TESIS DE LA UNIVERSIDAD DE ZARAGOZA
2025
199

Cristina Tania Lucia Campos

Diagnósticos complejos y AFF2 como gen causal del Espectro Cornelia de Lange

Director/es
Pié Juste, Juan
Latorre Pellicer, Ana

ISSN 2254-7606

Prensas de la Universidad
Universidad Zaragoza

Universidad de Zaragoza
Servicio de Publicaciones

ISSN 2254-7606

Universidad Zaragoza
1542

Tesis Doctoral

DIAGNÓSTICOS COMPLEJOS Y AFF2 COMO GEN CAUSAL DEL ESPECTRO CORNELIA DE LANGE

Autor

Cristina Tania Lucia Campos

Director/es

Pié Juste, Juan
Latorre Pellicer, Ana

UNIVERSIDAD DE ZARAGOZA
Escuela de Doctorado

Programa de Doctorado en Ciencias Biomédicas y Biotecnológicas

2025

Universidad Zaragoza
1542

Tesis Doctoral

Diagnósticos complejos y *AFF2* como gen causal del Espectro Cornelia de Lange

Autora

Cristina Tania Lucia Campos

Directores

Ana Latorre Pellicer

Juan Pié Juste

Universidad de Zaragoza

Facultad de Medicina

2024

La presente tesis doctoral se ha elaborado como un compendio de publicaciones en diversas revistas científicas, cuyas referencias completas son las siguientes:

1. ***Molecular Basis of the Schuurs-Hoeijmakers Syndrome: What We Know about the Gene and the PACS-1 Protein and Novel Therapeutic Approaches***

Arnedo M, Ascaso Á, Latorre-Pellicer A, **Lucia-Campos C**, Gil-Salvador M, Ayerza-Casas A, Pablo MJ, Gómez-Puertas P, Ramos FJ, Bueno-Lozano G, Pié J, Puisac B.

International Journal of Molecular Sciences. 2022 Aug 25. 23(17):9649
doi: 10.3390/ijms23179649. PMID: 36077045

2. ***Things are not always what they seem: From Cornelia de Lange to KBG phenotype in a girl with genetic variants in NIPBL and ANKRD11***

Latorre-Pellicer A, Ascaso Á, **Lucia-Campos C**, Gil-Salvador M, Arnedo M, Antoñanzas R, Ayerza-Casas A, Marcos-Alcalde Í, Gómez-Puertas P, Ramos FJ, Pié J, Puisac B

Molecular Genetics & Genomic Medicine. 2021 Nov. 9(11):e1826
doi: 10.1002/mgg3.1826. PMID: 34617417

3. ***A Novel Intragenic Duplication in the HDAC8 Gene Underlying a Case of Cornelia de Lange Syndrome***

Lucia-Campos C, Valenzuela I, Latorre-Pellicer A, Ros-Pardo D, Gil-Salvador M, Arnedo M, Puisac B, Castel N, Plaja A, Tenes A, Cuscó I, Trujillano L, Ramos FJ, Tizzano EF, Gómez-Puertas P, Pié J

Genes (Basel). 2022 Aug 8. 13(8):1413
doi: 10.3390/genes13081413. PMID: 36011323

4. ***An intragenic duplication in the AFF2 gene associated with Cornelia de Lange syndrome phenotypeSyndrome***

Lucia-Campos C, Parenti I, Latorre-Pellicer A, Gil-Salvador M, Bestetti I, Finelli P, Larizza L, Arnedo M, Ayerza-Casas A, Del Rincón J, Trujillano L, Morte B, Pérez-Jurado LA, Lapunzina P, Leitão E, Beygo J, Lich C, Kilpert F, Kaya S, Depienne C, Kaiser FJ, Ramos FJ, Puisac B, Pié J

Frontiers in Genetics. 2024 Nov 1. 15:1472543
doi: 10.3389/fgene.2024.1472543. PMID: 39553472

Don Juan Pié Juste, Profesor Catedrático del Departamento de Farmacología, Fisiología y Medicina Legal y Forense de la Facultad de Medicina de la Universidad de Zaragoza.

DECLARO:

Que la memoria de Tesis Doctoral presentada por la graduada doña **Cristina Tania Lucia Campos**, con el título "**Diagnósticos complejos y *AFF2* como gen causal del Espectro Cornelia de Lange**", corresponde a un trabajo realizado bajo mi dirección en la Unidad de Genética Clínica y Genómica Funcional del Departamento de Farmacología, Fisiología y Medicina Legal y Forense de la Facultad de Medicina de la Universidad de Zaragoza.

Terminada en esta fecha y revisado su contenido, estimo que se corresponde con el Proyecto de Tesis Doctoral aprobado por la Comisión de Doctorado, por lo que autorizo su presentación en la modalidad de compendio de publicaciones. Además, considero que reúne las condiciones requeridas para que su autora pueda optar al Grado de Doctor con Mención Internacional por la Universidad de Zaragoza.

Fdo.: Juan Pié Juste

Doña Ana Latorre Pellicer, Profesora Permanente Laboral del Departamento de Farmacología, Fisiología y Medicina Legal y Forense de la Facultad de Medicina de la Universidad de Zaragoza.

DECLARO:

Que la memoria de Tesis Doctoral presentada por la graduada doña **Cristina Tania Lucia Campos**, con el título "**Diagnósticos complejos y *AFF2* como gen causal del Espectro Cornelia de Lange**", corresponde a un trabajo realizado bajo mi dirección en la Unidad de Genética Clínica y Genómica Funcional del Departamento de Farmacología, Fisiología y Medicina Legal y Forense de la Facultad de Medicina de la Universidad de Zaragoza.

Terminada en esta fecha y revisado su contenido, estimo que se corresponde con el Proyecto de Tesis Doctoral aprobado por la Comisión de Doctorado, por lo que autorizo su presentación en la modalidad de compendio de publicaciones. Además, considero que reúne las condiciones requeridas para que su autora pueda optar al Grado de Doctor con Mención Internacional por la Universidad de Zaragoza.

Fdo.: Ana Latorre Pellicer

FINANCIACIÓN

Este trabajo ha sido posible gracias al apoyo económico prestado por diferentes instituciones en forma de:

- **Contrato Predoctoral de Formación en Investigación en Salud (PFIS)** otorgada a Dña. Cristina Tania Lucia Campos durante los años 2021-2024 (ref.: FI20/00290, Instituto de Salud Carlos III).

- **Proyecto FIS: "Redefinición del espectro Cornelia de Lange mediante la integración de nuevos datos clínicos, genómicos, transcriptómicos y metabolómicos"**, 2020-2023 (ref.: PI19/01860, Instituto de Salud Carlos III).

- **Proyecto FIS: "Espectro Cornelia de Lange, hacia una Medicina de Precisión: Integración de datos ómicos, seguimiento longitudinal y estudio ontogénico"**, 2024-2026 (ref.: PI23/01370, Instituto de Salud Carlos III).

- **Proyecto OTRI: Análisis genéticos de enfermedades raras"** 2021-2024 (ref.: 2021-2024/1106, Oficina de Transferencia de Resultados de la Investigación, Universidad de Zaragoza).

- **Grupo de Investigación Reconocido por el Gobierno de Aragón, "Genética Clínica y Genómica Funcional"**, 2020-2022 y 2023-2025 (ref.: B32_20R y B32_23R, Diputación General de Aragón).

PUBLICACIONES CIENTÍFICAS

A continuación, se indican las publicaciones científicas realizadas durante el período de formación predoctoral.

- ***Evaluating Face2Gene as a Tool to Identify Cornelia de Lange Syndrome by Facial Phenotypes***

  Latorre-Pellicer A, Ascaso A, Trujillano L, Gil-Salvador M, Arnedo M, **Lucia-Campos C**, Antoñanzas-Pérez R, Marcos-Alcalde I, Parenti I, Bueno-Lozano G, Musio A, Puisac B, Kaiser FJ, Ramos FJ, Gómez-Puertas P, Pié J

  International Journal of Molecular Science. 2020 Feb 4;21(3):1042. doi: 10.3390/ijms21031042.

- ***Clinical relevance of postzygotic mosaicism in Cornelia de Lange syndrome and purifying selection of NIPBL variants in blood***

  Latorre-Pellicer A, Gil-Salvador M, Parenti I, **Lucia-Campos C**, Trujillano L, Marcos-Alcalde I, Arnedo M, Ascaso Á, Ayerza-Casas A, Antoñanzas-Pérez R, Gervasini C, Piccione M, Mariani M, Weber A, Kanber D, Kuechler A, Munteanu M, Khuller K, Bueno-Lozano G, Puisac B, Gómez-Puertas P, Selicorni A, Kaiser FJ, Ramos FJ, Pié J.

  Scientific Reports. 2021 Jul 29.11(1):15459, doi: 10.1038/s41598-021-94958-z.

- ***Things are not always what they seem: From Cornelia de Lange to KBG phenotype in a girl with genetic variants in NIPBL and ANKRD11***

  Latorre-Pellicer A, Ascaso Á, **Lucia-Campos C**, Gil-Salvador M, Arnedo M, Antoñanzas R, Ayerza-Casas A, Marcos-Alcalde I, Gómez-Puertas P, Ramos FJ, Pié J, Puisac B

  Molecular Genetics & Genomic Medicine. 2021 Nov. 9(11):e1826. doi: 10.1002/mgg3.1826.

- ***Molecular Basis of the Schuurs–Hoeijmakers Syndrome: What We Know about the Gene and the PACS-1 Protein and Novel Therapeutic Approaches***

  Arnedo M, Ascaso Á, Latorre-Pellicer A, **Lucia-Campos C**, Gil-Salvador M, Ayerza-Casas A, Pablo MJ, Gómez-Puertas P, Ramos FJ, Bueno-Lozano G, Pié J, Puisac B

  International Journal of Molecular Science. 2022 Aug 25. 23(17):9649. doi: 10.3390/ijms23179649.

- ***A Novel Intragenic Duplication in the HDAC8 Gene Underlying a Case of Cornelia de Lange Syndrome***

  Lucia-Campos C, Valenzuela I, Latorre-Pellicer A, Ros-Pardo D, Gil-Salvador M, Arnedo M, Puisac B, Castells N, Plaja A, Tenes A, Cuscó I, Trujillano L, Ramos FJ, Tizzano EF, Gómez-Puertas P, Pié J

  Genes (Basel). 2022 Aug 8. 13(8):1413. doi: 10.3390/genes13081413.

- ***Case report: A novel case of parental mosaicism in SMC1A gene causes inherited Cornelia de Lange syndrome***

  Gil-Salvador M, Latorre-Pellicer A, **Lucia-Campos C**, Arnedo M, Darnaude MT, Díaz de Bustamante A, Villares R, Palma Milla C, Puisac B, Musio A, Ramos FJ, Pié J

  Frontiers in Genetics. 2022 Sep 28. 13:993064. doi: 10.3389/fgene.2022.993064.

- ***Subclinical myocardial dysfunction is revealed by speckle tracking echocardiography in patients with Cornelia de Lange syndrome***

  Trujillano L, Ayerza-Casas A, Puisac B, García GG, Ascaso Á, Latorre-Pellicer A, Arnedo M, **Lucia-Campos C**, Gil-Salvador M, Kaiser FJ, Ramos FJ, Pié J, Bueno-Lozano G

  The International Journal of Cardiovascular Imaging. 2022 Nov. 38(11):2291-2302. doi: 10.1007/s10554-022-02612-0. Epub 2022 May 19.

- ***Assessment of Quality of Life Using the Kidslife Scale in Individuals With Cornelia de Lange Syndrome***

  Trujillano L, Ayerza-Casas A, Puisac B, Latorre-Pellicer A, Arnedo M, **Lucia-Campos C**, Gil-Salvador M, Parenti I, Kaiser FJ, Ramos FJ, Trujillano J, Pié J

  Cureus. 2024 Apr 1. 16(4):e57378. doi: 10.7759/cureus.57378.

- ***Heart Disease Characterization and Myocardial Strain Analysis in Patients with PACS1 Neurodevelopmental Disorder***

  Latorre-Pellicer A, Trujillano L, Del Rincón J, Peña-Marco M, Gil-Salvador M, **Lucia-Campos C**, Arnedo M, Puisac B, Ramos FJ, Ayerza-Casas A, Pié

  Journal of Clinical MedicineJ Clin Med. 2023 Jun 14. 12(12):4052. doi: 10.3390/jcm12124052.

- ***An intragenic duplication in the AFF2 gene associated with Cornelia de Lange syndrome phenotype***

  Lucia-Campos C, Parenti I, Latorre-Pellicer A, Gil-Salvador M, Bestetti I, Finelli P, Larizza L, Arnedo M, Ayerza-Casas A, Del Rincón J, Trujillano L, Morte B, Pérez-Jurado LA, Lapunzina P, Leitão E, Beygo J, Lich C, Kilpert F, Kaya S, Depienne C, Kaiser FJ, Ramos FJ, Puisac B, Pié J

  Frontiers in Genetics. 2024 Nov 1. 15:1472543. doi: 10.3389/fgene.2024.1472543.

Agradecimientos

Me gustaría comenzar dando las gracias a mis directores. Gracias Juan por creer en mí desde el principio, darme la oportunidad de unirme a tu grupo y acogerme con tanto cariño. Como dices, "es como si fueras mi hija científica", y no tengo más que agradecer todos los consejos que me has dado a lo largo de estos años, siempre pensando en mi futuro. Gracias Ana por tu constante ayuda y guía durante todo este tiempo. Tu gran capacidad resolutiva ante cualquier situación, y tu gran pasión por la ciencia, ha hecho que aprendiese muchísimo de ti. Además, tu trato tan delicado y cercano, ha hecho que el trabajo haya sido muy cómodo contigo. ¡Muchas gracias a los dos!

Gracias a mis compañeras del laboratorio, con las que he compartido tantas horas. Gracias Beatriz, por estar siempre (aún sin estar presente), tu generosidad es inmensa y eres un gran ejemplo de persona, en todos los sentidos. Gracias María, por estar constantemente dispuesta a ayudar. Gracias Rebeca, por ser la memoria del laboratorio y saber dónde está siempre todo. Y finalmente, decir que he sido una gran afortunada de encontrar a una persona que me ha escuchado y comprendido a la perfección, y con la que he podido conversar de cualquier cosa durante todos estos años. Ella es Marta, mi compañera imprescindible durante la tesis. Todos los buenos momentos que me llevo de esta etapa, han sido en su compañía, y quería darle las gracias por ganar con ella una bonita amistad. ¡Gracias a todas por estar durante estos años y animarme tanto!

Quiero expresar mi más sincero agradecimiento al equipo clínico. Mi gratitud a Laura, Julia, Ariadna, Pilar, Chus y Ángela por su dedicación y excelente labor con los pacientes. Asimismo, agradezco especialmente a Feliciano, cuyo destacado trabajo como pediatra, y su gran experiencia en investigación, han sido una fuente de aprendizaje para mí.

Gracias también a todos los colaboradores y colaboradoras que han contribuido a enriquecer la calidad de nuestros trabajos. Un agradecimiento especial a Paulino, gracias por tu excelente trabajo, tu constante disposición para colaborar y tu gran eficiencia.

Gracias a todo el personal del departamento de Fisiología por su atención y cariño. También quiero agradecer a los servicios del Instituto de Investigación Sanitaria de Aragón, por su paciencia conmigo y la ayuda que me han ofrecido durante estos años.

Un GRACIAS inmenso y lleno de cariño a la Asociación Cornelia de España. El contacto cercano con las familias ha sido muy enriquecedor para mí. Su generosidad y compromiso con la investigación, no solo nos permiten avanzar cada día, sino que también nos motiva a seguir trabajando con dedicación.

No podía olvidarme de Milán, el lugar donde realicé mi estancia predoctoral, y donde pasé uno de los mejores meses de mi tesis. Gracias en primer lugar a Valentina Massa, por ofrecerme la oportunidad de conocerla y trabajar con ella. Es una gran científica que ha formado un excelente grupo de investigación, del cual me llevo un gran aprendizaje que sumo a mi experiencia. Además, estoy profundamente agradecida por el trato que recibí durante mi estancia. Me acogieron, me ayudaron y me enseñaron con una delicadeza, que jamás olvidaré.

Gracias a mis amigos y amigas, me habéis regalado momentos increíbles durante estos años, y me habéis permitido desconectar (y a veces reconectar) de mi tesis en los momentos que más lo necesitaba.

Sin duda, una parte fundamental ha sido mi familia, mis padres y mis hermanas, que están siempre al pie del cañón. Gracias mamá y papá, los valores y educación que me habéis dado, me han permitido llegar hasta aquí; he sido una gran afortunada de contar siempre con los dos. Gracias Leticia, por estar a mi lado y valorar tanto mi trabajo. Gracias Nerea, me has dado un apoyo incondicional animándome día tras día, y tu reiterada frase "Cris vas a llegar muy lejos", no te haces idea de lo que me ha ayudado. Sin el apoyo de los cuatro, no lo hubiera conseguido. Por ello, os dedico esta tesis.

El tiempo de calidad, es aquel que pasas con las personas que quieres y te quieren, con las que eres tú misma. Finalmente, quiero dar las GRACIAS, con todo mi cariño, a todas las personas que están siempre conmigo, apoyándome y arropándome, y que día a día, me ayudan a convertirme en una mejor versión de mí.

ÍNDICE

Resumen / Summary

El síndrome Cornelia de Lange (SCdL, OMIM #122470, #300590, #300882, #610759 y #614701) es un trastorno raro del neurodesarrollo que afecta a múltiples aparatos y sistemas, caracterizado por una facies distintiva, baja talla, anomalías en las extremidades y discapacidad intelectual. Desde la identificación del primer gen causal, *NIPBL*, hace veinte años, se han descrito multitud de variantes genéticas y se han asociado otros siete genes, la mayoría relacionados al complejo cohesina. Sin embargo, a pesar de los avances conseguidos, actualmente aproximadamente el 20% de los pacientes aún continúa sin un diagnóstico genético claro. Por ello, el objetivo principal de esta tesis es lograr un diagnóstico preciso de los casos clínicos susceptibles de ser incluidos en el Espectro Cornelia de Lange (ECdL).

El diagnóstico de enfermedades raras enfrenta numerosos desafíos. Los retrasos suelen deberse al desconocimiento de los genes causales implicados, o a limitaciones técnicas en la detección e interpretación de variantes genéticas complejas. Además, los diagnósticos imprecisos pueden estar relacionados con la evolución fenotípica de los pacientes, o con el solapamiento de características clínicas entre diferentes entidades nosológicas.

La identificación de un nuevo gen causal del SCdL se presenta como la contribución más significativa de esta tesis. Tras una odisea diagnóstica de más de dieciocho años, se describe por primera vez una variante patogénica en el gen *AFF2* en una familia afectada por una forma clásica del SCdL. Esta variante corresponde a una duplicación intragénica que provoca una disminución significativa del tránscrito del gen. Además, la similitud funcional en la regulación de la expresión génica entre este gen, y los ya conocidos, respalda la propuesta de *AFF2* como nuevo gen causal del SCdL.

En los últimos años, ha aumentado considerablemente la identificación de individuos con SCdL y variantes en el número de copias (CNVs). En este trabajo, se describe la segunda duplicación intragénica del gen *HDAC8*, detectada en una paciente de ocho años con fenotipo clásico del SCdL. Este hallazgo no solo aporta una mejor caracterización clínica y molecular de pacientes con variantes en *HDAC8*, sino que también destaca la importancia de implementar técnicas moleculares avanzadas y algoritmos de análisis que permitan identificar este tipo de alteraciones genéticas.

El seguimiento longitudinal y la reevaluación de datos genéticos, son fundamentales en pacientes con trastornos del neurodesarrollo. Un ejemplo claro de esta necesidad, es el presentado en el tercer trabajo, una niña cuya evolución clínica transitó de un diagnóstico inicial de SCdL a síndrome de KBG. Este cambio fenotípico alertó a los clínicos, lo que llevó a ampliar el estudio genético, que permitió la identificación de una variante patogénica en el gen *ANKRD11*.

El descubrimiento de nuevas entidades nosológicas, está mejorando la precisión diagnóstica. Un ejemplo es el caso de algunos pacientes inicialmente diagnosticados en la infancia con SCdL que, gracias a la identificación del síndrome de PACS1, han sido reclasificados. Aunque no existe una relación molecular entre ambos síndromes, su similitud fenotípica motivó la realización de una revisión detallada del síndrome de PACS1, proporcionando un contexto valioso para comprender las similitudes y diferencias con trastornos estrechamente relacionados. Además, dada la gran similitud clínica y molecular observada entre los síndromes PACS1, PACS2 y WDR37, se sugiere su agrupación en una nueva entidad nosológica más amplia, siguiendo el ejemplo del ECdL.

En conjunto, este trabajo profundiza en los mecanismos subyacentes al SCdL, mejorando la comprensión del cuadro clínico. Se propone un nuevo gen causal, se describen nuevas variantes genéticas complejas, y se amplía el conocimiento sobre el diagnóstico diferencial del síndrome. La caracterización de variantes patogénicas y su asociación con fenotipos clínicos complejos, resulta clave para optimizar la precisión diagnóstica y, en última instancia, ofrecer una atención de calidad a los pacientes y familias afectadas.

Cornelia de Lange syndrome (CdLS, OMIM #122470, #300590, #300882, #610759, and #614701) is a rare neurodevelopmental disorder affecting multiple systems and organs, characterised by distinctive facial features, short stature, limb anomalies and intellectual disability. Since the discovery of the first causal gene, *NIPBL*, 20 years ago, numerous genetic variants have been described and seven additional genes have been associated with this disorder, most of them related to the cohesin complex. Despite the advances achieved, to date approximately 20% of patients still lack a genetic diagnosis. Therefore, the main objective of this thesis is to establish an accurate diagnosis of clinical cases potentially included within Cornelia de Lange Spectrum.

The diagnosis of rare diseases faces numerous challenges. Delays are often attributable to insufficient knowledge of the causal genes involved or to technical limitations in the detection and interpretation of complex genetic variants. Furthermore, inaccurate diagnoses may arise from the phenotypic shift of patients or the overlap of clinical features among different nosological entities.

The identification of a new causal gene for CdLS emerges as the most significant contribution of this thesis. After a diagnostic odyssey spanning over 18 years, a pathogenic variant in the *AFF2* gene has been described for the first time in a family presenting with classic CdLS. This variant consists of an intragenic duplication resulting in a significant dose reduction of the gene's transcript. Furthermore, the functional similarity in gene expression regulation between *AFF2* and those genes already associated with the syndrome strongly supports the proposal of *AFF2* as a new causal gene for CdLS.

In the last few years, there has been a considerable increase in the identification of copy number variations (CNVs) in individuals with CdLS. This study describes the second intragenic duplication identified in *HDAC8* gene, detected in an eight-year-old patient with classic CdLS phenotype. This finding not only provides a better clinical and molecular characterization of patients with variants in *HDAC8*, but also underscores the importance of implementing advanced molecular techniques and analytical algorithms to identify such genetic alterations.

Longitudinal follow-up and reevaluation of genetic data are crucial in patients with neurodevelopmental disorders. A clear example of this is presented in the third article of this work, describing a girl whose clinical progression shifted from an initial CdLS diagnosis to KBG syndrome. This phenotypic change alerted the clinicians, prompting an expansion of the genetic study, which led to the identification of a pathogenic variant in the *ANKRD11* gene.

The discovery of new nosological entities is improving diagnostic accuracy. For example, the classification of several patients initially diagnosed with CdLS in childhood have been modified thanks to the identification of PACS1 syndrome. Although no molecular relationship exists between the two conditions, their phenotypic similarity prompted a thorough review of PACS1 syndrome, providing insights into the similarities and differences with closely related disorders. In addition, the notable clinical and molecular similarity observed between PACS1, PACS2, and WDR37 syndromes suggests that they may be grouped into a new nosological entity, following the example of Cornelia de Lange Spectrum.

Overall, this thesis explores the etiopathogenesis of CdLS, enhancing the understanding of its clinical presentation. In this work, a new causal gene is proposed, novel complex genetic variants are described and the knowledge of the differential diagnosis of the syndrome is expanded. The characterization of pathogenic variants and their relationship with complex clinical phenotypes is essential for improving diagnostic accuracy and, ultimately, for providing quality care to patients and their families.

Abreviaturas

aCGH: array de hibridación genómica comparativa, array *Comparative Genomic Hybridization.*

ACMG: Colegio Americano de Genética Médica, *American College of Medical Genetics*.

ADN: ácido desoxirribonucleico.

ARN: ácido ribonucleico.

ATP: adenosín trifosfato.

CNV: variación del número de copias, *Copy Number Variation.*

ECdL: Espectro Cornelia de Lange.

EGR: enfermedad genética rara.

ER: enfermedad rara.

FEDER: Federación Española de Enfermedades Raras.

FISH: hibridación fluorescente *in situ*, *Fluorescence In Situ Hybridization.*

HPO: Ontología del Fenotipo Humano, *Human Phenotype Ontology.*

IA: inteligencia artificial.

INDELS: pequeñas inserciones o deleciones, *Insertions and Deletions*.

iPSCs: células madre pluripotentes inducidas, *Induced Pluripotent Stem Cells*.

LRS: secuenciación de lectura larga, *Long Read Sequencing.*

MLPA: amplificación de sondas tras ligación múltiple, *Multiplex Ligation dependent Probe Amplification.*

mTOR: *mammalian Target of Rapamycin.*

NGS: secuenciación de nueva generación, *Next Generation Sequencing*.

OGM: mapeo óptico del genoma, *Optical Genome Mapping*.

PACS1: proteína 1 clasificadora de grupos ácidos de fosfofurina, *Phosphofurin Acidic Cluster Sorting Protein 1*.

SCdL: síndrome Cornelia de Lange.

SHMS: síndrome de Schuurs-Hoeijmakers, *Schuurs-Hoeijmakers syndrome*.

SKBG: síndrome de KBG.

SMC: mantenimiento estructural de los cromosomas, *Structural Maintenance of Chromosomes.*

SNV: variante de un solo nucleótido, S*ingle Nucleotide Variant.*

STR: repeticiones cortas en tándem, *Short Tandem Repeats*.

TND: trastorno del neurodesarrollo.

VUS: variantes de significado incierto, *Variants of Uncertain Significance*.

WES: secuenciación del exoma completo, *Whole Exome Sequencing*.

WGS: secuenciación del genoma completo, *Whole Genome Sequencing.*

Introducción

Desafíos en el diagnóstico de las enfermedades raras

Las enfermedades raras (ERs) constituyen un grupo heterogéneo de patologías que abarca entre 6.000 y 8.000 tipos diferentes, afectando aproximadamente al 5% de la población mundial [1]. Se estima que alrededor del 80% de los casos tiene un origen genético, generalmente con un inicio congénito, y en muchos casos, con un impacto potencialmente letal [1,2]. Dentro de las enfermedades genéticas raras (EGRs), el 70% de los pacientes manifiestan sus síntomas en la infancia, con una tasa de supervivencia inferior a cinco años en el 30% de los casos [1]. Uno de los principales problemas en este campo, que contribuye a la incertidumbre sobre las cifras actuales, es la falta de diagnóstico. Actualmente, se calcula que entre el 25% y el 50% de los pacientes carecen de un diagnóstico genético confirmado [3,4]. Además, según la Federación Española de Enfermedades Raras (FEDER), el tiempo promedio de espera para un diagnóstico es de cuatro años y en el 20% de los casos puede superar los diez años.

A pesar de los avances tecnológicos, el diagnóstico de las EGRs sigue siendo un desafío importante, debido a la amplia variabilidad clínica y genética que presentan [5]. La colaboración internacional ha mejorado notablemente la recopilación y el análisis de datos a nivel mundial, lo que favorece el avance del conocimiento sobre estas patologías. El fenotipado detallado, junto con la aplicación de nuevas técnicas diagnósticas, facilita la identificación y comprensión de nuevas variantes patogénicas [6–8]. Además, el desarrollo de nuevos modelos experimentales, contribuye a esclarecer los mecanismos moleculares subyacentes a estas enfermedades [9]. En ciertos casos, se están explorando enfoques terapéuticos con el fin de revertir el origen de los trastornos [10].

Las dificultades para alcanzar un diagnóstico son atribuibles a diversas causas, que se pueden agrupar en problemas estructurales del sistema de salud y limitaciones técnicas. En España, una de las principales carencias ha sido la ausencia de una especialidad médica en genética clínica y de laboratorio, la cual fue aprobada recientemente por la Comisión de Recursos Humanos del Sistema Nacional de Salud (5 de diciembre de 2024). Esta falta de especialización ha afectado en la capacidad del sistema para diagnosticar y gestionar adecuadamente estas enfermedades, debido a la escasez de profesionales con formación específica en este ámbito. Desde el punto de vista técnico, los desafíos más relevantes incluyen la complejidad genética de estas condiciones, la dificultad para interpretar variantes en regiones no codificantes y la limitada disponibilidad de bases de datos que relacionen genotipo y fenotipo. En consecuencia, la combinación de estos factores estructurales y técnicos, resalta la necesidad de implementar enfoques diagnósticos integrales y de desarrollar plataformas colaborativas que faciliten la interpretación de datos genómicos en el contexto de las EGRs [11–13].

Los consorcios y asociaciones, tanto nacionales como internacionales, juegan un papel crucial en la ampliación del conocimiento sobre las EGRs [14]. Su labor se enfoca en fomentar la colaboración global, facilitando el intercambio de datos y la coordinación de esfuerzos para afrontar los desafíos del diagnóstico. En España, destacan iniciativas en red como el CIBERER [15] o el reciente proyecto IMPaCT-GENóMICA, que constituye una infraestructura de apoyo al diagnóstico mediante tecnologías genómicas avanzadas. Además, a nivel internacional, están surgiendo herramientas innovadoras, como la plataforma *GPAP RD-Connect*, que proporciona información actualizada sobre el fenoma, genoma, y su correlación [16]. Por tanto, la combinación de esfuerzos globales, la gestión eficiente de datos y la creación de consorcios, están siendo esenciales para el avance en la investigación de las EGRs. Estas redes de trabajo también promueven el desarrollo de políticas de salud pública más inclusivas, tan necesarias para las personas y familias afectadas.

Para interpretar adecuadamente el síndrome en cuestión, es fundamental realizar un análisis exhaustivo del cuadro clínico. El fenotipado de precisión busca emplear una terminología única, codificada y estandarizada, a nivel internacional, para definir y describir detalladamente las características de la patología. Este enfoque permite identificar patrones fenotípicos, establecer correlaciones entre el genotipo y fenotipo, y delimitar posibles diagnósticos diferenciales. En concreto, se pretende implementar el fenotipado profundo (*deep phenotyping*), que consiste en un análisis preciso y completo de los componentes individuales del fenotipo [17,18]. Este nuevo planteamiento requiere de un estudio multidisciplinario que integre herramientas tecnológicas, analíticas y computacionales, facilitando así una comprensión más profunda de la complejidad fenotípica. El proyecto de Ontología del Fenotipo Humano (*Human Phenotype Ontology*, HPO) ha tenido un impacto significativo en el *deep phenotyping*, al permitir, mediante algoritmos de similitud semántica y aprendizaje automático, respaldar los análisis genómicos y fenotípicos [19]. Además, dado que muchos síndromes presentan rasgos faciales distintivos [20,21], cada vez se suman más herramientas capaces de discernir estas características [22].

Las variantes genéticas asociadas a las EGRs muestran una gran variabilidad. Estas abarcan desde cambios puntuales (*Single Nucleotide Variants*, SNVs), pequeñas inserciones o deleciones (*Insertions and deletions*, INDELS), hasta repeticiones cortas en tándem (*Short Tandem Repeats*, STR) y alteraciones estructurales más complejas, como variantes en el número de copias (*Copy Number Variants*, CNVs) y reordenamientos cromosómicos [23,24]. La evolución de las técnicas diagnósticas ha sido clave para mejorar la identificación de estas variantes genómicas. Durante mucho tiempo, la genética clínica se ha basado en pruebas que se enfocan en un solo gen o en análisis citogenéticos del genoma completo [25]. Sin embargo, estas metodologías presentaban una baja sensibilidad

y resolución, permitiendo el diagnóstico solo en aproximadamente el 10% de los pacientes [26]. La secuenciación Sanger ha sido pionera en la detección de variantes puntuales e INDELS, aunque su enfoque se limita a regiones específicas. La llegada de la secuenciación de nueva generación (*Next Generation Sequencing*, NGS) ha transformado la investigación genómica. El análisis masivo que ofrecen estas tecnologías, permite identificar un mayor número de variantes patogénicas [27,28]. En concreto, la secuenciación de exoma (*Whole Exome Sequencing*, WES) y genoma completo (*Whole Genome Sequencing*, WGS), están facilitando la detección de muchas variantes, incluidas alteraciones estructuralmente complejas o localizadas en regiones no codificantes [13,29,30]. Sin embargo, las técnicas NGS presentan, en ocasiones, limitaciones para resolver regiones genómicas altamente repetitivas o complejas. Esta deficiencia se está abordando con la incorporación de las nuevas tecnologías de secuenciación de lectura larga (*long read sequencing*, LRS), como *Nanopore* y *PacBio* [31–34]. Estas técnicas ofrecen una mayor precisión en la resolución de lectura de regiones genómicas difíciles, ampliando significativamente el alcance de las aplicaciones genómicas actuales. Simultáneamente, las técnicas citogenéticas han evolucionado desde el cariotipo convencional, hacia tecnologías más avanzadas como la hibridación fluorescente *in situ* (*Fluorescence In Situ Hybridization*, FISH), los arrays de hibridación genómica comparativa (array *Comparative Genomic Hybridization*, aCGH) de alta resolución y el mapeo óptico del genoma (*Optical Genome Mapping*, OGM) [35–37]. Estos avances en las técnicas diagnósticas están mejorando la identificación de variantes patogénicas, especialmente en la detección de CNVs y variantes en regiones no codificantes, que anteriormente podían pasar desapercibidas [13,29,38].

La incorporación de las nuevas metodologías, ha incrementado las tasas de diagnóstico genético en casi un 50% en los últimos años [39,40]. La gran capacidad actual de secuenciación genómica, está permitiendo identificar numerosas variantes en regiones no codificante asociadas a trastornos del neurodesarrollo (TND) [41–43]. Sin embargo, el manejo del gran volumen de datos generados por estas tecnologías representa un reto significativo, ya que una proporción considerable de las variantes detectadas se clasifica como de significado incierto (*Variants of Uncertain Significance*, VUS). Esta situación puede deberse a la falta de evidencias que permitan establecer su relación con la enfermedad en el momento de su identificación [44,45]. Para abordar estas dificultades, se aplican criterios de clasificación establecidos por el Colegio Americano de Genética Médica (*American College of Medical Genetic*s, ACMG), que proporciona un marco estandarizado para evaluar la relevancia clínica de las variantes genéticas [46]. Además, los repositorios de datos a gran escala, tanto poblacionales, p.ej., *GnomAD*, como asociados a enfermedad, p.ej., *ClinVar, LOVD* o *DECIPHER*, ofrecen información genética valiosa que ayuda en la caracterización de las variantes [47–49]. El desarrollo de herramientas

bioinformáticas para la predicción de patogenicidad, p.ej., *VarSome*, también contribuye significativamente en la interpretación de los datos genómicos [50,51]. En este contexto, la reciente aplicación *AlphaFold*, basada en inteligencia artificial, está teniendo un impacto relevante al proporcionar análisis avanzados para predecir la repercusión a nivel proteico [52].

El desarrollo de modelos preclínicos ha representado un gran avance en el estudio de la etiología y patogénesis de las EGRs [53]. Los modelos animales, como el pez cebra, se utilizan ampliamente para investigar el impacto de variantes genéticas específicas [54]. Estos modelos no solo permiten estudiar la afectación a nivel molecular, sino también el desarrollo de enfoques terapéuticos. Además, la reciente introducción de nuevos modelos de investigación, como los organoides y los modelos celulares derivados de células madre pluripotentes inducidas (*Induced Pluripotent Stem Cells*, iPSCs), están ofreciendo mayor versatilidad para explorar estas enfermedades [9].

Abordar estos desafíos, permitirá que todas las personas afectadas por una EGR obtengan un diagnóstico temprano, reciban una atención médica adecuada y tengan la posibilidad de acceder a terapias disponibles.

Bases moleculares de los trastornos del neurodesarrollo

Uno de los fenotipos más prevalente de las ERs, son las anomalías en el sistema nervioso, presentes en aproximadamente el 75% de los casos [55,56]. En este contexto, el término de trastorno del neurodesarrollo (TND), se refiere a un grupo diverso de enfermedades que afectan al desarrollo del sistema nervioso, y que pueden alterar funciones cognitivas, motoras, conductuales y emocionales [57]. Los TND abarcan una amplia gama de problemas neurológicos y psiquiátricos, que incluyen desde anomalías neuronales congénitas, hasta trastornos como esquizofrenia, autismo, epilepsia y trastorno por déficit de atención e hiperactividad [57,58]. Aunque estos trastornos comparten ciertas características clínicas, su origen es altamente diverso. Por ello, comprender los mecanismos patogénicos específicos de cada TND constituye un desafío significativo, debido a la elevada complejidad de estas enfermedades [59,60].

La clasificación molecular de los TND de origen monogénico, se fundamenta en la comprensión de las vías biológicas y genes implicados. Estudios recientes han evidenciado la conexión entre las bases genéticas, los patrones de interacción proteica y las manifestaciones clínicas de estas enfermedades [61]. Así, a pesar de la amplia heterogeneidad genética de los TND, las consecuencias funcionales de las diversas variantes patogénicas tienden a afectar a vías moleculares interrelacionadas. En estas vías

se incluyen la función sináptica, la regulación de la expresión génica, la mielinización, así como la proliferación y migración neuronal, y la señalización de aminoácidos, entre otras [62,63] (Figura 1). Por tanto, los genes asociados a los TND no se distribuyen de manera aleatoria en las redes moleculares, sino que se agrupan en módulos específicos [64].

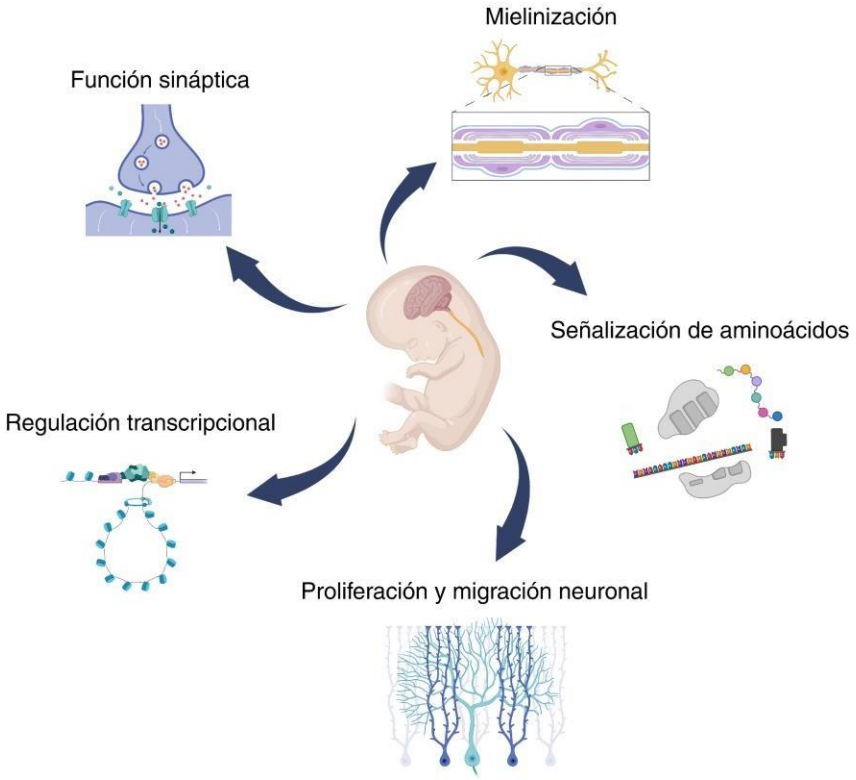


Figura 1. Principales vías afectadas en los trastornos del neurodesarrollo. (Imagen creada mediante BioRender)

Los trastornos que afectan a la función sináptica, engloban diversas patologías caracterizadas por alteraciones en la transmisión sináptica y la plasticidad neuronal [65]. Se identifican dos grupos principales de proteínas sinápticas, que son fundamentales para la formación de circuitos neuronales dependientes de la actividad durante el desarrollo. Por un lado, las moléculas de adhesión celular median la organización bidireccional de los compartimentos pre y postsinápticos a través la señalización transcelular [66]. Por otro lado, las proteínas asociadas al andamiaje y la señalización sináptica, forman grandes redes moleculares de receptores y proteínas vinculadas a la actina. Las variantes patogénicas en los genes que codifican estas proteínas, alteran significativamente el proceso de desarrollo cerebral, dando lugar a los cuadros clínicos observados en los pacientes afectados [67]. Además, se ha demostrado que alteraciones en genes que codifican proteínas reguladoras de estos procesos, como las proteínas SHANK (*SHANK2* y *SHANK3*), neuroliginas (*NLGN*) y neurexinas (*NRXN*), son comunes en muchos de los TND [68–71]. Un ejemplo de

trastorno relacionado con la función sináptica, es el síndrome de *Schuurs-Hoeijmakers* o PACS1 (SHMS, OMIM #615009). Investigaciones recientes, han evidenciado alteraciones en la función sináptica de las dendritas, correlacionando esta afectación con la discapacidad intelectual asociada a esta patología. Curiosamente, se está explorando un posible tratamiento mediante oligonucleótidos antisentido, que ha mostrado restaurar la estructura neuronal y la transmisión sináptica en modelos de ratón [72,73].

Otro tipo de TND, que afecta principalmente a las funciones motoras y cognitivas, es la alteración en la homeostasis de la mielina, afectando a su formación, mantenimiento y reparación [74]. Estos trastornos pueden originarse por defectos en las células de *Schwann* o en los oligodendrocitos [75]. Los mecanismos moleculares que se encuentran implicados, son los responsables de regular diversos procesos como la señalización celular, la diferenciación de los oligodendrocitos, la regulación del metabolismo lipídico y la respuesta a factores tróficos [76]. Entre los genes asociados a este fenotipo, se encuentran *GDAP1*, *MTMR2* y *PRX*, cuyas proteínas desempeñan un papel crucial en la señalización celular y el metabolismo lipídico [77–79]. Un ejemplo representativo de este subgrupo de TND, es el síndrome Charcot-Marie-Tooth (OMIM #607831, #118220, #118200, #302800, #609260), una de las neuropatías hereditarias más comunes.

La alteración en la señalización de aminoácidos durante la síntesis de proteínas, es una causa asociada a los TND, destacando la quinasa mTOR (*mammalian Target of Rapamycin*) como un elemento clave en este subgrupo de patologías. Durante el desarrollo embrionario, mTOR regula la proliferación y diferenciación de las neuronas progenitoras, así como el crecimiento y elongación de las neuritas [80]. En la edad adulta, esta quinasa también participa en procesos como la neurogénesis, el aprendizaje, la memoria y la plasticidad sináptica [81]. Además, mTOR está implicada en tres vías de señalización: la vía del factor de crecimiento, el sensor de energía, y el sensor de aminoácidos, lo que significa que su alteración puede tener múltiples consecuencias indirectas [82]. También se han descrito diversas variantes patogénicas en factores reguladores de estas vías, asociadas a TND, incluyendo variantes en los genes *TSC1*, *TSC2* o *PTEN* [83].

El crecimiento de la corteza cerebral humana es un proceso complejo y altamente regulado. La alteración de cualquiera de sus etapas puede dar lugar a la aparición de manifestaciones clínicas graves [84]. La adecuada proliferación y migración neuronal durante el desarrollo cerebral, son fundamentales para la organización cortical y la formación de redes neuronales funcionales. Se han identificado variantes patogénicas en genes que regulan el citoesqueleto, la adhesión celular y las señales químicas que guían la migración neuronal [85]. Estos trastornos, conocidos como malformaciones del desarrollo cortical, se clasifican en tres grupos: malformaciones en la proliferación celular y

apoptosis, malformaciones en la migración neuronal, y malformaciones en el desarrollo postmigratorio. Estas patologías se caracterizan por una estructura cortical irregular y/o por la presencia de materia gris heterotópica, a menudo asociada a un tamaño cerebral inusual [86–88].

Otro proceso clave durante el desarrollo cerebral, es la regulación transcripcional. Este delicado proceso debe ser preciso y coordinado durante esta etapa, lo que requiere de una adecuada regulación de la expresión génica para modular los niveles de transcripción de genes específicos del desarrollo cerebral. Esta regulación garantiza la correcta maduración de las conexiones excitatorias e inhibitorias en la corteza cerebral, así como la diferenciación de las neuronas, elementos esenciales para establecer un equilibrio funcional en los circuitos neuronales [89]. Existen numerosos genes asociados a este tipo de trastornos, responsables de codificar proteínas involucradas en la regulación de la transcripción o el remodelado de la cromatina [90]. Entre ellos se encuentran *MECP2*, *KMT2A*, *CHD8*, *SETD5*, *ANKRD11* o *NIPBL*, cuya desregulación puede causar efectos neurológicos muy graves [91,92]. Diversos estudios han investigado el efecto de la haploinsuficiencia de estos genes, destacando la importancia de su dosis adecuada durante el neurodesarrollo [93–96]. Un trastorno que se clasifica dentro de esta categoría, es el síndrome Cornelia de Lange (SCdL; OMIM #122470, #300590, #610759, #614701 y #300882), una enfermedad multisistémica que se caracteriza principalmente por su dismorfismo craneofacial distintivo, afectación en extremidades, retraso de crecimiento pre y postnatal, y discapacidad intelectual [97].

El efecto patogénico relacionado con la alteración de los reguladores de la cromatina, dentro del grupo de las cromatinopatías, se asocia con una desregulación global de la transcripción. Esta desregulación afecta, principalmente, a genes involucrados en el desarrollo neuronal, la organización tridimensional de la cromatina y la regulación del ARN [98–101]. Como resultado, se producen disfunciones celulares que se manifiestan en una variedad de fenotipos característicos de estos trastornos. Estos fenotipos reflejan la compleja interacción entre las vías moleculares afectadas y los procesos esenciales del desarrollo neuronal. Así, el impacto observado no se limita únicamente al efecto directo de la variante patogénica, sino que también incluye las consecuencias derivadas de la desregulación de múltiples genes afectados de manera indirecta.

Clasificación de las cromatinopatías

Las cromatinopatías constituyen un grupo de trastornos que se caracterizan por un desequilibrio en el estado de la cromatina. Este grupo engloba una amplia variedad de enfermedades con diversas manifestaciones clínicas. En la actualidad, se han clasificado ciento setenta y nueve síndromes como cromatinopatías [102], siendo el SCdL un ejemplo representativo [103,104]. Los fenotipos asociados a estos trastornos, resultan de alteraciones en los genes que regulan la cromatina. Estas variantes patogénicas suelen dar lugar a síntomas complejos que afectan a múltiples sistemas y aparatos, con un impacto notable en el sistema nervioso [102].

Las proteínas afectadas, se clasifican en cuatro grupos principales según su función: modificadores de la cromatina, remodeladores de la cromatina, moduladores de las modificaciones químicas del ADN y ARN, y proteínas accesorias [105,106]. Recientemente, esta clasificación se ha ampliado para detallar aún más las funciones que desempeñan. Como resultado, se han identificado hasta diecisiete grupos, que incluyen modificadores y cofactores de histonas, remodeladores de cromatina y sus cofactores, proteínas de andamiaje, proteínas de la familia *Polycomb*, factores de transcripción, así como cofactores y modificadores de ADN y ARN [107] (Figura 2).

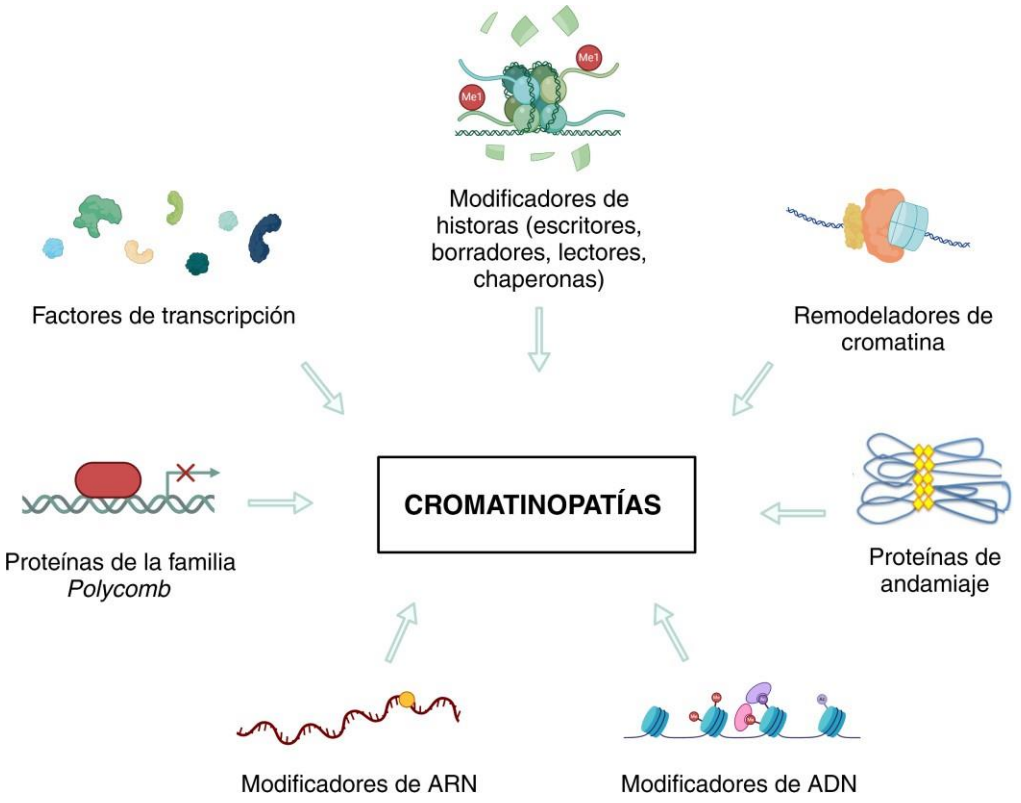


Figura 2. Representación gráfica de los principales grupos en los que se clasifican las proteínas afectadas en las cromatinopatías. (Imagen creada mediante BioRender)

La alteración de cualquiera de estas proteínas, provoca una disrupción del estado de la cromatina, que influye en la capacidad celular para regular la expresión génica de manera precisa [108]. Dado que estas proteínas suelen estar involucradas en la regulación de múltiples genes, las manifestaciones clínicas afectan a varios sistemas, como el sistema nervioso central, el sistema óseo y el sistema cardiaco [102]. La mayoría de las cromatinopatías cursan con retraso del desarrollo neurológico (HP:0012758), discapacidad intelectual (HP:0001249), convulsiones (HP:0001250), forma facial anormal (HP:0001999), retraso del crecimiento (HP:0001510), dificultades en la alimentación (HP:0011968) y anomalías de la morfología cardiaca (HP:0001627) y del sistema esquelético (HP:0000924) [102,109].

