



Fundació
La Marató de TV3
XVI SIMPOSIUM
Enfermedades minoritarias

Estudio de las funciones del complejo de cohesinas en el síndrome de Cornelia de Lange

Dra. Ethelvina Queralt Badia

Institut d'Investigació Biomèdica de Bellvitge

Dra. Susana Rodríguez Navarro

Centro de Investigación Príncipe Felipe. València





1. Resumen del proyecto

El síndrome de Cornelia de Lange (CdLS) es una enfermedad rara grave caracterizada por un retraso mental y de crecimiento, anomalías craneofaciales i microcefalia. En un 75% de personas diagnosticadas clínicamente de CdLS se han identificado mutaciones en cinco genes: *NIPBL*, *SMC1A*, *SMC3*, *RAD21* i *HDAC8*, que corresponden a subunidades del complejo de cohesinas o proteínas reguladoras de las cohesinas. El complejo de cohesinas es esencial para el establecimiento y mantenimiento de la cohesión entre las cromátidas hermanas. Células derivadas de enfermos de CdLS no presentan defectos obvios en la cohesión de las cromátidas hermanas, sugiriendo que las cohesinas también están implicadas en otras funciones desconocidas. Últimamente se han descrito numerosas observaciones que sugieren que las cohesinas también desempeñan un importante papel regulando la expresión génica, aunque las bases moleculares de esta nueva función todavía se desconocen. En el presente proyecto, nuestro objetivo principal es estudiar cómo las mutaciones de las cohesinas identificadas en pacientes de CdLS provocan los fenotipos de la enfermedad.

Metodología. Nuestra hipótesis es que los fenotipos característicos de los pacientes de CdLS son causa de alteraciones en la nueva función del complejo de cohesinas en la regulación de la expresión génica. Normalmente, los estudios en cohesinas se realizan en células ciclando o incluso en células sincronizadas, ya que su función de cohesión es necesaria en cada ciclo celular. Sin embargo, la función de las cohesinas en la regulación génica es más evidente en células en interfase. Por este motivo, en este proyecto utilizaremos células en interfase derivadas de pacientes de CdLS.

Resultados esperados. Esperamos profundizar en la función del complejo de cohesinas en la regulación de la expresión génica y cómo las mutaciones en dicho complejo provocan la aparición de enfermedades humanas como el CdLS.

2. Resultados

1. Aproximación genómica para el estudio de las funciones del complejo de cohesinas en expresión génica

1.1. Determinar las mutaciones genéticas de los pacientes de CdLS

Con el fin de identificar las mutaciones genéticas de los pacientes de CdLS, realizamos experimentos de secuenciación masiva (*next-generation sequencing*) tras realizar un enriquecimiento de los exones mediante captura con sondas de oligonucleótidos. Por este método hemos identificado la mutación presente en diferentes pacientes, incluyendo pacientes que presentaban la mutación en mosaicismo. El análisis bioinformático de las muestras se realizó en colaboración con el bioinformático Benjamín Rodríguez. Algunos de los resultados obtenidos pueden consultarse en las publicaciones realizadas (ver al final del documento).

1.2. Determinar el patrón de unión a lo largo de los cromosomas de las cohesinas en células derivadas de pacientes de CdLS

Para determinar cómo las cohesinas regulan la expresión génica, una de las cuestiones importantes es conocer dónde se unen las cohesinas a lo largo de los cromosomas y qué podemos aprender de dicho patrón de unión. Durante este trabajo analizamos el patrón de unión de las cohesinas en células derivadas de pacientes de CdLS mediante técnicas de inmunoprecipitación de la cromatina seguidas de secuenciación masiva (ChIP-seq). Para los experimentos se partió de 40 millones de células ciclando o en condiciones de quiescencia. En células ciclando, la distribución genómica del complejo de cohesinas es muy similar a las células controles. Sin embargo, en células quiescentes se observó un aumento en el número de picos encontrados. Posteriormente, los resultados obtenidos se validaron mediante técnicas cuantitativas de ChIP-qPCR.

