciación de diferentes regiones del genoma humano para muestras de ADN, utilizando la tecnología de secuenciación de Illumina ([www.illumina.com](http://www.illumina.com)). Las regiones a secuenciar fueron definidas y diseñadas internamente y proporcionadas por Agilent.



**Figura 5. Flujo de trabajo para la secuenciación y análisis de diferentes regiones del genoma humano.** Se describen los diferentes pasos del análisis primario, secundario y terciario que conducen a la elaboración final del informe de resultados de un estudio genético basado en NGS. *VCF: variant call format.*

La tecnología de secuenciación de Illumina se basa en la amplificación clonal de moléculas de ADN individuales. En este proceso de secuenciación de síntesis, se utilizan nucleótidos marcados con fluorescencia con terminación reversible. La confiabilidad de la incorporación de nucleótidos es muy alta, ya que en cada ciclo de secuenciación existe una competencia entre las 4 bases. Para cada biblioteca (línea de secuenciación), se obtienen varios millones de

fragmentos secuenciados (lecturas), lo que puede garantizar una alta cobertura en cada región y fortalecer la calidad de las secuencias consenso generadas. Cada ejecución de secuenciación permite el análisis de una celda de flujo (soporte físico de las muestras en el proceso de secuenciación). Para la captura de los fragmentos de ADN con las secuencias de interés se utilizó el sistema de hibridación SureSelect de Agilent Technologies (compatible con los protocolos de secuenciación de Illumina). Se diseñaron sondas específicas de las regiones de secuencia de interés para cada panel de genes. Estas sondas incluían exones y límites intrón-exón de un grupo de genes preseleccionado, incluidos todos los transcritos descritos para cada uno de ellos. La herramienta de diseño de sonda basada en web Sure Design fue la que se utilizó para este propósito (<https://earray.chem.agilent.com/suredesign/>). Para garantizar una representación óptima de todas las regiones de interés, se diseñaron diferentes subgrupos de sondas según el contenido porcentual de GC y la presencia de secuencias repetitivas.

El protocolo de secuenciación se realizó siguiendo los protocolos aprobados y utilizando el equipo y los reactivos recomendados oficialmente por Illumina.

### **3.5.1 Recepción de muestras y extracción de ADN**

Las muestras de sangre de los pacientes y sus progenitores se recogieron previo consentimiento informado en un tubo anticoagulado con ácido etilendiaminotetraacético (EDTA) en cada una de las UCINs participantes y fueron enviados a la Unidad de Diagnóstico y Tratamiento de Enfermedades Metabólicas Congénitas del Hospital Clínico de Santiago de Compostela junto con los datos clínicos anteriormente descritos.. A continuación, se extrajo el ADN genómico utilizando el kit comercial de purificación de ADN genómico Wizard (Promega), siguiendo las instrucciones del fabricante. La

concentración del ADN obtenido se cuantificó utilizando el Espectrofotómetro NanoDrop 1000 (Nanodrop Technologies) y el fluorímetro Qubit®. Además, la integridad del ADN se comprobó mediante electroforesis en gel de agarosa al 0,8%. En todos los casos se obtuvieron al menos 3µg de ADN completo, libre de impurezas ( $Abs_{260} / 280 > 1.8$  y  $Abs_{260} / 230 > 1.5$ ).

### **3.5.2 Librerías de ADN con captura de secuencia**

Una vez comprobadas las muestras de ADN genómico de los pacientes, se prepararon las bibliotecas a secuenciar. El kit SureSelect (Agilent) utilizado fue: kit de reactivos SureSelectXT, HSQ y SureSelect Capture Library Custom 2,9 Mb (16 reacciones). El kit de preparación de muestras de ADN TrueSeq (Illumina) se utilizó para el análisis del panel NeuroExome. Se denomina biblioteca al conjunto de muestras que se secuenciaron juntas. Se elaboró un documento de Excel que especifica las muestras que se incluyeron en cada una de las bibliotecas, así como la biblioteca de captura utilizada para cada muestra. Para la preparación de las bibliotecas, se siguió el protocolo indicado por Agilent, SureSelectXT Target Enrichment System for Illumina Paired-End Sequencing Library (Versión 1.5, noviembre de 2012), utilizando el material adicional recomendado.

El ADN genómico se fragmentó mediante sonicación, utilizando un proceso mecánico isotérmico de Covaris (Covaris, Inc) para el enriquecimiento de fragmentos en un rango de 150-200 pb. Para el control de calidad, las muestras se verificaron utilizando el sistema Agilent 2100 Bioanalyzer, una solución de electroforesis automatizada establecida para el control de calidad de las muestras de biomoléculas. Las muestras que mostraban una fragmentación insuficiente se sonicaron de nuevo en las mismas condiciones pero durante tiempos más breves (entre 10 y 25 minutos según el electroferograma obtenido por el Bioanalizador).

La purificación de las muestras fragmentadas se realizó mediante la utilización de microesferas Agencourt AMPure XP en un proceso basado en la tecnología de esferas magnéticas.

El control de calidad de las muestras fue realizado nuevamente por el Bioanalizador Agilent 2100, y luego los extremos de los fragmentos generados fueron reparados por una ADN polimerasa T4, Klenow, T4 PNK. La purificación del ADN reparado se comprobó de nuevo mediante perlas Agencourt AMPure XP. El paso de purificación se repitió entre varios pasos del protocolo para garantizar la calidad del ADN durante todo el proceso de secuenciación.

Una vez que se repararon los extremos del fragmento de ADN y se completó el proceso de purificación, se agregó dATP a los extremos 3' de los fragmentos de ADN. Luego se realizó la ligadura de los adaptadores específicos a los fragmentos de la plataforma de secuenciación de Illumina y posteriormente, se realizó la hibridación de las bibliotecas con una solución de captura (para capturar oligonucleótidos), según el protocolo Agilent SureSelect. De cada muestra se utilizaron 750 ng en 3,4ul, pero como las muestras estaban más diluidas en este paso fue necesario medir el volumen adecuado para la cantidad de 750ng, la evaporación de la muestra se realizó en un Speedvac para posteriormente traer las muestras a 34ul. El tiempo de hibridación fue de alrededor de 20-23 horas y la recuperación de las regiones capturadas se realizó utilizando las esferas magnéticas Dynal MyOne Streptavidin T1.

Se aplicó una PCR de 14 ciclos a la biblioteca capturada utilizando la ADN polimerasa de Fussion Herculase II. Los índices que identificarán cada muestra dentro del grupo se incluyeron en este paso de amplificación.

### **3.5.3 Dilución de muestras, agrupación y preparación de bibliotecas para secuenciación**

Las muestras se diluyeron a 10 nM y se prepararon conjuntos que incluían 5 µl de cada muestra. Estos conjuntos de muestras son las bibliotecas y cada una de ellas se secuenció en un secuenciador MiSeq. Las bibliotecas se desnaturalizaron con NaOH y luego se diluyeron para su posterior incorporación en la celda de flujo.

### **3.5.4 Secuenciación en la plataforma MiSeq (Illumina)**

En la plataforma MiSeq, el proceso de secuenciación se realiza y se completa. Primero se generaron millones de grupos de secuencias de ADN (representativos de los fragmentos de ADN que componen cada biblioteca) en la superficie de la celda de flujo. Después, las moléculas de ADN funcionaron como amplicones de secuenciación y en cada ciclo se produce la incorporación de un nucleótido siendo registrada la base incorporada. Se utilizaron los reactivos correspondientes al Illumina MiSeq® Reagent Kit v2 (300 ciclos). Se realizó una serie de secuenciación de 75 ciclos de “final pareado”, con 6 ciclos para la secuenciación de cada índice.

## **3.6 ANÁLISIS BIOINFORMÁTICO**

Se analizan los resultados de secuenciación masiva obtenidos de las muestras procesadas con MiSeq.

El objetivo principal de este análisis es leer los fragmentos de 75 bases de las lecturas generadas por el secuenciador y alinearlos con la secuencia del genoma humano de referencia, versión 'hg19'. Posteriormente, se procesarán los datos para detectar posibles variantes genéticas en las regiones de captura definidas. Para lograr esto, el análisis computacional siguió estos pasos:

- a) Análisis de imágenes y transformación de intensidades de fluorescencia en secuencias ("Llamada de base"); Eliminación de adaptadores de secuencia de Illumina, demultiplexación de muestras, generación de archivos FASTQ utilizando algoritmos de Illumina. Este paso lo realiza el software Real Time Analysis (RTA), incorporado en la plataforma MiSeq.
- b) Control de calidad de los datos obtenidos con el programa FastQC.
- c) Alineación de los fragmentos secuenciados con la secuencia de referencia con el software BWA.
- d) Detección de variantes genéticas y anotación, utilizando Samtools y Annotvar.

### 3.6.1 Análisis primario

#### 3.6.1.1 Análisis de imágenes y llamadas de base

El programa Real Time Analysis (RTA) v.1.8.70 (Illumina) realiza los primeros pasos del análisis de secuenciación, que comúnmente se denomina análisis primario. En cada ciclo de secuenciación, el secuenciador produce una foto para cada una de las áreas que componen la celda de flujo, y en cada una de estas áreas se tomaron fotos con los 4 filtros fluoróforos específicos para cada nucleótido.

A partir de estos archivos, RTA ubica los grupos generados y extrae sus intensidades, sus coordenadas y una estimación del ruido de fondo para cada grupo. El resultado del análisis de imágenes son los datos de entrada para la asignación de las bases. RTA permite transformar las señales de intensidad en secuencia de ADN.

Dependiendo de la eficiencia de los flujos de reactivos y de las reacciones de secuenciación, se puede incorporar una pequeña cantidad de moléculas para acelerar (avanzadas) o retrasar su ciclo de incorporación, lo que se conoce como pasos de prefase y escalonamiento. Este efecto se puede corregir durante el paso de asignación de base. Como la estimación estadística para el paso de corrección se basa en la estandarización de numerosos grupos para evaluar la correlación de la señal en diferentes ciclos, esto es más confiable para mosaicos con un gran número de grupos y una mezcla de diferentes secuencias. Las muestras con sólo un número "pequeño" de secuencias diferentes no producen estimaciones tan fiables. Al secuenciar una región de 0.5-1.5 Mb (dependiendo de la biblioteca de captura utilizada), el número de secuencias diferentes puede ser menor que un experimento de secuenciación del genoma completo, por lo que esta corrección es importante.

El resultado es un archivo de texto en formato "FASTQ" que contiene la secuencia de cada uno de los fragmentos secuenciados e información sobre su calidad.

### 3.6.1.2 Control de calidad

Para confirmar la calidad de los datos generados, se utilizó el programa FastQC v0.10.1 (<http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc>). Este programa permite comprobar la composición y la calidad de la secuenciación. La calidad de los datos de secuenciación se mide utilizando la medida Phred Score, empleada de forma rutinaria en la metodología de secuenciación de Sanger, e indica la confiabilidad de la base asignada en el proceso de secuenciación. Los datos de calidad generalmente se codifican en formato ASCII y se encuentran en los archivos originales de Illumina en formato .FASTQ.

## 3.6.2 Análisis secundario

### 3.6.2.1 Alineamiento de secuencias

La alineación se realizó con el programa BWA (130) con los parámetros predeterminados. Para detectar posibles problemas de contaminación y alineación con regiones de secuencias similares en otras secciones del genoma (duplicaciones segmentarias, etc.), los fragmentos generados se alinean contra todo el genoma humano (versión hg19). De esta forma podemos seleccionar aquellos fragmentos que se alinean en una única posición genómica y en el área de interés. Las áreas de interés comprenden las regiones capturadas especificadas anteriormente.

A partir de los datos de alineación y las coordenadas genómicas de captura (datos incluidos en el archivo target.bed del diseño experimental), se utilizó el programa NGSrich (<http://ngsrich.sourceforge.net>) para evaluar la eficiencia de captura.

Dado que no se realizó ningún filtrado en función de la cobertura obtenida, existe el riesgo de falsos positivos y / o falsos negativos en las regiones cubiertas con menos de 20X. Para cada muestra se generó un archivo, denominado “GRXXX\_lowCoveragePositions.txt” (donde XXX corresponde al número específico de cada muestra), que incluye aquellas regiones con cobertura menor a 20X en cada una de las muestras analizadas. En estas regiones la probabilidad de no detectar ninguna variante es mucho mayor que en el resto.

Después de la alineación, se eliminan aquellos fragmentos que se han alineado con una calidad menor que un valor de Phred de 20 unidades, lo que correspondería a una probabilidad de error de alineación de una base en 100. Estas lecturas de alineación de alta calidad son las que se utilizan para la fase de identificación de

variantes. Este paso reduce el número de falsos positivos en los resultados.

Los datos filtrados por calidad de alineación se guardaron en los archivos “GRXXX\_covered\_QCalign.bam” y “GRXXX\_covered\_QCalign.bam.bai”.

### 3.6.2.2 Llamada de variante

El proceso de detección de variantes, como sustituciones e INDELS (pequeñas inserciones-delecciones), incluidas en las regiones de captura se realizó con el programa Samtools 0.1.19 (130) y el programa VarScan (131) programas de uso común en la literatura para la detección de variantes.

A partir de los datos de alineación obtenidos, los programas de detección de variantes establecen la existencia de una variante en una posición al encontrar en la muestra bases diferentes a las observadas en la secuencia de referencia. La evaluación final de una variante dependerá del número de lecturas que respalden la existencia de esa variante, así como de la calidad de la alineación y la llamada. En este caso, para la detección de variantes, se utilizaron los parámetros predeterminados de Samtools a excepción del límite de calidad de alineación de lectura de 20 unidades Phred. Asignar una variante requiere suficiente confiabilidad de la base asignada de los fragmentos que se alinean en esa posición. Esta confiabilidad estuvo determinada principalmente por 2 factores: por un lado, el valor QUAL que nos informa sobre la calidad Phred para la base alternativa propuesta y, por otro lado, el valor DP4 que reporta el número de fragmentos de alta calidad en los detectados por base como variante tanto en la secuencia directa como en la inversa. En general, para considerar una detección precisa de la variante, se recomienda obtener un valor de QUAL (calidad Phred de la base alternativa propuesta) mayor o igual a 20. Asimismo, para asignar una variante, se recomienda que el

número del total de lecturas encontradas en esa posición es suficiente para una confiabilidad de cobertura aceptable.

Para cada muestra se generó un archivo vcf con todas las variantes encontradas, denominado “GRXXX.vcf” para cada muestra, el cual se guardó en la carpeta Resultados. Además, se elaboró un archivo de Excel denominado “GRXXX.xlsx” y en cada una hay dos hojas, una para SNP y otra para INDEL. En cada archivo de Excel, los resultados se muestran de la siguiente manera:

- CROMO: Cromosoma en el que se encuentra la variante.
- POS: Posición cromosómica en la que se encuentra la variante.
- REF: Base en la secuencia de referencia.
- ALT: base alternativa en la muestra después de la alineación.
- QUAL: Valor Phred de calidad para la base propuesta en ALT. Cuanto mayor sea el valor, más confiable será la base establecida en ALT.
- DP: Cobertura total de crudo.
- DP4: Número de lecturas que se alinean con la base de referencia en dirección de avance (ref F), en dirección inversa (ref R) y número de lecturas que se alinean con la base alternativa en dirección de avance (alt F) y en dirección inversa (alt R).
- MQ: cuadrático medio de la calidad de alineación de las lecturas que cubren la región.
- GT: tipo de genotipo, 0 para el alelo de referencia y 1 para el alelo alternativo.
- GQ: Calidad del genotipo dada en la columna GT, en escala Phred.

En la hoja INDELS, también hay una columna adicional:

- IS: número máximo de lecturas que admiten un INDEL y la fracción del total de lecturas.

### 3.6.3 Análisis terciario

#### 3.6.3.1 Identificación y anotación de las variantes

El conjunto de las variantes obtenidas anteriormente se anotó utilizando el programa Annovar (132) y el programa VarScan (131) para la anotación contra dbSNP (base de datos de polimorfismo de nucleótido único) y RefSeq (colección de secuencia de referencia). Además, los resultados se anotaron contra la base de datos del proyecto 1000 Genomes (133) la base de datos ESP (134) y GnomAD (“GnomAD Resource Introduced at ASHG Meeting, Double ExAC Conjunto de datos”, nd). En este caso, para la población general se utilizó la versión de la base de datos 2012\_04, mientras que para la población europea se utilizó la versión fase 1\_release3. Se generó un archivo VCF que contiene metainformación en líneas, una línea de encabezado, seguida de los datos, cada uno de los cuales contiene información sobre la posición del cromosoma, la base de referencia, la base o bases alternativas identificadas. Los resultados se muestran de la siguiente manera:

- Variant ID: Código que identifica la variante en la base de datos dbSNP 137 SNP.
- Región y gen: indica si la variante afecta a exones, intrones, regiones intergénicas o áreas de ARN no codificante. Si el área es un exón, intrón o región no codificante, Annovar indica en la segunda columna el nombre del gen o genes afectados. Por el contrario, si la variante no afecta a ninguna de estas regiones, Annovar indica los nombres de los dos genes vecinos

junto con la distancia de esos genes. Este parámetro se utiliza cuando una variante se ajusta a varias categorías. El valor de la primera columna tomó el siguiente orden de preferencia: exonic = splicing> ncRNA> UTR5 / UTR3> intron> upstream / downstream> intergénico. Cuando una variante se clasificó como exónica, se refiere solo a la parte codificante, pero no a la región UTR. El empalme se definió como una variante que tiene 1 o 2 pares de bases de un límite exón / intrón. Si una variante estaba ubicada en la región 5'UTR y 3'UTR, el resultado se mostraba como "UTR5, UTR3", respectivamente. Los términos "aguas arriba / aguas abajo" se definen como una distancia de 1 kb desde el sitio de inicio o final de la transcripción. Si una variante se ubicó tanto en sentido ascendente como descendente (debido a la existencia de dos genes diferentes), entonces aparecerá "ascendente / descendente" en la primera columna.

- Consecuencia: Identifica los cambios presentados como consecuencia de las variantes exónicas. Solo se anotaron las variantes exónicas, por lo que las variantes restantes no tenían esta información incluida en esta columna. Los valores posibles para este campo incluyen: SNV no sinónimo, SNV sinónimo, inserción de desplazamiento de fotograma, eliminación de desplazamiento de fotograma, inserción sin desplazamiento de fotograma, eliminación sin desplazamiento de fotograma, sustitución de bloque de desplazamiento de fotograma, sustitución de bloque sin desplazamiento de fotograma.
- Información adicional: información adicional sobre la variante.
- MAF\_1000g\_ALL: frecuencia del alelo menor en la base de datos general del proyecto 1000 Genomes

- MAF\_1000g\_EUR: frecuencia del alelo

### 3.6.3.1 Filtrado y priorización de las variantes

La priorización de las variantes identificadas en los pasos descritos anteriormente, y su clasificación como patógena (P), probablemente patógena (LP), variante de significado incierto (VUS), probable benigna (LB) o benigna (B), se realizó con base en los criterios del ACMG (79) y su registro en la base de datos ClinVar (<https://www.ncbi.nlm.nih.gov/clinvar/>).

La frecuencia de las variantes ha sido confirmada en bases de datos poblacionales como 1000Genomes, ExAC, gnomAD y en nuestra base de datos interna creada con el almacenamiento y registro de las variantes identificadas en nuestra población. Las variantes se descartaron si presentaban un minor allele frequency (MAF) > 0,01.

La posición de la variante identificada ha sido evaluada según diferentes criterios:

- a) Si se encuentra en una región exónica: valorar la posible consecuencia funcional de la variante en el siguiente orden de patogenicidad:
  1. Variantes que eliminan o duplican uno o más exones de un gen, o incluso el gen completo (CNV).
  2. Variantes truncadas: pequeñas deleciones o inserciones que cambian el patrón de lectura (cambio de marco), o variantes que provocan el cambio de un aminoácido a un codón de parada (parada).
  3. Pequeñas eliminaciones o inserciones que no cambian el patrón de lectura (sin desplazamiento de fotogramas).

4. Variantes que cambian el aminoácido (missense). Para predecir el efecto de las variantes de codificación no sinónimas se utilizó la herramienta CONsensus DELeterioriousness (CONDEL) (79) donde las variantes de missense se clasificaron como Deletéreas (D) o Neutrales (N) según el impacto funcional que podrían tener en las proteínas resultantes.
5. Variantes que no cambian el aminoácido ni los sinónimos pero que podrían influir en el empalme y el empalme del ARNm.
  - b) Si la variante está ubicada en una región intrónica: se evaluó la posición con respecto a las secuencias consenso de empalme.

El impacto de las variantes intrónicas o sinónimas en el *splicing* de ARN se evaluó utilizando el software Human Splicing Finder (HSF) (135). Utilizando esta herramienta, se puede predecir el efecto de las variantes en los sitios de empalme, incluido el sitio de empalme 5' o donante (dinucleótido AG inmediatamente antes del exón), el sitio de empalme 3' o sitio aceptor (dinucleótido GT inmediatamente después del exón), el punto de ramificación (ubicado en 18-40 nucleótidos cadena arriba del sitio aceptor y otras secuencias reguladoras a lo largo del pre-ARNm, como: secuencias exónicas que estimulan el empalme (ESE - Exonic Splicing Enhancer) y secuencias exónicas que reprimen el empalme (ESS - Exonic Splicing Silenciador).

El grado de patogenicidad de cada variante se analizó mediante las herramientas DANN (136) (el rango de valores es de 0 a 1, atribuyéndose 1 a las variantes que se predice que serán más dañinas) y Mutation Taster (137) que predice el potencial de la enfermedad mediante un algoritmo clasificador de Bayes.

Para calcular el grado de conservación de una base de nucleótidos contra una variante posiblemente patógena, se utilizó la herramienta GERP (Genomic Evolutionary Rate Profiling), que predice la conservación cuantificando los déficits de sustitución en múltiples alineaciones ortólogas utilizando los genomas de 35 mamíferos. Su rango es de -12,3 a 6,17, siendo 6,17 el más conservado (138).

### 3.6.4 Informe de variantes

En este estudio se consideró un diagnóstico positivo en los casos en que:

- a) Una variante de novo (P o LP) en un gen dominante cuyo fenotipo asociado coincide con el del paciente.
- b) Dos variantes trans (P o LP) en un gen recesivo cuyo fenotipo asociado coincide con el del paciente.
- c) Variantes hemicigotas clasificadas como patógenas.

E consideró un diagnóstico no concluyente cuando:

- a) Un VUS en un dominante o dos VUS en un gen recesivo cuyo fenotipo asociado coincide con el del paciente.
- b) Cuando la no realización de pruebas en familiares no permitió aceptar o rechazar la hipótesis de asociación positiva con la patología.
- c) Cuando la información clínica disponible del paciente no fue suficiente para establecer un vínculo directo entre genotipo y fenotipo.
- d) Cuando estuvieran pendientes los estudios de expresión para confirmar o descartar la patogenicidad de una variante.

- e) Cuando las variantes se encontraron en genes dominantes, pero también se encontraron en individuos no afectados. Podría existir la probabilidad de una penetrancia incompleta de la variante y que dependiera de la concurrencia de otras variantes en genes relacionados, o existe la probabilidad de mosaicismo en el progenitor de quien la variante ha heredado, lo que explicaría la falta de expresión fenotípica en él, sin embargo, no fue posible demostrarlo durante este trabajo.
- d) Cuando se identifica una variante (P o LP) en un gen recesivo cuyo fenotipo asociado coincide con el del paciente, pero no se detectó una segunda variante.

### **3.7. VALORACIÓN CLÍNICA DEL RESULTADO GENÉTICO CONSIDERADO POSITIVO**

Se evaluó el impacto del diagnóstico genético en cuatro aspectos distintos del tratamiento médico durante un período de tres meses mediante la actualización de la historia clínica con el médico de referencia del paciente:

- (a) Redireccionamiento de la atención (hacia la retirada de tratamientos invasivos o fútiles y el redireccionamiento de cuidados intensivos a cuidados paliativos),
- (b) Inicio de nuevas valoraciones por parte de distintos subespecialistas, incluyendo estudios adicionales de extensión y seguimiento con sus pertinentes pruebas complementarias.
- (c) Cambios en la medicación o la dieta, incluyendo tanto tratamientos curativos como sintomáticos.
- (d) Asesoramiento genético a ambos progenitores incluyendo cómo la herencia contribuye a la patología de su hijo, el

riesgo de ocurrencia o recurrencia y las alternativas para disminuir el mismo en futuros embarazos.



## 4. RESULTADOS

### 4.1. CARACTERÍSTICAS DE LOS PACIENTES A ESTUDIO

Entre el 1 de enero de 2018 y el 31 de diciembre de 2019 ingresaron en la UCIN de los centros hospitalarios participantes en el estudio 51 recién nacidos con sospecha de enfermedad genética. De los recién nacidos identificados, 10 de ellos fueron excluidos debido a indicaciones claras de un síndrome específico que podría probarse mediante un análisis dirigido de genes conocidos o variaciones estructurales. Se contactó con un total de 41 familias y se les ofreció un análisis genómico con NeoSeq. De ellos, 37 familias consintieron en unirse al estudio. De los pacientes reclutados, cuatro recién nacidos no fueron secuenciados porque el estudio genético de los padres no estaba disponible. Finalmente, 33 recién nacidos (16 mujeres y 17 hombres) cumplieron nuestros criterios de inclusión y fueron incluidos en el estudio (Tabla 26).

La mediana de edad en el momento de la inclusión fue de 26 días (rango, cuatro días a dos meses). Dos pacientes fallecieron antes de la conclusión del estudio, uno antes del diagnóstico y el otro después de que se estableció el diagnóstico y se iniciaron los cuidados paliativos.

La mediana del número de términos HPO empleada por probando fue de cuatro, teniendo el 58% de los pacientes al menos cuatro términos. Sobre la base de los síntomas clínicos predominantes, los pacientes se clasificaron en los siguientes 4 grupos:

- Anomalías congénitas (n = 8)

- Síntomatología neurológica, incluida encefalopatía, regresión del desarrollo, convulsiones e hipotonía (n = 17),
- Sospecha de enfermedades metabólicas hereditarias (n = 7)
- Retraso grave del crecimiento intrauterino (n = 1).

Para cada uno de los 33 pacientes incluidos en este estudio, los datos demográficos, la presentación clínica y los fenotipos estandarizados que utilizan términos HPO se muestran en la Tabla 26.

#### **4.2. ESTUDIOS GENÉTICOS. TIEMPO DE RESPUESTA Y TASA DIAGNÓSTICA DEL PANEL NEOSEQ**

La mediana del tiempo de respuesta al diagnóstico fue de 7.5 días (rango, 4 a 11 días). Se alcanzó un diagnóstico genético para el 36.4% (n = 12) de nuestros 33 pacientes (Tabla 3). De los cuatro grupos establecidos previamente el rendimiento diagnóstico fue el siguiente:

- Anomalías congénitas, 3 casos de los 8 pacientes incluidos, lo que supone un 23%.
- Sintomatología neurológica, 6 casos de los 17 pacientes incluidos, lo que supone un 35%.
- Sospecha de enfermedades metabólicas hereditarias, 3 casos de los 7 pacientes incluidos, lo que supone un 43%.

Aunque las enfermedades metabólicas y las enfermedades neurológicas tuvieron aproximadamente el doble del rendimiento diagnóstico que las anomalías congénitas, no observamos diferencias significativas en las tasas de diagnóstico entre los grupos ( $p = 0.572$ ). El análisis de los términos HPO mediante la herramienta Phenomizer correlacionó el fenotipo con el diagnóstico molecular obtenido de forma estadísticamente significativa ( $p < 0.05$ ) en el 41.6% de los pacientes.

#### 4.2.1 Pacientes con variantes definitivamente/probablemente causales

**CASO 1.** Se trata de un recién nacido de sexo masculino reclutado a los 25 días de vida. Entre sus antecedentes personales destacan un retraso de crecimiento intrauterino evidenciado desde el segundo trimestre gestación. A la exploración física al nacimiento se evidencia también criptorquidia unilateral y microcefalia. Durante los primeros días de vida presenta dificultad respiratoria que requiere soporte respiratorio con ventilación mecánica no invasiva así como tendencia a la hipoglucemia con acidosis metabólica e hiperlactacidemia, precisando perfusión de suero glucosado por vía intravenosa. Ante estos hallazgos se realizó estudio genético mediante panel NeoSeq trio cuyos resultados se obtuvieron en los 7 días posteriores. Se identificaron las variantes c.640A>T/c.1199G>C en heterocigosis compuesta en el gen *GFMI* (MIM \*606639) (Tabla 4), el cual se encuentra localizado en el cromosoma 3 (3q25.32). Las variantes patogénicas en el gen *GFMI* están asociados a la deficiencia combinada de fosforilación oxidativa 1 (MIM #609060). Se trata de un trastorno multisistémico autosómico recesivo con manifestaciones variables que resultan de un defecto en el sistema de fosforilación oxidativa mitocondrial (OXPHOS). El inicio de la sintomatología clínica se produce en el nacimiento o poco después y las características pueden incluir retraso del crecimiento, microcefalia, hipertonía, hipotonía axial, encefalopatía, miocardiopatía y disfunción hepática. Nuestro paciente presenta un fenotipo compatible con el cuadro clínico descrito, evidenciado mediante el análisis por Phenomizer de forma estadísticamente significativa (p 0.0343). El estudio en trio permitió confirmar

la herencia autosómica recesiva siendo el padre portador de la variante c.640A>T y la madre portadora de la variante c.1199G>C. Ante este diagnóstico se añadió al tratamiento un cóctel antioxidante y se inició seguimiento por otros subespecialistas, incluyendo neurología, hepatología y cardiología. La deficiencia combinada de la fosforilación oxidativa tipo 1 es una enfermedad incurable neurodegenerativa sin tratamiento específico, por lo que este paciente se reorientó hacia cuidados paliativos.

Tabla 2. Variantes identificadas en el caso 1

<i>Gen</i>	<i>Variante</i>	<i>Estado</i>	<i>Frecuencia *</i>	<i>Patogenicidad **</i>
<b>GFM1</b>	NM_024996.7:c.640A>T (p.Ile214Phe)	Heterocigosis	No descrita	Variante de significado incierto
<b>GFM1</b>	NM_024996.7: c.1199G>C (p.Arg400Pro)	Heterocigosis	0.0000796	Variante de significado incierto

\*GnomAD Exomes: Frecuencia descrita en Genome Aggregation Database

\*\*ACMG: American College of Medical Genetics

Tabla 3. Análisis in silico de la patogenicidad en el caso 1

<i>Variante</i>	<i>General</i>		<i>Conservación</i>	<i>Funcional</i>	
	<i>DANN</i>	<i>Mutation Tester</i>	<i>GERP</i>	<i>SIFT</i>	<i>Provean</i>
<b>GFM1 c.640A&gt;T</b>	0.9809	Causante de enfermedad	3.77	Deletérea	Deletérea
<b>GFM1 c.1199G&gt;C</b>	0.9973	Causante de enfermedad	4.01	Deletérea	Deletérea

DANN: El intervalo de valores es de 0 a 1, siendo 1 atribuido a las variantes previstas como más perjudiciales

GERP: varía de -12,3 a 6,17, siendo 6,17 el correspondiente a mayor conservación

**CASO 2.** Se trata de un paciente de 1 mes y 15 días de vida que ingresó en la UCIN para estudio de miocardiopatía y acidosis metabólica con hiperlactacidemia. En el cribado metabólico neonatal se evidenció una elevación de ácido metilmalónico en la orina. Ante estos hallazgos se realizó estudio genético mediante panel NeoSeq obteniéndose los resultados a los 7 días tras la recepción de la muestra. Se identificaron las variantes c.850C>T/c.859C>T en heterocigosis compuesta en el gen *SUCLA2* (MIM \*603921) (Tabla 5), el cual se encuentra localizado en el cromosoma 13 (13q14.2). Las variantes patógenas en el gen *SUCLA2* están asociadas al síndrome de depleción del ADN mitocondrial 5 (MIM # 612073), también conocido como encefalomiopatía con/sin aciduria metilmalónica. Se trata de un trastorno autosómico recesivo que se caracteriza por la aparición infantil de hipotonía, deterioro neurológico progresivo, trastorno del movimiento hipercinético-distónico, oftalmoplejía externa, sordera y disfunción tubular renal variable. Nuestro paciente presenta un fenotipo compatible con el cuadro clínico descrito, evidenciado mediante el análisis por Phenomizer de forma estadísticamente significativa (p 0.0271). El estudio en trio permitió confirmar la herencia autosómica recesiva siendo el padre portador de la variante c.850C>T y la madre portadora de la variante c.859C>T. Ante este diagnóstico se añadió al tratamiento un cóctel antioxidante y se inició seguimiento por otros subespecialistas, incluyendo neurología, oftalmología, audiología y nefrología. El síndrome de depleción del ADN mitocondrial 5 es una enfermedad incurable neurodegenerativa sin tratamiento específico, por lo que este paciente se reorientó hacia cuidados paliativos.