Un nuevo enfoque que se plantea para continuar investigando los aspectos aún desconocidos de las cromatinomatías, es el estudio multiómico [107]. Este enfoque permite realizar un análisis integral de todas las ciencias ómicas, lo que puede resultar en hallazgos más completos. A través de este análisis, es posible identificar las vías alteradas que contribuyen a la clínica observada y descubrir biomarcadores que faciliten el diagnóstico. El objetivo de este enfoque es promover el descubrimiento de nuevas cromatinopatías y, para aquellas ya identificadas, ofrecer una descripción más precisa.

Las posibilidades terapéuticas que se plantean para estas enfermedades incluyen terapias dirigidas en síndromes donde es posible revertir las anomalías presentes. Esto se puede lograr mediante la inhibición de enzimas que modifican la cromatina o restaurando la función normal de las proteínas mutadas [108]. Estas estrategias ofrecen nuevos tratamientos posibles para estas condiciones, que anteriormente no se habían considerado como opciones terapéuticas.

Síndrome Cornelia de Lange

El síndrome Cornelia de Lange (SCdL; OMIM #122470 (SCdL1), #300590 (SCdL2), #610759 (SCdL3), #614701 (SCdL4) y #300882 (SCdL5)), es una de las cromatinopatías más conocidas [103,104]. Este trastorno genético del neurodesarrollo afecta a múltiples órganos y sistemas. Aunque su prevalencia exacta es desconocida, se estima que oscila entre 1:10.000 y 1:30.000 nacidos vivos [110]. Sus manifestaciones clínicas presentan una amplia variabilidad, abarcando desde casos leves hasta formas graves de la enfermedad. La mayoría de los casos se asocian a variantes patogénicas en genes relacionados con el complejo cohesina, aunque se están identificando cada vez más casos vinculados a reguladores de la cromatina. La alta variabilidad fenotípica y genotípica hace que el diagnóstico de este síndrome sea un desafío significativo.

Características clínicas

El fenotipo del SCdL se caracteriza por una serie de signos y síntomas distintivos. Las manifestaciones clínicas más comunes incluyen asimetría craneofacial (HP:0004484), retraso de crecimiento intrauterino (HP:0001511) y postnatal (HP:0008897), discapacidad intelectual (HP:0001249), anomalías de las extremidades (HP:0040064), reflujo gastroesofágico (HP:0002020) e hirsutismo (HP:0001007). Además, el sistema nervioso central, musculoesquelético, cardíaco y gastrointestinal, son los más afectados en los pacientes [97,110,111].

Una de las características distintivas de este síndrome, es su facies, fácilmente reconocible desde el nacimiento (Figura 3). Los rasgos faciales incluyen microcefalia (HP:0000252), sinofridia (HP:0000664), cejas arqueadas (HP:0002553), pestañas largas (HP:0000527), surco nasolabial largo (HP:0000343) y/o liso (HP:0000319), bermellón del labio superior fino (HP:0000219), comisuras de la boca de oblicuidad descendente (HP:0002714), paladar alto (HP:0000218) y orejas de implantación baja (HP:0000369) [97,110,111] (Figura 3).

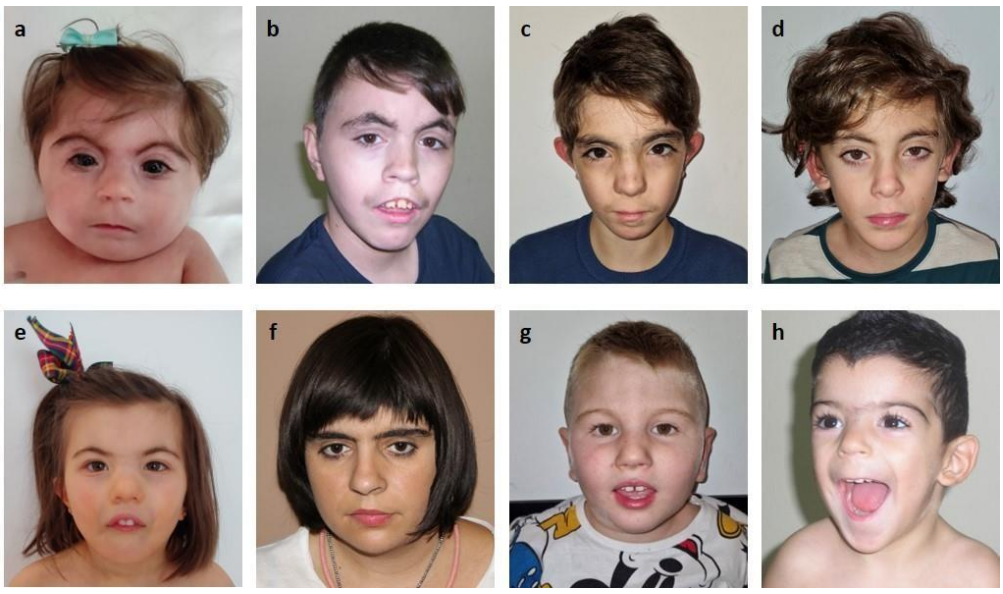


Figura 3. Fenotipo facial de individuos de la cohorte española del síndrome Cornelia de Lange. Fenotipo de una variante genética en a| *NIPBL*, b| *SMC1A*, c| *SMC3*, d| *RAD21*, e| *HDAC8*, f| *ANKRD11*, g| *BRD4*, h| *MAU2*. (Imágenes originales)

La afectación de las extremidades superiores es una característica típica del síndrome. Esta puede manifestarse de diversas maneras, desde manos pequeñas (HP:0200055), pliegue palmar transverso único (HP:0000954) y clinodactilia del quinto dedo (HP:0004209), hasta diversas formas de oligodactilia (HP:0012165) y/o monodactilia (HP:0004058). La manifestación más grave es la ausencia casi completa de las extremidades superiores [112].

Una de las presentaciones clínicas más frecuentes en los individuos con SCdL es el retraso de crecimiento, tanto intrauterino (HP:0001511) como postnatal (HP:0008897) [97,111]. Algunos estudios describen alteraciones hormonales o del desarrollo puberal, que se asocian con una mayor tendencia a la obesidad y resistencia a la insulina, así como con un nivel de actividad física reducido [113,114]. Además, el exceso de prolactina (HP:0000870) aparece con frecuencia en estos individuos [114]. En el sistema nervioso periférico, se han descrito déficits sensitivos y alta tolerancia al dolor, así como alteraciones en la respuesta de fibra fina [115]. Sin embargo, existen pocos estudios que analicen la afectación de este sistema.

Las anomalías congénitas cardiacas son una complicación común de este síndrome, afectando aproximadamente a un tercio de los individuos, y contribuyendo a una alta morbimortalidad [116]. También pueden aparecer manifestaciones cardiacas a posteriori, como alteraciones del músculo cardiaco que, aunque poco frecuentes, representan hasta el 3% de los fallecimientos [117]. Además, se han observado cambios en el tamaño y función ventricular en algunos individuos sin anomalías congénitas [118].

Otros fenotipos clínicos que puede aparecer, son la hernia diafragmática congénita (HP:0000776), la discapacidad auditiva (HP:0000365) y la miopía (HP:0000545) [111]. Relacionado a las conductas, este síndrome se caracteriza por comportamiento autista (HP:0000729), especialmente conductas repetitivas y déficits del lenguaje. Además, muchos pacientes muestran comportamientos auto dañinos (HP:0100716) [119–122]. Sin embargo, los estudios de imágenes cerebrales, necesarios para comprender las conexiones entre genes, cerebro y conducta, son limitados [123].

El fenotipo del SCdL se clasifica como un espectro, que abarca desde la presentación clásica, con una clínica más marcada, hasta características más leves o no clásicas del síndrome. El fenotipo clásico es distintivo y fácilmente reconocible, destacándose por dismorfismo craneofacial, retraso en el crecimiento y desarrollo, así como malformaciones en las extremidades. Por el contrario, los individuos con fenotipo no clásico pueden presentar solo algunas de las características comunes y mostrar grados variables de gravedad [97,111].

Bases genéticas y moleculares

El complejo cohesina es un regulador esencial de diversos aspectos de la biología celular. Pertenece al grupo de proteínas de mantenimiento estructural de los cromosomas (*Structural Maintenance of Chromosomes,* SMC) y participa en la unión de las cromátidas hermanas, la cohesión mitótica y meiótica, el mantenimiento de pliegues específicos del ADN, la reparación de roturas de doble cadena y la regulación transcripcional [119,124]

(Figura 4). Comprender su conformación es crucial para entender su actividad. Al igual que otras proteínas SMC, el complejo cohesina consta de una subunidad ATPasa y un complejo proteico de unión al ADN, no específico de secuencia [125]. La subunidad ATPasa utiliza ATP para activarse y plegar el genoma en bucles, lo cual es fundamental para la organización espacial y la regulación génica [126].

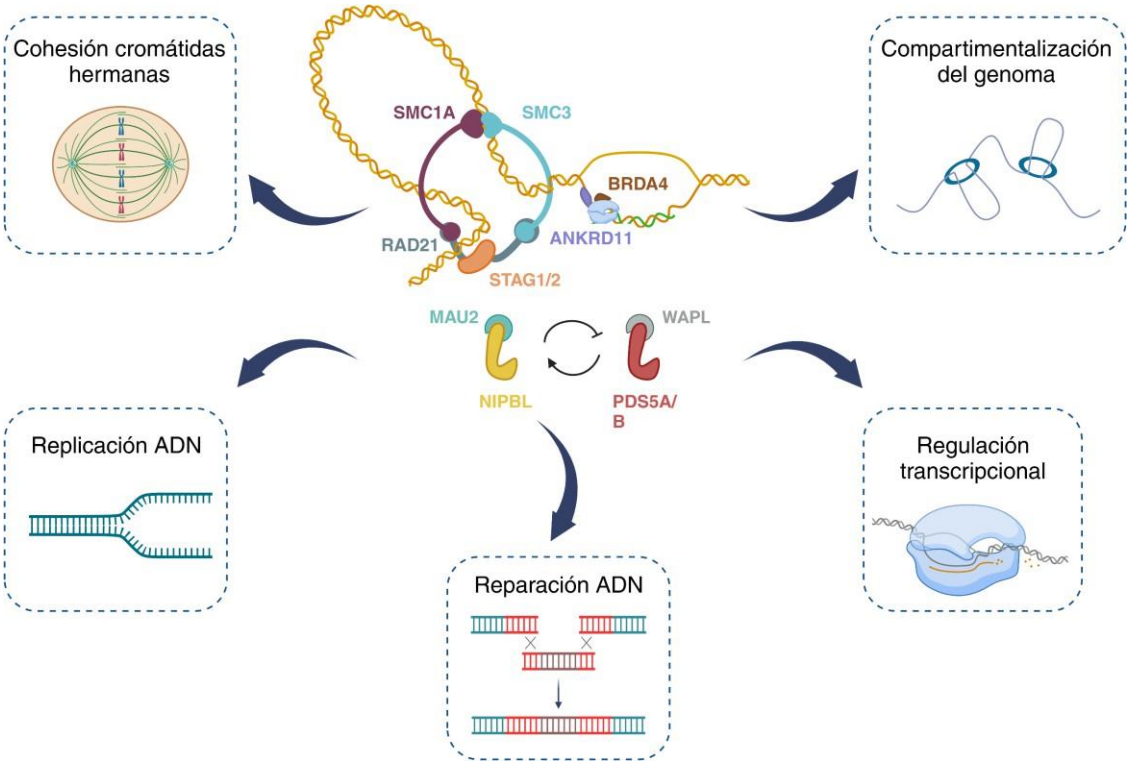


Figura 4. Componentes proteicos del complejo cohesina y funciones principales que desarrolla. (Imagen creada mediante BioRender)

La cohesina es un complejo en forma de anillo, compuesto por tres proteínas: SMC1A, SMC3 y RAD21. A este anillo tripartito, se unen diversas proteínas, dando lugar al complejo cohesina. Las subunidades auxiliares STAG1 y STAG2 son esenciales para la unión eficiente de la cohesina al ADN [127]. Otros reguladores importantes del complejo son NIPBL y MAU2, que facilitan la carga de la cohesina en el ADN [128]. Por otro lado, WAPL desempeña una función opuesta al unirse a la proteína PDS5, lo que permite la liberación de la cohesina de la cromatina [129].

El estudio de la estructura del complejo cohesina, es fundamental para comprender las diversas conformaciones que adopta al unirse al ADN y desplazarse a lo largo del genoma. Existen varios modelos que describen la conformación y función de este complejo [124,130,131]. Un análisis mediante microscopía electrónica ha identificado las conformaciones de los complejos de cohesina al unirse al ADN, así como las posiciones

de sus subunidades individuales [132]. Uno de los modelos sugiere que el complejo cohesina experimenta un proceso escalonado, que requiere de múltiples cambios conformacionales, que incluyen la oscilación del dominio bisagra, los movimientos de las cabezas de la ATPasa y la carga del ADN [131].

Hasta la fecha, se han identificado ocho genes asociados con el SCdL, la mayoría de los cuales participan en la estructura o regulación del complejo cohesina (*NIPBL*, *SMC1A*, *SMC3*, *RAD21*, *HDAC8*, *MAU2*) o, de manera directa o indirecta, en la regulación transcripcional (*ANKRD11*, *BRD4*) [111,128].

→ ***NIPBL.*** Fue el primer gen identificado como causante del SCdL, y es uno de los principales responsables de esta patología, afectando aproximadamente al 70 % de los casos [133,134]. Este gen codifica la proteína delangina, también conocida como NIPBL, una molécula de gran tamaño, que se expresa principalmente en dos isoformas diferenciadas únicamente en su región C-terminal [135]. Aunque no se ha logrado cristalizar la proteína completa, se han caracterizado ciertos dominios que incluyen regiones no estructuradas y múltiples repeticiones HEAT (*Huntingtin-elongation factor 3-protein phosphatase 2A-TOR1 repeats*), organizadas en una conformación similar a un gancho [136].

NIPBL juega un papel crucial en la carga del complejo cohesina y en el mantenimiento de la estabilidad genómica [124,137]. Varios estudios han demostrado que esta proteína es esencial para la cohesión y organización tridimensional del genoma, mediada por el complejo cohesina. Además, se han destacado las consecuencias perjudiciales de su haploinsuficiencia durante el desarrollo embrionario [138–141]. Los estudios de correlación genotipo-fenotipo indican que la mayoría de casos de variantes en *NIPBL*, generan un fenotipo clásico, más grave en comparación con el resto de genes asociados al síndrome [142]. Además, diversos estudios evidencian que las variantes de pérdida de función se relacionan con cuadros clínicos más graves [78,143–145].

→ ***SMC1A.*** Un número menor de individuos afectados, aproximadamente el 5%, presentan variantes patogénicas en el gen *SMC1A*, localizado en el cromosoma X, y que puede escapar de la inactivación [146]. Las variantes detectadas, mayoritariamente *missense*, *nonsense* o *frameshift*, se distribuyen uniformemente a lo largo del gen [147–149]. La proteína SMC1A consta de cinco regiones principales, destacando su motivo de unión al ATP en su extremo N-terminal y el dominio de unión al ADN en su extremo C-terminal (Uniprot: Q14683). La clínica asociada con este gen, tiende a ser más leve, el crecimiento de los pacientes está menos alterado, y sus rasgos faciales suelen ser menos pronunciados [147]. Además, las alteraciones en *SMC1A* pueden dar

lugar a diferentes fenotipos. La mayoría de las variantes genéticas sin sentido, están asociadas con manifestaciones clínicas leves del SCdL, mientras que las variantes de pérdida de función, se relacionan con epilepsia o fenotipos similares al síndrome Rett (RTS, OMIM #312750) [147,148,150–152].

→ ***SMC3* y *RAD21*.** *SMC3*, al igual que *SMC1A*, pertenece a la subfamilia de proteínas SMC, y se asocia con fenotipos no clásicos y más leves. En cuanto a las características faciales de estos pacientes, el surco nasolabial generalmente no es liso, y no se observan malformaciones significativas en las extremidades. Aunque la discapacidad intelectual es común, los problemas de comportamiento son menos frecuentes. Además, muchos de los individuos con este gen afectado, tienen una personalidad amable [153]. Por otro lado, el gen *RAD21*, codifica otro de los componentes estructurales del complejo cohesina [154]. La disfunción de esta proteína se asocia con un fenotipo no clásico, caracterizado por una clínica entre leve y moderada, con una afectación cognitiva más sutil y un dismorfismo facial menos pronunciado [149,155].

→ ***HDAC8.*** Está localizado en el cromosoma X, y a diferencia del gen *SMC1A*, no se ha descrito que escape de la inactivación [156]. Este gen representa aproximadamente el 6% de los casos del síndrome, con una amplia heterogeneidad de variantes descritas [111,149]. La proteína HDAC8 actúa como regulador clave del complejo cohesina, participando en la desacetilación de SMC3, y facilitando la disociación del complejo de la cromatina [157]. El fenotipo de los individuos es particularmente llamativo, destacando su facies característica, que incluye retraso del cierre de la fontanela anterior (HP:0001476), párpados caídos (HP:0030820), hiperterolismo (HP:0000316) y nariz bulbosa (HP:0000414). Dado que este gen se encuentra en el cromosoma X, los varones hemicigotos suelen presentar una afectación más grave, mientras que en las mujeres, la gravedad de la condición varía según el grado de inactivación del cromosoma X que contiene el alelo mutado [156].

→ ***BRD4* y *ANKRD11*.** Un pequeño porcentaje de individuos con fenotipos solapantes al SCdL, presenta variantes en los genes *BRD4* [158] y *ANKRD11* [159–162]. Ambos genes codifican proteínas involucradas en la regulación de la cromatina. El gen *ANKRD11*, en particular, se ha identificado como uno de los genes más frecuentemente mutados en individuos con trastornos graves del neurodesarrollo [55]. Además, se relaciona con el síndrome de KBG (SKBG, OMIM #148050) [163,164], lo que sugiere que variantes en este gen podrían estar asociadas a un espectro más amplio de síndromes del neurodesarrollo.

→ ***MAU2.*** El gen *MAU2* ha sido recientemente asociado al SCdL [128]. Se ha identificado al descubrir la interacción entre las proteínas NIPBL y MAU2, ambas implicadas en la carga del complejo cohesina en la cromatina. Esta interacción promueve el atrapamiento topológico del ADN, un proceso esencial para establecer la cohesión [165–167]. La primera variante patogénica identificada en *MAU2*, se relacionó con un fenotipo grave de un paciente que presentaba un dismorfismo facial característico del SCdL [128]. Además, se ha reportado un segundo caso que muestra como la interacción entre NIPBL-MAU2 se ve comprometida [168].

A pesar del creciente número de genes asociados al SCdL, los mecanismos patológicos que conducen al desarrollo de este trastorno aún no se comprenden completamente. La interacción entre NIPBL y el complejo cohesina se conoce cada vez mejor, así como los efectos de las alteraciones en estas proteínas [124,127]. Los individuos con SCdL no presentan alteraciones en el proceso de cohesión de las cromátidas hermanas [169], lo que llevó a que este síndrome dejara de clasificarse como una cohesinopatía. En cambio, se ha confirmado una desregulación transcripcional que parece ser responsable de la amplia variabilidad fenotípica observada en estos pacientes [170–172]. Esta desregulación afecta a genes implicados en el desarrollo de diversos tejidos y sistemas, como el neurológico, cardíaco, esquelético e inmunológico, lo que podría explicar las afectaciones en dichos sistemas [173–177]. Además, se ha observado un impacto en la organización tridimensional del genoma en células neuronales y fibroblastos de pacientes, estrechamente relacionado con la alteración en la expresión génica [175,178,179]. Recientemente, se han identificado mecanismos patológicos adicionales, como el daño del ADN y la senescencia celular, hallazgos que podrían ayudar a comprender el desarrollo de este síndrome [180–182].

Diagnóstico

Diagnóstico clínico

El diagnóstico clínico del SCdL se basa, principalmente, en sus características fenotípicas distintivas. Estas incluyen rasgos faciales típicos, retraso del crecimiento, discapacidad intelectual y malformaciones en las extremidades, así como otros hallazgos clínicos variables.

Desde el año 2007, se han desarrollo diferentes guías diagnósticas para establecer la clasificación clínica y molecular del síndrome. En un primer estudio, se definieron los criterios clínicos fundamentales para el diagnóstico, destacando la relevancia de los

rasgos faciales, junto con las malformaciones asociadas, como la reducción de miembros superiores y los defectos en órganos internos, categorizados según la gravedad de la afectación [110]. Posteriormente, en 2018, un consenso internacional actualizó estas directrices e introdujo una puntuación clínica estandarizada para facilitar y homogeneizar el diagnóstico (Tabla 1) [111]. Este sistema de puntuación, utilizado a día de hoy, integra características clínicas más y menos frecuentes (características cardinales y sugerentes, respectivamente), otorgando una puntuación específica a cada una de ellas. Los criterios con mayor puntuación, incluyen rasgos faciales distintivos, anomalías en extremidades y hernia diafragmática congénita, mientras que los de menor puntuación abarcan manifestaciones menos específicas, como microcefalia y retraso del crecimiento pre y postnatal (Tabla 1).

Tabla 1. Puntuación clínica para clasificar el síndrome Cornelia de Lange [111].

Características cardinales (2 puntos cada una si está presente)
→ Sinofridia (HP:0000664) y/o cejas arqueadas (HP:0002553) → Nariz corta (HP:0003196), cresta nasal cóncava (HP:0011120) y/o narinas antevertidas (HP:0000463) → Surco nasolabial largo (HP:0000343) y/o liso (HP:0000319) → Bermellón del labio superior fino (HP:0000219) y/o comisuras de la boca de oblicuidad descendente (HP:0002714) → Oligodactilia de la mano (HP:0012165) y/o adactilia (HP:0009776) → Hernia diafragmática congénita (HP:0000776)
Características sugerentes (1 punto cada una si está presente)
→ Retardo global del desarrollo (HP:0001263) y/o discapacidad intelectual (HP:0001249) → Retraso del crecimiento intrauterino (HP:0001511) → Retraso del crecimiento posnatal (HP:0008897) → Microcefalia (HP:0000252) (prenatal y/o postnatal) → Manos pequeñas (HP:0200055) y/o pies (HP:0001773) → Quinto dedo de la mano corto (HP:0009237) → Hirsutismo (HP:0001007)
Puntuación clínica
→ ≥11 puntos, de los cuales al menos 3 son cardinales: SCdL clásico → 9 o 10 puntos, de los cuales al menos 2 son cardinales: SCdL no clásico → 4-8 puntos, de los cuales al menos 1 es cardinal: se indican pruebas moleculares para SCdL → <4 puntos: insuficiente para indicar pruebas moleculares para SCdL

Este sistema permite clasificar a los pacientes en fenotipo clásico o no clásico, además de indicar la necesidad de realizar pruebas moleculares [111]. Por lo tanto, la implementación de la puntuación clínica internacional, establecida en 2018, ha supuesto la estandarización de la caracterización y diagnóstico del síndrome.

En la actualidad, la inteligencia artificial se está integrando en el proceso diagnóstico. Al tratarse de una patología con rasgos faciales reconocibles, las tecnologías de análisis facial, basadas en inteligencia artificial y algoritmos de aprendizaje profundo, son capaces de crear un prototipo facial (*Gestalt*) específico de cada trastorno [22]. La aplicación *Face2Gene* ha demostrado ser útil en el apoyo al diagnóstico clínico del SCdL [183]. Además, la reciente actualización de esta herramienta, que incorpora el algoritmo *GestaltMatcher*, basado en redes neuronales profundas, amplía significativamente su utilidad al identificar nuevas relaciones con otros síndromes que presentan morfologías faciales similares, las cuales podrían haber sido pasadas por alto en el proceso diagnóstico [184].

Diagnóstico genético

El diagnóstico genético del SCdL presenta una gran complejidad debido al elevado número de genes involucrados, la diversidad de variantes patogénicas y la alta prevalencia de mosaicismo [111]. Esta variabilidad genética es la responsable de la amplia heterogeneidad fenotípica, que ha hecho que el término haya evolucionado a Espectro Cornelia de Lange (ECdL) para incluir todos los fenotipos identificados del síndrome.

El abordaje del diagnóstico genético de este síndrome, requiere considerar varios aspectos clave, como la localización del gen afectado, el patrón de herencia, la posibilidad de presencia de mosaicismo y el tipo específico de variante genética que puede causarlo. Actualmente, se han identificado ocho genes responsables, seis de ellos con patrón de herencia autosómica dominante (*NIPBL*, *SMC3*, *RAD21*, *ANKRD11*, *BRD4* y *MAU2*), y dos ligados a cromosoma X (*SMC1A* y *HDAC8*). Estos últimos, tienen un impacto más grave en los varones, mientras que las mujeres portadoras pueden mostrar una sintomatología variable dependiendo de la inactivación del cromosoma X [156].

La mayoría de las variantes reportadas del SCdL son *de novo*, aunque hay un número considerable de recurrencias familiares, principalmente debido a la transmisión de variantes patogénicas de genes ligados al cromosoma X [156,185,186]. Un aspecto interesante, es la presencia de un porcentaje significativo de variantes postcigóticas que dan lugar a una mosaicismo, fenómeno que debe tenerse en cuenta en el diagnóstico genético [187–189]. A pesar de su relevancia, la implicación del mosaicismo en la clínica y patogénesis, no es bien comprendida. Generalmente, las variantes postcigóticas tienden a causar fenotipos más leves, sin embargo, en el caso del SCdL, los individuos con mosaicismo pueden presentar manifestaciones clínicas tan graves como las observadas en pacientes con variantes germinales [189]. Además, se ha observado una eliminación en

sangre del alelo mutado en estos pacientes, lo que subraya la importancia de analizar muestras biológicas de diferentes orígenes [189].

Las variantes genéticas vinculadas al SCdL abarcan un amplio espectro, reflejando la gran complejidad molecular del trastorno [111,149]. Predominan las variantes puntuales (SNVs), entre ellas las de pérdida de función, como variantes sin sentido (*nonsense*) y desplazamientos del marco de lectura (*frameshift*) [149,156,190,191]. También se han identificado variantes de corte y empalme (*splicing*), incluso en regiones intrónicas profundas [139,192–194]. Aunque menos frecuentes, se han relacionado también alteraciones estructurales, principalmente variantes en el número de copias (CNVs) que afectan a regiones intrágenicas [140,156,195–198]. Por tanto, resulta fundamental emplear técnicas complementarias que garanticen un análisis completo.

En la actualidad, el enfoque diagnóstico del SCdL se basa en el uso de paneles de genes dirigidos, que incluyen, no solo los genes principales, sino también genes relacionados con trastornos que presentan fenotipos solapantes. Este enfoque permite un estudio multigénico más amplio y eficiente, a diferencia del tradicional método de Sanger. Cuando los resultados de los paneles de genes son negativos, se inicia el análisis de ADN procedente de muestras biológicas alternativas a la sangre, como células bucales o fibroblastos, para evaluar un posible mosaicismo genético [189]. En caso de continuar con un resultado negativo, se plantean técnicas avanzadas para detectar variantes estructurales, como el aCGH de alta resolución, o el mapeo óptico [199]. Finalmente, se puede recurrir a la WGS para realizar un análisis más exhaustivo de variantes en secuencias no codificantes [13,29,30]. Además, en estudios recientes, se propone incluir la secuenciación de ARN (*RNA-sequencing*) en el proceso diagnóstico [200,201].

El diagnóstico prenatal puede estar indicado en determinadas situaciones. Se recomienda en familias con antecedentes de SCdL, en nuevos embarazos donde uno de los progenitores porta una variante patogénica, o cuando en la ecografía el feto presenta características sugerentes del síndrome. En cualquiera de estos casos, se extrae ADN de células fetales, que se somete a secuenciación para evaluar los riesgos potenciales [111].

Los avances en las técnicas de citogenética y secuenciación, están facilitando la detección de nuevas variantes y genes causales del síndrome (Figura 5). A medida que se identifiquen nuevos genes implicados, el enfoque del diagnóstico genético evolucionará hacia estudios más integrales, como el mapeo óptico y la *long read sequencing* [199]. Además, estas nuevas tecnologías se complementarán con herramientas bioinformáticas que permitirán interpretar, de manera más precisa, las variantes genéticas recién identificadas.

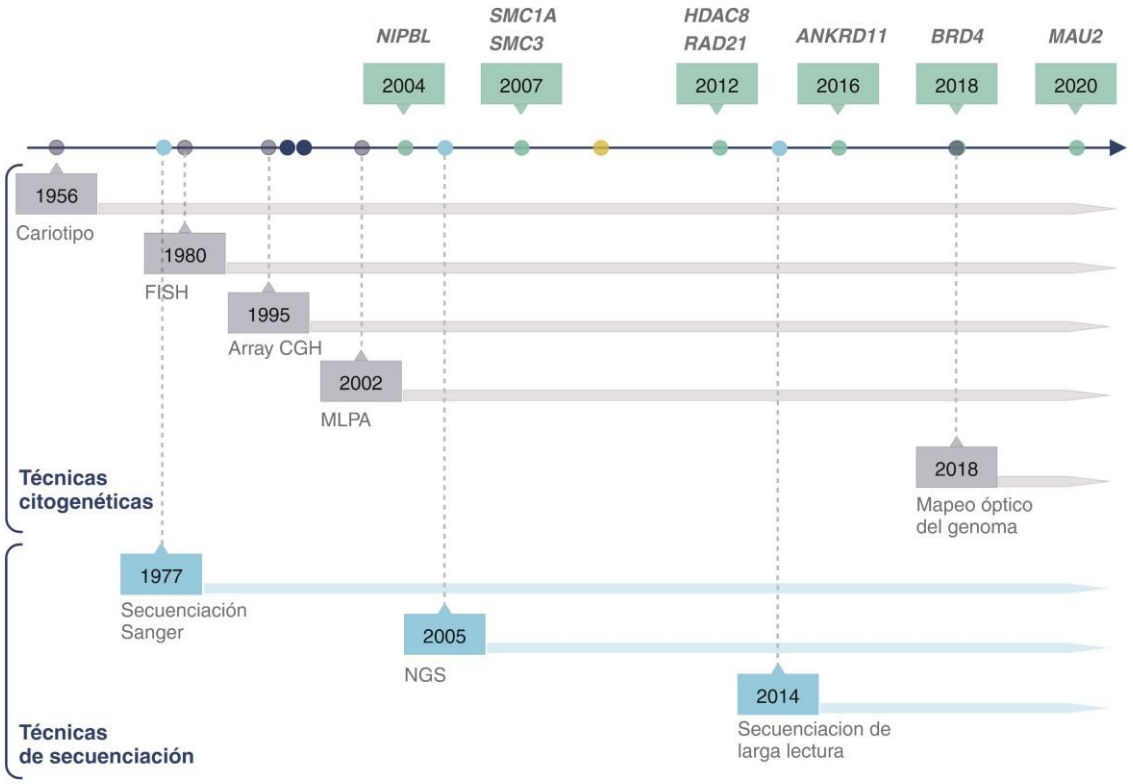


Figura 5. Relación temporal entre la identificación de genes relacionados al síndrome Cornelia de Lange y la evolución de técnica se secuenciación. (Imagen creada mediante BioRender)
· Genes: *NIPBL* [133,134] > *SMC1A* y *SMC3* [202] > *HDAC8* [157] > *RAD21* [155] > *ANKRD11* [159] > *BRD4* [158] > *MAU2* [128].
· Técnicas citogenéticas: Cariotipo [203] > Hibridación fluorescente in situ (FISH) [204] > Array de hibridación genómica comparativa (array CGH) [205] > Amplificación de sondas tras ligación múltiple (MLPA) [206] > Mapeo óptico del genoma [207].
· Técnicas de secuenciación: Secuenciación Sanger [208] > Secuenciación de nueva generación (NGS) [209,210] > Secuenciación de larga lectura [211].

Diagnóstico diferencial

La variedad de fenotipos asociados al SCdL, en ocasiones, puede complicar el diagnóstico diferencial. En los últimos años, se han descrito numerosos TND que han llevado a la creación de redes funcionales que facilitasen su clasificación [63]. El SCdL, conocido a nivel clínico y molecular, requiere del establecimiento de diagnósticos diferenciales con otros cuadros clínicos afines. Un ejemplo relevante es el síndrome de PACS1 (SHMS, OMIM #615009), cuyos primeros pacientes fueron clasificados como SCdL antes de su caracterización [212]. La causa genética de este síndrome difiere de la base molecular del SCdL, lo que facilita su diagnóstico genético. Además, recientemente se ha descrito una función alterada de la sinapsis [72,73], a diferencia del SCdL, relacionado con una alteración en la regulación transcripcional y clasificado dentro del grupo de las cromatinopatías [103,104].

En la actualidad, el diagnóstico diferencial del SCdL debe considerar las cromatinopatías que presentan similitudes clínicas con este trastorno (Figura 6). Es necesario evaluar aquellos síndromes en los que ya se han identificado casos con fenotipos solapantes al SCdL, como el síndrome de CHOPS (OMIM #616368) [213] y el síndrome de Rubinstein-Taybi (OMIM #180849, #613684) [161]. Además, el diagnóstico diferencial entre el SCdL y el SKBG es fundamental, debido a sus similitudes fenotípicas y al hecho de que el gen *ANKRD11* se asocie a ambos síndromes. Ambos trastornos presentan características faciales particulares, como sinofridia (HP:0000664), narinas antevertidas (HP:0000463) y bermellón del labio superior fino (HP:0000219), además de retardo global del desarrollo (HP:0001263) y anomalías del sistema esquelético (HP:0000924). Sin embargo, el SCdL se diferencia por rasgos distintivos como microcefalia (HP:0000252) y anomalías de las extremidades (HP:0040064), mientras que el SKBG, se caracteriza por un fenotipo que incluye un desarrollo intelectual variable, anomalías dentales y un crecimiento generalmente normal [159,164]. Además, el SCdL puede presentar un espectro más amplio de anomalías congénitas, lo que puede no ser tan característico del SKBG. Por lo tanto, un análisis genético y una evaluación clínica detallada, son esenciales para establecer un diagnóstico preciso y diferenciar entre estos dos síndromes.

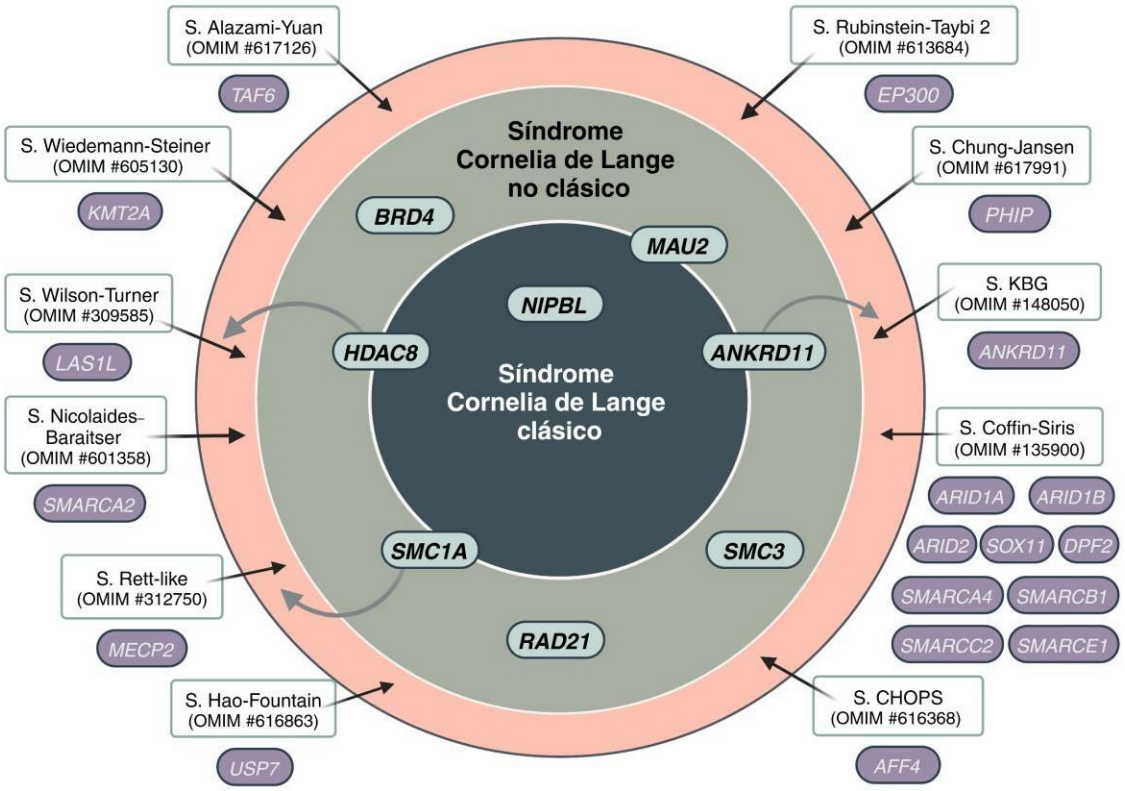


Figura 6. Espectro Cornelia de Lange y síndromes solapantes. Mediante círculos, se agrupan los genes que se incluyen dentro del SCdL clásico y no clásico. El círculo exterior naranja abarca los síndromes que se solapan clínicamente. En cada uno de estos trastornos se indica el gen/genes responsables, así como su número OMIM. Con flechas grises, se indican los genes del SCdL que se han relacionado también con otros síndromes. (Imagen creada mediante BioRender)

Objetivos

El objetivo principal es lograr un diagnóstico clínico y genético preciso de las entidades nosológicas susceptibles de ser incluidas en el Espectro Cornelia de Lange.

Como objetivos secundarios se persigue:

- Profundizar en la caracterización clínica y molecular que posibilite la identificación y clasificación de nuevas entidades nosológicas.
- Establecer protocolos para la reevaluación de datos clínicos y genéticos a fin de precisar el diagnóstico y conocer la evolución fenotípica con la edad.
- Detectar e interpretar variantes genéticas complejas.
- Identificar nuevos genes causales asociados al síndrome Cornelia de Lange.

Trabajos

TRABAJO 1:

Molecular Basis of the Schuurs-Hoeijmakers Syndrome: What We Know about the Gene and the PACS-1 Protein and Novel Therapeutic Approaches

María Arnedo M, Ángela Ascaso, Ana Latorre-Pellicer, **Cristina Lucia-Campos**, Marta Gil-Salvador, Ariadna Ayerza-Casas, María Jesús Pablo, Paulino Gómez-Puertas, Feliciano J Ramos, Gloria Bueno-Lozano, Juan Pié, Beatriz Puisac

Molecular Basis of the Schuurs-Hoeijmakers Syndrome: What We Know about the Gene and the PACS-1 Protein and Novel Therapeutic Approaches

International Journal of Molecular Sciences

Agosto 2022

doi: 10.3390/ijms23179649

PMID: 36077045

International Journal of
Molecular Sciences

MDPI

Review

Molecular Basis of the Schuurs–Hoeijmakers Syndrome: WhatWe Know about the Gene and the PACS-1 Protein and Novel Therapeutic Approaches

María Arnedo [1], Ángela Ascaso [1], Ana Latorre-Pellicer [1], Cristina Lucia-Campos [1], Marta Gil-Salvador [1], Ariadna Ayerza-Casas [1,2], María Jesús Pablo [1], Paulino Gómez-Puertas [3], Feliciano J. Ramos [1,4], Gloria Bueno-Lozano [4], Juan Pié [1,*] and Beatriz Puisac [1,*]

1 Unit of Clinical Genetics and Functional Genomics, Department of Pharmacology-Physiology, School of Medicine, University of Zaragoza, CIBERER-GCV02 and IIS-Aragon, E-50009 Zaragoza, Spain
2 Unit of Paediatric Cardiology, Service of Paediatrics, Hospital Universitary Miguel Servet,E-50009 Zaragoza, Spain
3 Molecular Modelling Group, Center of Molecular Biology "Severo Ochoa" (CSIC-UAM), Cantoblanco, E-28049 Madrid, Spain
4 Department of Pediatrics, Hospital Clínico Universitario "Lozano Blesa", E-50009 Zaragoza, Spain
* Correspondence: juanpie@unizar.es (J.P.); puisac@unizar.es (B.P.); Tel.: +34-976-761677 (J.P.); +34-976-761765 (B.P.); Fax: +34-976-761700 (J.P. & B.P.)

Citation: Arnedo, M.; Ascaso, Á.; Latorre-Pellicer, A.; Lucia-Campos, C.; Gil-Salvador, M.; Ayerza-Casas, A.; Pablo, M.J.; Gómez-Puertas, P.; Ramos, F.J.; Bueno-Lozano, G.; et al. Molecular Basis of the Schuurs–Hoeijmakers Syndrome: What We Know about the Gene and the PACS-1 Protein and Novel Therapeutic Approaches. *Int. J. Mol. Sci.* **2022**, *23*, 9649. https://doi.org/10.3390/ijms23179649

Academic Editors: Lidia Larizza and Maria Vittoria Cubellis

Received: 28 July 2022
Accepted: 23 August 2022
Published: 25 August 2022

Abstract: The Schuurs–Hoeijmakers syndrome (SHMS) or *PACS1* Neurodevelopment Disorder (*PACS1*-NDD) is a rare autosomal dominant disease caused by mutations in the *PACS1* gene. To date, only 87 patients have been reported and, surprisingly, most of them carry the same variant (c.607C>T; p.R203W). The most relevant clinical features of the syndrome include neurodevelopment delay, seizures or a recognizable facial phenotype. Moreover, some of these characteristics overlap with other syndromes, such as the PACS2 or Wdr37 syndromes. The encoded protein phosphofurin acidcluster sorting 1 (PACS-1) is able to bind to different client proteins and direct them to their subcellular final locations. Therefore, although its main function is protein trafficking, it could perform other roles related to its client proteins. In patients with *PACS1*-NDD, a gain-of-function or a dominant negative mechanism for the mutated protein has been suggested. This, together with the fact that most of the patients carry the same genetic variant, makes it a good candidate for novel therapeuticapproaches directed to decreasing the toxic effect of the mutated protein. Some of these strategiesinclude the use of antisense oligonucleotides (ASOs) or targeting of its client proteins.

Keywords: Schuurs–Hoeijmakers syndrome; SHMS; *PACS1*-NDD; trafficking protein; targeted therapy

1. Introduction

The Schuurs–Hoeijmakers syndrome (SHMS), or *PACS1* Neurodevelopmental Disor- der (*PACS1*-NDD) (MIM# 615009), is a rare autosomal dominant disease [1], which has recently been included in a group of genetic disorders of cellular trafficking [2]. It was firstly described in 2012 in two patients with intellectual disability (ID) and a striking facial resemblance; they both carried the same mutation in the *PACS1* gene [1]. To date, less than a hundred patients with *PACS1*-NDD have been genetically diagnosed, and, surprisingly,only three pathogenic variants have been reported [1,3–6]. Furthermore, most of the pa-tients have been described in the last few years, when the *PACS1* gene was included in the gene panel analysis for neurodevelopmental disorders; therefore, it is possible that the prevalence of this syndrome could be underestimated.

Although the clinical characteristics of *PACS1*-NDD patients are well described, there is little evidence on the pathomolecular mechanisms of this disease. The affected protein,phosphofurin acid cluster sorting 1 (PACS-1), belongs to a family of two members of membrane and protein trafficking regulators [7]. It was initially discovered in 1998 [8], and it has been reported in metazoans, invertebrates [9] and

vertebrates [10]. Lower species possess only a single gene of the *PACS* family; however, the *PACS* gene was duplicatedwith the appearance of vertebrates, resulting in the *PACS1* and *PACS2* genes [10].

PACS-1 is a multifunctional membrane traffic regulator that plays an important rolein cellular homeostasis [8,10]. An initial function of PACS-1 was the transport of several proteins between endosomes and the trans-Golgi network (TGN) [7]. However, in the lastfew years, novel roles have been proposed, such as a Ca^{2+} flux regulator or its probable implication in genomic stability [10–12]. The knowledge of the molecular mechanisms implied in the development of the *PACS1*-NDD is essential to make proper phenotype– genotype correlations and to propose therapies that could help and improve the quality of life of these patients. In this review, we will carry out a detailed study of the functionsalready known of the PACS-1 protein, of its role in the progression of the disorder and the latest improvements in the development of successful therapeutic strategies.

2. Clinical Characteristics of PACS1-NDD

The *PACS1*-NDD was first described in two unrelated male patients in 2012 [1]. Since then, about 100 patients have been reported in the literature [1,3–6]. All patients are described have a neurodevelopmental delay with an intellectual disability and psychomotor retardation. The range varies from mild to severe, although most of them have a moderate delay [13]. Language skills are universally affected, more severely than motor skills. Hypotonia is reported in approximately one third of the subjects, and it improves over time [13]. However, one patient has been described as having an impairment in their walking abilities over time [3]. Individuals with the *PACS1* mutation interact with others and appreciate receiving personal affection, but also show behavioral difficulties. Seizures are a common clinical feature, affecting 60 percent of patients; most of them are generalized as tonic-clonic, and are well controlled with antiepileptic drugs. Brain abnormalities have been found, mostly cerebellar. Other findings include ventriculomegaly, hydrocephalus or atrophy of the corpus callosum [3,5,14–18].

In addition to the neurological disorders, patients show a characteristic facial phe- notype, which is easily recognizable by a clinician. It is characterized by full and arched eyebrows, hypertelorims with downslanting palpebral fissures, long eyelashes, a bulbousnose, a flat philtrum and large low-set ears [13].

Patients also suffer from other congenital anomalies, from cardiac anomalies to ocular alterations, where septal defects and coloboma stand out. They may also have skele-tal abnormalities (abnormal skull shape), cryptorchidism or feeding problems, among others [3,19]. These clinic characteristics are summarized in Table 1.

Table 1. Clinical characteristics of the *PACS1*-NDD, PACS2, Wdr37, Kabuki and Cornelia de Lange Syndromes.

Clinical Feature	HPO ID *	*PACS1*-NDD	PACS2 Syndrome	Wdr37 Syndrome	Kabuki Syndrome	SCdL
Neurodevelopmental features						
Intellectual disability	0001249	Obligate	Very frequent	Obligate	Obligate	Very frequent
Autism spectrum disorder	000729	Occasional	Occasional	Occasional	Occasional	Frequent
Development delay	0012758	Obligate	Very frequent	Obligate	Obligate	Occasional
Speech delay	0000750	Very frequent	Very frequent	Frequent	Occasional	Frequent
Hypotonia	0001252	Frequent	Frequent	Frequent	Frequent	Occasional
Seizures	0001250	Frequent	Very frequent	Very frequent	Occasional	Occasional

Table 1. *Cont.*

Clinical Feature	HPO ID *	*PACS1*-NDD	PACS2 Syndrome	Wdr37 Syndrome	Kabuki Syndrome	SCdL
Congenital malformations						
Dysmorphic facial features						
Full and arched eyebrows	0002553	Frequent	Frequent	Frequent	Very frequent	Very frequent
Hypertelorism	0000316	Frequent	Frequent	Frequent	Occasional	Very rare
Downslanting palpebral fissures	0000494	Frequent	Frequent	Frequent	Very frequent	Very rare
Bulbous nasal tip	0000414	Frequent	Very frequent	Obligate	Frequent	Very frequent
Downturned mouth	0002714	Frequent	Frequent	Frequent	Very rare	Very frequent
Thin upper lip	0000219	Frequent	Very frequent	Very frequent	Occasional	Very frequent
Brain abnormalities						
Hypoplasia or partial agenesis of the cerebellar dermis	0006817	Frequent	Frequent	Obligate	Occasional	Occasional
Ophthalmologic						
Coloboma	0000589	Occasional	Occasional	Very frequent	Occasional	Very rare
Congenital heart anomalies						
Atrial or ventricular septal defects	0001671	Frequent	Occasional	Frequent	Frequent	Frequent
Others						
Feeding/GI issues	0011968	Occasional	Occasional	Very frequent	Frequent	Frequent
Skeletal anomalies	0000924	Occasional	Frequent	Frequent	Frequent/ very frequent	Frequent
Cryptorchidism	0000028	Frequent	Frequent	Very frequent	Occasional	frequent

* HPO ID, Human Phenotype Ontology Identifier. Obligate 100%; Very frequent 80–99%; Frequent 30–79%;Occasional 5–29%; Very rare 1–4%; Excluded 0%. Light grey: same disease spectrum.

The strong similarity between *PACS1*-NDD, PACS2 syndrome and the recently diag-nosed Wdr37 syndrome is striking. The three disorders share a very similar facial gestalt,intellectual disability, neurodevelopmental delay and seizures, suggesting that they might be included in the same disease spectrum [20–25]. Moreover, a new group of diseases, caused by cellular trafficking defects and characterized by neurodevelopmental disor- ders and skeletal abnormalities, has been recently proposed [2,20–26]. On the other hand, there are other closely related clinical syndromes, including Cornelia de Lange, Kabuki or Coffin-Siris, although they do not share the facial gestalt (Table 1).

3. Molecular Basis of the Disease

3.1. Genetic Update

To date, 87 patients with a *PACS1* deficiency have been genetically diagnosed, and,surprisingly, most of them carry the same pathogenic variant, the missense c.607C>T (p.R203W). This variant has been clearly demonstrated to be pathogenic because none ofthe parents tested were carriers [3], which means that all patients had a de novo mutation. There is one other pathogenic variant, c.608G>A, reported in only one patient, which results in a change in the same position, but the amino acid change is different, p.R203Q [5]. These patients have most of the typical features of the *PACS1*-NDD (Table 1). Lately, twonovel missense variants have been reported, one in the ClinVar database, the c.1574G>A(p.R525K), whose relationship with the *PACS1*-NDD has to be studied [27]. The other is in a broad study about the autism spectrum disorder (p.R245W) without a specific phenotype of the disease [28].

More recently, a multi-exon deletion of *PACS1* has been reported [6]. Liu and Cols found in four members of a three-generation pedigree the deletion of the exons 12 to 24 in the *PACS1* gene. However, the phenotype of this variant was milder, with slight speech and cognitive delay, only affecting two members of the last studied generation [6]. Furthermore, several databases have reported chromosomal reorganizations where the *PACS1* gene is involved (ClinVar, DECIPHER) [27,29]. Nevertheless, the huge number of affected genes does not allow us to directly relate the patients'

phenotype with the *PACS* gene.

3.2. PACS1 Gene Regulation

PACS1 is a gene that is broadly expressed in human tissues (GTEx database) [30]. According to the BrainSpan and EvoDevo databases, its mRNA expression is upregulated during fetal brain and cerebellum development, and it decreases after birth to slightly increase in puberty [31,32]. Its expression level is also important in pubertal testis tis-sues [31,32]. This specific tissue distribution could be the reason for some of the clinical characteristics of *PACS1*-NDD patients.