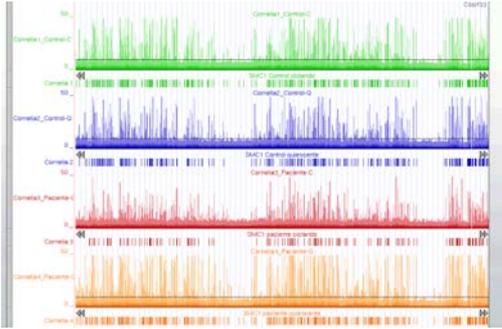


Figura 1. Análisis de ChIP-seq de SMC1A en los pacientes de CdLS. El número de picos de cohesinas aumenta en células quiescentes comparado con las células ciclando.

1.3. Analizar las alteraciones transcripcionales en muestras derivadas de pacientes de CdLS mediante *arrays* de expresión

En primer lugar realizamos experimentos con mutantes en NIPBL/Scc2 en células de la levadura *S. cerevisiae*. Diversos genes relacionados con el ciclo celular presentan una expresión alterada en mutantes de cohesinas (figura 2). Resultó interesante encontrar el gen *MED9* con expresión reducida en los mutantes en *scc2*. Med9 es una subunidad del complejo del mediador, un coactivador involucrado en la regulación de la transcripción de la mayoría de genes dependientes de RNA polimerasa II. El mediador actúa como un puente de unión entre proteínas reguladoras específicas de genes con la maquinaria de transcripción basal de RNA polimerasa II. Por tanto, este resultado sugiere que las cohesinas pueden estar regulando la expresión génica a través del complejo del mediador. Por otra parte, también se detectó una interacción física entre subunidades del complejo de cohesinas con diferentes subunidades del complejo del mediador mediante experimentos de coimmunoprecipitación, corroborando así nuestra hipótesis de que las cohesinas pueden estar regulando la expresión génica a través de la regulación del complejo del mediador.

AFFYMETRIX_3PRIME_IVT_ID	GENE NAME	Related Genes	Species
1778433_at	ARF3-interacting protein 1	RG	Saccharomyces cerevisiae
1774087_at	Accumulates dyads protein 3	RG	Saccharomyces cerevisiae
1779372_at	Cell division control protein KAR1	RG	Saccharomyces cerevisiae
1772570_at	DASH complex subunit DAD2	RG	Saccharomyces cerevisiae
1772709_at	DASH complex subunit HSK3	RG	Saccharomyces cerevisiae
1777453_at	DNA polymerase zeta processivity subunit	RG	Saccharomyces cerevisiae
1776729_at	Factor arrest protein 3	RG	Saccharomyces cerevisiae
1772962_at	Factor arrest protein 7	RG	Saccharomyces cerevisiae
1769607_at	Karyogamy protein KAR9	RG	Saccharomyces cerevisiae
1777576_at	Kinetochores-associated protein MTW1	RG	Saccharomyces cerevisiae
1773148_at	Mediator of RNA polymerase II transcription subunit 9	RG	Saccharomyces cerevisiae
1773241_at	Meiosis-specific protein ISC10	RG	Saccharomyces cerevisiae
1773240_at	Meiosis-specific protein SPO11	RG	Saccharomyces cerevisiae
1777726_at	Meiotic recombination 1 protein	RG	Saccharomyces cerevisiae
1773376_at	Meiotic recombination protein REC114	RG	Saccharomyces cerevisiae
1770249_at	Protein LDB18	RG	Saccharomyces cerevisiae
1779764_at	Protein ZIP2	RG	Saccharomyces cerevisiae
1779155_at	SET domain-containing protein 3	RG	Saccharomyces cerevisiae
1777421_at	Sporulation protein RMD6	RG	Saccharomyces cerevisiae
1780117_at	Sporulation-specific protein 1	RG	Saccharomyces cerevisiae
1778536_at	YAL047C	RG	Saccharomyces cerevisiae

Figura 2. Genes implicados en el control del ciclo celular y del complejo del mediador presentan su expresión disminuida en mutantes de *scc2*.