Tabla 4. Variantes identificadas en el caso 2

Gen	Variante	Estado	Frecuencia *	Patogenicidad **
SUCLA2	NM_003850.3: c.850C>T (p.Arg284Cys)	Homocigosis	No descrita	Patogénica

\*GnomAD Exomes: Frecuencia descrita em Genome Aggregation Database

\*\*ACMG: American College of Medical Genetics

Tabla 5. Análisis in silico de la patogenicidad en el caso 2

Variante	General		Conservación	Funcional	
	DANN	Mutation Tester	GERP	SIFT	Provean
SUCLA2 c.850C>T	0.9991	Causante de enfermedad	5.67	Deletérea	Deletérea

DANN: El intervalo de valores es de 0 a 1, siendo 1 atribuido a las variantes previstas como más perjudiciales

GERP: varía de -12,3 a 6,17, siendo 6,17 el correspondiente a mayor conservación

**CASO 3.** Se trata de un paciente de 20 días de vida y sexo femenino, fruto de un embarazo a término sin incidencias y que ingresó en la UCIN a los 15 días por convulsiones. En el registro electroencefalográfico (EEG) se evidenciaba un trazado con múltiples paroxismos y que evolucionó a brote supresión a pesar del tratamiento con fenobarbital. Durante su ingreso se observaron datos sugestivos de encefalopatía epiléptica. Ante estos hallazgos se realizó estudio genético mediante panel NeoSeq, obteniéndose los resultados a los 6 días tras la recepción de la muestra. Se identificó la variante c.778C>T en heterocigosis en el gen *KCNQ2* (MIM \*602235) (Tabla 6), el cual está localizado en el cromosoma 20 (20q13.33) y codifica un canal de potasio dependiente de voltaje que se expresa en el cerebro (139). Las características clínicas de los trastornos relacionados con *KCNQ2* van desde la epilepsia neonatal familiar benigna *KCNQ2*-BFNE tipo 1 (MIM #121200) en el extremo leve

hasta la encefalopatía epiléptica infantil precoz KCNQ2-NEE tipo 7 (MIM #613720), en el extremo grave. KCNQ2-BFNE(139-141) se caracteriza por un amplio espectro de tipos de convulsiones (episodios tónicos o apneicos, actividad clónica focal o cambios autonómicos) que comienzan en neonatos por lo demás sanos entre el segundo y el octavo día de vida y desaparecen espontáneamente entre el primero y el sexto a duodécimo mes de vida. La actividad motora puede limitarse a una parte del cuerpo, migrar a otras regiones del cuerpo o generalizarse. KCNQ2-NEE (142, 143) se caracteriza por múltiples convulsiones diarias que comienzan en la primera semana de vida y son en su mayoría tónicas, con características motoras focales y autonómicas asociadas. Las convulsiones generalmente cesan entre los nueve meses y los cuatro años. Al inicio, el EEG muestra un patrón de supresión de estallidos o actividad epileptiforme multifocal; la resonancia magnética cerebral temprana puede mostrar ganglios basales e hiperintensidades talámicas que luego se resuelven. Hay un deterioro del desarrollo de moderado a grave. En cuanto al tratamiento de las manifestaciones. Los niños con KCNQ2-NEE tienen múltiples convulsiones diarias resistentes al inicio del fenobarbital y otros fármacos antiepilépticos comunes de vieja y nueva generación, solos o en combinación. Se demostró que los bloqueantes de los canales de sodio como la fenitoína o la carbamazepina controlan las convulsiones mejor las convulsiones y deben considerarse un tratamiento de primera línea. Nuestro paciente presenta un fenotipo grave compatible con encefalopatía epiléptica infantil precoz KCNQ2-NEE tipo 7, evidenciado también mediante el análisis por Phenomizer de forma estadísticamente significativa ( $p < 0.04$ ). El estudio

en trio permitió confirmar que se trataba de una variante de novo en un gen asociado a un trastorno autosómico dominante. Ante este diagnóstico se inició tratamiento con fenitoina, con un mejor control de las convulsiones.

**Tabla 6. Variantes identificadas en el caso 3**

<i>Gen</i>	<i>Variante</i>	<i>Estado</i>	<i>Frecuencia *</i>	<i>Patogenicidad **</i>
<i>KCNQ2</i>	NM_172107.4: c.778C>T (p.His260Tyr)	Heterocigosis	No descrita	Patogénica

\*GnomAD Exomes: Frecuencia descrita en Genome Aggregation Database

\*\*ACMG: American College of Medical Genetics

**Tabla 7. Análisis in silico de la patogenicidad en el caso 3**

<i>Variante</i>	<i>General</i>		<i>Conservación</i>	<i>Funcional</i>	
	<i>DANN</i>	<i>Mutation Tester</i>	<i>GERP</i>	<i>SIFT</i>	<i>Provean</i>
<i>KCNQ2 c.778C&gt;T</i>	0.9927	Polimorfismo	3.38	Deletérea	Deletérea

DANN: El intervalo de valores es de 0 a 1, siendo 1 atribuido a las variantes previstas como más perjudiciales

GERP: varía de -12,3 a 6,17, siendo 6,17 el correspondiente a mayor conservación

**CASO 4.** Se trata de un paciente de 18 días de vida y sexo masculino, fruto de un embarazo sin incidencias. Ingresó a los 3 días de vida por convulsiones tónicas generalizadas y apneas de difícil control y con mala respuesta a midazolam y fenobarbital. En los días posteriores se evidencian también datos de encefalopatía con disminución del nivel de conciencia y pobre succión. Ante estos hallazgos se realizó estudio genético mediante panel NeoSeq, obteniéndose los resultados a los 7 días tras la recepción de la muestra. Se identificó la variante c.1016T>A en heterocigosis en el gen *KCNQ2* (MIM \*602235) (Tabla 7), que como se explicó anteriormente está localizado en el cromosoma 20 (20q13.33) y codifica un canal de potasio dependiente de

voltaje que se expresa en el cerebro (139). El cambio alternativo c.1016T>G (p.Leu339Arg), que afecta al mismo aminoácido, está descrito en la literatura como causante de enfermedad (144). Nuestro paciente presenta un fenotipo grave compatible con encefalopatía epiléptica infantil precoz KCNQ2-NEE tipo 7 (MIM #613720), el cual es también compatible con el resultado del análisis por Phenomyzer de forma estadísticamente significativa (p 0.04). El estudio en trio evidenció que el padre portaba la misma variante también en heterocigosis, habiendo presentado episodios convulsivos focales y apneicos sin datos de encefalopatía durante el primer mes de vida, con buena respuesta al fenobarbital y que permitieron suspender la medicación el año de vida sin recaída. El fenotipo paterno más benigno es compatible con la KCNQ2-BBFNE tipo 1.

Tabla 8. Variantes identificadas en el caso 4

<i>Gen</i>	<i>Variante</i>	<i>Estado</i>	<i>Frecuencia *</i>	<i>Patogenicidad **</i>
<b>KCNQ2</b>	NM_172107.4:c.1016T>A (p.Leu339Gln)	Heterocigosis	No descrita	Patogénica

\*GnomAD Exomes: Frecuencia descrita em Genome Aggregation Database

\*\*ACMG: American College of Medical Genetics

Tabla 9. Análisis in silico de la patogenicidad en el caso 4

<i>Variante</i>	<i>General</i>		<i>Conservación</i>	<i>Funcional</i>	
	<i>DANN</i>	<i>Mutation Tester</i>	<i>GERP</i>	<i>SIFT</i>	<i>Provean</i>
<b>KCNQ2 c.1016T&gt;A</b>	0.9932	Causante de enfermedad	4.34	Deletérea	Deletérea

DANN: El intervalo de valores es de 0 a 1, siendo 1 atribuido a las variantes previstas como más perjudiciales

GERP: varía de -12,3 a 6,17, siendo 6,17 el correspondiente a mayor conservación

**CASO 5.** Se trata de un paciente de 25 días de vida y sexo femenino, fruto de un embarazo a término sin incidencias y que ingresó en la UCIN a los 20 días de vida por convulsiones con un trazado electroencefalográfico que evolucionó a hipsarritmia resistente a múltiples anticonvulsivantes. Durante su evolución asoció encefalopatía epiléptica. Tras estos hallazgos se realizó estudio genético mediante panel NeoSeq obteniéndose los resultados a los 6 días tras la recepción de la muestra. Se identificó la variante c.616G>T en el gen *CDKL5* (OMIM \* 300203) (Tabla 8), localizado en el cromosoma X (Xp22.13) y variantes patógenas en el mismo se han asociado a la encefalopatía epiléptica precoz tipo 2 (MIM # 300672). Se trata de trastorno neurológico grave dominante ligado al cromosoma X que se caracteriza por la aparición de convulsiones en los primeros meses de vida y un retraso grave del desarrollo con discapacidad intelectual. Otras características incluyen rasgos faciales dismórficos sutiles, alteraciones del sueño, problemas gastrointestinales y movimientos estereotipados de las manos. Existe cierta superposición fenotípica con el síndrome de Rett (MIM # 312750), pero se considera que la encefalopatía epiléptica precoz tipo 2 es una entidad distinta. Nuestro paciente presenta un fenotipo grave compatible con esta patología, si bien el análisis por Phenomizer no fue significativo (p 0.1386). El estudio en trio permitió confirmar que se trataba de una variante heredada de la madre, la cual presenta un cuadro más leve de epilepsia sin encefalopatía, lo cual confirma la herencia dominante ligada a X. Ante este diagnóstico y la mala respuesta prevista al tratamiento anticonvulsivante se decidió iniciar dieta cetogénica y seguimiento con un especialista en dietética.

Tabla 10. Variantes identificadas en el caso 5

Gen	Variante	Estado	Frecuencia *	Patogenicidad **
CDKL5	NM_003159.2:c.616G>T (p.Asp206Tyr)	Heterocigosis	No descrita	Probablemente patogénica

\*GnomAD Exomes: Frecuencia descrita en Genome Aggregation Database

\*\*ACMG: American College of Medical Genetics

Tabla 11. Análisis in silico de la patogenicidad en el caso 5

Variante	General		Conservación	Funcional	
	DANN	Mutation Tester	GERP	SIFT	Provean
CDKL5 c.616G>T	0.9958	Causante de enfermedad	6.0999	Deletérea	Deletérea

DANN: El intervalo de valores es de 0 a 1, siendo 1 atribuido a las variantes previstas como más perjudiciales

GERP: varía de -12,3 a 6,17, siendo 6,17 el correspondiente a mayor conservación

**CASO 6.** Se trata de un paciente de 6 días de vida de sexo masculino que nació a término fruto de una gestación sin incidencias. Ingresó en la UCIN por distrés respiratorio neonatal al nacimiento precisando soporte respiratorio con ventilación mecánica no invasiva con buena evolución. Sin embargo, a los 3 días de vida inicia convulsiones con mal respuesta al tratamiento con midazolam y datos de encefalopatía epiléptica, acompañada de acidosis metabólica crónica. Ante estos hallazgos se realiza estudio genético mediante panel NeoSeq obteniéndose los resultados en un plazo de 7 días tras la recepción de las muestras. Se identificó la variante c.1258G>A en heterocigosis en el gen *KCNQ2* (MIM \*602235) (Tabla 9), el cual está localizado en el cromosoma 20 (20q13.33) y codifica un canal de potasio dependiente de voltaje que se expresa en el cerebro (139). La variante p.Arg553Gln en el gen *KCNQ2* se ha identificado previamente en individuos con convulsiones neonatales familiares (145). Los estudios funcionales *in vitro* realizados sugieren que este cambio altera

la función del canal KCNQ2 (145, 146). La variante p.Arg553Gln es una sustitución de aminoácidos semiconservativa, que puede afectar la estructura de la proteína secundaria ya que estos residuos difieren en algunas propiedades. Este residuo está dentro del dominio citoplásmico C-terminal. Los análisis in-silico, incluidos los predictores de proteínas y la conservación evolutiva, respaldan un efecto deletéreo. Por lo tanto, este resultado es consistente con el diagnóstico de un trastorno relacionado con KCNQ2 en este individuo. El fenotipo del paciente es compatible con el fenotipo grave del espectro, si bien el análisis por Phenomizer no fue estadísticamente significativo (p 0.471). El estudio en trío permitió confirmar que se trataba de una variante de novo en un gen asociado a un trastorno autosómico dominante. Dado el efecto beneficioso de los antiepilépticos que bloquean los canales de sodio en la encefalopatía epiléptica infantil precoz KCNQ2-NEE tipo 7 (MIM #613720), ante los resultados del estudio genético se decidió iniciar tratamiento con carbamacepina.

**Tabla 12. Variantes identificadas en el caso 6**

<i>Gen</i>	<i>Variante</i>	<i>Estado</i>	<i>Frecuencia *</i>	<i>Patogenicidad **</i>
<b>KCNQ2</b>	NM_172107.4: c.1658G>A (p.Arg553Gln)	Heterocigosis	No descrita	Patogénica

\*GnomAD Exomes: Frecuencia descrita en Genome Aggregation Database

\*\*ACMG: American College of Medical Genetics

**Tabla 13. Análisis in silico de la patogenicidad**

<i>Variante</i>	<i>General</i>		<i>Conservación</i>	<i>Funcional</i>	
	<i>DANN</i>	<i>Mutation Tester</i>	<i>GERP</i>	<i>SIFT</i>	<i>Provean</i>
<b>KCNQ2 c.1658G&gt;A</b>	0.9995	Causante de enfermedad	4.9899	Deletérea	Deletérea

DANN: El intervalo de valores es de 0 a 1, siendo 1 atribuido a las variantes previstas como más perjudiciales

GERP: varía de -12,3 a 6,17, siendo 6,17 el correspondiente a mayor conservación

**CASO 7.** Se trata de un neonato de 10 días de vida y sexo masculino fruto de un embarazo y parto a término sin incidencias. Durante las ecografías pregestacionales se había evidenciado una ectasia piélica fetal. A los 4 días de vida inicia un cuadro de vómitos con hiperamonemia e hipoglucemia hipocetósica que precisa ingreso en UCIN. Ante estos hallazgos se realizó estudio genético mediante panel NeoSeq que identificaron en un plazo de 8 días las variantes c.1201G>C/c.2810Y>A en heterocigosis compuesta en el gen *CPS1* (MIM \* 608307) (Tabla 10), localizado en el cromosoma 2 (2q34). Variantes patógenas en el gen *CPS1* se han asociado a la deficiencia de carbamoilfosfato sintetasa 1 (MIM #237300), un error innato autosómico recesivo del metabolismo del ciclo de la urea que causa hiperamonemia. Ambas variantes han sido descritas previamente en la literatura como patógenas (147). Hay 2 formas principales: un tipo neonatal letal y un tipo menos grave de aparición tardía (148). Los trastornos del ciclo de la urea se caracterizan por la tríada de hiperamonemia, encefalopatía y alcalosis respiratoria. En cuanto a las formas de inicio precoz de la deficiencia de CPS1, este trastorno fue reportado por primera vez por Freeman, et al. (149) en un paciente con hiperamonemia congénita y niveles disminuidos de carbamoil fosfato sintetasa, así como daño cerebral severo. El tratamiento consiste en administración de quelantes del amonio, suplementación con aminoácidos esenciales (arginina), alimentación restringida en proteína natural y, cuando fallan otras medidas, diálisis. En estos pacientes es necesario evitar ciertas medicaciones como valproato, haloperidol y corticosteroides sistémicos, así como

el ayuno y el estrés físico y psicológico. El fenotipo del paciente es compatible con el fenotipo grave del espectro, si bien el análisis por Phenomizer no fue estadísticamente significativo (p 0.0624). El estudio en trío evidenció que el padre era portador de la variante c.1201G>C y la madre de la variante c.2810Y>A, confirmado la herencia autosómica recesiva de esta patología. Ante este diagnóstico se instauró tratamiento con dieta restringida en proteínas, suplementación con arginina y citrulina y administración de fármacos quelantes del amonio. Además, se derivó al paciente a la Unidad de Diagnóstico y Tratamiento de Enfermedades Metabólicas Congénitas.

Tabla 14. Variantes identificadas en el caso 7

<i>Gen</i>	<i>Variante</i>	<i>Estado</i>	<i>Frecuencia *</i>	<i>Patogenicidad **</i>
<i>CPS1</i>	NM_001875.5:c.1201G>C (p.Gly401Arg)	Heterocigosis	0.00002	Probablemente patogénica
<i>CPS1</i>	NM_001875.5):c.2810T>A (p.Ile937Asn)	Heterocigosis	0.0000199	Probablemente patogénica

\*GnomAD Exomes: Frecuencia descrita en Genome Aggregation Database

\*\*ACMG: American College of Medical Genetics

Tabla 15. Análisis in silico de la patogenicidad en el caso 7

<i>Variante</i>	<i>General</i>		<i>Conservación</i>	<i>Funcional</i>	
	<i>DANN</i>	<i>Mutation Tester</i>	<i>GERP</i>	<i>SIFT</i>	<i>Provean</i>
<i>CPS1</i> c.1201G>C	0.9982	Causante de enfermedad	5.88	Tolerada	Deletérea
<i>CPS1</i> c.1199G>C	0.9902	Causante de enfermedad	5.4699	Deletérea	Delétèrea

DANN: El intervalo de valores es de 0 a 1, siendo 1 atribuido a las variantes previstas como más perjudiciales

GERP: varía de -12,3 a 6,17, siendo 6,17 el correspondiente a mayor conservación

**CASO 8.** Se trata de un paciente de un mes de vida fruto de un embarazo y parto a término que al nacimiento presenta una hipotonía generalizada con movimientos espontáneos disminuidos y pobre succión. A los dos días de vida destaca disminución de conciencia con letargia por lo que se conecta a monitor de función cerebral evidenciándose convulsiones. Asimismo, se realiza una ecografía cerebral que impresiona de cerebelo disminuido de tamaño, confirmándose una atrofia cerebelosa mediante resonancia magnética cerebral a los 15 días de vida. Ante estos hallazgos se realiza estudio genético mediante panel NeoSeq obteniéndose los resultados en el plazo 10 días tras la recepción de las muestras. Se identifican las variantes c.202G>C/c.718C>T en el gen *COQ4* (MIM \* 612898) en heterocigosis compuesta (Tabla 11). Este gen está localizado en el cromosoma 9 (9q34.11) y codifica la enzima COQ4, la cual está implicada en la biosíntesis de la coenzima Q10 (CoQ10) (150). Variantes patógenas en *COQ4* se asocian a la deficiencia primaria de coenzima Q10 tipo 7 (MIM #616276), un trastorno autosómico recesivo resultante de una disfunción mitocondrial. Ambas variantes c.202G>C (151) y c.718C>T (152) han sido previamente descrita en la literatura como patogénicas. La mayoría de los pacientes presentan síntomas neurológicos graves incluyendo las convulsiones y la encefalopatía epiléptica con retraso global del desarrollo poco después del nacimiento, que por lo general conducen a la muerte, asociando con frecuencia atrofia o hipoplasia del cerebelo y alteraciones cardíacas incluyendo miocardiopatía hipertrófica. En individuos con deficiencia primaria de CoQ10, el tratamiento temprano con suplementos de CoQ10 oral en dosis altas (que van de 5 a 50 mg / kg / día)

puede limitar la progresión de la enfermedad y revertir algunas manifestaciones; sin embargo, el daño neurológico grave establecido no se puede revertir. El fenotipo del paciente es compatible con la deficiencia primaria de coenzima Q10 tipo 7, si bien el análisis por Phenomizer no arrojó ningún resultado como patología posible. El estudio en trío evidenció que el padre era portador de la variante c.718C>T y la madre de la variante c.202G>C, confirmado la herencia autosómica recesiva de esta patología. Ante este diagnóstico se instauró tratamiento con altas dosis de coenzima Q10 y se inició seguimiento por cardiología pediátrica.

Tabla 16. Variantes identificadas en el caso 8

<i>Gen</i>	<i>Variante</i>	<i>Estado</i>	<i>Frecuencia</i> *	<i>Patogenicidad</i> **
<b>COQ4</b>	NM_016035.5:c.202G>C (p.Asp68His)	Heterocigosis	0.0000432	Patogénica
<b>COQ4</b>	NM_016035.5:c.718C>T (p.Arg240Cys)	Heterocigosis	0.000198	Patogénica

\*GnomAD Exomes: Frecuencia descrita en Genome Aggregation Database

\*\*ACMG: American College of Medical Genetics

Tabla 17. Análisis in silico de la patogenicidad en el caso 8

<i>Variante</i>	<i>General</i>		<i>Conservación</i>	<i>Funcional</i>	
	<i>DANN</i>	<i>Mutation Tester</i>	<i>GERP</i>	<i>SIFT</i>	<i>Provean</i>
<b>COQ4</b> <b>c.202G&gt;C</b>	0.9952	Causante de enfermedad	5.4899	Deletérea	Deletérea
<b>COQ4</b> <b>c.718C&gt;T</b>	0.9994	Causante de enfermedad	5.5199	Deletérea	Deletérea

DANN: El intervalo de valores es de 0 a 1, siendo 1 atribuido a las variantes previstas como más perjudiciales

GERP: varía de -12,3 a 6,17, siendo 6,17 el correspondiente a mayor conservación

**CASO 9.** Se trata de un recién nacido prematuro de 29 semanas de edad gestacional y sexo masculino de 3 días de vida. Desde el nacimiento destaca hipotonía neonatal generalizada más marcada en cara y a nivel axial con debilidad muscular, requiriendo soporte respiratorio invasivo. Asimismo, destaca escasa conexión con el medio en relación a encefalopatía. Ante estos hallazgos se realiza estudio genético mediante panel NeoSeq obteniéndose los resultados a los 10 días tras la recepción de las muestras. Se identificó la variante c.614C>A en heterocigosis en el gen *ACTA1* (MIM \*102610) (Tabla 12), el cual se encuentra localizado en el cromosoma 1 (1q42.13) y codifica la alfa-actina del músculo esquelético, la principal isoforma de actina en el músculo esquelético adulto. Variantes patógenas en este gen están asociadas con la miopatía nemalínica 3 (MIM #161800). La miopatía nemalínica es una forma de miopatía congénita caracterizada por estructuras anómalas en forma de hilo o varilla en las fibras musculares en el examen histológico ("nema" en griego significa "hilo"). El fenotipo clínico es muy variable, con diferentes edades de inicio y gravedad. La debilidad muscular suele afectar a los músculos proximales, incluyendo los músculos facial, bulbar y respiratorio(153). En general, se pueden distinguir fácilmente 2 grupos clínicos: "típico" y "grave". La miopatía nemalínica típica es la forma más común y se presenta como hipotonía infantil y debilidad muscular. Es lentamente progresivo o no progresivo y la mayoría de los adultos logran la deambulación. La forma grave del trastorno se caracteriza por ausencia de movimiento espontáneo o respiración al nacer, artrogriposis y muerte en los primeros meses de vida. Algunos pacientes tienen bastones clásicos, mientras que otros también pueden mostrar bastones intranucleares, filamentos agrupados, núcleos o

desproporción de tipo de fibra, todos los cuales son hallazgos patológicos inespecíficos y no patognomónicos de una miopatía congénita específica. El fenotipo del paciente es compatible con la forma grave de este trastorno. El estudio en trio identificó que se trata de una variante de novo con herencia autosómica dominante. Ante este diagnóstico se realizó una biopsia muscular que objetivó estructuras en forma de bastones en el sarcolema (conocidos como “cuerpos nemalínicos”). Se inició tratamiento con piridostigmina y carnitina sin mejoría en la evolución. Dado el pronóstico infausto de la patología y la ausencia de tratamientos específicos el paciente se reorientó hacia cuidados paliativos.

Tabla 18. Variantes identificadas en el caso 9

<i>Gen</i>	<i>Variante</i>	<i>Estado</i>	<i>Frecuencia *</i>	<i>Patogenicidad **</i>
<b>ACTA1</b>	NM_001100.4:c.614C>A (p.Thr205Lys)	Heterocigosis	ND	Probablemente patogénica

\*GnomAD Exomes: Frecuencia descrita en Genome Aggregation Database

\*\*ACMG: American College of Medical Genetics

Tabla 19. Análisis in silico de la patogenicidad en el caso 9

<i>Variante</i>	<i>General</i>		<i>Conservación</i>	<i>Funcional</i>	
	<i>DANN</i>	<i>Mutation Tester</i>	<i>GERP</i>	<i>SIFT</i>	<i>Provean</i>
<b>ACTA1 c.614C&gt;A</b>	0.9772	Causante de enfermedad	4.4	Deletérea	Deletérea

DANN: El intervalo de valores es de 0 a 1, siendo 1 atribuido a las variantes previstas como más perjudiciales

GERP: varía de -12,3 a 6,17, siendo 6,17 el correspondiente a mayor conservación

**CASO 10.** Se trata de un recién nacido a término de 9 días de vida y sexo femenino que presenta cataratas bilaterales congénitas y quistes renales, completándose el estudio con una resonancia magnética cerebral que evidencia hallazgos compatibles con

leucoencefalopatía. Ante estos hallazgos se realiza estudio genético mediante panel NeoSeq obteniéndose los resultados a los 5 días tras la recepción de las muestras. Se identificó la variante c.2906 G>A en heterocigosis en el gen *COL4A1* (MIM \*120130) (Tabla 13). Este gen se encuentra localizado en el cromosoma 13 (13q14) y codifica la subunidad alfa-1 del colágeno tipo IV. La molécula de colágeno IV es un heterotrímero de 2 cadenas alfa-1 y 1 cadena alfa-2(154). La cadena alfa-2 está codificada en un locus distinto (MIM \*120090). El espectro de trastornos relacionados con *COL4A1* incluye: enfermedad cerebral de vasos pequeños (MIM #175780) de diversa gravedad (155, 156), incluida la porencefalia, asociada de forma variable con defectos oculares (tortuosidad arterial retiniana, anomalía de Axenfeld-Rieger, cataratas) y hallazgos sistémicos (afectación renal, calambres musculares, aneurismas cerebrales, fenómeno de Raynaud, arritmia cardíaca y anemia hemolítica). En los estudios de imagen, la enfermedad cerebral de vasos pequeños se manifiesta como leucoencefalopatía periventricular difusa, infartos lacunares, microhemorragia, espacios perivasculares dilatados y hemorragias intracerebrales profundas. Clínicamente, la enfermedad cerebral de vasos pequeños se manifiesta como hemiparesia infantil, convulsiones, accidente cerebrovascular hemorrágico único o recurrente, accidente cerebrovascular isquémico y migraña aislada con aura. La porencefalia (cavidades llenas de líquido en el cerebro detectadas por tomografía computerizada o resonancia magnética) se manifiesta típicamente como hemiparesia infantil, convulsiones y discapacidad intelectual; sin embargo, en ocasiones puede ser un hallazgo incidental. El síndrome angiopatía hereditaria con nefropatía, aneurismas y

calambres musculares (Síndrome de Hanac) (MIM #611773) suele asociar enfermedad cerebral asintomática de vasos pequeños, afectación de grandes vasos cerebrales (es decir, aneurismas) y hallazgos sistémicos que comprometen el riñón, el músculo y los vasos pequeños del ojo (157, 158). Tres fenotipos adicionales incluyen la tortuosidad aislada de la arteria retiniana (MIM #180000), la microangiopatía con leucoencefalopatía pontina (MIM #618564) y la susceptibilidad a hemorragias intracerebrales (MIM #614519). El tratamiento de las manifestaciones de los trastornos asociados a *COL4A1* incluye el apoyo emocional para las personas afectadas y sus familias. La hipertensión debe tratarse para reducir el riesgo general de accidente cerebrovascular. También deben evitarse el consumo de tabaco, los traumatismos craneoencefálicos y la exposición a anticoagulantes para disminuir el riesgo de hemorragia intracerebral. El fenotipo del paciente es compatible con la enfermedad vascular cerebral de vaso pequeño, si bien el análisis por Phenomizer no fue estadísticamente significativo (p 0.5810). El estudio en trio identificó que se trata de una variante de novo con herencia autosómica dominante. A la semana de recibir el diagnóstico el paciente fue éxitus en relación a una hemorragia digestiva con shock hipovolémico.

**Tabla 20. Variantes identificadas en el caso 10**

<i>Gen</i>	<i>Variante</i>	<i>Estado</i>	<i>Frecuencia</i> *	<i>Patogenicidad</i> **
<i>COL4A1</i>	NM_001845.6:c.2906G>A (p.Gly969Glu)	Heterocigosis	No descrita	Probablemente patogénica

\*GnomAD Exomes: Frecuencia descrita en Genome Aggregation Database

\*\*ACMG: American College of Medical Genetics

Tabla 21. Análisis in silico de la patogenicidad en el caso 10

Variante	General		Conservación	Funcional	
	DANN	Mutation Tester	GERP	SIFT	Provean
<b>COL4A1 c.2906G&gt;A</b>	0.9982	Causante de enfermedad	6.0599	Deletérea	Deletérea

DANN: El intervalo de valores es de 0 a 1, siendo 1 atribuido a las variantes previstas como más perjudiciales

GERP: varía de -12,3 a 6,17, siendo 6,17 el correspondiente a mayor conservación

**CASO 11.** Se trata de un recién nacido de 12 días de vida fruto de un embarazo a término. Durante la gestación la madre refería escasos movimientos fetales y al nacimiento precisó reanimación por respiración poco eficaz con presión positiva intermitente, evidenciándose hipotonía y contracturas articulares, así como microretrognatia, microtia y micropene. Tras completar el estudio de extensión por sospecha de síndrome polimalformativo también se evidenció una hipoacusia neurosensorial. Ante estos hallazgos se realiza estudio genético mediante panel NeoSeq obteniéndose los resultados a los 6 días tras la recepción de la muestra. Se identificó la variante c.1884-9A>G en homocigosis en el gen *ERBB3* (MIM \* 190151) (Tabla 14), la cual se encuentra descrita en la literatura como causante de enfermedad (159) . Dicho gen se encuentra localizado en el cromosoma 12 (12q13.2) y es un miembro de la familia de tirosina quinasas del receptor del factor de crecimiento epidérmico(160). Variantes patogénicas en este gen se han asociado con el síndrome de contractura congénita letal (LCCS) 2 (MIM # 607598), un trastorno autosómico recesivo bien definido descrito originalmente en familias finlandesas. Los criterios diagnósticos del LCCS son hidropesía fetal temprana y

acinesia, fenotipo de Pena-Shokeir (MIM #208150), neuropatía específica con degeneración de las neuronas del asta anterior y atrofia extrema del músculo esquelético. Landau, et al. (161). describieron en 2003 un pedigrí consanguíneo israelí-beduino extendido con contracturas congénitas y anomalías fenotípicas únicas adicionales, lo que sugiere que representa una nueva variante de LCCS autosómico recesivo. Las características que distinguen el nuevo trastorno, LCCS2, del tipo finlandés de LCCS incluyen hallazgos craneofaciales / oculares adicionales, falta de hidropesía, pterigion múltiple y fracturas, así como una duración normal del embarazo. La principal característica clínica única y no descrita anteriormente en el trastorno beduino israelí fue la vejiga urinaria marcadamente distendida; También se observaron otras anomalías urinarias. La mayoría de los bebés con LCCS2 murieron poco después del nacimiento. El diagnóstico prenatal ecográfico fue posible a las 15 semanas de gestación mediante la demostración de acinesia fetal, contracturas de las extremidades, hidramnios y distensión de la vejiga urinaria. Dos niñas, de 12 y 13 años, seguían vivas en el momento del informe. Además de artrogriposis y atrofia muscular, ambos tenían parálisis facial izquierda, problemas oftalmológicos graves con miopía alta y vitreorretinopatía degenerativa. El fenotipo del paciente es compatible con el síndrome de contractura letal tipo 2, si bien el análisis por Phenomizer no resultó estadísticamente significativo (p.04243). El estudio en trio identificó que ambos padres eran portadores heterocigotos de la variante c.1884-9A>G, confirmando la herencia autosómica recesiva. Dada la ausencia de tratamiento específico y el

mal pronóstico de la patología de base el paciente se reorientó hacia cuidados paliativos.

**Tabla 22. Variantes identificadas en el caso 11**

<i>Gen</i>	<i>Variante</i>	<i>Estado</i>	<i>Frecuencia *</i>	<i>Patogenicidad **</i>
<b>ERBB3</b>	NM_001982.4):c.1184-9A>G	Homocigosis	No descrita	Probablemente patogénica

\*GnomAD Exomes: Frecuencia descrita en Genome Aggregation Database<sup>1</sup>

\*\*ACMG: American College of Medical Genetics<sup>2</sup>

**Tabla 23. Análisis in silico de la patogenicidad en el caso 11**

<i>Variante</i>	<i>General</i>		<i>Conservación</i>	<i>Funcional</i>	
	<i>DANN</i>	<i>Mutation Tester</i>	<i>GERP</i>	<i>SIFT</i>	<i>Provean</i>
<b>ERBB3 c.1184-9A&gt;G</b>	0.7883	-	1.4299	-	-

DANN: El intervalo de valores es de 0 a 1, siendo 1 atribuido a las variantes previstas como más perjudiciales<sup>3</sup>.

GERP: varía de -12,3 a 6,17, siendo 6,17 el correspondiente a mayor conservación<sup>4</sup>.

**CASO 12.** Se trata de un recién nacido de 15 días de vida tras un embarazo y parto sin incidencias. Realizó meconiorrexia a las 72 horas de vida, presentando distensión abdominal por lo que se ingresó para completar estudio, evidenciándose un hábito intestinal estreñido precisando estimulación rectal periódica. Tras realización de enema con contraste y biopsia rectal se confirmó una enfermedad de Hirschsprung. A los 10 días de vida inició movimientos oculares conjugados caóticos, multidireccionales y rápidos que se filieron como opsoclonus, completándose estudio con ecografía abdominal y resonancia magnética cerebral que descartaron malformaciones del sistema nervioso central y neuroblastoma. Asimismo, se constató una hipoacusia neurosensorial. Ante estos hallazgos se realizó estudio

genético mediante panel NeoSeq obteniéndose los resultados a los 7 días tras la recepción de las muestras. Se identificó la variante c.850G>T en heterocigosis en el gen *SOX10* (\*602229), este gen está localizado en el cromosoma 22 (22q13.1) (Tabla 15) y codifica un factor de transcripción caracterizado por un motivo de unión al ADN conocido como dominio HMG (grupo de alta movilidad). La familia de genes *SOX*, que incluye *SOX10*, consta de genes relacionados con *SRY*, con una identidad de secuencia de más del 60% con la caja *SRY* HMG. *SOX10* es un factor de transcripción que funciona en la cresta neural y el desarrollo de oligodendrocitos (162, 163). Variantes patógenas en este gen están asociadas a tres trastornos de herencia autosómica dominante: síndrome de Waardenburg tipo 4C (MIM # 613266), síndrome PCWH (MIM #609136) y síndrome de Waardenburg tipo 2E con o sin manifestaciones neurológicas asociadas (MIM # 611584). El síndrome de Waardenburg-Shah, también conocido como síndrome de Waardenburg tipo 4 se caracteriza por sordera, anomalías pigmentarias y enfermedad de Hirschsprung (164, 165). Todas estas características son causadas por defectos en la cresta neural embrionaria. El síndrome PCHW incluye neuropatía desmielinizante periférica, dismielinización central, síndrome de Waardenburg y enfermedad de Hirschsprung. De 12 pacientes con síndrome de Waardenburg-Shah, Touraine, et al.(166) describieron 3 pacientes con retraso del crecimiento y un fenotipo neurológico no reportado previamente con deterioro de los sistemas nerviosos central y autónomo y ocasionalmente hipotonía y artrogriposis neonatal. El síndrome de Waardenburg tipo 2E consiste en hipogonadismo hipogonadotrofo con anosmia e hiposmia asociado en forma variable a sordera, anomalías

pigmentarias, enfermedad de Hirschsprung y afectación neurológica (165). El fenotipo del paciente es compatible con el síndrome de Waardenburg-Shah y el análisis por Phenomizer fue significativo (p.00323). El estudio en trio identificó que se trataba de una variante de novo, confirmando la herencia autosómica dominante. Dada la ausencia de tratamiento específico y el mal pronóstico de la patología de base el paciente se reorientó hacia cuidados paliativos.