The *PACS1* gene is located on 11q13.1 and contains 24 exons and at least 16 transcripts (ensembl) [33]. Its expression regulation has been poorly studied, but some proteins have been proposed as regulators of PACS-1 expression, such as the P300/CBP-associated factor (PCAF) or the transcriptional adaptor protein 3 (ADA3). Chromatin inmuno precipitation (ChIP) experiments showed an enrichment in the promoter of *PACS1* of the proteins PCAF and ADA3. The downregulation of both factors decreased the relative *PACS1* expression level in HeLa and HCT116 cell lines. These facts point to a plausible role of PCAF andADA3 as gene regulators of *PACS1* [34].

In addition to that, the study of its 3′-UTR sequence showed two putative binding sites for the miRNAs, 34a and 449a. In tumor tissues, an overexpression of these miRNAsand a downexpression of the PACS-1 protein has been reported [35]. Furthermore, the overexpression of another miRNA, miR-485-5p, induces a decrease of PACS-1 in pericytes that has been related to Alzheimer's disease progression [36].

3.3. Characteristics of the PACS-1 Protein

PACS-1 is a protein of 963 amino acids with several domains and key regions located in the cytosol and nucleus [12,37]. Its N-terminal region, called ARR (atrophin-1-related region), is followed by the FBR (furin-binding), the MR (middle) and the CTR (C-terminal) domains (Figure 1). The FBR binds client proteins, such as furin, as well as the cytoplas-mic membrane trafficking machinery. Several specific sequences have been described, which are responsible for the binding of PACS-1 to the clathrin adaptors AP-1 and AP-3 (E_{168}TELQLTF) or to the monomeric adaptor GGA3 (K_{249}IY). Moreover, in the FBR is the binding sequence of the protein kinase CK2 (R_{196}RKRY) that phosphorylates the S_{278} residue located in the MR autoregulatory domain (S_{278}EEEEE). Furthermore, the MR also has a nuclear localization sequence (NLS) (V_{311}KKTRRKL) and a nuclear export sequence (NES) (L_{366}DELYDSLEM) [37]. The entry and exit from the nucleus of PACS-1 depends on the receptors importin alpha 5 (IPO5/KPNA1) and exportin 1 (XPO1) [37].

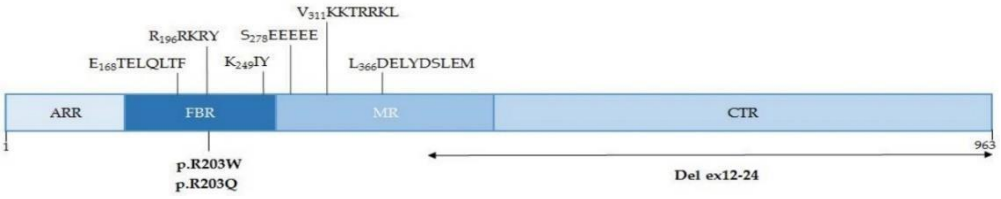


Figure 1. Schematic representation of the PACS-1 protein. Key sequences and pathological variants are located in the sequence. ARR, atrophin-1-related region; FBR, furin-binding region; MR, middle region; CTR, *C*-terminal region.

The three-dimensional (3D) structure of PACS-1 is still unknown. Some attempts have been made using ab-initio modeling, such as the one described for residues located between V_{117} and D_{300}, which are predicted as a globular domain [7]. Different unpublished analyses performed using structure prediction systems using hidden Markov models [38,39] roughly coincide in predicting a globular structure for the C-terminal subdomain (residues from R_{622} to L_{956}), similar to a regulatory subunit of phosphoinositide 3-kinase. They also predicta structure similar to a C2 domain (calcium/lipid-binding domain) for the segment between amino acids P_{98} and E_{261}, which would include the amino acid R_{203} (variant p.R203W). Unfortunately, not all of these data are conclusive. Future knowledge of the structure of PACS-1 would be

key to knowing more accurately how the mutations in the protein affectits structure and to developing therapies that are more precise for these patients.

3.4. Functions of PACS-1

Nowadays, more than 100 proteins have been described that can putatively bind toPACS-1 [11]. This could explain, in part, the number of different processes where PACS-1might be involved. Therefore, although the main function of PACS-1 is related to proteintrafficking, sometimes it is difficult to discern between the PACS-1 function and the role of some of its client proteins (Figure 2) (Supplementary Table S1).

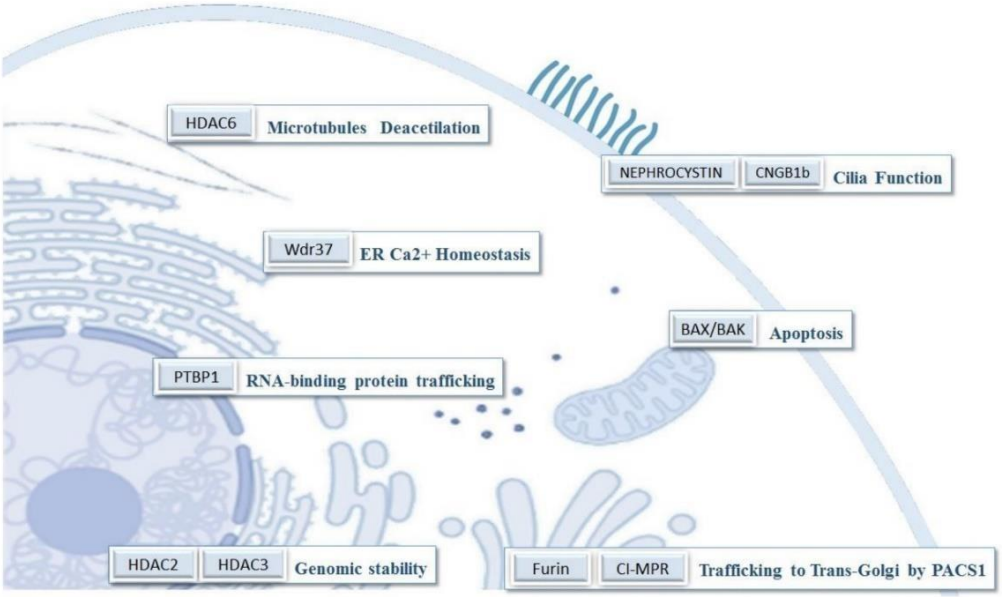


Figure 2. Some of PACS-1's client proteins and their subcellular location and function.

The first described function of the PACS-1 protein was its role in the regulation of membrane traffic proteins. This specific function is well known and conserved among vertebrates in PACS proteins. The activation mechanism involves the phosphorylation ofits S_{278} residue by CK2. The activated PACS-1 protein is able to bind its client proteins, transferring them to their final location. It mediates the trafficking of proteins from the plasma membrane to the early endosomes (SorLA) [40] and from the late endosomes to the trans-Golgi network (TGN) (furin, CI-MPR) [8,41]. Moreover, some proteins related to the function of the primary cilium, such as nephrocystin or CNGB1b, are also traffick- ing [42,43]. Besides, PACS-1 has been linked to the cytosolic HDAC6, which is involved in the deacetilation of microtubules, in the Golgi integrity and cilium retraction [44,45].

Another function for the PACS family of proteins is connected to the trafficking of proteins related to the apoptosis pathway [10]. PACS2 allows the transferring of Bid to the mitochondria, where it is cleaved to tBid, facilitating the triggering of apoptosis [46]. On the other hand, the downexpression of PACS-1 increases cell survival. The mechanism implies a failure in the BAX/BAK oligomerization, avoiding the mitochondrial outer membrane permeabilization (MOMP) [34]. The increase in cell survival when PACS-1 is downregulated has been linked with a worst prognosis in gastric cancers, and it has beenproposed as a biomarker [47].

Moreover, it has been reported that the complex PACS1–Wdr37 facilitates the reg- ulation of calcium flux between the endoplasmic reticulum and the cytosol [11]. The endoplasmic reticulum's Ca^{2+} release, mediated by the inositol 1,4,5-trisphosphate receptor (IP3R), is regulated by this complex. The deletion of PACS-1 provoked a decrease in the Wdr37 protein, and the deficiency of both caused a reduction in the IP3R expression level. This could explain the overlapping phenotypic characteristics

found between the PACS1 and Wdr37 deficiency patients (Table 1).

In the nucleus, PACS-1 binds to PTBP1, a protein involved in the binding and traf- ficking of RNA [36]. Moreover, it interacts with and could stabilize HDAC2 and HDAC3, contributing to genomic stability [12,35].

4. Relationship between PACS-1 Function and the PACS1-NDD Patient's Phenotype

Nowadays, there are a few studies regarding the relationship between PACS-1 func-tion and the *PACS1*-NDD phenotype. It is proposed that the most frequent PACS-1 variant, p.R203W, causes a gain-of-function (GOF) or a dominant negative mechanism [1,48]. How- ever, the truncation of the protein (del ex12-24) with a loss-of-function has been described in two patients with a mild and not characteristic *PACS1*-NDD clinic [6].

The mutated amino acid (R203W) is in the furin cargo-binding domain and in close proximity to the CK2 binding motif (Figure 1), so it is possible that the correct binding and phosphorylation function of CK2 could be compromised [14]. Several experimental approaches have been carried out with a mutant PACS-1 to explore the differences in the binding of PACS-1 to its client proteins. In this sense, some protein–protein interactionsare not affected (AP3D1, CLCN7, HDAC2/3) [1,12], whilst others decrease their bindinglevel (TRPV4v2) [1] and others increase it (HDAC6) [45]. On the other hand, experimentsin zebrafish embryos where the mutant mRNA of *PACS1* was injected showed a decreasein cranial cartilaginous structures compared to the control. This could be due to the fact that the migration of cranial-neural-crest cells (CNCCs) is related to PACS1 [1].

An experimental model of the *PACS1*-NDD of forebrain organoids has been developed (PACS1$^{(+/R203W)}$), besides the knock-out of the *PACS1* gene (PACS1$^{(-/-)}$). Gene expression pattern experiments showed differences between the PACS1$^{(+/R203W)}$ and the control model,but not between the PACS1$^{(-/-)}$ and the control, supporting the idea that this specific mutation confers a GOF and a toxic effect on the protein. Genes directly related to the autism spectrum disorder (ASD) and to the development of GABAergic synapses are upregulated in the disease model [48]. These results support the previous findings that about 40% of the *PACS1*-NDD patients have been formally diagnosed with autism [49].

These results, in zebrafish and forebrain organoids, could explain, in part, the neuro- logic impairment development of *PACS1*-NDD patients. Moreover, other proteins related to intracellular trafficking have also been associated with craniofacial diseases [26] and neurodevelopmental disorders [2].

5. Therapy for PACS1 Deficiency Patients

To date, the *PACS1*-NDD syndrome has been symptomatically treated by a multidis-ciplinary team. The neurodevelopmental alterations recommend an early interventionalprogram, which includes occupational, physical and speech therapy. Anxiety and behav- ioral problems have been managed with psychotropic drugs. Seizures respond well to theclassical epilepticus treatment. It is suggested that early physical therapy treatment for motor dysfunction can improve mobility and decrease the risk of later orthopedic compli-cations. Feeding problems might need nutritional intervention therapy and, in severe cases, a gastrostomy tube [13].

In recent years, more targeted therapy approaches are being developed for some rare diseases. In this sense, the *PACS1*-NDD is a good candidate [50] because most of the patients carry the same mutation, and a gain-of-function or a dominant negative mechanism has been proposed [1,3,48]. However, the highest expression level of PACS-1 during fetal brain development could limit the effectiveness of the treatment [31,32]. Nowadays, there are four approaches which can be performed; two of them are in more advanced steps due to the use of antisense oligonucleotides (ASOs) (Figure 3) or inhibitors against HDAC6 [45,51]. The other two are subject to advances in the knowledge of the 3D PACS-1 structure [7], as the proteolysis-targeting chimeras (PROTACs) or molecules, which specifically target the mutated protein.

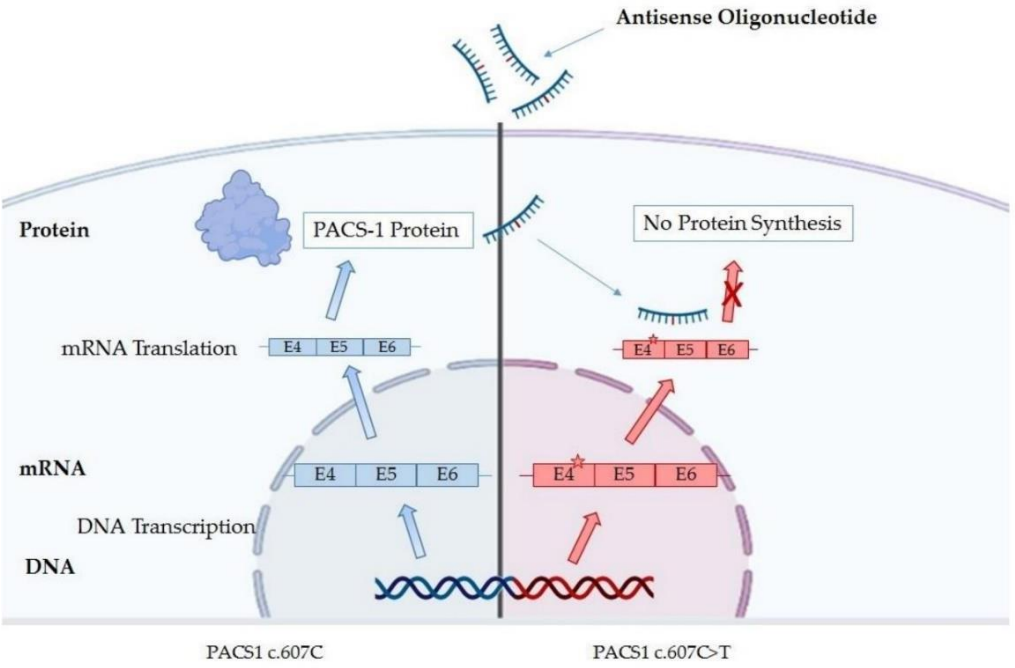


Figure 3. Schematic mechanism of antisense oligonucleotides (ASOs) therapy in order to avoid thetranslation of the mutant PACS-1 protein.

The antisense oligonucleotides (ASOs) can specifically target the mutated mRNA of*PACS1*, avoiding the translation of the pathologic protein [52] (Figure 3). The company IONIS is collaborating with the PACS1 foundation in order to develop an ASO targeting the p.R203W variant [51], and, nowadays, it is an active research field. However, ASOs are incapable of crossing the blood brain barrier (BBB) and require direct central nervous system (CNS) administration, so intrathecal delivery is recommended [53].

Other approaches propose the targeting of some of the client proteins of PACS-1, whose function is altered in these patients. In this sense, the research carried out by Dr. Thomas's group is interesting [45]. They suggest the specific targeting of HDAC6 throughits inhibition. They have delivered a patent application in which, in fibroblasts derived from patients, the stronger binding between the mutated PACS-1 and HDAC6 could be associated with a fragmented Golgi and a different microtubule network compared to the controls. They confirm that the use of general (TSA) or selective (tubacin, ACY-1215 or SW-100) HDAC6 inhibitors rescues the normal cellular phenotype [45].

Finally, two more approaches can be developed that focus on the specific targeting of the mutated PACS-1 protein. The proteolysis-targeting chimeras (PROTACs) bind simultaneously to the target protein and to an E3 ligase, forming a ternary complex, which promotes the ubiquitination of the protein of interest, thereby, inducing proteasomal degradation [54]. On the other hand, the use of molecules that specifically bind to the mutant PACS-1 protein is another way to produce protein degradation. However, the proper development of both technologies needs knowledge of the 3D structure and of the differences between the PACS-1 wild-type and mutated proteins [7].

6. Conclusions

In recent years, an extraordinary advance in the study of cellular protein traffickinghas been made. PACS-1 is a connector multifunctional protein with a role that trespasses the trafficking between endosomes and the TGN and could be key in cellular homeosta- sis, interacting with proteins related to apoptosis, genomic stability or calcium flux in the endoplasmic reticulum. Since 2012, when *PACS1*-NDD was first characterized, deep phenotyping analysis has allowed us to establish a potential relationship with other neu- rodevelopmental disorders, such as the PACS2 or Wdr37 syndromes. However, it is necessary to find a more profound approach that allows us to connect the physiological mechanisms of cellular trafficking and clinical features in order to reclassify and properlyunderstand these genetic disorders.

Exceptionally, most of the *PACS1*-NDD patients share the same mutation, and a gain- of-function or a dominant negative mechanism has been proposed. This makes it a greatmodel for the therapy of rare diseases caused by this mechanism. In this sense, strategies focused on decreasing the toxic effect of the mutated protein by inhibiting its expressionwith ASOs or its client proteins seem to be promising.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ijms23179649/s1. References [55–62] are cited in the supplementary materials.

Author Contributions: M.A. and Á.A., manuscript writing, collection and assembly of data; A.L.-P., C.L.-C., M.G.-S., A.A.-C., M.J.P., G.B.-L. and F.J.R., Literature review, collection and assembly of data; P.G.-P., structural studies; J.P. and B.P., conception and design, manuscript editing and final approval of the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

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

References

1. Schuurs-Hoeijmakers, J.H.; Oh, E.C.; Vissers, L.E.; Swinkels, M.E.; Gilissen, C.; Willemsen, M.A.; Holvoet, M.; Steehouwer, M.; Veltman, J.A.; de Vries, B.B.; et al. Recurrent De Novo Mutations in PACS1 Cause Defective Cranial-Neural-Crest Migration andDefine a Recognizable Intellectual-Disability Syndrome. *Am. J. Hum. Genet.* **2012**, *91*, 1122–1127. [CrossRef] [PubMed]
2. García-Cazorla, A.; Oyarzábal, A.; Saudubray, J.-M.; Martinelli, D.; Dionisi-Vici, C. Genetic disorders of cellular trafficking. *Trends Genet.* **2022**, *38*, 724–751. [CrossRef] [PubMed]
3. Schuurs-Hoeijmakers, J.H.M.; Landsverk, M.L.; Foulds, N.; Kukolich, M.K.; Gavrilova, R.H.; Greville-Heygate, S.; Hanson- Kahn, A.; Bernstein, J.A.; Glass, J.; Chitayat, D.; et al. Clinical delineation of thePACS1-related syndrome-Report on 19 patients. *Am. J. Med. Genet. Part A* **2016**, *170*, 670–675. [CrossRef] [PubMed]
4. Tenorio-Castaño, J.; Morte, B.; Nevado, J.; Martinez-Glez, V.; Santos-Simarro, F.; García-Miñaúr, S.; Palomares-Bralo, M.; Pacio- Míguez, M.; Gómez, B.; Arias, P.; et al. Schuurs–Hoeijmakers Syndrome (PACS1 Neurodevelopmental Disorder): Seven Novel Patients and a Review. *Genes* **2021**, *12*, 738. [CrossRef]
5. Miyake, N.; Ozasa, S.; Mabe, H.; Kimura, S.; Shiina, M.; Imagawa, E.; Miyatake, S.; Nakashima, M.; Mizuguchi, T.; Takata, A.; et al.A novel missense mutation affecting the same amino acid as the recurrent PACS1 mutation in Schuurs-Hoeijmakers syndrome.*Clin. Genet.* **2017**, *93*, 929–930. [CrossRef]
6. Liu, Y.; Ding, H.; Yan, T.; Liu, L.; Yu, L.; Huang, Y.; Li, F.; Zeng, Y.; Huang, W.; Zhang, Y.; et al. A Novel Multi-Exon Deletion of PACS1 in a Three-Generation Pedigree: Supplements to PACS1 Neurodevelopmental Disorder Spectrum. *Front. Genet.* **2021**, *12*, 690216. [CrossRef]
7. Youker, R.T.; Shinde, U.; Day, R.; Thomas, G. At the crossroads of homoeostasis and disease: Roles of the PACS proteins in membrane traffic and apoptosis. *Biochem. J.* **2009**, *421*, 1–15. [CrossRef]

8. Wan, L.; Molloy, S.S.; Thomas, L.; Liu, G.; Xiang, Y.; Rybak, S.L.; Thomas, G. PACS-1 Defines a Novel Gene Family of CytosolicSorting Proteins Required for trans-Golgi Network Localization. *Cell* **1998**, *94*, 205–216. [CrossRef]
9. Sieburth, D.; Ch'Ng, Q.; Dybbs, M.; Tavazoie, M.; Kennedy, S.; Wang, D.; Dupuy, D.; Rual, J.-F.; Hill, D.E.; Vidal, M.; et al.Systematic analysis of genes required for synapse structure and function. *Nature* **2005**, *436*, 510–517. [CrossRef]
10. Thomas, G.; Aslan, J.; Thomas, L.; Shinde, P.; Shinde, U.; Simmen, T. Caught in the act—Protein adaptation and the expanding roles of the PACS proteins in tissue homeostasis and disease. *J. Cell Sci.* **2017**, *130*, 1865–1876. [CrossRef]
11. Nair-Gill, E.; Bonora, M.; Zhong, X.; Liu, A.; Miranda, A.; Stewart, N.; Ludwig, S.; Russell, J.; Gallagher, T.; Pinton, P.; et al. Calcium flux control by PACS1-Wdr37 promotes lymphocyte quiescence and lymphoproliferative diseases. *EMBO J.* **2021**, *40*, e104888. [CrossRef]
12. Mani, C.; Tripathi, K.; Luan, S.; Clark, D.W.; Andrews, J.F.; Vindigni, A.; Thomas, G.; Palle, K. The multifunctional protein PACS-1 is required for HDAC2- and HDAC3-dependent chromatin maturation and genomic stability. *Oncogene* **2020**, *39*, 2583–2596. [CrossRef]
13. Lusk, L.; Smith, S.; Martin, C.; Taylor, C.; Chung, W. PACS1 Neurodevelopmental Disorder. In *GeneReviews®*; Adam, M.P., Mirzaa,G.M., Pagon, R.A., Wallace, S.E., Bean, L.J., Gripp, K.W., Amemiya, A., Eds.; University of Washington: Seattle, WA, USA, 1993. Available online: http://www.ncbi.nlm.nih.gov/books/NBK559434/ (accessed on 23 June 2022).
14. Stern, D.; Cho, M.; Chikarmane, R.; Willaert, R.; Retterer, K.; Kendall, F.; Deardorff, M.; Hopkins, S.; Bedoukian, E.; Slavotinek, A.; et al. Association of the missense variant p.Arg203Trp in PACS1 as a cause of intellectual disability and seizures.*Clin. Genet.* **2017**, *92*, 221–223. [CrossRef]
15. Martinez-Monseny, A.; Bolasell, M.; Arjona, C.; Martorell, L.; Yubero, D.; Arsmtrong, J.; Maynou, J.; Fernandez, G.; Salgado, M.D.C.; Palau, F.; et al. Mutation of PACS1: The milder end of the spectrum. *Clin. Dysmorphol.* **2018**, *27*, 148–150. [CrossRef]
16. Pefkianaki, M.; Schneider, A.; Capasso, J.E.; Wasserman, B.N.; Bardakjian, T.; Levin, A.V. Ocular manifestations of PACS1 mutation. *J. Am. Assoc. Pediatr. Ophthalmol. Strabismus* **2018**, *22*, 323–325. [CrossRef]
17. Dutta, A.K. Schuurs-Hoeijmakers syndrome in a patient from India. *Am. J. Med. Genet. Part A* **2019**, *179*, 522–524. [CrossRef]
18. Hoshino, Y.; Enokizono, T.; Imagawa, K.; Tanaka, R.; Suzuki, H.; Fukushima, H.; Arai, J.; Sumazaki, R.; Uehara, T.; Takenouchi, T.; et al. Schuurs-Hoeijmakers syndrome in two patients from Japan. *Am. J. Med. Genet. Part A* **2018**, *179*, 341–343.[CrossRef]
19. Seto, M.T.; Bertoli-Avella, A.M.; Cheung, K.W.; Chan, K.Y.; Yeung, K.S.; Fung, J.L.; Beetz, C.; Bauer, P.; Luk, H.M.; Lo, I.F.; et al. Prenatal and postnatal diagnosis of Schuurs-Hoeijmakers syndrome: Case series and review of the literature. *Am. J. Med. Genet. Part A* **2020**, *185*, 384–389. [CrossRef]
20. Hay, E.; Henderson, R.H.; Mansour, S.; Deshpande, C.; Jones, R.; Nutan, S.; Mankad, K.; Young, R.M.; Moosajee, M.; Genomics England Research Consortium; et al. Expanding the phenotypic spectrum consequent upon de novo *WDR37* missense variants. *Clin. Genet.* **2020**, *98*, 191–197. [CrossRef]
21. Kanca, O.; Andrews, J.C.; Lee, P.-T.; Patel, C.; Braddock, S.R.; Slavotinek, A.M.; Cohen, J.S.; Gubbels, C.S.; Aldinger, K.; Williams, J.; et al. De Novo Variants in WDR37 Are Associated with Epilepsy, Colobomas, Dysmorphism, Developmental Delay,Intellectual Disability, and Cerebellar Hypoplasia. *Am. J. Hum. Genet.* **2019**, *105*, 413–424. [CrossRef]
22. Reis, L.M.; Sorokina, E.A.; Thompson, S.; Muheisen, S.; Velinov, M.; Zamora, C.; Aylsworth, A.S.; Semina, E.V. De Novo Missense Variants in WDR37 Cause a Severe Multisystemic Syndrome. *Am. J. Hum. Genet.* **2019**, *105*, 425–433. [CrossRef]
23. Dentici, M.L.; Barresi, S.; Niceta, M.; Ciolfi, A.; Trivisano, M.; Bartuli, A.; Digilio, M.C.; Specchio, N.; Dallapiccola, B.; Tartaglia, M. Expanding the clinical spectrum associated with PACS2 mutations. *Clin. Genet.* **2019**, *95*, 525–531. [CrossRef]
24. Olson, H.E.; Jean-Marçais, N.; Yang, E.; Heron, D.; Tatton-Brown, K.; van der Zwaag, P.A.; Bijlsma, E.K.; Krock, B.L.; Backer, E.; Kamsteeg, E.-J.; et al. A Recurrent De Novo PACS2 Heterozygous Missense Variant Causes Neonatal-Onset Developmental Epileptic Encephalopathy, Facial Dysmorphism, and Cerebellar Dysgenesis. *Am. J. Hum. Genet.* **2018**, *102*, 995–1007. [CrossRef]
25. Terrone, G.; Marchese, F.; Vari, M.S.; Severino, M.; Madia, F.; Amadori, E.; Del Giudice, E.; Romano, A.; Gennaro, E.; Zara, F.; et al.A further contribution to the delineation of epileptic phenotype in PACS2-related syndrome. *Seizure* **2020**, *79*, 53–55. [CrossRef] [PubMed]
26. Lu, C.-L.; Kim, J. Craniofacial Diseases Caused by Defects in Intracellular Trafficking. *Genes* **2021**, *12*, 726. [CrossRef] [PubMed]
27. ClinVar. Available online: https://www.ncbi.nlm.nih.gov/clinvar/?term=PACS1%5Bgene%5D&redir=gene (accessed on 22 August 2022).
28. Lim, E.T.; Uddin, M.; De Rubeis, S.; Chan, Y.; Kamumbu, A.S.; Zhang, X.; D'Gama, A.M.; Kim, S.N.; Hill, R.S.; Goldberg, A.P.; et al. Rates, distribution and implications of postzygotic mosaic mutations in autism spectrum disorder. *Nat. Neurosci.* **2017**, *20*,1217–1224. [CrossRef] [PubMed]

29. DECIPHER. Available online: https://www.deciphergenomics.org/gene/PACS1/overview/clinical-info (accessed on 22 August 2022).
30. GTEx. Available online: https://gtexportal.org/home/gene/PACS1 (accessed on 22 August 2022).
31. Cardoso-Moreira, M.; Halbert, J.; Valloton, D.; Velten, B.; Chen, C.; Shao, Y.; Liechti, A.; Ascenção, K.; Rummel, C.; Ovchinnikova, S.; et al. Gene expression across mammalian organ development. *Nature* **2019**, *571*, 505–509. [CrossRef] [PubMed]
32. EvoDevo. Available online: https://apps.kaessmannlab.org/evodevoapp/ (accessed on 22 August 2022).
33. Ensembl. Available online: https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000175115;r=11:66070272-66244744 (accessed on 22 August 2022).
34. Brasacchio, D.; Alsop, A.E.; Noori, T.; Lufti, M.; Iyer, S.; Simpson, K.J.; Bird, P.I.; Kluck, R.M.; Johnstone, R.W.; Trapani, J.A. Epigenetic control of mitochondrial cell death through PACS1-mediated regulation of BAX/BAK oligomerization. *Cell Death Differ.* **2017**, *24*, 961–970. [CrossRef] [PubMed]
35. Veena, M.S.; Raychaudhuri, S.; Basak, S.K.; Venkatesan, N.; Kumar, P.; Biswas, R.; Chakrabarti, R.; Lu, J.; Su, T.; Gallagher-Jones, M.; et al. Dysregulation of hsa-miR-34a and hsa-miR-449a leads to overexpression of PACS-1 and loss of DNA damage response (DDR) in cervical cancer. *J. Biol. Chem.* **2020**, *295*, 17169–17186. [CrossRef]
36. He, C.; Su, C.; Zhang, W.; Wan, Q. miR-485-5p alleviates Alzheimer's disease progression by targeting PACS1. *Transl. Neurosci.* **2021**, *12*, 335–345. [CrossRef]
37. Trothen, S.M.; Zang, R.X.; Lurie, A.; Dikeakos, J.D. PACS-1 contains distinct motifs for nuclear-cytoplasmic transport and interacts with the RNA-binding protein PTBP1 in the nucleus and cytosol. *FEBS Lett.* **2021**, *596*, 232–248. [CrossRef]
38. Zheng, W.; Wuyun, Q.; Zhou, X.; Li, Y.; Freddolino, P.L.; Zhang, Y. LOMETS3: Integrating deep learning and profile alignment foradvanced protein template recognition and function annotation. *Nucleic Acids Res.* **2022**, *50*, W454–W464. [CrossRef]
39. Kelley, L.A.; Mezulis, S.; Yates, C.M.; Wass, M.N.; Sternberg, M.J.E. The Phyre2 web portal for protein modeling, prediction andanalysis. *Nat. Protoc.* **2015**, *10*, 845–858. [CrossRef]
40. Schmidt, V.; Sporbert, A.; Rohe, M.; Reimer, T.; Rehm, A.; Andersen, O.M.; Willnow, T.E. SorLA/LR11 Regulates Processing ofAmyloid Precursor Protein via Interaction with Adaptors GGA and PACS-1. *J. Biol. Chem.* **2007**, *282*, 32956–32964. [CrossRef]
41. Scott, G.K.; Fei, H.; Thomas, L.; Medigeshi, G.R.; Thomas, G. A PACS-1, GGA3 and CK2 complex regulates CI-MPR trafficking. *EMBO J.* **2006**, *25*, 4423–4435. [CrossRef]
42. Schermer, B.; Höpker, K.; Omran, H.; Ghenoiu, C.; Fliegauf, M.; Fekete, A.; Horvath, J.; Köttgen, M.; Hackl, M.; Zschiedrich, S.; et al. Phosphorylation by casein kinase 2 induces PACS-1 binding of nephrocystin and targeting to cilia.*EMBO J.* **2005**, *24*, 4415–4424. [CrossRef]
43. Jenkins, P.M.; Zhang, L.; Thomas, G.; Martens, J.R. PACS-1 Mediates Phosphorylation-Dependent Ciliary Trafficking of the Cyclic-Nucleotide-Gated Channel in Olfactory Sensory Neurons. *J. Neurosci.* **2009**, *29*, 10541–10551. [CrossRef]
44. Seidel, C.; Schnekenburger, M.; Dicato, M.; Diederich, M. Histone deacetylase 6 in health and disease. *Epigenomics* **2015**, *7*, 103–118. [CrossRef]
45. Thomas, G.; Thomas, L.; Villar-Pazos, S. Methods of Treating PACS1 and Pacs2 Syndromes. WO 2020/018647 A1. Available online: https://patents.google.com/patent/WO2020018647A1/en (accessed on 19 August 2022).
46. Brasacchio, D.; Noori, T.; House, C.; Brennan, A.J.; Simpson, K.; Susanto, O.; Bird, P.; Johnstone, R.W.; Trapani, J.A. A functional genomics screen identifies PCAF and ADA3 as regulators of human granzyme B-mediated apoptosis and Bid cleavage. *Cell Death Differ.* **2014**, *21*, 748–760. [CrossRef]
47. Brasacchio, D.; Busuttil, R.A.; Noori, T.; Johnstone, R.; Boussioutas, A.; Trapani, J.A. Down-regulation of a pro-apoptotic pathwayregulated by PCAF/ADA3 in early stage gastric cancer. *Cell Death Dis.* **2018**, *9*, 442. [CrossRef]
48. Rylaarsdam, L.; Guemez-Gamboa, A. A gain-of-function recurrent missense variant leads to a GABAergic/glutamatergic imbalance in a forebrain organoid model of PACS1 syndrome. *bioRxiv* **2022**. [CrossRef]
49. Van Nuland, A.; Reddy, T.; Quassem, F.; Vassalli, J.-D.; Berg, A.T. PACS1-Neurodevelopmental disorder: Clinical features andtrial readiness. *Orphanet J. Rare Dis.* **2021**, *16*, 386. [CrossRef]
50. Rylaarsdam, L.; Reddy, T.; Guemez-Gamboa, A. In search of a cure: PACS1 Research Foundation as a model of rare diseasetherapy development. *Trends Genet.* **2021**, *38*, 109–112. [CrossRef]
51. PACS1 Foundation. Available online: https://www.PACS1foundation.org/research (accessed on 18 July 2022).
52. Rinaldi, C.; Wood, M.J.A. Antisense oligonucleotides: The next frontier for treatment of neurological disorders. *Nat. Rev. Neurol.* **2017**, *14*, 9–21. [CrossRef]
53. Mendonça, M.C.P.; Kont, A.; Aburto, M.R.; Cryan, J.F.; O'Driscoll, C.M. Advances in the Design of (Nano)Formulations for Delivery of Antisense Oligonucleotides and Small Interfering RNA: Focus on the Central Nervous System. *Mol. Pharm.* **2021**, *18*,1491–1506. [CrossRef]
54. Deshaies, R. Prime time for PROTACs. *Nat. Chem. Biol.* **2015**, *11*, 634–635. [CrossRef]
55. Sun, M.; Zhang, H. Par3 and aPKC regulate BACE1 endosome-to-TGN trafficking through PACS1. *Neurobiol. Aging* **2017**, *60*, 129–140. [CrossRef]
56. Hinners, I.; Wendler, F.; Fei, H.; Thomas, L.; Thomas, G.; Tooze, S.A. AP-1 recruitment to VAMP4 is modulated by

phosphorylation- dependent binding of PACS-1. *EMBO Rep.* **2003**, *4*, 1182–1189. [CrossRef]
57. Piguet, V.; Wan, L.; Borel, C.; Mangasarian, A.; Demaurex, N.; Thomas, G.; Trono, D. HIV-1 Nef protein binds to the cellular protein PACS-1 to downregulate class I major histocompatibility complexes. *Nat. Cell Biol.* **2000**, *2*, 163–167. [CrossRef]
58. Liu, H.; Hu, P.-W.; Budhiraja, S.; Misra, A.; Couturier, J.; Lloyd, R.E.; Lewis, D.E.; Kimata, J.T.; Rice, A.P. PACS1 is an HIV-1 cofactor that functions in Rev-mediated nuclear export of viral RNA. *Virology* **2019**, *540*, 88–96. [CrossRef]
59. Chiu, Y.-F.; Sugden, B.; Chang, P.-J.; Chen, L.-W.; Lin, Y.-J.; Lan, Y.-C.; Lai, C.-H.; Liou, J.-Y.; Liu, S.-T.; Hung, C.-H. Characterization and Intracellular Trafficking of Epstein-Barr Virus BBLF1, a Protein Involved in Virion Maturation. *J. Virol.* **2012**, *86*, 9647–9655. [CrossRef] [PubMed]
60. Köttgen, M.; Benzing, T.; Simmen, T.; Tauber, R.; Buchholz, B.; Feliciangeli, S.; Huber, T.B.; Schermer, B.; Kramer-Zucker, A.; Höpker, K.; et al. Trafficking of TRPP2 by PACS proteins represents a novel mechanism of ion channel regulation. *EMBO J.* **2005**, *24*, 705–716. [CrossRef] [PubMed]
61. Chen, J.; Wang, J.; Meyers, K.R.; Enns, C.A. Transferrin-Directed Internalization and Cycling of Transferrin Receptor 2. *Traffic* **2009**, *10*, 1488–1501. [CrossRef] [PubMed]
62. Xiang, Y.; Molloy, S.S.; Thomas, L.; Thomas, G. The PC6B Cytoplasmic Domain Contains Two Acidic Clusters That Direct Sortingto distinct trans-Golgi Network/Endosomal Compartments. *Mol. Biol. Cell* **2000**, *11*, 1257–1273. [CrossRef] [PubMed

TRABAJO 2:

Things are not always what they seem: From Cornelia de Lange to KBG phenotype in a girl with genetic variants in NIPBL and ANKRD11

Molecular Genetics & Genomic Medicine
Open Access

Ana Latorre-Pellicer, Ángela Ascaso, **Cristina Lucia-Campos**, Marta Gil-Salvador, María Arnedo, Rebeca Antoñanzas, Ariadna Ayerza-Casas, Iñigo Marcos-Alcalde, Paulino Gómez-Puertas, Feliciano J Ramos, Juan Pié, Beatriz Puisac

Things are not always what they seem: From Cornelia de Lange to KBG phenotype in a girl with genetic variants in NIPBL and ANKRD11

Molecular Genetics & Genomic Medicine

Noviembre 2021

doi: 10.1002/mgg3.1826

PMID: 34617417

Received: 2 August 2021 | Revised: 9 September 2021 | Accepted: 16 September 2021

DOI: 10.1002/mgg3.1826

Molecular Genetics & Genomic Medicine Open Access WILEY

LETTER TO THE EDITOR

Things are not always what they seem: From Cornelia de Lange to KBG phenotype in a girl with genetic variants in *NIPBL* and *ANKRD11*

Dear Editor,

The diagnosis success rates for developmental disordershave greatly improved in the last years mainly due tothe widespread use of DNA next-generation sequencing. Nevertheless, several studies have stressed the importance of a critical reconsideration of genetic results and a further implementation of protocols for variant-level reevaluation and case-level reanalysis (Deignan et al., 2019). This is especially relevant in the context of syndromes, such as the chromatinopathies Cornelia de Lange syndrome (CdLS, OMIM#122470) and KBG syndrome (KBGS, OMIM #148050), with overlapping phenotypes that may evolveover time (Parenti et al., 2021). Here, we present a chal-lenging familiar case reanalyzed in which phenotypicfeatures of both KBGS and CdLS are observed, and where genetic variants in *ANKRD11* and *NIPBL* were identified.

Our case is a female, first child of a non-consanguineous couple, born at 37 weeks of gestation after an uneventful pregnancy. Birth length (47.2 cm, −0.55 SD), body weight (2.680 kg, −0.52 SD), and head circumference(34 cm, 0.26 SD) were all normal. She was referred toour hospital at the age of 3 years because of facial dysmorphism, gastroesophageal reflux, and motor developmental delay. After a comprehensive physical evaluationby our clinical geneticist, she was clinically diagnosed as CdLS with a clinical score of 10, mainly due to the synophrys (HP:0000664), thick eyebrows (HP:0000574), concave nasal ridge (HP:0011120), downturned cor-ners of mouth (HP:0002714), global developmental delay (HP:0001263), small hands (HP:0200055) and feet (HP:0001773), short fifth finger (HP:0009237), and hirsutism (HP:0001007)(Kline et al., 2018). At that time, molecular diagnosis was performed by sequencing the *NIPBL*gene and a potential disease-causing variant was identi- fied [*NIPBL*:NM_133433.3:c.7553A>G, p.(Asp2518Gly)].

The variant, that induces changes in the surface charge of the protein (Figure 1c), was classified as likely pathogenic (PM2, PP2, and PP3) according to the ACMG/AMP 2015 guidelines (Richards et al., 2015). Segregation studies revealed that the patient's mother (II.4), one aunt (II.2), and one uncle (II.1) carried the variant. The mother (II.4) was clinically evaluated and dysmorphic facial features fully consistent with her daughter's phenotype were observed. Facial photographs of individuals II.2, II.3, and I.2 were available and checked with the Face2Gene application (Latorre-Pellicer et al., 2020). A medium-low probability of CdLS was assigned for the aunt (II.2) that was the only one with the c.7553A>G *NIPBL* variant (Figure 1a). Altogether, clinical and molecular findings supported the CdLS diagnosis in the patient.

However, clinical follow-up of the proband revealed an evolution from CdLS to KBGS features. At the age of six, the patient showed an evident KBGS gestalt with macrodontia (HP:0001572), triangular face (HP:0000325), or bulbous nasal tip (HP:0000414). An additional clinical analysis was carried out with Face2Gene. At 3 years old, KBGS and CdLS were the first and second syndromes suggested, respectively, whereas at 6 years old, CdLS diagno- sis did not appear between the top-5 diagnosis provided by Face2Gene (Figure 1b). A molecular diagnosis reevalua-tion was performed by using a targeted gene panel includ- ing *ANKRD11*, in which a pathogenic nonsense variant was identified [*ANKRD11*:NM_001256183.1:c.2512C>T, p.(Arg838*)] (Figure 1d). This variant was maternally inherited, and neither the aunts (II.2 and II.3) nor the uncle (II.1) had it (Figure 1a). The variant was classifiedas pathogenic (PVS1, PM2, PP3, and PP5).

Considerable efforts have been made to standardize the interpretation of genetic variants in the laboratory (Latorre-Pellicer et al., 2020). However, once again, a clear example of the existing limitations is shown here, reinforcing the relevance of the implementation of protocols for peri- odic reevaluation of phenotypic and genetic information in clinical laboratories. Furthermore, this case highlights the clinical challenges in interpreting multiple pathogenic variants in single patients. An incorrect genetic diagnosis

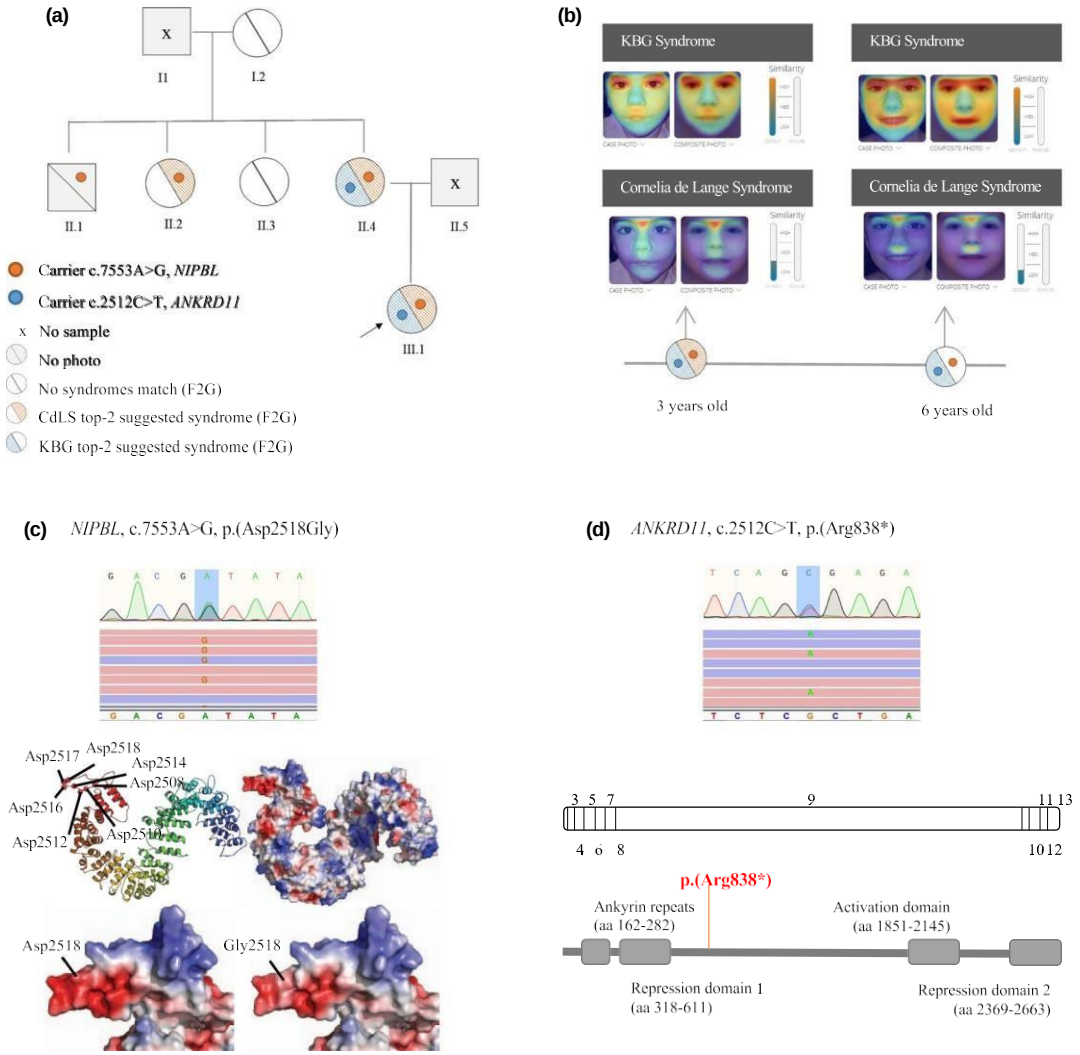


FIGURE 1 (a) Pedigree chart of the family showing Mendelian segregation of the heterozygous variants, c.7553A>G in *NIPBL* and c.2512C>T in *ANKRD11*, as well as the results of the Face2Gene (F2G) of the facial photographs of the individuals. (b) Face2Gene evaluation of facial photographs of the patient at 3 and 6 years of age, respectively. (c) Sanger chromatogram and Integrative Genomics Viewer (IGV) view of sequencing results of *NIPBL*: c.7553A>G in III.1, and structural modeling of NIPBL Asp2518Gly mutant. Top: predicted HEAT-repeat arrangement of NIPBL residues 1538 to 2544. Position of negatively charged amino acids Asp2508, Asp2510, Asp2512, Asp2514, Asp2516, Asp2517. and Asp2518 is indicated. Bottom: mutation of Asp2518 to Gly promotes a decrease in the negative charge on the surface of the local patch. (d) Sanger chromatogram and Integrative Genomics Viewer (IGV) view of sequencing results of *ANKRD11*: c.2512C>T in III.1. Localization of the *ANKRD11* variant at protein and DNA levels

can have severe consequences for prognostic and therapeutic management of the patient, and, as in this case, for the reproductive advice. Moreover, our findings seem to support recent evidences of age-dependent phenotypic evolution in individuals harboring *ANKRD11* variants (Parenti et al., 2021), and demonstrate the importance of including KBGS in the differential diagnosis of young children with CdLS features. Therefore, it is crucial to include *ANKRD11* in gene panels used for molecular testing of individuals presenting with clinical characteristics of CdLS.

FUNDING INFORMATION

European Social Fund, Grant/Award Number: B32_17R / B32_20R; Ministerio de Ciencia e Innovación; Instituto de Salud Carlos III, Grant/Award Number: PI19/01860

ACKNOWLEDGEMENTS

Spanish Ministry of Health (MH)-ISCIII (FIS) [Ref. PI19/01860, to F.J.R. and J.P.]; DGA- FEDER: European Social Fund [B32_17R / B32_20R, to J.P.]. MICIU-"Juan de la Cierva-Incorporación" (to A.L.P.).

CONFLICT OF INTEREST

The authors declare no competing interests.

AUTHOR CONTRIBUTIONS

A.L.-P., C.L.-C., M.G.-S., M.A., R.A., J.P., and B.P.: molecular analysis; A.A., A.A.-C., and F.J.R.: patients' recruitment, clinical evaluation, and clinical score calculation; I.M.-A. and P.G.-P.: bioinformatics studies and variants interpretation; A.L.-P., J.P., and B.P.: manuscript writing, collection, and assembly of data; F.J.R., J.P., and B.P.: manuscript editing and approval of the manuscript. All authors have read and agreed to the submitted version of the manuscript.

ETHICS APPROVAL

The employed procedure was reviewed and approved by the Ethics Committee of Clinical Research from the Government of Aragón (CEICA; PI15/00707). All human subjects participating in the research (or their legal guardians) signed the informed consent. An additional informed consent was collected and signed for the publication of subjects' photographs.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding authors upon reasonable request.

Ana Latorre-Pellicer[1]
Ángela Ascaso[1]
Cristina Lucia-Campos[1]
Marta Gil-Salvador[1]
María Arnedo[1]
Rebeca Antoñanzas[1]
Ariadna Ayerza-Casas[1,2]
Iñigo Marcos-Alcalde[3,4]
Paulino Gómez-Puertas[4]
Feliciano J. Ramos[5]
Juan Pié[1]
Beatriz Puisac[1]

[1]*Unit of Clinical Genetics and Functional Genomics, Department of Pharmacology-Physiology, School of Medicine, Universidad de Zaragoza, Zaragoza, Spain*
[2]*Unit of Paediatric Cardiology, Service of Paediatrics, Hospital Universitario Miguel Servet, Zaragoza, Spain*
[3]*Molecular Modeling Group, Centro de Biología Molecular Severo Ochoa, CBMSO (CSIC-UAM), Madrid, Spain*
[4]*Biosciences Research Institute, School of Experimental Sciences, Universidad Francisco de Vitoria, Madrid, Spain*
[5]*Unit of Clinical Genetics, Service of Paediatrics, Hospital Clínico Universitario Lozano Blesa, Department of Paediatrics, School of Medicine, Universidad de Zaragoza, Zaragoza, Spain*

Correspondence
Juan Pié and Beatriz Puisac, Unit of Clinical Genetics and Functional Genomics, Department of Pharmacology-Physiology, School of Medicine, Universidad de Zaragoza, CIBERER-GCV02 and IIS-Aragon, E-50009 Zaragoza, Spain.
Email: juanpie@unizar.es (J. P.); puisac@unizar.es (B. P.)

ORCID

Beatriz Puisac https://orcid.org/0000-0003-0170-7326

REFERENCES

Deignan, J. L. et al (2019). Points to consider in the reevaluation and reanalysis of genomic test results: A statement of the American College of Medical Genetics and Genomics (ACMG). *Genetics in Medicine*, *21*(6), 1267–1270.

Kline, A. D. et al (2018). Diagnosis and management of Cornelia de Lange syndrome: First international consensus statement. *Nature Reviews Genetics*, *19*(10), 649–666. https://doi.org/10.1038/s41576-018-0031-0.

Latorre-Pellicer, A. et al (2020). Evaluating Face2Gene as a tool to identify cornelia de lange syndrome by facial phenotypes. *International Journal of Molecular Sciences*, *21*(3), 1042. https://doi.org/10.3390/ijms21031042.

Parenti, I. et al (2021). *ANKRD11* variants: KBG syndrome and beyond. *Clinical Genetics*, *100*(2), 187–200.

Richards, S. et al (2015). 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. *Genetics in Medicine*, *17*(5), 405–424. https://doi.org/10.1038/gim.2015.30.