Con el fin de profundizar en la investigación de las funciones del complejo de cohesinas en la expresión génica, estudiamos los cambios de expresión génica mediante experimentos con *arrays* de expresión, utilizando fibroblastos dermales derivados de pacientes de CdLS. Como controles se usaron fibroblastos dermales de padres no afectados. En primer lugar, observamos que los perfiles de expresión génica se agrupan en los pacientes con mutación en *NIPBL*, y al mismo tiempo difieren de los perfiles de expresión de los pacientes con mutación en *SMC1A*. Por tanto, los patrones de expresión génica son distintos dependiendo del gen que esté Mutado, sugiriendo que el mecanismo molecular afectado sería diferente según el gen mutado, *NIPBL* o *SMC1A*. A continuación estudiamos los perfiles de expresión de los pacientes con mutación en *NIPBL*. Resultó interesante detectar genes implicados en el desarrollo como la categoría funcional más alterada en dichos pacientes de CdLS. Por tanto, dichas alteraciones transcripcionales podrían explicar los defectos de desarrollo que presentan los pacientes de CdLS.

Posteriormente validamos nuestros resultados mediante técnicas cuantitativas de PCR a tiempo real qPCR. Finalmente, comparamos bioinformáticamente nuestros resultados con datos publicados y observamos que en el perfil de expresión de pacientes de CdLS se encuentran genes comunes a los perfiles de expresión de muestras con sobreexpresión de *cmyc* o los perfiles de expresión de CTCF.

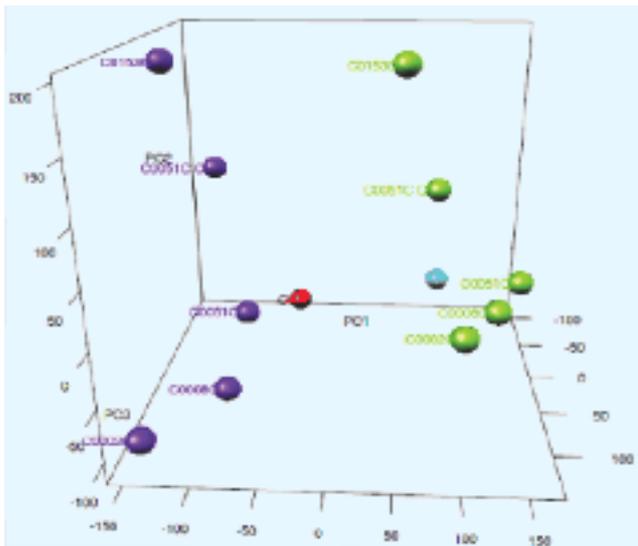


Figura 3. Las muestras de pacientes con mutación en *NIPBL* se agrupan y diferencian de las muestras con mutación en *SMC1A*. Perfiles de expresión de los pacientes de CdLS con mutación en *NIPBL* representados en un diagrama PCDA 3D. Se han realizado triplicados de los experimentos. C0051, C0051C y C0008 corresponden pacientes con mutación en el gen *NIPBL*.

pValue	GO Term
5,79E-16	multicellular organismal process
8,69E-12	system development
1,04E-11	developmental process
2,12E-11	multicellular organismal development
6,75E-11	anatomical structure development
2,36E-10	multicellular organismal process
1,86E-09	system development
4,56E-09	anatomical structure development
1,39E-08	regulation of multicellular organismal development
2,02E-08	skeletal system development

Cycling CdLS fibroblasts vs Cycling Wild-type fibroblasts

pValue	GO Term
1,77E-13	multicellular organismal process
2,27E-11	cell adhesion
2,51E-11	biological adhesion
8,00E-11	system development
4,60E-10	multicellular organismal development
3,39E-13	system development
2,06E-12	multicellular organismal development
2,56E-12	anatomical structure development
5,94E-12	organ morphogenesis
2,14E-11	developmental process

Quiescent CdLS fibroblasts vs Quiescent Wild-type fibroblasts

Figura 4. Perfiles de expresión de los pacientes de CdLS. El análisis de GO revela los procesos biológicos con expresión elevada (rojo) o reducida (azul) en células derivadas de CdLS comparadas con células controles.

