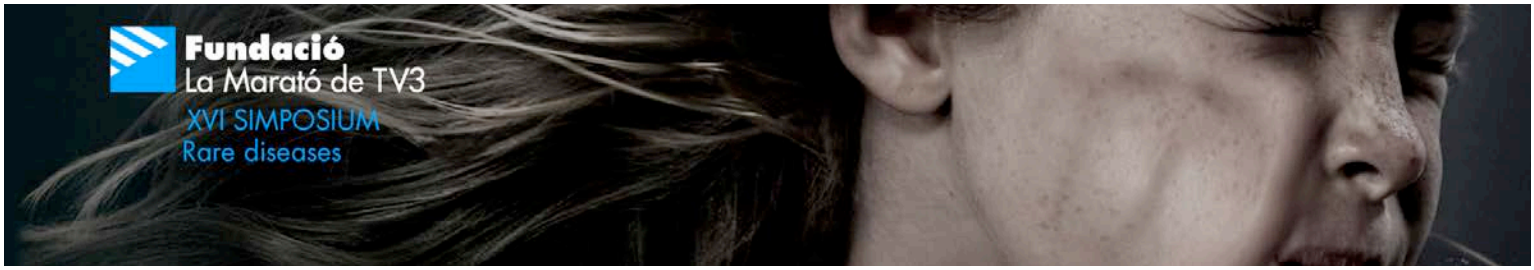
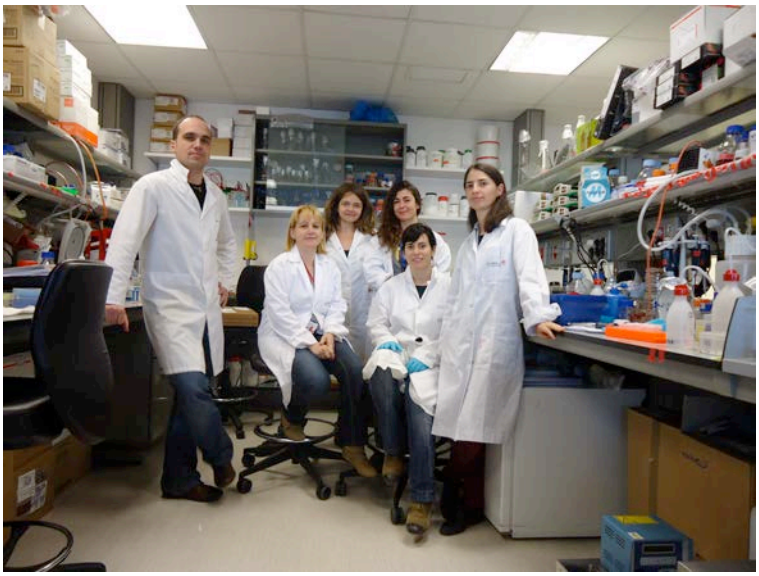


Study of cohesin functions in Cornelia de Lange Syndrome

Dr Ethelvina Queralt Badia
IDIBELL Fundació Institut d'Investigació Biomèdica de Bellvitge

Dr Susana Rodríguez Navarro
Fundación Valenciana Centro de Investigación Príncipe Felipe



1. Project Summary

Cornelia de Lange syndrome (CdLS) is a severe disease characterized by growth and mental retardation, craniofacial anomalies and microcephaly. Mutations in five genes that correspond to 5 cohesin-related proteins have been identified in 75% of individuals with clinically diagnosed CdLS: *NIPBL*, *SMC1A*, *SMC3, RAD21* and *HDAC8*. Cohesin complex acts as the chromosomal "glue" and is essential for sister chromatid cohesion and their subsequent segregation. Cells derived from CdLS patients show only mild and in some cases no obvious defects in sister chromatid cohesion, suggesting that cohesins are also involved in other unknown functions. Recently, numerous observations that suggest that cohesin has an important role in regulating gene expression have been described; however, the mechanistic basis of this function is still very poorly understood. In this project our main objective is to study how the cohesin mutations found in CdLS patients cause the disease phenotypes.