How to cite this article: Latorre-Pellicer, A., Ascaso, Á., Lucia-Campos, C., Gil-Salvador, M., Arnedo, M., Antoñanzas, R., Ayerza-Casas, A., Marcos-Alcalde, I., Gómez-Puertas, P., Ramos, F. J., Pié, J., & Puisac, B. (2021). Things are not always what they seem: From Cornelia de Lange to KBG phenotype in a girl with genetic variants in *NIPBL* and *ANKRD11*. *Molecular Genetics & Genomic Medicine*, 9, e1826. https://doi.org/10.1002/mgg3.1826

TRABAJO 3:

A Novel Intragenic Duplication in the HDAC8 Gene Underlying a Case of Cornelia de Lange Syndrome

Cristina Lucia-Campos, Irene Valenzuela, Ana Latorre-Pellicer, David Ros-Pardo, Marta Gil-Salvador, María Arnedo, Beatriz Puisac, Neus Castel, Alberto Plaja, Anna Tenes, Ivon Cuscó, Laura Trujillano, Feliciano J Ramos, Eduardo F Tizzano, Paulino Gómez-Puertas, Juan Pié

A Novel Intragenic Duplication in the HDAC8 Gene Underlying a Case of Cornelia de Lange Syndrome

Genes (Basel)

Agosto 2022

doi: 10.3390/genes13081413

PMID: 36011323

genes

MDPI

Article

A Novel Intragenic Duplication in the *HDAC8* Gene Underlying a Case of Cornelia de Lange Syndrome

Cristina Lucia-Campos [1], Irene Valenzuela [2,3], Ana Latorre-Pellicer [1], David Ros-Pardo [4], Marta Gil-Salvador [1], María Arnedo [1], Beatriz Puisac [1], Neus Castells [2,3], Alberto Plaja [2,3], Anna Tenes [2,3], Ivon Cuscó [2,3,5], Laura Trujillano [2,3], Feliciano J. Ramos [6], Eduardo F. Tizzano [2,3], Paulino Gómez-Puertas [4] and Juan Pié [1,*]

[1] Unit of Clinical Genetics and Functional Genomics, Department of Pharmacology-Physiology, School of Medicine, Universidad de Zaragoza, CIBERER-GCV02 and IIS-Aragon, E-50009 Zaragoza, Spain
[2] Department of Clinical and Molecular Genetics Hospital Vall d'Hebron, E-08035 Barcelona, Spain
[3] Medicine Genetics Group, Vall Hebron Research Institute, E-08035 Barcelona, Spain
[4] Centro de Biología Molecular Severo Ochoa, CBMSO (CSIC-UAM), Molecular Modeling Group, E-28049 Madrid, Spain
[5] Department of Genetics, Hospital Sant Pau, E-08041 Barcelona, Spain
[6] Unit of Clinical Genetics, Department of Paediatrics, Service of Paediatrics, Hospital Clínico Universitario Lozano Blesa, School of Medicine, Universidad de Zaragoza, CIBERER-GCV02 and IIS-Aragon, E-50009 Zaragoza, Spain
* Correspondence: juanpie@unizar.es; Tel.: +34-976-76167

check for updates

Citation: Lucia-Campos, C.; Valenzuela, I.; Latorre-Pellicer, A.; Ros-Pardo, D.; Gil-Salvador, M.; Arnedo, M.; Puisac, B.; Castells, N.; Plaja, A.; Tenes, A.; et al. A Novel Intragenic Duplication in the *HDAC8* Gene Underlying a Case of Cornelia de Lange Syndrome. *Genes* **2022**, *13*, 1413. https://doi.org/10.3390/genes13081413

Academic Editor: Stefania Zampatti

Received: 18 July 2022
Accepted: 6 August 2022
Published: 8 August 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Abstract: Cornelia de Lange syndrome (SCdL) is a multisystemic genetic disorder characterized by distinctive facial features, growth retardation, and intellectual disability, as well as various systemic conditions. It is caused by genetic variants in genes related to the cohesin complex. Single-nucleotide variations are the best-known genetic cause of SCdL; however, copy number variants (CNVs) clearly underlie a substantial proportion of cases of the syndrome. The *NIPBL* gene was thought to be the locus within which clinically relevant CNVs contributed to SCdL. However, in the last few years, pathogenic CNVs have been identified in other genes such as *HDAC8*, *RAD21*, and *SMC1A*. Here, we studied an affected girl presenting with a classic SCdL phenotype heterozygous for a de novo ~32 kbp intragenic duplication affecting exon 10 of *HDAC8*. Molecular analyses revealed an alteration in the physiological splicing that included a 96 bp insertion between exons 9 and 10 of the main transcript of *HDAC8*. The aberrant transcript was predicted to generate a truncated protein whose accessibility to the active center was restricted, showing reduced ease of substrate entry into the mutated enzyme. Lastly, we conclude that the duplication is responsible for the patient's phenotype, highlighting the contribution of CNVs as a molecular cause underlying SCdL.

Keywords: Cornelia de Lange syndrome; genetic disorder; copy number variants; *HDAC8*; intragenic duplication; array CGH; genetic diagnosis

1. Introduction

Cornelia de Lange syndrome (SCdL, OMIM #122470, #300590, #610759, #614701, #300882) is a genetically heterogeneous multisystemic disorder with an estimated prevalence of one in 10,000 to 30,000 live births. Since severity and clinical manifestations can vary widely among individuals, the SCdL phenotype has been defined as a spectrum that includes the so-called "classic" and "nonclassic" clinical presentations. The most clinically recognizable findings include distinctive facies with synophrys (HP:0000664), concave nasal bridge (HP:0011120), upturned nasal tip (HP:0000463), smooth philtrum (HP:0000319), and thin upper lip vermilion (HP:0000219), all of which are helpful in the diagnostic approach. Patients also commonly show intellectual disability (HP:0001249), prenatal and postnatal growth retardation (HP:0001511, HP:0008897), microcephaly (HP:0000252), limb reduction defects (HP:0001180, HP:0009237), and hirsutism (HP:0001007) [1].

From a molecular point of view, SCdL has been linked to eight genes involved in the structural or regulatory function of the cohesin complex. The most frequent causal gene is *NIPBL*, followed by *SMC1A* and *HDAC8*, and less frequently *SMC3*, *RAD21*, *BRD4*, *MAU2*, and *ANKRD11* [1]. Although a consistent link between the severity of the phenotype and the type of the genetic change has not been revealed so far, some clinical features have been associated more frequently with genetic variants in specific genes [2–5]. For example, patients with pathogenic variants in *HDAC8* have facial and clinical distinguishable features such as ocular hypertelorism (HP:0000316), delayed fontanelle closure (HP: 0000270), bulbous nasal tip (HP:0000414), hooding or redundant overfolded skin of the upper eyelids, dental anomalies, and mosaic skin pigmentation [6,7].

Currently, the widespread use of sequencing targeted panels, including causative and related SCdL genes, has significantly improved the diagnosis success rate, as well as reducing the time to achieve it [8]. However, although the causal variant of a SCdL case involves only one of the related genes, genetic diagnosis may still be challenging due to difficulties in interpretation such as allele frequency or even mosaicism, which appear to be quite recurrent in SCdL [9]. Furthermore, the genetic variant type can range from single-nucleotide variants (SNVs) to small insertions and deletions (INDELs) or copy number variants (CNVs). In fact, the presence of pathogenic CNVs in *NIPBL* may account for up to 3% of SCdL cases [10]. Therefore, the international guidelines recommend multi- plex ligation-dependent probe amplification (MLPA) approaches when panels and Sanger sequencing cannot detect any variant in this gene [1], but commercial MLPA assays only cover the *NIPBL* gene. Since next-generation sequencing panels have been implemented, structural variants involving other causal genes such as *SMC1A* [11], *HDAC8* [12], and *RAD21* [13] have been described in some individuals with SCdL. These cases indicate that pathogenic CNVs in SCdL-related genes may be more common than previously thought.

In this study, we report for the first time a molecular functional study of an intragenic duplication in the *HDAC8* gene identified in a girl with classic SCdL phenotype. We present genotype data and assess the pathogenicity of the intragenic variant through a combination of clinical phenotype evaluation, array CGH with exonic coverage of several genes involved with SCdL, splicing analysis, and structural prediction with protein modeling.

2. Materials and Methods

2.1. Clinical Diagnosis

The study was performed according to the Declaration of Helsinki protocols and was approved by each Regional Ethics Committee of Clinical Research. Informed consent was obtained from parents or guardians of all individuals included in this study. Additional in- formed consent was collected for the publication of photographs of the patient. Clinical data were collected by a clinical geneticist at the Vall d'Hebron Hospital (Barcelona) following a standard restricted-term questionnaire. The clinical score was calculated by SCdL clinical geneticist specialists according to the international consensus guidelines [1]. Face2Gene (https://www.face2gene.com (accessed on 10 January 2022)) was used to determine the most probable clinical diagnoses for the patient [14].

2.2. Isolation of DNA and RNA

Sequencing analyses were carried out on DNA from the patient and her progenitors. The DNA was isolated from non-cultivated blood samples using a Gentra® Puregene® Kit (Qiagen, Hilden, Germany) following the recommendations of the manufacturer.

Total RNA was extracted from 10 mL of peripheral blood lymphocytes using Trisure reagent (Sigma-Aldrich, St. Louis, MO, USA) following the manufacturer's

protocol. RNA was cleaned up using an RNeasy Mini Kit (Qiagen) with an additional step of DNase digestion using an RNase-Free DNase Set (Qiagen). Purity and integrity of the RNA were assessed by electrophoresis and spectrophotometry using a NanoDrop ND-2000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA).

2.3. Next-Generation Sequencing

The patient's DNA was analyzed on a panel of gene amplicons specifically designed for SCdL in the Clinical Genetics and Functional Genomics Group at the University of Zaragoza, as previously described [9]. The variants were classified according to the ACMG recommendations and detailed information provided in the public databases gnomAD (https://gnomad.broadinstitute.org/ (accessed on 7 March 2022)), ClinVar (https://www. ncbi.nlm.nih.gov/clinvar/ (accessed on 7 March 2022)), dbSNP (https://www.ncbi.nlm. nih.gov/snp/ (accessed on 7 March 2022)), LOVD (https://www.lovd.nl/ (accessed on 7 March 2022)), and relevant scientific literature. The in silico analyses were performed using the following online tools: Polyphen-2 (http://genetics.bwh.harvard.edu/pph2/ (accessed on 7 March 2022)), SIFT (https://sif.bii.a-star.edu.sg/ (accessed on 7 March 2022)), and the integration tool VarSome (https://varsome.com/ (accessed on 7 March 2022).

2.4. Array Comparative Genomic Hybridization

DNA extracted from an uncultured blood sample was analyzed with CytoSure Con- stitutional 8× 60K v3 (Oxfordshire, UK) array comparative genomic hybridization (array CGH) following the recommendations of the manufacturer. CytoSure Constitutional 8× 60K v3 has exonic resolution in 354 genes selected by the ClinGen Dosage Sensitiv- ity Map192, including *SMC1A*, *HDAC8*, *RAD21*, and *ANKRD11*. CNVs were classified following recommendations of the American College of Medical Genetics and Genomics standards [15] and reevaluated with actualized guidelines [16].

2.5. cDNA Synthesis and Analysis

RNA isolated from blood lymphocytes was reverse-transcribed using random hex- amer primers with an Invitrogen SuperScript™ First-Strand Synthesis System Kit for qPCR. For the analysis of physiological splicing of *HDAC8* (ENST00000373573.9, NM_018486.3), the cDNA was amplified with different pairs of primers using a Thermo Scientific™ Dream- Taq PCR Master Mix (2×) Kit in Applied Biosystems equipment. Primers were designed using the Primer3Plus in silico tool (https://www.bioinformatics.nl/cgi-bin/primer3plus/ primer3plus.cgi (accessed on 7 March 2022)) and checked using the UCSC in silico PCR tool (https://genome.ucsc.edu/cgi-bin/hgPcr (accessed on 7 March 2022)). The reverse primer spanned the junction between exons 10 and 11, and it was the same for all PCRs with the sequence 5'GCTTCAGATTCCCTTTGATGTAG 3' (Reverse 1). The forward primer was dif- ferent for each PCR: Forward 1 bound within exons 1 and 2, 5'CAAACGGGCCAGTATGGT G 3'; Forward 2 bound within exons 7 and 8, 5'GATTTTTCCCAGGAACAGGT 3'; Forward 3 hybridized with exon 9, 5'GAGGCTATAACCTTGCCAAC 3'.

PCR products were purified by NZYtech NZYGelpure Kit and screened by Sanger sequencing on ABI3730xl Capillary Electrophoresis Sequencing System according to the manufacturer's protocol.

2.6. Real-Time Quantitative PCR (qPCR)

Relative quantification of *HDAC8* expression was performed by qPCR. In this experiment, we used six cDNA samples as control to compare with *HDAC8* expression. After cDNA synthesis, qPCR amplification was carried out using the

Applied Biosystems SYBR™ Green PCR Master Mix Kit on an Applied Biosystems™ QuantStudio™ 5 System. We used the following amplification conditions: 95 °C for 10 min for the hold stage; 40 cycles at 95 °C for 15 s, 60 °C for 1 min; finally, 95 °C for 15 s, 60 °C for 1 min, and 95 °C for 20 s for the melt curve stage. Samples were assessed in triplicate.

The primer sequences were as follows: Forward 8–9 5'TTGGGAGGAGGAGGCTATA AC 3' and Reverse 9–10 5'GCTGTGAAAAACTCATGATCTGG 3'; Forward 9-Insertion 5'CCAGATCATGAGAATATGCCTG 3' and Reverse Insertion-10 5'CTGTGAAAAACTGCACATCACA 3'; Forward Exon 1 5' CGCTGGTCCCGGTTTATATC 3' and Reverse Exon 2 5' TGCAGTGCATATGCTTCAATC 3'. Gene expression levels were calculated normal-lizing with respect to the housekeeping gene β-actin, using the forward β-actin 5' CTTC- CTGGGCATGGAGTC 3' and reverse β-actin 5' AGCACTGTGTTGGCGTACAG 3' primers. The Ct values for each sample were determined with amplification plots in the loga- rithmic phase. The PCR outcome and efficiency of amplification were determined using QuantStudio™ Design and Analysis Software (v1.5.1, Applied Biosystems, Waltham, MA, USA) using the $2^{-\Delta\Delta C}$t method. GraphPad Prism was used for the graphics.

2.7. Structure Modeling of HDAC8 Variant and Molecular Dynamics Simulation

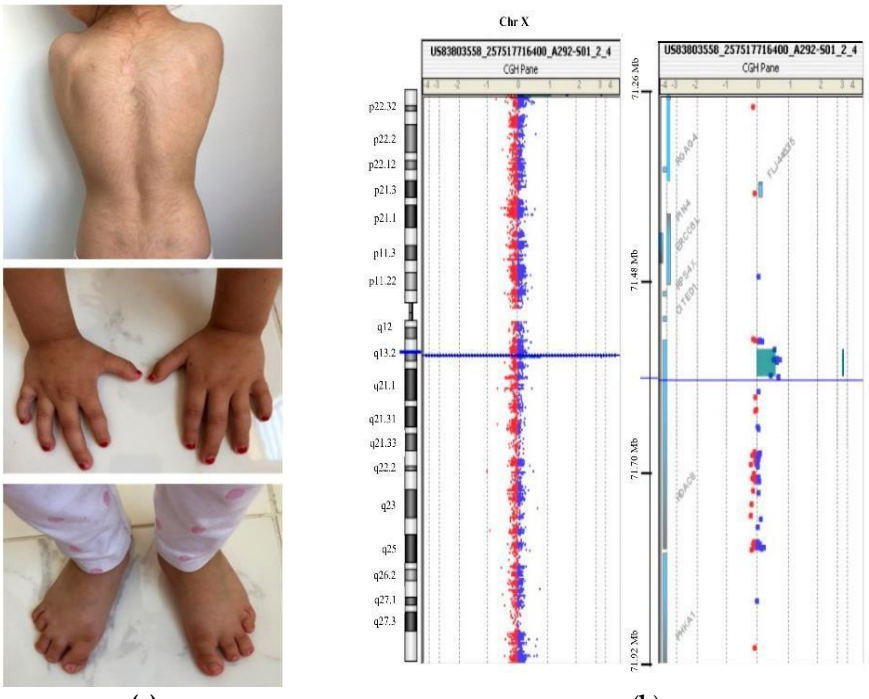
The 3D structure of the *HDAC8* variant protein was modeled using the crystal struc- ture of human *HDAC8* (PDB ID: 1T64 [17]) as a template. The model was built using the SWISS-MODEL server (http://swissmodel.expasy.org (accessed on 5 April 2022)), its structural quality being within the range accepted for homology-based structures (Anolea/Gromos/QMEAN4). Structures for wildtype and variant *HDAC8* proteins were subjected to 200 ns of unrestrained molecular dynamics (MD) simulation using the AM- BER18 molecular dynamics package (http://ambermd.org/ (accessed on 5 April 2022); University of California—San Francisco, CA, USA), essentially as previously described [18]. In brief, 3D models were first solvated with a periodic octahedral pre-equilibrated solvent box using the LEaP module of AMBER, with 12 Å as the shortest distance between any atom in the protein domain and the periodic box boundaries. Free MD simulation was performed using the PMEMD program of AMBER18 and the ff14SB force field (http://ambermd.org/ (accessed on 5 April 2022)), applying the SHAKE algorithm, a time step of 2 fs, and a nonbonded cutoff of 12 Å. Systems were initially relaxed over 10,000 steps of energy min- imization, using 1000 steps of steepest descent minimization followed by 9000 steps of conjugate-gradient minimization. Simulations were then started with a 20 ps heating phase, raising the temperature from 0 to 300 K in 10 temperature change steps, after each of which velocities were reassigned. During minimization and heating, the Cα trace dihedrals were restrained with a force constant of 500 $kcal \cdot mol^{-1} \cdot rad^{-2}$ and gradually released into an equilibration phase in which the force constant was progressively reduced to 0 over 200 ps. After the equilibration phase, 200 ns of unrestricted MD simulation was obtained for the structures. MD trajectories were analyzed using VMD software (v1.9.3., University of Illinois, Urbana, USA) [19]. Figures were generated using the Pymol Molecular Graph- ics System (https://pymol.org/ (accessed on 5 April 2022); Schrödinger, LLC, Portland, OR, USA).

3. Results

3.1. Clinical Report

The proband is a 5 year old girl, the first child of nonconsanguineous healthy parents. She has a healthy younger sibling. During pregnancy, intrauterine growth

restriction (IUGR) was revealed in the second trimester. The patient was born at 41 weeks gestational age via spontaneous vaginal delivery (SVD). At that moment, it was noted that she showed congenital microcephaly (HP:0000252) (head circumference 32.0 cm; −2.07 SD) and sym- metrical IUGR (HP:0001511) with a birthweight of 2.230 kg (−2.79 SD) and length of 45 cm (−3.22 SD). Growth retardation (HP:0008897) and microcephaly (HP: 0000252) persisted after birth, and, at the age of 22 months, head circumference was 43.5 cm (−3.62 SD), weight was 9.5 kg (−1.92 SD), and length was 76.5 cm (−2.81 SD). General motor development was not significantly delayed, and she achieved sitting and walking independently at 9 and 16 months, respectively. However, at the present age of 5 years, she is currently still nonverbal and shows global developmental delay (HP:0001263) and behavioral problems. Regarding dysmorphic facial features, she showed bulbous nasal tip (HP:0000414), long philtrum (HP:0000319), synophrys (HP:0000664), microdontia (HP:0000691), and widely spaced maxillary central incisors (HP:0001566). Furthermore, she presented delayed clo- sure of fontanels (more than 2 years) (HP: 0000270), hirsutism (HP:001007), small hands (HP:0200055) and feet (HP:0001773), short fifth finger (HP:0009237), clinodactyly of the fifth finger (HP:0004209), sensorineural hearing impairment (HP:0000407), and gastroesophageal reflux (HP:0002020) (Figure 1A). On the basis of these features, an expert clinical geneticist assigned the clinical diagnosis of SCdL during early childhood with a clinical score of 12. An additional clinical analysis was carried out with Face2Gene® at the age of 5, and KBGS and SCdL were the first and second syndromes suggested, respectively, with a medium–high probability.



(a) **(b)**

Figure 1. Clinical and genetic description of the patient. (**a**) Patient at 5 years showing her back with hirsutism and small hands and feet with clinodactyly of the fifth finger (more details in the text); (**b**) CytoSure Constitutional v3 array 8 × 60K array-CGH result showing a 0.032 Mb duplication at Xq13.1 in the patient.

3.2. DNA Molecular Analyses

An initial genetic test with a SCdL deep targeted gene panel (>1000×) did not de- tect any potentially constitutive and/or mosaic causative genetic variant in DNA from blood. Somatic mosaic variants could not be ruled out totally since it was not possible to obtain another biological sample from the patient. A variant of

uncertain significance was identified in the ANKRD11 gene ((NM_001256183.1; c.890C > T, p.(Thr297Met)) (G = 1019, A = 976). However, although some in silico predictors such as SIFT (0.0) and Polyphen-2 (0.999) suggested a possible damaging effect, this variant is reported as likely benign in ClinVar, and the allele frequency in gnomAD is greater than 0.001. Furthermore, familiar co-segregation studies revealed that the healthy mother carried the variant. Therefore, despite the compatible genotype–phenotype correlation in the patient, this variant was reclassified as likely benign according to ACMG criteria.

Despite the compatible genotype–phenotype correlation, familiar segregation studies revealed that the healthy mother carried the variant; therefore, the causality of this one was ruled out. Oligonucleotide-based array CGH was subsequently performed on genomic DNA using CytoSure Constitutional v3 array 8 × 60 K, which offers enhanced exon- level coverage of 354 developmental disorder genes. A duplication spanning ~32 kbp at Xq13.1 was identified in the patient (arr(GRCh38) Xq13.1(72,340,096–72,389,392) × 3) (Figures 1B and 2A). This duplication implicates at least exon 10 (NM_018486.3) of the *HDAC8* gene and was not present in leucocyte-derived DNA of the parents.

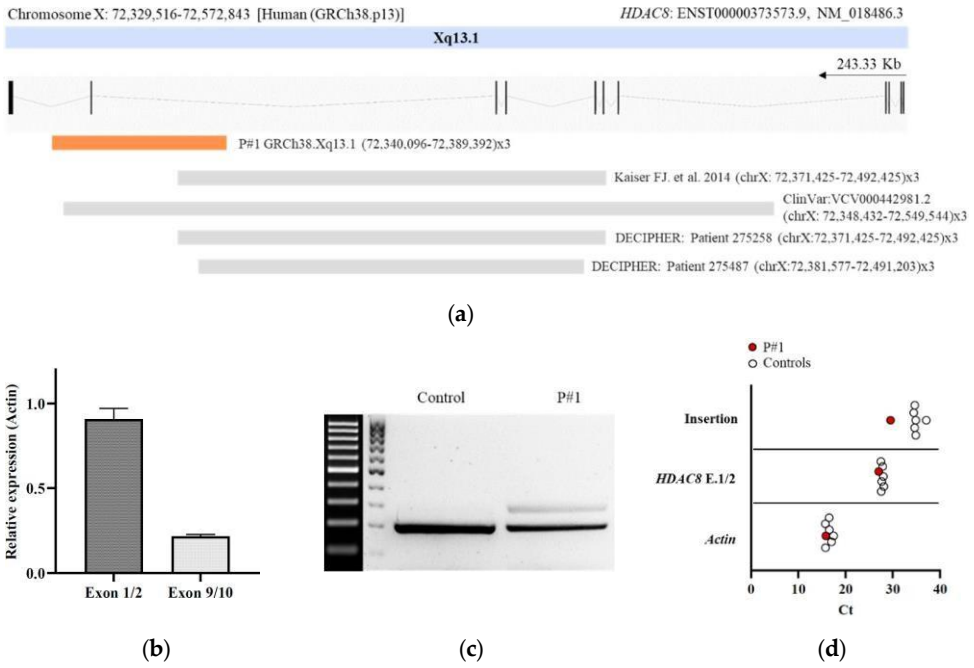


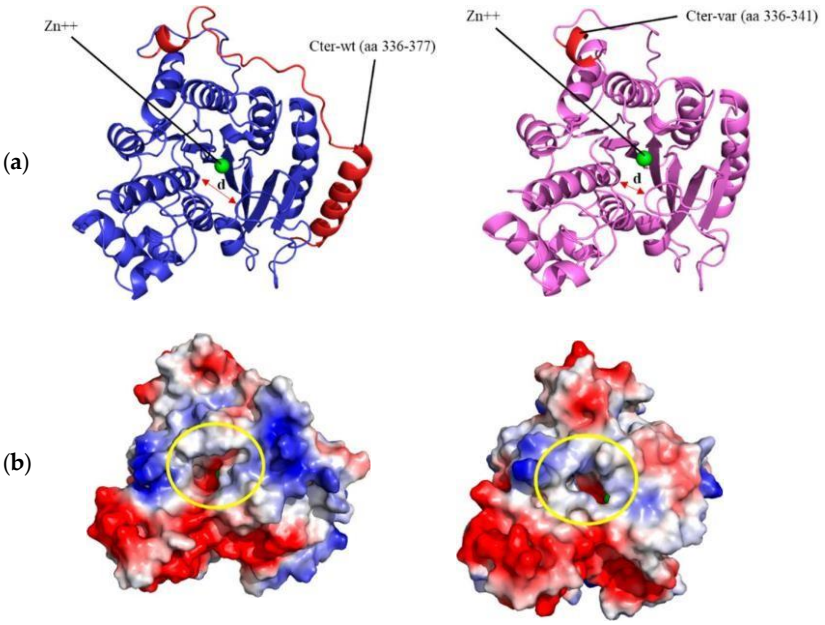
Figure 2. Molecular characterization of the intragenic duplication in *HDAC8*. (**a**) Schematic rep- resentation of intragenic duplications on *HDAC8* related to SCdL reported to date. The genomic region duplicated for each case is shown. Localizations of duplications on *HDAC8* are indicated by chromosome band and position (human (GRCh38.p13)). The arrow indicates the direction of transcription. The region duplicated in the patient is marked by an orange-filled box [6]; (**b**) qPCR analysis of exons 1–2 and 9–10 of *HDAC8* (NM_018486.3). Gene expression levels were normalized to actin. The expression level in controls was arbitrarily set to 1.0; (**c**) agarose gel of the cDNA PCR products. After amplification of a fragment compromising exons 9 and 10, cDNA of the patient yielded the expected PCR product of 213 bp, as well as an aberrant fragment of 309 bp corresponding to an insertion of 96 bp between exons 9 and 10; (**d**) Cts from qPCR analysis of exons 1–2 and the Δ96 bp region in the patient and controls. Gene expression levels were normalized to *Actin*.

3.3. RNA Molecular Analyses

To assess the functional impact of the intragenic duplication on *HDAC8* transcription, we performed qPCR and Sanger sequencing from blood cDNA. Specific quantitative amplification of the pair spanning exons 1–2 and 9–10 revealed a significant reduction in exon junction 9–10 in the patient (Figure 2B). We also performed a conventional PCR amplification from exon 9 to exon junction 10–11. An aberrant transcript of 309 bp in addition to the expected wildtype PCR product of 213 bp was observed. Sequencing analysis of the aberrantly spliced product revealed an insertion of a 96 bp fragment between exons 9 and 10 (Figure 2C and Table S1). This sequence aligns with an intronic region located between 225,627 bp and 225,722 bp of intron 9 of *HDAC8*, GRCh38.p13 chrX: 72,352,477–72,352,382. In addition, we confirmed the presence of the 96 bp insertion in the patient by qPCR (Figure 2D).

3.4. Structural Prediction of HDAC8 Variant

A theoretical 3D structure for the *HDAC8* variant was obtained by homology modeling. As a result of the 96 bp insertion in the nucleotide sequence, at the structural level, amino acids 336 to 377 of the C-terminus of the protein were replaced, including a complete α helix (in red in Figure 3A, left) by a much shorter segment of six amino acids (in red in Figure 3A, right), followed by a stop codon. To analyze the effect that this C-terminal deletion could have on the structure and function of the enzyme, the wildtype and variant protein models were simulated for 200 ns of unrestricted molecular dynamics, comparing the behavior of the two structures. The main differences between them were, on the one hand, a marked change the surface electrostatic charge around the entrance of the active center, from being mostly electropositive in the wildtype protein (blue-colored patch in Figure 3B, left) to neutral in the variant protein (light-blue and white colors in the equivalent position in Figure 3B, right). In addition to this shift in surface charge, a notable difference observed between the two was a displacement in the variant protein, starting at approximately 140 ns, of the loop containing amino acids Gly206 to Gly220, located at the entrance of the active center, shortening the distance between this loop and the opposite wall of the substrate entry site (distance marked with an arrow in Figure 3A, left and right). As a quantification of this movement, Figure 3C shows the variation in the distance between the α carbon of the Pro209 and Gly151 residues, located in the shifting loop and in a loop on the opposite wall of the entry site, respectively, decreasing from an average of 8.5 Å in the wildtype protein to 5.5 Å in the variant *HDAC8* protein.



(**c**)

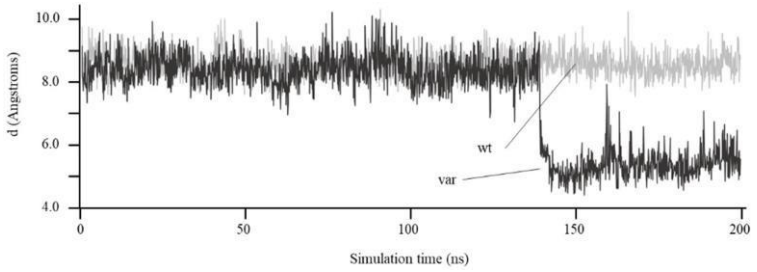


Figure 3. Structural modeling of *HDAC8* variant. (**a**) Structural model of wildtype (**left**) and variant (**right**) human *HDAC8* proteins after 200 ns of unrestricted molecular dynamics simulation. The dif- ferential C-terminal end between the two proteins is colored in red. The position of the Zn^{++} atom in the active center and the distance between residues Pro209 and Gly151 (arrow) are indicated; (**b**) surface of both proteins after 200 ns of molecular dynamics. The entrance of the active center is circled in yellow. The surface is colored according to the electrostatic charge (red: negative, blue: positive); (**c**) distance in Å between the α carbon atoms of residues Pro209 and Gly151 over the 200 ns simulation of the wildtype (wt) and variant (var) *HDAC8* protein.

4. Discussion

Identification of the genetic cause sometimes remains a challenge in SCdL. Although this is a phenotypically recognizable syndrome, the huge clinical and genetic heterogeneity makes it difficult to establish genotype–phenotype correlations. Here, we described the genetic diagnosis procedure of a girl with classic SCdL phenotype and highlighted the limitations faced during this process: (i) interpretation of VUS, (ii) technical difficulty in detecting small CNVs and somatic mosaic variants, and (iii) functional interpretation of intragenic duplications. In this study, we presented a novel intragenic ~ 32 kbp duplication affecting exon 10 of *HDAC8* gene. SCdL-associated variants in the *HDAC8* gene (Xq13.1) are estimated to account for the 4% of the cases of this disorder (SCDL5, OMIM #300882). The majority of known disease-causing changes in *HDAC8* are SNVs, including non- sense, missense, or splice site variants [6,7,20–25]. Nevertheless, recently, several cases of changes involving larger regions of *HDAC8* have been reported in individuals with SCdL, especially intragenic deletions ranging from single to multiple exons [6,12]. Interestingly, the presence of two to three pairs of microhomology at the breakpoints was found in these cases [12]. To date, only one intragenic duplication in *HDAC8* has been reported in the literature (GRCh38.p13 (chrX: 72,371,425–72,731,334) × 3), whose region affected includes exons 6 to 9 (p.Phe336Leufs*1) (6). In addition, ClinVar reports one likely pathogenic duplication (ClinVar: VCV000442981.2, GRCh38.p13 (chrX: 72,348,432–72,549,544) × 3), and DECIPHER reports two duplications associated with patients with overlapping SCdL phenotypes (Patient 275,487, Patient 275,258) (Figure 2A). However, to our knowledge, no report of the effect of an intragenic duplication in *HDAC8* at a molecular level has been provided.

Haploinsufficiency for genes within a deletion CNV is a well-recognized cause of genetic disease. However, duplication CNVs might cause disease through triplosensitivity, gene disruption, or gene fusion at breakpoints. Undoubtedly, interpreting the genetic consequences of the duplication is essential to understand the etiology of the genetic disease. In this report, we presented the case of an intragenic duplication in *HDAC8* that disrupted the reading frame of at least the most common *HDAC8* gene isoform present in blood. By conventional splicing analysis and qPCR, we detected an aberrant *HDAC8* transcript in the patient. Surprisingly, this transcript included an insertion between

exons 9 and 10, corresponding to a 96 bp fragment of intron 9. By analyzing this intronic sequence, we found that it contained noncanonical splice site sequences, AG donor (5′) and GC acceptor (3′) [26]. As a consequence, the translation was impaired, resulting in a protein of 342 amino acids instead of the 377 comprising the wildtype protein (Table S2 and Table S3) due to a break in the reading frame caused by the insertion of a premature stop codon.

Although the direct correlation of clinical severity with the activity of *HDAC8* variants is complex, the structure–function relationship has been previously proven. Depending on the structural location of the different variants studied in this enzyme, the degree of loss of enzyme activity is variable, being greater when the catalytic center is affected or when the mutations affect an amino acid located in a conserved region of the protein [27–30]. All the mutations previously analyzed in *HDAC8* were randomly distributed throughout the entire protein structure, and, in the vast majority of them, the enzyme activity was compromised [6,27–31]. The de novo intragenic duplication presented in the current studio involved only exon 10, which, at the structural level, mainly affected an α helix at the C-terminal end of the protein, far from the active center.

With such a large deletion within the carboxyl-terminal end, a drastic effect was expected in the molecular dynamics simulation, entailing a large disorganization of the protein. However, it seems striking that the overall structure of the wildtype and variant proteins remained almost constant throughout 200 ns of simulation (Figure 3A). Despite this, notable differences were observed between them, such as a change in the surface charge next to the entry of the active center in the variant protein (Figure 3B), which would compromise the correct binding between the mutated enzyme and the substrate. This alteration would also trigger a displacement of a loop near the entrance of the catalytic site, which remains stable at the new position (Figure 3C) and results in reduced accessibility of the substrate to the active center (Figure 3B). Experimental *HDAC8* activity measurements would be necessary to verify this hypothesis.

In summary, the progressive increase in reports of patients with SCdL and intragenic CNVs makes it necessary to cover a wider genetic diagnosis scenario. For this purpose, as NGS becomes routine for genetic testing, we must incorporate specific designs and pipelines that analyze and detect this type of variant, not only in the *NIPBL* gene but also in the other causal genes of SCdL such as *HDAC8*. Lastly, a comprehensive molecular characterization of the intragenic duplications is essential for the investigation of their pathogenicity, with the ultimate goal of providing a better understanding of the disease.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/genes13081413/s1: Table S1. Insertion sequence 96 pb; Table S2. WT protein sequence; Table S3. Variant protein sequence.

Author Contributions: Molecular analysis, C.L.-C., A.L.-P., M.G.-S., M.A., B.P. and J.P.; genetic analysis, I.V., N.C., A.P., A.T., I.C., L.T. and E.F.T.; clinical diagnosis and clinical score calculation F.J.R., I.V., N.C., A.P., A.T., I.C., L.T. and E.F.T.; protein modeling, D.R.-P. and P.G.-P.; writing—original draft preparation, C.L.-C., A.L.-P. and J.P.; writing—review and editing C.L.-C., A.L.-P., J.P., E.F.T., I.V., A.P. and P.G.-P. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Spanish Ministry of Health-ISCIII Fondo de Investigación Sanitaria (FIS) (Ref. PI19/01860, to F.J.R. and J.P.) and Diputación General de Aragón-FEDER: European Social Fund (Grupo de Referencia B32_17R/B32_20R, to J.P.). A.L.-P. is supported by a "Juan de la Cierva-Incorporación" postdoctoral grant from MICIU (Spanish Ministry of Science and Universities), M.G.-S. is supported by a

Predoctoral Fellowship from the Diputación General de Aragón, and C.L.-C. is supported by a Predoctoral Fellowship from the MH-ISCIII. This work was also supported by Spanish government grants RTI2018-094434-B-I00 (MCIU/AEI/FEDER, UE) and DTS20-00024 (ISCIII) to P.G.-P., as well as funds from the European JPIAMR network "EPIC-Alliance" to P.G.-P. The computational support of the "Centro de Computación Científica CCC-UAM" is gratefully recognized. This work was also partially supported by Spanish Instituto de Salud Carlos III, Fondo de Investigaciones Sanitarias co-funded with ERDF funds, Grant No. FIS PI20/01767) to A.P. and by Spanish Instituto de Salud Carlos III, Fondo de Investigaciones Sanitarias co-funded with ERDF funds, Grant No. FIS PI18/000687 to E.F.T.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of Clinical Research from the Government of Aragón (CEICA; PI15/00707).

Informed Consent Statement: Informed consent was obtained from parents or guardians of all individuals included in this study. Additional informed consent was collected for the publication of photographs of the patient.

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

References

1. Kline, A.D.; Moss, J.F.; Selicorni, A.; Bisgaard, A.M.; Deardorff, M.A.; Gillett, P.M.; Ishman, S.L.; Kerr, L.M.; Levin, A.V.; Mulder, P.A.; et al. Diagnosis and management of Cornelia de Lange syndrome: First international consensus statement. *Nat. Rev. Genet.* **2018**, *19*, 649–666. [CrossRef] [PubMed]
2. Mannini, L.; Cucco, F.; Quarantotti, V.; Krantz, I.D.; Musio, A. Mutation Spectrum and Genotype-Phenotype Correlation in Cornelia de Lange Syndrome. *Hum. Mutat.* **2013**, *34*, 1589–1596. [CrossRef] [PubMed]
3. Krab, L.C.; Marcos-Alcalde, I.; Assaf, M.; Balasubramanian, M.; Andersen, J.B.; Bisgaard, A.-M.; Fitzpatrick, D.R.; Gudmundsson, S.; Huisman, S.A.; Kalayci, T.; et al. Delineation of phenotypes and genotypes related to cohesin structural protein RAD21. *Hum. Genet.* **2020**, *139*, 575–592. [CrossRef] [PubMed]
4. Huisman, S.; Mulder, P.A.; Redeker, E.; Bader, I.; Bisgaard, A.M.; Brooks, A.; Cereda, A.; Cinca, C.; Clark, D.; Cormier-Daire, V.; et al. Phenotypes and genotypes in individuals with SMC1A variants. *Am. J. Med. Genet. A* **2017**, *173*, 2108–2125. [CrossRef] [PubMed]
5. Gil-Rodríguez, M.C.; Deardorff, M.A.; Ansari, M.; Tan, C.A.; Parenti, I.; Baquero-Montoya, C.; Ousager, L.B.; Puisac, B.; Hernández-Marcos, M.; Teresa-Rodrigo, M.E.; et al. De Novo Heterozygous Mutations in SMC3 Cause a Range of Cornelia de Lange Syndrome-Overlapping Phenotypes. *Hum. Mutat.* **2015**, *36*, 454–462. [CrossRef]
6. Kaiser, F.J.; Ansari, M.; Braunholz, D.; Gil-Rodríguez, M.C.; Decroos, C.; Wilde, J.J.; Fincher, C.T.; Kaur, M.; Bando, M.; Amor, D.J.; et al. Loss-of-function HDAC8 mutations cause a phenotypic spectrum of Cornelia de Lange syndrome-like features, ocular hypertelorism, large fontanelle and X-linked inheritance. *Hum. Mol. Genet.* **2014**, *23*, 2888–2900. [CrossRef]
7. Parenti, I.; Gervasini, C.; Pozojevic, J.; Wendt, K.S.; Watrin, E.; Azzollini, J.; Braunholz, D.; Buiting, K.; Cereda, A.; Engels, H.; et al. Expanding the clinical spectrum of the 'HDAC8 -phenotype'—Implications for molecular diagnostics, counseling and risk prediction: Expanding the clinical spectrum of the 'HDAC8 -phenotype'. *Clin. Genet.* **2016**, *89*, 564–573. [CrossRef]
8. Latorre-Pellicer, A.; Ascaso, Á.; Lucia-Campos, C.; Gil-Salvador, M.; Arnedo, M.; Antoñanzas, R.; Ayerza-Casas, A.; Marcos-Alcalde, I.; Gómez-Puertas, P.; Ramos, F.J.; et al. Things are not always what they seem: From Cornelia de Lange to KBG phenotype in a girl with genetic variants in NIPBL and ANKRD11. *Mol. Genet. Genom. Med.* **2021**, *9*, e1826. [CrossRef]
9. Latorre-Pellicer, A.; Gil-Salvador, M.; Parenti, I.; Lucia-Campos, C.; Trujillano, L.; Marcos-Alcalde, I.; Arnedo, M.; Ascaso, Á.; Ayerza-Casas, A.; Antoñanzas-Pérez, R.; et al. Clinical relevance of postzygotic mosaicism in Cornelia de Lange syndrome and purifying selection of NIPBL variants in blood. *Sci. Rep.* **2021**, *11*, 15459. [CrossRef]
10. Russo, S.; Masciadri, M.; Gervasini, C.; Azzollini, J.; Cereda, A.; Zampino, G.; Haas, O.; Scarano, G.; Di Rocco, M.; Finelli, P.; et al. Intragenic and large NIPBL rearrangements revealed by MLPA in Cornelia de Lange patients. *Eur. J. Hum. Genet.* **2012**, *20*, 734–741. [CrossRef]
11. Baquero-Montoya, C.; Gil-Rodríguez, M.C.; Teresa-Rodrigo, M.E.; Hernández-Marcos, M.; Bueno-Lozano, G.; Bueno-Martínez, I.; Remeseiro, S.; Fernández-Hernández, R.; Bassecourt-Serra, M.; de Alba, M.R.; et al. Could a patient with SMC1A duplication be classified as a human cohesinopathy?: Could a patient with SMC1A duplication be classified as a human cohesinopathy? *Clin. Genet.* **2014**, *85*, 446–451. [CrossRef] [PubMed]
12. Helgeson, M.; Keller-Ramey, J.; Knight Johnson, A.; Lee, J.A.; Magner, D.B.; Deml, B.; Deml, J.; Hu, Y.-Y.; Li, Z.;

Donato, K.; et al. Molecular characterization of HDAC8 deletions in individuals with atypical Cornelia de Lange syndrome. *J. Hum. Genet.* **2018**, *63*, 349–356. [CrossRef] [PubMed]
13. Cheng, H.; Zhang, N.; Pati, D. Cohesin subunit RAD21: From biology to disease. *Gene* **2020**, *758*, 144966. [CrossRef] [PubMed]
14. Latorre-Pellicer, A.; Ascaso, Á.; Trujillano, L.; Gil-Salvador, M.; Arnedo, M.; Lucia-Campos, C.; Antoñanzas-Pérez, R.; Marcos- Alcalde, I.; Parenti, I.; Bueno-Lozano, G.; et al. Evaluating Face2Gene as a Tool to Identify Cornelia de Lange Syndrome by Facial Phenotypes. *Int. J. Mol. Sci.* **2020**, *21*, 1042. [CrossRef] [PubMed]
15. Kearney, H.M.; Thorland, E.C.; Brown, K.K.; Quintero-Rivera, F.; South, S.T. American College of Medical Genetics standards and guidelines for interpretation and reporting of postnatal constitutional copy number variants. *Genet. Med.* **2011**, *13*, 680–685. [CrossRef]
16. Riggs, E.R.; Andersen, E.F.; Cherry, A.M.; Kantarci, S.; Kearney, H.; Patel, A.; Raca, G.; Ritter, D.I.; South, S.T.; Thorland, E.C.; et al. Technical standards for the interpretation and reporting of constitutional copy-number variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics (ACMG) and the Clinical Genome Resource (ClinGen). *Genet. Med.* **2020**, *22*, 245–257. [CrossRef]
17. Somoza, J.R.; Skene, R.J.; Katz, B.A.; Mol, C.; Ho, J.D.; Jennings, A.J.; Luong, C.; Arvai, A.; Buggy, J.J.; Chi, E.; et al. Structural Snapshots of Human HDAC8 Provide Insights into the Class I Histone Deacetylases. *Structure* **2004**, *12*, 1325–1334. [CrossRef]
18. Marcos-Alcalde, Í.; Mendieta-Moreno, J.I.; Puisac, B.; Gil-Rodríguez, M.C.; Hernández-Marcos, M.; Soler-Polo, D.; Ramos, F.J.; Ortega, J.; Pie, J.; Mendieta, J.; et al. Two-step ATP-driven opening of cohesin head. *Sci. Rep.* **2017**, *7*, 3266. [CrossRef]
19. Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.* **1996**, *14*, 33–38. [CrossRef]
20. Deardorff, M.A.; Bando, M.; Nakato, R.; Watrin, E.; Itoh, T.; Minamino, M.; Saitoh, K.; Komata, M.; Katou, Y.; Clark, D.; et al. HDAC8 mutations in Cornelia de Lange syndrome affect the cohesin acetylation cycle. *Nature* **2012**, *489*, 313–317. [CrossRef]
21. Sun, L. A commentary on exome sequencing identifies a de novo mutation in HDAC8 associated with Cornelia de Lange syndrome. *J. Hum. Genet.* **2014**, *59*, 479. [CrossRef] [PubMed]
22. Feng, L.; Zhou, D.; Zhang, Z.; Liu, Y.; Yang, Y. Exome sequencing identifies a de novo mutation in HDAC8 associated with Cornelia de Lange syndrome. *J. Hum. Genet.* **2015**, *60*, 165. [CrossRef] [PubMed]
23. Gao, X.; Huang, Z.; Fan, Y.; Sun, Y.; Liu, H.; Wang, L.; Gu, X.-F.; Yu, Y. A Functional Mutation in HDAC8 Gene as Novel Diagnostic Marker for Cornelia De Lange Syndrome. *Cell. Physiol. Biochem.* **2018**, *47*, 2388–2395. [CrossRef] [PubMed]
24. Jezela-Stanek, A.; Murcia, P.V.; Jurkiewicz, D.; Iwanicka-Pronicka, K.; Je˛drzejowska, M.; Krajewska-Walasek, M.; Płoski, R. Novel variant in HDAC8 gene resulting in the severe Cornelia de Lange phenotype. *Clin. Dysmorphol.* **2019**, *28*, 124–128. [CrossRef]
25. Mio, C.; Passon, N.; Fogolari, F.; Cesario, C.; Novelli, A.; Pittini, C.; Damante, G. A novel de novo HDAC8 missense mutation causing Cornelia de Lange syndrome. *Mol. Genet. Genom. Med.* **2021**, *9*, e1612. [CrossRef]
26. Anna, A.; Monika, G. Splicing mutations in human genetic disorders: Examples, detection, and confirmation. *J. Appl. Genet.* **2018**, *59*, 253–268. [CrossRef]
27. Decroos, C.; Bowman, C.M.; Moser, J.A.S.; Christianson, K.E.; Deardorff, M.A.; Christianson, D.W. Compromised Structure and Function of HDAC8 Mutants Identified in Cornelia de Lange Syndrome Spectrum Disorders. *ACS Chem. Biol.* **2014**, *9*, 2157–2164. [CrossRef]
28. Decroos, C.; Christianson, N.H.; Gullett, L.E.; Bowman, C.M.; Christianson, K.E.; Deardorff, M.A.; Christianson, D.W. Biochem- ical and Structural Characterization of HDAC8 Mutants Associated with Cornelia de Lange Syndrome Spectrum Disorders. *Biochemistry* **2015**, *54*, 6501–6513. [CrossRef]
29. Deardorff, M.A.; Porter, N.J.; Christianson, D.W. Structural aspects of HDAC8 mechanism and dysfunction in Cornelia de Lange syndrome spectrum disorders: Structural Aspects of HDAC8 Mechanism. *Protein Sci.* **2016**, *25*, 1965–1976. [CrossRef]
30. Osko, J.D.; Porter, N.J.; Decroos, C.; Lee, M.S.; Watson, P.R.; Raible, S.E.; Krantz, I.D.; Deardorff, M.A.; Christianson, D.W. Structural analysis of histone deacetylase 8 mutants associated with Cornelia de Lange Syndrome spectrum disorders. *J. Struct. Biol.* **2021**, *213*, 107681. [CrossRef]
31. Porter, N.J.; Christianson, N.H.; Decroos, C.; Christianson, D.W. Structural and Functional Influence of the Glycine-Rich Loop G 302 GGGY on the Catalytic Tyrosine of Histone Deacetylase 8. *Biochemistry* **2016**, *55*, 6718–6729. [CrossRef] [PubMed]

TRABAJO 4:

An intragenic duplication in the AFF2 gene associated with Cornelia de Lange syndrome phenotype

frontiers | Frontiers in Genetics

Cristina Lucia-Campos, Ilaria Parenti, Ana Latorre-Pellicer, Marta Gil-Salvador, Ilaria Bestetti, Palma Finelli, Lidia Larizza, Maria Arnedo, Ariadna Ayerza-Casas, Julia Del Rincón, Laura Trujillano, Beatriz Morte, Luis A. Pérez-Jurado, Pablo Lapunzina, Elsa Leitão, Jasmin Beygo, Christina Lich, Fabian Kilpert, Sabine Kaya, Christel Depienne, Frank J. Kaiser, Feliciano J. Ramos, Beatriz Puisac, Juan Pié

An intragenic duplication in the AFF2 gene associated with Cornelia de Lange syndrome phenotype

Frontiers in Genetics

Noviembre 2024

doi: 10.3389/fgene.2024.1472543

PMID: 39553472

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 01 November 2024
DOI 10.3389/fgene.2024.1472543

OPEN ACCESS

EDITED BY
Jordi Pérez-Tur,
Spanish National Research Council (CSIC), Spain

REVIEWED BY
Jolanta Wierzba,
Medical University of Gdansk, Poland
Rowena Ng,
Kennedy Krieger Institute, United States

*CORRESPONDENCE
Feliciano J. Ramos,
framos@unizar.es
Beatriz Puisac,
puisac@unizar.es
Juan Pié, juanpie@unizar.es

RECEIVED 29 July 2024
ACCEPTED 11 October 2024
PUBLISHED 01 November 2024

CITATION
Lucia-Campos C, Parenti I, Latorre-Pellicer A, Gil-Salvador M, Bestetti I, Finelli P, Larizza L, Arnedo M, Ayerza-Casas A, Del Rincón J, Trujillano L, Morte B, Pérez-Jurado LA, Lapunzina P, Leitão E, Beygo J, Lich C, Kilpert F, Kaya S, Depienne C, Kaiser FJ, Ramos FJ, Puisac B and Pié J (2024) An intragenic duplication in the AFF2 gene associated with Cornelia de Lange syndrome phenotype. *Front. Genet.* 15:1472543.
doi: 10.3389/fgene.2024.1472543

An intragenic duplication in the *AFF2* gene associated with Cornelia de Lange syndrome phenotype

Cristina Lucia-Campos [1], Ilaria Parenti[2], Ana Latorre-Pellicer[1], Marta Gil-Salvador[1], Ilaria Bestetti[3], Palma Finelli[3,4], Lidia Larizza [5],María Arnedo [1], Ariadna Ayerza-Casas[1,6], Julia Del Rincón[1], Laura Trujillano [1,7], Beatriz Morte[8], Luis A. Pérez-Jurado[8,9,10], Pablo Lapunzina[8,11,12], Elsa Leitão[2], Jasmin Beygo[2], Christina Lich[2], Fabian Kilpert[2], Sabine Kaya[2], Christel Depienne[2],Frank J. Kaiser[2,13], Feliciano J. Ramos [14]*, Beatriz Puisac[1]* and Juan Pié[1]*

[1]Department of Pharmacology and Physiology, Unit of Clinical Genetics and Functional Genomics, School of Medicine, University of Zaragoza, CIBERER-GCV02 and IIS-Aragon, Zaragoza, Spain, [2]Institute of Human Genetics, University Hospital Essen, University of Duisburg-Essen, Essen, Germany, [3]SS Medical Genetics Laboratory, SC Clinical Pathology, Foundation IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy, [4]Department of Medical Biotechnology and Translational Medicine, Università degli Studi di Milano, Milan, Italy, [5]Experimental Research Laboratory of Medical Cytogenetics and Molecular Genetics, IRCCS Istituto Auxologico Italiano, Milan, Italy, [6]Unit of Paediatric Cardiology, Service of Paediatrics, University Hospital "Miguel Servet", Zaragoza, Spain, [7]Clinical and Molecular Genetics Area, Vall d'Hebron Hospital, Medicine Genetics Group, Vall d'Hebron Research Institute (VHIR),Barcelona, Spain, [8]Centre for Biomedical Network Research on Rare Diseases (CIBERER), Instituto de Salud Carlos III, Madrid, Spain, [9]Genetics Service, Hospital del Mar Research Institute (IMIM), Barcelona,Spain, [10]Genetics Unit, University Pompeu Fabra, Barcelona, Spain, [11]Institute of Medical and Molecular Genetics (INGEMM), University Hospital "La Paz"-IdiPAZ, Madrid, Spain, [12]ERN-ITHACA, University Hospital La Paz, Madrid, Spain, [13]Center for Rare Diseases (Essener Zentrum für Seltene Erkrankungen, EZSE), University Hospital Essen, Essen, Germany, [14]Department of Paediatrics, Unit of Clinical Genetics,Service of Paediatrics, University Hospital "Lozano Blesa", School of Medicine, University of Zaragoza,CIBERER-GCV02 and IIS-Aragon, Zaragoza, Spain

Cornelia de Lange syndrome (CdLS, OMIM #122470, #300590, #300882, #610759, and #614701) is a rare congenital disorder that affects the development of multiple organs and is characterized by physical abnormalitiesand cognitive and behavioral disabilities. Its molecular basis is mainly based on alterations in genes encoding structural and regulatory proteins related to thecohesin complex. Moreover, other transcriptional regulatory factors have beenlinked to this syndrome. However, additional causative genes are still unknown, since many patients still lack a molecular diagnosis. Herein, we describe a case with multiple affected family members presenting with an intragenic duplication in the AFF2 gene. The direct tandem intragenic duplication of exons 10, 11 and 12 was detected through high-resolution array Comparative Genomic Hybridization and next-generation sequencing technologies. Confirming the X-linked inheritance pattern, the duplication was found in the patient, his mother and his maternal aunt affected (dizygotic twins). Targeted sequencing with Oxford Nanopore Technologies revealed an aberrant transcript which is predominantly expressed in the patient and his aunt. Along with these results, a significant reduction in AFF2 gene expression levels was detected in these two individuals. Clinically both subjects exhibit a classic CdLS phenotype, whereas the mother is mostly unaffected. Consistent with the phenotypical

differences observed between the mother and the aunt, there is a marked difference in X-inactivation patterns skewing. Given the crucial role of AFF2 in transcriptional regulation, it is not surprising that AFF2 variants can give rise to CdLS phenotypes.Therefore, the AFF2 gene should be considered for the molecular diagnosis of thissyndrome.