**Tabla 24. Variantes identificadas en el caso 12**

<i>Gen</i>	<i>Variante</i>	<i>Estado</i>	<i>Frecuencia *</i>	<i>Patogenicidad **</i>
<i>SOX10</i>	NM_006941.4):c.850G>T (p.Glu284Ter)	Heterocigosis	No descrita	Probablemente patogénica

\*GnomAD Exomes: Frecuencia descrita en Genome Aggregation Database

\*\*ACMG: American College of Medical Genetics

**Tabla 25. Análisis in silico de la patogenicidad en el caso 12**

<i>Variante</i>	<i>General</i>		<i>Conservación</i>	<i>Funcional</i>	
	<i>DANN</i>	<i>Mutation Tester</i>	<i>GERP</i>	<i>SIFT</i>	<i>Provean</i>
<i>SOX10 c.850G&gt;T</i>	0.9969	Causante de enfermedad	4.73	-	-

DANN: El intervalo de valores es de 0 a 1, siendo 1 atribuido a las variantes previstas como más perjudiciales

GERP: varía de -12,3 a 6,17, siendo 6,17 el correspondiente a mayor conservación

En resumen, en doce identificamos la causa molecular de su enfermedad llegando a un diagnóstico genético. En cinco de ellos su patología estaba asociada a una herencia autosómica recesiva y en siete a una herencia autosómica dominante. De estos siete pacientes, cinco portaban variantes de novo y dos habían heredado la variante de un solo padre, que presentaba un fenotipo menos grave.

#### 4.2.2 Pacientes con variantes no definitivamente/ probablemente causales

En tres casos se identificaron se identificaron variantes potencialmente asociadas a enfermedad pero se descartaron como causales en los neonatos estudiados por incompatibilidad del cuadro clínico y/o de la herencia asociada:

**CASO 13.** Se trata de un neonato de 12 días de vida y sexo femenino que nació mediante cesárea a las 37 semanas de edad gestacional por preeclampsia materna. Ingresó en la UCIN por distrés respiratorio que precisó ventilación mecánica invasiva presentando hipoglucemias con requerimientos de glucosa intravenosa de hasta 12 mg/kg/min durante los primeros 5 días de vida. Tras realizar estudio genético con panel NeoSeq se identificó la variante c.126G>T en heterocigosis en el gen *MTOR* (MIM \* 601231). Este gen se encuentra localizado en 1p36.22 y variantes patógenas en el mismo se han asociado al síndrome de Smith-Kingsmore (MIM # 616638) y a la displasia focal cortical tipo (MIM # 607341), ambos de herencia autosómica recesiva. El primero es un síndrome de discapacidad intelectual autosómica dominante poco común caracterizado por macrocefalia, convulsiones, hernia umbilical y características dismórficas faciales que incluyen protuberancia frontal, hipoplasia del tercio medio facial, mentón pequeño, hipertelorismo con fisuras palpebrales inclinadas hacia abajo, puente nasal deprimido, surco nasolabial liso y labio superior delgado (167, 168). La displasia focal cortical tipo 2 es una malformación del desarrollo cerebral que resulta en un fenotipo clínico de epilepsia intratable y que generalmente requiere cirugía de resección del tejido cerebral afectado para

mejorar la frecuencia y gravedad de las convulsiones (169, 170). En el caso del paciente n° 14 ninguna de las patologías asociadas con variantes patógenas en el gen *MTOR* explicarían su fenotipo, además de que las patologías asociadas a este gen siempre presentan una herencia autosómica recesiva. En este sentido, el estudio en trio evidenció que el padre asintomático también portaba la misma mutación en heterocigosis.

**CASO 14.** Se trata de un neonato de 29 días de vida y sexo femenino que nació a término e ingresó en la UCIN por distrés respiratorio que requirió soporte ventilatorio no invasivo y bajo peso. En la radiografía de tórax al ingreso se observaron 10 pares de costillas. A la exploración física destacaba hipotonía generalizada neonatal y un click de cadera bilateral, completándose estudio con ecografía de caderas que constató displasia congénita de caderas. Presentó además distensión abdominal progresiva y vómitos, con meconiorrexis a los 4 días de vida, estableciéndose el diagnóstico de íleo meconial. Tras realizar estudio genético con panel NeoSeq se identificó la variante c.2507C>G en heterocigosis en el gen *NALCN* (MIM \* 611549). Este gen se encuentra localizado en 13q32.3-q33.1 y variantes patógenas en el mismo se han asociado a hipotonía infantil con retraso psicomotor y facies característica 1 (MIM # 615419) y el síndrome de contracturas congénitas de extremidades y cara, hipotonía y retraso en el desarrollo (MIM # 616266). El primero de estos cuadros es un trastorno neurológico autosómico recesivo grave que se inicia al nacer o en la primera infancia. Los individuos afectados suelen presentar retraso

psicomotor y discapacidad intelectual grave. Algunos pacientes nunca aprenden a sentarse o caminar de forma independiente (171, 172). Por otra parte, el síndrome de contracturas congénitas de extremidades y cara, hipotonía y retraso en el desarrollo es un trastorno autosómico dominante que se caracteriza por contracturas congénitas de las extremidades y la cara, que dan como resultado rasgos faciales característicos (fisuras palpebrales inclinadas hacia abajo, puente nasal ancho con punta nasal antevertida, narinas grandes, columela corta, surco nasolabial largo, micrognatia, hoyuelos en el mentón en forma de H), hipotonía y grados variables de retraso en el desarrollo (173, 174). En el caso del paciente nº 15 ninguna de las patologías asociadas con variantes patógenas en el gen *MTOR* explicarían su fenotipo. Además, el estudio en trio evidenció que el padre asintomático también portaba la misma mutación en heterocigosis.

**CASO 15.** Se trata de un neonato de 4 días de vida y sexo masculino que a los 2 días de vida inició convulsiones focales, sin otros hallazgos, con buena respuesta al fenobarbital. Tras realizar el estudio genético mediante panel NeoSeq se identificó la variante c.1724T>C en heterocigosis en el gen *TSC2* (MIM \* 191092) que se localiza en 16p13.3 y codifica la proteína tuberina. Variantes patógenas en este gen se han asociado al complejo de esclerosis tuberosa 2 (MIM # 613254), con la linfangioleiomiomatosis familiar (MIM # 606690) y la displasia cortical focal tipo 2 (MIM # 607341). El complejo de esclerosis tuberosa 2 es un trastorno multisistémico de herencia autosómica dominante caracterizado por hamartomas en múltiples órganos,

incluidos el cerebro, la piel, el corazón, los riñones y los pulmones. Estas alteraciones pueden dar lugar a epilepsia, dificultades de aprendizaje, problemas de comportamiento e insuficiencia renal, entre otras complicaciones(175, 176). Aproximadamente del 10 al 30% de los casos de esclerosis tuberosa se deben a mutaciones en el gen *TSC1* y la frecuencia de casos debidos a variantes en el gen *TSC2* es consistentemente más alta y se asocian con una enfermedad más grave(177, 178). La linfangiomatosis pulmonar es una enfermedad rara de herencia recesiva que se presenta casi exclusivamente en mujeres con una edad promedio inicio de los síntomas, que incluyen dificultad para respirar (67%), colapso pulmonar (25%), tos (12%) y dolor en el pecho (10%), es de 33 años. Las radiografías de tórax suelen mostrar un infiltrado intersticial difuso, sin predominio en ninguna zona pulmonar(179, 180). La displasia cortical focal tipo 2 es una patología de herencia autosómica recesiva que consiste en una malformación del desarrollo cerebral que resulta en un fenotipo clínico de epilepsia intratable, que generalmente requiere neurocirugía. Los individuos afectados tienen convulsiones refractarias, generalmente con inicio en la primera infancia, y pueden tener discapacidad intelectual persistente (181, 182). En el paciente n° 16 se completó estudio con resonancia magnética cerebral, ecocardiografía y ecografía renal, sin evidenciarse nuevos hallazgos sugestivos de esclerosis tuberosa, en el estudio en tríos se evidenció que el padre era portador asintomático de la misma variante, completándose igualmente estudio de extensión (valoración cardiológica, oftalmológica, dermatológica y resonancia magnética cerebral) que descartó también una esclerosis tuberosa.

**Tabla 26.** Características clínicas y demográficas de los pacientes incluidos en el estudio.

Casos	Edad	Sexo	Fenotipo: HPO
1	25d	M	Retraso del crecimiento intrauterino Distrés respiratorio neonatal Microcefalia congénita Criptorquidia unilateral Hipoglucemia neonatal Acidosis metabólica crónica Acidosis láctica
2	1m 15d	M	Aciduria metilmalónica Anomalía del miocardio Acidosis metabólica crónica Acidosis láctica
3	20d	F	Convulsiones Encefalopatía epiléptica
4	18d	M	Convulsiones Encefalopatía epiléptica
5	25d	F	Convulsiones Encefalopatía epiléptica
6	6d	M	Distrés respiratorio neonatal Convulsiones Encefalopatía epiléptica Acidosis metabólica crónica
7	10d	M	Vómitos Hiperamoniemia Hipoglucemia hipocetósica Pielectasia fetal Exceso de hormona estimulante de tiroides
8	1m	M	Hipotonía neonatal generalizada Convulsiones Letargia Atrofia cerebelosa
9	3d	F	Parto prematuro Insuficiencia respiratoria debida a debilidad muscular Hipotonía generalizada Encefalopatía
10	9d	F	Leucoencefalopatía Cataratas Quistes renales Sangrado digestivo
11	12d	M	Contracturas articulares Microretrognatia Microtia Micropene

Casos	Edad	Sexo	Fenotipo: HPO
			Hipotonía Pérdida auditiva
12	15d	F	Enfermedad de Hirschsprung Pérdida auditiva Opsoclonus
13	22d	F	Preeclampsia Distrés respiratorio neonatal Hipoglucemia
14	29d	M	Pequeño para edad gestacional Ileo meconial Distrés respiratorio neonatal Hipotonía neonatal generalizada Luxación congénita de cadera 10 pares de costillas
15	4d	F	Convulsiones focales
16	2m	F	Fallo de medro Pérdida de peso Irritabilidad Hipotonía axial infantil Regresión neurodesarrollo Anomalía del metabolismo de la serina Hiperglicemia
17	1m 15d	M	Hepatomegalia Hipoglucemia Ascitis
18	4m	M	Hipotonía neonatal generalizada Debilidad muscular
19	2m	M	Bajo peso al nacimiento Hipotonía neonatal generalizada Íleo meconial Fallo de medro Luxación congénita de cadera
20	10d	F	Bajo peso al nacimiento Hipotonía neonatal generalizada Microretrognatia Anquilobefaro Hipotelorismo Distonía de extremidades
21	12d	F	Ductus arterioso persistente Hipertensión pulmonar Hipoplasia pulmonar Parálisis diafragmática Encefalopatía Pie equinovaro bilateral

Casos	Edad	Sexo	Fenotipo: HPO
22	16d	M	Retraso del crecimiento intrauterino Hipotonía neonatal generalizada Criptorquidia bilateral EEG con descargas epilépticas focales Trombosis venosa cerebral
23	7d	M	Ausencia de radio izquierdo Aplasia del primer dedo Defecto septal atrial Pérdida auditiva
24	2m	F	Bajo peso al nacimiento Hernia diafragmática congénita derecha Arteria subclavia derecha aberrante Displasia pancreática Atrofia bilateral coriorretiniana Aplasia/hipoplasia vertebral Clinodactilia bilateral Quiste pineal
25	10d	F	Distrés respiratorio neonatal Piel redundante en cuello Quistes subependimarios Cuello corto Paladar ojival
26	26d	M	Pie equinovaro bilateral Dificultades para la alimentación Alimentación por gastrostomía Reflujo gastroesofágico Hipotonía neonatal generalizada
27	1m	M	Retraso del crecimiento intrauterino severo
28	1m 15d	F	Convulsiones Encefalopatía epiléptica
29	9d	F	Hipotonía generalizada neonatal Encefalopatía Aumento ácido láctico en LCR Edema corneal estromal Fallo hepático
30	13d	F	Hiperinsulinismo difuso Hiperamonemia
31	21d	M	Encefalopatía Cardiomiopatía hipertrófica
32	1m 4d	F	Convulsiones Encefalopatía epiléptica
33	2m	M	Convulsiones Encefalopatía epiléptica

d: días;; F: femenino; HPO: human phenotype ontology; LCR: líquido cefalorraquídeo; m: meses; M: masculino

Tabla 27. Hallazgos demográficos, clínicos y genéticos de los pacientes con variantes causales/probables.

Casos	Edad	Sexo	HPO	TAT	Gen	Variantes	Phenomizer	Patología	MIM	Padres/Herencia
1	25d	M	Retraso del crecimiento intrauterino Distrés respiratorio neonatal Microcefalia congénita Criptorquidia unilateral Hipoglucemia neonatal Acidosis metabólica crónica Acidosis láctica	7 d	<i>GFM1</i>	c.640A>T/ c.1199G>C	p 0.0343	Defecto combinado de la fosforilación oxidativa I	#609060	Portadores Autosómico recesivo
2	1m 15d	M	Aciduria metilmalónica Anomalía del miocardio Acidosis metabólica crónica Acidosis láctica	7 d	<i>SUCLA2</i>	c.850C>T/ c.859C>T	p 0.0271	Síndrome de depleción mitocondrial 5	#612073	Portadores Autosómico recesivo
3	20d	F	Convulsiones Encefalopatía epiléptica	6 d	<i>KCNQ2</i>	c.778C>T	p 0.04	Encefalopatía epiléptica infantil precoz 7	#613720	De novo Autosómico dominante
4	18d	M	Convulsiones Encefalopatía epiléptica	7 d	<i>KCNQ2</i>	c.1016 T>A	p 0.04	Encefalopatía epiléptica infantil precoz 7	#613720	Padre, epilepsia en la infancia Autosómica dominante
5	25d	F	Convulsiones Encefalopatía epiléptica	6 d	<i>CDKL5</i>	c.616G>T	n.s. p 0.1386	Encefalopatía epiléptica infantil precoz 2	#300672	Madre, epilepsia Autosómica dominante ligada a X
6	6d	M	Distrés respiratorio neonatal Convulsiones Encefalopatía epiléptica Acidosis metabólica crónica	7 d	<i>KCNQ2</i>	c.1658G>A	n.s p 0.471	Encefalopatía epiléptica infantil precoz 7	#613720	De novo Autosómica dominante

Casos	Edad	Sexo	HPO	TAT	Gen	Variantes	Phenomizer	Patología	MIM	Padres/Herencia
7	10d	M	Vómitos Hiperamoniemia Hipoglucemia hipocetósica Pielectasia fetal Exceso de hormona estimulante de tiroides	8 d	<i>CPS1</i>	c.1201G>C/ c.2810T>A	n.s. p 0.0624	Deficiencia de CPSI	#608307	Portadores Autosómica recesiva
8	1m	M	Hipotonía neonatal generalizada Convulsiones Letargia Atrofia cerebelosa	10 d	<i>COQ4</i>	c.202G>C/ c.718C>T	n.d.	Deficiencia primaria de Coenzima Q10 tipo 7	#616276	Portadores Autosómica recesiva
9	3d	F	Parto prematuro Insuficiencia respiratoria debida a debilidad muscular Hipotonía generalizada Encefalopatía	10 d	<i>ACTA1</i>	c.614C>A	n.s. p 0.582	Miopatía nemalínica 3	#161800	De novo Autosómica dominante
10	9d	F	Leucoencefalopatía Cataratas Quistes renales Sangrado digestivo	5 d	<i>COL4A1</i>	c.2906 G>A	n.s. p 0.5810	Enfermedad vascular cerebral de vaso pequeño	#175780	De novo Autosómica dominante
11	12 d	M	Contracturas articulares Microretrognatia Microtia Micropene Hipotonía Pérdida auditiva	6 d	<i>ERBB3</i>	c.1184-9A>G/ c.1184-9A>G	n.s. p.0,4243	Síndrome de contractura letal tipo 2	#607598	Portadores Autosómica recesiva
12	15 d	F	Enfermedad de Hirschsprung Pérdida auditiva Opsoclonus	7 d	<i>SOX10</i>	c.850G>T	p.0,0323	Síndrome Waardenburg- Shah	#142623	De novo Autosómica dominante

ACMG: clasificación de la variante según criterios del American College of Medical Genetics; d: día; HPO: human phenotype ontology; F: femenino; m: mes; M: masculino; n.d.: no detectado; n.s: no significativo; TAT: tiempo hasta el diagnóstico

**Tabla 28.** Hallazgos demográficos, clínicos y genéticos de los pacientes con variantes no definitivamente/no probablemente causales.

Casos	Edad	Sexo	HPO	Gen	Variantes	Phenomizer	Patología	MIM	Progenitores/herencia
13	22 d	F	Preeclampsia Distrés respiratorio neonatal Hipoglucemia	<i>MTOR</i>	c.126G>T	n.d.	- Displasia cortical focal tipo II - Síndrome de Smith- Kingsmore	#607341 #616638	Padre portador asintomático
14	29 d	F	Pequeño para edad gestacional Ileo meconial Distrés respiratorio neonatal Hipotonía generalizada neonatal Displasia congénita de cadera 10 pares de costillas	<i>NALCN</i>	c.2507C>G	n.d.	- Contracturas congénitas de manos y cara - Hipotonía y retraso del desarrollo	#616266	Padre portador asintomático
15	4 d	M	Convulsiones focales	<i>TSC2</i>	c.1724T>C	n.d.	- Complejo esclerosis tuberosa - Linfangioleiomatosis pulmonar - Displasia cortical focal tipo II	# 613254	Padre portador asintomático

ACMG: clasificación de la variante según criterios del American College of Medical Genetics; d: día; F: femenino; HPO: human phenoype ontology; M: masculino; n.d.:no detectado; TAT: tiempo hasta el diagnóstico.

### 4.2.3 Pacientes con enfermedad genética no confirmada mediante el panel NeoSeq

Los diagnósticos de síndrome de Prader-Willi y distrofia de Steinert no se detectaron en los casos 20 y 21, respectivamente, pero se establecieron en base a pruebas específicas realizadas más adelante al progresar la sintomatología. El paciente n° 20 era un neonato de sexo femenino y 5 días de vida que nació a las 36+3 semanas de edad gestacional con un peso de 2500 g. A la exploración física destacaba una hipotonía generalizada con insuficiencia respiratoria leve. El estudio genético mediante panel NeoSeq fue negativo. Evolutivamente se observó una diplejía facial con boca en acento circunflejo, presentando la madre también una facies hipomímica en menor grado así como dificultad para la relajación muscular de las manos tras una contracción voluntaria. A los 2 meses de vida se sospecha una distrofia miotónica de Steinert, confirmándose el diagnóstico mediante PCR y TP-PCR específica de la región 3'-UTR del gen *DMPK* (MIM \*605377) que evidenció una expansión de 2333 copias del triplete CTG. La distrofia miotónica de Steinert (MIM #160900) es un trastorno autosómico dominante caracterizado principalmente por miotonía, distrofia muscular, cataratas, hipogonadismo, calvicie frontal y cambios en el ECG(183). La gravedad de la enfermedad varía con el número de repeticiones(184): los individuos normales tienen de 5 a 37 repeticiones, las personas levemente afectadas tienen de 50 a 150 repeticiones, los pacientes con DM clásica tienen de 100 a 1,000 repeticiones y aquellos con inicio congénito pueden tener más de 2,000 repeticiones(185), siendo este el caso del paciente n° 20.

El caso 21 era un neonato a término de sexo masculino y 10 días de vida que nació a las 40+3 semanas de edad gestacional con un peso de 2700 g. Durante el último trimestre de la gestación se evidenció un retraso del crecimiento intrauterino. A la exploración física destacaba

una hipotonía generalizada, microrretrognatia, hipotelorismo, anquilosis palpebral y distonía de extremidades. El estudio genético mediante panel NeoSeq fue negativo. Evolutivamente se observó obesidad, hipogonadismo y criptorquidia bilateral, por lo que a los 4 meses de vida se estableció la sospecha de síndrome de Prader-Willi confirmándose el diagnóstico mediante MS-MLPA, detectándose una deleción de la región 15q11-q13. El síndrome de Prader-Willi (MIM# 176270) se caracteriza por disminución de la actividad fetal, obesidad, hipotonía muscular, retraso mental, baja estatura, hipogonadismo hipogonadotrópico y manos y pies pequeños(186, 187). Puede considerarse un trastorno autosómico dominante y es causado por la deleción o alteración de un gen o de varios genes en el brazo largo proximal del cromosoma 15 paterno o disomía 15 uniparental materna, porque los genes en el cromosoma materno 15 son virtualmente inactivos por impronta.

#### **4.2. IMPACTO DEL DIAGNÓSTICO GENÉTICO SOBRE EL MANEJO CLÍNICO**

Al evaluar el impacto del diagnóstico genético en la evolución clínica, encontramos que el establecimiento de un diagnóstico molecular afectó directamente el manejo médico en 11 de los 12 pacientes (Tabla 6), resultando en cambios en la medicación o dieta (9 pacientes), inicio de nuevas valoraciones por subespecialistas específicos que no se habían considerado previamente (5 pacientes) o retirada del tratamiento de cuidados intensivos con reorientación hacia cuidados paliativos (5 pacientes). En el caso de los pacientes con enfermedades mitocondriales (casos 1 y 2), tras el diagnóstico se instauró un nuevo tratamiento médico con cóctel antioxidante en todos ellos. El manejo de este tipo de patología sigue siendo en gran parte sintomático y no altera significativamente el curso de la enfermedad. Sin embargo, sobre la base de un número limitado de ensayos clínicos,

se han utilizado varios agentes destinados a mejorar la función mitocondrial o tratar las consecuencias de la disfunción mitocondrial, a los que se denomina cóctel antioxidante y que incluye en combinaciones variables coenzima Q10, idebenona, riboflavina, dicloroacetato, tiamina, vitamina C, vitamina E, ácido lipoico, arginina, citrulina...(188, 189) etc. Además, al tratarse de enfermedades multisistémicas, en función del fenotipo previsto se completó estudio y seguimiento por parte de nuevos subespecialistas incluyendo la valoración de otros órganos diana como el ojo, el corazón, el hígado, el oído y el riñón(190). Además, dado el mal pronóstico y el componente degenerativo de estas patologías los pacientes se reorientaron hacia cuidados paliativos, rechazando las familias medidas invasivas de tratamiento.

En cuanto a los tres pacientes con encefalopatía epiléptica infantil precoz 7 (casos 4, 5 y 7) por mutaciones en el gen *KCNQ2* el diagnóstico molecular supuso el inicio de tratamiento anticonvulsivante con agentes bloqueantes de los canales de sodio (carbamacepina y fenitoína). El mencionado grupo de antiepilépticos debe considerarse la primera línea de tratamiento en esta patología, ya que se ha demostrado que permiten un mejor control de las convulsiones refractarias a otras terapias y mejoran el desarrollo neurológico de los pacientes (191, 192). En relación al paciente con encefalopatía epiléptica infantil precoz tipo 2 (caso 6), tras obtener el diagnóstico molecular se decidió iniciar dieta cetogénica (193). En el paciente que se confirmó el diagnóstico de déficit de CPS I se pudo iniciar precozmente el tratamiento con dieta restringida en proteínas naturales, quelantes del amonio y aminoácidos esenciales, además de evitarse las situaciones de estrés metabólico (ayuno, enfermedad concomitante, fiebre, cirugía) que desencadenan episodios de descompensación hiperamoniémica y comprometen el desarrollo neurológico de estos pacientes(194). En cuanto al paciente con

deficiencia primaria de coenzima Q10 tipo 7 (caso 9), tras el diagnóstico molecular se inició un nuevo tratamiento médico con coenzima Q10 a dosis altas, ya que se ha descrito que este tratamiento en los pacientes con formas primarias de deficiencia de coenzima Q10 puede detener la progresión tanto de síndrome nefrótico córticorresistente (no reportado en el tipo 7) como de la encefalopatía, y disminuir el riesgo de accidentes cerebrovasculares *stroke like*, de ahí la importancia crítica de un diagnóstico rápido(195). Además, se instauró seguimiento por cardiología, ya que la miocardiopatía hipertrófica es otra de las manifestaciones principales de esta enfermedad (152, 196).

En relación al paciente con miopatía nemalínica (caso 10), tras alcanzar el diagnóstico molecular se inició terapia empírica con piridostigmina y carnitina, si bien se trata de una patología que carece de tratamiento específico o curativo(197), exceptuando los casos de debut rápidamente progresivo en la edad adulta que suelen deberse a procesos inmunes (196). Debido a la falta de respuesta al tratamiento médico y al mal pronóstico de la patología, el paciente se reorientó hacia cuidados paliativos y la familia rechazó terapias invasivas. El paciente con vasculopatía cerebral de pequeño vaso (caso 11) falleció a la semana de recibir el diagnóstico molecular sin instaurarse ningún tratamiento específico ni iniciarse nuevos estudios de extensión o seguimiento por parte de otros especialistas. En los casos 12 y 13 con los diagnósticos de síndrome de contractura letal tipo 2 y Síndrome Waardenburg-Shah el diagnóstico genético motivó la reorientación de su manejo hacia cuidados paliativos, al tratarse de dos patologías muy graves sin tratamiento específico o curativo.

Además, en todos los casos los pacientes y sus familias se beneficiaron de un asesoramiento genético oportuno basado en un diagnóstico molecular concreto.

Tabla 29. Impacto del diagnóstico molecular en el manejo médico

Casos	Patología	Cambios en medicación o dieta	Seguimiento por un nuevo subespecialista	Cuidados paliativos	Consejo genético
1	Defecto de la fosforilación oxidativa tipo 1	Cóctel antioxidante	Neurólogo Hepatólogo Cardiólogo	Si	Si
2	Síndrome de depleción mitocondrial 5	Cóctel antioxidante	Neurólogo Audiólogo Oftalmólogo Nefrólogo	Si	Si
3	Encefalopatía epiléptica infantil precoz 7	Bloqueante de los canales de sodio: fenitoína Evitar retigabina and ezogabina	No	No	Si
4	Encefalopatía epiléptica infantil precoz 7	Bloqueante de los canales de sodio: carbamacepina Evitar retigabina and ezogabina	No	No	Si
5	Encefalopatía epiléptica infantil precoz 2	Dieta cetogénica	Dietista	No	Si
6	Encefalopatía epiléptica infantil precoz 7	Bloqueante de los canales de sodio: carbamacepina Evitar retigabina y ezogabina	No	No	Si
7	Deficiencia CPS I	Dieta restringida en proteínas Arginina/citrulina Quelantes de amonio Evitar ácido valproato	Especialista en errores congénitos del metabolismo	No	Si
8	Deficiencia primaria de coenzima Q10 tipo 7	CoenzimaQ10	Cardiólogo	No	Si
9	Miopatía nemalínica	Piridostigmina Carnitina	No	Si	Si
10	Vasculopatía cerebral de pequeño vaso	- (exitus)	- (exitus)	- (exitus)	Si
11	Síndrome de contractura letal tipo 2	No	No	Si	Si
12	Síndrome Waardenburg-Shah	No	No	Si	Si

## 5. DISCUSIÓN

La secuenciación genómica se ha convertido en una de las herramientas de diagnóstico más prometedoras en el campo de la neonatología. Las características que presenta este tipo de abordaje, con capacidad para secuenciar gran cantidad de ADN (exomas clínicos, exomas completos o genomas completos), además de hacerlo con un tiempo de respuesta de días, han demostrado ser de especial importancia para su aplicación en recién nacidos gravemente enfermos con sospecha de enfermedad genética. Varios estudios realizados hasta la fecha evidencian que la incorporación de esta tecnología como prueba de primer nivel para el diagnóstico genético en este grupo de pacientes aumenta la tasa de diagnóstico y reduce el tiempo hasta el mismo, mejorando los resultados y reduciendo los costes sanitarios (198, 199). Debe señalarse que los trastornos genéticos que debutan en el período neonatal tienen un impacto significativo en la morbilidad y la mortalidad de los recién nacidos, de modo que el diagnóstico precoz que ofrece la secuenciación genómica facilita también el asesoramiento pronóstico oportuno a las familias y permite realizar una medicina de precisión. A medida que se amplía la disponibilidad de pruebas de diagnóstico genético, aumenta también de forma significativa la complejidad de las mismas así como la necesidad de un conocimiento especializado por parte del neonatólogo, incluyendo la aplicación y comprensión de las modalidades de pruebas genéticas y de sus resultados. Esto es especialmente cierto en aquellos centros donde la disponibilidad de consultas de genética clínica es limitada.

En el presente estudio, evaluamos la eficacia diagnóstica y el impacto clínico de NeoSeq, un panel que consta de 1870 genes y está diseñado específicamente para facilitar el diagnóstico de recién nacidos críticamente enfermos con sospecha de trastorno genético. Finalmente, comparamos nuestros resultados con los publicados previamente en estudios prospectivos en los que se aplicaron técnicas de secuenciación masiva (WGS, WES, CES) a los recién nacidos ingresados en UCIN con sospecha de trastorno genético (Tabla 17). Nuestra comparación incluyó 14 estudios publicados entre 2016 y 2020. Doce eran estudios de cohortes y dos eran ensayos clínicos aleatorios. Los estudios seleccionados aplicaron las siguientes técnicas de secuenciación: CES (cuatro estudios) (102-104, 114), WES (seis estudios) (58, 115, 116, 120, 121, 126), WGS (tres estudios) (122, 127, 128) y WES + WGS (un estudio) (129). El número total de pacientes incluidos fue 1038. La mediana de la tasa de diagnóstico fue 37.2% (rango, 13.2% -56%). Después de ajustar por la tecnología NGS empleada, se obtuvieron las siguientes tasas de diagnóstico: CES 43.2% (rango 32.4% - 50.8%), WES 31.2% (15.6% - 56%), WGS con filtrado posterior por CES 30.4% y WGS 29.8% (rango 13.2% - 45.8%). La mediana del tiempo de respuesta al diagnóstico fue de 30.2 días.

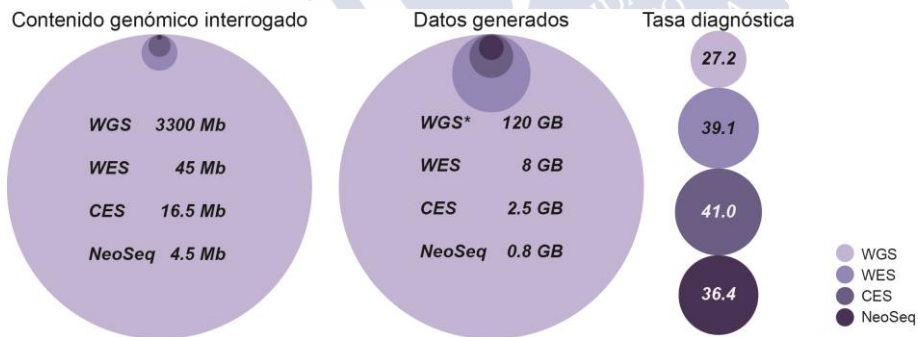
**Tabla 30.** Estudios prospectivos del rendimiento diagnóstico de WES, WGS y CES en recién nacidos con sospecha de trastorno genético

Referencia	Fecha	Tipo de estudio	Test de secuenciación	Tamaño muestral	Tasa diagnóstica	Tiempo de respuesta
Stark, et al. Genetic in Medicine (116)	2016	Cohortes	WES	n=33	16/33 (48%)	134 d
Daoud, et al. CMAJ (102)	2016	Cohortes	CES	n=8	4/8 (50%)	15.2 s
Meng, et al. JAMA Pediatrics (103)	2017	Cohortes	CES Trio CES Critical trio CES	n=178 n=37 n=63	58/178 (32.6%) 12/37 (32.4%) 32/63 (50.8%)	13 d
van Diemen, et al. Pediatrics (114)	2017	Cohortes	WGS filtrado por CES	n=23	7/23 (30.4%)	12 d
Petrikina, et al. Genomic Medicine (127)	2018	RCT	Trio rWGS	n=64	21/64 (32.8%)	13 d
Stark, et al Genetics in Medicine (115, 116)	2018	Cohortes	rWES	n=40	21/40 (52.5%)	16 d
Brunelli, et al. MGGM (104)	2019	Cohortes	rCES	n=20	10/20 (50%)	9.6 d
Ceyhan-Birsoy, et al. AJHG (58)	2019	Cohortes	WES	n=32	5/32 (15.6%)	16 s
Elliot, et al. EJP (120)	2019	Cohortes	WES	n=25	14/25 (56%)	7.2 d
French, et al. Intensive Care Medicine (122)	2019	Cohortes	Trio WGS	n=106	14/106 (13.2%)	21 d
Kingsmore, et al. AJHG (129)	2019	RCT	rWGS rWES urWGS	n=94 n=95 n=24	18/94 (19.1%) 19/95 (20%) 11/24 (45.8%)	11 d 11.2 d 4.6 d
Wang, et al. Human Genetics(128)	2020	Cohortes	Trio WGS	n=84	32/84 (38%)	94 h
Gubbels, et al. Genetic Medicine (121)	2020	Cohortes	Trio WES	n=50	29/50 (26%)	4.9 d
Lunke, et al .JAMA (126)	2020	Cohortes	Trio WES	n=62	35/62 (56%)	3.3 d

CES: secuenciación de exoma clínico; d: días; h: horas; RCT: ensayo clínico randomizado; r: rápido; s: semanas; ur: ultrarápido; WES: secuenciación de exoma completo; WGS: secuenciación de genoma completo.

En total, 33 recién nacidos y sus respectivos progenitores se sometieron a secuenciación en trío utilizando el panel NeoSeq. Los recién nacidos con síntomas neurológicos constituyeron el grupo más grande (n = 17), seguidos por aquellos con múltiples anomalías congénitas (n = 8) y sospecha de enfermedades metabólicas (n = 7). Estos grupos fenotípicos se corresponden con los descritos previamente en otros estudios que utilizaron estrategias NGS rápidas para diagnosticar a los pacientes de la UCIN (102, 121, 126).

La tasa diagnóstica resultante (12/33 pacientes [36,4%]) fue comparable a la descrita anteriormente para enfoques más amplios de secuenciación de trío o de probando, como WES y WGS con unas tasas diagnósticas medias de 31.2% y 29.8% respectivamente, e incluso más alto que el informado para WGS en pacientes ingresados en UCIN(122).



**Figura 6. Comparación del panel NeoSeq con otras estrategias diagnósticas de secuenciación genética.** Se representa la porción genómica interrogada, la cantidad de datos generados y las tasas de diagnóstico logradas. Para facilitar la comparación visual, los datos se representan como el área relativa de cada círculo. Para WGS y WES, la información sobre la porción genómica interrogada y los datos generados se extrajeron de [www.illumina.com/wes-wgs](http://www.illumina.com/wes-wgs). Para CES, la información sobre la porción genómica interrogada y los datos generados se estimó para 6000 genes. Las tasas de diagnóstico se calcularon como la media de los resultados de los estudios citados. Adaptado de de Castro MJ, et al. J Clin Med 2020 23;9(8):2362 (200).