Methodology: Our hypothesis is that the CdLS phenotypes are due to alterations in the putative cohesin complex role in gene expression. Commonly, cohesin studies are performed in cycling cells or even synchronous cultures, since sister chromatid cohesion is required during each cell cycle. However, in this project to study the cohesin functions in gene expression we will use non-dividing cells derived from CdLS patients.

Expected results: In this project we aim to get a new insight into the cohesin roles in non-diving cells. We expect to further confirm the involvement of cohesin complex in gene expression regulation and how its mutations induce human malignancies such as CdLS.

2. Results

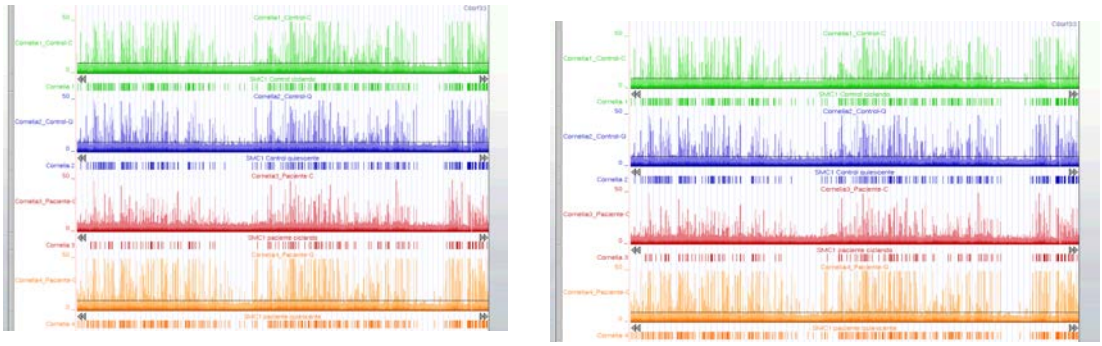
1. Genomic approach to studying the roles of cohesin in gene expression

1.1 Determining the genetic mutation of the CdLS patients

In order to identify the genetic mutation of CdLS patients, we did a Next-generation sequencing approach after an exome enrichment of the DNA samples. We identified mutation in several patients, including some patients with mosaic mutation. This was performed in collaboration with the bioinformatician Benjamin Rodriguez. Some example of the results obtained can be found in the publication section of this report (see below).

1.2 Determining the cohesin binding pattern along chromosomes in cells derived from CdLS patients.

To determine how cohesin controls gene expression, important questions include where along chromosomes cohesin binds and what we can learn from its binding pattern. In this study we analyzed the cohesin binding pattern along chromosomes in cells derived from CdLS patients by ChIP-seq experiments. We performed ChIP-seq experiments starting with 40 million cells in cycling or quiescent conditions. We saw that in the patients the genomic distribution of the cohesin complex in cycling cells is pretty similar to control cells. However, in quiescent cells we found an increase in the number of cohesin peaks. We validated the results obtained from the ChIP-seq by ChIP-qPCR.



Control CdLS Patient

Figure 1. **SMC1A ChIP-seq analysis of CdLS patients**. Non-diving cells from CdLS patients have more peaks compared to cycling cells and control cells.

1.3 Analyzing transcription alteration in samples derived from CdLS patients using human gene expression arrays.

We first performed experiments with model organisms such as budding yeast. We did gene expression arrays from cohesin mutants in *S. cerevisiae*. From these experiments we have seen several cell cycle genes whose expression is affected in cohesin mutants cells (Figure 2). More interesting, in *scc2* mutant cells the *MED9* gene is downregulated. Med9 is a subunit of the mediator complex; a coactivator involved in the regulated transcription of nearly all RNA polymerase II-dependent genes. Mediator functions as a bridge to convey information from gene-specific regulatory proteins to the basal RNA polymerase II transcription machinery. Therefore, this result is in accordance with our hypothesis that cohesin can regulate gene expression, and this regulation will be performed by the control of the mediator complex. Moreover, we also detected a physical interaction among cohesin subunits and different subunits of the mediator complex by co-immunoprecipitation assays.