KEYWORDS

AFF2, CdLS, intragenic duplication, familial case, X-inactivation, Oxford Nanopore Technologies

1 Introduction

Regulation of gene expression is an essential process in cell biology, which occurs precisely and coordinately from the early developmental stages (Asami et al., 2022). Recently, the term chromatinopathies has been introduced to describe disorders affecting transcription and chromatin remodeling. These disorders are often accompanied by developmental delay, facial dysmorphism, and intellectual disability (Avagliano et al., 2020; Parenti and Kaiser, 2021).

One of the best characterized chromatinopathies is Cornelia deLange syndrome (CdLS, OMIM #122470, #300590, #300882, #610759, #614701), a congenital disorder affecting the development of multiple organs. Its prevalence is estimated to be about 1/10,000-30,000 live births (Kline et al., 2007). Clinically, it is characterized by distinctive facial features, prenatal and postnatal growth retardation, as well as cognitive impairment and behavioraldisturbances (Ajmone et al., 2014; Pié et al., 2016; Grados et al., 2017;Kline et al., 2018). However, not all individuals with the syndromeexhibit the typical phenotype. The clinical features of the disorder can vary widely, ranging from mild to severe, and even overlap withthe phenotypes of other similar syndromes, such as Rubinstein- Taybi, KBG, Coffin-Siris, or Wiedemann-Steiner (Parenti et al., 2017; Cucco et al., 2020).

To date, CdLS has been associated with several types of genetic variations, including missense, nonsense, splicing, and copy number variations (CNVs) (Teresa-Rodrigo et al., 2014; Kline et al., 2018). These variants affect eight main genes: *NIPBL*, *SMC1A*, *SMC3*, *RAD21*, *HDAC8*, *BRD4*, *MAU2*, and *ANKRD11* (Gil-Rodríguez et al., 2015; Huisman et al., 2017; Kline et al., 2018; Krab et al., 2020; Parenti et al., 2020; Selicorni et al., 2021). Variants in the *NIPBL* gene account for approximately 70% of cases, with more than 10% of them involving postzygotic mosaicism (Latorre-Pellicer et al., 2021). The proteins encoded by the aforementioned genes are all involved intranscriptional regulation as well as chromatin remodeling processes (Kline et al., 2018). In terms of molecular implications, there is increasing evidence linking CdLS to global alterations of gene expression (Yuan B et al., 2015; Izumi, 2016; Mills et al., 2018; Garcia et al., 2021). Recent studies in cortical neurons from CdLS patients have revealed transcriptional alterations in genes associated with neuronal functions, including synaptic transmission or signaling processes (Weiss et al., 2021).

Given the extensive phenotypic and genetic variability of the syndrome, consensus criteria have been established to aid the clinical diagnosis and to classify the condition into classical and non-classical forms, both falling within the CdLS spectrum (Kline et al., 2018). However, 15% of patients with a CdLS-like phenotypestill remain without a clear genetic diagnosis.

In this study, we report an intragenic duplication within the *AFF2* gene in three individuals of the same family. CCG triplet repeats of the 5′ untranslated region (UTR) of this gene have been described as causative for Fragile XE syndrome (FRAXE, OMIM #309548) (Gu et al., 1996; Hillman and Gecz, 2001). Additionally, intragenic deletions have also been described (Gecz et al., 1996; Stettner et al., 2011). However, the full spectrum of clinical features associated with variants in *AFF2* still needs to be assessed.

In this work, we analyzed the consequences of the *AFF2* duplication through the study of gene expression in three family members. We subsequently compare the clinical features of the reported individuals with those of FRAXE syndrome, *AFF2* variants and CdLS. Our data suggest that *AFF2* may be considered a novel causal gene within the broad CdLS spectrum.

1. 2 Materials and method

2.1. Clinical diagnosis

Clinical data have been collected at the University Hospital “Lozano Blesa” using a standard restricted-term questionnaire. Detailed phenotypes of the individuals have been evaluated bythe patient’ clinician using the Human Phenotype Ontology (HPO) nomenclature for better comparison. The severity was determined according to the clinical score published by Klineand colleagues in 2018 (Kline et al., 2018). Additionally, as a complementary tool, Face2Gene technology was used to analyze the facial photographs of the individuals in the study (Gurovich Y et al., 2019; Latorre-Pellicer et al., 2020).

The biological samples have been collected from the sample collection C.0002316, registered in the Institute of Health Carlos III(ISCIII) for the investigation on the Cornelia de Lange syndrome, whose principal investigator is Prof. Feliciano J. Ramos.

The study has been performed according to the Declaration of Helsinki protocols and approved by the Regional Ethics Committee of Clinical Research from the Government of Aragón (CEICA; PI15/ 00707, PI19/01860). Informed consent was obtained from all subjects included in this study. Additional permission andinformed consent were obtained for the publication of the photographs.

2.2. Analysis of the AFF2-related phenotypic spectrum

This analysis was performed by reviewing all the relevant clinical information about *AFF2* variants reported in the literature and databases, including missense, nonsense, small deletions, small insertions variants and intragenic deletions and duplications, as well as all associated phenotypes. All cases were retriev from PubMed (https://pubmed.ncbi.nlm.nih.gov/), ClinVar (https:// www.ncbi.nlm.nih.gov/clinvar/), and DECIPHER (https://www. deciphergenomics.org/). We only took into consideration variants associated with clinical descriptions.

2.3. Molecular diagnosis

2.3.1. Cell culture

Human primary fibroblasts were obtained from a skin biopsy from the anterior thigh. Fibroblasts were cultured in Dulbecco's Modified Eagle Medium (DMEM, HyClone™) supplemented with 10% fetal bovine serum (FBS, Gibco™), 1% streptomycin/penicillin(HyClone™), under the conditions of a humidified atmosphere at37°C with 5% CO_2 in the incubator. The same protocol was used for all the cell lines. Cells were cultured between passages 1 and 4. Beforeisolating the total DNA and RNA, the cell lines were tested to verifythe absence of *Mycoplasma* contamination.

2.3.2. Isolation of DNA and RNA

Genomic DNA and RNA isolation were performed from bloodcells and fibroblast samples. Lymphocytes derived' DNA was extracted using a conventional phenol–chloroform isoamyl alcohol method. The DNA from fibroblast was isolated using the PureLink™ Genomic DNA kit (Invitrogen, Thermo Fisher Scientific), according to the manufacturer's protocol. Isolation of RNA was performed from both blood and fibroblast lines using the Total RNA Purification kit (Norgen Biotech) according to the manufacturer's protocol. The purity and integrity of the nucleic acids were assessed by electrophoresis and spectrophotometry using NanoDrop ND-2000 spectrophotometer (Thermo Fisher Scientific).

2.3.3. DNA analyses

2.3.3.1. Exome sequencing and reanalysis

Whole-exome sequencing was performed using blood derived'genomic DNA of the patient (II-1) in a private laboratory (qGenomicsS.L.). Then, the clinical and genomic raw data from the patient (II-1) were submitted to the Undiagnosed Rare Diseases Programme (EnoD; https://www.ciberer.es/en/transversal-programmes/scientific-projects/undiagnosed-rare-diseases- programme-enod) for more in-depth investigation. This programme is an initiative of the Center for Biomedical Network Research on Rare Diseases (CIBERER) to identify the genetic cause of rare diseases in selected undiagnosed cases (Luque et al., 2022). Results are periodically reassessed by an expert team using the software associated withthe EnoD Programme (Bullich et al., 2022).

2.3.3.2. Array Comparative Genomic Hybridization

Human Genome Comparative Genomic Hybridization (CGH) Microarray kit 1 × 400 K was performed on genomic DNA from blood of the patient (II-1) according to the manufacturer's instructions (Agilent Technologies). Data were extracted and analyzed for copy number changes using Agilent CytoGenomics v.3.0 software (Agilent Technologies). Copy number variations (CNVs) were analyzed and mapped using the Human Genome assembly GRCh37/hg19. CNV classification was performedaccording to the Database of Genomic Variants (DGV) (http://projects.tcag.ca/variation/) to exclude common polymorphic CNVswith a frequency >1% in healthy population.

2.3.3.3. Deep-sequencing Cornelia de Lange syndromegene panel

The fibroblast´s genomic DNA from the patient (II-1) and blood´s genomic DNA from the mother (I- 2) and the maternal aunt (I-3) were sequenced using a deep-sequencing panel including 35 genes related with CdLS, with a total target region of 249.25 kb. Library preparation, emulsion PCR, bead enrichment, and chip loading were automatically performed on an Ion Chef™ instrument (Thermo Fisher Scientific) using Ion AmpliSeq™ Kit for Chef DL8 and 530™Kits (Thermo Fisher Scientific) according to the manufacture's protocols. Templates were sequenced on an Ion S5™ XL sequencer (Thermo Fisher Scientific) using 530™ Kits with a read length of 200. Raw data analysis was carried out with Torrent Suite SoftwareTM (v.5.12.3). After alignment to the hg19 human reference genome, the annotation of SNVs, indels and CNVs was performed using the Ion Reporter SoftwareTM (v.5.20.2.0). The sequencing data were evaluated using Integrative Genomics Viewer (IGV) (http://software.broadinstitute.org/software/igv/).

2.3.3.4. Real-time quantitative PCR (qPCR)

Confirmation of the coding regions involved in the CNV was performed on genomic DNA from blood from all available family members (II-1, I-1, I-2, I-3, I-4, I-5). The analysis was screened byreal-time quantitative PCR (qPCR) using an NZYSupreme qPCR Green Master Mix (2x) ROX (NZYtech) on QuantStudio™5 System(Applied Biosystems). Samples were assessed in triplicate in each experiment and two biological replicates were performed. Data were analysed through the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001). Statistical analyses and graphics were generated with GraphPad Prism 9 software. The primer pairs bind to inner regions of exons inthe gene of interest, and their sequences and PCR conditions are available upon request.

2.3.4. RNA analyses

2.3.4.1. cDNA synthesis and analysis

Total RNA isolated from blood and fibroblast lines was retrotranscribed into single-strand cDNA using a SuperScript™ First-Strand Synthesis System for RT-PCR kit (Invitrogen, Thermo Fisher Scientific) following the manufacturer's protocol. cDNA synthesis was performed using 2 µg of total RNA and random hexamers as primers.

Sanger sequencing was performed to localize the duplication breakpoints. This analysis was carried out on cDNA from blood and fibroblasts from the patient (II-1), as well as from his father (I-1), his mother (I-2) and his maternal aunt (I-3). PCRs were performed using a KAPA2G Robust PCR kit (KAPA BIOSYSTEMS). Primer sequences and PCR conditions are available upon request. Amplicons were sequenced using a BigDye™ Terminator v.3.1 Cycle

Sequencing kit (Thermo Fisher Scientific). Deletion junction sequences were aligned to the human reference genome sequence (GRCh38/hg38), and electropherograms were analyzed with the ChromasPro 1.5 software (Technelysium Pty Ltd.).

2.3.4.2. Targeted sequencing with Oxford Nanopore Technologies (ONT)

RNA from fibroblasts was retrotranscribed with the LunaScript® RT SuperMix Kit (NEB, Basdorf, Germany), following manufacturer' instruction. The full length *AFF2* transcript (ENST00000370460.7, NM_002025.4) was amplified from cDNAof the patient (II-1), his mother (I-2) and his maternal aunt (I-3) bytouch-down PCR using PrimeSTAR GXL DNA Polymerase (Takara Bio) with 10% DMSO (Primer sequences are available upon request). The PCR products were purified using the DNA Clean & Concentrator (Zymo Research). A library of multiplexed sampleswas prepared using the SQK-LSK109 ligation-based sequencing kit (Oxford Nanopore) and the native barcoding protocol (EXP- NBD196). In summary, 200 fmol of each purified amplicon were incubated with NEBNext Ultra II End Repair/dA-tailing Module Reagents (NEB) at 20°C for 5 min and 65°C for 5 min, after which,barcodes were ligated using the NEB Blunt/TA Ligase Master Mix at 20°C for 20 min and 65°C for 10 min. Barcoded amplicons were pooled and purified using AMPure XP beads (Beckman Coulter), adapters were ligated incubating during 10 min with Adapter Mix IIExpansion/NEBNext Quick Ligation Reaction Module followed byclean-up with AMPure XP beads. Approximately 15 ng of the library were loaded onto a MinION R9.4.1 FLO-MIN106 flow cell sequencing for up to 24 h while monitoring using the MinKNOW software. All steps were performed following the manufacturer's protocols. Guppy (version 6.5.7) (Community N. Downloads - Release notes. Oxford Nanopore Technologies) was used for base-calling, pycoQC (version 2.5.2) (Leger and Leonardi, 2019) and NanoPlot (version 1.41.6) (De Coster and Rademakers, 2023) for quality control. The command line version of Guppy, guppy_basecaller, used the "sup" model and the parameters: "--recursive -- compress_fastq --do_read_splitting --calib_detect-- records_per_fastq 0 -- enable_trim_barcodes." Reads were alignedwith minimap2 (version 2.28-r1209) (Li, 2018) to a reference consisting of NM_002025.4 concatenated exons. Variants were called with Sniffles2 (version 2.2) (Smolka et al., 2024) as well as the reference number and variant reads.

2.3.4.3. Gene expression analysis

Global expression levels of AFF2 (NM_002025.4) in fibroblasts were assessed by qPCR using NZYSupreme qPCR Green Master Mix (2x) ROX (NZYtech) on QuantStudio™ 5 System (Applied Biosystems). Six healthy controls (three male and three female) were used for comparison. 1 µL of cDNA and 2 µM of primers were used in a volume of 10 µL for each reaction. Samples were assessed in triplicate in each experiment and three biological replicates were performed. Primer sequences are available upon request. The global gene expression levels were calculated normalizing to the housekeeping β-actin gene and compared with the controls. The Ct values for each sample were determined with amplification plots in the logarithmic phase. The PCR outcome and efficiency of amplification were determined using QuantStudio™ Design and Analysis Software v1.5.1 using the $2^{-\Delta\Delta CT}$ method. Statistical analyses and graphics were generated with GraphPad Prism 9 software.

2.3.5. Assessment of X chromosome inactivation

The level of skewing of X chromosome inactivation (XCI) wasevaluated by studying the methylation status at two polymorphic regions of the X chromosome. Specifically, the highly polymorphicregion (CAG repeats) in exon 1 of the androgen receptor (*AR*) genewas explored. Digestions with methylation-sensitive restriction enzymes HpaII and HhaI were carried out at both blood and fibroblas genomic DNA from the mother (I-2) and maternal aunt affected (I-3). In addition, the highly polymorphic region (CGG repeats) of the *FMR1* gene was also evaluated. Digestion with the methylation-sensitive restriction enzyme HpaII was carried out at both blood and fibroblast genomic DNA from the mother (I-2) andthe maternal aunt (I-3). For both loci, PCR amplification of digestedand undigested DNA was performed. Finally, the allele ratios werecalculated (Allen et al., 1992; Carrel and Willard, 1999). We considered alleles separated by more than two trinucleotide repeats to be informative. For skewed XCI, we used a cutoffof >80:20%, and for extremely skewed XCI, a cutoff 241 of >95:5%.

2.

3. 3 Results

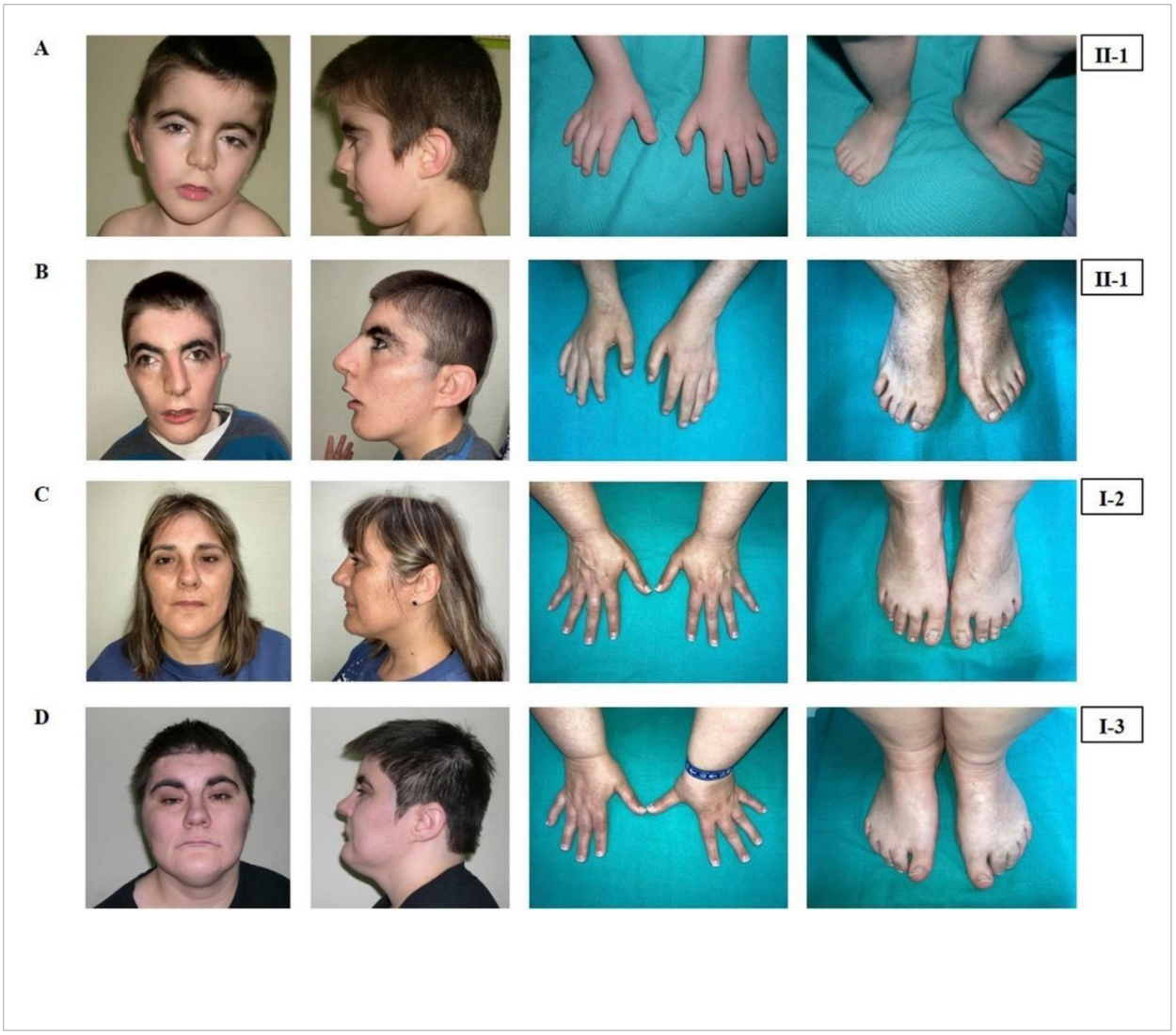
3.1. Clinical report

3.1.1. Patient (II-1)

The index patient of this study is an adolescent male (II-1, see Figures 1A, B) born from a non-consanguineous couple at 35 weeksof gestational age by Cesarean section. His birth weight was 1.640 kg (−2.29 SD), length 44 cm (−1.1 SD) and head circumference 29.5 cm (−1.36 SD). Due to the low birth weight and the presence of a congenital atrial septal defect, he was admitted to the Neonatal Intensive Care Unit (NICU). He was clinically diagnosed with CdLS at the age of 4 years (Figure 1A). The following features supported CdLS diagnosis: synophrys (HP:0000664), thick eyebrows (HP:0000574), long philtrum (HP:0000343), thin upper lip vermilion (HP:0000219) with downturned corners of the mouth (HP:0002714), microcephaly (HP:0000252), and short fifth finger (HP:0009237) (Figures 1A, B). He additionally presented with cognitive impairments, abnormal heart morphology (HP:0001627), absent speech (HP:0002300), hyperactivity (HP:0000752), and a short attention span (HP:0000736). Psychiatric assessment was carried out at the age of 13 years using the Brunet-Lezine BL-EC Scale and suggested a developmental quotient of around 24 months. A summary of the clinical features of the patient (II-1) is shownin Table 1.

The latest evaluation was performed at 19 years of age. At this time, his clinical features remained consistent with the initial CdLSdiagnosis (Figure 1B). He additionally presented with gastroesophageal reflux (HP:0002020) and self-injurious behavior (HP:0100716).

The calculated clinical CdLS score was 13 (Supplementary Table S1). According to the consensus criteria, he would be included in the classic CdLS phenotype group. Further clinical analysis was conducted with Face2Gene at ages 4 (Figure 2A) and 19 year old (Figure 2B), and the prediction showed high and medium CdLS Gestalt similarity, respectively.



3.1.2. Mother (I-2)

The mother of the index patient presented with some milder features such as brachycephaly (HP:0000248), long philtrum (HP: 0000343), and thin upper lip vermilion (HP:0000219) (Figure 1C).The clinical evaluation yielded a CdLS score of 6 (Supplementary Table S1), which is below the threshold for a classification of non-classic CdLS. Face2Gene revealed that she exhibits a low similarity tothe CdLS Gestalt features (Figure 2C).

3.1.3. Maternal aunt (I-3)

The maternal aunt affected of the index patient (I-3, mother's dizygotic twin) also exhibits a CdLS phenotype (Figure 1D). The clinical diagnosis of CdLS was primarily based on the following dysmorphic facial features: synoprhys (HP:0000664), thick eyebrows (HP:0000574), long and smooth philtrum (HP:0000343,HP:0000343), thin upper lip vermilion (HP:0000219), downturned corners of the mouth (HP:0002714), and microcephaly (HP: 0000252) (Figure 1D).

Moreover, she presented with moderate intellectual disability (HP:0002342), global developmental delay (HP:0001263), postnatal growth retardation (HP:0008897), small hands (HP:0200055), small feet (HP:0001773), and hirsutism (HP: 0001007) (more detailed clinical features are shown in Table 1).

The clinical evaluation yielded a CdLS score of 14 (Supplementary Table S1), consistent with a diagnosis of classic CdLS, as per the consensus criteria. Face2Gene result indicates a medium similarity with CdLS Gestalt (Figure 2D). In contrast, the two other sisters (I-4 and I-5) did not display any signs compatiblewith the diagnosis of CdLS (Figure 2E).

3.2. Molecular diagnosis

3.2.1. DNA molecular analyses

Exome sequencing of the patient (II-1) did not reveal any pathogenic or likely-pathogenic single nucleotide variant. High-resolution array-CGH analysis of the patient (II-1)

TABLE 1 Clinical features of the patient (II-1), his mother (I-2) and his maternal aunt affected (I-3) compared with the clinical features of FRAXE syndrome, copy number variants affecting the *AFF2* gene and Cornelia de Lange syndrome.

Clinical features	FRAXE syndrome	Variants *AFF2*	CdL syndrome	Patient (II-1)	Mother (I-2)	Maternal aunt (I-3)
Global developmental delay HP:0001263						
Growth delay HP:0001510						
Prenatal growth retardation HP:0001511						
Postnatal growth retardation HP:0008897						
Abnormal skull morphology HP:0000929						
Microcephaly HP:0000252						
Brachycephaly HP:0000248						
Dysmorphic facial features HP:0001999						
Synophrys HP:0000664						
Arched, thick eyebrows HP:0002553, HP:0000574						
Long eyelashes HP:0000527						
Proptosis HP:0000520						
Palpebral ptosis HP:0000508						
Ocular hypertelorism HP:0000316						
Upturned nasal tip HP:0000463						
Long, smooth philtrum HP:0000343, HP:0000319						
Thin upper lip vermilion HP:0000219						
Downturned corners of the mouth HP:0002714						
High arched palate HP:0000218						
Dental crowding HP:0000678						
Low anterior hairline HP:0000294						
Narrow face HP:0000275						
Abnormality of limbs HP:0040064 and musculoskeletal system HP:0033127						
Small hands HP:0200055						
Clinodactyly HP:0004209						
Short fifth finger HP:0009237						
Small feet HP:0001773						
Vertebral anomalies HP:0003468						
Broad chest HP:0000914						
Neurodevelopmental abnormality HP:0012759						
Intellectual disability HP:0001249						
Specific learning disability HP:0001328						
Delayed speech and language development HP:0000750						
Abnormal social communication behavior HP:0034434						

(Continued on following page)

TABLE 1 (*Continued*) Clinical features of the patient (II-1), his mother (I-2) and his maternal aunt affected (I-3) compared with the clinical features of FRAXE syndrome, copy number variants affecting the *AFF2* gene and Cornelia de Lange syndrome.

Clinical features	FRAXE syndrome	Variants *AFF2*	CdL syndrome	Patient (II-1)	Mother (I-2)	Maternal aunt (I-3)
Behavior disorders HP:0000708						
Attention déficit disorder HP:0007018						
Hyperactivity HP:0000752						
Aggressive behavior HP:0000718						
Autistic behavior HP:0000729						
Other phenotypic abnormality HP:0000118						
Seizures/Epilepsy HP:0001250						
Hearing impairment HP:0000365						
Cardiovascular anomalies HP:0002564						
Apnea HP:0002104						
Gastroesophageal reflux HP:0002020						
Vesicoureteral reflux HP:0000076						
Recurrent infections HP:0002719						
Hirsutism HP:0001007						

While blue cells indicate presence of the feature, beige cells indicate its absence.

revealed an intragenic duplication in the *AFF2* gene (NM_002025.4). Thisduplication is described neither in population genomic databases nor in patients with neurodevelopmental disorders or our internal database of patients with CdLS. The chromosomal coordinates of the duplication are chrX:148,934,576-148,960,040 (GRCh38/hg38) (Figure 3A), spanning from intron 9 to intron 12 and harboring exons 10, 11, and 12 of the *AFF2* gene. The reanalysis of exome data within the Program for Undiagnosed Rare Diseases (ENOD) detected the intragenic duplication in the *AFF2* gene (NM_ 002025.4). In addition, segregation analyses in the family were carried out through deep-sequencing (>1,000x) using CdLS gene panel (Figure 3B) and qPCR (Supplementary Figure S1), and bothconfirmed the intragenic duplication in the *AFF2* gene in the patient(II-1) in hemizygosity, as well as in the mother (I-2) and thematernal aunt (I-3) who are heterozygous carriers.

3.2.2. Breakpoint identification

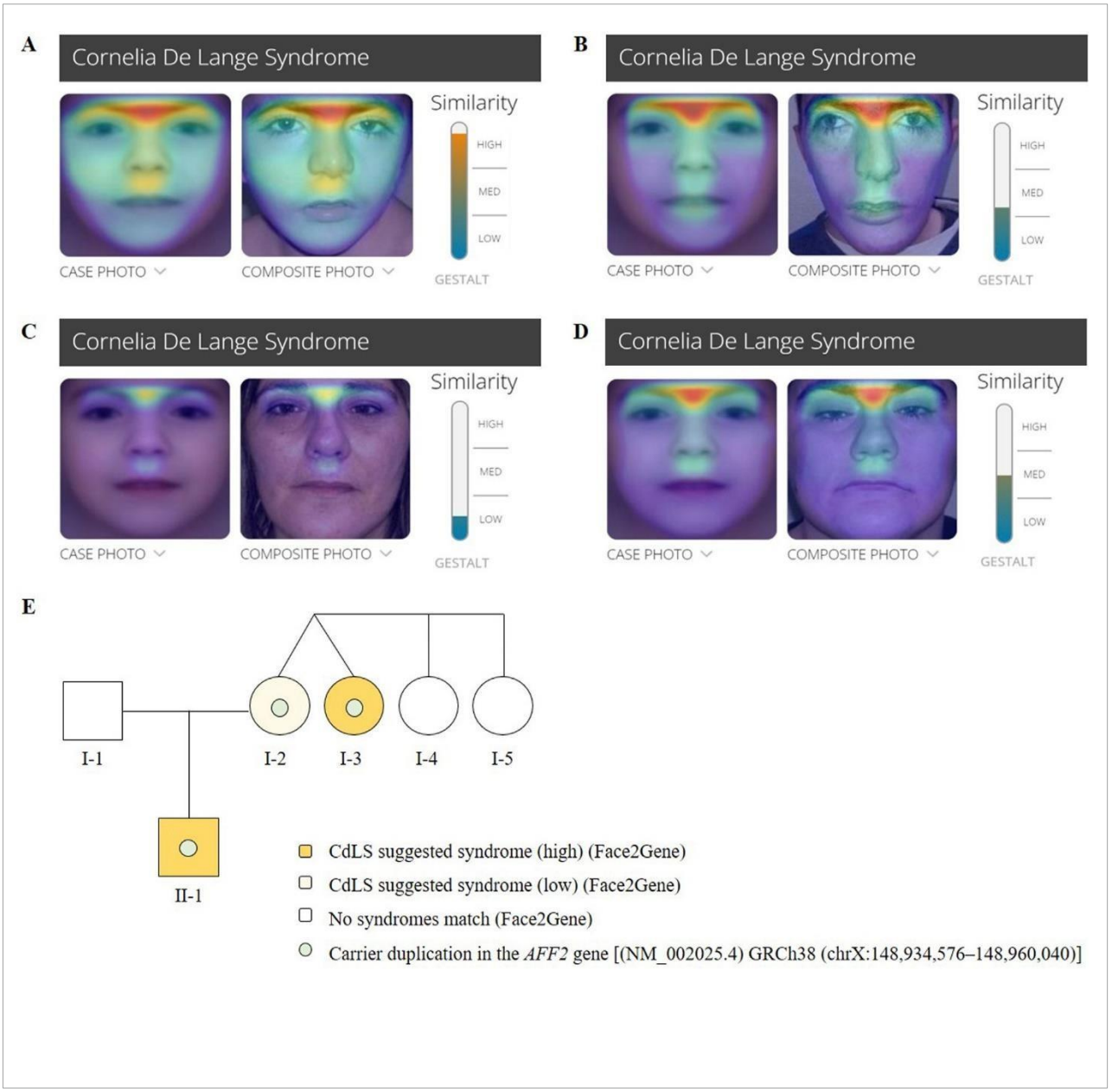
To investigate if the duplicated sequence was arranged in tandem, amplification of a cDNA fragment using a forward primer located distal to exon 12 and a reverse primer located proximal to exon 10 wasperformed. By doing this, a PCR product could be detected in the patient (II-1), his mother (I-2) and his maternal aunt (I-2), but not in the father (I-1) (Supplementary Figure S2A). Sanger sequencing of theaberrant product confirmed that the duplication is arranged in a direct tandem orientation (Supplementary Figure S2B), consisting of exons10, 11, and 12 followed by a repeated identical sequence of exons 10, 11 and 12.

The analysis of the aberrant transcript sequence revealed the additional presence of the following missense substitution: p.(Glu897Asp) at the breakpoint. Notably, the aberrant transcript could potentially be translated into a mutant protein carrying 431 additional amino acids. This would include a duplication of the serine-rich domain, the two bipartite nuclear localization signals and the AF4 interaction motif (Supplementary Figure S3).

3.2.3. Relative expression of the AFF2 transcripts

The full-length *AFF2* transcript was amplified by PCR. A single band of 3,933 bp corresponding to the wild-type *AFF2* construct was evident in the healthy control (Figure 4A). Individuals II-1, I-2, andI-3, however, showed an additional band, indicating the presence ofan aberrant transcript (1,293 bp larger, with exons 10, 11 and12 duplicated). The ratios between these two transcripts varied among the analyzed individuals. Specifically, the patient (II-1)and his maternal aunt affected (I-3) predominantly express the aberrant transcript, with relative expression levels calculated at 95% and 94% of the total, respectively, based on band intensities (Figure 4B). In contrast, the unaffected mother (I-2) primarily expresses the wild-type allele, with the aberrant allele accounting for only 12% of the total (Figure 4B).

Furthermore, an additional intermediate band was detected in the samples from both the mother (I-2) and the maternal aunt (I-3).To elucidate its origin, we utilized Oxford Nanopore Technology for comprehensive analysis of the amplified full-length transcripts. This approach confirmed the presence of the wild-type transcript (Figure 4C, peak around 4 kb) and confirmed the exons composition of the aberrant transcript (peak around 5 kb). However, reads corresponding to the intermediate amplicon (peak around 4.5 kb) mapped to the centromeric region of chromosome 16, indicating that this amplicon is rather a nonspecific PCR artifact. Analysis of relative read numbers corroborated the previous findings from the PCR band intensity



analysis, showing that the patient (II-1) and his maternal aunt (I-3)predominantly express the aberrant transcript (82% and 78%, respectively), while the unaffected mother (I-2) primarily expresses the wild-type allele (89%) (Figure 4D).

3.2.4. Total AFF2 expression

The total *AFF2* mRNA levels were assessed by qPCR. Expression levels of the patient (II-1) were normalized on the average expression levels of sex-matched healthy controls. The mean *AFF2* expression level in each individual was calculated as a percentage relative to the controls. While no significant changes were observed in *AFF2* mRNA levels in the mother (I-2 = 90,60%), a significant decrease of approximately 45% in total *AFF2* mRNA levels was noticed in the maternal aunt (I-3 = 56,16%), and a significant decrease of approximately 70% in total *AFF2* mRNA levels was observed in the patient (II-1 =33.80%) (Figure 5).

3.2.5. X-inactivation

The X-inactivation analysis was non-informative at the *AR* locus due to homozygosity. Analysis of the *FMR1* locus in the mother (I-2) revealed borderline skewing in favor of the wild-type allele (Allele 2) in blood (81/19) and random X-inactivation in fibroblast (36/64) (Table 2). On the other hand, the maternal aunt (I-3) similarly preferentially expresses the wild-type allele (Allele 2) in blood (83/17), but exclusively expresses the mutant allele (Allele 1) in fibroblasts (100/0) (Table 2).

3.2.6. Systematic review of AFF2 variants and their clinical features

A total of 50 *AFF2* variants were reviewed from ClinVar, DECIPHER and Pubmed and included in Supplementary Table S2. Among them, of which 7 were reported in case reports, 27 were

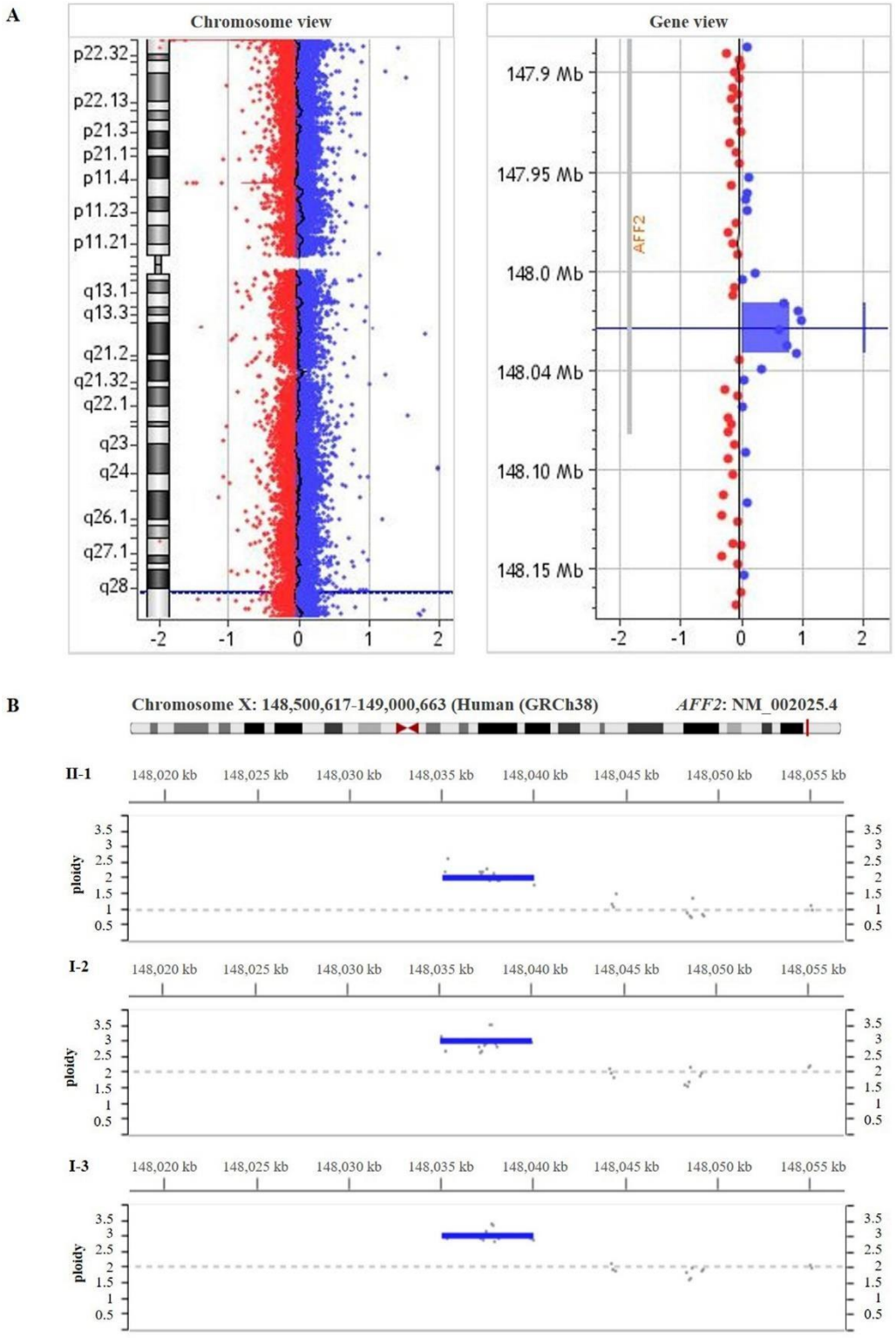


FIGURE 3
(A) High-resolution array-CGH 400 K performed on genomic DNA from the patient's blood (II-1) showed a duplication spanning 25.5 kb [arr(GRCh38) Xq28(148,934,576 – 148,960,040)x2)] involving part of the coding region of *AFF2* gene (NM_002025.4). (B) The deep-sequencing panel ofgene amplicons specific for Cornelia de Lange syndrome result showed the duplication in *AFF2* [(NM_002025.4) GRCh38 (chrX:148,95,3536- 148,958,504)] in the patient (II-1), his mother (I-2) and his maternal aunt affected (I-3).

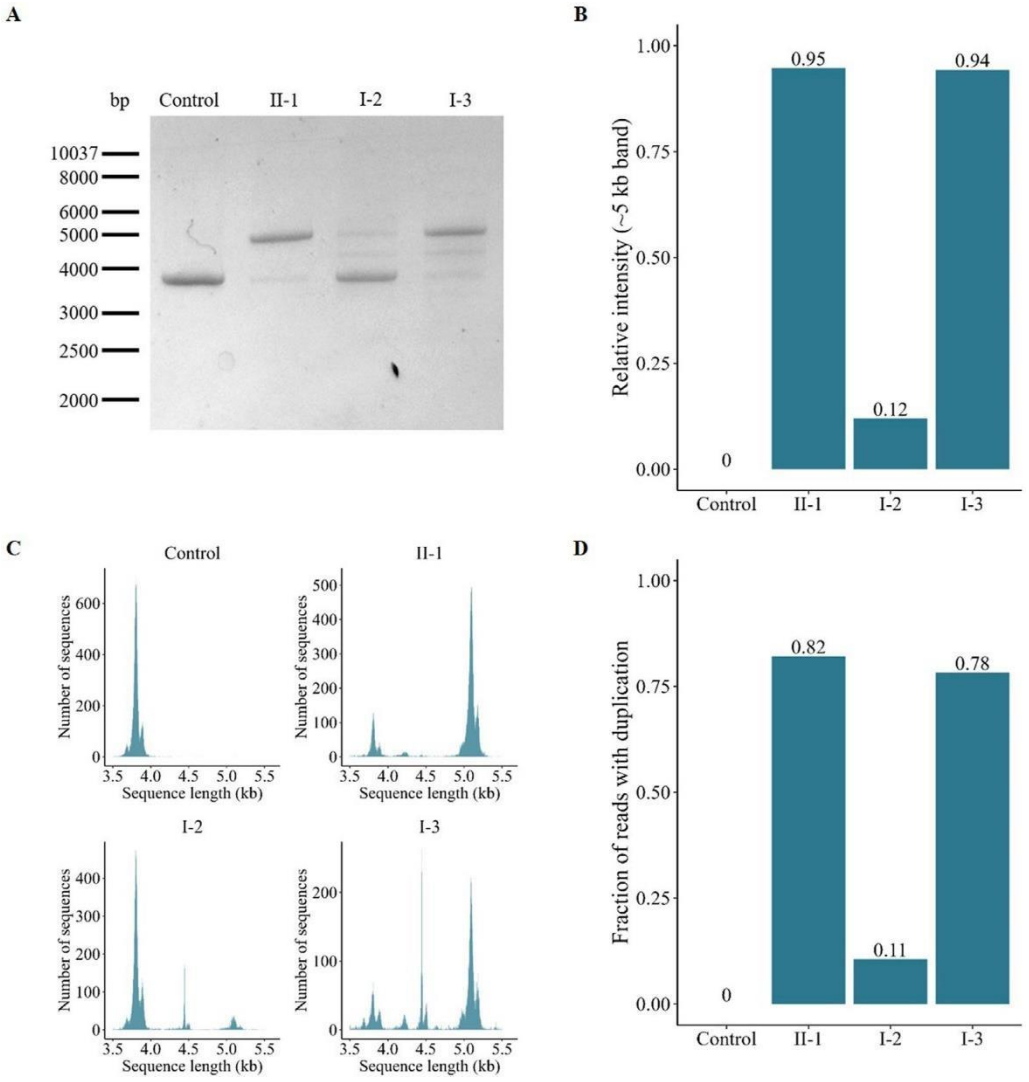


FIGURE 4
Characterization and quantification of the AFF2 aberrant transcript. (A) Gel electrophoresis of the amplified full-length AFF2 transcript. The control simple exclusively expresses the wild-type allele (≈4 kb). In contrast, the patient (II-1), his mother (I-2) and his maternal aunt affected (I-3) express both the wild-type and the duplicated allele (≈5 kb) at varying ratios. (B) Relative intensity of the PCR bands of the duplicated allele. The patient (II-1) and his maternal aunt (I-3) predominantly express the duplicated allele, whereas the unaffected mother (I-2) mainly expresses the wild-type allele. (C) Sizebased distribution of the number of reads obtained after targeted sequencing with Oxford Nanopore Technologies. The peak around 4 kb represents the wild-type allele, while the peak around 5 kb corresponds to the mutant allele with the duplication of exons 10, 11, and 12. The middle peak around 4.5 kb, observed primarily on the mother (I-2) and the maternal aunt (I-3), is an unspecific product that maps at the centrosome region of chromosome 16. (D) Fraction of reads with the duplication obtained after targeted sequencing with Oxford Nanopore Technologies.

published in studies of large cohorts of families or patients with intellectual disability, autism or epilepsy, 11 were reported in ClinVar as likely pathogenic or pathogenic and five were reported in DECIPHER. The type of the mentioned variants were intragenic duplications (6 cases), intragenic deletions (12 cases), small deletions (1 case), small insertions (3 cases) and SNVs such as nonsense (4 cases), missense (21 cases), and splicing (3 cases) variants (Supplementary Table S2).

Although there is no detailed clinical description of these cases, most are associated with intellectual disability (HP:0001249) (23/50) or autism (HP:0000717) (30/50). In cases in which data on facial dysmorphism were available, some of the traits that the individuals manifest were epicanthal folds (HP:0000286) (Gedeon et al., 1995), divergent strabismus (HP:0020049) (Gedeon et al., 1995), synophrys (HP:0000664) (Valentino et al., 2021), prominent nasal bridge (HP:0000426) (Gedeon et al., 1995; Sahoo et al., 2011), anteverted nares (HP:0000463) (Valentino et al., 2021), cleft palate (HP:0000175) (DECIPHER ID 273134, ClinVar VCV000986210.2) or widely spaced teeth (HP:0000687) (Gedeon et al., 1995) (Supplementary Table S2).

4. 4 Discussion

The genetic etiology of Cornelia de Lange syndrome (CdLS) ismainly attributed to variants affecting the cohesin

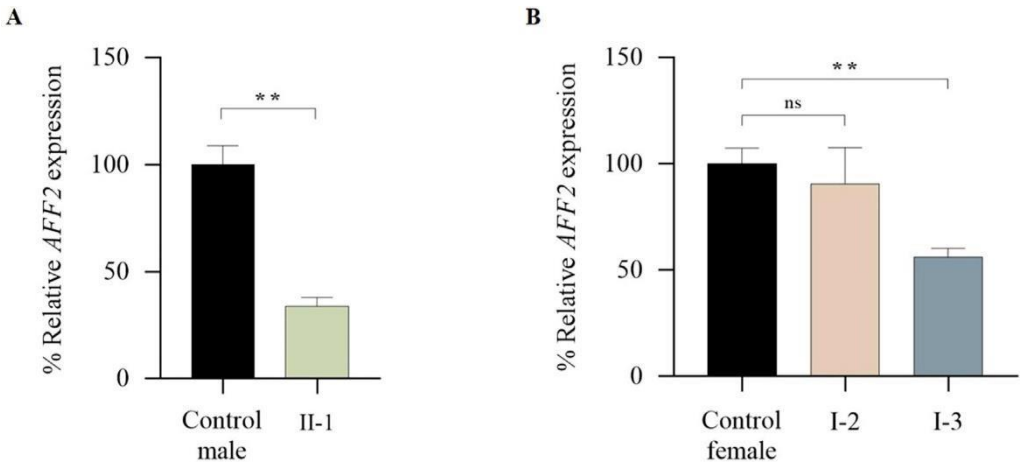


FIGURE 5
qPCR analysis of global expression levels of the AFF2 gene (NM_002025.4). (A) Global expression levels of AFF2 in the patient (II-1). Expression levels of AFF2 (NM_002025.4) were normalized to actin. The expression level in male controls (n = 3) was arbitrarily set at 100%. (B) Global expression levels of AFF2 in the mother (I-2) and the maternal aunt affected (I-3). Expression levels of AFF2 (NM_002025.4) were normalized to actin. The expression level in female controls (n = 3) was arbitrarily set at 100%.

TABLE 2 X-inactivation analysis.

Individual	DNA source	Locus	Allele 1:Allele 2
Mother (I-2)	Blood	*AR*	NI
I-2	Fibroblast	*AR*	NI
I-2	Blood	*FMR1*	81:19
I-2	Fibroblast	*FMR1*	36:64
Maternal aunt (I-3)	Blood	*AR*	NI
I-3	Fibroblast	*AR*	NI
I-3	Blood	*FMR1*	83:17
I-3	Fibroblast	*FMR1*	0:100

Individual studied, DNA, source, locus and the ratio of X-inactivation for each allele are specified. NI, indicates not informative (skewing.could not be determined due to the homozygosity of the tested loci).

complex, and more specifically to heterozygous variants in the *NIPBL* gene (Klineet al., 2018; Alonso-Gil and Losada, 2023). However, an increasingnumber of patients with pathogenic variants in non-cohesin relatedgenes with phenotypes that clinically overlap with CdLS are beingdescribed (Parenti et al., 2017; Avagliano et al., 2020; Eigenhuis et al.,2022; Shangguan and Chen, 2022). Thus, the term CdLS spectrumhas emerged to include both classical and non-classical phenotypesof the syndrome (Kline et al., 2018).