Si la única variable asociada al rendimiento diagnóstico fuera la cantidad de ADN secuenciado sería esperable que las tasas diagnósticas más altas se obtuvieran con el WGS, seguidas por WES y CES, y correspondiendo las más bajas a paneles genéticos específicos de una sospecha diagnóstica. Sin embargo, revisando los estudios publicados hasta la fecha encontramos que dicho rendimiento no se correlaciona con el tamaño de la porción del genoma secuenciado (Figura 6) ya que NeoSeq proporcionó tasas de diagnóstico similar a las obtenidas con los enfoques WGS y WES, a pesar de interrogar 10 veces (WES) y 500 veces (WGS) menos pares de bases. Esto indica con claridad que deben existir otras variables que están influyendo en la tasa diagnóstica.

Tres factores pueden ayudar a explicar esta aparente paradoja. Primero, durante la priorización de las variantes en función de su potencial relación con el fenotipo del paciente generalmente se excluyen aquellas para las que no se puede predecir una consecuencia funcional (5, 86, 201, 202). En este punto es preciso tener presente que cualquiera de estas aproximaciones de secuenciación genómica va a generar, después de los procesos de anotación y filtrado, un listado con cientos o miles de variantes genéticas. Priorizar entre estas variantes aquellas que podrían estar relacionadas con el fenotipo del paciente va a requerir de entrada seleccionar aquellas que el conocimiento actual nos permite asociar con un efecto biológico plausible. Esto necesariamente va a limitar el análisis a variantes exónicas e intrónicas en genes conocidos asociados a enfermedades o variantes con patogenicidad demostrada localizadas en regiones no codificantes. En consecuencia, independientemente de la cantidad de ADN interrogada, el fenotipo del recién nacido va a orientar el análisis a un grupo reducido de genes candidatos, y dentro de estos genes, a aquellas regiones en las que es posible predecir el efecto de la variante. Por lo tanto, independientemente de cuán extenso sea el

análisis genómico, el diagnóstico genético se ciñe a una lista corta de genes (114, 129) y de tipos de variantes (128). Nuestros resultados y los de otros autores que utilizan una aproximación similar (104), demuestran que un enfoque más simple, basado en el análisis de paneles de genes con funciones conocidas y asociaciones de enfermedades, puede ser una alternativa útil a los enfoques WGS o WES, produciendo tasas de diagnóstico comparables. En la práctica, lo que propone este tipo de aproximación es un filtrado previo aplicado al diseño del panel, limitándolo a aquellos genes y regiones potencialmente asociadas a enfermedad de debut neonatal, en lugar de aplicar un filtrado posterior a la secuenciación, como sucede al aplicar aproximaciones que interrogan una cantidad de ADN mayor, como el WGS o el WES.

El segundo factor a considerar es la asociación entre la tasa de diagnóstico y la gravedad del fenotipo clínico de los participantes del estudio. Todos los pacientes reclutados en nuestro estudio eran recién nacidos críticamente enfermos que requerían cuidados intensivos. En dos estudios anteriores en los que se asignó a pacientes con enfermedad grave y menos grave a distintas cohortes, se observaron mejores resultados para el grupo gravemente enfermo en términos de tasa de diagnóstico y de impacto en el manejo clínico. En el ensayo NSIGHT2 (127), los pacientes se asignaron al azar a dos grupos, que se analizaron utilizando WES o rWGS, excepto para aquellos pacientes críticamente enfermos, que fueron evaluados mediante WGS ultrarrápido (urWGS). La tasa de diagnóstico resultante fue del 46% en el grupo urWGS, en comparación con 19% y 20% en los grupos rWGS y rWES, respectivamente. Meng, et al. (103) utilizaron WES para realizar la secuenciación de exoma probando, exoma trío y exoma trío crítico (este último test más rápido se aplicó a los pacientes más gravemente enfermos) en tres grupos de lactantes dentro de los primeros 100 días de vida. La tasa de diagnóstico en el grupo de

exoma trío crítico fue mayor que la obtenida en el resto de los grupos. Además, los diagnósticos moleculares afectaron directamente al manejo médico en un 76.9% de los pacientes en la cohorte de exoma trío crítico, en comparación con el 42.9% en los otros grupos. Estos resultados sugieren que una estrategia de diagnóstico basada en la secuenciación de un panel de genes asociados con fenotipos neonatales graves produce mejores resultados porque en estos pacientes la probabilidad de padecer un trastorno genético es mayor.

El tercer factor a tener en cuenta y que aún no está completamente definido es cómo sistematizar la selección de los pacientes que tienen más probabilidades de padecer una enfermedad de causa genética y beneficiarse de técnicas de NGS y, por lo tanto, cuales son los criterios de inclusión de los pacientes en estos estudios. A pesar de que estos criterios suelen estar avalados por la opinión de expertos y la evidencia científica, a menudo se encuentra una gran variabilidad en los mismos. Se ha propuesto que la selección de pacientes basada en el fenotipo podría ser eficaz para identificar a aquellos recién nacidos críticamente enfermos con una alta probabilidad de padecer una enfermedad genética y por lo tanto, aquellos recién nacidos que más se podrían beneficiar de un estudio genómico. A partir de los resultados del estudio BabySeq, en el que se utilizaba la secuenciación genómica WES de manera aleatorizada a un grupo de pacientes ingresados en UCIN sin la aplicación de ningún criterio de inclusión basado en el fenotipo y que únicamente consiguió alcanzar un diagnóstico definitivo en 1 de los 29 casos reclutados (58), los autores Gubbels, et al. (121) decidieron desarrollar unos criterios basados en el fenotipo para la selección de neonatos en la UCI que se iban a someter a pruebas de diagnóstico molecular. De este manera, los criterios de elegibilidad propuestos incluyeron una edad corregida <6 meses, ingreso en la UCI (o en espera de transferencia a la UCI) y presentación reciente con una o más de las siguientes características

clínicas inexplicables: (1) convulsiones, (2) hipotonía, (3) anomalías congénitas múltiples, (4) fenotipo metabólico complejo (definido como la presencia de características clínicas o bioquímicas sospechosas de un trastorno metabólico, con más de una condición posible en el diagnóstico diferencial, que no podría distinguirse con una prueba de laboratorio fácilmente disponible). Estos cuatro criterios representaron los criterios de elegibilidad fenotípicos básicos utilizados para la inscripción de los primeros 20 casos. Después de los primeros 20 casos, se añadieron fenotipos clínicos adicionales: (5) trastornos del desarrollo sexual, (6) enfermedad pulmonar intersticial e (7) inmunodeficiencia. Los criterios de exclusión comprendían los siguientes i. presencia de una posible explicación no genética para el fenotipo (por ejemplo, asfixia perinatal en relación a hipotonía); ii. características clínicas patognomónicas de una anomalía cromosómica reconocible, incluidas las trisomías 13, 18 y 21; y iii asociaciones ya conocidas por tener un bajo rendimiento de diagnóstico genético, como la asociación de defectos vertebrales, atresia anal, defectos cardíacos, fístula tráqueo esofágica, anomalías renales y anomalías de las extremidades (VATER / VACTERL) y el complejo onfalocele, extrofia vesical y rectal, ano imperforado y defectos espinales (OEIS). Un criterio de exclusión adicional para los propósitos del estudio fue iv. una contraindicación para la recolección de muestras en el niño o en uno de los padres. Aplicando estos criterios a la selección de neonatos críticamente enfermos obtuvieron una de las tasas diagnósticas más altas en este tipo de estudios (56%) concluyendo que la selección de pacientes basada en el fenotipo es uno de los factores que más influyen en el rendimiento de las pruebas de diagnóstico molecular.

Así, podemos concluir que la tasa diagnóstica de las pruebas de secuenciación genómica aplicadas a neonatos ingresados en UCIN está influida por varios factores que deben ser tenidos en cuenta a la

hora de planificar la implementación de este tipo de pruebas en la rutina clínica. A partir del análisis de la presentación clínica del paciente, se debe priorizar la realización de este tipo de pruebas en aquellos neonatos en los que la probabilidad de padecer una enfermedad genética es mayor. Para esto es de gran ayuda la sistematización de la información fenotípica mediante el uso de herramientas como los términos HPO (*Human Phenotype Ontology*). Por otro lado, se debe priorizar el análisis de las variantes en aquellos genes asociados a enfermedad grave de debut neonatal. Esta priorización se podrá realizar antes de la secuenciación, como en el caso del panel NeoSeq, o en la fase de priorización de variantes, con posterioridad a la secuenciación en el caso de utilizar WES o WGS.

En nuestro estudio, aplicamos un enfoque de análisis en trío (secuenciación del recién nacidos y de ambos progenitores), ya que se trata de una estrategia que ha demostrado su eficacia a la hora de gestionar la riqueza de variantes genéticas identificadas por secuenciación genómica. Este tipo de enfoque permite estudiar de manera rápida el patrón de herencia asociado a las variantes identificadas en el paciente índice y confirmar su concordancia con el tipo de herencia asociado a cada gen. De esta manera la estrategia en tríos permite, por ejemplo, identificar fácilmente las variantes *de novo*, filtrar variantes familiares benignas raras y establecer patrones de herencia en los trastornos recesivos (58, 86). En nuestro estudio identificamos cinco variantes causales *de novo* entre nuestros pacientes e, igualmente importante, en tres casos (pacientes 14, 15 y 16), descartamos posibles variantes asociadas al fenotipo después de detectarlos en un solo padre asintomático.

En nuestra cohorte, hubo dos pacientes en los que el empleo de NeoSeq no pudo establecer una causa molecular y que fueron diagnosticados posteriormente con síndrome de Prader-Willi y distrofia miotónica tipo I o enfermedad de Steinert, respectivamente.

Estos dos pacientes son dos ejemplos de las limitaciones diagnósticas que presentan las herramientas basadas en secuenciación NGS. Por una parte, la secuenciación NGS es incapaz de detectar defectos en la impronta genética como el causante del síndrome de Prader-Willi, dado que el mecanismo molecular detrás del mismo es la metilación de la región crítica Prader-Willi en el cromosoma 15 paterno, requiriéndose técnicas específicas para su detección. Por otro lado, el hecho de que estas herramientas se basen en la secuenciación masiva en paralelo de fragmentos cortos de ADN, hacen que no sea útil en la identificación de fragmentos altamente repetitivos, como ocurre en las enfermedades asociadas a expansiones de tripletes, como es el caso de la distrofia miotónica tipo I, causada por la expansión del triplete CTG en el gen *DMPK*(105, 203).

Además de la tasa de diagnóstico del 36,4% (12/33 pacientes), el diagnóstico molecular rápido logrado con NeoSeq (tiempo medio de respuesta al diagnóstico, 7,5 días) supuso un cambio en el tratamiento médico en el 91,7% de los pacientes, como se refleja en cambios en la medicación o la dieta (11/12 pacientes), inicio de nuevo seguimiento por un subespecialista (5/12 pacientes), o retirada del tratamiento de cuidados intensivos / redireccionamiento a cuidados paliativos (5/12 pacientes). Los resultados clínicos de este estudio se pueden clasificar en cuatro categorías de acuerdo con el pronóstico de los pacientes.

El primer tipo es el resultado clínico más optimista en el que se puede iniciar un tratamiento curativo o que mejora significativamente el pronóstico del paciente evitando morbilidad y discapacidad. Este es el caso de algunas enfermedades genéticas que afectan al metabolismo, incluso aquellas con pronóstico grave, que pueden lograr mejores resultados con la instauración de una dieta específica al establecerse un diagnóstico temprano. Se trata de los resultados más alentadores para los médicos y familias de pacientes cuando se recibe el diagnóstico genético. El caso 8 era un recién nacido de 10 días de

vida con hiperamoniemia, vómitos e hipoglucemias hipocetósicas en relación con un déficit de CPS 1. En este paciente tanto la leche materna como la fórmula de lactancia artificial o la nutrición parenteral con proteínas eran deletéreas, desencadenando episodios de hiperamoniemia que comprometen el desarrollo neurológico. La clínica se resolvió con la instauración de dieta restringida en proteínas, suplementación con arginina y citrulina y administración de fármacos quelantes del amonio. Además, su diagnóstico permitió evitar las situaciones desencadenantes de estrés metabólico como el ayuno o instaurar regímenes de emergencia apropiados en situaciones de riesgo de descompensación (por ejemplo, vómitos, fiebre y cirugías programadas). De la misma manera, los casos 4, 5 y 7, con encefalopatía epiléptica secundaria a alteraciones en el gen *KCNQ2* se pudieron beneficiar del tratamiento con anticonvulsivantes específicos, ya que los bloqueantes de los canales de sodio han demostrado ser de utilidad en el control de las convulsiones en la encefalopatía epiléptica asociada a *KCNQ2*, disminuyendo además la afectación a nivel del neurodesarrollo que se asocia a esta patología y mejorando significativamente, por tanto, el pronóstico de estos pacientes cuando reciben un diagnóstico temprano que permite el inicio precoz del tratamiento.

El segundo tipo de resultados clínicos incluiría aquellos que pueden guiar al equipo médico hacia un plan de tratamiento rápido y efectivo para controlar o tratar de mejorar de manera efectiva los resultados antes de que la sintomatología se haga manifiesta. No es el caso de ninguno de los pacientes de este estudio porque incluimos a neonatos críticamente enfermos que ya presentaban signos y síntomas de una enfermedad grave.

El tercer tipo de resultados clínicos comprende aquellos que pueden guiar una intervención clínica, seguimiento especializado y/ o tratamiento sintomático, pero con un mal pronóstico. Este es el caso

de los recién nacidos con enfermedades mitocondriales (casos 1 y 2) en los que el cóctel antioxidante con diferentes vitaminas y cofactores pueden aliviar algunos de los síntomas clínicos asociados a la enfermedad y, teóricamente, puede ofrecer un medio para evitar la progresión de la enfermedad. Los objetivos clínicos actuales para el tratamiento de estas patologías son mejorar el funcionamiento del sistema de fosforilación oxidativa (OXPHOS), aumentando por un lado la producción de energía en forma de adenosina trifosfato (ATP) y reduciendo al mismo tiempo la producción de radicales libres de oxígeno (ROS) en un esfuerzo por mejorar, o al menos estabilizar, los signos y síntomas de la enfermedad. Sin embargo, la gran mayoría de los estudios que evalúan la eficacia antioxidante en la enfermedad mitocondrial han sido ensayos a corto plazo no aleatorizados y todavía no se dispone de ningún medio definitivo para detener la progresión de la enfermedad, y la capacidad clínica para predecir la respuesta terapéutica, lograr la dosificación óptima de medicamentos y cuantificar los beneficios en pacientes individuales sigue siendo limitada (204, 205). Además, el diagnóstico permite evitar situaciones de descompensación en este tipo de pacientes, especialmente cuando se exponen a factores estresantes médicos, incluido el ayuno y medicaciones tóxicas como el valproato, las estatinas, la metformina, altas dosis de paracetamol o determinados antibióticos (206, 207).

Finalmente, el cuarto tipo de resultados clínicos incluye aquellos que conducen a una acción médica que dirige al paciente hacia cuidados paliativos o hacia la limitación del esfuerzo terapéutico, dada la gravedad de la patología de base y el pronóstico infausto de la misma, sin posibilidad de instaurar medidas modificadoras de la enfermedad. Este es el caso de los pacientes 11, 12 y 13 de nuestro estudio, afectados de vasculopatía cerebral de pequeño vaso, síndrome de contractura letal tipo 2 y síndrome de Waardenburg-Shah en los que tras conocer el diagnóstico se optimizaron los cuidados centrados

en el confort, evitándose procedimientos invasivos. Sin embargo, incluso para este tipo de diagnósticos, la familia se beneficia de recibir una respuesta a por qué su hijo estaba enfermo, poniendo fin a una odisea diagnóstica. Además, el diagnóstico molecular permite realizar pruebas a otros miembros de la familia en riesgo y prestar asesoramiento sobre reproducción cuando sea necesario para disminuir o evitar el riesgo de afectación en gestaciones futuras.

El análisis de la tasa de diagnóstico por indicación muestra que los neonatos con clínica neurológica y sospecha de errores innatos del metabolismo fueron los que se beneficiaron en mayor medida del uso del panel NeoSeq (rendimiento diagnóstico del 35.3% y el 42.9%, respectivamente). Por el contrario, solo en dos pacientes del grupo de anomalías congénitas se alcanzó un diagnóstico mediante la identificación de una variante causal definitiva (28.6%). Estos resultados son similares a los alcanzados en otros estudios según los grupos fenotípicos incluidos, con un mayor rendimiento para los neonatos con clínica neurológica y un rendimiento inferior para aquellos con malformaciones congénitas múltiples. Así, en los dos estudios con tasas diagnósticas más altas (56%) (121, 126) el mejor rendimiento diagnóstico se observó en pacientes con manifestaciones neurológicas. En el estudio de Gubbels et al. (121) se alcanzó un diagnóstico molecular en el 80% de los pacientes que presentaban hipotonía y en el 90% de los que presentaban convulsiones, mientras que el grupo de pacientes con malformaciones congénitas tuvo un rendimiento diagnóstico del 64%.

Como hemos comentado con anterioridad, una de las herramientas que más utilidad presentan en la sistematización del análisis fenotípico de los pacientes con sospecha de enfermedad genética en general y de los neonatos críticamente enfermos en particular, es el uso de los términos HPO para clasificar los síntomas. El uso de dichos términos permite realizar un análisis informático de asociación con los

diagnósticos más probables mediante el uso de programas informáticos como Phenomizer. En nuestra cohorte, Phenomizer predijo un diagnóstico genético basado en fenotipo con significación estadística ( $p < 0,05$ ) en cinco pacientes (41.6%). Esta tasa de predicción del diagnóstico es similar a la reportada por Brunelli, et al.(104) utilizando un panel de genes específicos (40%), y superior al informado por French et al. (122) (10%) utilizando WGS. En base a sus datos, French et al. propusieron que el análisis de la información generada mediante secuenciación genómica debía ser agnóstico, dado que al no ser posible dirigir el análisis desde el fenotipo podría perderse información valiosa que llevase al diagnóstico genético de los pacientes. Sin embargo, los análisis mediante el uso de WGS o WES no puede ser completamente agnósticos debido a las limitaciones antes mencionadas que afectan a la interpretación de las variantes.

En nuestro estudio no se identificó ningún hallazgo incidental ya que al ceñirse la secuenciación a un grupo de genes causantes de patología grave en neonatos se evita la posibilidad de diagnosticar patología de debut en la edad adulta. Sin embargo, los hallazgos accidentales son mucho más frecuentes en técnicas de secuenciación no focalizadas, como CES, WES y especialmente WGS, lo que supone una dificultad ética añadida a la hora de informar sobre dichos hallazgos en pacientes menores de edad y vulnerables. Esto se puso en evidencia durante el desarrollo del Proyecto BabySeq, el cual fue un estudio financiado por los Institutos Nacionales de Salud de Estados Unidos y tenía como objetivo explorar los impactos médicos, conductuales y económicos de la integración de la secuenciación genómica en la atención de recién nacidos sanos y recién nacidos enfermos. El protocolo y el consentimiento especificaron que solo se informarían las variantes asociadas a enfermedades de debut en la infancia. Cuando se descubrió que un niño portaba una mutación en el

gen *BRCA2* a pesar de tener antecedentes familiares negativos, el equipo de investigación se enfrentó a un dilema ético dado que no se trataba de uno de los diagnósticos informables según el protocolo del estudio y solicitó el permiso de la junta de revisión institucional para poder informarlo. A partir de este caso, se modificaron tanto el protocolo como el consentimiento informado para recoger la aceptación de los participantes de ser informados de las afecciones graves de debut en la edad adulta.

En el Proyecto Babyseq, se reclutaron 127 neonatos sanos y 32 críticamente enfermos, a los que se aplicó secuenciación genómica mediante WES. En estos sujetos, los resultados de dicha secuenciación revelaron la presencia de variantes genéticas asociadas a riesgo de enfermedad de inicio en la niñez en 15/159 (9,4%) de los participantes; en ninguno de los casos dicho riesgo se podría haber anticipado en función de los antecedentes clínicos o familiares conocidos de los bebés. También se identificaron variantes genéticas asociadas con riesgo de enfermedad de inicio en la edad adulta en 3/85 (3,5%) recién nacidos cuyos padres dieron su consentimiento para recibir esta información. El estado de portador de enfermedades de herencia autosómica recesiva y la presencia de variantes de naturaleza farmacogenómica se informó en el 88% y el 5% de los recién nacidos, respectivamente.

De la misma manera en el proyecto NEXUS (que realizó análisis de exoma completo a 45 recién nacidos con enfermedades metabólicas y con hipoacusia detectadas mediante cribado neonatal) se identificó a un subconjunto de padres que fue elegible para recibir información adicional sobre afecciones de inicio en la infancia con baja o nula capacidad de acción clínica, afecciones de inicio en la edad adulta y estado de portador de afecciones autosómicas recesivas. Se encontraron variantes patogénicas asociadas con cáncer de mama y / o de ovario hereditario en dos niños, una variante patogénica probable

en el gen asociado con el síndrome de Lowe en un niño y un promedio de 1.8 variantes notificables por niño para los resultados de portador (208).

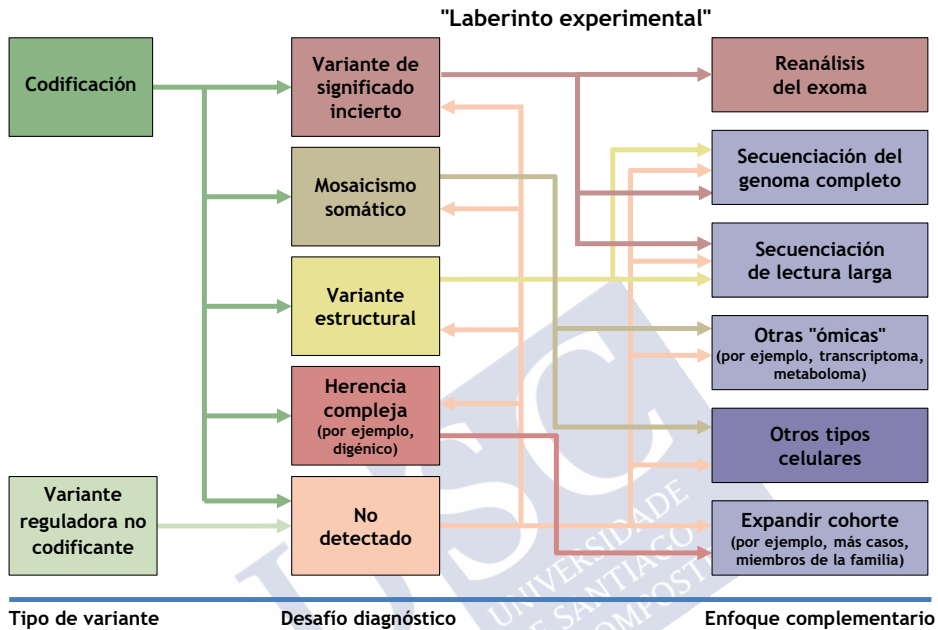
A raíz del Proyecto Babyseq y sus hallazgos, se generó una gran controversia con respecto a las pruebas genéticas predictivas en niños para detectar condiciones de inicio en la edad adulta. En este sentido se concluyó finalmente que los investigadores deben diseñar sus estudios pediátricos para evitar, cuando sea posible, la identificación de variantes genéticas asociadas a enfermedades de inicio en la edad adulta y que no se debe ofrecer a los padres esta información si se descubre, a menos que sea relevante para la salud actual o inminente del niño. Además de los menores requerimientos técnicos y económicos de los paneles de genes como NeoSeq, otra ventaja es el menor riesgo de encontrar hallazgos incidentales, ya que los genes se han seleccionado previamente para correlacionarse con las enfermedades de interés con manifestación en el periodo neonatal.

Traducir la experiencia de un estudio de investigación realizado, como en este caso, en un solo centro a su posterior implementación generalizada en todos los sistemas de atención médica del sistema público de salud, requiere del desarrollo de protocolos clínicos y de laboratorio sólidos, que permitan ofrecer pruebas genómicas rápidas de manera confiable y ofreciendo resultados consistentes y reproducibles. Además, la difusión y ampliación con éxito de las innovaciones en los sistemas de atención de la salud es compleja y precisa de una estructura y adaptabilidad al contexto local. La aplicación de un panel de fácil interpretación para neonatos críticamente enfermos ingresados en UCIN que además requiere de menores recursos técnicos podría constituir una alternativa a pruebas genéticas más complejas.

Uno de los retos que se presentan en el uso de la secuenciación genómica es como aumentar del rendimiento diagnóstico de estas

pruebas en los neonatos críticamente enfermos con sospecha de enfermedad genética, ya que como se refleja en la Figura 7, independientemente de cada aproximación utilizada, la tasa diagnóstica en ningún caso supera el 60%. Los dos estudios con mayor rendimiento diagnóstico hasta la fecha realizados en esta población de pacientes confirmaron la causa molecular de la enfermedad en el 56% de los neonatos empleando en ambos casos una aproximación con WES (120, 126). En el caso del estudio de Lunke et al. once diagnósticos (20% de los pacientes en los que se encontró una causa molecular de su patología) resultaron del uso de los siguientes enfoques para aumentar la capacidad de análisis de secuenciación del WES: análisis de secuenciación del genoma mitocondrial, análisis del número de copias basado en la secuenciación del exoma, uso de bases de datos internacionales para identificar nuevas asociaciones entre genes y enfermedades, y análisis adicional de fenotipado y ARN. Hoy en día es cada vez más frecuente la incorporación del ADN mitocondrial a las pruebas de diagnóstico genético por secuenciación genómica, sobre todo en pacientes en edad pediátrica, aunque este tipo de aproximación puede verse limitada por el hecho de que el ADN extraído de sangre periférica no sea la muestra más apropiada para realizar estudios de ADN mitocondrial. En cuanto al análisis de variantes de número de copias (CNV), en la actualidad los pipelines bioinformáticos suelen incorporar este tipo de análisis, que permite identificar ganancias o pérdidas de material genético a partir de los datos de profundidad de lectura para cada posición secuenciada. Es fundamental, como refleja por otro lado este estudio, la existencia de bases de datos de asociación genotipo fenotipo actualizadas que permitan compartir el conocimiento generado y ayuden en la interpretación de los resultados de la secuenciación genómica. En este sentido, para los casos en los que las técnicas de secuenciación masiva no proporcionan un diagnóstico, se pueden aplicar múltiples

enfoques, pero elegir el siguiente paso en un "laberinto experimental" sigue siendo un gran desafío (209).



**Figura 7. Retos y aproximaciones diagnósticas en pacientes con sospecha de enfermedad genética sin confirmación mediante NGS.** Dependiendo de la variante causal hay diferentes opciones disponibles, pero sin conocer la causa de la enfermedad la aproximación para seleccionar la herramienta diagnóstica adecuada es más compleja. Figura adaptada de Frésard L, et al. Cold Spring Harb Mol Case Stud 2018 17;4(6):a003392 (209).

Los resultados negativos de las técnicas de secuenciación masiva se pueden explicar de varias formas (209). La variante causal puede pasarse por alto en casos de mosaicismo somático (210); si se encuentra en regiones de codificación pero no se detecta adecuadamente(211); si es de importancia desconocida y no se seleccionó (212, 213); si múltiples variantes son responsables de la patología, en dos o más genes (214); si se trata de una variante estructural que no se detecta adecuadamente mediante la

secuenciación del exoma (215); y en los casos en los que la variante se encuentra en una zona no codificante.

Además, para identificar qué variantes tienen más probabilidades de ser patógenas es fundamental la existencia de grandes conjuntos de datos de control que nos permita establecer la variación genética común normal (216, 217), o datos de secuenciación basados en la familia, como la secuenciación de tríos, particularmente en el contexto del descubrimiento de variantes de novo (5, 218). Como ya se ha comentado, en un contexto diagnóstico solo podemos interpretar aquellas variantes sobre las que tenemos información previa sobre sus potenciales consecuencias biológicas. Por tanto, una de las áreas de desarrollo más activas ha sido la expansión y el desarrollo continuo de grandes recursos de bases de datos de variación genética común y asociada a enfermedad, tanto en zonas codificantes (exoma) como en aquellas no codificantes, donde a día de hoy reside la mayor dificultad a la hora de interpretar variantes generadas por ejemplo en los estudios de genoma completo. Bases de datos como ExAC (217, 219) o gnomAD (220, 221) y Discover EHR (216) actualmente proporcionan datos sobre más de 100.000 exomas (ExAC) y genomas (gnomAD) de diversas poblaciones combinadas. La Unión Europea está trabajando actualmente en esta dirección a través de la iniciativa colaborativa “1+ Million Genomes”. Se trata de un proyecto incluido en el programa Horizonte 2020 que involucra a 22 estados miembros y Noruega con el objetivo de secuenciar 1.000.000 de genomas hasta febrero de 2022, vinculando el acceso a las bases de datos genómicas existentes y futuras en toda la unión europea y proporcionando una escala suficientemente amplia para nuevas asociaciones clínicamente significativas en la investigación. De la misma manera, otras iniciativas semejantes están siendo desarrolladas en otras partes del mundo como el proyecto de investigación “All of us” en Estados Unidos.

Por otra parte, utilizando la base de datos ExAC, Lek et al. (217) desarrolló la puntuación pLI que refleja la intolerancia de cada gen a la variación de pérdida de función. Esta puntuación proporciona antecedentes sobre la probabilidad de ver variantes con impacto funcional en genes específicos dentro de una muestra de población sana y ha sido particularmente informativo para el descubrimiento de genes que contribuyen a los trastornos de herencia mendeliana dominantes(222, 223) Recientemente, Coban-Akdemir et al. (223) desarrollaron una puntuación de intolerancia al escape del mecanismo *nonsense mediated decay* (NMD) para identificar genes en los que los alelos sin sentido pueden contribuir a la ganancia de función. A medida que estos recursos informativos crecen, se espera que vayan surgiendo nuevas métricas de puntuación de variantes que tengan en cuenta las diferentes formas de herencia y el impacto funcional de las variantes y permitan una mejor clasificación de dichas variantes.

Además de la ampliación y mejora, de las bases de datos de pacientes y controles, otra área donde las bases de datos permiten mejorar el conocimiento y la interpretación biológica del efecto de las variantes son aquellas bases que integran la información fenotípica y genotípica. El fenotipado extenso y preciso es esencial para establecer un vínculo adecuado entre los genes candidatos potenciales y las características de la enfermedad. Herramientas como DECIPHER (224), Matchmaker Exchange (225), GeneMatcher (226) o PhenomeCentral (227) permiten emparejar casos con fenotipos similares y / o perfiles genotípicos para ayudar en el diagnóstico de enfermedades raras. Bases de datos como HPO (88), OMIM y Orphanet proporcionan información útil para delimitar genes candidatos. Sin embargo, es importante subrayar que estas bases de datos pueden basarse en una semántica diferente, y la variabilidad de los términos clínicos al describir características clínicas similares puede conducir a análisis menos precisos.

Independientemente del enfoque de secuenciación del ADN, se estima que la tasa de diagnóstico utilizando la secuenciación del exoma o del genoma se estabiliza en general para el diferente espectro de patologías entre el 35% y el 50% (98, 199, 228, 229). Hay una necesidad de otros enfoques para identificar la firma molecular de la enfermedad y, por lo tanto, apuntar de manera más eficiente al gen o genes causales. Una aproximación potencial es centrarse en las consecuencias de las variantes patogénicas en diferentes productos celulares, entre los que se encuentran los transcritos de genes, proteínas o metabolitos. Al agregar esas capas de información, es posible identificar productos o actividades aberrantes que pueden reducir aún más la lista de genes y variantes candidatos. De hecho se ha propuesto que el aumento en el rendimiento diagnóstico vendrá de la integración de información proveniente de diferentes fuentes, genómicas, fenómicas, metabolómicas y transcriptómicas.

La secuenciación de ARN (ARN-seq), por ejemplo, se ha convertido ahora en el gold estándar para cuantificar los niveles de ARN y estimar la diversidad de transcritos. El ARN-seq proporciona una medida de las consecuencias de las variantes codificantes y no codificantes en los niveles de expresión génica y el empalme (splicing) alternativo(230). Bajo el supuesto de que solo uno o muy pocos genes se ven afectados, ARN-seq permite detectar y restringir la investigación al subconjunto de genes con expresión aberrante o defectos del splicing en los individuos afectados en comparación con los controles no afectados. Existe una evidencia cada vez mayor de que las variantes genéticas raras influyen en estos eventos aberrantes, lo que proporciona más información para localizar variantes causales (231-233). El análisis del transcriptoma ha demostrado ser una herramienta complementaria útil para la mejora del diagnóstico genético, particularmente en el caso de enfermedades raras (209). Su uso permite la identificación de eventos potencialmente patogénicos

relacionados con la expresión génica, causados por variantes no identificadas por WES. Uno de los ejemplos de la utilidad de este tipo de aproximación es el trabajo de Cummings et al (209) en un grupo de pacientes con enfermedad muscular rara no diagnosticada. En este trabajo, a partir de ARN extraído de musculo de los pacientes, se pudieron identificar defectos de la expresión en un 35% de los pacientes. Entre estos defectos, se identificó una variante en el gen *COL6A1*, afectando al splicing, responsable del 25% de los casos de Distrofia Muscular Congénita de Déficit de Colágeno VI. Una limitación de este tipo de estudios es la especificidad de tejido en relación a la regulación de la expresión génica, lo que en principio obliga al estudio del tejido diana de cada patología. Sin embargo, un reciente estudio basado en el análisis del transcriptoma en sangre ha demostrado su utilidad en el diagnóstico de enfermedades raras (234). En este estudio se analizaron mediante RNA-seq 94 pacientes, aumentando la tasa diagnóstica entre un 7.5% y un 16.7%. Además, demostraron que el 76% de los genes registrados en la base de datos de OMIM se expresaban en sangre, así como un 66% de los genes con intolerancia alta a la pérdida de función y que las variantes con consecuencias más graves solían ocurrir en genes cuya expresión no era específica de tejido. Así, la aplicación de las tecnologías de secuenciación masiva al estudio del transcriptoma se está revelando como una herramienta de enorme utilidad a la hora de mejorar el diagnóstico genético en pacientes con enfermedades raras, al permitir la identificación de defectos moleculares que, mediante la secuenciación de ADN, no es posible identificar.

El uso de la transcriptómica en el diagnóstico de neonatos enfermos no ha sido evaluado hasta la fecha, pero dada la importancia del diagnóstico genético en esta población, la posibilidad de aumentar la tasa diagnóstica mediante este tipo de estudio es de especial interés como investigación y más, si cabe, como aplicación clínica.