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	Kinetochore-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

Figure 2. **Cell cycle genes and genes of the mediator complex are downregulated in a *scc2* mutant.**

In order to further investigate the role of cohesin in gene expression, we studied the influence in genome-wide gene expression in human dermal fibroblast derived from CdLS patients and some unaffected parents as a control. Firstly, the gene expression profiles of NIPBL mutant patients cluster together pretty well, and they present different expression profiles from patients with SMC1A. Therefore, this result suggests that the altered gene expression is different in patients with different mutations, indicating that the molecular mechanism behind the cohesin mutation in these patients could have a different nature. Interestingly, we detected genes related to development as the most deregulated functional category in CdLS patients. That is quite promising since CdLS is mostly a developmental disorder. We validated the results by real-time qPCR experiments. In addition, we did an extensive bioinformatic analysis of these samples with experts in gene expression in order to compare the gene expression profile with published data related to the cohesin complex such as CTCF expression profile, myc regulated genes and mediator regulator genes, among others. In CdLS patients the gene expression profile enrichment correlates with the profile obtained on upregulation of c-myc and CTCF regulated genes.

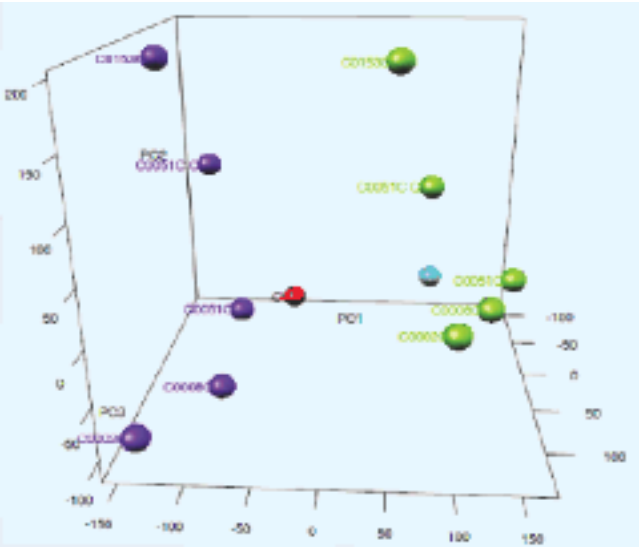


Figure 3. **NIPBL-mutated CdLS patients' gene expression profiles cluster together**. Expression profiles of CdLS patients with NIPBL mutation in two independent experiments are plotted in a PCDA

3D diagram. Technical triplicates have been performed. C0051, C0051C and C0008 correspond to patients with known mutation inside the NIPBL gene.

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 Wilde-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

Figure 4. **Gene expression profiles of CdLS patients**. GO analysis reveals biological processes upregulated (red) or downregulated (blue) in CdLS derived cells compared with control cells.

2. Investigating the role of cohesin in gene expression by proteomic approaches

2.1 Studying the integrity of the cohesin complex in cells derived from CdLS patients.

Cohesin is an evolutionally conserved multisubunit protein complex consisting of an SMC1A and SMC3 heterodimer, and at least two non-SMC proteins RAD21 and SA1/SA2. Other subunits bind to the core cohesin complex and regulate its activity. NIPBL is the loader of cohesin onto the DNA, ESCO2 is involved in cohesion establishment, Pds5A and Pds5B are required to maintain sister chromatid cohesion and Wapl is needed to remove cohesin from DNA. The integrity of the cohesin complex and its interaction with its regulatory subunits have always been studied in cycling cells. For this reason we investigated whether the other subunits are also important for cells in interphase and, therefore, for the non-canonical role of cohesin in gene expression. To do that, we first analyzed whether these other cohesin subunits are also expressed in non-diving cells. In fact, we observed that all the cohesin-related proteins are expressed in interphase cells, although in some cases the protein levels are reduced to half compared with cycling cells.