In this study, we report a familial case with multiple affected family members presenting with a previously unreported intragenicduplication in the *AFF2* gene. After the first molecular evaluation, the causative CdLS genes known at that time were analyzed by Sanger sequencing but no pathogenic variation was revealed. The high-resolution array-CGH and the reanalysis of exome detected the intragenic duplication in the *AFF2* gene that could explain the patient's phenotype. Familial segregation studies, performed by deep sequencing and qPCR supported the possible pathogenicity of the variant. The patient (II-1), his mother (I-2) and his maternal aunt affected (I-3) carried the duplication, while none of the maternal aunts unaffected harbored it. It was not possible to conduct the study in the maternal grandparents. In this case, theuse of stepwise genetic diagnostic techniques enabled us to definitively confirm the identified intragenic duplication in the *AFF2* gene. This is an example of the difficulty of reaching a molecular diagnosis in patients with rare genetic diseases.

Studies at the mRNA level revealed a direct tandem duplicationof exons 10, 11 and 12, which is common in these types of variants (Newman et al., 2015; Richardson et al., 2019). Implementing long- read sequencing using Oxford Nanopore Technologies allowed us to sequence and quantify the aberrant transcripts (Glinos et al., 2022). Through this technique, we detected an aberrant transcript that is more expressed in the patient (II-1) and his maternal aunt (I-3) in comparison to the mother (I-2), who exhibits minimal expression of the aberrant transcript. Conversely, global expression of the *AFF2* gene showed a significant decrease in total transcript levels in the patient (II-1) and his maternal aunt (I-3). This overall decrease in

expression may result from mRNA quality control mechanisms thatdegrade aberrant transcripts (Łabno et al., 2016; Peck et al., 2019).The mother (I-2), who is largely unaffected, does not showed a reduction of the global expression. Thus, the preferential expressionof the aberrant transcript, coupled with the decrease in globalexpression, accounts for the more severe phenotype of the patient (II-1) and his maternal aunt (I-3). This phenomenon has been described in other reported cases of intragenic duplications (Takahashi S et al., 2015; Zepeda-Mendoza et al., 2019).

Surprisingly, the patient (II-1) expresses 10%–20% of the wild- type transcript, which is not expected given his hemizygous status for the duplication. This observation suggests the presence of a mechanism that preserves wild-type mRNA during alternative splicing (Cetin et al., 2016). One possible underlying mechanism could be DNA methylation, which may influence alternative splicing and facilitate the rescue of the normal size transcript (Lev et al., 2015). On the other hand, the functional impact at the protein levelin this case is challenging to predict. Intragenic duplications frequently result in a loss of function (Bertini et al., 2023). The additional 431 amino acids may alter the AFF2 structure due to thegain of functional domains and as a consequence affect its RNA-binding function (Bensaid et al., 2009).

Clinically, our index patient was classified as classic CdLS during his first evaluation at the age of four, with the diagnosis confirmed at his most recent evaluation. Furthermore, the application of artificial intelligence databases such as Face2Gene (Gurovich Y et al., 2019)further supports the CdLS diagnosis. Notably, an age-related evolution of his dysmorphic features was observed, involving particularly changes in the shape of his nose and the pronounced appearance of his eyes. Similar phenotypic changes with age have been described in other patients with CdLS (Latorre-Pellicer et al.,2021b; Jouret G et al., 2022).

Significant clinical differences are observed between his motherand his maternal aunt affected who are dizygotic twins. The maternal aunt (I-3), with a clinical CdLS score of 14, presents with classic phenotype. On the contrary, the mother (I-2) does not fall within theCdLS spectrum, even though some mild isolated features were observed. The difference in severity could be influenced by the variable level of skewing of the X-chromosome inactivation, as documented in other X-linked diseases (Sun et al., 2022). Although skewed X-chromosome inactivation was observed in both the mother (I- 2) and the maternal aunt (I-3) in blood, different inactivation rates were observed in fibroblasts. In the mother (I-2), X-chromosome inactivation was found to be random in fibroblasts, whereas in the maternal aunt (I-3) it was found to be fully skewed towards the silenced allele in blood. Furthermore, decreased AFF2 expression wasobserved in the maternal aunt (I-3) but not in the mother (I-2). Therefore, the skewed X-chromosome inactivation and the differential expression between the twins could potentially explaintheir distinct phenotypes.

The literature review showed that variants affecting the *AFF2* gene are mainly associated with the Intellectual Developmental Disorder X-Linked 109 (FRAXE, OMIM, #309548), a rare type of intellectual disability without well-characterized specific phenotypic abnormalities (Flynn et al., 1993; Mulley et al., 1995; Gecz et al., 1996). Repeated expansion of more than 200 copies of the CCG triplet of the 5′ untranslated region at the fragile X site is the main genetic cause of FRAXE syndrome. However, the literature reports other intragenic variants of the *AFF2* gene associated with clinical features that do not involve FRAXE syndrome. Strikingly, some ofthese individuals exhibit dysmorphic features that resemble those often seen in patients with CdLS, such as synophrys, anteverted nares, prominent nasal tip, cleft palate, microcephaly, clinodactyly,or abnormality of the digestive system (Valentino et al., 2021; Sahoo et al., 2011; ClinVar: VCV000986210.2; DECIPHER: 273134 and 323383). In addition, other features beyond those typically expectedin CdLS are also reported, such as macrosomia, tall stature, or macrocephaly (da Rocha et al., 2014; ClinVar: VCV000986210.2; DECIPHER: 323383). Therefore, all these features suggest that the underlying clinical features in patients carrying intragenic variants in the *AFF2* gene would not be explained by the clinical picture described for FRAXE syndrome. However, an exhaustive comparison has been proved to be difficult as the clinical description of these cases is often fragmentary. Based on these distinctions, it can be concluded that our case does not align with FRAXE syndrome but with the spectrum of CdLS.

Notably, there is a functional association between *AFF2* and theother CdLS genes, all involved in transcriptional regulation. Cohesinplays a crucial role in the recruitment of chromatin remodelers and transcription factors to promoters. It also interacts with the mediator and stabilizes its interaction with RNA polymerase II (Singh et al., 2023). Similarly, the AFF2 protein participates in transcriptional regulation, forming the superelongation complex like-2 (SEC like-2) along with other factors. This complex is known to regulate RNA polymerase II activity (Luo et al., 2012; Luo et al., 2012b), particularly by preventing its paused state (Eigenhuis et al., 2022). Therefore, the association of CdLS phenotype with a *AFF2* variant can be also explained at the molecular level.

Based on the comprehensive clinical and molecular findings of this study, the intragenic duplication in the *AFF2* gene appears to be responsible for the clinical phenotype observed in our affected patients. The phenotype, combined with the known functional role of AFF2, underscores its relevance within the CdLSspectrum. Future research should focus on analyzing larger cohorts of patients with *AFF2* variants in combination with detailed clinical phenotyping to establish definitive genotype- phenotype correlations. Furthermore, the findings from this study suggest that individuals exhibiting a CdLS spectrum clinical presentation but lacking a molecular diagnosis, should be considered for *AFF2* gene analysis.

Data availability statement

The original contributions presented in the study are publicly available. This data can be found here: Leiden Open Variation Database (LOVD), variant ID: 00456114; https://databases.lovd.nl/shared/individuals/00456114.

Ethics statement

The studies involving humans were approved by Comité de Ética de la Investigación de la Comunidad Autónoma de Aragón. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants'; legal guardians/next of kin. Written informed consent was obtained from the individual(s), and minor(s)'; legal guardian/next of kin, for the publication of any potentially identifiable images or data included in this article.

Author contributions

CL-C: Conceptualization, Formal Analysis, Investigation, Methodology, Writing–original draft, Writing–review and editing. IP: Formal Analysis, Investigation, Methodology, Writing–reviewand editing. AL-P: Investigation, Methodology, Writing–review and editing. MG-S: Investigation, Methodology, Writing–review and editing. IB: Formal Analysis, Investigation, Methodology, Writing–review and editing, Resources. PF: Formal Analysis, Investigation, Methodology, Writing–review and editing, Resources. LL: Formal Analysis, Investigation, Methodology, Writing–review and editing, Resources. MA: Investigation, Methodology, Writing–review and editing. AA-C: Investigation, Methodology, Writing–review and editing. JD: Investigation, Methodology, Writing–review and editing. LT: Investigation, Methodology, Writing–review and editing. BM: Investigation, Methodology, Writing–review and editing, Resources. LP-J: Formal Analysis, Investigation, Methodology, Resources, Writing–review and editing. PL: Investigation, Methodology, Writing–review and editing. EL: Investigation, Methodology, Writing–review and editing. JB: Investigation, Methodology, Writing–review and editing. CL: Investigation, Methodology, Writing–review and editing. FAK: Investigation, Methodology,Writing–review and editing. SK: Investigation, Methodology, Writing–review and editing. CD: Investigation, Methodology, Writing–review and editing. FRK: Formal Analysis, Funding acquisition, Investigation, Methodology, Resources, Writing–review and editing. FR: Formal Analysis, Funding acquisition, Investigation, Methodology, Resources, Writing–review and editing. BP: Formal Analysis, Investigation, Methodology, Writing–original draft, Writing–review and editing. JP: Formal Analysis, Fundingacquisition, Investigation, Methodology, Resources, Writing–original draft, Writing–review and editing.

Funding

The author(s) declare that financial support was received for theresearch, authorship, and/or publication of this article. This work was supported by the: Institute of Health Carlos III (ISCIII), Spanish Ministry of Science, Innovation and Universities Fondo de Investigación Sanitaria (FIS) [Ref. PI19/01860 and PI23/01370 to F.J.R. and J.P.], the Diputación General de Aragón-FEDER: European Social 369 Fund [Grupo de Referencia B32_23R toJ.P.] and the Generalitat de Catalunya [GenCat (PERIS Program) SLT002/16/00 to L.A.P.-J.]. C.L.-C. by a Predoctoral Fellowship from the ISCIII (FI20/00290) and M.G.-S. by a Predoctoral Fellowship from the Diputación General de Aragón. This workhas been generated within the European Reference Network on Rare Congenital Malformations and Rare Intellectual Disability (ERN-ITHACA) [EU Framework Partnership Agreement ID: 3HP- HP-FPA ERN-01-2016/739516]. Nanopore sequencing was supported by the University Hospital Essen (personal allocation to CD).

Acknowledgments

We are grateful to the patient and his family for participating in this study. The authors thank the Face2Gene tool for contributing to DeepGestalt analysis.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2024.1472543/full#supplementary-material

References

Ajmone, P. F., Rigamonti, C., Dall'Ara, F., Monti, F., Vizziello, P., Milani, D., et al.(2014). Communication, cognitive development and behavior in children with Corneliade Lange Syndrome (CdLS): preliminary results. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* 165B (3), 223–229. doi:10.1002/ajmg.b.32224

Allen, R. C., Zoghbi, H. Y., Moseley, A. B., Rosenblatt, H. M., and Belmont, J. W. (1992). Methylation of HpaII and HhaI sites near the polymorphic CAG repeat in the human androgen-receptor gene correlates with X chromosome inactivation. *Am. J. Hum. Genet.* 51 (6), 1229–1239.

Alonso-Gil, D., and Losada, A. (2023). NIPBL and cohesin: new take on a classic tale. *Trends Cell Biol. S0962-* 8924 (23), 860–871. doi:10.1016/j.tcb.2023.03.006

Asami, M., Lam, B. Y. H., Ma, M. K., Rainbow, K., Braun, S., VerMilyea, M. D., et al. (2022). Human embryonic genome activation initiates at the one-cell stage. *Cell StemCell* 3 (2), 209–216.e4. doi:10.1016/j.stem.2021.11.012

Avagliano, L., Parenti, I., Grazioli, P., Di Fede, E., Parodi, C., Mariani, M., et al. (2020).Chromatinopathies: A focus on Cornelia de Lange syndrome. *Clin. Genet.* 97 (1), 3–11. doi:10.1111/cge.13674

Bensaid, M., Melko, M., Bechara, E. G., Davidovic, L., Berretta, A., Catania, M. V.,et al. (2009). FRAXE-associated mental retardation protein (FMR2) is an RNA-binding protein with high affinity for G-quartet RNA forming structure. *Nucleic acids Res.* 37 (4), 1269–1279. doi:10.1093/nar/gkn1058

Bertini, V., Baldinotti, F., Parma, P., Tyutyusheva, N., Sepich, M., Bertolucci, G., et al. (2023). In tandem intragenic duplication of doublesex and mab- 3-related transcriptionfactor 1 (DMRT1) in an SRY-negative boy with a 46,XX disorder of sex development.*Genes* 14 (11), 2067. doi:10.3390/genes14112067

Bullich, G., Matalonga, L., Pujadas, M., Papakonstantinou, A., Piscia, D., Tonda, R., et al. (2022). Systematic collaborative reanalysis of genomic data improves diagnosticyield in neurologic rare diseases. *J. Mol. Diagn* 24 (5), 529–542. doi:10.1016/j.jmoldx.2022.02.003

Carrel, L., and Willard, H. F. (1999). An assay for X inactivation based on differential methylation at the fragile X locus,FMR1. *AmJ Med. Genet.* 64 (1), 27–30. doi:10.1002/ (SICI)1096-8628(19960712)64:1<27::AID-AJMG3>3.0.CO;2-O

Cetin, H., Wöhrer, A., Rittelmeyer, I., Gencik, M., Zulehner, G., Zimprich, F., et al.(2016). The c.65-2A>G splice site mutation is associated with a mild phenotype in Danon disease due to the transcription of normal LAMP2 mRNA. *Clin. Genet.* 90 (4),366–371. doi:10.1111/cge.12724

Cucco, F., Sarogni, P., Rossato, S., Alpa, M., Patimo, A., Latorre, A., et al. (2020). Pathogenic variants in EP300 and ANKRD11 in patients with phenotypes overlappingCornelia de Lange syndrome. *Am. J. Med. Genet. A* 182 (7), 1690–1696. doi:10.1002/ajmg.a.61611

da Rocha, N. B., dos Santos, P. A., Safatle, H. P., de Melo, R. M., Pereira, R. W., deOliveira, S. F., et al. (2014). Partial AFF2 microduplication in a patient with auditory processing disorder, emotional impairment and macrosomia. *Am. J. Med. Genet. A* 164A (12), 3206–3208. doi:10.1002/ajmg.a.36768

De Coster, W., and Rademakers, R. (2023). NanoPack2: population-scale evaluation of long-read sequencing data. *Bioinformatics* 39 (5), btad311. doi:10.1093/ bioinformatics/btad311

Eigenhuis, K. N., Somsen, H. B., and van den Berg, D. L. C. (2022). Transcription pause and escape in neurodevelopmental disorders. *Front. Neurosci.* 16, 846272. doi:10. 3389/fnins.2022.846272

Flynn, G. A., Hirst, M. C., Knight, S. J., Macpherson, J. N., Barber, J. C., Flannery, A. V., et al. (1993). Identification of the FRAXE fragile site in two families ascertained for X linked mental retardation. *J. Med. Genet.* 30 (2), 97–100. doi:10.1136/jmg.30.2.97

Garcia, P., Fernandez-Hernandez, R., Cuadrado, A., Coca, I., Gomez, A., Maqueda, M., et al. (2021). Disruption of NIPBL/Scc2 in Cornelia de Lange Syndrome provokescohesin genome-wide redistribution with an impact in the transcriptome. *Nat. Commun.* 12 (1), 4551. doi:10.1038/s41467-021-24808-z

Gecz, J., Gedeon, A. K., Sutherland, G. R., and Mulley, J. C. (1996). Identification ofthe gene FMR2, associated with FRAXE mental retardation. *Nat. Genet.* 13 (1), 105–108. doi:10.1038/ng0596-105

Gedeon, A. K., Meinänen, M., Adès, L. C., Kääriäinen, H., Gécz, J., Baker, E., et al.(1995). Overlapping submicroscopic deletions in Xq28 in two unrelated boys with developmental disorders: identification of a gene near FRAXE. *Am. J. Hum. Genet.* 56(4), 907–914.

Gil-Rodríguez, M. C., Deardorff, M. A., Ansari, M., Tan, C. A., Parenti, I., Baquero-Montoya, C., et al. (2015). *De novo* heterozygous mutations in SMC3 cause a range of Cornelia de Lange syndrome-overlapping phenotypes. *Hum. Mutat.* 36 (4), 454–462. doi:10.1002/humu.22761

Glinos, D. A., Garborcauskas, G., Hoffman, P., Ehsan, N., Jiang, L., Gokden, A., et al. (2022). Transcriptome variation in human tissues revealed by long-read sequencing. *Nature* 608 (7922), 353–359. doi:10.1038/s41586-022-05035-y

Grados, M. A., Alvi, M. H., and Srivastava, S. (2017). Behavioral and psychiatric manifestations in Cornelia de Lange syndrome. *Curr. Opin. Psychiatry* 30 (2), 92–96.doi:10.1097/YCO.0000000000000311

Gu, Y., Shen, Y., Gibbs, R. A., and Nelson, D. L. (1996). Identification of FMR2, anovel gene associated with the FRAXE CCG repeat and CpG island. *Nat. Genet.* 13 (1), 109–113. doi:10.1038/ng0596-109

Gurovich, Y., Hanani, Y., Bar, O., Nadav, G., Fleischer, N., Gelbman, D., et al. (2019). Identifying facial phenotypes of genetic disorders using deep learning. *Nat. Med.* 25 (1), 60–64. doi:10.1038/s41591-018-0279-0

Hillman, M. A., and Gecz, J. (2001). Fragile XE-associated familial mental retardation protein 2 (FMR2) acts as a potent transcription activator. *J. Hum. Genet.* 46 (5), 251–259. doi:10.1007/s100380170074

Huisman, S., Mulder, P. A., Redeker, E., Bader, I., Bisgaard, A. M., Brooks, A., et al.(2017). Phenotypes and genotypes in individuals with SMC1A variants. *Am. J. Med. Genet. A* 173 (8), 2108–2125. doi:10.1002/ajmg.a.38279

Izumi, K. (2016). Disorders of transcriptional regulation: an emerging category of multiple malformation syndromes. *Mol. Syndromol.* 7 (5), 262–273. doi:10.1159/000448747

Jouret, G., Heide, S., Sorlin, A., Faivre, L., Chantot-Bastaraud, S., Beneteau, C., et al. (2022). Understanding the new BRD4-related syndrome: clinical and genomic delineation with an international cohort study. *Clin. Genet.* 102 (2), 117–122. doi:10.1111/cge.14141

Kline, A. D., Krantz, I. D., Sommer, A., Kliewer, M., Jackson, L. G., FitzPatrick, D. R.,et al. (2007). Cornelia de Lange syndrome: clinical review, diagnostic and scoring systems, and anticipatory guidance. *Am. J. Med. Genet. A* 143A (12), 1287–1296. doi:10. 1002/ajmg.a.31757

Kline, A. D., Moss, J. F., Selicorni, A., Bisgaard, A. M., Deardorff, M. A., Gillett, P. M.,et al. (2018). Diagnosis and management of Cornelia de Lange syndrome: first international consensus statement. *Nat. Rev. Genet.* 19 (10), 649–666. doi:10.1038/s41576-018-0031-0

Krab, L. C., Marcos-Alcalde, I., Assaf, M., Balasubramanian, M., Andersen, J. B., Bisgaard, A. M., et al. (2020). Delineation of phenotypes and genotypes related to cohesin structural protein RAD21. *Hum. Genet.* 139 (5), 575–592. doi:10.1007/s00439-020-02138-2

Łabno, A., Tomecki, R., and Dziembowski, A. (2016). Cytoplasmic RNA decay pathways - enzymes and mechanisms. *Biochim. Biophys. Acta* 1863 (12), 3125–3147. doi:10.1016/j.bbamcr.2016.09.023

Latorre-Pellicer, A., Ascaso, Á., Trujillano, L., Gil-Salvador, M., Arnedo, M., Lucia-Campos, C., et al. (2020). Evaluating Face2Gene as a Tool to Identify Cornelia de Lange Syndrome by Facial Phenotypes. *Int. J. Mol. Sci.* 21 (3), 1042. doi:10.3390/ijms21031042

Latorre-Pellicer, A., Gil-Salvador, M., Parenti, I., Lucia-Campos, C., Trujillano, L.,Marcos-Alcalde, I., et al. (2021a). Clinical relevance of postzygotic mosaicism in Cornelia de Lange syndrome and purifying selection of NIPBL variants in blood. *Sci. Rep.* 11 (1), 15459. doi:10.1038/s41598-021-94958-z

Latorre-Pellicer, A., Ascaso, Á., Lucia-Campos, C., Gil-Salvador, M., Arnedo, M., Antoñanzas, R., et al. (2021b). Things are not always what they seem: From Cornelia de Lange to KBG phenotype in a girl with genetic variants in NIPBL and ANKRD11. *Mol. Genet. Genomic Med.* 9 (11), e1826. doi:10.1002/mgg3.1826

Leger, A., and Leonardi, T. (2019). pycoQC, interactive quality control for Oxford Nanopore Sequencing. *J. Open Source Softw.* 4 (34), 1236. doi:10.21105/joss.01236

Lev, M. G., Yearim, A., and Ast, G. (2015). The alternative role of DNA methylation in splicing regulation. *Trends Genet.* 31 (5), 274–280. doi:10.1016/j.tig.2015.03.002

Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34 (18), 3094–3100. doi:10.1093/bioinformatics/bty191

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression datausing real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 25 (4), 402–408. doi:10.1006/meth.2001.1262

Luo, Z., Lin, C., Guest, E., Garrett, A. S., Mohaghegh, N., Swanson, S., et al. (2012b). The super elongation complex family of RNA polymerase II elongation factors: gene target specificity and transcriptional output. *Mol. Cell Biol.* 32 (13), 2608–2617. doi:10.1128/MCB.00182-12

Luo, Z., Lin, C., and Shilatifard, A. (2012). The super elongation complex (SEC) familyin transcriptional control. *Nat. Rev. Mol. Cell Biol.* 13 (9), 543–547. doi:10.1038/nrm3417

Luque, J., Mendes, I., Gómez, B., Morte, B., López de Heredia, M., Herreras, E., et al. (2022). CIBERER: Spanish national network for research on rare diseases: a highly productive collaborative initiative. *Clin. Genet.* 101 (5-6), 481–493. doi:10.1111/cge.14113

Mills, J. A., Herrera, P. S., Kaur, M., Leo, L., McEldrew, D., Tintos-Hernandez, J. A.,et al. (2018). NIPBL+/- haploinsufficiency reveals a constellation of transcriptome disruptions in the pluripotent and cardiac states. *Sci. Rep.* 8 (1), 1056. doi:10.1038/s41598-018-19173-9

Mulley, J. C., Yu, S., Loesch, D. Z., Hay, D. A., Donnelly, A., Gedeon, A. K., et al. (1995). FRAXE and mental retardation. *J. Med. Genet.* 32 (3), 162–169. doi:10.1136/jmg.32.3.162

Newman, S., Hermetz, K. E., Weckselblatt, B., and Rudd, M. K. (2015). Next-generation sequencing of duplication CNVs reveals that most are tandem and some create fusion genes at breakpoints. *Am. J. Hum. Genet.* 96 (2), 208–220. doi:10.1016/j.ajhg.2014.12.017

Parenti, I., Diab, F., Gil, S. R., Mulugeta, E., Casa, V., Berutti, R., et al. (2020). MAU2 and NIPBL Variants Impair the Heterodimerization of the Cohesin Loader Subunits and Cause Cornelia de Lange Syndrome. *Cell Rep.* 31 (7), 107647. doi:10.1016/j.celrep.2020.107647

Parenti, I., and Kaiser, F. J. (2021). Cornelia de Lange Syndrome as Paradigm of Chromatinopathies. *Front. Neurosci.* 15, 774950. doi:10.3389/fnins.2021. 774950

Parenti, I., Teresa-Rodrigo, M. E., Pozojevic, J., Ruiz Gil, S., Bader, I., Braunholz, D.,et al. (2017). Mutations in chromatin regulators functionally link Cornelia de Lange syndrome and clinically overlapping phenotypes. *Hum. Genet.* 136 (3), 307–320. doi:10.1007/s00439-017-1758-y

Peck, S. A., Hughes, K. D., Victorino, J. F., and Mosley, A. L. (2019). Writing a wrong:coupled RNA polymerase II transcription and RNA quality control. *Wiley Interdiscip. Rev. RNA* 10 (4), e1529. doi:10.1002/wrna.1529

Pié, J., Puisac, B., Hernández-Marcos, M., Teresa-Rodrigo, M. E., Gil-Rodríguez, M., Baquero-Montoya, C., et al. (2016). Special cases in Cornelia de Lange syndrome: The Spanish experience. *Am. J. Med. Genet. C Semin. Med. Genet.* 172 (2), 198–205. doi:10.1002/ajmg.c.31501

Richardson, M. E., Chong, H., Mu, W., Conner, B. R., Hsuan, V., Willett, S., et al. (2019). DNA breakpoint assay reveals a majority of gross duplications occur in tandemreducing VUS classifications in breast cancer predisposition genes. *Genet. Med.* 21 (3),683–693. doi:10.1038/s41436-018-0092-7

Sahoo, T., Theisen, A., Marble, M., Tervo, R., Rosenfeld, J. A., Torchia, B. S., et al.(2011). Microdeletion of Xq28 involving the AFF2 (FMR2) gene in two unrelated males with developmental delay. *Am. J. Med. Genet. A* 155A (12), 3110–3115. doi:10.1002/ajmg.a.34345

Selicorni, A., Mariani, M., Lettieri, A., and Massa, V. (2021). Cornelia de Lange Syndrome: From a Disease to a Broader Spectrum. *Genes (Basel)* 12 (7), 1075. doi:10.3390/genes12071075

Shangguan, H., and Chen, R. (2022). Phenotypes of Cornelia de Lange syndrome caused by non-cohesion genes: Novel variants and literature review. *Front. Pediatr.* 10, 940294. doi:10.3389/fped.2022.940294

Singh, A. K., Chen, Q., Nguyen, C., Meerzaman, D., and Singer, D. S. (2023). Cohesin regulates alternative splicing. *Sci. Adv.* 9 (9), eade3876. doi:10.1126/sciadv.ade3876

Smolka, M., Paulin, L. F., Grochowski, C. M., Horner, D. W., Mahmoud, M., Behera, S., et al. (2024). Publisher Correction: detection of mosaic and population-level structural variants with Sniffles2. *Nat. Biotechnol.* 42, 1616. doi:10.1038/s41587-024-02141-2

Stettner, G. M., Shoukier, M., Hoger, C., Brockmann, K., and Auber, B. (2011). Familial intellectual disability and autistic behavior caused by a small FMR2 gene deletion. *Am. J. Med. Genet.* 155A (8), 2003–2007. doi:10.1002/ajmg.a.34122

Sun, Z., Fan, J., and Wang, Y. (2022). X-chromosome inactivation and related diseases. *Genet. Res. (Camb).* 2022, 1–8. doi:10.1155/2022/1391807

Takahashi, S., Tanaka, R., Okano, S., Okayama, A., Suzuki, N., and Azuma, H. (2015). Characterization of intragenic tandem duplication in the PAFAH1B1 gene leading to isolated lissencephaly sequence. *Mol. Cytogenet* 8, 84. doi:10.1186/s13039-015-0186-8

Teresa-Rodrigo, M. E., Eckhold, J., Puisac, B., Dalski, A., Gil-Rodríguez, M. C., Braunholz, D., et al. (2014). Functional characterization of NIPBL physiological splice variants and eight splicing mutations in patients with Cornelia de Lange syndrome. *Int. J. Mol. Sci.* 15 (6), 10350–10364. doi:10.3390/ijms150610350

Valentino, F., Bruno, L. P., Doddato, G., Giliberti, A., Tita, R., Resciniti, S., et al. (2021). Exome sequencing in 200 intellectual disability/autistic patients: new candidates and atypical presentations. *Brain Sci.* 11 (7), 936. doi:10.3390/brainsci11070936

Weiss, F. D., Calderon, L., Wang, Y. F., Georgieva, R., Guo, Y., Cvetesic, N., et al.(2021). Neuronal genes deregulated in Cornelia de Lange Syndrome respond to removal and re-expression of cohesin. *Nat. Commun.* 12 (1), 2919. doi:10.1038/s41467-021- 23141-9

Yuan, B., Pehlivan, D., Karaca, E., Patel, N., Charng, W. L., Gambin, T., et al. (2015). Global transcriptional disturbances underlie Cornelia de Lange syndrome and related phenotypes. *J. Clin. Invest* 125 (2), 636–651. doi:10.1172/JCI77435

Zepeda-Mendoza, C. J., Cousin, M. A., Basu, S., Jenkinson, G., Oliver, G., Pittock, S. T., et al. (2019). An intragenic duplication of TRPS1 leading to abnormal transcripts and causing trichorhinophalangeal syndrome type I. *Cold Spring Harb. Mol. Case Stud.* 5 (6), a004655. doi:10.1101/mcs.a004655

Discusión

El diagnóstico de los trastornos del neurodesarrollo (TND) es habitualmente un proceso complejo que exige un enfoque multidisciplinar [58]. Los avances en las ciencias ómicas y la identificación de nuevos genes, están ayudando a entender y describir, de manera más precisa, la etiología de estas enfermedades [107]. Sin embargo, aunque el uso de las nuevas tecnologías y el *big data* facilita una comprensión más profunda, sigue habiendo casos donde el abordaje diagnóstico es complicado.

Los TND presentan un amplio espectro de síntomas, así como distintos patrones de progresión y edades de aparición, que dificultan su identificación clínica. El solapamiento fenotípico entre ciertos síndromes, como el síndrome Cornelia de Lange (SCdL, OMIM #122470, #300590, #610759, #614701, #300882) y el síndrome de KBG (SKBG, OMIM #148050), hace que, en muchas ocasiones, establecer un diagnóstico diferencial no sea fácil (160,164). Además, la aparición de nuevas entidades nosológicas, como el síndrome de PACS1 (SHMS, OMIM #615009), recientemente caracterizado, hace necesaria la reclasificación y descripción clínica detallada de los pacientes.

Una vez establecido el diagnóstico clínico, es fundamental identificar la base molecular. En algunos casos, la heterogeneidad causal implica que diferentes genes puedan originar cuadros clínicos similares [63]. En otros casos, las variantes patogénicas pueden quedar fuera del alcance de las técnicas diagnósticas disponibles, especialmente cuando afectan a regiones complejas del genoma. Hasta el momento, las variantes se detectaban en el 1,5% del genoma codificante [214], sin embargo, en la actualidad, la exploración del 98,5% restante, está permitiendo resolver numerosos casos previamente sin diagnóstico [12]. Un ejemplo destacado es la reciente identificación de un ARN no codificante (*RNU4-2 snRNA*) asociado a un porcentaje significativo de pacientes con TND [215].

La Ontología del Fenotipo Humano (HPO), junto con la actualización constante de datos moleculares, está mejorando la clasificación fenotípica y genotípica, incrementando la precisión diagnóstica. Aunque existen múltiples recursos dedicados a enfermedades genéticas raras (EGRs), sigue siendo necesaria una integración y unificación de estos datos para optimizar su clasificación [216,217]. Los códigos HPO han contribuido significativamente a la descripción clínica, concretando los fenotipos según el grado de afectación [19,218–220]. Esta terminología, unida a diversas iniciativas destinadas a estandarizar el lenguaje médico, facilita la diferenciación entre grupos de enfermedades heterogéneas, mejorando la identificación de patrones de patogenicidad.

A pesar de los avances en las herramientas genómicas y analíticas, aproximadamente el 60% de los individuos con TND carecen de un diagnóstico genético claro [221]. En el caso del SCdL, y según los criterios de inclusión utilizados, se estima que

aproximadamente entre un 20-40% de los pacientes continúan sin diagnóstico confirmado [149]. En la cohorte española, hay un 20% de individuos en los que no se conoce la causa (datos no publicados). Este trabajo presenta los problemas de diagnóstico genético de tres pacientes con SCdL, destacándose las limitaciones encontradas como: la necesidad de la reevaluación de los casos debido a cambios fenotípicos, la detección de variantes complejas mediante técnicas de aparición reciente y la identificación de nuevos genes asociados al síndrome. Además, se incluye la primera revisión integral del síndrome de PACS1, profundizando en su diagnóstico diferencial y explorando posibles enfoques terapéuticos.

A pesar de los desafíos encontrados, el diagnóstico preciso de los TND es el primer paso para optimizar la atención al paciente. También es fundamental estudiar la recurrencia de las variantes genéticas para ofrecer un asesoramiento adecuado. Una vez identificada la patología, los especialistas pueden planificar intervenciones oportunas y evaluar posibles terapias. Además, las familias pueden acceder a asociaciones de pacientes, lo que les hace sentirse más acompañadas. En definitiva, obtener un diagnóstico representa el primer paso hacia la mejora de la calidad de vida de los pacientes y sus familias.

Investigación de nuevas entidades nosológicas

El análisis de las similitudes y diferencias entre los diversos TND, desde una perspectiva fenotípica y etiológica, es fundamental para mejorar la comprensión y manejo clínico. Este enfoque comparativo, permite identificar aspectos clave como la prevalencia, los rasgos distintivos y las alteraciones moleculares subyacentes. El solapamiento fenotípico de algunos pacientes, hace cuestionarse sobre los mecanismos responsables de la convergencia, a pesar de las diferencias existentes a nivel molecular.

El síndrome de PACS1 y el SCdL ejemplifican la complejidad del diagnóstico diferencial de los TND. En el año 2012 se describió una nueva entidad nosológica, el síndrome de PACS1, también conocido como el síndrome *Schuurs-Hoeijmakers* (SHMS, OMIM #615009) [212]. Su causa genética se asoció a una variante recurrente en el gen *PACS1* (c.607C>T; p.R203W), que generaba una ganancia de función a nivel proteico [212,222]. A pesar de la disparidad funcional entre esta proteína y las del SCdL, algunos individuos con síndrome de PACS1 fueron inicialmente clasificados como SCdL. Este fue el caso de los dos primeros pacientes holandeses donde se describió el cuadro [212], y se repitió en las cohortes española y estadounidense [149], en las que se identificaron individuos con este síndrome.

Esto ha llevado a la realización de una revisión del síndrome de PACS1, en la que se aborda en detalle sus manifestaciones clínicas, bases moleculares y aspectos relacionados con el diagnóstico diferencial. Se trata de la primera revisión centrada en este trastorno desde una perspectiva molecular, con el objetivo de establecer relaciones entre las vías patológicas implicadas. En ella se describen cinco TND relacionados fenotípicamente con el síndrome, de los cuales dos están vinculados al mecanismo molecular subyacente, y tres se correlacionan con una clínica solapante.

Aunque las características clínicas de este síndrome están bien definidas, las evidencias sobre sus mecanismos patológicos siguen siendo limitadas. En los últimos años, ha mejorado la comprensión funcional de la proteína PACS1, destacándose su papel esencial en el tráfico de proteínas entre los endosomas y el aparato de Golgi [223]. Además, se sugiere que podría desempeñar otras funciones relacionadas con las proteínas que transporta. PACS1 también es clave en la homeostasis celular al interactuar con proteínas vinculadas a la apoptosis, la estabilidad genómica y el flujo de calcio en el retículo endoplasmático [224–226]. Sin embargo, su estructura sigue siendo desconocida, lo que limita la comprensión del efecto de las variantes en su función.

El primer artículo presentado (Trabajo 1) destaca la estrecha correlación clínica y molecular entre los síndromes de PACS1, PACS2 (DEE66, OMIM #618067) y WDR37 (NOCGUS, OMIM #618652). La interconexión funcional de las proteínas implicadas, podría explicar el solapamiento fenotípico de estas tres patologías. PACS1 y PACS2 intervienen en el tráfico vesicular de proteínas a nivel intracelular [224]. PACS1 facilita la carga en la red trans-Golgi, mientras que PACS2 participa en el transporte vesicular desde el Golgi al retículo endoplasmático, las membranas celulares y las mitocondrias [227,228]. Además, ambas proteínas se relacionan con la regulación del proceso de apoptosis [229,230]. Por otro lado, la interacción entre PACS1 y WDR37 se asocia con el proceso de regulación del flujo de calcio, ya que ambas proteínas forman un complejo que modula la liberación de este ion entre el retículo endoplasmático y el citosol [226].

A nivel clínico, los tres síndromes comparten características faciales similares, junto con discapacidad intelectual (HP:0001249), retraso del desarrollo neurológico (HP:0012758) y convulsiones (HP:0001250) [231–236]. La relación funcional entre las proteínas afectadas y el solapamiento fenotípico, sugiere que estos tres trastornos podrían agruparse dentro de un único cuadro de amplio espectro. Recientemente, se ha propuesto un nuevo grupo de patologías causadas por defectos en el tráfico celular, que se acompañan de alteraciones del neurodesarrollo y anomalías esqueléticas [237,238], y que bien podría incluir a estos tres trastornos. Sin embargo, es necesario investigar más a fondo la relación entre la alteración del tráfico celular y los fenotipos, a fin de poder comprender y clasificar mejor estas patologías.

La revisión también compara los síndromes de PACS1, CdL, Kabuki (OMIM #147920, #300867) y Coffin-Siris (OMIM #135900). En todos ellos se observa un solapamiento fenotípico que incluye discapacidad intelectual (HP:0001249), autismo (HP:0000717), retardo global del desarrollo (HP:0001263), retraso del lenguaje y la comunicación (HP:0000750), hipotonía (HP:0001252) y convulsiones (HP:0001250). Además, aparece un dismorfismo facial común, con características solapantes como cejas arqueadas (HP:0002553) y bermellón del labio superior fino (HP:0000219). Aunque aún no se han identificado vías moleculares comunes, las proteínas implicadas desempeñan funciones clave y específicas durante el neurodesarrollo. Por todo ello, es fundamental investigar su relación a nivel molecular para comprender los mecanismos responsables del solapamiento clínico observado.

Aunque aún existen aspectos por determinar, la identificación del síndrome de PACS1 ha supuesto una mejora significativa en la atención de estos pacientes. Su diagnóstico preciso permite orientar los protocolos de intervención de manera más eficaz. Uno de los rasgos distintivos de este síndrome, es su naturaleza neurodegenerativa y regresiva, lo que subraya la urgencia de un diagnóstico temprano y la implementación de terapias capaces de ralentizar este deterioro. Además, la presencia de la misma variante genética en la mayoría de los pacientes, facilita el desarrollo de terapias específicas capaces de revertir la condición. Una propuesta que ha demostrado ser efectiva en modelos de ratón, es el uso de oligonucleótidos antisentido [72,73]. La posibilidad de desarrollar tratamientos específicos, ofrece perspectivas prometedoras para el abordaje terapéutico de esta enfermedad.

Seguimiento longitudinal y reevaluación de datos genómicos

Una cuestión clave en el campo de las enfermedades raras (ERs), es cómo optimizar la precisión del diagnóstico. Diversos factores pueden dificultar la detección de variantes patogénicas durante la evaluación de los datos genómicos. Entre ellos se encuentran las limitaciones inherentes a las técnicas del método diagnóstico, el desconocimiento de los genes implicados en la patología o la exclusión inapropiada de variantes durante el proceso de análisis [239]. Para abordar estos desafíos, se está implementando la reevaluación de los datos genómicos unido al seguimiento longitudinal de los pacientes [214,239,240,240–243]. Este enfoque reciente requiere de una comunicación iterativa entre familias y profesionales clínicos y de laboratorio.

El segundo artículo (Trabajo 2) resalta la relevancia de implementar este nuevo enfoque en el proceso diagnóstico. En su primera evaluación, en 2012, la paciente fue clasificada, con una puntuación clínica de 10, como SCdL no clásico de acuerdo a los

criterios de Kline et al. [111]. Esta clasificación se basó en las manifestaciones clínicas de la paciente, como sinofridia (HP:0000664), cejas arqueadas (HP:0002553), comisuras de la boca de oblicuidad descendente (HP:0002714), retardo global del desarrollo (HP:0001263), manos (HP:0200055) y pies (HP:0001773) pequeños, quinto dedo de la mano corto (HP:0009237) e hirsutismo (HP:0001007). En ese momento, se realizó el diagnóstico molecular mediante secuenciación Sanger del gen *NIPBL*, identificando la variante *NIPBL*:NM_133433.3:c.7553A>G, p.(Asp2518Gly), que fue clasificada como probablemente patogénica. Los hallazgos clínicos y la variante encontrada, parecían respaldar el diagnóstico del SCdL. Sin embargo, la evolución fenotípica observada durante el seguimiento longitudinal de la paciente, llevó a la necesidad de realizar una reevaluación del caso. La aparición de nuevos rasgos faciales, como macrodoncia (HP:0001572), cara triangular (HP:0000325) y nariz bulbosa (HP:0000414), características distintivas del SKBG, sugirió la importancia de realizar un análisis genético más amplio. A través de un panel de secuenciación profunda de genes específicos, se pudo examinar el gen *ANKRD11*. Los resultados confirmaron la variante previamente identificada en el gen *NIPBL*, y revelaron una nueva variante sin sentido en el gen *ANKRD11* [*ANKRD11*:NM_001256183.1:c.2512C>T, p.(Arg838*)]. Según los criterios ACMG/AMP [46], el análisis de ambas variantes utilizando predictores de patogenicidad, clasificó la variante de *NIPBL* como posiblemente patogénica (PM2, PP2, and PP3), mientras que la variante en *ANKRD11* fue clasificada como patogénica (PVS1, PM2, PP3 y PP5). El estudio de segregación familiar mostró la presencia de la variante del gen *NIPBL* en la madre y dos tíos maternos, mientras que la variante en el gen *ANKRD11* solo fue detectada en la madre. Una evaluación clínica posterior reveló que la madre tenía rasgos faciales dismórficos similares a los de la hija, mientras que los tíos no estaban afectados. Todo ello sugirió que la variante responsable del cuadro clínico correspondía a la identificada en el gen *ANKRD11*.

La evolución fenotípica dependiente de la edad en individuos portadores de variantes en *ANKRD11*, ya ha sido descrita. Se han reportado casos de pacientes diagnosticados con SCdL durante la primera infancia, que fueron reclasificados como SKBG tras una reevaluación fenotípica [164]. La relación entre estos síndromes se refleja a nivel molecular, ya que ambos pertenecen al grupo de las cromatinopatías y presentan alteración en la de regulación transcripcional. Debido a la identificación de variante en *ANKRD11* en individuos con un fenotipo no clásico de SCdL [159,188], y el solapamiento fenotípico entre los dos síndromes [164], resulta fundamental incluir este gen en las pruebas genéticas de los pacientes que presentan características clínicas compatibles con el SCdL. Además, estudios recientes han demostrado que un porcentaje significativo (0.68%) de individuos con TND presentan variantes en *ANKRD11* [215], lo que sugiere la necesidad de considerar este gen en la evaluación de estos trastornos.

El uso de la inteligencia artificial (IA) se está integrando cada vez más en la práctica clínica. Un ejemplo destacado es la aplicación *Face2Gene*, diseñada para el análisis y la comparación de rasgos faciales mediante los algoritmos *DeepGestalt* y *GestaltMatcher* [22,184]. Esta herramienta ha demostrado ser útil para interpretar el dismorfismo facial asociado al SCdL [183]. El estudio retrospectivo con *Face2Gene* de este caso, ha respaldado el cambio fenotípico observado durante la evaluación clínica. Por tanto, estos resultados subrayan la relevancia de utilizar la IA como apoyo para contextualizar y justificar las manifestaciones fenotípicas identificadas.

Otro desafío importante, es lograr una clasificación precisa de las variantes genéticas identificadas [244]. La predicción de la patogenicidad ha mejorado significativamente gracias a la aparición de nuevas herramientas que integran la información disponible. Un ejemplo de ello es *VarSome*, una plataforma que analiza múltiples bases de datos de variantes genéticas humanas, incluyendo herramientas predictivas y criterios de interpretación genética, lo que la convierte en un recurso de gran utilidad para investigadores y clínicos [245].

Tanto este caso, como otros previamente publicados en los que se observa una evolución fenotípica, plantean la necesidad de reflexionar sobre el enfoque diagnóstico de los TND. Como resultado de este trabajo, se concluye que es fundamental establecer protocolos de reevaluación, especialmente en casos donde se detecta un cambio fenotípico del paciente, o en los que todavía no se ha logrado identificar la variante patogénica. Además, la incorporación de nuevas técnicas basadas en IA, como herramientas complementarias en el proceso diagnóstico, contribuye a mejorar su precisión y validez.

Variación del número de copias: detección e interpretación

La identificación de las variantes patológicas asociadas al SCdL sigue siendo un desafío debido a su gran heterogeneidad, dejando alrededor del 20% de pacientes sin diagnóstico genético. Esta dificultad no viene dada solo por el elevado número de genes casuales, sino también por los distintos tipos de variantes que se pueden encontrar. Se han descrito desde cambios de un solo nucleótido (SNVs), inserciones o deleciones de pequeño tamaño (INDELs), hasta variaciones en el número de copias (CNVs) de la secuencia génica [111,149]. Actualmente, la puesta en marcha de nuevas técnicas diagnósticas, está facilitando el descubrimiento de nuevas CNVs asociadas al síndrome, especialmente en el gen *NIPBL* [140,246,247]. Se estima que este tipo de variantes son la causa de aproximadamente el 20% de los TND, lo que hace que su exploración sea necesaria [55,248]. Sin embargo, su interpretación clínica sigue siendo compleja, debido a su impacto variable a nivel proteico

[249]. Además, no todas las CNVs tienen efectos patogénicos, algunas son benignas o polimorfismos comunes en la población general [250].

En el caso concreto del gen causal *HDAC8*, la mayoría de las variantes son SNVs o INDELs [156,251,252]. Sin embargo, también se ha publicado algún caso de CNVs, especialmente deleciones intragénicas [156,197]. El tercer artículo (Trabajo 3) presenta la primera caracterización de una duplicación intragénica en este gen, identificada en una niña de ocho años con fenotipo clásico. Aunque previamente ya se había informado de una variante de este tipo, su efecto patogénico no había sido analizado [156]. Este trabajo incluye la caracterización funcional de la nueva CNV mediante técnicas *in silico* de dinámica molecular, que ayudan a comprender el mecanismo de acción e impacto clínico de la variante.

El array de hibridación genómica comparativa (aCGH) es la técnica estándar para la detección de CNVs [253,254]. En este estudio, la aplicación de esta prueba permitió identificar una duplicación intragénica *de novo* del exón 10 del gen *HDAC8* [arr(GRCh38) Xq13.1(72,340,096–72,389,392)×3]. Es importante destacar la alta profundidad de lectura y resolución del array utilizado (8x60K). La mejora de estos parámetros ha permitido que esta técnica detecte variaciones de tamaño pequeño o mediano que antes podían pasar desapercibidas [255]. Además, las técnicas de secuenciación del exoma (WES) y genoma completo (WGS), también han demostrado ser útiles en la detección de este tipo de variantes [256,257]. Sin embargo, su sensibilidad para identificar CNVs de gran tamaño o estructuralmente complejas, especialmente aquellas localizadas en regiones no codificantes o ricas en secuencias repetitivas, resulta limitada [256,258]. La aplicación de un panel de genes específicos para el SCdL (>1000×) no reveló, inicialmente, ninguna variante genética asociada al fenotipo de la paciente. Sin embargo, una vez identificada la duplicación mediante el aCGH, la aplicación de un nuevo algoritmo de análisis capaz de detectar CNVs, permitió confirmar la duplicación. Este caso plantea una reflexión sobre la capacidad diagnóstica de la técnica empleada, así como la relevancia de seleccionar adecuadamente el algoritmo para el análisis de datos.

Aunque resulta fácil de entender como la haploinsuficiencia que generan las CNVs de deleción da lugar a la enfermedad [259], el impacto de las variaciones por duplicación es más complejo de interpretar. En este caso, la duplicación del exón 10, hacía prever su posible repetición en el ARN mensajero [260]. Sin embargo, en lugar de la repetición esperada, se encontró una inserción de 96 pares de bases correspondiente al intrón 9, que se justificaba por la presencia de sitios de empalme no canónicos en la secuencia intrónica insertada [197,261]. Esta secuencia adicional en el ARN mensajero, provocaba una alteración del marco de lectura, dando lugar a la aparición de un codón de parada prematuro. La traducción de este tránscrito, generaba una proteína truncada con un tamaño reducido de

342 aminoácidos, en lugar de los 377 presentes en la de tipo salvaje. En general, las variantes que introducen un codón de *stop* se consideran patológicas, ya que interrumpen la síntesis proteica, dando lugar a un polipéptido acortado habitualmente disfuncional [262].

El conocimiento de la estructura y función de la proteína HDAC8 facilita los estudios funcionales de sus variantes [263–266]. A diferencia de otras proteínas causales del síndrome, su naturaleza enzimática ofrece la posibilidad de realizar ensayos de actividad. Sin embargo, en este caso, debido a la falta de material biológico, se recurrió a metodologías de modelado estructural y simulación de dinámica molecular. Estas estrategias *in silico* están ganando relevancia, ya que permiten evaluar aspectos que no pueden ser estudiados *in vivo* [267]. En este trabajo, se analizaron detalladamente la conformación de la proteína truncada. El modelado estructural mostró una pérdida de la hélice α del extremo C-terminal, que parecía provocar un cambio en la carga electrostática superficial y un desplazamiento del bucle cercano al centro activo. Posteriormente, estos efectos fueron confirmados por dinámica molecular, concluyendo que los cambios estructurales comprometían la unión correcta de la enzima al sustrato. La accesibilidad reducida del sustrato al centro activo, sugería una disminución de la actividad enzimática, y ayudaba a entender el desarrollo del cuadro clínico de la paciente.

El aumento en la detección de CNVs asociadas al SCdL, destaca la necesidad de implementar un enfoque diagnóstico más amplio y dirigido. Debido a las limitaciones técnicas de ciertos métodos diagnósticos, es probable que exista un mayor número de CNVs asociadas a este síndrome, que aún no han sido identificadas. En este sentido, se propone un enfoque analítico más específico que considere este tipo de variantes, más allá del gen *NIPBL*.

Descubrimiento de nuevos genes

El origen genético del SCdL está principalmente relacionado con variantes en genes vinculados al complejo cohesina [97,111,221]. Desde el año 2004, se han identificado ocho genes causales, lo que ha mejorado la comprensión de los mecanismos patogénicos del cuadro, sin embargo, otros genes podrían estar implicados. En 2004 se descubrió el primer gen causal, *NIPBL*, lo que marcó un antes y un después en el diagnóstico genético del síndrome [133,134,143,268,269]. En 2007, se reportaron variantes de los genes *SMC1A* y *SMC3*, en pacientes con manifestaciones clínicas más leves [202]. Unos años más tarde, en 2012, se identificó el gen *HDAC8*, asociado tanto a fenotipos clásicos como no clásicos [157]; y el gen *RAD21*, relacionado principalmente con manifestaciones clínicas de menor gravedad [155]. Más recientemente, se añadieron los genes *ANKRD11* [159] y *BRD4* [158], ambos con un papel clave en la regulación de la cromatina. Finalmente, en 2020, se identificó el gen

MAU2, cuyo descubrimiento fue posible al conocer la interacción funcional entre las proteínas NIPBL y MAU2 [128]. Esta cronología refleja cómo los avances de las técnicas diagnósticas, han facilitado una mejor comprensión de las bases moleculares del síndrome (Figura 5) [97,103,111,149,162]. La identificación de nuevos genes ha venido también acompañado de una mayor variabilidad clínica, que ha llevado al uso de un término más amplio, Espectro Cornelia de Lange (ECdL), en el que quedarían incluidos todos los fenotipos relacionados con el síndrome [97,111].