## 6. CONCLUSIONES

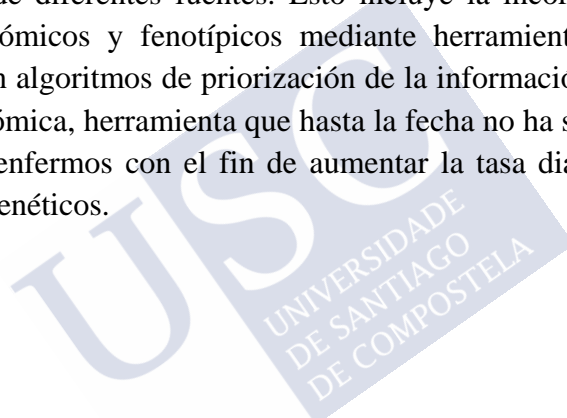
1. El panel NeoSeq, diseñado tras seleccionar específicamente aquellos genes asociados a enfermedad neonatal y aplicado en neonatos críticamente enfermos con sospecha de enfermedad genética neurológica, metabólica y dismórfica, presentó una tasa diagnóstica del 36.4% con un tiempo medio hasta el diagnóstico de 7.5 días, demostrando ser una alternativa real a otras herramientas diagnósticas de mayor complejidad y coste, como el exoma o el genoma completo.
2. El rendimiento diagnóstico del panel está fuertemente asociado con la especificidad de los hallazgos fenotípicos. Así, el panel presentó un mayor rendimiento diagnóstico en aquellos pacientes con sospecha de enfermedad metabólica y enfermedad neurológica. En estos casos se alcanzó un diagnóstico genético en el 43% y el 35% respectivamente, prácticamente el doble de rendimiento diagnóstico en el primer grupo que en el grupo de anomalías congénitas, con un 23%. Sin embargo, no observamos diferencias significativas en las tasas de diagnóstico entre los grupos fenotípicos.
3. El abordaje en trio demostró su utilidad al permitir descartar por un lado la asociación entre las variantes identificadas y la enfermedad del neonato en tres casos al identificarse un patrón de herencia incompatible con el diagnóstico genético asociado a los posibles diagnósticos, mientras que en un 41.6% de los casos permitió identificar variantes *de novo*, en un 41.7% confirmar el patrón de herencia recesiva y en un 58.3% confirmar el patrón de herencia dominante. El abordaje en trio permite obtener el

diagnóstico definitivo de manera más rápida, al hacer innecesarios los posteriores estudios de segregación de las variantes identificadas en los pacientes índice.

4. Al evaluar el impacto del diagnóstico genético en la evolución clínica, se encontró que el establecimiento de un diagnóstico molecular afectó directamente el manejo médico en 12 de los 13 pacientes, resultando en cambios en la medicación o dieta (10 pacientes), inicio de nuevas valoraciones por subespecialistas específicos que no se habían considerado previamente (cinco pacientes) o retirada del tratamiento de cuidados intensivos con reorientación hacia cuidados paliativos (cinco pacientes). Estos datos demuestran la utilidad clínica de establecer un diagnóstico genético rápido en neonatos críticamente enfermos.
5. La tasa diagnóstica resultante (12/33 pacientes [36,4%]) fue comparable a la descrita anteriormente para enfoques más amplios de secuenciación de trío o de probando, como WES y WGS con unas tasas diagnósticas medias de 31.2% y 29.8% respectivamente, e incluso más alto que el informado para WGS en pacientes ingresados en UCIN. Mientras que las tasas de diagnóstico más altas deberían esperarse de WGS que de WES, y de WES que CES o paneles específicos, encontramos que ésta no se correlaciona con el tamaño de la porción del genoma interrogado ya que NeoSeq proporcionó tasas de diagnóstico similar a las obtenidas con los enfoques WGS y WES, a pesar de interrogar 10 veces (WES) y 500 veces (WGS) menos pares de bases.
6. Independientemente del enfoque de secuenciación del ADN, se estima que la tasa de diagnóstico utilizando la secuenciación del exoma o del genoma se estabiliza en general para el diferente espectro de patologías y no alcanza en ningún caso una tasa superior al 60%. Existe por lo tanto una necesidad de encontrar nuevos enfoques que permitan identificar la causa molecular de la enfermedad en estos pacientes.

## 7. PERSPECTIVAS FUTURAS

Nuevos estudios y líneas de investigación deberían explorar la posibilidad de aumentar el rendimiento de las herramientas de diagnóstico genético mediante la integración de la información obtenida de diferentes fuentes. Esto incluye la incorporación de los datos genómicos y fenotípicos mediante herramientas informáticas basadas en algoritmos de priorización de la información y el uso de la transcriptómica, herramienta que hasta la fecha no ha sido evaluada en neonatos enfermos con el fin de aumentar la tasa diagnóstica de los estudios genéticos.





## 8. ANEXO I: Genes incluidos en el panel NeoSeq

AARS	ADAR	AHDC1	ANO10
AARS2	ADAT3	AHI1	ANO3
AASS	ADCK1	AIFM1	ANO5
ABAT	ADCK2	AIMP1	ANXA11
ABCB11	ADCK3	AIMP2	AP1S2
ABCB7	ADCK4	AKR1B15	AP3B2
ABCC8	ADCK5	AKR1D1	AP4B1
ABCD1	ADCY1	AKT1	AP4E1
ABCD3	ADCY5	AKT2	AP4M1
ABCD4	ADCY6	AKT3	AP4S1
ABHD12	ADD3	ALAD	AP5Z1
ABHD5	ADGRG1	ALAS2	APOL2
ACAA1	ADGRG6	ALDH18A1	APOL4
ACAD8	ADGRV1	ALDH3A2	APOPT1
ACAD9	ADHD1	ALDH4A1	APPL1
ACADM	ADHD2	ALDH5A1	APRT
ACADS	ADHD3	ALDH6A1	APTX
ACADSB	ADHD4	ALDH7A1	ARCN1
ACADVL	ADK	ALDOA	ARF1
ACAT1	ADNP	ALDOB	ARFGEF2
ACAT2	ADPRHL2	ALG13	ARG1
ACBD5	ADSL	ALG14	ARHGEF10
ACER3	ADSSL1	ALG2	ARHGEF15
ACMSD	AEBP1	ALMS1	ARHGEF2
ACO2	AFF2	ALPK3	ARHGEF6
ACOX1	AFG3L2	ALS2	ARHGEF9
ACP33	AGA	ALX4	ARID1A
ACSF3	AGK	AMACR	ARID1B
ACSL4	AGL	AMPD1	ARL13B
ACTA1	AGMO	AMPD2	ARL3
ACTB	AGPS	AMT	ARL6
ACTG1	AGRN	ANK3	ARL6IP1
ADA	AGXT	ANKH	ARMC4
ADAM22	AGXT2	ANKLE2	ARMC9
ADAMTS2	AHCY	ANKRD11	ARSA

ARSB	ATP7B	BSCL2	CC2D2A
ARSG	ATP8A2	BSND	CCDC174
ARSI	ATPAF1	BTD	CCDC28B
ARV1	ATPAF2	BVES	CCDC50
ARX	ATR	C10ORF2	CCDC78
ASAH1	ATRN	C11ORF73	CCDC88A
ASCC1	ATRX	C12ORF65	CCDC88C
ASH1L	AUH	C19ORF12	CCM1
ASL	B3GALNT2	C1QBP	CCND2
ASNS	B3GNT2	C2ORF25	CCT5
ASPA	B4GALNT1	C2ORF86	CD164
ASPM	B4GALT7	C5ORF42	CD320
ASS1	B9D1	C6ORF66	CDC14A
ATAD1	B9D2	C7ORF10	CDH15
ATAD3A	BAG3	C8ORF37	CDH23
ATAD3B	BBIP1	CA5A	CDK5
ATAD3C	BBS1	CA8	CDK5RAP2
ATCAY	BBS10	CABP2	CDK6
ATG5	BBS12	CACNA1A	CDKL5
ATL1	BBS2	CACNA1B	CDON
ATL3	BBS4	CACNA1C	CEACAM16
ATM	BBS5	CACNA1D	CEL
ATP13A2	BBS7	CACNA1G	CENPE
ATP1A2	BBS9	CACNA1H	CENPF
ATP1A3	BCAP31	CACNA1S	CENPJ
ATP2A1	BCKDHA	CACNA2D1	CEP104
ATP2B3	BCKDHB	CACNA2D2	CEP120
ATP5A1	BCKDK	CACNB4	CEP135
ATP5B	BCL11A	CACNG2	CEP152
ATP5C1	BCL11B	CAD	CEP164
ATP5D	BCS1L	CAMK2A	CEP290
ATP5E	BDP1	CAMK2B	CEP41
ATP5F1	BEAN	CAMTA1	CEP63
ATP5F1D	BICD2	CAPN1	CEP89
ATP5G1	BIN1	CAPN3	CERS1
ATP5G2	BLK	CARS2	CFL2
ATP5G3	BMP1	CASC5	CHAMP1
ATP5O	BMP2	CASK	CHAT
ATP6AP2	BOLA3	CASQ1	CHCHD10
ATP6V0A4	BRAF	CASR	CHD1
ATP6V1A	BRAT1	CAV3	CHD2
ATP6V1B1	BRF1	CBL	CHD7
ATP6V1B2	BRPF1	CBS	CHD7
ATP7A	BRWD3	CC2D1A	CHD8

CHI3L1	COCH	COX18	CYCS
CHKB	COL11A1	COX19	CYFIP2
CHMP1A	COL11A2	COX20	CYP26B1
CHMP1B	COL12A1	COX4I1	CYP27A1
CHRNA1	COL13A1	COX4I2	CYP2U1
CHRNA2	COL18A1	COX6A1	CYP7B1
CHRNA4	COL1A1	COX6A2	CZD12
CHRNB1	COL1A2	COX6B1	D2HGDH
CHRNB2	COL2A1	COX6B2	DAG1
CHRND	COL3A1	COX7B	DAO
CHRNE	COL4A1	COX8A	DAOA
CHRNG	COL4A2	CP	DARS
CHST14	COL4A3	CPA6	DARS2
CIB2	COL4A3BP	CPLX1	DBH
CISD2	COL4A4	CPOX	DBT
CIT	COL4A5	CPS1	DCAF8
CIZ1	COL4A6	CPT1A	DCDC2
CLCN1	COL5A1	CPT1C	DCHS1
CLCN2	COL5A2	CPT2	DCX
CLCN4	COL5A3	CRADD	DDC
CLCN6	COL6A1	CRAT	DDHD1
CLDN14	COL6A2	CRBN	DDHD2
CLIC2	COL6A3	CRTAP	DDX3X
CLIC5	COL6A6	CRYAB	DEAF1
CLN3	COL9A1	CRYM	DENND5A
CLN5	COL9A2	CSF1R	DEPDC5
CLN6	COLEC11	CSNK1G1	DES
CLN8	COLQ	CSNK2B	DFNA5
CLP1	COMT	CSPP1	DFNB31
CLPB	COQ10A	CSTB	DFNB59
CLPP	COQ10B	CTBP1	DFNX3
CLPX	COQ2	CTC1	DGAT2
CLRN1	COQ3	CTCF	DGUOK
CLTC	COQ4	CTH	DHCR24
CNNM2	COQ5	CTNNA2	DHCR7
CNPY3	COQ6	CTNNB1	DHDDS
CNTN1	COQ7	CTSA	DHFR
CNTN2	COQ9	CTSD	DHTKD1
CNTNAP1	COX10	CTSF	DHX30
CNTNAP2	COX11	CTU2	DIABLO
COA5	COX14	CUL4B	DIAPH1
COA6	COX15	CUX2	DISC2
COA7	COX16	CWF19L1	DISP1
COASY	COX17	CYC1	DLAT

<b>DLD</b>	<b>EDN3</b>	<b>ETHE1</b>	<b>FLAD1</b>
<b>DLG3</b>	<b>EDNRB</b>	<b>EXOSC3</b>	<b>FLNA</b>
<b>DLGAP4</b>	<b>EEF1A2</b>	<b>EXOSC9</b>	<b>FLNC</b>
<b>DLL1</b>	<b>EEF2</b>	<b>EXT2</b>	<b>FLVCR1</b>
<b>DMD</b>	<b>EFHC1</b>	<b>EYA1</b>	<b>FMN2</b>
<b>DNA2</b>	<b>EFNB1</b>	<b>EYA4</b>	<b>FMO3</b>
<b>DNAJB2</b>	<b>EFTUD2</b>	<b>EZH2</b>	<b>FMR1</b>
<b>DNAJB6</b>	<b>EGR2</b>	<b>FA2H</b>	<b>FOLR1</b>
<b>DNAJC12</b>	<b>EIF2B1</b>	<b>FAH</b>	<b>FOXA2</b>
<b>DNAJC19</b>	<b>EIF2B2</b>	<b>FAM111B</b>	<b>FOXG1</b>
<b>DNAJC3</b>	<b>EIF2B3</b>	<b>FAM126A</b>	<b>FOXH1</b>
<b>DNAJC5</b>	<b>EIF2B4</b>	<b>FAM36A</b>	<b>FOXI1</b>
<b>DNM1</b>	<b>EIF2B5</b>	<b>FAM65B</b>	<b>FOXP1</b>
<b>DNM1L</b>	<b>EIF2S3</b>	<b>FARS2</b>	<b>FOXP2</b>
<b>DNM2</b>	<b>EIF4E</b>	<b>FARSB</b>	<b>FOXRED1</b>
<b>DNMT1</b>	<b>ELAC2</b>	<b>FAS</b>	<b>FREM1</b>
<b>DNMT3A</b>	<b>ELMOD3</b>	<b>FASTKD2</b>	<b>FRMD4A</b>
<b>DOCK3</b>	<b>ELOVL4</b>	<b>FAT2</b>	<b>FRMPD4</b>
<b>DOCK7</b>	<b>ELOVL5</b>	<b>FAT4</b>	<b>FRRS1L</b>
<b>DOCK8</b>	<b>ELP2</b>	<b>FBLN5</b>	<b>FTCD</b>
<b>DOK7</b>	<b>EMD</b>	<b>FBN1</b>	<b>FTL</b>
<b>DPAGT1</b>	<b>EMILIN1</b>	<b>FBN2</b>	<b>FTSJ1</b>
<b>DPP6</b>	<b>EML1</b>	<b>FBP1</b>	<b>FUCA1</b>
<b>DPYD</b>	<b>EMX2</b>	<b>FBXL4</b>	<b>FUS</b>
<b>DPYS</b>	<b>ENTPD1</b>	<b>FBXO11</b>	<b>G6PC</b>
<b>DRD2</b>	<b>EOMES</b>	<b>FBXO31</b>	<b>GAA</b>
<b>DRD3</b>	<b>EPB41L1</b>	<b>FDFT1</b>	<b>GABBR2</b>
<b>DRD4</b>	<b>EPM2A</b>	<b>FDX1L</b>	<b>GABRA1</b>
<b>DRD5</b>	<b>EPRS</b>	<b>FECH</b>	<b>GABRA5</b>
<b>DRP2</b>	<b>EPS8</b>	<b>FGD1</b>	<b>GABRB1</b>
<b>DST</b>	<b>EPT1</b>	<b>FGD4</b>	<b>GABRB2</b>
<b>DSTYK</b>	<b>ERAL1</b>	<b>FGF12</b>	<b>GABRB3</b>
<b>DYM</b>	<b>ERBB3</b>	<b>FGF14</b>	<b>GABRD</b>
<b>DYNC1H1</b>	<b>ERBB4</b>	<b>FGF8</b>	<b>GABRG2</b>
<b>DYRK1A</b>	<b>ERF</b>	<b>FGFR1</b>	<b>GAD1</b>
<b>DYSF</b>	<b>ERGIC1</b>	<b>FGFR2</b>	<b>GALC</b>
<b>DYT1</b>	<b>ERLIN1</b>	<b>FGFR3</b>	<b>GALE</b>
<b>E4F1</b>	<b>ERLIN2</b>	<b>FH</b>	<b>GALK1</b>
<b>EARS2</b>	<b>ERMARD</b>	<b>FHL1</b>	<b>GALNS</b>
<b>EBF3</b>	<b>ESPN</b>	<b>FIG4</b>	<b>GALT</b>
<b>EBP</b>	<b>ESRRB</b>	<b>FKBP10</b>	<b>GAN</b>
<b>ECEL1</b>	<b>ETFA</b>	<b>FKBP14</b>	<b>GARS</b>
<b>ECHS1</b>	<b>ETFB</b>	<b>FKRP</b>	<b>GAS1</b>
<b>EDC3</b>	<b>ETFDH</b>	<b>FKTN</b>	<b>GATAD2B</b>

GBA	GNB4	HAL	HPD
GBA2	GNB5	HARS	HPRT1
GBE	GNE	HARS2	HRAS
GBE1	GNMT	HBA1	HSD17B10
GCDH	GNPAT	HCCS	HSD17B4
GCH1	GNPTAB	HCFC1	HSD3B7
GCK	GNPTG	HCN1	HSPB1
GCSH	GNS	HCN2	HSPB3
GDAP1	GOLGA2	HCN4	HSPB8
GDF6	GOSR2	HDAC8	HSPD1
GDI1	GOT2	HECW2	HSPE1
GFAP	GPA A1	HEPACAM	HSPG2
GFER	GPHN	HERC2	HTR2A
GFM1	GPR88	HESX1	HTRA1
GFM2	GPR98	HEXA	HTRA2
GFPT1	GPSM2	HEXB	HUWE1
GGT1	GPT2	HFE	HYAL1
GIPC3	GRHL2	HGD	HYOU1
GJB1	GRHPR	HGF	IARS
GJB2	GRIA3	HGSNAT	IARS2
GJB3	GRIA4	HIBCH	IBA57
GJB6	GRID2	HINT1	ICK
GJC2	GRIK2	HINT3	IDH2
GLA	GRIN1	HIVEP2	IDS
GLB1	GRIN2A	HK1	IDUA
GLDC	GRIN2B	HLCS	IER3IP1
GLDN	GRIN2D	HMBS	IFIH1
GLE1	GRM1	HMGCL	IFITM5
GLI2	GRN	HMGCS2	IFT122
GLI3	GRXCR1	HNF1A	IFT27
GLIS2	GRXCR2	HNF4A	IFT64
GLIS3	GSTZ1	HNF6	IGBP1
GLRA1	GTPBP3	HNMT	IGHMBP2
GLRB	GUF1	HNRNPA1	IKBKAP
GLRX5	GUSB	HNRNPA2B1	IL11RA
GLUD1	GYG1	HNRNPDL	IL1RAPL1
GLUL	GYG2	HNRNPH1	ILDR1
GLYCTK	GYS1	HNRNPH2	IMPA1
GM2A	GYS2	HNRNPU	INF2
GMPPB	GZF1	HOGA1	INPP5E
GNAL	HACE1	HOMER2	INPP5K
GNAO1	HADH	HOXA1	INS
GNAQ	HADHA	HOXD10	INSR
GNB1	HADHB	HPCA	INVS

<b>IQCB1</b>	<b>KDM5C</b>	<b>LEPRE1</b>	<b>MAPT</b>
<b>IQSEC1</b>	<b>KIAA0196</b>	<b>LETM1</b>	<b>MARS</b>
<b>IQSEC2</b>	<b>KIAA0226</b>	<b>LGI1</b>	<b>MARS2</b>
<b>IQSEQ2</b>	<b>KIAA0442</b>	<b>LGI4</b>	<b>MARVELD2</b>
<b>IRF2BPL</b>	<b>KIAA0556</b>	<b>LHFPL5</b>	<b>MASP1</b>
<b>ISCA1</b>	<b>KIAA0586</b>	<b>LIAS</b>	<b>MAT1A</b>
<b>ISCA2</b>	<b>KIAA1279</b>	<b>LIMS2</b>	<b>MAT2A</b>
<b>ISCU</b>	<b>KIAA2022</b>	<b>LINGO1</b>	<b>MATR3</b>
<b>ISPD</b>	<b>KIDINS220</b>	<b>LINS</b>	<b>MBD5</b>
<b>ITGA7</b>	<b>KIF11</b>	<b>LIPA</b>	<b>MBOAT7</b>
<b>ITPA</b>	<b>KIF14</b>	<b>LIPT1</b>	<b>MCCC1</b>
<b>ITPR1</b>	<b>KIF1A</b>	<b>LIPT2</b>	<b>MCCC2</b>
<b>IVD</b>	<b>KIF1B</b>	<b>LITAF</b>	<b>MCEE</b>
<b>JNK3</b>	<b>KIF1C</b>	<b>LMAN2L</b>	<b>MCM2</b>
<b>JPH1</b>	<b>KIF22</b>	<b>LMBRD1</b>	<b>MCM3AP</b>
<b>KIAA1109</b>	<b>KIF2A</b>	<b>LMNA</b>	<b>MCOLN1</b>
<b>KANK1</b>	<b>KIF4A</b>	<b>LMNB1</b>	<b>MCPH1</b>
<b>KARS</b>	<b>KIF5A</b>	<b>LMNB2</b>	<b>MDH1</b>
<b>KAT6A</b>	<b>KIF5C</b>	<b>LMOD3</b>	<b>MDH2</b>
<b>KATNB1</b>	<b>KIF7</b>	<b>LNPB</b>	<b>MECP2</b>
<b>KBTBD13</b>	<b>KIRREL3</b>	<b>LONP1</b>	<b>MECR</b>
<b>KCC2</b>	<b>KITLG</b>	<b>LOXHD1</b>	<b>MED12</b>
<b>KCNA1</b>	<b>KLC2</b>	<b>LRP2</b>	<b>MED17</b>
<b>KCNA2</b>	<b>KLF11</b>	<b>LRP4</b>	<b>MED20</b>
<b>KCNB1</b>	<b>KLHL15</b>	<b>LRPPRC</b>	<b>MED23</b>
<b>KCNC1</b>	<b>KLHL40</b>	<b>LRRC56</b>	<b>MED25</b>
<b>KCNC3</b>	<b>KLHL41</b>	<b>LRSAM1</b>	<b>MEF2C</b>
<b>KCND3</b>	<b>KMT2B</b>	<b>LRTOMT</b>	<b>MEGF10</b>
<b>KCNE1</b>	<b>KMT5B</b>	<b>LTC4S</b>	<b>MET</b>
<b>KCNH1</b>	<b>KPTN</b>	<b>LYRM4</b>	<b>METTL23</b>
<b>KCNH5</b>	<b>KRAS</b>	<b>LYRM7</b>	<b>MFF</b>
<b>KCNJ10</b>	<b>KY</b>	<b>LZTFL1</b>	<b>MFN1</b>
<b>KCNJ11</b>	<b>L1CAM</b>	<b>LZTR1</b>	<b>MFN2</b>
<b>KCNMA1</b>	<b>L2HGDH</b>	<b>MAFA</b>	<b>MFSD2A</b>
<b>KCNQ1</b>	<b>LAMA2</b>	<b>MAG</b>	<b>MFSD8</b>
<b>KCNQ2</b>	<b>LAMB1</b>	<b>MAGEL2</b>	<b>MGME1</b>
<b>KCNQ3</b>	<b>LAMB2</b>	<b>MAGI2</b>	<b>MICU1</b>
<b>KCNQ4</b>	<b>LAMC3</b>	<b>MAN1B1</b>	<b>MID2</b>
<b>KCNQ5</b>	<b>LAMP2</b>	<b>MAN2B1</b>	<b>MIPEP</b>
<b>KCNT1</b>	<b>LARGE</b>	<b>MANBA</b>	<b>MIR17HG</b>
<b>KCNT2</b>	<b>LARS2</b>	<b>MAOA</b>	<b>MITF</b>
<b>KCTD17</b>	<b>LAS1L</b>	<b>MAP18</b>	<b>MKKS</b>
<b>KCTD7</b>	<b>LDB3</b>	<b>MAP2</b>	<b>MKS1</b>
<b>KDM5B</b>	<b>LDHA</b>	<b>MAPK10</b>	<b>MLC1</b>

MLYCD	MVK	NDUFA3	NEXMIF
MMAA	MYBPC1	NDUFA4	NFIA
MMAB	MYCN	NDUFA4L2	NFS1
MMACHC	MYF6	NDUFA5	NFU1
MMADHC	MYH14	NDUFA6	NGF
MME	MYH2	NDUFA9	NGN3
MNF1	MYH3	NDUFAF1	NHEJ1
MOCOS	MYH7	NDUFAF2	NHLRC1
MOCS1	MYH8	NDUFAF3	NIN
MOCS2	MYH9	NDUFAF4	NIPA1
MORC2	MYO15A	NDUFAF5	NIPA2
MPV17	MYO18B	NDUFAF6	NIPBL
MPZ	MYO3A	NDUFB1	NKX6-2
MR1	MYO6	NDUFB10	NLGN3
MRE11A	MYO7A	NDUFB11	NLGN4
MRPL12	MYO9A	NDUFB2	NLGN4X
MRPL3	MYOT	NDUFB3	NONO
MRPL44	MYPN	NDUFB4	NOTCH3
MRPS16	MYT1L	NDUFB5	NPC1
MRPS22	NAA15	NDUFB6	NPC2
MRPS23	NACC1	NDUFB7	NPHP1
MRPS34	NADK2	NDUFB8	NPHP3
MRPS7	NAGA	NDUFB9	NPHP4
MRT4	NAGLU	NDUFS1	NPL
MRT60	NAGS	NDUFS2	NPRL2
MSMO1	NALCN	NDUFS3	NPRL3
MSRB3	NANS	NDUFS4	NRAS
MSTO1	NAPB	NDUFS5	NRXN1
MSX2	NARS2	NDUFS6	NSDHL
MTCH1	NAT8L	NDUFS7	NSUN2
MTERF	NAXE	NDUFS8	NT5C2
MTFMT	NCAPD2	NDUFV1	NT5C3A
MTHFD1L	NCAPD3	NDUFV2	NTNG1
MTHFR	NCAPH	NDUFV3	NTRK1
MTM1	NDE1	NEB	NTRK2
MTMR14	NDP	NECAP1	NUBPL
MTMR2	NDRG1	NEDD4L	NUP37
MTO1	NDST1	NEFH	NUS1
MTOR	NDUFA1	NEFL	OAT
MTPAP	NDUFA10	NEK1	OCLN
MTR	NDUFA11	NEK8	OFD1
MTRR	NDUFA12	NEK9	OGDH
MUSK	NDUFA13	NEU1	OPA1
MUT	NDUFA2	NEUROD1	OPA3

OPHN1	PDHB	PIGA	POLR3A
ORAI1	PDHX	PIGC	POLR3B
OSBPL2	PDK3	PIGG	POLRMT
OTC	PDP1	PIGH	POMGNT1
OTOA	PDP2	PIGN	POMGNT2
OTOF	PDSS1	PIGO	POMK
OTOG	PDSS2	PIGQ	POMT1
OTOGL	PDX1	PIGS	POMT2
OTUD4	PDYN	PIK3CA	POU3F4
OTUD6B	PDZD7	PIK3R2	POU4F3
OTX2	PEPD	PIK3R5	PIIB
OXCT1	PET100	PINK1	PPM1D
P2RX2	PET117	PIP5K1C	PPOX
P4HB	PEX1	PKLR	PPP1CB
PABPN1	PEX10	PLA2G6	PPP1R15B
PACS1	PEX11B	PLAA	PPP2R1A
PACS2	PEX12	PLCB1	PPP2R5D
PAFAH1B1	PEX13	PLD3	PPP3CA
PAH	PEX14	PLEC	PPT1
PAK1	PEX16	PLEKHG2	PQBP1
PAK3	PEX19	PLEKHG5	PRDM8
PANK2	PEX2	PLOD1	PRICKLE1
PARK2	PEX26	PLOD2	PRICKLE2
PARK7	PEX3	PLOD3	PRKCG
PARS2	PEX5	PLP1	PRKRA
PAX3	PEX6	PMM2	PRNP
PAX4	PEX7	PMP22	PRODH
PBX1	PFKM	PMPCA	PROSC
PC	PGAM2	PMPCB	PRPS1
PCBD1	PGAP1	PNKP	PRRT2
PCCA	PGAP2	PNP	PRSS12
PCCB	PGK1	PNPLA2	PRUNE1
PCDH12	PGM1	PNPLA4	PRX
PCDH15	PHC1	PNPLA6	PSAP
PCDH19	PHF6	PNPLA8	PTCD1
PCDH7	PHF8	PNPO	PTCH1
PCK1	PHIP	PNPT1	PTCHD1
PCK2	PHKA1	POGLUT1	PTEN
PCNA	PHKA2	POGZ	PTPLA
PDE10A	PHKB	POLG	PTPN11
PDE6D	PHKG2	POLG2	PTPRQ
PDGFB	PHYH	POLR1A	PTS
PDHA1	PIBF1	POLR1C	PUM1
PDHA2	PIEZO2	POLR1D	PURA

PUS1	RNASEH2A	SCN9A	SHH
PUS3	RNASEH2B	SCO1	SHOC2
PYCR2	RNASEH2C	SCO2	SHROOM4
PYGL	RNASET2	SCP2	SIGMAR1
PYGM	RNF170	SCYL1	SIK1
PYROXD1	RNF216	SCZD1	SIL1
QARS	ROBO1	SCZD11	SIN3A
QDPR	ROGDI	SCZD12	SIX1
QRSL1	RORA	SCZD2	SIX3
RAB11A	RORB	SCZD3	SIX5
RAB11B	RP2	SCZD5	SKI
RAB18	RPGRIP1L	SCZD6	SLC12A5
RAB23	RPIA	SCZD7	SLC12A6
RAB33B	RPL10	SCZD8	SLC13A5
RAB39B	RPS23	SDCCAG8	SLC16A1
RAB3GAP1	RPS6KA3	SDHA	SLC16A2
RAB3GAP2	RRM2B	SDHAF1	SLC17A5
RAB40AL	RTN2	SDHAF2	SLC17A8
RAB7A	RTN4IP1	SDHB	SLC18A2
RAF1	RTN4R	SDHC	SLC18A3
RANBP2	RTTN	SDHD	SLC19A3
RAPSN	RYR1	SEMA3E	SLC1A1
RAR39B	RYR3	SEPN1	SLC1A2
RARS	S1PR2	SEPSECS	SLC1A3
RARS2	SACS	SERAC1	SLC1A4
RBBP8	SAMD9L	SERPINB6	SLC22A5
RBCK1	SAMHD1	SERPINF1	SLC25A1
RBFOX1	SARS	SERPINH1	SLC25A12
RBFOX3	SARS2	SERPINI1	SLC25A13
RBMX	SASS6	SET	SLC25A15
RDX	SBF1	SETBP1	SLC25A19
RECQL4	SBF2	SETD2	SLC25A20
REEP1	SC5D	SETD5	SLC25A22
REEP2	SCA2	SETX	SLC25A26
RELN	SCAPER	SFXN4	SLC25A3
REPS1	SCARB2	SGCA	SLC25A32
RERE	SCN10A	SGCB	SLC25A4
RETREG1	SCN1A	SGCD	SLC25A42
RFT1	SCN1B	SGCE	SLC25A46
RHOBTB2	SCN2A	SGCG	SLC26A4
RIT1	SCN3A	SGSH	SLC26A5
RLIM	SCN4A	SH3TC2	SLC2A1
RMND1	SCN7A	SHANK2	SLC2A10
RNASEH1	SCN8A	SHANK3	SLC2A2

SLC30A10	SOX10	SYNE2	TGFB2
SLC30A9	SOX11	SYNE4	TGFB3
SLC33A1	SOX2	SYNGAP1	TGFBR1
SLC35A2	SOX3	SYNJ1	TGIF1
SLC35A3	SP7	SYP	TGM6
SLC37A4	SPAST	SYT14	TH
SLC39A13	SPATA5	SYT2	THAP1
SLC39A14	SPEG	SZT2	THOC2
SLC3A1	SPG11	TACO1	TIA1
SLC45A1	SPG20	TAF1	TIMM50
SLC46A1	SPG7	TAF2	TIMM8A
SLC52A1	SPR	TANGO2	TIMMDC1
SLC52A2	SPTAN1	TARDBP	TJP2
SLC52A3	SPTBN2	TARS2	TK2
SLC5A1	SPTBN4	TAT	TKT
SLC5A7	SPTLC1	TAZ	TLK2
SLC6A1	SPTLC2	TBC1D23	TM4SF20
SLC6A17	SQSTM1	TBC1D24	TMC1
SLC6A19	SRGAP2	TBC1D7	TMEM106B
SLC6A3	SRGAP2C	TBCD	TMEM107
SLC6A5	SRPX2	TBCE	TMEM126A
SLC6A8	SSPD1	TBCK	TMEM126B
SLC6A9	ST3GAL3	TBL1XR1	TMEM132E
SLC9A1	ST3GAL5	TCAP	TMEM138
SLC9A6	STAC3	TCF12	TMEM216
SLC9A9	STAG1	TCF2	TMEM231
SLITRK6	STIL	TCF4	TMEM237
SMARCA1	STIM1	TCN1	TMEM240
SMARCA4	STRADA	TCN2	TMEM38B
SMARCB1	STRC	TCOF1	TMEM43
SMC1A	STUB1	TCTN1	TMEM5
SMCHD1	STX1B	TCTN2	TMEM67
SMPD1	STXBP1	TCTN3	TMEM70
SMPX	STYXL1	TDGF1	TMIE
SMS	SUCLA2	TDP1	TMLHE
SNAI2	SUCLG1	TDP2	TMPRSS3
SNAP25	SUCO	TECPR2	TMTC3
SNAP29	SUFU	TECR	TNC
SNORD118	SUMF1	TECTA	TNIK
SNX14	SUOX	TELO2	TNNI2
SOD1	SURF1	TENM4	TNNT1
SON	SYN1	TFAM	TNNT3
SOS1	SYN2	TFG	TNPO3
SOS2	SYNE1	TGFB1	TNXB

TOP3A	TSFM	UQCC3	WDPCP
TOR1A	TSPAN7	UQCR10	WDR19
TOR1AIP1	TSPEAR	UQCR11	WDR45
TOR1AIP2	TTBK2	UQCRB	WDR45B
TPK1	TTC19	UQCRC1	WDR48
TPM2	TTC21B	UQCRC2	WDR62
TPM3	TTC8	UQCRES1	WDR73
TPP1	TTI2	UQCRH	WDR81
TPRN	TTN	UQCRHL	WFS1
TRAF3IP1	TTPA	UQCRQ	WNK1
TRAIP	TTR	UROCI	WVOX
TRAK1	TUBA1A	UROD	XDH
TRAPPC11	TUBA8	UROS	XRCC4
TRAPPC12	TUBB	USH1C	YARS
TRAPPC6B	TUBB2A	USH1G	YARS2
TRAPPC9	TUBB2B	USH2A	YME1L1
TREX1	TUBB3	USP27X	YWHAG
TRIM2	TUBB4A	USP8	ZAK
TRIM32	TUBB6	USP9X	ZBTB18
TRIM36	TUBG1	VAC14	ZBTB42
TRIM54	TUBGCP6	VAMP1	ZC3H14
TRIM63	TUFM	VARS	ZDHC15
TRIM8	TUSC3	VARS2	ZDHC9
TRIO	TWIST1	VCP	ZEB2
TRIOBP	TXN2	VDAC1	ZFR
TRIP12	TYMP	VDAC2	ZFYVE26
TRIP4	UBA1	VHL	ZFYVE27
TRIT1	UBA5	VLDLR	ZIC2
TRMT10A	UBE2A	VMA21	ZMYM3
TRMT10C	UBE3A	VPS11	ZMYND11
TRMT5	UBR4	VPS13A	ZNF148
TRMU	UBTF	VPS13D	ZNF335
TRNT1	UCHL1	VPS33A	ZNF41
TRPC3	UCP2	VPS37A	ZNF423
TRPV4	UFM1	VPS53	ZNF469
TSC1	UMOD	VRK1	ZNF592
TSC2	UMPS	VWA3B	ZNF674
TSEN15	UNC13A	WARS	ZNF711
TSEN2	UNC80	WARS2	ZNF81
TSEN34	UPB1	WASHC4	
TSEN54	UPF3B	WDFY3	



## 9. ANEXO II: Publicaciones

de Castro MJ, González-Vioque E, Barbosa-Gouveia S, Salguero E, Rite S, López-Suárez O, Pérez-Muñuzuri A, Couce M-L. Rapid Phenotype-Driven Gene Sequencing with the NeoSeq Panel: A Diagnostic Tool for Critically Ill Newborns with Suspected Genetic Disease. *Journal of Clinical Medicine*. 2020; 9(8):2362.