2. Investigar la función del complejo de cohesinas en expresión génica mediante aproximaciones proteómicas

2.1. Estudio de la integridad del complejo de cohesinas en células derivadas de pacientes de CdLS

Las cohesinas son un complejo proteico formado por un heterodímero de las proteínas SMC1A y SMC3, al que se unen por lo menos dos proteínas no-SMC RAD21 y SA1/2. Otras proteínas se unen al núcleo del complejo de cohesinas y regulan su actividad. La proteína NIPBL es la encargada de reclutar al complejo de cohesinas al DNA; ESCO2 y sororina están implicadas en el establecimiento de la cohesión, Pds5A y Pds5B se requieren para el mantenimiento de la cohesión y Wapl se necesita para eliminar el complejo de cohesinas de su unión al DNA. La formación y la integridad del complejo de cohesinas siempre se ha estudiado en células ciclando. Por este motivo investigamos si todos los componentes del complejo de cohesinas y sus proteínas reguladores están presentes en células en quiescencia. Todas las subunidades del complejo de cohesinas y sus proteínas reguladoras (a excepción de la sororina) se expresan en células ciclando; aunque en ocasiones los niveles de proteína se encuentran reducidos a la mitad.

Posteriormente, realizamos ensayos de fraccionamiento subcelular con el fin de estudiar la fracción del complejo de cohesinas que se halla unida al DNA. La distribución de las subunidades de las cohesinas en las 3 fracciones (citosol, nuclear soluble y unido a cromatina) es similar en las células derivadas de pacientes de CdLS y en las células controles. Finalmente, investigamos la integridad y formación del complejo de cohesinas en las diferentes fracciones mediante experimentos de coimmunoprecipitación. La interacción entre las diferentes subunidades del complejo de cohesinas se mantiene en todas las fracciones estudiadas en las células derivadas de pacientes de CdLS, indicando que la integridad del complejo de cohesinas se mantiene tanto en el citosol como en los complejos unidos a la cromatina

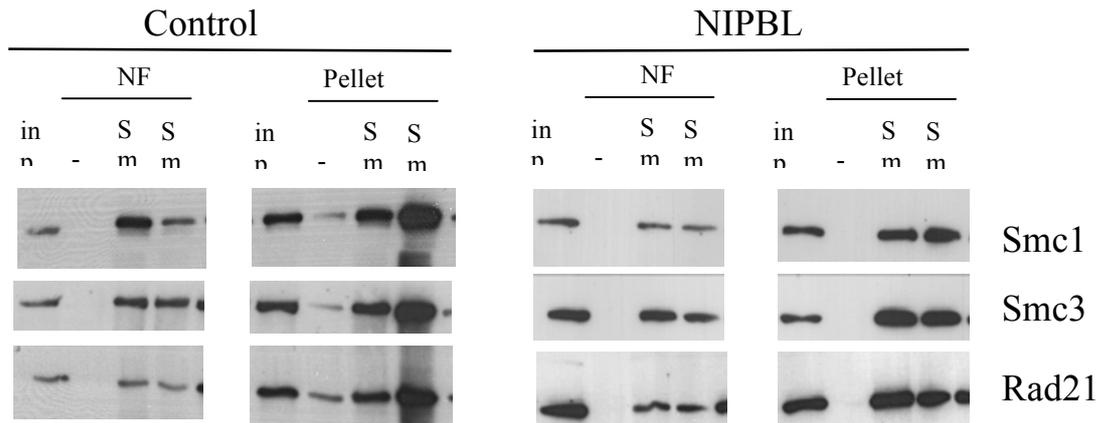


Figure 5. La integridad del complejo de cohesinas se mantiene en células derivadas de pacientes de CdLS. Los experimentos de coimmunoprecipitación se realizaron de las diferentes fracciones obtenidas de los fraccionamientos celulares.

2.2. Identificación de nuevas proteínas interactoras del complejo de cohesinas en células en interfase

Las interacciones proteína-proteína constituyen un método muy útil para estudiar la interacción entre diferentes factores. Por ello, realizamos un rastreo proteómico global para identificar nuevas proteínas que interaccionan con NIPBL mediante técnicas de espectrometría de masas. Observamos que NIPBL interacciona con histonas, DNA helicasas y con el complejo remodelador de la cromatina SWI/SNF.