El último artículo (Trabajo 4) aborda el desafío de caracterizar un nuevo gen causal de una ER. Se describe el diagnóstico genético del primer paciente de la cohorte española del SCdL, tras quince años de estudio. Su fenotipo clásico, evidente a los cuatro años, y mantenido durante su seguimiento longitudinal, justifica el estudio molecular realizado. Gradualmente, la técnica inicial de Sanger fue reemplazada por tecnologías más avanzadas, como la WGS. El uso de diversas metodologías, permitió confirmar la existencia de una duplicación intragénica en el gen *AFF2*. El retraso de este hallazgo se debió, no solo al hecho de que se trataba de un gen causal diferente, sino también a la naturaleza compleja de la variante encontrada. Fue un aCGH de alta resolución (1×400K) la técnica que permitió revelar la duplicación intragénica de los exones 10, 11 y 12 del gen *AFF2* [arr(GRCh38) Xq28(148,934,576 –148,960,040)x2)]. Este descubrimiento fue posteriormente confirmado mediante el reanálisis de los datos obtenidos con el panel de genes (>1000×) y la WES. Además, el estudio de segregación familiar mostró la presencia de la variante en la madre, con una afectación muy leve, y en la tía materna, quien presentaba un fenotipo similar al del paciente. Las diferencias clínicas entre los portadores, parecían deberse a que el gen estaba ligado al cromosoma X.

A nivel funcional, la conexión entre el nuevo gen causal y los previamente asociados al SCdL, podría deberse a la existencia de un mecanismo de acción común relacionado con la regulación de la expresión génica. La mayoría de los genes identificados hasta la fecha, se asocian a componentes relacionados al complejo cohesina, el cual desempeña un papel esencial en el reclutamiento y unión de factores de transcripción a zonas potenciadoras (*enhancers)* y promotoras [124,270,271]. Por su parte, *AFF2* codifica el segundo miembro de la familia AF4/FMR2 (AFF2), que actúa como un activador transcripcional [272,273]. Esta proteína, junto a otros factores, participa en la regulación de la ARN polimerasa II, al modular su actividad e impedir que entre en pausa [274–276]. Estudios en modelos de ratones *knock out* de *Aff2*, han evidenciado un deterioro significativo en las capacidades de aprendizaje y memoria, así como en el procesamiento cognitivo [277]. Experimentos en neuronas, donde *AFF2* ha sido silenciado, muestran una actividad sináptica reducida, hallazgo que resalta el papel de este gen en la regulación de la actividad neuronal [278]. Estas evidencias sugieren que la afectación de *AFF2*, podrían

ser compatible con un cuadro clínico similar al SCdL, probablemente debido a una disrupción compartida en los mecanismos de regulación transcripcional.

En la literatura, el gen *AFF2* se ha vinculado al síndrome de FRAXE (OMIM #309548), un tipo de discapacidad intelectual causada por la expansión repetida de más de 200 copias del triplete CCG de su región 5' no codificante [279,280]. Sin embargo, el fenotipo asociado a otras variantes intragénicas, se alejaría del síndrome de FRAXE, mostrando un solapamiento parcial con el SCdL. Así, se encuentran rasgos dismórficos como sinofridia (HP:0000664), narinas antevertidas (HP:0000463), paladar hendido (HP:0000175) o clinodactilia (HP:0030084) típicos del SCdL [281,282, ClinVar:VCV000986210.2, DECIPHER: 273134 y 323383]. Pero también aparecen peculiaridades, como talla alta (HP:0000098) o macrocefalia (HP:0000256) [283, ClinVar: VCV000986210.2, DECIPHER: 323383]. Es importante señalar que la descripción incompleta de muchos de estos casos, algunos únicamente reportados en bases de datos, dificulta la comparación y caracterización precisa de un fenotipo clínico asociado a este gen.

En este paciente, el diagnóstico diferencial de FRAXE no fue considerado, dado que, desde la primera consulta a los cuatro años, hasta la última a los diecinueve, ha presentado un fenotipo clásico del SCdL. El diagnóstico se apoyó tanto en la historia natural del paciente, como en la elevada puntuación clínica obtenida mediante los criterios de Kline et al. [111]. Además, la aplicación de *Face2Gene*, herramienta basada en IA, sirvió también para confirmar este diagnóstico. Aunque el seguimiento longitudinal reveló cambios morfológicos significativos, como la forma nasal y la aparición de estrabismo (HP:0000486) divergente, estas alteraciones no afectaron a la clasificación inicial. En cuanto a la tía materna, pese a que su discapacidad intelectual era menos grave que la del paciente, su puntuación clínica también la clasificó dentro del fenotipo clásico del síndrome [111]. Por el contrario, la madre, aunque presentaba algunas características típicas como braquicefalia (HP:0000248), surco nasolabial largo (HP:0000343) y bermellón del labio superior fino (HP:0000219), la falta de sintomatología específica, impedía su inclusión dentro del ECdL.

La similitud fenotípica entre el paciente y la tía, reforzaba la hipótesis de que el gen causal estuviera ligado al cromosoma X. Las diferencias clínicas observadas entre la madre y la tía, gemelas dicigóticas, podían atribuirse a la variabilidad en la inactivación del cromosoma X, como ha sido demostrado en casos similares [284]. El análisis de fibroblastos de ambas hermanas, reveló una inactivación balanceada de los alelos de la madre, mientras que en la tía se observó un sesgo hacia el alelo salvaje, con expresión exclusiva del alelo mutado. Este hallazgo ha permitido comprender las diferencias clínicas observadas entre ambas. Sin embargo, en la sangre las dos hermanas tenían el

alelo afectado silenciado, fenómeno que ya ha sido descrito en este tejido en otras ocasiones [156,285,286].

El efecto patológico de la duplicación intragénica parece explicarse por la marcada disminución en la expresión del gen *AFF2* en el paciente y la tía, quedando fuera de esta observación la madre. Esta reducción de la expresión ha sido descrita en casos similares, y se ha asociado con mecanismos de degradación de los tránscritos aberrantes [287,288]. Además, la aplicación de la técnica *Oxford Nanopore Technologies*, ha permitido una secuenciación y cuantificación precisa de los tránscritos. Gracias a ella, se ha podido confirmar una expresión dominante del tránscrito aberrante, con duplicación en tándem de los exones 10, 11 y 12, en el paciente y la tía, aunque no en la madre. Estos hallazgos contribuyen a justificar las diferencias clínicas observadas entre los diferentes miembros de la familia.

A nivel proteico, el análisis de la secuencia del tránscrito aberrante mostró el cambio del primer aminoácido del exón 10 duplicado, seguido por la repetición de 431 aminoácidos adicionales, producto de la duplicación de los exones 10, 11 y 12. Como resultado se genera una nueva proteína que incluye la duplicación de un dominio rico en serina, dos señales de localización nuclear y el motivo de interacción AF4. Estos cambios sugieren una alteración estructural significativa. Sin embargo, la ausencia de estudios de cristalización impide realizar un modelado tridimensional preciso. Además, la exclusión de las CNVs en la mayoría de predictores de patogenicidad, restringe el uso de estas herramientas para interpretar su impacto funcional. A pesar de estas limitaciones, la reducción de la expresión de *AFF2*, unida a la alteración de la proteína mutada, parecen suficientes para explicar la patogenicidad de esta duplicación intragénica.

Este trabajo describe por primera vez el hallazgo de una variante en el gen *AFF2* en una familia afectada por una forma clásica del SCdL. Además, también recoge el papel regulador de la transcripción que la proteína AFF2 tiene, y que la conecta funcionalmente con otras proteínas implicadas en el SCdL. Los resultados obtenidos sugieren que el gen *AFF2* podría ser un nuevo gen causal del SCdL, aunque es evidente que estudio adicionales deberán apoyar y refrendar este descubrimiento. A partir de ahora, se recomienda incluir el análisis de este gen en individuos con características clínicas compatibles con el síndrome, especialmente en casos con dificultades en el diagnóstico genético. Además, la existencia de variantes en *AFF2* reportadas en bases de datos, con clínica incompleta, sugiere la importancia de examinar cohortes más amplias, con un análisis fenotípico detallado que permita conocer el cuadro clínico asociado a este gen.

Perspectivas de futuro y limitaciones del estudio

A continuación, se exponen las futuras líneas de investigación que sugiere este estudio, así como las limitaciones encontradas.

- El avance en el conocimiento de las vías moleculares afectadas en el síndrome de PACS1, está permitiendo establecer nuevas relaciones con patologías del neurodesarrollo. Sin embargo, resulta necesario investigar la relación funcional a nivel molecular, para poder explicar el solapamiento clínico de estos trastornos.

- Es fundamental establecer protocolos de reevaluación de datos genómicos en los casos que se observa una evolución fenotípica del paciente, o en aquellos que todavía no se ha logrado identificar la variante patogénica, a fin de precisar el diagnóstico genético.

- El aumento de casos de CNVs asociadas al SCdL pone en evidencia la capacidad diagnóstica de la técnica empleada, así como la relevancia de la selección adecuada de los algoritmos a utilizar en el análisis de datos.

- Es necesario llevar a cabo un análisis de cohortes más amplias de pacientes con variantes en el gen *AFF2*, acompañada de un fenotipado profundo, que permita definir de manera precisa el cuadro clínico asociado a este gen.

- Los resultados obtenidos sugieren la importancia de incluir en el futuro los genes *AFF2*, *ANKRD11* y *PACS1* en los paneles diagnósticos del SCdL.

Conclusiones / Conclusions

1. Se describe por primera vez el hallazgo de una variante en el gen *AFF2* en una familia afectada por una forma clásica del SCdL. Además, se recoge el papel regulador de la transcripción que la proteína AFF2 tiene, y que la une funcionalmente con otras proteínas relacionadas con el SCdL. Estos resultados sugieren que *AFF2* podría ser un nuevo gen causal de este síndrome.

2. El estudio longitudinal de una paciente con una variante patogénica en el gen *ANKRD11*, muestra su evolución clínica de SCdL a síndrome de KBG. Este hecho, induce la reflexión sobre si la similitud fenotípica de ambos cuadros, ha llevado, a veces, a considerar *ANKRD11* como un gen causal del SCdL.

3. El hallazgo de la segunda duplicación intragénica en el gen *HDAC8* asociada al SCdL, subraya la importancia de implementar un enfoque diagnóstico más amplio y dirigido a la detección de variantes en el número de copias (CNVs).

4. La gran similitud clínica y molecular observada entre los síndromes PACS1, PACS2 y WDR37 sugiere que podrían incluirse en una nueva entidad nosológica más amplia siguiendo el ejemplo del Espectro Cornelia de Lange. Además, las coincidencias fenotípicas de estos cuadros con el SCdL apoyan la necesidad de realizar un correcto diagnóstico diferencial.

1. The discovery of a pathogenic variant in the *AFF2* gene in a family affected with classic Cornelia de Lange syndrome (CdLS) is reported for the first time. Furthermore, the regulatory role of AFF2 in transcription is emphasised, which links it functionally with other causal proteins of CdLS. These findings suggest that *AFF2* may represent a novel causal gene for this syndrome.

2. A longitudinal study of a patient with a pathogenic variant in the *ANKRD11* gene reveals her clinical shift from CdLS to KBG syndrome. This fact prompts reflection on whether the phenotypic similarity between the two conditions has sometimes led to *ANKRD11* being considered as a causal gene for CdLS.

3. The identification of the second intragenic duplication reported in the *HDAC8* gene associated with CdLS underscores the importance of implementing a broader diagnostic approach to detect copy number variations (CNVs).

4. The considerable clinical and molecular similarity observed among PACS1, PACS2, and WDR37 syndromes suggests that they could be included in a new, broader nosological entity, following the example of the Cornelia de Lange Spectrum. In addition, the phenotypic overlaps of these conditions with CdLS support the need for accurate differential diagnosis.

Bibliografía

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 EJHG. (2020). 28(2):165-73. doi: 10.1038/s41431-019-0508-0. PMID: 31527858.

2. Richter T, Nestler-Parr S, Babela R, Khan ZM, Tesoro T, Molsen E, et al. *Rare Disease Terminology and Definitions-A Systematic Global Review: Report of the ISPOR Rare Disease Special Interest Group*. Value Health J Int Soc Pharmacoeconomics Outcomes Res. (2015). 18(6):906-14. doi: 10.1016/j.jval.2015.05.008. PMID: 26409619.

3. Kerr K, McAneney H, Smyth LJ, Bailie C, McKee S, McKnight AJ. *A scoping review and proposed workflow for multi-omic rare disease research*. Orphanet J Rare Dis. (2020). 15(1):107. doi: 10.1186/s13023-020-01376-x. PMID: 32345347.

4. van Karnebeek CDM, O'Donnell-Luria A, Baynam G, Baudot A, Groza T, Jans JJM, et al. *Leaving no patient behind! Expert recommendation in the use of innovative technologies for diagnosing rare diseases*. (2024). 19(1):357. doi: 10.1186/s13023-024-03361-0. PMID: 39334316.

5. Ramos-Fuentes F, González-Meneses A, Ars E, Hernández-Jaras J. *Genetic Diagnosis of Rare Diseases: Past and Present*. Adv Ther. (2020). 37(Suppl 2):29-37. doi: 10.1007/s12325-019-01176-1. Epub 2020 Mar 31. PMID: 32236876.

6. Valenzuela I, Codina-Solà M, Vazquez E, Cueto-González A, Leno-Colorado J, Lasa-Aranzasti A, et al. *Deep phenotyping of 11 individuals with pathogenic variants in RNU4-2 reveals a clinically recognizable syndrome*. Genet Med Off J Am Coll Med Genet. (2024). 26(12):101288. doi: 10.1016/j.gim.2024.101288. Epub 2024 Oct 2. PMID: 39369315.

7. Sullivan JA, Schoch K, Spillmann RC, Shashi V. *Exome/Genome Sequencing in Undiagnosed Syndromes*. Annu Rev Med. (2023). 74:489-502. doi: 10.1146/annurev-med-042921-110721. PMID: 36706750.

8. Wojcik MH, Lemire G, Berger E, Zaki MS, Wissmann M, Win W, et al. *Genome Sequencing for Diagnosing Rare Diseases*. N Engl J Med. (2024). 390(21):1985-97. doi: 10.1056/NEJMoa2314761. PMID: 38838312.

9. Lancaster MA, Knoblich JA. *Organogenesis in a dish: modeling development and disease using organoid technologies*. Science. (2014). 345(6194):1247125. doi: 10.1126/science.1247125. PMID: 25035496.

10. Martini PGV, Guey LT. *A New Era for Rare Genetic Diseases: Messenger RNA Therapy*. *Hum Gene Ther*. (2019). 30(10):1180-9. doi: 10.1089/hum.2019.090. PMID: 31179759.

11. Wright CF, FitzPatrick DR, Firth HV. *Paediatric genomics: diagnosing rare disease in children*. Nat Rev Genet. (2018). 19(5):253-68. doi: 10.1038/nrg.2018.12. PMID: 29398702.

12. 100,000 Genomes Project Pilot Investigators, Smedley D, Smith KR, Martin A, Thomas EA, McDonagh EM, et al. *100,000 Genomes Pilot on Rare-Disease Diagnosis in Health Care - Preliminary Report*. N Engl J Med. (2021). 385(20):1868-80. doi: 10.1056/NEJMoa2035790. PMID: 34758253.

13. Marwaha S, Knowles JW, Ashley EA. *A guide for the diagnosis of rare and undiagnosed disease: beyond the exome*. Genome Med. (2022). 14(1):23. doi: 10.1186/s13073-022-01026-w. PMID: 35220969.

14. Julkowska D, Austin CP, Cutillo CM, Gancberg D, Hager C, Halftermeyer J, et al. *The importance of international collaboration for rare diseases research: a European perspective*. Gene Ther. (2017). 24(9):562-71. doi: 10.1038/gt.2017.29. PMID: 28440796.

15. Luque J, Mendes I, Gómez B, Morte B, López de Heredia M, Herreras E, et al. *CIBERER: Spanish national network for research on rare diseases: A highly productive collaborative initiative*. Clin Genet. (2022). 101(5-6):481-93. doi: 10.1111/cge.14113. PMID: 35060122.

16. Thompson R, Johnston L, Taruscio D, Monaco L, Béroud C, Gut IG, et al. *RD-Connect: an integrated platform connecting databases, registries, biobanks and clinical bioinformatics for rare disease research*. J Gen Intern Med. (2014). 29 Suppl 3(Suppl 3):S780-787. doi: 10.1007/s11606-014-2908-8. PMID: 25029978.

17. Robinson PN. *Deep phenotyping for precision medicine*. Hum Mutat. (2012). 33(5):777-80. doi: 10.1002/humu.22080. PMID: 22504886.

18. Li G, Xu K, Yin X, Yang J, Cai J, Yang X, et al. *Integrating deep phenotyping with genetic analysis: a comprehensive workflow for diagnosis and management of rare bone diseases*. Orphanet J Rare Dis. (2024). 19(1):371. doi: 10.1186/s13023-024-03367-8. PMID: 39380097.

19. Gargano MA, Matentzoglu N, Coleman B, Addo-Lartey EB, Anagnostopoulos AV, Anderton J, et al. *The Human Phenotype Ontology in 2024: phenotypes around the world*. Nucleic Acids Res. (2024). 52(D1):D1333-46. doi: 10.1093/nar/gkad1005. PMID: 37953324.

20. Kalaydjieva L. *Congenital cataracts-facial dysmorphism-neuropathy*. Orphanet J Rare Dis. (2006). 1:32. doi: 10.1186/1750-1172-1-32. PMID: 16939648.

21. Hart TC, Hart PS. *Genetic studies of craniofacial anomalies: clinical implications and applications*. Orthod Craniofac Res. (2009). 12(3):212-20. doi: 10.1111/j.1601-6343.2009.01455.x. PMID: 19627523.

22. Gurovich Y, Hanani Y, Bar O, Nadav G, Fleischer N, Gelbman D, et al. *Identifying facial phenotypes of genetic disorders using deep learning*. Nat Med. (2019). 25(1):60-4. doi: 10.1038/s41591-018-0279-0. PMID: 30617323.

23. Turro E, Astle WJ, Megy K, Gräf S, Greene D, Shamardina O, et al. *Whole-genome sequencing of patients with rare diseases in a national health system*. Nature. (2020). 583(7814):96-102. doi: 10.1038/s41586-020-2434-2. PMID: 32581362.

24. Pagnamenta AT, Camps C, Giacopuzzi E, Taylor JM, Hashim M, Calpena E, et al. *Structural and non-coding variants increase the diagnostic yield of clinical whole genome sequencing for rare diseases*. Genome Med. (2023). 15(1):94. doi: 10.1186/s13073-023-01240-0. PMID: 37946251.

25. Katsanis SH, Katsanis N. *Molecular genetic testing and the future of clinical genomics*. Nat Rev Genet. (2013). 14(6):415-26. doi: 10.1038/nrg3493. PMID: 23681062.

26. Sagoo GS, Butterworth AS, Sanderson S, Shaw-Smith C, Higgins JPT, Burton H. *Array CGH in patients with learning disability (mental retardation) and congenital anomalies: updated systematic review and meta-analysis of 19 studies and 13,926 subjects*. Genet Med Off J Am Coll Med Genet. (2009). 11(3):139-46. doi: 10.1097/GIM.0b013e318194ee8f. PMID: 19367186.

27. Vissers LELM, van Nimwegen KJM, Schieving JH, Kamsteeg EJ, Kleefstra T, Yntema HG, et al. *A clinical utility study of exome sequencing versus conventional genetic testing in pediatric neurology*. Genet Med Off J Am Coll Med Genet. (2017). 19(9):1055-63. doi: 10.1038/gim.2017.1. PMID: 28333917.

28. Srivastava S, Love-Nichols JA, Dies KA, Ledbetter DH, Martin CL, Chung WK, et al. *Meta-analysis and multidisciplinary consensus statement: exome sequencing is a first-tier clinical diagnostic test for individuals with neurodevelopmental disorders*. Genet Med Off J Am Coll Med Genet. (2019). 21(11):2413-21. doi: 10.1038/s41436-020-0913-3. PMID: 31182824.

29. Ewans LJ, Minoche AE, Schofield D, Shrestha R, Puttick C, Zhu Y, et al. *Whole exome and genome sequencing in mendelian disorders: a diagnostic and health economic analysis.* Eur J Hum Genet EJHG. (2022). 30(10):1121-31. doi: 10.1038/s41431-022-01162-2. Epub 2022 Aug 15. PMID: 35970915.

30. Félix TM, Fischinger Moura de Souza C, Oliveira JB, Rico-Restrepo M, Zanoteli E, Zatz M, et al. *Challenges and recommendations to increasing the use of exome sequencing and whole genome sequencing for diagnosing rare diseases in Brazil: an expert perspective. Int J Equity Health*. (2023). 22(1):11. doi: 10.1186/s12939-022-01809-y. PMID: 36639662.

31. Mitsuhashi S, Matsumoto N. *Long-read sequencing for rare human genetic diseases*. J Hum Genet.(2020). 65(1):11-9. doi: 10.1038/s10038-019-0671-8. PMID: 31558760.

32. Karst SM, Ziels RM, Kirkegaard RH, Sørensen EA, McDonald D, Zhu Q, et al. *High-accuracy long-read amplicon sequences using unique molecular identifiers with Nanopore or PacBio sequencing*. Nat Methods. (2021). 18(2):165-9. doi: 10.1038/s41592-020-01041-y. PMID: 33432244.

33. Warburton PE, Sebra RP. *Long-Read DNA Sequencing: Recent Advances and Remaining Challenges. Annu Rev Genomics Hum Genet*. (2023). 24:109-32. doi: 10.1146/annurev-genom-101722-103045. PMID: 37075062.

34. Olivucci G, Iovino E, Innella G, Turchetti D, Pippucci T, Magini P. *Long read sequencing on its way to the routine diagnostics of genetic diseases*. Front Genet. (2024). 15:1374860. doi: 10.3389/fgene.2024.1374860. PMID: 38510277.

35. Evangelidou P, Alexandrou A, Moutafi M, Ioannides M, Antoniou P, Koumbaris G, et al. *Implementation of high resolution whole genome array CGH in the prenatal clinical setting: advantages, challenges, and review of the literature*. BioMed Res Int. (2013). 2013:346762. doi: 10.1155/2013/346762. PMID: 23555083.

36. Tsoutsou E, Tzetis M, Giannikou K, Braoudaki M, Mitrakos A, Amenta S, et al. *Application of high-resolution array comparative genomic hybridization in children with unknown syndromic microcephaly*. Pediatr Res. (2017). 82(2):253-60. doi: 10.1038/pr.2017.65. PMID: 28422950.

37. Bestetti I, Castronovo C, Sironi A, Caslini C, Sala C, Rossetti R, et al. *High-resolution array-CGH analysis on 46,XX patients affected by early onset primary ovarian insufficiency discloses new genes involved in ovarian function*. Hum Reprod Oxf Engl. (2019). 34(3):574-83. doi: 10.1093/humrep/dey389. PMID: 30689869.

38. Wenger AM, Peluso P, Rowell WJ, Chang PC, Hall RJ, Concepcion GT, et al. *Accurate circular consensus long-read sequencing improves variant detection and assembly of a human genome*. Nat Biotechnol. (2019). 37(10):1155-62. doi: 10.1038/s41587-019-0217-9. PMID: 31406327.

39. Hartin SN, Means JC, Alaimo JT, Younger ST. *Expediting rare disease diagnosis: a call to bridge the gap between clinical and functional genomics*. Mol Med Camb Mass. (2020). 26(1):117. doi: 10.1186/s10020-020-00244-5. PMID: 33238891.

40. Shreeve N, Sproule C, Choy KW, Dong Z, Gajewska-Knapik K, Kilby MD, et al. *Incremental yield of whole-genome sequencing over chromosomal microarray analysis and exome sequencing for congenital anomalies in prenatal period and infancy: systematic review and meta-analysis*. Ultrasound Obstet Gynecol Off J Int Soc Ultrasound Obstet Gynecol. (2024). 63(1):15-23. doi: 10.1002/uog.27491. PMID: 37725747.

41. Short PJ, McRae JF, Gallone G, Sifrim A, Won H, Geschwind DH, et al. *De novo mutations in regulatory elements in neurodevelopmental disorders*. Nature. (2018). 555(7698):611-6. doi: 10.1038/nature25983. PMID: 29562236.

42. Bonnert M, Olén O, Bjureberg J, Lalouni M, Hedman-Lagerlöf E, Serlachius E, et al. T*he role of avoidance behavior in the treatment of adolescents with irritable bowel syndrome: A mediation analysis*. Behav Res Ther. (2018). 105:27-35. doi: 10.1016/j.brat.2018.03.006. PMID: 29614378.

43. Zhou J, Park CY, Theesfeld CL, Wong AK, Yuan Y, Scheckel C, et al. *Whole-genome deep-learning analysis identifies contribution of noncoding mutations to autism risk*. Nat Genet. (2019). 51(6):973-80. doi: 10.1038/s41588-019-0420-0. PMID: 31133750.

44. Reuter C, Chun N, Pariani M, Hanson-Kahn A. *Understanding variants of uncertain significance in the era of multigene panels: Through the eyes of the patient*. J Genet Couns. (2019). 28(4):878-86. doi: 10.1002/jgc4.1130. PMID: 31050105.

45. Chen E, Facio FM, Aradhya KW, Rojahn S, Hatchell KE, Aguilar S, et al. *Rates and Classification of Variants of Uncertain Significance in Hereditary Disease Genetic Testing*. JAMA Netw Open. (2023). 6(10):e2339571. doi: 10.1001/jamanetworkopen.2023.39571. PMID: 37878314.

46. 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 Off J Am Coll Med Genet. (2015) 17(5):405-24. doi: 10.1038/gim.2015.30. PMID: 25741868.

47. Landrum MJ, Lee JM, Benson M, Brown GR, Chao C, Chitipiralla S, et al. *ClinVar: improving access to variant interpretations and supporting evidence*. Nucleic Acids Res. (2018). 46(D1):D1062-7. doi: 10.1093/nar/gkx1153. PMID: 29165669.

48. Hartman P, Beckman K, Silverstein K, Yohe S, Schomaker M, Henzler C, et al. *Next generation sequencing for clinical diagnostics: Five year experience of an academic laboratory*. Mol Genet Metab Rep. (2019). 19:100464. doi: 10.1016/j.ymgmr.2019.100464. PMID: 30891420.

49. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, et al. *The mutational constraint spectrum quantified from variation in 141,456 humans*. Nature. (2020). 581(7809):434-43. doi: 10.1038/s41586-021-03758-y. PMID: 32461654.

50. Wu Y, Li R, Sun S, Weile J, Roth FP. *Improved pathogenicity prediction for rare human missense variants*. Am J Hum Genet. (2021). 108(10):1891-906. doi: 10.1016/j.ajhg.2021.11.010. PMID: 34551312.

51. Frazer J, Notin P, Dias M, Gomez A, Min JK, Brock K, et al. *Disease variant prediction with deep generative models of evolutionary data*. Nature. (2021). 599(7883):91-5. doi: 10.1038/s41586-021-04207-6. PMID: 34707284.

52. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. *Highly accurate protein structure prediction with AlphaFold*. Nature. agosto de 2021;596(7873):583-9.

53. Khodosevich K, Sellgren CM. *Neurodevelopmental disorders-high-resolution rethinking of disease modeling*. Mol Psychiatry. (2023). 28(1):34-43. doi: 10.1038/s41586-021-03819-2. PMID: 34265844.

54. Crouzier L, Richard EM, Sourbron J, Lagae L, Maurice T, Delprat B. *Use of Zebrafish Models to Boost Research in Rare Genetic Diseases*. Int J Mol Sci. 12 de diciembre de 2021;22(24):13356. doi: 10.3390/ijms222413356. PMID: 34948153.

55. Deciphering Developmental Disorders Study. *Prevalence and architecture of de novo mutations in developmental disorders*. Nature. (2017). 542(7642):433-8. doi: 10.1038/nature21062. PMID: 28135719.

56. Sanders SJ, Sahin M, Hostyk J, Thurm A, Jacquemont S, Avillach P, et al. *A framework for the investigation of rare genetic disorders in neuropsychiatry*. Nat Med. (2019). 25(10):1477-87. doi: 10.1038/s41591-019-0581-5. PMID: 31548702.

57. Thapar A, Cooper M, Rutter M. *Neurodevelopmental disorders*. Lancet Psychiatry. (2017). 4(4):339-46. doi: 10.1016/S2215-0366(16)30376-5. PMID: 27979720.

58. Morris-Rosendahl DJ, Crocq MA. *Neurodevelopmental disorders-the history and future of a diagnostic concept*. Dialogues Clin Neurosci. (2020). 22(1):65-72. doi: 10.31887/DCNS.2020.22.1/macrocq. PMID: 32699506.

59. Cristino AS, Williams SM, Hawi Z, An JY, Bellgrove MA, Schwartz CE, et al. *Neurodevelopmental and neuropsychiatric disorders represent an interconnected molecular system*. Mol Psychiatry. (2014). 19(3):294-301. doi: 10.1038/mp.2013.16. PMID: 23439483.

60. Hormozdiari F, Penn O, Borenstein E, Eichler EE. *The discovery of integrated gene networks for autism and related disorders*. Genome Res. (2015). 25(1):142-54. doi: 10.1101/gr.178855.114. PMID: 25378250.

61. Barabási AL, Gulbahce N, Loscalzo J. *Network medicine: a network-based approach to human disease*. Nat Rev Genet. (2011). 12(1):56-68. doi: 10.1038/nrg2918. PMID: 21164525.

62. De Rubeis S, He X, Goldberg AP, Poultney CS, Samocha K, Cicek AE, et al. *Synaptic, transcriptional and chromatin genes disrupted in autism*. Nature. (2014). 515(7526):209-15. doi: 10.1038/nature13772. PMID: 25363760.

63. Parenti I, Rabaneda LG, Schoen H, Novarino G. *Neurodevelopmental Disorders: From Genetics to Functional Pathways*. Trends Neurosci. (2020). 43(8):608-21. doi: 10.1016/j.tins.2020.05.004. PMID: 32507511.

64. Buphamalai P, Kokotovic T, Nagy V, Menche J. *Network analysis reveals rare disease signatures across multiple levels of biological organization*. Nat Commun. (2021). 12(1):6306. doi: 10.1038/s41467-021-26674-1. PMID: 34753928.

65. Südhof TC. Synaptic Neurexin Complexes: A Molecular Code for the Logic of Neural Circuits. Cell. (2017). 171(4):745-69. oi: 10.1016/j.cell.2017.10.024. PMID: 29100073.

66. Südhof TC. *Towards an Understanding of Synapse Formation*. Neuron. 24 (2018). 100(2):276-93. doi: 10.1016/j.neuron.2018.09.040. PMID: 30359597.

67. Bourgeron T. *From the genetic architecture to synaptic plasticity in autism spectrum disorder*. Nat Rev Neurosci. (2015). 16(9):551-63. doi: 10.1038/nrn3992. PMID: 26289574.

68. Zaslavsky K, Zhang WB, McCready FP, Rodrigues DC, Deneault E, Loo C, et al. *SHANK2 mutations associated with autism spectrum disorder cause hyperconnectivity of human neurons*. Nat Neurosci.(2019). 22(4):556-64. doi: 10.1038/s41593-019-0365-8. PMID: 30911184.

69. Marro SG, Chanda S, Yang N, Janas JA, Valperga G, Trotter J, et al. N*euroligin-4 Regulates Excitatory Synaptic Transmission in Human Neurons*. Neuron. (2019). 103(4):617-626.e6. doi: 10.1016/j.neuron.2019.05.043. PMID: 31257103.

70. Pak C, Danko T, Zhang Y, Aoto J, Anderson G, Maxeiner S, et al. *Human Neuropsychiatric Disease Modeling using Conditional Deletion Reveals Synaptic Transmission Defects Caused by Heterozygous Mutations in NRXN1*. Cell Stem Cell. (2015). 7(3):316-28. doi: 10.1016/j.stem.2015.07.017. PMID: 26279266.

71. Flaherty E, Zhu S, Barretto N, Cheng E, Deans PJM, Fernando MB, et al. *Neuronal impact of patient-specific aberrant NRXN1α splicing*. Nat Genet. (2019). 51(12):1679-90. doi: 10.1038/s41588-019-0539-z. PMID: 31784728.

72. Villar-Pazos S, Thomas L, Yang Y, Chen K, Lyles JB, Deitch BJ, et al. *RNA-targeted therapy corrects neuronal deficits in PACS1 syndrome mice*. Res Sq. (2023). rs.3.rs-2440581. doi: 10.1038/s41467-023-42176-8. PMID: 36747781.

73. Villar-Pazos S, Thomas L, Yang Y, Chen K, Lyles JB, Deitch BJ, et al. *Neural deficits in a mouse model of PACS1 syndrome are corrected with PACS1- or HDAC6-targeting therapy*. Nat Commun. (2023). 14(1):6547. doi: 10.1038/s41467-023-42176-8. PMID: 37848409.

74. Simons M, Nave KA. *Oligodendrocytes: Myelination and Axonal Support*. Cold Spring Harb Perspect Biol. (2015). 8(1):a020479. doi: 10.1101/cshperspect.a020479. PMID: 26101081.

75. Knowles JK, Batra A, Xu H, Monje M. *Adaptive and maladaptive myelination in health and disease*. Nat Rev Neurol. (2022). 18(12):735-46. doi: 10.1038/s41582-022-00737-3. PMID: 36376595.

76. Emery B. *Regulation of oligodendrocyte differentiation and myelination*. Science. (2010). 330(6005):779-82. doi: 10.1126/science.1190927. PMID: 21051629.

77. Rzepnikowska W, Kochański A. *A role for the GDAP1 gene in the molecular pathogenesis of Charcot-Marie-Tooth disease*. Acta Neurobiol Exp (Warsz). (2018). 78(1):1-13. PMID: 29694336.

78. Wang H, Kaçar Bayram A, Sprute R, Ozdemir O, Cooper E, Pergande M, et al. *Genotype-Phenotype Correlations in Charcot-Marie-Tooth Disease Due to MTMR2 Mutations and Implications in Membrane Trafficking*. Front Neurosci. (2019). 13:974. doi: 10.3389/fnins.2019.00974. PMID: 31680794.

79. Ma X, Liu X, Duan X, Fan D. *Screening for PRX mutations in a large Chinese Charcot-Marie-Tooth disease cohort and literature review*. Front Neurol. (2023). 14:1148044. doi: 10.3389/fneur.2023.1148044. PMID: 37470010.

80. Girodengo M, Ultanir SK, Bateman JM. *Mechanistic target of rapamycin signaling in human nervous system development and disease*. Front Mol Neurosci. (2022). 15:1005631. doi: 10.3389/fnmol.2022.1005631. PMID: 36226315.

81. Graber TE, McCamphill PK, Sossin WS. *A recollection of mTOR signaling in learning and memory*. *Learn Mem Cold Spring Harb N.* (2013). 20(10):518-30. doi: 10.1101/lm.027664.112. PMID: 24042848.

82. Lipton JO, Sahin M. *The neurology of mTOR*. Neuron. (2014). 84(2):275-91. doi: 10.1016/j.neuron.2014.09.034. PMID: 25374355.

83. Curatolo P, Moavero R, de Vries PJ. *Neurological and neuropsychiatric aspects of tuberous sclerosis complex*. Lancet Neurol. (2015).14(7):733-45. doi: 10.1016/S1474-4422(15)00069-1. PMID: 26067126.

84. Raybaud C, Widjaja E. *Development and dysgenesis of the cerebral cortex: malformations of cortical development*. Neuroimaging Clin N Am. (2011). 21(3):483-543, vii. doi: 10.1016/j.nic.2011.05.014. PMID: 21807310.

85. Poirier K, Lebrun N, Broix L, Tian G, Saillour Y, Boscheron C, et al. *Mutations in TUBG1, DYNC1H1, KIF5C and KIF2A cause malformations of cortical development and microcephaly*. Nat Genet. (2013);45(6):639-47. doi: 10.1038/ng.2613. PMID: 23603762.

86. Guerrini R, Dobyns WB. *Malformations of cortical development: clinical features and genetic causes*. Lancet Neurol. (2014). 13(7):710-26. doi: 10.1016/S1474-4422(14)70040-7. PMID: 24932993.

87. Roberts B. *Neuronal Migration Disorders*. Radiol Technol. (2018). 89(3):279-95. PMID: 29298944.

88. Severino M, Geraldo AF, Utz N, Tortora D, Pogledic I, Klonowski W, et al. *Definitions and classification of malformations of cortical development: practical guidelines*. Brain J Neurol. (2020). 143(10):2874-94. doi: 10.1093/brain/awaa298. PMID: 32779696.

89. Sokpor G, Brand-Saberi B, Nguyen HP, Tuoc T. *Regulation of Cell Delamination During Cortical Neurodevelopment and Implication for Brain Disorders*. Front Neurosci. (2022). 16:824802. doi: 10.3389/fnins.2022.824802. PMID: 35281509.

90. Ronan JL, Wu W, Crabtree GR. *From neural development to cognition: unexpected roles for chromatin*. Nat Rev Genet. (2013). 14(5):347-59. doi: 10.1038/nrg3413. PMID: 23568486.

91. Stessman HAF, Xiong B, Coe BP, Wang T, Hoekzema K, Fenckova M, et al. *Targeted sequencing identifies 91 neurodevelopmental-disorder risk genes with autism and developmental-disability biases*. Nat Genet. (2017). 49(4):515-26. doi: 10.1038/ng.3792. PMID: 28191889.

92. Goodman JV, Bonni A. *Regulation of neuronal connectivity in the mammalian brain by chromatin remodeling*. Curr Opin Neurobiol. (2019). 59:59-68. doi: 10.1016/j.conb.2019.04.010. PMID: 31146125.

93. Katayama Y, Nishiyama M, Shoji H, Ohkawa Y, Kawamura A, Sato T, et al. *CHD8 haploinsufficiency results in autistic-like phenotypes in mice*. Nature. (2016). 537(7622):675-9. doi: 10.1038/nature19357. PMID: 27602517.

94. Jung EM, Moffat JJ, Liu J, Dravid SM, Gurumurthy CB, Kim WY. *Arid1b haploinsufficiency disrupts cortical interneuron development and mouse behavior*. Nat Neurosci. (2017). 20(12):1694-707. doi: 10.1038/s41593-017-0013-0. PMID: 29184203.

95. Deliu E, Arecco N, Morandell J, Dotter CP, Contreras X, Girardot C, et al. *Haploinsufficiency of the intellectual disability gene SETD5 disturbs developmental gene expression and cognition*. Nat Neurosci. (2018). 21(12):1717-27. doi: 10.1038/s41593-018-0266-2. PMID: 30455454.

96. Zaghi M, Longo F, Massimino L, Rubio A, Bido S, Mazzara PG, et al. *SETD5 haploinsufficiency affects mitochondrial compartment in neural cells*. Mol Autism. (2023). 14(1):20. doi: 10.1186/s13229-023-00550-9. PMID: 37264456.

97. Ascaso Á, Arnedo M, Puisac B, Latorre-Pellicer A, Del Rincón J, Bueno-Lozano G, et al. *Cornelia de Lange Spectrum*. An Pediatr. (2024). 100(5):352-62. doi: 10.1016/j.anpede.2024.04.012. PMID: 38735830.

98. Li X, Zhang Y, Wang L, Lin Y, Gao Z, Zhan X, et al. *Integrated Analysis of Brain Transcriptome Reveals Convergent Molecular Pathways in Autism Spectrum Disorder*. Front Psychiatry. (2019). 10:706. doi: 10.3389/fpsyt.2019.00706. PMID: 31649562.

99. Prashad S, Gopal PP. *RNA-binding proteins in neurological development and disease*. RNA Biol. (2021).18(7):972-87. doi: 10.1080/15476286.2020.1809186. PMID: 32865115.

100. Lalli MA, Avey D, Dougherty JD, Milbrandt J, Mitra RD. *High-throughput single-cell functional elucidation of neurodevelopmental disease-associated genes reveals convergent mechanisms altering neuronal differentiation*. Genome Res. (2020). 30(9):1317-31. doi: 10.1101/gr.262295.120. PMID: 32887689.

101. Mullegama SV, Klein SD, Williams SR, Innis JW, Probst FJ, Haldeman-Englert C, et al. *Transcriptome analysis of MBD5-associated neurodevelopmental disorder (MAND) neural progenitor cells reveals dysregulation of autism-associated genes*. Sci Rep. (2021). 11(1):11295. doi: 10.1038/s41598-021-90798-z. PMID: 34050248.

102. Bukowska-Olech E, Majchrzak-Celińska A, Przyborska M, Jamsheer A. *Chromatinopathies: insight in clinical aspects and underlying epigenetic changes*. J Appl Genet. (2024). 65(2):287-301. doi: 10.1007/s13353-023-00824-1. PMID: 38180712.

103. Avagliano L, Parenti I, Grazioli P, Di Fede E, Parodi C, Mariani M, et al. *Chromatinopathies: A focus on Cornelia de Lange syndrome*. Clin Genet. (2020). 97(1):3-11. doi: 10.1007/s13353-023-00824-1. PMID: 38180712.

104. Parenti I, Kaiser FJ. Cornelia de Lange Syndrome as Paradigm of Chromatinopathies. Front Neurosci. (2021). 15:774950. doi: 10.3389/fnins.2021.774950. PMID: 34803598.

105. Javaid N, Choi S. *Acetylation- and Methylation-Related Epigenetic Proteins in the Context of Their Targets*. Genes. (2017). 8(8):196. doi: 10.3390/genes8080196. PMID: 28783137.

106. Sadikovic B, Levy MA, Aref-Eshghi E. *Functional annotation of genomic variation: DNA methylation episignatures in neurodevelopmental Mendelian disorders*. Hum Mol Genet. (2020). 29(R1):R27-32. doi: 10.1093/hmg/ddaa144. PMID: 32644126.

107. Nava AA, Arboleda VA. *The omics era: a nexus of untapped potential for Mendelian chromatinopathies*. Hum Genet. (2024). 143(4):475-95. doi: 10.1007/s00439-023-02560-2. PMID: 37115317.

108. Di Fede E, Grazioli P, Lettieri A, Parodi C, Castiglioni S, Taci E, et al. *Epigenetic disorders: Lessons from the animals-animal models in chromatinopathies*. Front Cell Dev Biol. (2022). 10:979512. doi: 10.3389/fcell.2022.979512. PMID: 36225316.

109. Ciptasari U, van Bokhoven H. *The phenomenal epigenome in neurodevelopmental disorders.* Hum Mol Genet. (2020). 29(R1):R42-50. doi: 10.1093/hmg/ddaa175. PMID: 32766754.

110. Kline AD, Krantz ID, Sommer A, Kliewer M, Jackson LG, FitzPatrick DR, et al. *Cornelia de Lange syndrome: clinical review, diagnostic and scoring systems, and anticipatory guidance*. Am J Med Genet A. (2007). 143A(12):1287-96. doi: 10.1002/ajmg.a.31757. PMID: 17508425.

111. Kline AD, Moss JF, Selicorni A, Bisgaard AM, Deardorff MA, Gillett PM, et al. *Diagnosis and management of Cornelia de Lange syndrome: first international consensus statement*. Nat Rev Genet. (2018). 19(10):649-66. doi: 10.1038/s41576-018-0031-0. PMID: 29995837.

112. Mehta D, Vergano SAS, Deardorff M, Aggarwal S, Barot A, Johnson DM, et al. *Characterization of limb differences in children with Cornelia de Lange Syndrome*. Am J Med Genet C Semin Med Genet. (2016). 172(2):155-62. doi: 10.1002/ajmg.c.31498. PMID: 27120260.

113. Schwartz ID, Schwartz KJ, Kousseff BG, Bercu BB, Root AW. *Endocrinopathies in Cornelia de Lange syndrome*. J Pediatr. (1990).117(6):920-3. doi: 10.1016/s0022-3476(05)80137-4. PMID: 2246693.

114. Ascaso Á, Latorre-Pellicer A, Puisac B, Trujillano L, Arnedo M, Parenti I, et al. *Endocrine Evaluation and Homeostatic Model Assessment in Patients with Cornelia de Lange Syndrome*. J Clin Res Pediatr Endocrinol. (2024). 16(2):211-7. doi: 10.4274/jcrpe.galenos. PMID: 36482071.

115. Pablo MJ, Pamplona P, Haddad M, Benavente I, Latorre-Pellicer A, Arnedo M, et al. *High rate of autonomic neuropathy in Cornelia de Lange Syndrome*. Orphanet J Rare Dis. (2021). 16(1):458. doi: 10.1186/s13023-021-02082-y. PMID: 34717699.

116. Ayerza Casas A, Puisac Uriol B, Teresa Rodrigo ME, Hernández Marcos M, Ramos Fuentes FJ, Pie Juste J. *Cornelia de Lange syndrome: Congenital heart disease in 149 patients*. Med Clin (Barc). (2017). 149(7):300-2. doi: 10.1016/j.medcli.2017.03.051. PMID: 28629661.

117. Schrier SA, Sherer I, Deardorff MA, Clark D, Audette L, Gillis L, et al. *Causes of death and autopsy findings in a large study cohort of individuals with Cornelia de Lange syndrome and review of the literature*. Am J Med Genet A. (2011). 155A(12):3007-24. doi: 10.1002/ajmg.a.34329. PMID: 22069164.

118. Ayerza-Casas A, Puisac-Uriol B, Pie-Juste J. *Cornelia de Lange syndrome: Ventricular size and function in six children without congenital heart defects*. Med Clin (Barc). (2020). 154(2):67-8. doi: 10.1016/j.medcli.2018.10.001. PMID: 30470439.

119. Ajmone PF, Rigamonti C, Dall'Ara F, Monti F, Vizziello P, Milani D, et al. *Communication, cognitive development and behavior in children with Cornelia de Lange Syndrome (CdLS): preliminary results*. Am J Med Genet Part B Neuropsychiatr Genet Off Publ Int Soc Psychiatr Genet. (2014). 165B(3):223-9. doi: 10.1002/ajmg.b.32224. PMID: 24706566.

120. Parisi L, Di Filippo T, Roccella M. *Behavioral Phenotype and Autism Spectrum Disorders in Cornelia de Lange Syndrome*. Ment Illn. (2015). 7(2):5988. doi: 10.1097/YCO.0000000000000311. PMID: 28125439.

121. Groves L, Oliver C, Moss J. *Behaviour across the lifespan in Cornelia de Lange syndrome*. Curr Opin Psychiatry. (2021). 34(2):112-7. doi: 10.1097/YCO.0000000000000671. PMID: 33278154.

122. Srivastava S, Clark B, Landy-Schmitt C, Offermann EA, Kline AD, Wilkinson ST, et al. *Repetitive and Self-injurious Behaviors in Children with Cornelia de Lange Syndrome*. J Autism Dev Disord. (2021). 51(5):1748-58. doi: 10.1007/s10803-020-04617-x. PMID: 32809170.

123. Grados MA, Alvi MH, Srivastava S. *Behavioral and psychiatric manifestations in Cornelia de Lange syndrome*. Curr Opin Psychiatry. (2017). 30(2):92-6. doi: 10.1097/YCO.0000000000000311. PMID: 28125439.

124. Alonso-Gil D, Losada A. *NIPBL and cohesin: new take on a classic tale*. Trends Cell Biol. (2023). 33(10):860-71. doi: 10.1016/j.tcb.2023.03.006. PMID: 37062615.

125. Yatskevich S, Rhodes J, Nasmyth K. *Organization of Chromosomal DNA by SMC Complexes*. Annu Rev Genet. (2019). 53:445-82. doi: 10.1146/annurev-genet-112618-043633. PMID: 31577909.

126. Davidson IF, Peters JM. *Genome folding through loop extrusion by SMC complexes*. Nat Rev Mol Cell Biol. (2021). 22(7):445-64. doi: 10.1038/s41580-021-00349-7. PMID: 33767413.

127. Alonso-Gil D, Cuadrado A, Giménez-Llorente D, Rodríguez-Corsino M, Losada A. *Different NIPBL requirements of cohesin-STAG1 and cohesin-STAG2*. Nat Commun. (2023). 14(1):1326. doi: 10.1038/s41467-023-36900-7. PMID: 36898992.

128. Parenti I, Diab F, Gil SR, Mulugeta E, Casa V, Berutti R, et al. *MAU2 and NIPBL Variants Impair the Heterodimerization of the Cohesin Loader Subunits and Cause Cornelia de Lange Syndrome*. Cell Rep. (2020). 31(7):107647. doi: 10.1016/j.celrep.2020.107647. PMID: 32433956.

129. Rittenhouse NL, Dowen JM. *Cohesin regulation and roles in chromosome structure and function*. Curr Opin Genet Dev. (2024). 85:102159. doi: 10.1016/j.gde.2024.102159. PMID: 38382406.

130. Davidson IF, Bauer B, Goetz D, Tang W, Wutz G, Peters JM. *DNA loop extrusion by human cohesin*. Science. (2019). 366(6471):1338-45. doi: 10.1126/science.aaz3418. PMID: 31753851.

131. Bauer BW, Davidson IF, Canena D, Wutz G, Tang W, Litos G, et al. *Cohesin mediates DNA loop extrusion by a «swing and clamp» mechanism*. Cell. (2021). 184(21):5448-5464.e22. doi: 10.1016/j.cell.2021.09.016. PMID: 34624221.

132. Shi Z, Gao H, Bai XC, Yu H. *Cryo-EM structure of the human cohesin-NIPBL-DNA complex. Science*. (2020). 368(6498):1454-9. doi: 10.1126/science.abb0981. PMID: 32409525.

133. Krantz ID, McCallum J, DeScipio C, Kaur M, Gillis LA, Yaeger D, et al. *Cornelia de Lange syndrome is caused by mutations in NIPBL, the human homolog of Drosophila melanogaster Nipped-B*. Nat Genet.(2004). 36(6):631-5. doi: 10.1038/ng1364. PMID: 15146186.

134. Tonkin ET, Wang TJ, Lisgo S, Bamshad MJ, Strachan T. *NIPBL, encoding a homolog of fungal Scc2-type sister chromatid cohesion proteins and fly Nipped-B, is mutated in Cornelia de Lange syndrome*. Nat Genet. (2004). 36(6):636-41. doi: 10.1038/ng1363. PMID: 15146185.