<https://doi.org/10.3390/jcm9082362>

Factor de impacto: 3.303



Categoría y cuartil JCR: 36/165 (Q1) en 'Medicine, General & Internal'

<https://www.mdpi.com/2077-0383/9/8/2362>



Article

# Rapid Phenotype-Driven Gene Sequencing with the NeoSeq Panel: A Diagnostic Tool for Critically Ill Newborns with Suspected Genetic Disease

María José de Castro <sup>1,2,3,4</sup>, Emiliano González-Vioque <sup>1,2,3</sup>, Sofia Barbosa-Gouveia <sup>1,2</sup>, Enrique Salguero <sup>5</sup>, Segundo Rite <sup>6</sup> , Olalla López-Suárez <sup>1,4</sup>, Alejandro Pérez-Muñuzuri <sup>1,4</sup> and María-Luz Couce <sup>1,2,3,4,\*</sup> 

<sup>1</sup> Diagnosis and Treatment of Congenital Metabolic Diseases Unit (UDyTEMC), Neonatology Division, Department of Pediatrics, Clinical University Hospital of Santiago de Compostela, IDIS-Health Research Institute of Santiago de Compostela, 15706 Santiago de Compostela, Spain;

maria.jose.de.castro.lopez@sergas.es (M.J.d.C.); emiliano.gonzalez.vioque@sergas.es (E.G.-V.); sofia.bsg@gmail.com (S.B.-G.); olalla.elena.lopez.suarez@sergas.es (O.L.-S.);

alejandro.perez.muñuzuri@sergas.es (A.P.-M.)

<sup>2</sup> Centro de Investigación Biomedica en Red (CIBER)—CIBER DE ENFERMEDADES RARAS (CIBERER), Pabellón 11, 28029 Madrid, Spain

<sup>3</sup> MetabERN, Via Pozzuolo, 330, 33100 Udine, Italy

<sup>4</sup> Faculty of Medicine, University of Santiago de Compostela, 15706 Santiago de Compostela, Spain

<sup>5</sup> Neonatology Department, Malaga Regional Hospital, Malaga Biomedical Research Institute-IBIMA, 29011 Malaga, Spain; ensalgar@yahoo.es

<sup>6</sup> Neonatology Unit, Miguel Servet University Hospital, 50009 Zaragoza, Spain; sriteg@salud.aragon.es

\* Correspondence: maria.luz.couce.pico@sergas.es; Tel.: +34-981950151

Received: 27 June 2020; Accepted: 21 July 2020; Published: 23 July 2020



**Abstract:** New genomic sequencing techniques have shown considerable promise in the field of neonatology, increasing the diagnostic rate and reducing time to diagnosis. However, several obstacles have hindered the incorporation of this technology into routine clinical practice. We prospectively evaluated the diagnostic rate and diagnostic turnaround time achieved in newborns with suspected genetic diseases using a rapid phenotype-driven gene panel (NeoSeq) containing 1870 genes implicated in congenital malformations and neurological and metabolic disorders of early onset (<2 months of age). Of the 33 newborns recruited, a genomic diagnosis was established for 13 (39.4%) patients (median diagnostic turnaround time, 7.5 days), resulting in clinical management changes in 10 (76.9%) patients. An analysis of 12 previous prospective massive sequencing studies (whole genome (WGS), whole exome (WES), and clinical exome (CES) sequencing) in newborns admitted to neonatal intensive care units (NICUs) with suspected genetic disorders revealed a comparable median diagnostic rate (37.2%), but a higher median diagnostic turnaround time (22.3 days) than that obtained with NeoSeq. Our phenotype-driven gene panel, which is specific for genetic diseases in critically ill newborns is an affordable alternative to WGS and WES that offers comparable diagnostic efficacy, supporting its implementation as a first-tier genetic test in NICUs.

**Keywords:** critically ill newborn; genomic sequencing; genetic diagnosis; trio sequencing

## 1. Introduction

While genetic disease is suspected in over 50% of the children who are admitted to neonatal intensive care units (NICUs) and die during the first year of life, diagnosis is confirmed in only 20%–30% of cases, often post-mortem [1,2]. Multiple factors hinder genetic diagnosis in neonates,

including genetic heterogeneity, there are over 5587 known genetic diseases [3], clinical heterogeneity (e.g., the appearance of formes frustes of classical phenotypes) [4,5], and comorbidity due to the increased fragility of neonates [6,7]. Moreover, disease tends to progress faster in neonates than in other stages of life.

Technological advances in gene sequencing have enabled rapid reading of any part of the genome at an affordable price [8,9]. NICUs are a key target for the implementation of genomic tools [10,11]. Timely and specific diagnosis of newborns can have critical implications for health and wellbeing for the remainder of an infant's life. The few pilot studies focused on the use of this technology that have been conducted in NICU settings [12–23] have yielded highly promising results. Nonetheless, further advances in this field will be required to make precision, personalized, and predictive medicine a reality. In particular, integration and interpretation of the data produced by genomic sequencing is a key obstacle to the incorporation of these strategies into clinical routine practice [24,25]. However, phenotype-based filtering and prioritization could greatly facilitate the interpretation of genetic variants detected by genome sequencing [26,27].

In this pilot study, we evaluated the utility of rapid selective gene panel trio sequencing in critically ill newborns with suspected genetic disorders. We created a specific gene panel, NeoSeq, consisting of 1870 human genes associated with neurological or metabolic disorders and congenital malformations of early onset (i.e., during the first two months of life). We hypothesized that a rapid gene panel consisting of genes of known function, with established associations with diseases of early onset, could simplify data interpretation while offering diagnostic rates comparable to those achieved using whole genome sequencing (WGS), whole exome sequencing (WES), or clinical exome sequencing (CES). Here, we describe the diagnostic rate and turnaround time achieved using this approach, the prevalence and inheritance patterns of the diagnosed diseases, and the effect of a positive molecular diagnosis on patient clinical management. Furthermore, we compare our results with those of similar prospective studies using next generation sequencing (NGS) technologies in cohorts of critically ill newborns.

## 2. Methods

### 2.1. Study Design

This multicenter prospective study included consecutive patients who were admitted to a level IIIB/C NICU in three Spanish reference hospitals during a 24 months period. Patients underwent selective genetic screening for early diagnosis. Participants for this study were recruited through the Clinical University Hospital of Santiago, Malaga Regional Hospital, and the Miguel Servet Clinical University Hospital, with the approval of the Santiago-Lugo Research Ethics Committee (2018/366). Written informed consent was obtained from the legal guardians of all participating patients upon enrolment in the study. Participants for this study were recruited through the Clinical University Hospital of Santiago de Compostela, Malaga Regional Hospital, and the Miguel Servet Clinical University Hospital, with the approval of the Santiago-Lugo Research Ethics Committee (2018/366).

### 2.2. Study Population

Critically ill patients of less than two months of age who met one or more of the following criteria were considered eligible for inclusion: (i) Congenital malformations not obviously related to a clinically identifiable genetic syndrome, (ii) metabolic decompensation not associated with biochemical parameters and/or neonatal screening indicating suspected hereditary metabolic disease, (iii) epilepsy or neurodevelopmental diseases of probable genetic origin. Individuals with clear indications of a specific syndrome that could be tested by targeted analysis of known genes or structural variations were excluded from the study.

### 2.3. Study Variables

For each patient, the following variables were evaluated: Family history, consanguinity data, maternal obstetric history, sex, anthropometric parameters, age of symptom onset, clinical signs and symptoms presented, laboratory test parameters, imaging variables and additional tests carried out to identify the underlying disease, treatments administered (pharmacological, nutritional, respiratory support, dialysis, or other invasive measures), genetic analyses of patient and parents. A list of possible differential diagnoses was generated based on the clinical findings obtained using the Phenomizer clinical diagnostics application ([compbio.charite.de/phenomizer/](http://compbio.charite.de/phenomizer/)) [28]. Upon obtaining the results of the genetic study, the clinical and molecular variables were correlated.

### 2.4. Statistical Analysis

The Chi-squared test was used to compare diagnostic rates between groups. Phenomizer was used to identify candidate diseases based on the clinical features of each patient. *p*-values were estimated by Monte Carlo random sampling and corrected for multiple testing using the Benjamini and Hochberg method. A *p*-value < 0.05 was considered statistically significant.

### 2.5. Procedures

**Blood sample collection from patient and parents.** A 3-mL sample of blood per individual from the trio (patient and parents) to be analyzed was collected in an ethylenediaminetetraacetic (EDTA) tube (BD, Franklin Lakes, NJ, US) and transported by express courier from the patient's reference hospital at room temperature in a padded, hermetically sealed envelope.

**Genetic testing.** DNA was isolated from 400 µL of fresh blood in EDTA following standard procedures. Enrichment for focused exome analysis covering the exons and flanking introns ( $\pm 10$  pb) of 1870 genes associated with metabolic, neurological, and dysmorphic diseases of infants (NeoSeq, Table S1) was carried out using the SureDesign tool (Agilent Technologies Inc., Santa Clara, CA, USA). Candidate genes were selected by a multidisciplinary team meeting (MDT), which included research bioinformatics analysts, clinical geneticists, neonatal intensivists, neurologists and pediatricians specialized in inborn errors of metabolism. Enriched libraries were sequenced on the NextSeq platform (Illumina Inc., San Diego, CA, USA) following the manufacturer recommendations to achieve an average coverage of 100X using the NextSeq 500/550 Mid Output V2 kit, (Illumina Inc., San Diego, CA, USA), (150 cycles), which provides 18 Gb of sequence data. For each NextSeq run, a single trio was sequenced.

Variant annotation and filtering were performed using pipelines developed in-house (see Appendix A for detailed description).

## 3. Results

Between the period of January 2018 and December 2019, 51 newborns were admitted to the NICU with a suspected diagnosis of genetic disease. Of the identified newborns, 10 of them were excluded due to clear indications of a specific syndrome that could be tested by targeted analysis of known genes or structural variations. A total of 41 families were approached and offered genomic analysis with NeoSeq. Of them, 37 families consented to join the study. Of the recruited patients, four newborns were not sequenced because the genetic study of the parents was not available. Finally, 33 newborns (16 females and 17 males) met our inclusion criteria and were included in the study (Supplementary Figure S1). The median age at inclusion was 26 days (range, four days to two months). Two patients died before conclusion of the study, one prior to diagnosis and the other after diagnosis was established and palliative care initiated. For all patients, DNA was obtained from both parents. All probands were phenotyped using human phenotype ontology (HPO) terms extracted manually from electronic health records. The median number of HPO terms per proband was four, and 58% of patients had at least four terms. Based on the predominant presenting clinical symptoms, patients were classified into the

following groups: Congenital anomalies ( $n = 8$ ), neurological symptoms, including encephalopathy, developmental regression, seizures, and hypotonia ( $n = 17$ ), suspected metabolic diseases ( $n = 7$ ), and severe intrauterine growth retardation ( $n = 1$ ). We observed no significant differences in rates of diagnosis between groups ( $p = 0.572$ ). For each of the 33 patients included in this study, demographic data, clinical presentation, and standardized phenotypes using HPO terms are shown in Supplementary Table S2. Median diagnostic turnaround time was 7.5 days (range, 4–11 days).

A genetic diagnosis was established for 39.4% ( $n = 13$ ) of our 33 patients (Table 1). Analysis of inheritance patterns revealed compound heterozygous variants in six patients for the following genes (associated autosomal recessive disorders are shown in parentheses): GFM1 (combined oxidative phosphorylation deficiency 1), SUCLA2 (mitochondrial DNA depletion syndrome 5), PNPT1 (combined oxidative phosphorylation deficiency 13), CPS1 (Carbamoyl-phosphate synthetase 1[CPS1] deficiency), COQ4 (primary coenzyme Q10 deficiency), and ERBB3 (lethal congenital contracture syndrome type 3). We identified seven patients with heterozygous variants in the following genes (associated autosomal dominant disorders are shown in parentheses): KCNQ2 (early infantile epileptic encephalopathy 7, three patients), CHD2 (nemaline myopathy), COL4A1 (cerebral small vessel disease), and SOX10 (Waardenburg syndrome-Hirschsprung's disease). Of these seven patients, five carried de novo variants and two had inherited the variant from a single parent, who presented a less severe phenotype. Only in 38.4% of cases was the phenotype of the newborn predicted based on the molecular diagnosis returned by Phenomizer ( $p < 0.05$ ).

Likely pathogenic variants were identified in three patients but were ruled out owing to their presence in a healthy parent. These cases are presented in detail in Table 2.

Diagnoses of Prader–Willi syndrome and Steinert dystrophy were not detected in patients #20 and #21, respectively, but were established based on specific tests performed later in life as the respective diseases progressed. Patient #20 carried a deletion in the paternal allele in 15q11–q13 and patient #21 harbored a pathological expansion of 2333 CTG copies in the DMPK gene.

We next evaluated the impact of genetic diagnosis on four distinct aspects of medical management for a period of three months: (a) Redirection of care (towards withdrawal of intensive care or initiation of palliative care), (b) initiation of new subspecialist care, including additional extension/follow up studies, and (c) changes in medication or diet. We found that establishing a genomic diagnosis directly affected medical management in 12 of the 13 patients (Table 3), resulting in changes in medication or diet (10 patients), initiation of new subspecialist care (six patients), or withdrawal of intensive care treatment/initiation of palliative care (six patients). Moreover, all patients and families benefited from timely genetic counseling based on a concrete diagnosis.

Finally, we compared our results with those of previously published in prospective studies in which massive sequencing techniques (WGS, WES, CES) were applied to newborns admitted to NICUs with a suspected genetic disorder (Table 4). Our comparison included 12 studies published between 2016 and 2020. Nine were cohort studies, and three were randomized clinical trials. The selected studies applied the following sequencing techniques: CES (three studies) [12,13,17], WES (four studies) [16,18,19,23], WGS with subsequent filtering by CES (1 study) [14], WGS (three studies) [15,20,22], and WES + WGS (one study) [21]. The total number of patients included was 1073. The median diagnostic rate was 37.2% (range, 13.2%–58%). After adjusting for the NGS technology employed, the following diagnostic rates were obtained: CES, 43.2% (range, 32.4%–50.8%), WES, 40.4% (15.6%–56%), WGS with subsequent filtering by CES, 30.4%, and WGS, 32.8% (range, 13.2%–47.7%). The median diagnostic turnaround time was 22.3 days.

**Table 1.** Demographic and clinical data and genetic findings in patients with definitive/probable causative variants.

Cases	Age	Sex	HPO	TAT	Gene	Variants	Phenotizer	Disorder	MIM	Parents/Inheritance
1	25 d	M	Intenstive growth retardation	7 d	GFMI	c.[60A>T][1199G>C] (p.[Ile214Phe][Arg400Pro])	p = 0.0343	Combined oxidative phosphorylation deficiency 1	#60960	Carriers Autosomal recessive
			Neonatal respiratory distress Congenital microcephaly Unilateral cryptorchidism Neonatal hypoglycemia Chronic metabolic acidosis Lactic acidosis							
2	1 m 15 d	M	Methylmalonic aciduria	7 d	SUCLA2	c.[850C>T][1850C>T] (p.[Arg284Cys][Arg294Cys])	p = 0.0271	Mitochondrial DNA depletion syndrome 5	#612073	Carriers Autosomal recessive
			Abnormality of myocardium Chronic metabolic acidosis Lactic acidosis							
3	2 m	F	Failure to thrive	4 d	PNPT1	c.[1495-55_1495+80del] [1495-55_1495+80del]	n.s. p = 0.119	Combined oxidative phosphorylation deficiency 15	#614932	Carriers Autosomal recessive
			Weight loss Irritability Infantile axial hypostonia Developmental regression Hypoglycemia							
4	20 d	F	Seizures	6 d	KCNQ2	c.778C>T (p.H2620Ile)	p = 0.04	Epileptic encephalopathy, early infantile 7	#613720	De novo Autosomal dominant
			Epileptic encephalopathy							
5	18 d	M	Seizures	7 d	KCNQ2	c.1016T>A (p.Leu339Gln)	p = 0.04	Epileptic encephalopathy, early infantile 7	#613720	Father, epilepsy in infancy Autosomal dominant
			Epileptic encephalopathy							
6	25 d	F	Seizures	6 d	CDKL5	c.616G>T (p.Asp216Ile)	n.s. p = 0.1386	Epileptic encephalopathy, early infantile 2	#300672	Mother, epilepsy Autosomal dominant X-linked
			Epileptic encephalopathy							
7	6 d	M	Neonatal respiratory distress	7 d	KCNQ2	c.1686G>A (p.Arg553Gln)	n.s. p = 0.471	Familial neonatal seizures 1	#613720	De novo Autosomal dominant
			Seizures Epileptic encephalopathy Chronic metabolic acidosis							
8	10 d	M	Vomiting	8 d	CPS1	c.[1201G>C][2810T>A] (p.[Gly401Arg][Ile937Asn])	n.s. p = 0.0624	CPS1 deficiency	#68307	Carriers Autosomal recessive
			Hyperammonemia Hypoglycemia Fetal hyaline membrane disease Thyroid-stimulating hormone excess							
9	1 m	M	Generalized neonatal hypotonia	10 d	COQ4	c.[202C>C][718C>T] (p.[Asp68His][Arg40Cys])	n.d.	Primary coenzyme Q10 deficiency 7	#616276	Carriers Autosomal recessive
			Seizures Cerebellar atrophy							
10	3 d	F	Premature birth	10 d	ACTA1	c.614C>A (p.Trp215Ile)	n.s. p = 0.582	Nemaline myopathy 3	#161800	De novo Autosomal dominant
			Respiratory insufficiency due to muscle weakness Generalized hypotonia, Encephalopathy							

Table 1. Contd.

Cases	Age	Sex	HPO	TAT	Gene	Variants	Phenomizer	Disorder	MIM	Parents/Inheritance
11	9 d	F	Leukoencephalopathy Calanacts Renal cysts Bleeding digestive	5 d	COL4A1	c.2906 G>A (p.G1969Glu)	n.s. p = 0.3810	Small vessel vascular brain disease	#175780	De novo Autosomal dominant
12	12 d	M	Joint contractures Microretrognathia Microtia Microptosis Hypotonia Hearing loss	6 d	ERBB3	c.[1184-9A>G][1184-9A>C]	n.s. p = 0.4243	Lethal contracture syndrome type 2	#607598	Carriers Autosomal recessive
13	15 d	F	Hirschsprung disease Congenital hearing loss Opsoclonus	7 d	SOX10	c.850C>T (p.G14284Ter)	p = 0.0323	Wanderburg-Shah syndrome	#142623	De novo Autosomal dominant

Abbreviations: d, day; HPO, human phenotype ontology; F, female; m, month; M, male; MIM, Mendelian inheritance in Man; n.d., not detected; n.s., not significant; TAT, turnaround time to diagnosis.

Table 2. Demographic and clinical data and genetic findings in patients with non-definitive/non probable causative variants.

Cases	Age	Sex	HPO	TAT	Gene	Mutations	Phenomizer	Disorder	MIM	Parents/Inheritance
14	22 d	F	Preeclampsia Neonatal respiratory distress Hypoglycemia	7 d	MTOR	c.126G>T (p.Lys42Asn)	n.d.	Focal cortical dysplasia type II - Smith-Kingsmore Syndrome	#607341 #616638	Father asymptomatic carrier
15	29 d	F	Intrauterine growth retardation Meconium ileus Neonatal respiratory distress Generalized neonatal hypotonia Congenital hip dislocation 10 pairs of ribs	10 d	NALCN	c.2507C>G (p.Pro836Arg)	n.d.	Congenital contractures of the limbs and face, hypotonia and developmental delay	#616266	Father asymptomatic carrier
16	4 d	M	Focal seizures	9 d	TSC2	c.1724T>C (p.Leu573Pro)	n.d.	Tuberous sclerosis complex	#613254	Father asymptomatic carrier

d, day; F, female; HPO, human phenotype ontology; M, male; MIM, Mendelian inheritance in Man; n.d., not detected; TAT, turnaround time.

Table 3. Impact of molecular diagnosis on medical management.

Cases	Disorder	Changes in Medication or Diet	Initiation of New Subspecialist Care	Withdrawal of Intensive Care Treatment/Initiation of Palliative Care	Genetic Counseling
1	Combined oxidative phosphorylation deficiency 1	Mitochondrial cocktail Avoid valproate	Neurologist Hepatologist Audiologist Ophthalmologist	Yes	Yes
2	Mitochondrial DNA depletion syndrome 5	Mitochondrial cocktail	Neurologist Audiologist Ophthalmologist	Yes	Yes
3	Combined oxidative phosphorylation deficiency 13	Mitochondrial cocktail Avoid valproate	Cardiologist Audiologist Ophthalmologist	Yes	Yes
4	Epileptic encephalopathy early infantile 7	Sodium channel blocker (phenytoin) Avoid retigabine and ezogabine	No	No	Yes
5	Epileptic encephalopathy early infantile 7	Sodium channel blocker (carbamazepine) Avoid retigabine and ezogabine	No	No	Yes
6	Epileptic encephalopathy early infantile 2	Ketogenic diet	Gastroenterologist	No	Yes
7	Familial neonatal seizures 1	Sodium channel blocker (phenobarbital) Avoid retigabine and ezogabine	No	No	Yes
8	CPS1 deficiency	Protein restricted diet Arginine/citrulline	Specialist in inborn errors of metabolism	No	Yes
9	Primary coenzyme Q10 deficiency 7	CoQ10	Cardiologist Ophthalmologist Audiologist	No	Yes
10	Nemaline myopathy	Pyridostigmine Carnitine	No	Yes	Yes
11	Small vessel brain disease	-(exitus)	-(exitus)	-(exitus)	Yes
12	Lethal contracture syndrome type 2	No	No	Yes	Yes
13	Waaerdenburg-Shah syndrome	No	No	Yes (exitus)	Yes

**Table 4.** Previous prospective studies of the diagnostic performance of WES, WGS, and CES in newborns with a suspected genetic disorder.

Reference	Date	Study Type	Sequencing Type	Study Population	Rate of Diagnosis	Diagnostic Turnaround Time
Daoud, et al. [12]	2016	Cohort	CES	n = 8	4/8 (50%)	15.2 w
Meng, et al. [13]	2017	Cohort	CES	n = 178	58/178 (33%)	95 d
			Trio CES	n = 37	12/37 (32%)	51 d
			Critical trio CES	n = 63	32/63 (51%)	13 d
van Diemen, et al. [14]	2017	Cohort	WGS filtered by CES	n = 23	7/23 (30%)	12 d
Petrikin, et al. [15]	2018	RCT	Trio rWGS	n = 64	21/64 (33%)	13 d
Stark, et al. [16]	2018	Cohort	rWES	n = 40	21/40 (52%)	16 d
Brunelli, et al. [17]	2019	Cohort	rCES	n = 20	10/20 (50%)	9,6 d
Ceynab-Birsoy, et al. [18]	2019	Cohort	WES	n = 32	5/32 (16%)	-
Elliot, et al. [19]	2019	Cohort	WES	n = 25	14/25 (56%)	7.2 d
French, et al. [20]	2019	Cohort	Trio WGS	n = 106	14/106 (13%)	21 d
			rWGS	n = 94	18/94 (19%)	11 d
			rWES	n = 95	19/95 (20%)	11.2 d
Kingsmore, et al. [21]	2019	RCT	urWGS	n = 24	11/24 (46%)	4.6 d
			Trio WGS	n = 84	32/84 (38%)	4 d
Wang, et al. [22]	2020	RCT	Trio WGS	n = 84	32/84 (38%)	4 d
Gubbels, et al. [23]	2020	Cohort	Trio WES	n = 50	29/50 (56%)	4.9 d

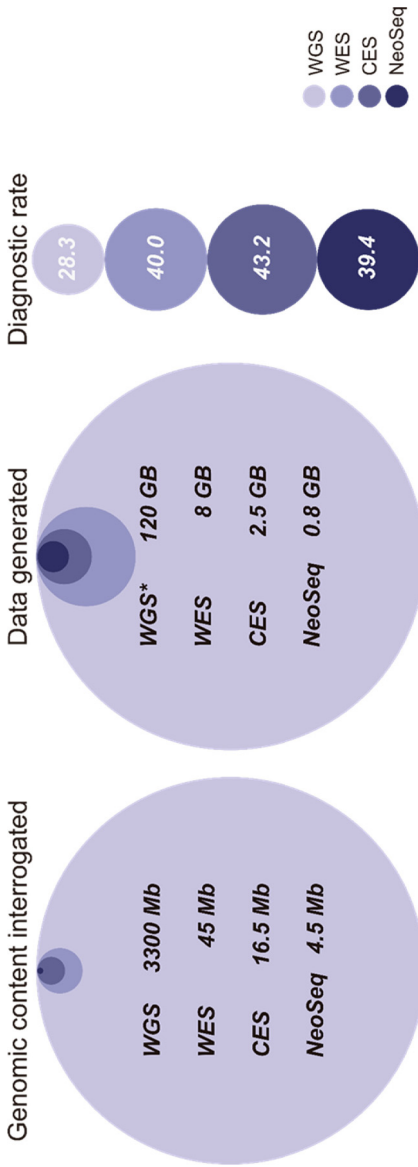
Abbreviations: CES, clinical exome sequencing; RCT, randomized clinical trial; r, rapid; ur, ultrarapid; WES, whole exome sequencing; WGS, whole genome sequencing.

#### 4. Discussion

Genomic sequencing has emerged as one of the most promising diagnostic tools in the field of neonatology. Studies have demonstrated that the adoption of this technology as a first-tier test for genetic diagnosis of severely ill newborns increases the diagnostic rate and reduces the time to diagnosis, improving outcomes while reducing health costs [29,30]. In the present study, we evaluated the diagnostic efficacy and clinical impact of NeoSeq, a gene panel that consists of 1870 genes and is specifically designed to facilitate the diagnosis of critically ill newborns with suspected genetic disorders. In total, 33 newborns and their parents underwent trio sequencing using the NeoSeq panel. Newborns with neurological symptoms comprised the largest group ( $n = 17$ ), followed by those with multiple congenital anomalies ( $n = 8$ ) and suspected metabolic diseases ( $n = 7$ ). These phenotypic groupings corresponded to those previously described in other studies that used rapid NGS strategies to diagnose NICU patients [12,23].

The resulting diagnostic rate (13/33 patients [39.4%]) was comparable to that previously described for more extensive trio or singleton sequencing approaches, such as WES [16,21] and WGS [14,15,21,22], and even higher than that reported for WGS in NICU patients [20]. While higher diagnostic rates should be expected from WGS than WES, and from WES than CES or targeted panels, we found that the diagnostic rate does not correlate with the size of genome portion interrogated (Figure 1): NeoSeq provided diagnostic rates similar to those obtained with WGS and WES approaches, despite interrogating 10 times (WES) and 500 times (WGS) fewer bases.

Two factors may help explain these discrepancies. First, variants are usually excluded if they have no predicted or known functional consequences [31–34]. This limits the analysis to exonic or intronic variants in known disease-associated genes or variants with demonstrated pathogenicity located in noncoding regions. Consequently, the phenotype of the newborn guides the analysis to a small group of candidate genes. Therefore, regardless of how extensive the genomic analysis, genetic diagnosis is limited to a shortlist of genes [14,21] and variant types [22]. Our results, and those of others using a similar approach [17], demonstrate that a simpler approach, based in the analysis of panels of genes with known functions and disease associations, can be a useful and more cost-effective alternative to WGS or WES approaches, yielding comparable diagnostic rates. The second factor to consider is the association between the diagnostic rate and the severity of the clinical phenotype of the study participants. All patients recruited in our study were critically ill newborns that required intensive care. In two previous studies in which patients with severe and less severe disease were assigned to distinct cohorts, better outcomes were observed for the severely ill group in terms of diagnostic rate and the impact on clinical management. In the NSIGHT2 trial [21], patients were randomized into two groups, which were analyzed using rapid WES (rWES) or rapid WGS (rWGS), except for severely ill patients, who were screened using ultra-rapid WGS (urWGS). The resulting diagnostic rate was 46% in the urWGS group, as compared with 19% and 20% in the rWGS in rWES groups, respectively. Meng et al. [13] used WES to perform proband exome, trio exome, and critical trio exome sequencing (a rapid genomic assay for seriously ill infants) in three groups of infants within the first 100 days of life. The diagnostic rate in the critical trio exome group was higher than that obtained for the proband and trio exome groups. Moreover, molecular diagnoses directly affected medical management in 76.9% of patients in the critical trio exome group, as compared with 42.9% in the other groups. These results suggest that a diagnosis strategy based on the sequencing of a panel of genes associated with severe neonatal phenotypes yields better outcomes because those patients likely have a genetic disorder and can greatly benefit from accurate and early diagnosis.



\* indicates 30X coverage. For all other sequencing approaches coverage = 100X.

**Figure 1.** Schematic comparing NeoSeq and other diagnostic genome sequencing strategies (WGS<sup>14,15,20,22</sup>, WES<sup>16,18,19,23</sup>, and CES<sup>12,13,17</sup>) showing the size of the genomic portion interrogated, the amount of data generated, and the median diagnostic rates achieved for each approach. To facilitate visual comparison, data are represented as the relative area of each circle. For WGS and WES, information on the genomic portion interrogated and data generated data were extracted from [www.illumina.com/wes-wgs](http://www.illumina.com/wes-wgs). For CES, information on the genomic portion interrogated and data generated was estimated for 6000 genes. Diagnostic rates were calculated as the mean of the results of the cited studies. Total number of patients for each genome sequencing strategy: CES, n = 306; WES, n = 242; WGS, n = 395.

In our study, we applied a family trio analysis approach (testing newborns and both parents), as this is an effective strategy to manage the wealth of genetic variants identified by NGS. This approach can be used to easily identify de novo variants, filter out rare benign familial variants, and establish inheritance patterns in recessive disorders [18,34]. We identified five de novo causative variants among our patients and, most importantly, in three cases, we ruled out potential phenotype-associated variants after detecting them in a single asymptomatic parent (see Table 2). In our cohort, two patients for whom NeoSeq failed to establish a molecular diagnosis were later diagnosed with Prader–Willi syndrome and Steinert dystrophy, respectively. This is a common limitation of current NGS technologies, including WGS and WES, which lack the capacity to identify repetitive sequences, homologous genes, and epigenetic modifications [35,36].

In addition to the diagnostic rate of 39.4% (13/33 patients), the rapid molecular diagnosis achieved with NeoSeq (mean diagnostic turnaround time, 7.5 days) impacted medical management in 92.3% of patients, as reflected in changes in medication or diet (10/13 patients), initiation of new subspecialist care (6/13 patients), or withdrawal of intensive care treatment/redirection to palliative care (6/13 patients). In two patients, NeoSeq enabled diagnosis of inborn errors of metabolism: CPS1 deficiency and primary coenzyme Q10 deficiency, allowing the early initiation of ammonia scavenger therapy combined with a protein-restricted diet [37] and coenzyme Q10 supplementation [38], respectively. In CPS1 deficiency treatment can prevent hyperammonemic crisis and adverse neurological outcomes, while in primary forms of coenzyme Q10 deficiency treatment can prevent the progression of both steroid-resistant nephrotic syndrome and encephalopathy, hence the critical importance of a prompt diagnosis. In three other patients who presented primarily with epilepsy, gene sequencing identified de novo *KCNQ2* variants, for which anticonvulsant treatment with sodium channel blockers is highly recommended [39] and reduces the neurodevelopmental impairment associated with the disease. Although some of the established diagnoses did not lead to institution of effective treatments, they provided important information regarding prognosis and disease management, putting an end to a potentially lengthy diagnostic odyssey. Furthermore, a genetic diagnosis allows for testing of other at-risk family members and provision of reproductive counseling where necessary [24,40].

Analysis of diagnostic rate according to phenotype shows that neonates with neurological symptoms and suspected inborn errors of metabolism benefited most from the NeoSeq panel (diagnostic yield, 47.1% and 42.9%, respectively). By contrast, only in two patients with congenital anomalies was a definitive causative variant identified (28.6%). Phenomizer predicted a genetic diagnosis based on phenotype with statistical significance ( $p < 0.05$ ) in only five patients. This rate of diagnosis prediction is similar to that reported by Brunelli et al. [17] using a targeted gene panel (40%), and higher than that reported by French et al. [20] (10%) using WGS. Based on their data, French et al. proposed agnostic analysis of genomic data, as opposed to phenotype-driven analysis. However, analyses using WGS or WES cannot be completely agnostic owing to the aforementioned limitations affecting variant interpretation.

## 5. Conclusions

In summary, rapid trio sequencing with our phenotype-driven gene panel specific for genetic diseases in critically ill newborns is an affordable alternative to WGS and WES that offers comparable diagnostic efficacy, supporting its implementation in routine clinical practice as a first-tier genetic test in NICUs.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2077-0383/9/8/2362/s1>, Figure S1: Flow chart depicting recruitment summary of cohort, Table S1: List of genes included in NeoSeq; Table S2: Patient demographic and clinical data.