Next, we performed subcellular fractionation in order to study the pool of cohesin subunits that are bound to chromatin. We detected that the distribution of cohesin subunits among cytosol fraction, nuclear soluble fraction and chromatin bound fraction in fibroblast derived from CdLS patients, is similar to the control cells. Next, we investigated the integrity of the cohesin complex in each fraction by co-immunoprecipitation experiments. The cohesin complex integrity was preserved in the three fractions studied. Therefore, we can conclude that the integrity of the cohesin complex is maintained in fibroblast cells derived from CdLS patients.

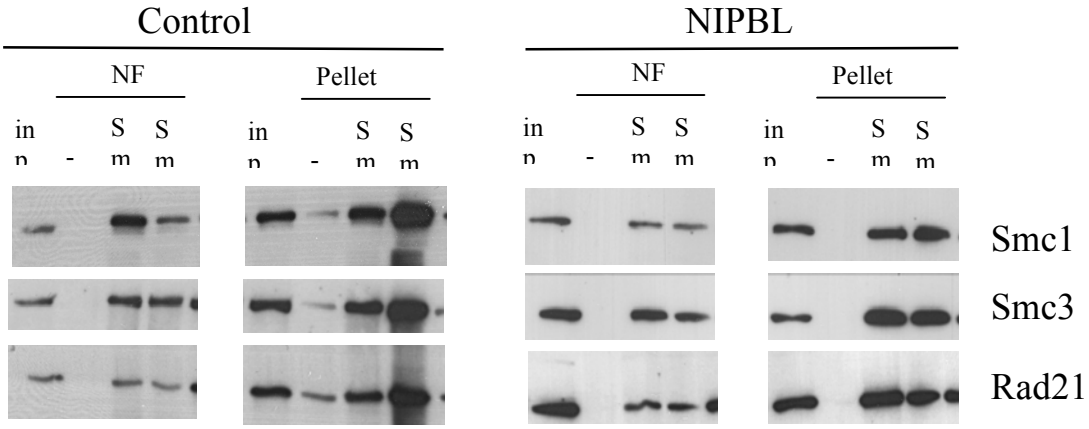


Figure 5. **Cohesin complex integrity is preserved in CdLS-derived fibroblasts.** Cell fractionations followed by immunoprecipitation experiments were performed. Copurified proteins were present at similar levels in controls and in CdLS derived fibroblasts.

2.2 Identifying new cohesin-interacting proteins in protein extracts prepared from cells in interphase

Protein-protein interaction is an excellent way to functionally link different factors. We did a global screening to identify new cohesion/NIPBL interacting partners in eukaryotic cells using the highly sensitive mass spectroscopy methods currently available. We found that NIPBL interacts with histones, DNA helicase and SWI/SNF chromatin remodeling complex.

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. Relevance and possible clinical implications

During the time spent on this project we gained experience in the exome-sequencing of CdLS patients. We tested the next-generation sequence techniques compared with traditional Sanger sequencing. We observed that the next-generation approaches could detect mosaic mutation even for mutations found in only 25% of the sequences. Usually, Sanger sequencing does not detect those mutations. In fact, we will recommend clinicians to establish a custom exome-sequencing panel including NIPBL and cohesion related genes to perform the molecular diagnosis of the disease. A custom panel will allow this state-of-the-art technique to be applied at a lower cost than whole exome-sequencing.

In our gene expression assays we have determined that mostly developmental genes are deregulated in CdLS samples. Some of these genes should be silenced after embryonic development, however they are abnormally expressed in the CdLS

patients. In addition, some genes related to hormone stimulus and assimilation are downregulated, suggesting that CdLS patients have hormonal problems. In fact, the lower height of these patients could be related to the hormonal problems. However, the administration of common oral hormones will not be sufficient by itself to solve the hormonal deregulation since higher than normal dosages will be needed due to hormonal assimilation problems.

4. Publications

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.