135. Puisac B, Teresa-Rodrigo ME, Hernández-Marcos M, Baquero-Montoya C, Gil-Rodríguez MC, Visnes T, et al. mRNA *Quantification of NIPBL Isoforms A and B in Adult and Fetal Human Tissues, and a Potentially Pathological Variant Affecting Only Isoform A in Two Patients with Cornelia de Lange Syndrome*. Int J Mol Sci. (2017). 18(3):481. doi: 10.3390/ijms18030481. PMID: 28241484.

136. Kikuchi S, Borek DM, Otwinowski Z, Tomchick DR, Yu H. *Crystal structure of the cohesin loader Scc2 and insight into cohesinopathy*. Proc Natl Acad Sci U S A. (2016).113(44):12444-9. doi: 10.1073/pnas.1611333113. PMID: 27791135.

137. Gao D, Zhu B, Cao X, Zhang M, Wang X. *Roles of NIPBL in maintenance of genome stability*. Semin Cell Dev Biol. (2019). 90:181-6. doi: 10.1016/j.semcdb.2018.08.005. PMID: 30096364.

138. Ratajska M, Wierzba J, Pehlivan D, Xia Z, Brundage EK, Cheung SW, et al. *Cornelia de Lange syndrome case due to genomic rearrangements including NIPBL*. Eur J Med Genet. (2010). 53(6):378-82. doi: 10.1016/j.ejmg.2010.08.002. PMID: 20727427.

139. Teresa-Rodrigo ME, Eckhold J, Puisac B, Dalski A, Gil-Rodríguez MC, Braunholz D, et al. Functional characterization of NIPBL physiological splice variants and eight splicing mutations in patients with Cornelia de Lange syndrome. Int J Mol Sci. (2014). 15(6):10350-64. doi: 10.3390/ijms150610350. PMID: 24918291.

140. Cheng YW, Tan CA, Minor A, Arndt K, Wysinger L, Grange DK, et al. *Copy number analysis of NIPBL in a cohort of 510 patients reveals rare copy number variants and a mosaic deletion.* Mol Genet Genomic Med. (2014). 2(2):115-23. doi: 10.1002/mgg3.48. PMID: 24689074.

141. Krawczynska N, Kuzniacka A, Wierzba J, Parenti I, Kaiser FJ, Wasag B. *Mosaic Intronic NIPBL Variant in a Family With Cornelia de Lange Syndrome*. Front Genet. (2018). 9:255. doi: 10.3389/fgene.2018.00255. PMID: 30057591.

142. Kaur M, Mehta D, Noon SE, Deardorff MA, Zhang Z, Krantz ID. *NIPBL expression levels in CdLS probands as a predictor of mutation type and phenotypic severity*. Am J Med Genet C Semin Med Genet. (2016). 172(2):163-70. doi: 10.1002/ajmg.c.31495. PMID: 27125329.

143. Gillis LA, McCallum J, Kaur M, DeScipio C, Yaeger D, Mariani A, et al. *NIPBL mutational analysis in 120 individuals with Cornelia de Lange syndrome and evaluation of genotype-phenotype correlations*. Am J Hum Genet. (2004). 75(4):610-23. doi: 10.1086/424698. PMID: 15318302.

144. Yan J, Saifi GM, Wierzba TH, Withers M, Bien-Willner GA, Limon J, et al. *Mutational and genotype-phenotype correlation analyses in 28 Polish patients with Cornelia de Lange syndrome*. Am J Med Genet A. (2006). 140(14):1531-41. doi: 10.1002/ajmg.a.31305. PMID: 16770807.

145. Selicorni A, Russo S, Gervasini C, Castronovo P, Milani D, Cavalleri F, et al. *Clinical score of 62 Italian patients with Cornelia de Lange syndrome and correlations with the presence and type of NIPBL mutation*. Clin Genet. (2007). 72(2):98-108. doi: 10.1111/j.1399-0004.2007.00832.x. PMID: 17661813.

146. Carrel L, Willard HF. *X-inactivation profile reveals extensive variability in X-linked gene expression in females*. Nature. (2005). 434(7031):400-4. doi: 10.1038/nature03479. PMID: 15772666.

147. Huisman S, Mulder PA, Redeker E, Bader I, Bisgaard AM, Brooks A, et al. *Phenotypes and genotypes in individuals with SMC1A variants*. Am J Med Genet A. (2017). 173(8):2108-25. doi: 10.1002/ajmg.a.38279. PMID: 28548707.

148. Barañano KW, Kimball A, Fong SL, Egense AS, Hudon C, Kline AD. *Further Characterization of SMC1A Loss of Function Epilepsy Distinct From Cornelia de Lange Syndrome*. J Child Neurol. (2022). 37(5):390-6. doi: 10.1177/08830738221081244. PMID: 35238682.

149. Kaur M, Blair J, Devkota B, Fortunato S, Clark D, Lawrence A, et al. *Genomic analyses in Cornelia de Lange Syndrome and related diagnoses: Novel candidate genes, genotype-phenotype correlations and common mechanisms*. Am J Med Genet A. (2023). 191(8):2113-31. doi: 10.1002/ajmg.a.63247. PMID: 37377026.

150. Musio A. *The multiple facets of the SMC1A gene*. Gene. (2020). 743:144612. doi: 10.1016/j.gene.2020.144612. PMID: 32222533.

151. Parmeggiani L, Stanzial F, Menna E, Boni E, Manzoni F, Benedicenti F, et al. *Early onset developmental and epileptic encephalopathy and Rett-like phenotype in a 15-year-old girl affected by Cornelia de Lange syndrome type 2 due to a SMC1A gene mutation*. Epilepsy Behav Rep. (2023). 24:100634. doi: 10.1016/j.ebr.2023.100634. PMID: 38076278.

152. Gibellato E, Cianci P, Mariani M, Parma B, Huisman S, Śmigiel R, et al. *SMC1A epilepsy syndrome: clinical data from a large international cohort*. Am J Med Genet A. (2024). 194(7):e63577. doi: 10.1002/ajmg.a.63577. PMID: 38421079.

153. Gil-Rodríguez MC, Deardorff MA, Ansari M, Tan CA, Parenti I, Baquero-Montoya C, et al. *De novo heterozygous mutations in SMC3 cause a range of Cornelia de Lange syndrome-overlapping phenotypes*. Hum Mutat. (2015). 36(4):454-62. doi: 10.1002/humu.22761. PMID: 25655089.

154. Gligoris TG, Scheinost JC, Bürmann F, Petela N, Chan KL, Uluocak P, et al. *Closing the cohesin ring: structure and function of its Smc3-kleisin interface*. Science. (2014). 346(6212):963-7. doi: 10.1126/science.1256917. PMID: 25414305.

155. Deardorff MA, Wilde JJ, Albrecht M, Dickinson E, Tennstedt S, Braunholz D, et al. *RAD21 mutations cause a human cohesinopathy*. Am J Hum Genet. (2012) 90(6):1014-27. doi: 10.1016/j.ajhg.2012.04.019. PMID: 22633399.

156. Kaiser FJ, Ansari M, Braunholz D, Concepción Gil-Rodríguez M, Decroos C, Wilde JJ, et al. *Loss-of-function HDAC8 mutations cause a phenotypic spectrum of Cornelia de Lange syndrome-like features, ocular hypertelorism, large fontanelle and X-linked inheritance*. Hum Mol Genet. (2014). 23(11):2888-900. doi: 10.1093/hmg/ddu002. PMID: 24403048.

157. Deardorff MA, Bando M, Nakato R, Watrin E, Itoh T, Minamino M, et al. *HDAC8 mutations in Cornelia de Lange syndrome affect the cohesin acetylation cycle*. Nature. (2012). 489(7415):313-7. doi: 10.1038/nature11316. PMID: 22885700.

158. Olley G, Ansari M, Bengani H, Grimes GR, Rhodes J, von Kriegsheim A, et al. *BRD4 interacts with NIPBL and BRD4 is mutated in a Cornelia de Lange-like syndrome*. Nat Genet. (2018). 50(3):329-32. doi: 10.1038/s41588-019-0448-1. PMID: 29379197.

159. Parenti I, Gervasini C, Pozojevic J, Graul-Neumann L, Azzollini J, Braunholz D, et al. *Broadening of cohesinopathies: exome sequencing identifies mutations in ANKRD11 in two patients with Cornelia de Lange-overlapping phenotype*. Clin Genet. (2016). 89(1):74-81. doi: 10.1111/cge.12564. PMID: 25652421.

160. Parenti I, Teresa-Rodrigo ME, Pozojevic J, Ruiz Gil S, Bader I, Braunholz D, et al. *Mutations in chromatin regulators functionally link Cornelia de Lange syndrome and clinically overlapping phenotypes*. Hum Genet. (2017). 136(3):307-20. doi: 10.1007/s00439-017-1758-y. PMID: 28120103.

161. Cucco F, Sarogni P, Rossato S, Alpa M, Patimo A, Latorre A, et al. *Pathogenic variants in EP300 and ANKRD11 in patients with phenotypes overlapping Cornelia de Lange syndrome*. Am J Med Genet A. (2020). 182(7):1690-6. doi: 10.1002/ajmg.a.61611. PMID: 32476269.

162. Shangguan H, Chen R. *Phenotypes of Cornelia de Lange syndrome caused by non-cohesion genes: Novel variants and literature review*. Front Pediatr. (2022). 10:940294. doi: 10.3389/fped.2022.940294. PMID: 35935361.

163. Sirmaci A, Spiliopoulos M, Brancati F, Powell E, Duman D, Abrams A, et al. *Mutations in ANKRD11 cause KBG syndrome, characterized by intellectual disability, skeletal malformations, and macrodontia*. Am J Hum Genet. (2011). 89(2):289-94. doi: 10.1016/j.ajhg.2011.06.007. PMID: 21782149.

164. Parenti I, Mallozzi MB, Hüning I, Gervasini C, Kuechler A, Agolini E, et al. *ANKRD11 variants: KBG syndrome and beyond*. Clin Genet. (2021). 100(2):187-200. doi: 10.1111/cge.13977. PMID: 33955014.

165. Murayama Y, Uhlmann F. *Biochemical reconstitution of topological DNA binding by the cohesin ring*. Nature. (2014). 505(7483):367-71. doi: 10.1038/nature12867. PMID: 24291789.

166. Gutierrez-Escribano P, Newton MD, Llauró A, Huber J, Tanasie L, Davy J, et al. *A conserved ATP- and Scc2/4-dependent activity for cohesin in tethering DNA molecules*. Sci Adv. (2019). 5(11):eaay6804. doi: 10.1126/sciadv.aay6804. PMID: 31807710.

167. Kurokawa Y, Murayama Y. *DNA Binding by the Mis4Scc2 Loader Promotes Topological DNA Entrapment by the Cohesin Ring*. Cell Rep. (2020). 33(6):108357. doi: 10.1016/j.celrep.2020.108357. PMID: 33176147.

168. Peng Y, Zhu Y, Wu L, Deng F. *Clinical study and genetic analysis of Cornelia de Lange syndrome caused by a novel MAU2 gene variant in a Chinese boy*. Mol Genet Genomic Med. (2024). 12(1):e2318. doi: 10.1002/mgg3.2318. PMID: 37962004.

169. Remeseiro S, Cuadrado A, Kawauchi S, Calof AL, Lander AD, Losada A. *Reduction of Nipbl impairs cohesin loading locally and affects transcription but not cohesion-dependent functions in a mouse model of Cornelia de Lange Syndrome*. Biochim Biophys Acta. (2013). 1832(12):2097-102. doi: 10.1016/j.bbadis.2013.07.020. PMID: 23920377.

170. Kawauchi S, Calof AL, Santos R, Lopez-Burks ME, Young CM, Hoang MP, et al. *Multiple organ system defects and transcriptional dysregulation in the Nipbl(+/-) mouse, a model of Cornelia de Lange Syndrome*. PLoS Genet. (2009). 5(9):e1000650. doi: 10.1371/journal.pgen.1000650. PMID: 19763162.

171. Yuan B, Pehlivan D, Karaca E, Patel N, Charng WL, Gambin T, et al. *Global transcriptional disturbances underlie Cornelia de Lange syndrome and related phenotypes*. J Clin Invest. (2015). 125(2):636-51. doi: 10.1172/JCI77435. PMID: 25574841.

172. García-Gutiérrez P, García-Domínguez M. *BETting on a Transcriptional Deficit as the Main Cause for Cornelia de Lange Syndrome*. Front Mol Biosci. (2021). 8:709232. doi: 10.3389/fmolb.2021.709232. PMID: 34386522.

173. Jyonouchi S, Orange J, Sullivan KE, Krantz I, Deardorff M. *Immunologic features of Cornelia de Lange syndrome.* Pediatrics. (2013). 132(2):e484-489. doi: 10.1542/peds.2012-3815. PMID: 23821697.

174. Björkman A, Du L, van der Burg M, Cormier-Daire V, Borck G, Pié J, et al. *Reduced immunoglobulin gene diversity in patients with Cornelia de Lange syndrome*. J Allergy Clin Immunol. (2018). 141(1):408-411.e8. doi: 10.1016/j.jaci.2017.06.043. PMID: 28826797.

175. Mills JA, Herrera PS, Kaur M, Leo L, McEldrew D, Tintos-Hernandez JA, et al. *NIPBL+/-haploinsufficiency reveals a constellation of transcriptome disruptions in the pluripotent and cardiac states*. Sci Rep. (2018). 8(1):1056. doi: 10.1038/s41598-018-19173-9. PMID: 29348408.

176. Weiss FD, Calderon L, Wang YF, Georgieva R, Guo Y, Cvetesic N, et al. *Neuronal genes deregulated in Cornelia de Lange Syndrome respond to removal and re-expression of cohesin*. Nat Commun. (2021). 12(1):2919. doi: 10.1038/s41467-021-23141-9. PMID: 34006846.

177. Gu W, Wang L, Gu R, Ouyang H, Bao B, Zheng L, et al. *Defects of cohesin loader lead to bone dysplasia associated with transcriptional disturbance*. J Cell Physiol. (2021). 236(12):8208-25. doi: 10.1002/jcp.30491. Epub 2021 Jun 25. PMID: 34170011.

178. Perea-Resa C, Wattendorf L, Marzouk S, Blower MD. *Cohesin: behind dynamic genome topology and gene expression reprogramming*. Trends Cell Biol. (2021). 31(9):760-73. doi: 10.1016/j.tcb.2021.03.005. Epub 2021 Mar 22. PMID: 33766521.

179. Pileggi S, La Vecchia M, Colombo EA, Fontana L, Colapietro P, Rovina D, et al. *Cohesin Mutations Induce Chromatin Conformation Perturbation of the H19/IGF2 Imprinted Region and Gene Expression Dysregulation in Cornelia de Lange Syndrome Cell Lines*. Biomolecules. (2021). 11(11):1622. doi: 10.3390/biom11111622. PMID: 34827619.

180. Singh VP, McKinney S, Gerton JL. *Persistent DNA Damage and Senescence in the Placenta Impacts Developmental Outcomes of Embryos*. Dev Cell. (2020). 54(3):333-347.e7. doi: 10.1016/j.devcel.2020.05.025. PMID: 32800293.

181. Arnould C, Rocher V, Finoux AL, Clouaire T, Li K, Zhou F, et al. *Loop extrusion as a mechanism for formation of DNA damage repair foci*. Nature. (2021). 590(7847):660-5. doi: 10.1038/s41586-021-03193-z. PMID: 33597753.

182. Olley G, Pradeepa MM, Grimes GR, Piquet S, Polo SE, FitzPatrick DR, et al. *Cornelia de Lange syndrome-associated mutations cause a DNA damage signalling and repair defect*. Nat Commun. (2021). 12(1):3127. doi: 10.1038/s41467-021-23500-6. PMID: 34035299.

183. Latorre-Pellicer A, Ascaso Á, Trujillano L, Gil-Salvador M, Arnedo M, Lucia-Campos C, et al. *Evaluating Face2Gene as a Tool to Identify Cornelia de Lange Syndrome by Facial Phenotypes*. Int J Mol Sci. (2020). 21(3):1042. doi: 10.3390/ijms21031042. PMID: 32033219.

184. Hsieh TC, Bar-Haim A, Moosa S, Ehmke N, Gripp KW, Pantel JT, et al. *GestaltMatcher facilitates rare disease matching using facial phenotype descriptors*. Nat Genet. (2022). 54(3):349-57. doi: 10.1038/s41588-021-01010-x. PMID: 35145301.

185. Niu DM, Huang JY, Li HY, Liu KM, Wang ST, Chen YJ, et al. *Paternal gonadal mosaicism of NIPBL mutation in a father of siblings with Cornelia de Lange syndrome*. Prenat Diagn. (2006). 26(11):1054-7. doi: 10.1002/pd.1554. PMID: 16958143.

186. Masciadri M, Ficcadenti A, Milani D, Cogliati F, Divizia MT, Larizza L, et al. *Recurrence and Familial Inheritance of Intronic NIPBL Pathogenic Variant Associated With Mild CdLS*. Front Neurol. (2018). 9:967. doi: 10.3389/fneur.2018.00967. PMID: 30538663.

187. Huisman SA, Redeker EJW, Maas SM, Mannens MM, Hennekam RCM. *High rate of mosaicism in individuals with Cornelia de Lange syndrome*. J Med Genet. (2013). 50(5):339-44. doi: 10.1136/jmedgenet-2012-101477. PMID: 23505322.

188. 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. doi: 10.1136/jmedgenet-2014-102573. PMID: 25125236.

189. Latorre-Pellicer A, Gil-Salvador M, Parenti I, Lucia-Campos C, Trujillano L, Marcos-Alcalde I, et al. *Clinical relevance of postzygotic mosaicism in Cornelia de Lange syndrome and purifying selection of NIPBL variants in blood*. Sci Rep. (2021). 11(1):15459. doi: 10.1038/s41598-021-94958-z. PMID: 34326454.

190. Mannini L, Cucco F, Quarantotti V, Krantz ID, Musio A. *Mutation spectrum and genotype-phenotype correlation in Cornelia de Lange syndrome*. Hum Mutat. (2013) 34(12):1589-96. doi: 10.1002/humu.22430. PMID: 24038889.

191. Kang MJ, Ahn SM, Hwang IT. *A Novel Frameshift Mutation (c.5387_5388insTT) in NIPBL in Cornelia de Lange Syndrome with Severe Phenotype*. Ann Clin Lab Sci. (2018). 48(1):106-9. PMID: 29531005.

192. Pié J, Puisac B, Hernández-Marcos M, Teresa-Rodrigo ME, Gil-Rodríguez M, Baquero-Montoya C, et al. *Special cases in Cornelia de Lange syndrome: The Spanish experience*. Am J Med Genet C Semin Med Genet. (2016). 172(2):198-205. doi: 10.1002/ajmg.c.31501. PMID: 27164022.

193. Krawczynska N, Wierzba J, Jasiecki J, Wasag B. *Molecular characterization of two novel intronic variants of NIPBL gene detected in unrelated Cornelia de Lange syndrome patients*. BMC Med Genet. (2019). 20(1):1. doi: 10.1186/s12881-018-0738-y. PMID: 30606125.

194. Coursimault J, Cassinari K, Lecoquierre F, Quenez O, Coutant S, Derambure C, et al. *Deep intronic NIPBL de novo mutations and differential diagnoses revealed by whole genome and RNA sequencing in Cornelia de Lange syndrome patients*. Hum Mutat. (2022). 43(12):1882-97. doi: 10.1002/humu.24438. PMID: 35842780.

195. Pehlivan D, Hullings M, Carvalho CMB, Gonzaga-Jauregui CG, Loy E, Jackson LG, et al. *NIPBL rearrangements in Cornelia de Lange syndrome: evidence for replicative mechanism and genotype-phenotype correlation*. Genet Med Off J Am Coll Med Genet. (2012). 14(3):313-22. doi: 10.1038/gim.2011.13. PMID: 22241092.

196. Baquero-Montoya C, Gil-Rodríguez MC, Teresa-Rodrigo ME, Hernández-Marcos M, Bueno-Lozano G, Bueno-Martínez I, et al. *Could a patient with SMC1A duplication be classified as a human cohesinopathy?* Clin Genet. (2014). 85(5):446-51. doi: 10.1111/cge.12194. PMID: 23683030.

197. Helgeson M, Keller-Ramey J, Knight Johnson A, Lee JA, Magner DB, Deml B, et al. *Molecular characterization of HDAC8 deletions in individuals with atypical Cornelia de Lange syndrome*. J Hum Genet. (2018). 63(3):349-56. doi: 10.1038/s10038-017-0387-6. PMID: 29279609.

198. Bestetti I, Crippa M, Sironi A, Bellini M, Tumiatti F, Ballabio S, et al. *Long-read sequencing reveals chromothripsis in a molecularly unsolved case of Cornelia de Lange syndrome*. Front Genet. (2024).15:1358334. doi: 10.3389/fgene.2024.1358334. PMID: 38544803.

199. Kernohan KD, Boycott KM. *The expanding diagnostic toolbox for rare genetic diseases*. Nat Rev Genet. (2024). 25(6):401-15. doi: 10.1038/s41576-023-00683-w. PMID: 38238519.

200. Rentas S, Rathi KS, Kaur M, Raman P, Krantz ID, Sarmady M, et al. *Diagnosing Cornelia de Lange syndrome and related neurodevelopmental disorders using RNA sequencing*. Genet Med Off J Am Coll Med Genet. (2020). 22(5):927-36. doi: 10.1038/s41436-019-0741-5. PMID: 31911672.

201. Seyama R, Uchiyama Y, Ceroni JRM, Kim VEH, Furquim I, Honjo RS, et al. *Pathogenic variants detected by RNA sequencing in Cornelia de Lange syndrome*. Genomics. (2022). 114(5):110468. doi: 10.1016/j.ygeno.2022.110468. PMID: 36041635.

202. Deardorff MA, Kaur M, Yaeger D, Rampuria A, Korolev S, Pie J, et al. *Mutations in cohesin complex members SMC3 and SMC1A cause a mild variant of cornelia de Lange syndrome with predominant mental retardation*. Am J Hum Genet. (2007). 80(3):485-94. doi: 10.1086/511888. PMID: 17273969.

203. Ford CE, Hamerton JL. *The chromosomes of man*. Nature. (1956). 178(4541):1020-3. doi: 10.1038/1781020a0. PMID: 13378517.

204. Bauman JG, Wiegant J, Borst P, van Duijn P. *A new method for fluorescence microscopical localization of specific DNA sequences by in situ hybridization of fluorochromelabelled RNA*. Exp Cell Res. (1980).128(2):485-90. doi: 10.1016/0014-4827(80)90087-7. PMID: 6157553.

205. Schena M, Shalon D, Davis RW, Brown PO. *Quantitative monitoring of gene expression patterns with a complementary DNA microarray*. Science. (1995). 270(5235):467-70. doi: 10.1126/science.270.5235.467. PMID: 7569999.

206. Schouten JP, McElgunn CJ, Waaijer R, Zwijnenburg D, Diepvens F, Pals G. *Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification*. Nucleic Acids Res. (2002). 30(12):e57. doi: 10.1093/nar/gnf056. PMID: 12060695.

207. Deschamps S, Zhang Y, Llaca V, Ye L, Sanyal A, King M, et al. *A chromosome-scale assembly of the sorghum genome using nanopore sequencing and optical mapping*. Nat Commun. (2018). 9(1):4844. doi: 10.1038/s41467-018-07271-1. PMID: 30451840.

208. Sanger F, Nicklen S, Coulson AR. *DNA sequencing with chain-terminating inhibitors*. Proc Natl Acad Sci U S A. (1977). 74(12):5463-7. doi: 10.1073/pnas.74.12.5463. PMID: 271968.

209. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, et al. *Genome sequencing in microfabricated high-density picolitre reactors*. Nature. (2005). 437(7057):376-80. doi: 10.1038/nature03959. PMID: 16056220.

210. Shendure J, Porreca GJ, Reppas NB, Lin X, McCutcheon JP, Rosenbaum AM, et al. *Accurate multiplex polony sequencing of an evolved bacterial genome*. Science. (2005). 309(5741):1728-32. doi: 10.1126/science.1117389. PMID: 16081699.

211. Rosenstein J. *The promise of nanopore technology: Nanopore DNA sequencing represents a fundamental change in the way that genomic information is read, with potentially big savings*. IEEE Pulse. (2014). 5(4):52-4. doi: 10.1109/MPUL.2014.2321214. PMID: 25029683.

212. Schuurs-Hoeijmakers JHM, Oh EC, Vissers LELM, Swinkels MEM, Gilissen C, Willemsen MA, et al. *Recurrent de novo mutations in PACS1 cause defective cranial-neural-crest migration and define a recognizable intellectual-disability syndrome*. Am J Hum Genet. (2012). 91(6):1122-7. doi: 10.1016/j.ajhg.2012.10.013. PMID: 23159249.

213. Izumi K, Nakato R, Zhang Z, Edmondson AC, Noon S, Dulik MC, et al. *Germline gain-of-function mutations in AFF4 cause a developmental syndrome functionally linking the super elongation complex and cohesin*. Nat Genet. (2015). 47(4):338-44. doi: 10.1038/ng.3229. PMID: 25730767.

214. Ellingford JM, Ahn JW, Bagnall RD, Baralle D, Barton S, Campbell C, et al. *Recommendations for clinical interpretation of variants found in non-coding regions of the genome*. Genome Med. (2022). 14(1):73. doi: 10.1186/s13073-022-01073-3. PMID: 35850704.

215. Chen Y, Dawes R, Kim HC, Ljungdahl A, Stenton SL, Walker S, et al. *De novo variants in the RNU4-2 snRNA cause a frequent neurodevelopmental syndrome*. Nature. (2024). 632(8026):832-40. doi: 10.1038/s41586-024-07773-7. PMID: 38991538.

216. Boycott KM, Rath A, Chong JX, Hartley T, Alkuraya FS, Baynam G, et al. *International Cooperation to Enable the Diagnosis of All Rare Genetic Diseases*. Am J Hum Genet. (2017). 100(5):695-705. doi: 10.1016/j.ajhg.2017.04.003. PMID: 28475856.

217. Zhu Q, Nguyen DT, Alyea G, Hanson K, Sid E, Pariser A. *Phenotypically Similar Rare Disease Identification from an Integrative Knowledge Graph for Data Harmonization: Preliminary Study*. JMIR Med Inform. (2020). 8(10):e18395. doi: 10.2196/18395. PMID: 33006565.

218. Robinson PN, Köhler S, Bauer S, Seelow D, Horn D, Mundlos S. *The Human Phenotype Ontology: a tool for annotating and analyzing human hereditary disease*. Am J Hum Genet. (2008). 83(5):610-5. doi: 10.1016/j.ajhg.2008.09.017. PMID: 18950739.

219. Groza T, Köhler S, Moldenhauer D, Vasilevsky N, Baynam G, Zemojtel T, et al. T*he Human Phenotype Ontology: Semantic Unification of Common and Rare Disease*. Am J Hum Genet. (2015). 97(1):111-24. doi: 10.1016/j.ajhg.2015.05.020. PMID: 26119816.

220. Köhler S, Vasilevsky NA, Engelstad M, Foster E, McMurry J, Aymé S, et al. *The Human Phenotype Ontology in 2017*. Nucleic Acids Res. (2017). 45(D1):D865-76. doi: 10.1093/nar/gkw1039. PMID: 27899602.

221. Wright CF, Campbell P, Eberhardt RY, Aitken S, Perrett D, Brent S, et al. *Genomic Diagnosis of Rare Pediatric Disease in the United Kingdom and Ireland*. N Engl J Med. (2023). 388(17):1559-71. doi: 10.1056/NEJMoa2209046. PMID: 37043637.

222. Schuurs-Hoeijmakers JHM, Landsverk ML, Foulds N, Kukolich MK, Gavrilova RH, Greville-Heygate S, et al. *Clinical delineation of the PACS1-related syndrome--Report on 19 patients*. Am J Med Genet A. (2016). 170(3):670-5. doi: 10.1002/ajmg.a.37476. PMID: 26842493.

223. Youker RT, Shinde U, Day R, Thomas G. *At the crossroads of homoeostasis and disease: roles of the PACS proteins in membrane traffic and apoptosis*. Biochem J. (2009). 421(1):1-15. doi: 10.1042/BJ20081016. PMID: 19505291.

224. Thomas G, Aslan JE, Thomas L, Shinde P, Shinde U, Simmen T. *Caught in the act - protein adaptation and the expanding roles of the PACS proteins in tissue homeostasis and disease*. J Cell Sci. (2017). 130(11):1865-76. doi: 10.1242/jcs.199463. PMID: 28476937.

225. Mani C, Tripathi K, Luan S, Clark DW, Andrews JF, Vindigni A, et al. T*he multifunctional protein PACS-1 is required for HDAC2- and HDAC3-dependent chromatin maturation and genomic stability*. Oncogene. (2020). 39(12):2583-96. doi: 10.1038/s41388-020-1167-x. PMID: 31988453.

226. Nair-Gill E, Bonora M, Zhong X, Liu A, Miranda A, Stewart N, et al. *Calcium flux control by Pacs1-Wdr37 promotes lymphocyte quiescence and lymphoproliferative diseases*. EMBO J. 3 (2021). 40(9):e104888. doi: 10.15252/embj.2020104888. PMID: 33630350.

227. Wan L, Molloy SS, Thomas L, Liu G, Xiang Y, Rybak SL, et al. *PACS-1 defines a novel gene family of cytosolic sorting proteins required for trans-Golgi network localization*. Cell. (1998). 94(2):205-16. doi: 10.1016/s0092-8674(00)81420-8. PMID: 9695949.

228. Simmen T, Aslan JE, Blagoveshchenskaya AD, Thomas L, Wan L, Xiang Y, et al. *PACS-2 controls endoplasmic reticulum-mitochondria communication and Bid-mediated apoptosis*. EMBO J. (2005). 24(4):717-29. doi: 10.1038/sj.emboj.7600559. PMID: 15692567.

229. Brasacchio D, Alsop AE, Noori T, Lufti M, Iyer S, Simpson KJ, et al. *Epigenetic control of mitochondrial cell death through PACS1-mediated regulation of BAX/BAK oligomerization*. Cell Death Differ. (2017). 24(6):961-70. doi: 10.1038/cdd.2016.119. PMID: 28060382.

230. Li C, Li L, Yang M, Zeng L, Sun L. *PACS-2: A key regulator of mitochondria-associated membranes (MAMs)*. Pharmacol Res. (2020). 160:105080. doi: 10.1016/j.phrs.2020.105080. PMID: 32673704.

231. Olson HE, Jean-Marçais N, Yang E, Heron D, Tatton-Brown K, van der Zwaag PA, et al. *A Recurrent De Novo PACS2 Heterozygous Missense Variant Causes Neonatal-Onset Developmental Epileptic Encephalopathy, Facial Dysmorphism, and Cerebellar Dysgenesis*. Am J Hum Genet. (2018). 102(5):995-1007. doi: 10.1016/j.ajhg.2018.09.002. PMID: 29656858.

232. Kanca O, Andrews JC, Lee PT, Patel C, Braddock SR, Slavotinek AM, et al. *De Novo Variants in WDR37 Are Associated with Epilepsy, Colobomas, Dysmorphism, Developmental Delay, Intellectual Disability, and Cerebellar Hypoplasia*. Am J Hum Genet. (2019). 105(2):413-24. doi: 10.1016/j.ajhg.2019.07.017. PMID: 31327508.

233. Reis LM, Sorokina EA, Thompson S, Muheisen S, Velinov M, Zamora C, et al. *De Novo Missense Variants in WDR37 Cause a Severe Multisystemic Syndrome*. Am J Hum Genet. (2019). 105(2):425-33. doi: 10.1016/j.ajhg.2019.06.015. PMID: 31327510.

234. Dentici ML, Barresi S, Niceta M, Ciolfi A, Trivisano M, Bartuli A, et al. *Expanding the clinical spectrum associated with PACS2 mutations*. Clin Genet. (2019). 95(4):525-31. doi: 10.1111/cge.13516. PMID: 30684285.

235. Terrone G, Marchese F, Vari MS, Severino M, Madia F, Amadori E, et al. *A further contribution to the delineation of epileptic phenotype in PACS2-related syndrome*. Seizure. (2020). 79:53-5. doi: 10.1016/j.seizure.2020.05.001. PMID: 32416568.

236. Hay E, Henderson RH, Mansour S, Deshpande C, Jones R, Nutan S, et al. *Expanding the phenotypic spectrum consequent upon de novo WDR37 missense variants*. Clin Genet. (2020). 98(2):191-7. doi: 10.1111/cge.13795. PMID: 32530092.

237. Lu CL, Kim J. *Craniofacial Diseases Caused by Defects in Intracellular Trafficking*. Genes. (2021). 12(5):726. doi: 10.3390/genes12050726. PMID: 34068038.

238. García-Cazorla A, Oyarzábal A, Saudubray JM, Martinelli D, Dionisi-Vici C. *Genetic disorders of cellular trafficking*. Trends Genet TIG. (2022). 38(7):724-51. doi: 10.1016/j.tig.2022.02.012. PMID: 35367088.

239. 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 Off J Am Coll Med Genet. (2018). 20(10):1216-23. doi: 10.1038/gim.2017.246. PMID: 29323667.

240. Bullich G, Matalonga L, Pujadas M, Papakonstantinou A, Piscia D, Tonda R, et al. *Systematic Collaborative Reanalysis of Genomic Data Improves Diagnostic Yield in Neurologic Rare Diseases*. J Mol Diagn JMD. (2022). 24(5):529-42. doi: 10.1016/j.jmoldx.2022.02.003. PMID: 35569879.

241. Deignan JL, Chung WK, Kearney HM, Monaghan KG, Rehder CW, Chao EC, et al. *Points to consider in the reevaluation and reanalysis of genomic test results: a statement of the American College of Medical Genetics and Genomics (ACMG)*. Genet Med Off J Am Coll Med Genet.(2019). 21(6):1267-70. doi: 10.1038/s41436-019-0478-1. PMID: 31015575.

242. Baker SW, Murrell JR, Nesbitt AI, Pechter KB, Balciuniene J, Zhao X, et al. *Automated Clinical Exome Reanalysis Reveals Novel Diagnoses*. J Mol Diagn JMD. (2019). 21(1):38-48. doi: 10.1016/j.jmoldx.2018.07.008. PMID: 30577886.

243. Dai P, Honda A, Ewans L, McGaughran J, Burnett L, Law M, et al. *Recommendations for next generation sequencing data reanalysis of unsolved cases with suspected Mendelian disorders: A systematic review and meta-analysis*. Genet Med Off J Am Coll Med Genet. (2022). 24(8):1618-29. doi: 10.1016/j.gim.2022.04.021. PMID: 35550369.

244. Harrison SM, Dolinksy JS, Chen W, Collins CD, Das S, Deignan JL, et al. *Scaling resolution of variant classification differences in ClinVar between 41 clinical laboratories through an outlier approach*. Hum Mutat. (2018). 39(11):1641-9. doi: 10.1002/humu.23643. PMID: 30311378.

245. Kopanos C, Tsiolkas V, Kouris A, Chapple CE, Albarca Aguilera M, Meyer R, et al. *VarSome: the human genomic variant search engine*. Bioinforma Oxf Engl. (2019). 35(11):1978-80. doi: 10.1093/bioinformatics/bty897. PMID: 30376034.

246. Yan J, Zhang F, Brundage E, Scheuerle A, Lanpher B, Erickson RP, et al. *Genomic duplication resulting in increased copy number of genes encoding the sister chromatid cohesion complex conveys clinical consequences distinct from Cornelia de Lange*. J Med Genet. (2009). 46(9):626-34. doi: 10.1136/jmg.2008.062471. Epub 2008 Dec 3. PMID: 19052029.

247. Aoi H, Mizuguchi T, Ceroni JR, Kim VEH, Furquim I, Honjo RS, et al. *Comprehensive genetic analysis of 57 families with clinically suspected Cornelia de Lange syndrome*. J Hum Genet. (2019). 64(10):967-78. doi: 10.1038/s10038-019-0643-z. PMID: 31337854.

248. Dong X, Liu B, Yang L, Wang H, Wu B, Liu R, et al. *Clinical exome sequencing as the first-tier test for diagnosing developmental disorders covering both CNV and SNV: a Chinese cohort*. J Med Genet. (2020). 57(8):558-66. doi: 10.1136/jmedgenet-2019-106377. PMID: 32005694.

249. Lemire G, Sanchis-Juan A, Russell K, Baxter S, Chao KR, Singer-Berk M, et al. *Exome copy number variant detection, analysis and classification in a large cohort of families with undiagnosed rare genetic disease*. MedRxiv Prepr Serv Health Sci. (2023). 2023.10.05.23296595. doi: 10.1016/j.ajhg.2024.03.008. PMID: 38565148.

250. Hehir-Kwa JY, Wieskamp N, Webber C, Pfundt R, Brunner HG, Gilissen C, et al. *Accurate distinction of pathogenic from benign CNVs in mental retardation*. PLoS Comput Biol. (2010). 6(4):e1000752. doi: 10.1371/journal.pcbi.1000752. PMID: 20421931.

251. Parenti I, Gervasini C, Pozojevic J, Wendt KS, Watrin E, Azzollini J, et al. *Expanding the clinical spectrum of the «HDAC8-phenotype» - implications for molecular diagnostics, counseling and risk prediction*. Clin Genet. mayo de 2016;89(5):564-73. doi: 10.1111/cge.12717. PMID: 26671848.

252. Mio C, Passon N, Fogolari F, Cesario C, Novelli A, Pittini C, et al. *A novel de novo HDAC8 missense mutation causing Cornelia de Lange syndrome*. Mol Genet Genomic Med. (2021). 9(9):e1612. doi: 10.1002/mgg3.1612. PMID: 34342180.

253. Bi W, Borgan C, Pursley AN, Hixson P, Shaw CA, Bacino CA, et al. *Comparison of chromosome analysis and chromosomal microarray analysis: what is the value of chromosome analysis in today's genomic array era?* Genet Med Off J Am Coll Med Genet. (2013). 15(6):450-7. doi: 10.1038/gim.2012.152. PMID: 23238528.

254. Levy B, Wapner R. *Prenatal diagnosis by chromosomal microarray analysis*. Fertil Steril. (2018) 109(2):201-12. doi: 10.1038/gim.2012.152. PMID: 23238528.

255. Siggberg L, Ala-Mello S, Linnankivi T, Avela K, Scheinin I, Kristiansson K, et al. *High-resolution SNP array analysis of patients with developmental disorder and normal array CGH results*. BMC Med Genet. (2012). 13:84. doi: 10.1186/s12881-014-0124-3. PMID: 22984989.

256. Mantere T, Kersten S, Hoischen A. *Long-Read Sequencing Emerging in Medical Genetics*. Front Genet. (2019) 10:426. doi: 10.3389/fgene.2019.00426. PMID: 31134132.

257. Gabrielaite M, Torp MH, Rasmussen MS, Andreu-Sánchez S, Vieira FG, Pedersen CB, et al. *A Comparison of Tools for Copy-Number Variation Detection in Germline Whole Exome and Whole Genome Sequencing Data*. Cancers. (2021). 13(24):6283. doi: 10.3390/cancers13246283. PMID: 34944901.

258. Porubsky D, Höps W, Ashraf H, Hsieh P, Rodriguez-Martin B, Yilmaz F, et al. *Recurrent inversion polymorphisms in humans associate with genetic instability and genomic disorders*. Cell. (2022). 185(11):1986-2005.e26. doi: 10.1016/j.cell.2022.04.017. PMID: 35525246.

259. Cassinari K, Rovelet-Lecrux A, Tury S, Quenez O, Richard AC, Charbonnier C, et al. *Haploinsufficiency of the Primary Familial Brain Calcification Gene SLC20A2 Mediated by Disruption of a Regulatory Element*. Mov Disord Off J Mov Disord Soc. (2020). 35(8):1336-45. doi: 10.1002/mds.28090. PMID: 32506582.

260. Takahashi S, Tanaka R, Okano S, Okayama A, Suzuki N, Azuma H. *Characterization of intragenic tandem duplication in the PAFAH1B1 gene leading to isolated lissencephaly sequence*. Mol Cytogenet. 2015;8:84. doi: 10.1186/s13039-015-0186-8. PMID: 26523152.

261. Anna A, Monika G. *Splicing mutations in human genetic disorders: examples, detection, and confirmation*. J Appl Genet.(2018). 59(3):253-68. doi: 10.1007/s13353-019-00493-z. PMID: 29680930.

262. Potapova NA. *Nonsense Mutations in Eukaryotes*. Biochem Biokhimiia. (2022). 87(5):400-12. doi: 10.1134/S0006297922050029. PMID: 35790376.

263. Decroos C, Bowman CM, Moser JAS, Christianson KE, Deardorff MA, Christianson DW. *Compromised structure and function of HDAC8 mutants identified in Cornelia de Lange Syndrome spectrum disorders*. ACS Chem Biol. (2014). 9(9):2157-64. doi: 10.1021/cb5003762. PMID: 25075551.

264. Decroos C, Christianson NH, Gullett LE, Bowman CM, Christianson KE, Deardorff MA, et al. *Biochemical and structural characterization of HDAC8 mutants associated with Cornelia de Lange syndrome spectrum disorders*. Biochemistry. (2015). 54(42):6501-13. doi: 10.1021/acs.biochem.5b00881. PMID: 26463496.

265. Deardorff MA, Porter NJ, Christianson DW. *Structural aspects of HDAC8 mechanism and dysfunction in Cornelia de Lange syndrome spectrum disorders*. Protein Sci Publ Protein Soc. (2016). 25(11):1965-76. doi: 10.1002/pro.3030. PMID: 27576763.

266. Osko JD, Porter NJ, Decroos C, Lee MS, Watson PR, Raible SE, et al. *Structural analysis of histone deacetylase 8 mutants associated with Cornelia de Lange Syndrome spectrum disorders*. J Struct Biol. (2021). 213(1):107681. doi: 10.1016/j.jsb.2020.107681. PMID: 33316326.

267. Wu X, Xu LY, Li EM, Dong G. *Application of molecular dynamics simulation in biomedicine*. Chem Biol Drug Des. (2022). 99(5):789-800. doi: 10.1111/cbdd.14038. PMID: 35293126.

268. Ben-Asher E, Lancet D. *NIPBL gene responsible for Cornelia de Lange syndrome, a severe developmental disorder*. Isr Med Assoc J IMAJ. (2004). 6(9):571-2. PMID: 15373324.

269. Borck G, Redon R, Sanlaville D, Rio M, Prieur M, Lyonnet S, et al. *NIPBL mutations and genetic heterogeneity in Cornelia de Lange syndrome*. J Med Genet. (2004). 41(12):e128. doi: 10.1136/jmg.2004.026666. PMID: 15591270.

270. Rao SSP, Huang SC, Glenn St Hilaire B, Engreitz JM, Perez EM, Kieffer-Kwon KR, et al. *Cohesin Loss Eliminates All Loop Domains*. Cell. (2017). 171(2):305-320.e24. doi: 10.1016/j.cell.2017.09.026. PMID: 28985562.

271. Singh AK, Chen Q, Nguyen C, Meerzaman D, Singer DS. *Cohesin regulates alternative splicing*. Sci Adv. marzo de 2023;9(9):eade3876. doi: 10.1126/sciadv.ade3876. PMID: 36857449.

272. Hillman MA, Gecz J. *Fragile XE-associated familial mental retardation protein 2 (FMR2) acts as a potent transcription activator*. J Hum Genet. (2001). 46(5):251-9. doi: 10.1007/s100380170074. PMID: 11355014.

273. Melko M, Douguet D, Bensaid M, Zongaro S, Verheggen C, Gecz J, et al. *Functional characterization of the AFF (AF4/FMR2) family of RNA-binding proteins: insights into the molecular pathology of FRAXE intellectual disability*. Hum Mol Genet. (2011). 20(10):1873-85. doi: 10.1093/hmg/ddr069. PMID: 21330300.

274. Luo Z, Lin C, Guest E, Garrett AS, Mohaghegh N, Swanson S, et al. *The super elongation complex family of RNA polymerase II elongation factors: gene target specificity and transcriptional output*. Mol Cell Biol. (2012). 32(13):2608-17. doi: 10.1128/MCB.00182-12. PMID: 22547686.

275. Luo Z, Lin C, Shilatifard A. *The super elongation complex (SEC) family in transcriptional control*. Nat Rev Mol Cell Biol. (2012). 13(9):543-7. doi: 10.1038/nrm3417. PMID: 22895430.

276. Eigenhuis KN, Somsen HB, van den Berg DLC. *Transcription Pause and Escape in Neurodevelopmental Disorders.* Front Neurosci. (2022). 16:846272. doi: 10.3389/fnins.2022.846272. PMID: 35615272.

277. Gu Y, McIlwain KL, Weeber EJ, Yamagata T, Xu B, Antalffy BA, et al. *Impaired conditioned fear and enhanced long-term potentiation in Fmr2 knock-out mice*. J Neurosci Off J Soc Neurosci. (2002). 22(7):2753-63. doi: 10.1523/JNEUROSCI.22-07-02753.2002. PMID: 11923441.

278. Deneault E, White SH, Rodrigues DC, Ross PJ, Faheem M, Zaslavsky K, et al. *Complete Disruption of Autism-Susceptibility Genes by Gene Editing Predominantly Reduces Functional Connectivity of Isogenic Human Neurons*. Stem Cell Rep. (2018). 11(5):1211-25. doi: 10.1016/j.stemcr.2019.01.008. PMID: 30392976.

279. Mulley JC, Yu S, Loesch DZ, Hay DA, Donnelly A, Gedeon AK, et al. *FRAXE and mental retardation*. J Med Genet. (1995). 32(3):162-9. doi: 10.1136/jmg.32.3.162. PMID: 7783162.

280. Gecz J, Gedeon AK, Sutherland GR, Mulley JC. *Identification of the gene FMR2, associated with FRAXE mental retardation*. Nat Genet. (1996). 13(1):105-8. doi: 10.1038/ng0596-105. PMID: 8673085.

281. Sahoo T, Theisen A, Marble M, Tervo R, Rosenfeld JA, Torchia BS, et al. *Microdeletion of Xq28 involving the AFF2 (FMR2) gene in two unrelated males with developmental delay*. Am J Med Genet A. (2011). 155A(12):3110-5. doi: 10.1002/ajmg.a.34345. PMID: 22065534.

282. Valentino F, Bruno LP, Doddato G, Giliberti A, Tita R, Resciniti S, et al. *Exome Sequencing in 200 Intellectual Disability/Autistic Patients: New Candidates and Atypical Presentations*. Brain Sci. (2021). 11(7):936. doi: 10.3390/brainsci11070936. PMID: 34356170.

283. da Rocha NB, dos Santos PAC, Safatle HPN, de Melo RM, Pereira RW, de Oliveira SF, et al. *Partial AFF2 microduplication in a patient with auditory processing disorder, emotional impairment and macrosomia*. Am J Med Genet A. (2014). 164A(12):3206-8. doi: 10.1002/ajmg.a.36768. PMID: 25256661.

284. Sun Z, Fan J, Wang Y. *X-Chromosome Inactivation and Related Diseases*. Genet Res. (2022). 2022:1391807. doi: 10.1155/2022/1391807. PMID: 35387179.

285. Fieremans N, Van Esch H, Holvoet M, Van Goethem G, Devriendt K, Rosello M, et al. *Identification of Intellectual Disability Genes in Female Patients with a Skewed X-Inactivation Pattern*. Hum Mutat. (2016). 37(8):804-11. doi: 10.1002/humu.23012. PMID: 27159028.

286. Vianna EQ, Piergiorge RM, Gonçalves AP, Dos Santos JM, Calassara V, Rosenberg C, et al. *Understanding the Landscape of X-linked Variants Causing Intellectual Disability in Females Through Extreme X Chromosome Inactivation Skewing*. Mol Neurobiol. (2020). 57(9):3671-84. doi: 10.1007/s12035-020-01981-8. PMID: 32564284.

287. Łabno A, Tomecki R, Dziembowski A. *Cytoplasmic RNA decay pathways - Enzymes and mechanisms. Biochim Biophys Acta*. (2016). 1863(12):3125-47. doi: 10.1016/j.bbamcr.2016.09.023. PMID: 27713097.

288. Peck SA, Hughes KD, Victorino JF, Mosley AL. *Writing a wrong: Coupled RNA polymerase II transcription and RNA quality control*. Wiley Interdiscip Rev RNA. (2019). 10(4):e1529. doi: 10.1002/wrna.1529. PMID: 30848101.

Apéndices

Apéndice 1: Factor de impacto y áreas temáticas de las revistas

Tabla 2. Características de las revistas en las que han sido publicados los artículos que conforman esta tesis doctoral.

Trabajo	Año de publicación	Nombre de la revista	Área temática	Índice de impacto	Cuartil
1	2022	*International Journal of Molecular Sciences*	Biología Molecular	5,6	Q1
2	2021	*Molecular Genetics & Genomic Medicine*	Genética	2,473	Q3
3	2022	*Genes (Basel)*	Genética	3,5	Q2
4	2024	*Frontiers in Genetics*	Genética	2,8 (2023)	Q2 (2023)

Apéndice 2: Contribución de la doctoranda

Tabla 3. Contribución de la doctoranda en cada uno de los trabajos que incluye la tesis doctoral.

Trabajo	Título	Contribución
1	*Molecular Basis of the Schuurs-Hoeijmakers Syndrome: What We Know about the Gene and the PACS-1 Protein and Novel Therapeutic Approaches.*	▪ Concepción y diseño del estudio ▪ Revisión de la literatura ▪ Redacción del manuscrito
2	*Things are not always what they seem: F Cornelia de Lange to KBG phenotype in a with genetic variants in NIPBL and ANKR*	▪ Concepción y diseño del estudio ▪ Diseño, realización e interpretación de los análisis moleculares ▪ Revisión de la literatura ▪ Redacción del manuscrito
3	*Novel Intragenic Duplication in the HDAC Gene Underlying a Case of Cornelia de La Syndrome.*	▪ Concepción y diseño del estudio ▪ Diseño, realización e interpretación de los análisis moleculares ▪ Revisión de la literatura ▪ Redacción del manuscrito
4	*An intragenic duplication in the AFF2 gene associated with Cornelia de Lange syndrome phenotype*	▪ Concepción y diseño del estudio ▪ Diseño, realización e interpretación de los análisis moleculares ▪ Revisión de la literatura ▪ Redacción del manuscrito

Apéndice 3: Licencia BioRender

Las figuras de esta tesis doctoral se elaboraron utilizando BioRender (https://www.biorender.com/), bajo la licencia académica otorgada a Cristina Tania Lucia Campos por el Instituto de Investigación Sanitaria de Aragón, con el número de registro AA27NR4JCD.

Dicha licencia permite emplear el contenido de BioRender, incluyendo íconos, plantillas y diseños originales, en publicaciones académicas como artículos científicos, libros de texto y plataformas web. Todos los derechos sobre este material son propiedad exclusiva de BioRender. El uso del contenido en esta tesis doctoral está limitado estrictamente a fines académicos, quedando excluidos propósitos comerciales.

Todas las imágenes presentadas en este documento han sido debidamente referenciadas como “Imagen creada mediante BioRender”, en conformidad con los términos de la licencia académica de la plataforma.