**Author Contributions:** M.-L.C., E.G.-V. and M.J.d.C. conceptualized and designed the study. E.S., S.R., O.L.-S. and A.P.-M. participated in data collection and clinical evaluation. E.G.-V. and M.-L.C. coordinated and supervised data collection. M.J.d.C., S.B.-G., E.G.-V. drafted the initial manuscript. M.-L.C. critically reviewed the manuscript for important intellectual content. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study was supported by the Spanish Ministry of Science and Innovation, and the Instituto de Salud Carlos III (PI17/01294). Sofia Gouveia was supported by a Health Research Institute of Santiago (IDIS) pre-doctoral grant.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A. Bioinformatic Analysis

**Primary:** alignment, filtering quality readings, variant detection and annotation. Illumina SCS 2.6.26 / RTA 1.8.70 + CASAVA 1.8.2, FastQC v0.11.5 software were used for quality control. The alignment software used was BWA v0.7. Statistical analyses were performed using the R statistical package. TEQC v3.18 was used for pre-variant detection control, BEDTools v2.25 and Picard v2.17.10 for intermediate steps, SAMtools v0.1.19, GATK v3.7, Scalpel v0.5.4 and Varscan v2.3 for detection of variants, and SnpEff v4.3 software for variant annotation against dbSNP and RefSeq, in addition to other programs developed by the Bioinformatics Unit of our Center (including the PattREC program for CNV detection).

**Secondary:** variant filtering by frequency. R scripts designed in house were used to filter variants in each individual based on their frequency in known databases (1000 genomes, ExAC, ClinVar, EVS, dbSNP) and in our own database in order to rule out polymorphisms.

**Secondary:** variant filtering according to different inheritance hypotheses. Filtering programs (R-based programming) were also used to analyze data from exomes for members of the same family according to the different inheritance hypotheses. Finally, the variants considered most likely to cause the pathological phenotype were prioritized for each family and categorized according to the recommendations of the American College of Medical Genetics (ACMG). Once variants were prioritized according to frequency and the different inheritance hypotheses, a theoretical evaluation of their possible implications and functional effects was conducted using the in-silico prediction programs CONDEL and Human Splicing Finder.

## References

1. Wojcik, M.H.; Schwartz, T.S.; Thiele, K.E.; Paterson, H.; Stadelmaier, R.; Mullen, T.E.; Vannoy, G.E.; Genetti, C.A.; Madden, J.A.; Gubbels, C.S.; et al. Infant mortality: The contribution of genetic disorders. *J. Perinatol.* **2019**, *39*, 1611–1619. [[CrossRef](#)] [[PubMed](#)]
2. Wojcik, M.H.; Schwartz, T.S.; Yamin, I.; Edward, H.L.; Genetti, C.A.; Towne, M.C.; Agrawal, P.B. Genetic disorders and mortality in infancy and early childhood: Delayed diagnoses and missed opportunities. *Genet. Med.* **2018**, *20*, 1396–1404. [[CrossRef](#)] [[PubMed](#)]
3. Online Mendelian Inheritance in Man. Available online: <https://omim.org/> (accessed on 20 February 2020).
4. Arya, V.B.; Guemes, M.; Nessa, A.; Alam, S.; Shah, P.; Gilbert, C.; Senniappan, S.; Flanagan, S.E.; Ellard, S.; Hussain, K. Clinical and histological heterogeneity of congenital hyperinsulinism due to paternally inherited heterozygous ABCC8/KCNJ11 mutations. *Eur. J. Endocrinol.* **2014**, *171*, 685–695. [[CrossRef](#)] [[PubMed](#)]
5. Inoue, S.; Mangat, C.; Rafe’E, Y.; Sharman, M. Forme Fruste of HLH (haemophagocytic lymphohistiocytosis): Diagnostic and therapeutic challenges. *BMJ Case Rep.* **2015**, *29*. [[CrossRef](#)] [[PubMed](#)]
6. Kieran, E.; Sara, R.; Claydon, J.; Hait, V.; De Salaberry, J.; Osioviich, H.; Shivananda, S. Outcomes of Neonates With Complex Medical Needs. *Adv. Neonatal Care* **2019**, *19*, 275–284. [[CrossRef](#)] [[PubMed](#)]
7. Knijnenburg, T.A.; Vockley, J.G.; Chambwe, N.; Gibbs, D.L.; Humphries, C.; Huddleston, K.C.; Klein, E.; Kothiyal, P.; Tasseff, R.; Dhankani, V.; et al. Genomic and molecular characterization of preterm birth. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 5819–5827. [[CrossRef](#)]
8. Reuter, J.A.; Spacek, D.V.; Snyder, M.P. High-Throughput Sequencing Technologies. *Mol. Cell* **2015**, *58*, 586–597. [[CrossRef](#)]
9. Gaff, C.L.; Winship, I.M.; Forrest, S.M.; Hansen, D.; Clark, J.; Waring, P.M.; South, M.; Sinclair, A.H. Preparing for genomic medicine: A real world demonstration of health system change. *NPJ Genom. Med.* **2017**, *2*, 16. [[CrossRef](#)]

10. Holm, I.A.; Agrawal, P.B.; Ceyhan-Birsoy, O.; Christensen, K.D.; Fayer, S.; Frankel, L.A.; Genetti, C.A.; Krier, J.B.; LaMay, R.C.; Levy, H.L.; et al. The BabySeq project: Implementing genomic sequencing in newborns. *BMC Pediatr.* **2018**, *18*, 225. [[CrossRef](#)]
11. Friedman, J.M.; Bombard, Y.; Cornel, M.C.; Fernandez, C.V.; Junker, A.K.; Plon, S.E.; Stark, Z.; Knoppers, B.M. Genome-wide sequencing in acutely ill infants: Genomic medicine's critical application? *Genet. Med.* **2019**, *21*, 498–504. [[CrossRef](#)]
12. Daoud, H.; Luco, S.M.; Li, R.; Bareke, E.; Beaulieu, C.; Jarinova, O.; Carson, N.; Nikkel, S.M.; Graham, G.E.; Richer, J.; et al. Next-generation sequencing for diagnosis of rare diseases in the neonatal intensive care unit. *Can. Med. Assoc. J.* **2016**, *188*, E254–E260. [[CrossRef](#)] [[PubMed](#)]
13. Meng, L.; Pamm, M.; Saronwala, A.; Magoulas, P.; Ghazi, A.R.; Vetrini, F.; Zhang, J.; He, W.; Dharmadhikari, A.; Qu, C.; et al. Use of Exome Sequencing for Infants in Intensive Care Units: Ascertainment of Severe Single-Gene Disorders and Effect on Medical Management. *JAMA Pediatr.* **2017**, *171*, e173438. [[CrossRef](#)] [[PubMed](#)]
14. Van Diemen, C.C.; Kerstjens-Frederikse, W.S.; Bergman, K.A.; De Koning, T.J.; Sikkema-Raddatz, B.; Van Der Velde, J.K.; Abbott, K.M.; Herkert, J.C.; Löhner, K.; Rump, P.; et al. Rapid Targeted Genomics in Critically Ill Newborns. *Pediatrics* **2017**, *140*, e20162854. [[CrossRef](#)] [[PubMed](#)]
15. Petrikin, J.E.; Cakici, J.A.; Clark, M.M.; Willig, L.K.; Sweeney, N.M.; Farrow, E.; Saunders, C.J.; Thiffault, I.; Miller, N.A.; Zellmer, L.; et al. The NSIGHT1-randomized controlled trial: Rapid whole-genome sequencing for accelerated etiologic diagnosis in critically ill infants. *NPJ Genom. Med.* **2018**, *3*, 6. [[CrossRef](#)]
16. Stark, Z.; Tan, T.Y.; Chong, B.; Brett, G.R.; Yap, P.; Walsh, M.; Yeung, A.; Peters, H.; A Mordaunt, D.; Cowie, S.; et al. A prospective evaluation of whole-exome sequencing as a first-tier molecular test in infants with suspected monogenic disorders. *Genet. Med.* **2016**, *18*, 1090–1096. [[CrossRef](#)]
17. Brunelli, L.; Jenkins, S.M.; Gudgeon, J.M.; Bleyl, S.B.; Miller, C.E.; Tvrđik, T.; Dames, S.A.; Ostrander, B.; Daboub, J.A.F.; Zielinski, B.A.; et al. Targeted gene panel sequencing for the rapid diagnosis of acutely ill infants. *Mol. Genet. Genom. Med.* **2019**, *7*, e00796. [[CrossRef](#)]
18. Ceyhan-Birsoy, O.; Murry, J.B.; Machini, K.; Lebo, M.S.; Yu, T.W.; Fayer, S.; Genetti, C.A.; Schwartz, T.S.; Agrawal, P.B.; Parad, R.B.; et al. Interpretation of Genomic Sequencing Results in Healthy and Ill Newborns: Results from the BabySeq Project. *Am. J. Hum. Genet.* **2019**, *104*, 76–93. [[CrossRef](#)]
19. Elliott, A.M.; du Souich, C.; Lehman, A.; Guella, I.; Evans, D.M.; Candido, T.; Tooman, L.; Armstrong, L.; Clarke, L.A.; Gibson, W.; et al. RAPIDOMICS: Rapid genome-wide sequencing in a neonatal intensive care unit—successes and challenges. *Eur. J. Pediatr.* **2019**, *178*, 1207–1218. [[CrossRef](#)]
20. French, C.E.; Disease, N.B.; Delon, I.; Dolling, H.; Sanchis-Juan, A.; Shamardina, O.; Mégy, K.; Abbs, S.; Austin, T.; Bowdin, S.; et al. Whole genome sequencing reveals that genetic conditions are frequent in intensively ill children. *Intensiv. Care Med.* **2019**, *45*, 627–636. [[CrossRef](#)]
21. Kingsmore, S.F.; Cakici, J.A.; Clark, M.M.; Gaughran, M.; Feddock, M.; Batalov, S.; Bainbridge, M.N.; Carroll, J.; Caylor, S.A.; Clarke, C.; et al. A Randomized, Controlled Trial of the Analytic and Diagnostic Performance of Singleton and Trio, Rapid Genome and Exome Sequencing in Ill Infants. *Am. J. Hum. Genet.* **2019**, *105*, 719–733. [[CrossRef](#)]
22. Wang, H.; Lu, Y.; Dong, X.; Lu, G.; Cheng, G.; Qian, Y.; Ni, Q.; Zhang, P.; Yang, L.; Wu, B.; et al. Optimized trio genome sequencing (OTGS) as a first-tier genetic test in critically ill infants: Practice in China. *Hum. Genet.* **2020**, *139*, 473–482. [[CrossRef](#)]
23. Gubbels, C.S.; VanNoy, G.E.; Madden, J.A.; Copenheaver, D.; Yang, S.; Wojcik, M.H.; Gold, N.B.; Genetti, C.A.; Stoler, J.; Parad, R.B.; et al. Prospective, phenotype-driven selection of critically ill neonates for rapid exome sequencing is associated with high diagnostic yield. *Genet. Med.* **2020**, *22*, 736–744. [[CrossRef](#)] [[PubMed](#)]
24. Ross, L.F.; Clayton, E.W. Ethical Issues in Newborn Sequencing Research: The Case Study of BabySeq. *Pediatrics* **2019**, *144*, e20191031. [[CrossRef](#)] [[PubMed](#)]
25. Yang, H.; Robinson, P.N.; Wang, K. Phenolyzer: Phenotype-based prioritization of candidate genes for human diseases. *Nat. Methods* **2015**, *12*, 841–843. [[CrossRef](#)] [[PubMed](#)]
26. Wise, A.L.; A Manoliot, T.; A Mensah, G.; Peterson, J.F.; Roden, D.M.; Tamburro, C.; Williams, M.S.; Green, E.D. Genomic medicine for undiagnosed diseases. *Lancet* **2019**, *394*, 533–540. [[CrossRef](#)]
27. Smith, L.D.; Willig, L.K.; Kingsmore, S.F. Whole-Exome Sequencing and Whole-Genome Sequencing in Critically Ill Neonates Suspected to Have Single-Gene Disorders. *Cold Spring Harb. Perspect. Med.* **2015**, *6*, a023168. [[CrossRef](#)]

28. Köhler, S.; Schulz, M.H.; Krawitz, P.; Bauer, S.; Dölken, S.; Ott, C.E.; Mundlos, C.; Horn, D.; Mundlos, S.; Robinson, P.N. Clinical diagnostics in human genetics with semantic similarity searches in ontologies. *Am. J. Hum. Genet.* **2009**, *85*, 457–464. [[CrossRef](#)]
29. Farnaes, L.; Hildreth, A.; Sweeney, N.M.; Clark, M.M.; Chowdhury, S.; Nahas, S.; Cakici, J.; Benson, W.; Kaplan, R.H.; Kronick, R.; et al. Rapid whole-genome sequencing decreases infant morbidity and cost of hospitalization. *NPJ Genom. Med.* **2018**, *3*, 10. [[CrossRef](#)]
30. Clark, M.M.; Stark, Z.; Farnaes, L.; Tan, T.Y.; White, S.M.; Dimmock, D.; Kingsmore, S.F. Meta-analysis of the diagnostic and clinical utility of genome and exome sequencing and chromosomal microarray in children with suspected genetic diseases. *NPJ Genom. Med.* **2018**, *3*, 16. [[CrossRef](#)]
31. Wright, C.F.; FitzPatrick, D.R.; Firth, H.V. Paediatric genomics: Diagnosing rare disease in children. *Nat. Rev. Genet.* **2018**, *19*, 253–268. [[CrossRef](#)]
32. Worthey, E.A. Analysis and Annotation of Whole-Genome or Whole-Exome Sequencing Derived Variants for Clinical Diagnosis. *Curr. Protoc. Hum. Genet.* **2017**, *95*, 9.24.1–9.24.28. [[CrossRef](#)] [[PubMed](#)]
33. Bacchelli, C.; Williams, H.J. Opportunities and technical challenges in next-generation sequencing for diagnosis of rare pediatric diseases. *Expert Rev. Mol. Diagn.* **2016**, *16*, 1073–1082. [[CrossRef](#)] [[PubMed](#)]
34. Petrikin, J.E.; Willig, L.K.; Smith, L.D.; Kingsmore, S.F. Rapid whole genome sequencing and precision neonatology. *Semin. Perinatol.* **2015**, *39*, 623–631. [[CrossRef](#)] [[PubMed](#)]
35. Mantere, T.; Kersten, S.; Hoischen, A. Long-Read Sequencing Emerging in Medical Genetics. *Front. Genet.* **2019**, *10*, 426. [[CrossRef](#)] [[PubMed](#)]
36. Neu, M.B.; Bowling, K.M.; Cooper, G.M. Clinical utility of genomic sequencing. *Curr. Opin. Pediatr.* **2019**, *31*, 732–738. [[CrossRef](#)] [[PubMed](#)]
37. Häberle, J.; Burlina, A.; Chakrapani, A.; Dixon, M.; Karall, D.; Lindner, M.; Mandel, H.; Martinelli, D.; Pintos-Morell, G.; Santer, R.; et al. Suggested guidelines for the diagnosis and management of urea cycle disorders: First revision. *J. Inherit. Metab. Dis.* **2019**, *42*, 1192–1230. [[CrossRef](#)]
38. Desbats, M.A.; Lunardi, G.; Doimo, M.; Trevisson, E.; Salviati, L. Genetic bases and clinical manifestations of coenzyme Q10 (CoQ10) deficiency. *J. Inherit. Metab. Dis.* **2014**, *38*, 145–156. [[CrossRef](#)]
39. Pisano, T.; Numis, A.L.; Heavin, S.B.; Weckhuysen, S.; Angriman, M.; Suls, A.; Podestà, B.; Thibert, R.L.; Shapiro, K.A.; Guerrini, R.; et al. Early and effective treatment of KCNQ2encephalopathy. *Epilepsia* **2015**, *56*, 685–691. [[CrossRef](#)]
40. Vannoy, G.E.; Genetti, C.A.; McGuire, A.L.; Green, R.C.; Beggs, A.H.; Holm, I.A.; BabySeq Project Group. Challenging the Current Recommendations for Carrier Testing in Children. *Pediatrics* **2019**, *143*, S27–S32. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

## 10. ANEXO III: Permiso de reproducción de la Figura 1

----- Forwarded message -----

De: **Copyrights** <[copyrights@aafp.org](mailto:copyrights@aafp.org)>

Date: mér., 12 de maio de 2021 ás 18:17

Subject: RE: Permission to use a content from AFP Journal

To: María José de Castro López <[mj.decastrol@gmail.com](mailto:mj.decastrol@gmail.com)>

Good morning,

Thank you for your email. You may use these figures in your thesis.

Kind regards,

**From:** María José de Castro López <[mj.decastrol@gmail.com](mailto:mj.decastrol@gmail.com)>

**Sent:** Tuesday, May 11, 2021 11:56 AM

**To:** Copyrights <[copyrights@aafp.org](mailto:copyrights@aafp.org)>

**Subject:** Permission to use a content from AFP Journal

**CAUTION:** This email came from an EXTERNAL address. Use caution when clicking links or opening attachments.

To whom it may concern,

I would like to reuse a particular figure representing a child with Noonan Syndrome from an article published by AFP Journal (<https://www.aafp.org/aafp/2014/0101/p37.html>) in my doctoral thesis. Could you please provide me a permission to use this content with the above mentioned purpose?

Best regards,MJ



## 11. REFERENCIAS

1. Nguengang Wakap S, Lambert DM, Olry A, Rodwell C, Gueydan C, Lanneau V, et al. Estimating cumulative point prevalence of rare diseases: analysis of the Orphanet database. *Eur J Hum Genet.* 2020;28(2):165-73.
2. Angelis A, Tordrup D, Kanavos P. Socio-economic burden of rare diseases: A systematic review of cost of illness evidence. *Health Policy.* 2015;119(7):964-79.
3. Moffitt KB, Case AP, Farag NH, Canfield MA. Hospitalization charges for children with birth defects in Texas, 2001 to 2010. *Birth Defects Res A Clin Mol Teratol.* 2016;106(3):155-63.
4. Amberger JS, Bocchini CA, Schiettecatte F, Scott AF, Hamosh A. OMIM.org: Online Mendelian Inheritance in Man (OMIM(R)), an online catalog of human genes and genetic disorders. *Nucleic Acids Res.* 2015;43(Database issue):D789-98.
5. Wright CF, FitzPatrick DR, Firth HV. Paediatric genomics: diagnosing rare disease in children. *Nat Rev Genet.* 2018;19(5):325.
6. Metcalfe SA. Genetic counselling, patient education, and informed decision-making in the genomic era. *Semin Fetal Neonatal Med.* 2018;23(2):142-9.
7. EURORDIS. What is a rare disease? [Available from: <http://www.eurordis.org/about-rare-diseases>. .
8. Ramalle-Gomara E, Dominguez-Garrido E, Gomez-Eguilaz M, Marzo-Sola ME, Ramon-Trapero JL, Gil-de-Gomez J. Education and information needs for physicians about rare diseases in Spain. *Orphanet J Rare Dis.* 2020;15(1):18.
9. Evans WR, Rafi I. Rare diseases in general practice: recognising the zebras among the horses. *Br J Gen Pract.* 2016;66(652):550-1.

10. Austin CP, Cutillo CM, Lau LPL, Jonker AH, Rath A, Julkowska D, et al. Future of Rare Diseases Research 2017-2027: An IRDiRC Perspective. *Clin Transl Sci*. 2018;11(1):21-7.
11. Kochinke K, Zweier C, Nijhof B, Fenckova M, Cizek P, Honti F, et al. Systematic Phenomics Analysis Deconvolutes Genes Mutated in Intellectual Disability into Biologically Coherent Modules. *Am J Hum Genet*. 2016;98(1):149-64.
12. Grozeva D, Carss K, Spasic-Boskovic O, Tejada MI, Gez J, Shaw M, et al. Targeted Next-Generation Sequencing Analysis of 1,000 Individuals with Intellectual Disability. *Hum Mutat*. 2015;36(12):1197-204.
13. Iwase S, Berube NG, Zhou Z, Kasri NN, Battaglioli E, Scandaglia M, et al. Epigenetic Etiology of Intellectual Disability. *J Neurosci*. 2017;37(45):10773-82.
14. Del Campo M, Jones KL. A review of the physical features of the fetal alcohol spectrum disorders. *Eur J Med Genet*. 2017;60(1):55-64.
15. Torgerson PR, Mastroiacovo P. The global burden of congenital toxoplasmosis: a systematic review. *Bull World Health Organ*. 2013;91(7):501-8.
16. Eppes C, Rac M, Dunn J, Versalovic J, Murray KO, Suter MA, et al. Testing for Zika virus infection in pregnancy: key concepts to deal with an emerging epidemic. *Am J Obstet Gynecol*. 2017;216(3):209-25.
17. von der Lippe C, Diesen PS, Feragen KB. Living with a rare disorder: a systematic review of the qualitative literature. *Mol Genet Genomic Med*. 2017;5(6):758-73.
18. Wojcik MH, Schwartz TS, Thiele KE, Paterson H, Stadelmaier R, Mullen TE, et al. Infant mortality: the contribution of genetic disorders. *J Perinatol*. 2019;39(12):1611-9.
19. Hudome SM, Kirby RS, Senner JW, Cunniff C. Contribution of genetic disorders to neonatal mortality in a regional intensive care setting. *Am J Perinatol*. 1994;11(2):100-3.

20. Erratum. Jacob et al. Etiologies of NICU deaths. *Pediatrics*. 2015;135(1):e59-e65. *Pediatrics*. 2015;135(4):775-7.
21. Yoon PW, Olney RS, Khoury MJ, Sappenfield WM, Chavez GF, Taylor D. Contribution of birth defects and genetic diseases to pediatric hospitalizations. A population-based study. *Arch Pediatr Adolesc Med*. 1997;151(11):1096-103.
22. Meng M, Zhang YP. Impact of inborn errors of metabolism on admission in a neonatal intensive care unit: a 4-year report. *J Pediatr Endocrinol Metab*. 2013;26(7-8):689-93.
23. Ely DM, Driscoll AK, Matthews TJ. Infant Mortality by Age at Death in the United States, 2016. *NCHS Data Brief*. 2018(326):1-8.
24. Wen SW, Liu S, Joseph KS, Rouleau J, Allen A. Patterns of infant mortality caused by major congenital anomalies. *Teratology*. 2000;61(5):342-6.
25. Costa S, Rodrigues M, Centeno MJ, Martins A, Vilan A, Brandao O, et al. Diagnosis and cause of death in a neonatal intensive care unit--how important is autopsy? *J Matern Fetal Neonatal Med*. 2011;24(5):760-3.
26. Solomon BD, Muenke M. When to suspect a genetic syndrome. *Am Fam Physician*. 2012;86(9):826-33.
27. Hunter AG. Medical genetics: 2. The diagnostic approach to the child with dysmorphic signs. *CMAJ*. 2002;167(4):367-72.
28. Jones KL. *Smith's Recognizable Patterns of Human Malformation*: Philadelphia, Pa: Elsevier-Saunders; 2014.
29. Bishop BN, Ward R, Brady R. Velocardiofacial Syndrome. *StatPearls*. Treasure Island (FL)2020.
30. Moore CA, Krishnan K. Bone Marrow Failure. *StatPearls*. Treasure Island (FL)2020.
31. Feliciano DM. The Neurodevelopmental Pathogenesis of Tuberous Sclerosis Complex (TSC). *Front Neuroanat*. 2020;14:39.

32. Krauser AF, Schury MP. Holt Oram Syndrome. StatPearls. Treasure Island (FL)2020.
33. Moeschler JB, Shevell M, Committee on G. Comprehensive evaluation of the child with intellectual disability or global developmental delays. *Pediatrics*. 2014;134(3):e903-18.
34. Kahler SG, Fahey MC. Metabolic disorders and mental retardation. *Am J Med Genet C Semin Med Genet*. 2003;117C(1):31-41.
35. Wojcik MH, Schwartz TS, Yamin I, Edward HL, Genetti CA, Towne MC, et al. Genetic disorders and mortality in infancy and early childhood: delayed diagnoses and missed opportunities. *Genet Med*. 2018;20(11):1396-404.
36. Online Mendelian Inheritance in Man. Available online: <https://omim.org/> (accessed on 20 February 2020). [
37. Inoue S, Mangat C, Rafe'e Y, Sharman M. Forme Fruste of HLH (haemophagocytic lymphohistiocytosis): diagnostic and therapeutic challenges. *BMJ Case Rep*. 2015;2015.
38. Kieran E, Sara R, Claydon J, Hait V, de Salaberry J, Osiovich H, et al. Outcomes of Neonates With Complex Medical Needs. *Adv Neonatal Care*. 2019;19(4):275-84.
39. Katsanis SH, Katsanis N. Molecular genetic testing and the future of clinical genomics. *Nat Rev Genet*. 2013;14(6):415-26.
40. Brewington J, Clancy JP. Diagnostic Testing in Cystic Fibrosis. *Clin Chest Med*. 2016;37(1):31-46.
41. Aartsma-Rus A, Ginjaar IB, Bushby K. The importance of genetic diagnosis for Duchenne muscular dystrophy. *J Med Genet*. 2016;53(3):145-51.
42. Bick D, Jones M, Taylor SL, Taft RJ, Belmont J. Case for genome sequencing in infants and children with rare, undiagnosed or genetic diseases. *J Med Genet*. 2019;56(12):783-91.
43. Beliveau BJ, Joyce EF, Apostolopoulos N, Yilmaz F, Fonseka CY, McCole RB, et al. Versatile design and synthesis platform for

- visualizing genomes with Oligopaint FISH probes. *Proc Natl Acad Sci U S A*. 2012;109(52):21301-6.
44. Shaffer LG, Kashork CD, Saleki R, Rorem E, Sundin K, Ballif BC, et al. Targeted genomic microarray analysis for identification of chromosome abnormalities in 1500 consecutive clinical cases. *J Pediatr*. 2006;149(1):98-102.
  45. Zepeda-Mendoza CJ, Ibn-Salem J, Kammin T, Harris DJ, Rita D, Gripp KW, et al. Computational Prediction of Position Effects of Apparently Balanced Human Chromosomal Rearrangements. *Am J Hum Genet*. 2017;101(2):206-17.
  46. Ansorge WJ. Next-generation DNA sequencing techniques. *N Biotechnol*. 2009;25(4):195-203.
  47. Ece Solmaz A, Onay H, Atik T, Aykut A, Cerrah Gunes M, Ozalp Yuregir O, et al. Targeted multi-gene panel testing for the diagnosis of Bardet Biedl syndrome: Identification of nine novel mutations across BBS1, BBS2, BBS4, BBS7, BBS9, BBS10 genes. *Eur J Med Genet*. 2015;58(12):689-94.
  48. Cheng AY, Teo YY, Ong RT. Assessing single nucleotide variant detection and genotype calling on whole-genome sequenced individuals. *Bioinformatics*. 2014;30(12):1707-13.
  49. de Ligt J, Boone PM, Pfundt R, Vissers LE, Richmond T, Geoghegan J, et al. Detection of clinically relevant copy number variants with whole-exome sequencing. *Hum Mutat*. 2013;34(10):1439-48.
  50. Budworth H, McMurray CT. A brief history of triplet repeat diseases. *Methods Mol Biol*. 2013;1010:3-17.
  51. Halvorsen M, Petrovski S, Shellhaas R, Tang Y, Crandall L, Goldstein D, et al. Mosaic mutations in early-onset genetic diseases. *Genet Med*. 2016;18(7):746-9.
  52. Saudi Mendeliome G. Comprehensive gene panels provide advantages over clinical exome sequencing for Mendelian diseases. *Genome Biol*. 2015;16:134.

53. Sun Y, Ruivenkamp CA, Hoffer MJ, Vrijenhoek T, Kriek M, van Asperen CJ, et al. Next-generation diagnostics: gene panel, exome, or whole genome? *Hum Mutat.* 2015;36(6):648-55.
54. Wright CF, Fitzgerald TW, Jones WD, Clayton S, McRae JF, van Kogelenberg M, et al. Genetic diagnosis of developmental disorders in the DDD study: a scalable analysis of genome-wide research data. *Lancet.* 2015;385(9975):1305-14.
55. Wright CF, McRae JF, Clayton S, Gallone G, Aitken S, FitzGerald TW, et al. Making new genetic diagnoses with old data: iterative reanalysis and reporting from genome-wide data in 1,133 families with developmental disorders. *Genet Med.* 2018;20(10):1216-23.
56. Tan TY, Lunke S, Chong B, Phelan D, Fanjul-Fernandez M, Marum JE, et al. A head-to-head evaluation of the diagnostic efficacy and costs of trio versus singleton exome sequencing analysis. *Eur J Hum Genet.* 2019;27(12):1791-9.
57. Moorhith S, Hall A, Wright CF. Informatics and clinical genome sequencing: opening the black box. *Genet Med.* 2013;15(3):165-71.
58. Ceyhan-Birsoy O, Murry JB, Machini K, Lebo MS, Yu TW, Fayer S, et al. Interpretation of Genomic Sequencing Results in Healthy and Ill Newborns: Results from the BabySeq Project. *Am J Hum Genet.* 2019;104(1):76-93.
59. Telenti A, Pierce LC, Biggs WH, di Iulio J, Wong EH, Fabani MM, et al. Deep sequencing of 10,000 human genomes. *Proc Natl Acad Sci U S A.* 2016;113(42):11901-6.
60. Yen JL, Garcia S, Montana A, Harris J, Chervitz S, Morra M, et al. A variant by any name: quantifying annotation discordance across tools and clinical databases. *Genome Med.* 2017;9(1):7.
61. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al. The variant call format and VCFtools. *Bioinformatics.* 2011;27(15):2156-8.
62. Endrullat C, Glokler J, Franke P, Frohme M. Standardization and quality management in next-generation sequencing. *Appl Transl Genom.* 2016;10:2-9.

63. Salgado D, Bellgard MI, Desvignes JP, Beroud C. How to Identify Pathogenic Mutations among All Those Variations: Variant Annotation and Filtration in the Genome Sequencing Era. *Hum Mutat.* 2016;37(12):1272-82.
64. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GR, Thormann A, et al. The Ensembl Variant Effect Predictor. *Genome Biol.* 2016;17(1):122.
65. Shamsani J, Kazakoff SH, Armean IM, McLaren W, Parsons MT, Thompson BA, et al. A plugin for the Ensembl Variant Effect Predictor that uses MaxEntScan to predict variant spliceogenicity. *Bioinformatics.* 2019;35(13):2315-7.
66. Cingolani P, Platts A, Wang le L, Coon M, Nguyen T, Wang L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin).* 2012;6(2):80-92.
67. Bahcall OG. Genetic variation: ExAC boosts clinical variant interpretation in rare diseases. *Nat Rev Genet.* 2016;17(10):584.
68. Ulintz PJ, Wu W, Gates CM. Bioinformatics Analysis of Whole Exome Sequencing Data. *Methods Mol Biol.* 2019;1881:277-318.
69. Landry LG, Ali N, Williams DR, Rehm HL, Bonham VL. Lack Of Diversity In Genomic Databases Is A Barrier To Translating Precision Medicine Research Into Practice. *Health Aff (Millwood).* 2018;37(5):780-5.
70. Popejoy AB, Fullerton SM. Genomics is failing on diversity. *Nature.* 2016;538(7624):161-4.
71. Retterer K, Jussola J, Cho MT, Vitazka P, Millan F, Gibellini F, et al. Clinical application of whole-exome sequencing across clinical indications. *Genet Med.* 2016;18(7):696-704.
72. Ansari M, Poke G, Ferry Q, Williamson K, Aldridge R, Meynert AM, et al. Genetic heterogeneity in Cornelia de Lange syndrome (CdLS) and CdLS-like phenotypes with observed and predicted levels of mosaicism. *J Med Genet.* 2014;51(10):659-68.

73. Stenson PD, Mort M, Ball EV, Shaw K, Phillips A, Cooper DN. The Human Gene Mutation Database: building a comprehensive mutation repository for clinical and molecular genetics, diagnostic testing and personalized genomic medicine. *Hum Genet.* 2014;133(1):1-9.
74. Strande NT, Riggs ER, Buchanan AH, Ceyhan-Birsoy O, DiStefano M, Dwight SS, et al. Evaluating the Clinical Validity of Gene-Disease Associations: An Evidence-Based Framework Developed by the Clinical Genome Resource. *Am J Hum Genet.* 2017;100(6):895-906.
75. Ghouse J, Have CT, Skov MW, Andreasen L, Ahlberg G, Nielsen JB, et al. Numerous Brugada syndrome-associated genetic variants have no effect on J-point elevation, syncope susceptibility, malignant cardiac arrhythmia, and all-cause mortality. *Genet Med.* 2017;19(5):521-8.
76. Landrum MJ, Lee JM, Benson M, Brown G, Chao C, Chitipiralla S, et al. ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic Acids Res.* 2016;44(D1):D862-8.
77. Firth HV, Wright CF, Study DDD. The Deciphering Developmental Disorders (DDD) study. *Dev Med Child Neurol.* 2011;53(8):702-3.
78. Kohler S, Doelken SC, Mungall CJ, Bauer S, Firth HV, Bailleul-Forestier I, et al. The Human Phenotype Ontology project: linking molecular biology and disease through phenotype data. *Nucleic Acids Res.* 2014;42(Database issue):D966-74.
79. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med.* 2015;17(5):405-24.
80. Nykamp K, Anderson M, Powers M, Garcia J, Herrera B, Ho YY, et al. Correction: Sherlock: a comprehensive refinement of the ACMG-AMP variant classification criteria. *Genet Med.* 2020;22(1):240.
81. Bragin E, Chatzimichali EA, Wright CF, Hurles ME, Firth HV, Bevan AP, et al. DECIPHER: database for the interpretation of phenotype-linked plausibly pathogenic sequence and copy-number variation. *Nucleic Acids Res.* 2014;42(Database issue):D993-D1000.

82. Chatzimichali EA, Brent S, Hutton B, Perrett D, Wright CF, Bevan AP, et al. Facilitating collaboration in rare genetic disorders through effective matchmaking in DECIPHER. *Hum Mutat.* 2015;36(10):941-9.
83. Bowdin S, Gilbert A, Bedoukian E, Carew C, Adam MP, Belmont J, et al. Recommendations for the integration of genomics into clinical practice. *Genet Med.* 2016;18(11):1075-84.
84. Thevenon J, Duffourd Y, Masurel-Paulet A, Lefebvre M, Feillet F, El Chehadeh-Djebbar S, et al. Diagnostic odyssey in severe neurodevelopmental disorders: toward clinical whole-exome sequencing as a first-line diagnostic test. *Clin Genet.* 2016;89(6):700-7.
85. Sawyer SL, Hartley T, Dymant DA, Beaulieu CL, Schwartzenruber J, Smith A, et al. Utility of whole-exome sequencing for those near the end of the diagnostic odyssey: time to address gaps in care. *Clin Genet.* 2016;89(3):275-84.
86. Petrikin JE, Willig LK, Smith LD, Kingsmore SF. Rapid whole genome sequencing and precision neonatology. *Semin Perinatol.* 2015;39(8):623-31.
87. Saunders CJ, Miller NA, Soden SE, Dinwiddie DL, Noll A, Alnadi NA, et al. Rapid whole-genome sequencing for genetic disease diagnosis in neonatal intensive care units. *Sci Transl Med.* 2012;4(154):154ra35.
88. Kohler S, Vasilevsky NA, Engelstad M, Foster E, McMurry J, Ayme S, et al. The Human Phenotype Ontology in 2017. *Nucleic Acids Res.* 2017;45(D1):D865-D76.
89. Kohler S, Carmody L, Vasilevsky N, Jacobsen JOB, Danis D, Gourdine JP, et al. Expansion of the Human Phenotype Ontology (HPO) knowledge base and resources. *Nucleic Acids Res.* 2019;47(D1):D1018-D27.
90. Shefchek KA, Harris NL, Gargano M, Matentzoglou N, Unni D, Brush M, et al. The Monarch Initiative in 2019: an integrative data and analytic platform connecting phenotypes to genotypes across species. *Nucleic Acids Res.* 2020;48(D1):D704-D15.