Accession	Name
CHD6_HUMAN	Chromodomain-helicase-DNA-binding protein 6 OS=Homo sapiens GN=CHD6 PE=1 SV=4
H2A1H_HUMAN	Histone H2A type 1-H OS=Homo sapiens GN=HIST1H2AH PE=1 SV=3
H31_HUMAN	Histone H3.1 OS=Homo sapiens GN=HIST1H3A PE=1 SV=2
H32_HUMAN	Histone H3.2 OS=Homo sapiens GN=HIST2H3A PE=1 SV=3
RBP56_HUMAN	TATA-binding protein-associated factor 2N OS=Homo sapiens GN=TAF15 PE=1 SV=1
SMCA5_HUMAN	SWI/SNF-related matrix-associated actin-dependent regulator of chromatin GN=SMARCA5 PE=1 SV=1
RAD21_HUMAN	Double-strand-break repair protein rad21 homolog OS=Homo sapiens GN=RAD21 PE=1 SV=2

3. Relevancia y posibles implicaciones clínicas

Durante el desarrollo del presente proyecto hemos realizado experimentos de secuenciación masiva tras la captura de exoma (*exome-sequencing*) de muestras de pacientes de CdLS. Con nuestros estudios comparamos las técnicas de secuenciación masiva con la secuenciación tradicional Sanger y pudimos observar que las nuevas técnicas de secuenciación masiva son más eficientes para detectar las mutaciones. Además, dichas técnicas son capaces de detectar incluso mutaciones que se encuentran en mosaicismo, ya que se basan en la obtención de muchas secuencias cortas de cada región del gen. De hecho, en nuestro caso hemos detectado mutaciones que se encontraban en mosaicismo y que aparecían tan solo en un 25% de las secuencias. Por tanto, para el diagnóstico molecular de la enfermedad sugerimos la utilización de un panel personalizado de captura de exones de genes relacionados con las cohesinas. Esto permitiría aplicar las técnicas de secuenciación masiva pero a menor precio.

En nuestros estudios de expresión génica hemos observado que la mayor parte de los genes desregulados corresponden a genes implicados en el desarrollo. Dichos genes deberían silenciarse tras el desarrollo embrionario, pero sin embargo se expresan de manera anormal en los pacientes de CdLS. Por otra parte, también hemos detectado alterados genes implicados en la estimulación y asimilación hormonal. Así podrían explicarse los problemas de crecimiento asociados a los pacientes de CdLS. Pero, puesto que también presentan alteraciones de asimilación hormonal el tratamiento oral con hormonas, no sería suficiente para los pacientes de CdLS (al menos no a dosis normales).

4. Publicaciones

Gil-Rodríguez MC; Hernández-Marcos M; Teresa-Rodrigo ME; Vicente-Gabas A; Bernal ML; Casale CH; Bueno-Lozano G; Bueno-Martínez I; **Queralt E**; Villa O; Hernando-Davalillo C; Armengol L; Gómez-Puertas P; Puisac B; Selicorni A; Ramos FJ; Pié J.

De novo heterozygous mutations in SMC3 cause a range of Cornelia de Lange overlapping phenotypes.

Human Mutation. Feb. 5 Doi 10.1002/humu.22761/**2015**

Baquero-Montoya C; Gil-Rodríguez MC; Hernández-Marcos M; Teresa-Rodrigo ME; Vicente-Gabas A; Bernal ML; Casale CH; Bueno-Lozano G; Bueno-Martínez I; **Queralt E**; Villa O; Hernando-Davalillo C; Armengol L; Gómez-Puertas P; Puisac B; Selicorni A; Ramos FJ; Pié J.

Severe ipsilateral musculoskeletal involvement in a Cornelia de Lange patient with a novel NIPBL mutation.

European journal of medical genetics. 7212 - 14, pp. 118 - 119. 27/05/**2014**.

Baquero Montoya; M Gil Rodriguez; M Teresa Rodrigo; M Hernandez Marcos; G Bueno Lozano; I Bueno Martinez; Silvia Remeseiro; Rita Fernandez Hernandez; M

Bassecourt Serra; M Rodriguez de Alba; Ethel Queralt; Ana Losada; Beatriz Puisac;
Feliciano Ramos; Juan Pie.

Could a patient with SMC1A duplication be classified as a human cohesinopathy.

Clinical Genetics. 17/05/2013.