91. Kohler S, Oien NC, Buske OJ, Groza T, Jacobsen JOB, McNamara C, et al. Encoding Clinical Data with the Human Phenotype Ontology for Computational Differential Diagnostics. *Curr Protoc Hum Genet.* 2019;103(1):e92.
92. Zemojtel T, Kohler S, Mackenroth L, Jager M, Hecht J, Krawitz P, et al. Effective diagnosis of genetic disease by computational phenotype analysis of the disease-associated genome. *Sci Transl Med.* 2014;6(252):252ra123.
93. Maver A, Lovrecic L, Volk M, Rudolf G, Writzl K, Blatnik A, et al. Phenotype-driven gene target definition in clinical genome-wide sequencing data interpretation. *Genet Med.* 2016;18(11):1102-10.
94. Simeoni I, Stephens JC, Hu F, Deevi SV, Megy K, Bariana TK, et al. A high-throughput sequencing test for diagnosing inherited bleeding, thrombotic, and platelet disorders. *Blood.* 2016;127(23):2791-803.
95. Baynam G, Pachter N, McKenzie F, Townshend S, Slee J, Kiraly-Borri C, et al. The rare and undiagnosed diseases diagnostic service - application of massively parallel sequencing in a state-wide clinical service. *Orphanet J Rare Dis.* 2016;11(1):77.
96. Directors ABo. Clinical utility of genetic and genomic services: a position statement of the American College of Medical Genetics and Genomics. *Genet Med.* 2015;17(6):505-7.
97. Ng SB, Buckingham KJ, Lee C, Bigham AW, Tabor HK, Dent KM, et al. Exome sequencing identifies the cause of a mendelian disorder. *Nat Genet.* 2010;42(1):30-5.
98. Taylor JC, Martin HC, Lise S, Broxholme J, Cazier JB, Rimmer A, et al. Factors influencing success of clinical genome sequencing across a broad spectrum of disorders. *Nat Genet.* 2015;47(7):717-26.
99. Brunelli L, Mao R, Jenkins SM, Bleyl SB, Dames SA, Miller CE, et al. A rapid gene sequencing panel strategy to facilitate precision neonatal medicine. *Am J Med Genet A.* 2017;173(7):1979-82.
100. Belkadi A, Bolze A, Itan Y, Cobat A, Vincent QB, Antipenko A, et al. Whole-genome sequencing is more powerful than whole-exome

- sequencing for detecting exome variants. *Proc Natl Acad Sci U S A*. 2015;112(17):5473-8.
101. Meienberg J, Bruggmann R, Oexle K, Matyas G. Clinical sequencing: is WGS the better WES? *Hum Genet*. 2016;135(3):359-62.
  102. Daoud H, Luco SM, Li R, Bareke E, Beaulieu C, Jarinova O, et al. Next-generation sequencing for diagnosis of rare diseases in the neonatal intensive care unit. *CMAJ*. 2016;188(11):E254-E60.
  103. Meng L, Pammi M, Saronwala A, Magoulas P, Ghazi AR, Vetrini F, et al. Use of Exome Sequencing for Infants in Intensive Care Units: Ascertainment of Severe Single-Gene Disorders and Effect on Medical Management. *JAMA Pediatr*. 2017;171(12):e173438.
  104. Brunelli L, Jenkins SM, Gudgeon JM, Bleyl SB, Miller CE, Tvrdik T, et al. Targeted gene panel sequencing for the rapid diagnosis of acutely ill infants. *Mol Genet Genomic Med*. 2019;7(7):e00796.
  105. Green RC, Berg JS, Grody WW, Kalia SS, Korf BR, Martin CL, et al. ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. *Genet Med*. 2013;15(7):565-74.
  106. May T. On the justifiability of ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. *J Law Med Ethics*. 2015;43(1):134-42.
  107. Bick D, Fraser PC, Gutzeit MF, Harris JM, Hambuch TM, Helbling DC, et al. Successful Application of Whole Genome Sequencing in a Medical Genetics Clinic. *J Pediatr Genet*. 2017;6(2):61-76.
  108. Friedman JM, Bombard Y, Cornel MC, Fernandez CV, Junker AK, Plon SE, et al. Genome-wide sequencing in acutely ill infants: genomic medicine's critical application? *Genet Med*. 2019;21(2):498-504.
  109. Sweeny K, Ghane A, Legg AM, Huynh HP, Andrews SE. Predictors of genetic testing decisions: a systematic review and critique of the literature. *J Genet Couns*. 2014;23(3):263-88.
  110. Holm IA, McGuire A, Pereira S, Rehm H, Green RC, Beggs AH, et al. Returning a Genomic Result for an Adult-Onset Condition to the

- Parents of a Newborn: Insights From the BabySeq Project. *Pediatrics*. 2019;143(Suppl 1):S37-S43.
111. Ross LF, Clayton EW. Ethical Issues in Newborn Sequencing Research: The Case Study of BabySeq. *Pediatrics*. 2019;144(6).
  112. Holm IA, Agrawal PB, Ceyhan-Birsoy O, Christensen KD, Fayer S, Frankel LA, et al. The BabySeq project: implementing genomic sequencing in newborns. *BMC Pediatr*. 2018;18(1):225.
  113. Ceyhan-Birsoy O, Machini K, Lebo MS, Yu TW, Agrawal PB, Parad RB, et al. A curated gene list for reporting results of newborn genomic sequencing. *Genet Med*. 2017;19(7):809-18.
  114. van Diemen CC, Kerstjens-Frederikse WS, Bergman KA, de Koning TJ, Sikkema-Raddatz B, van der Velde JK, et al. Rapid Targeted Genomics in Critically Ill Newborns. *Pediatrics*. 2017;140(4).
  115. Stark Z, Lunke S, Brett GR, Tan NB, Stapleton R, Kumble S, et al. Meeting the challenges of implementing rapid genomic testing in acute pediatric care. *Genet Med*. 2018;20(12):1554-63.
  116. Stark Z, Tan TY, Chong B, Brett GR, Yap P, Walsh M, et al. A prospective evaluation of whole-exome sequencing as a first-tier molecular test in infants with suspected monogenic disorders. *Genet Med*. 2016;18(11):1090-6.
  117. VanNoy GE, Genetti CA, McGuire AL, Green RC, Beggs AH, Holm IA, et al. Challenging the Current Recommendations for Carrier Testing in Children. *Pediatrics*. 2019;143(Suppl 1):S27-S32.
  118. Pereira S, Robinson JO, Gutierrez AM, Petersen DK, Hsu RL, Lee CH, et al. Perceived Benefits, Risks, and Utility of Newborn Genomic Sequencing in the BabySeq Project. *Pediatrics*. 2019;143(Suppl 1):S6-S13.
  119. Frankel LA, Pereira S, McGuire AL. Potential Psychosocial Risks of Sequencing Newborns. *Pediatrics*. 2016;137 Suppl 1:S24-9.
  120. Elliott AM, du Souich C, Lehman A, Guella I, Evans DM, Candido T, et al. RAPIDOMICS: rapid genome-wide sequencing in a neonatal intensive care unit-successes and challenges. *Eur J Pediatr*. 2019;178(8):1207-18.

121. Gubbels CS, VanNoy GE, Madden JA, Copenheaver D, Yang S, Wojcik MH, et al. Prospective, phenotype-driven selection of critically ill neonates for rapid exome sequencing is associated with high diagnostic yield. *Genet Med.* 2020;22(4):736-44.
122. French CE, Delon I, Dolling H, Sanchis-Juan A, Shamardina O, Megy K, et al. Whole genome sequencing reveals that genetic conditions are frequent in intensively ill children. *Intensive Care Med.* 2019;45(5):627-36.
123. Loupatty FJ, Clayton PT, Ruiter JP, Ofman R, Ijlst L, Brown GK, et al. Mutations in the gene encoding 3-hydroxyisobutyryl-CoA hydrolase results in progressive infantile neurodegeneration. *Am J Hum Genet.* 2007;80(1):195-9.
124. Gonorazky HD, Naumenko S, Ramani AK, Nelakuditi V, Mashouri P, Wang P, et al. Expanding the Boundaries of RNA Sequencing as a Diagnostic Tool for Rare Mendelian Disease. *Am J Hum Genet.* 2019;104(5):1007.
125. Schmitz-Abe K, Li Q, Rosen SM, Nori N, Madden JA, Genetti CA, et al. Unique bioinformatic approach and comprehensive reanalysis improve diagnostic yield of clinical exomes. *Eur J Hum Genet.* 2019;27(9):1398-405.
126. Australian Genomics Health Alliance Acute Care F, Lunke S, Eggers S, Wilson M, Patel C, Barnett CP, et al. Feasibility of Ultra-Rapid Exome Sequencing in Critically Ill Infants and Children With Suspected Monogenic Conditions in the Australian Public Health Care System. *JAMA.* 2020;323(24):2503-11.
127. Petrikin JE, Cakici JA, Clark MM, Willig LK, Sweeney NM, Farrow EG, et al. The NSIGHT1-randomized controlled trial: rapid whole-genome sequencing for accelerated etiologic diagnosis in critically ill infants. *NPJ Genom Med.* 2018;3:6.
128. Wang H, Lu Y, Dong X, Lu G, Cheng G, Qian Y, et al. Optimized trio genome sequencing (OTGS) as a first-tier genetic test in critically ill infants: practice in China. *Hum Genet.* 2020;139(4):473-82.
129. Kingsmore SF, Cakici JA, Clark MM, Gaughran M, Feddock M, Batalov S, et al. A Randomized, Controlled Trial of the Analytic and

- Diagnostic Performance of Singleton and Trio, Rapid Genome and Exome Sequencing in Ill Infants. *Am J Hum Genet.* 2019;105(4):719-33.
130. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* 2009;25(14):1754-60.
  131. Koboldt DC, Chen K, Wylie T, Larson DE, McLellan MD, Mardis ER, et al. VarScan: variant detection in massively parallel sequencing of individual and pooled samples. *Bioinformatics.* 2009;25(17):2283-5.
  132. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* 2010;38(16):e164.
  133. Genomes Project C, Abecasis GR, Altshuler D, Auton A, Brooks LD, Durbin RM, et al. A map of human genome variation from population-scale sequencing. *Nature.* 2010;467(7319):1061-73.
  134. Genomes Project C, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, et al. A global reference for human genetic variation. *Nature.* 2015;526(7571):68-74.
  135. Desmet FO, Hamroun D, Lalande M, Collod-Beroud G, Claustres M, Beroud C. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res.* 2009;37(9):e67.
  136. Quang D, Chen Y, Xie X. DANN: a deep learning approach for annotating the pathogenicity of genetic variants. *Bioinformatics.* 2015;31(5):761-3.
  137. Schwarz JM, Rodelsperger C, Schuelke M, Seelow D. MutationTaster evaluates disease-causing potential of sequence alterations. *Nat Methods.* 2010;7(8):575-6.
  138. Davydov EV, Goode DL, Sirota M, Cooper GM, Sidow A, Batzoglou S. Identifying a high fraction of the human genome to be under selective constraint using GERP++. *PLoS Comput Biol.* 2010;6(12):e1001025.

139. Biervert C, Schroeder BC, Kubisch C, Berkovic SF, Propping P, Jentsch TJ, et al. A potassium channel mutation in neonatal human epilepsy. *Science*. 1998;279(5349):403-6.
140. Singh NA, Charlier C, Stauffer D, DuPont BR, Leach RJ, Melis R, et al. A novel potassium channel gene, KCNQ2, is mutated in an inherited epilepsy of newborns. *Nat Genet*. 1998;18(1):25-9.
141. Dedek K, Fusco L, Teloy N, Steinlein OK. Neonatal convulsions and epileptic encephalopathy in an Italian family with a missense mutation in the fifth transmembrane region of KCNQ2. *Epilepsy Res*. 2003;54(1):21-7.
142. Weckhuysen S, Mandelstam S, Suls A, Audenaert D, Deconinck T, Claes LR, et al. KCNQ2 encephalopathy: emerging phenotype of a neonatal epileptic encephalopathy. *Ann Neurol*. 2012;71(1):15-25.
143. Saitsu H, Kato M, Koide A, Goto T, Fujita T, Nishiyama K, et al. Whole exome sequencing identifies KCNQ2 mutations in Ohtahara syndrome. *Ann Neurol*. 2012;72(2):298-300.
144. Moulard B, Picard F, le Hellard S, Agulhon C, Weiland S, Favre I, et al. Ion channel variation causes epilepsies. *Brain Res Brain Res Rev*. 2001;36(2-3):275-84.
145. Soldovieri MV, Boutry-Kryza N, Milh M, Doummar D, Heron B, Bourel E, et al. Novel KCNQ2 and KCNQ3 mutations in a large cohort of families with benign neonatal epilepsy: first evidence for an altered channel regulation by syntaxin-1A. *Hum Mutat*. 2014;35(3):356-67.
146. Ambrosino P, Alaimo A, Bartollino S, Manocchio L, De Maria M, Mosca I, et al. Epilepsy-causing mutations in Kv7.2 C-terminus affect binding and functional modulation by calmodulin. *Biochim Biophys Acta*. 2015;1852(9):1856-66.
147. Haberle J, Shchelochkov OA, Wang J, Katsonis P, Hall L, Reiss S, et al. Molecular defects in human carbamoyl phosphate synthetase I: mutational spectrum, diagnostic and protein structure considerations. *Hum Mutat*. 2011;32(6):579-89.

148. Klaus V, Vermeulen T, Minassian B, Israelian N, Engel K, Lund AM, et al. Highly variable clinical phenotype of carbamylphosphate synthetase 1 deficiency in one family: an effect of allelic variation in gene expression? *Clin Genet.* 2009;76(3):263-9.
149. Freeman JM, Nicholson JF, Schimke RT, Rowland LP, Carter S. Congenital hyperammonemia. Association with hyperglycinemia and decreased levels of carbamyl phosphate synthetase. *Arch Neurol.* 1970;23(5):430-7.
150. Brusilow SW, Danney M, Waber LJ, Batshaw M, Burton B, Levitsky L, et al. Treatment of episodic hyperammonemia in children with inborn errors of urea synthesis. *N Engl J Med.* 1984;310(25):1630-4.
151. Chung WK, Martin K, Jalas C, Braddock SR, Juusola J, Monaghan KG, et al. Mutations in COQ4, an essential component of coenzyme Q biosynthesis, cause lethal neonatal mitochondrial encephalomyopathy. *J Med Genet.* 2015;52(9):627-35.
152. Brea-Calvo G, Haack TB, Karall D, Ohtake A, Invernizzi F, Carrozzo R, et al. COQ4 mutations cause a broad spectrum of mitochondrial disorders associated with CoQ10 deficiency. *Am J Hum Genet.* 2015;96(2):309-17.
153. Ilkovski B, Cooper ST, Nowak K, Ryan MM, Yang N, Schnell C, et al. Nemaline myopathy caused by mutations in the muscle alpha-skeletal-actin gene. *Am J Hum Genet.* 2001;68(6):1333-43.
154. Mayne R, Wiedemann H, Irwin MH, Sanderson RD, Fitch JM, Linsenmayer TF, et al. Monoclonal antibodies against chicken type IV and V collagens: electron microscopic mapping of the epitopes after rotary shadowing. *J Cell Biol.* 1984;98(5):1637-44.
155. Meuwissen ME, Halley DJ, Smit LS, Lequin MH, Cobben JM, de Coo R, et al. The expanding phenotype of COL4A1 and COL4A2 mutations: clinical data on 13 newly identified families and a review of the literature. *Genet Med.* 2015;17(11):843-53.
156. Verdura E, Herve D, Bergametti F, Jacquet C, Morvan T, Prieto-Morin C, et al. Disruption of a miR-29 binding site leading to COL4A1 upregulation causes pontine autosomal dominant

- microangiopathy with leukoencephalopathy. *Ann Neurol.* 2016;80(5):741-53.
157. Plaisier E, Chen Z, Gekeler F, Benhassine S, Dahan K, Marro B, et al. Novel COL4A1 mutations associated with HANAC syndrome: a role for the triple helical CB3[IV] domain. *Am J Med Genet A.* 2010;152A(10):2550-5.
  158. Ronco P, Plaisier E. [Hereditary angiopathy with nephropathy, aneurysms and muscle cramps (HANAC): a new basement membrane-disease associated with mutations of the COL4A1 gene]. *Bull Acad Natl Med.* 2008;192(5):971-84; discussion 84-6.
  159. Narkis G, Ofir R, Manor E, Landau D, Elbedour K, Birk OS. Lethal congenital contractural syndrome type 2 (LCCS2) is caused by a mutation in ERBB3 (Her3), a modulator of the phosphatidylinositol-3-kinase/Akt pathway. *Am J Hum Genet.* 2007;81(3):589-95.
  160. Braunstein EM, Li R, Sobreira N, Marosy B, Hetrick K, Doheny K, et al. A germline ERBB3 variant is a candidate for predisposition to erythroid MDS/erythroleukemia. *Leukemia.* 2016;30(11):2242-5.
  161. Landau D, Mishori-Dery A, Hershkovitz R, Narkis G, Elbedour K, Carmi R. A new autosomal recessive congenital contractural syndrome in an Israeli Bedouin kindred. *Am J Med Genet A.* 2003;117A(1):37-40.
  162. Pusch C, Hustert E, Pfeifer D, Sudbeck P, Kist R, Roe B, et al. The SOX10/Sox10 gene from human and mouse: sequence, expression, and transactivation by the encoded HMG domain transcription factor. *Hum Genet.* 1998;103(2):115-23.
  163. Chaoui A, Kavou A, Baral V, Watanabe Y, Lecerf L, Colley A, et al. Subnuclear re-localization of SOX10 and p54NRB correlates with a unique neurological phenotype associated with SOX10 missense mutations. *Hum Mol Genet.* 2015;24(17):4933-47.
  164. Pingault V, Bondurand N, Kuhlbrodt K, Goerich DE, Prehu MO, Puliti A, et al. SOX10 mutations in patients with Waardenburg-Hirschsprung disease. *Nat Genet.* 1998;18(2):171-3.

165. Pingault V, Bodereau V, Baral V, Marcos S, Watanabe Y, Chaoui A, et al. Loss-of-function mutations in SOX10 cause Kallmann syndrome with deafness. *Am J Hum Genet.* 2013;92(5):707-24.
166. Touraine RL, Attie-Bitach T, Manceau E, Korsch E, Sarda P, Pingault V, et al. Neurological phenotype in Waardenburg syndrome type 4 correlates with novel SOX10 truncating mutations and expression in developing brain. *Am J Hum Genet.* 2000;66(5):1496-503.
167. Baynam G, Overkov A, Davis M, Mina K, Schofield L, Allcock R, et al. A germline MTOR mutation in Aboriginal Australian siblings with intellectual disability, dysmorphism, macrocephaly, and small thoraces. *Am J Med Genet A.* 2015;167(7):1659-67.
168. Moosa S, Bohrer-Rabel H, Altmuller J, Beleggia F, Nurnberg P, Li Y, et al. Smith-Kingsmore syndrome: A third family with the MTOR mutation c.5395G>A p.(Glu1799Lys) and evidence for paternal gonadal mosaicism. *Am J Med Genet A.* 2017;173(1):264-7.
169. Mirzaa GM, Campbell CD, Solovieff N, Goold C, Jansen LA, Menon S, et al. Association of MTOR Mutations With Developmental Brain Disorders, Including Megalencephaly, Focal Cortical Dysplasia, and Pigmentary Mosaicism. *JAMA Neurol.* 2016;73(7):836-45.
170. Moller RS, Weckhuysen S, Chipaux M, Marsan E, Taly V, Bebin EM, et al. Germline and somatic mutations in the MTOR gene in focal cortical dysplasia and epilepsy. *Neurol Genet.* 2016;2(6):e118.
171. Al-Sayed MD, Al-Zaidan H, Albakheet A, Hakami H, Kenana R, Al-Yafee Y, et al. Mutations in NALCN cause an autosomal-recessive syndrome with severe hypotonia, speech impairment, and cognitive delay. *Am J Hum Genet.* 2013;93(4):721-6.
172. Fukai R, Saitsu H, Okamoto N, Sakai Y, Fattal-Valevski A, Masaaki S, et al. De novo missense mutations in NALCN cause developmental and intellectual impairment with hypotonia. *J Hum Genet.* 2016;61(5):451-5.
173. Chong JX, McMillin MJ, Shively KM, Beck AE, Marvin CT, Armenteros JR, et al. De novo mutations in NALCN cause a syndrome characterized by congenital contractures of the limbs and

- face, hypotonia, and developmental delay. *Am J Hum Genet.* 2015;96(3):462-73.
174. Bend EG, Si Y, Stevenson DA, Bayrak-Toydemir P, Newcomb TM, Jorgensen EM, et al. NALCN channelopathies: Distinguishing gain-of-function and loss-of-function mutations. *Neurology.* 2016;87(11):1131-9.
  175. Peron A, Northrup H. Tuberous sclerosis complex. *Am J Med Genet C Semin Med Genet.* 2018;178(3):274-7.
  176. Ding Y, Wang J, Zhou S, Zhou Y, Zhang L, Yu L, et al. Genotype and Phenotype Analysis of Chinese Children With Tuberous Sclerosis Complex: A Pediatric Cohort Study. *Front Genet.* 2020;11:204.
  177. Dufner Almeida LG, Nanhoe S, Zonta A, Hosseinzadeh M, Kom-Gortat R, Elfferich P, et al. Comparison of the functional and structural characteristics of rare TSC2 variants with clinical and genetic findings. *Hum Mutat.* 2020;41(4):759-73.
  178. Marom D. Genetics of tuberous sclerosis complex: an update. *Childs Nerv Syst.* 2020;36(10):2489-96.
  179. Liu J, Zhao W, Ou X, Zhao Z, Hu C, Sun M, et al. Mutation spectrums of TSC1 and TSC2 in Chinese women with lymphangi leiomyomatosis (LAM). *PLoS One.* 2019;14(12):e0226400.
  180. Xu KF, Xu W, Liu S, Yu J, Tian X, Yang Y, et al. Lymphangi leiomyomatosis. *Semin Respir Crit Care Med.* 2020;41(2):256-68.
  181. Lim JS, Gopalappa R, Kim SH, Ramakrishna S, Lee M, Kim WI, et al. Somatic Mutations in TSC1 and TSC2 Cause Focal Cortical Dysplasia. *Am J Hum Genet.* 2017;100(3):454-72.
  182. Lim JS, Kim WI, Kang HC, Kim SH, Park AH, Park EK, et al. Brain somatic mutations in MTOR cause focal cortical dysplasia type II leading to intractable epilepsy. *Nat Med.* 2015;21(4):395-400.
  183. Kamsteeg EJ, Kress W, Catalli C, Hertz JM, Witsch-Baumgartner M, Buckley MF, et al. Best practice guidelines and recommendations on

- the molecular diagnosis of myotonic dystrophy types 1 and 2. *Eur J Hum Genet.* 2012;20(12):1203-8.
184. Musova Z, Mazanec R, Krepelova A, Ehler E, Vales J, Jaklova R, et al. Highly unstable sequence interruptions of the CTG repeat in the myotonic dystrophy gene. *Am J Med Genet A.* 2009;149A(7):1365-74.
185. Kumar A, Agarwal S, Agarwal D, Phadke SR. Myotonic dystrophy type 1 (DM1): a triplet repeat expansion disorder. *Gene.* 2013;522(2):226-30.
186. Butler MG, Miller JL, Forster JL. Prader-Willi Syndrome - Clinical Genetics, Diagnosis and Treatment Approaches: An Update. *Curr Pediatr Rev.* 2019;15(4):207-44.
187. Cassidy SB, Schwartz S, Miller JL, Driscoll DJ. Prader-Willi syndrome. *Genet Med.* 2012;14(1):10-26.
188. El-Hattab AW, Zarante AM, Almannai M, Scaglia F. Therapies for mitochondrial diseases and current clinical trials. *Mol Genet Metab.* 2017;122(3):1-9.
189. Bottani E, Lamperti C, Prigione A, Tiranti V, Persico N, Brunetti D. Therapeutic Approaches to Treat Mitochondrial Diseases: "One-Size-Fits-All" and "Precision Medicine" Strategies. *Pharmaceutics.* 2020;12(11).
190. Muraresku CC, McCormick EM, Falk MJ. Mitochondrial Disease: Advances in clinical diagnosis, management, therapeutic development, and preventative strategies. *Curr Genet Med Rep.* 2018;6(2):62-72.
191. Pisano T, Numis AL, Heavin SB, Weckhuysen S, Angriman M, Suls A, et al. Early and effective treatment of KCNQ2 encephalopathy. *Epilepsia.* 2015;56(5):685-91.
192. Kuersten M, Tacke M, Gerstl L, Hoelz H, Stulpnagel CV, Borggraefe I. Antiepileptic therapy approaches in KCNQ2 related epilepsy: A systematic review. *Eur J Med Genet.* 2020;63(1):103628.

193. Olson HE, Demarest ST, Pestana-Knight EM, Swanson LC, Iqbal S, Lal D, et al. Cyclin-Dependent Kinase-Like 5 Deficiency Disorder: Clinical Review. *Pediatr Neurol*. 2019;97:18-25.
194. Haberle J, Burlina A, Chakrapani A, Dixon M, Karall D, Lindner M, et al. Suggested guidelines for the diagnosis and management of urea cycle disorders: First revision. *J Inherit Metab Dis*. 2019;42(6):1192-230.
195. Desbats MA, Lunardi G, Doimo M, Trevisson E, Salviati L. Genetic bases and clinical manifestations of coenzyme Q10 (CoQ 10) deficiency. *J Inherit Metab Dis*. 2015;38(1):145-56.
196. Sondheimer N, Hewson S, Cameron JM, Somers GR, Broadbent JD, Ziosi M, et al. Novel recessive mutations in COQ4 cause severe infantile cardiomyopathy and encephalopathy associated with CoQ10 deficiency. *Mol Genet Metab Rep*. 2017;12:23-7.
197. Sewry CA, Laitila JM, Wallgren-Pettersson C. Nemaline myopathies: a current view. *J Muscle Res Cell Motil*. 2019;40(2):111-26.
198. Farnaes L, Hildreth A, Sweeney NM, Clark MM, Chowdhury S, Nahas S, et al. Rapid whole-genome sequencing decreases infant morbidity and cost of hospitalization. *NPJ Genom Med*. 2018;3:10.
199. Clark MM, Stark Z, Farnaes L, Tan TY, White SM, Dimmock D, et al. Meta-analysis of the diagnostic and clinical utility of genome and exome sequencing and chromosomal microarray in children with suspected genetic diseases. *NPJ Genom Med*. 2018;3:16.
200. de Castro MJ, Gonzalez-Vioque E, Barbosa-Gouveia S, Salguero E, Rite S, Lopez-Suarez O, et al. Rapid Phenotype-Driven Gene Sequencing with the NeoSeq Panel: A Diagnostic Tool for Critically Ill Newborns with Suspected Genetic Disease. *J Clin Med*. 2020;9(8).
201. Worthey EA. Analysis and Annotation of Whole-Genome or Whole-Exome Sequencing Derived Variants for Clinical Diagnosis. *Curr Protoc Hum Genet*. 2017;95:9 24 1-9 8.
202. Bacchelli C, Williams HJ. Opportunities and technical challenges in next-generation sequencing for diagnosis of rare pediatric diseases. *Expert Rev Mol Diagn*. 2016;16(10):1073-82.

203. Mantere T, Kersten S, Hoischen A. Long-Read Sequencing Emerging in Medical Genetics. *Front Genet.* 2019;10:426.
204. Parikh S, Saneto R, Falk MJ, Anselm I, Cohen BH, Haas R, et al. A modern approach to the treatment of mitochondrial disease. *Curr Treat Options Neurol.* 2009;11(6):414-30.
205. Parikh S, Goldstein A, Karaa A, Koenig MK, Anselm I, Brunel-Guitton C, et al. Patient care standards for primary mitochondrial disease: a consensus statement from the Mitochondrial Medicine Society. *Genet Med.* 2017;19(12).
206. Gruosso F, Montano V, Simoncini C, Siciliano G, Mancuso M. Therapeutical Management and Drug Safety in Mitochondrial Diseases-Update 2020. *J Clin Med.* 2020;10(1).
207. Barcelos I, Shadiack E, Ganetzky RD, Falk MJ. Mitochondrial medicine therapies: rationale, evidence, and dosing guidelines. *Curr Opin Pediatr.* 2020;32(6):707-18.
208. Roman TS, Crowley SB, Roche MI, Foreman AKM, O'Daniel JM, Seifert BA, et al. Genomic Sequencing for Newborn Screening: Results of the NC NEXUS Project. *Am J Hum Genet.* 2020;107(4):596-611.
209. Fresard L, Montgomery SB. Diagnosing rare diseases after the exome. *Cold Spring Harb Mol Case Stud.* 2018;4(6).
210. Priest JR, Gawad C, Kahlig KM, Yu JK, O'Hara T, Boyle PM, et al. Early somatic mosaicism is a rare cause of long-QT syndrome. *Proc Natl Acad Sci U S A.* 2016;113(41):11555-60.
211. Cornish A, Guda C. A Comparison of Variant Calling Pipelines Using Genome in a Bottle as a Reference. *Biomed Res Int.* 2015;2015:456479.
212. McCarthy DJ, Humburg P, Kanapin A, Rivas MA, Gaulton K, Cazier JB, et al. Choice of transcripts and software has a large effect on variant annotation. *Genome Med.* 2014;6(3):26.
213. Hoffman-Andrews L. The known unknown: the challenges of genetic variants of uncertain significance in clinical practice. *J Law Biosci.* 2017;4(3):648-57.

214. Sambuughin N, Mungunsukh O, Ren M, Capacchione JF, Horkayne-Szakaly I, Chuang K, et al. Pathogenic and rare deleterious variants in multiple genes suggest oligogenic inheritance in recurrent exertional rhabdomyolysis. *Mol Genet Metab Rep.* 2018;16:76-81.
215. Merker JD, Wenger AM, Sneddon T, Grove M, Zappala Z, Fresard L, et al. Long-read genome sequencing identifies causal structural variation in a Mendelian disease. *Genet Med.* 2018;20(1):159-63.
216. Dewey FE, Murray MF, Overton JD, Habegger L, Leader JB, Fetterolf SN, et al. Distribution and clinical impact of functional variants in 50,726 whole-exome sequences from the DiscovEHR study. *Science.* 2016;354(6319).
217. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature.* 2016;536(7616):285-91.
218. Zhu X, Petrovski S, Xie P, Ruzzo EK, Lu YF, McSweeney KM, et al. Whole-exome sequencing in undiagnosed genetic diseases: interpreting 119 trios. *Genet Med.* 2015;17(10):774-81.
219. Kobayashi Y, Yang S, Nykamp K, Garcia J, Lincoln SE, Topper SE. Pathogenic variant burden in the ExAC database: an empirical approach to evaluating population data for clinical variant interpretation. *Genome Med.* 2017;9(1):13.
220. The Gnomad Consortium Releases First Studies of Human Genetic Variation. *Am J Med Genet A.* 2020;182(9):1999-2000.
221. Wang Q, Pierce-Hoffman E, Cummings BB, Alfoldi J, Francioli LC, Gauthier LD, et al. Landscape of multi-nucleotide variants in 125,748 human exomes and 15,708 genomes. *Nat Commun.* 2020;11(1):2539.
222. Eilbeck K, Quinlan A, Yandell M. Settling the score: variant prioritization and Mendelian disease. *Nat Rev Genet.* 2017;18(10):599-612.
223. Rao AR, Nelson SF. Calculating the statistical significance of rare variants causal for Mendelian and complex disorders. *BMC Med Genomics.* 2018;11(1):53.

224. Firth HV, Richards SM, Bevan AP, Clayton S, Corpas M, Rajan D, et al. DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. *Am J Hum Genet.* 2009;84(4):524-33.
225. Philippakis AA, Azzariti DR, Beltran S, Brookes AJ, Brownstein CA, Brudno M, et al. The Matchmaker Exchange: a platform for rare disease gene discovery. *Hum Mutat.* 2015;36(10):915-21.
226. Sobreira N, Schiettecatte F, Valle D, Hamosh A. GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. *Hum Mutat.* 2015;36(10):928-30.
227. Buske OJ, Girdea M, Dumitriu S, Gallinger B, Hartley T, Trang H, et al. PhenomeCentral: a portal for phenotypic and genotypic matchmaking of patients with rare genetic diseases. *Hum Mutat.* 2015;36(10):931-40.
228. Wortmann SB, Koolen DA, Smeitink JA, van den Heuvel L, Rodenburg RJ. Whole exome sequencing of suspected mitochondrial patients in clinical practice. *J Inherit Metab Dis.* 2015;38(3):437-43.
229. Powis Z, Farwell Hagman KD, Speare V, Cain T, Blanco K, Mowlavi LS, et al. Exome sequencing in neonates: diagnostic rates, characteristics, and time to diagnosis. *Genet Med.* 2018;20(11):1468-71.
230. Byron SA, Van Keuren-Jensen KR, Engelthaler DM, Carpten JD, Craig DW. Translating RNA sequencing into clinical diagnostics: opportunities and challenges. *Nat Rev Genet.* 2016;17(5):257-71.
231. Zhao J, Akinsanmi I, Arafat D, Cradick TJ, Lee CM, Banskota S, et al. A Burden of Rare Variants Associated with Extremes of Gene Expression in Human Peripheral Blood. *Am J Hum Genet.* 2016;98(2):299-309.
232. Pala M, Zappala Z, Marongiu M, Li X, Davis JR, Cusano R, et al. Population- and individual-specific regulatory variation in Sardinia. *Nat Genet.* 2017;49(5):700-7.

233. Li X, Kim Y, Tsang EK, Davis JR, Damani FN, Chiang C, et al. The impact of rare variation on gene expression across tissues. *Nature*. 2017;550(7675):239-43.
234. Fresard L, Smail C, Ferraro NM, Teran NA, Li X, Smith KS, et al. Identification of rare-disease genes using blood transcriptome sequencing and large control cohorts. *Nat Med*. 2019;25(6):911-9.

