UNIVERSITA' DEGLI STUDI DI MILANO

FACOLTÀ DI MEDICINA E CHIRURGIA

CORSO di DOTTORATO DI RICERCA in
PATOLOGIA E NEUROPATOLOGIA SPERIMENTALI

XXVIII CICLO

Settore Scientifico – Disciplinare MED/03

CORNELIA DE LANGE SYNDROME AND RELATED DISORDERS: NEW INSIGHTS INTO GLOBAL TRANSCRIPTIONAL DISTURBANCES DUE TO MUTATIONS IN CHROMATIN-ASSOCIATED FACTORS

Tesi di Dottorato di Ricerca di:
ILARIA PARENTI
Matricola R10281

Docente di Riferimento: Prof.ssa Palma FINELLI
Tutor: Dott.ssa Cristina GERVASINI
Coordinatore: Prof. Massimo LOCATI

Anno Accademico 2014-2015

To Ferruccio,

always

ABSTRACT

Cornelia de Lange syndrome (CdLS) is a rare disorder characterized by an extensive clinical heterogeneity. The main features of the syndrome are characteristic facial dysmorphisms and a variable level of intellectual disability, growth retardation and developmental delay. Though, the number and severity of the clinical signs vary among patients. An extensive genetic heterogeneity partially accounts for the reported clinical variability. Mutations in different cohesin-associated proteins are in fact responsible for the onset of the syndrome. The known CdLS-genes include *NIPBL*, *SMC1A*, *SMC3*, *RAD21* and *HDAC8*. Alterations in the cohesin loader *NIPBL* are found in more than half of CdLS cases and are associated with a classical phenotype and with a high frequency of limb malformations. In addition, mosaicism has been proven to play an important role in association with *NIPBL*. Mutations in the structural elements *SMC1A*, *SMC3* and *RAD21* and in the regulator *HDAC8* account for about 10-15% of CdLS cases. The phenotype of those patients who harbor mutations in these genes is usually milder or atypical. The five genes all together, also taking into account the role of mosaicism, can explain about 70% of CdLS cases.

In an internationally assembled cohort of patients we were able to identify 109 mutations in *NIPBL*, 8 mutations in *SMC1A*, 15 mutations in *SMC3* and 11 mutations in *HDAC8*, thus increasing the total number of mutations so far described for CdLS. In addition, by the use of next generation sequencing techniques we were able to identify mutations in five genes different from cohesin in six unrelated patients with a clinical diagnosis of CdLS. The five genes include those encoding for different subunits of the chromatin remodeling complex named SWI/SNF and for the transcriptional repressor ANKRD11. Mutations in these genes have been so far associated to Coffin-Siris syndrome and KBG syndrome, respectively. Protein-protein interaction experiments also showed a direct interaction of the SWI/SNF subunit SMARCB1 with the cohesin-related proteins NIPBL and SMC3. These direct link between cohesin and SWI/SNF subunits indicate that mutations affecting the two protein complexes might determine the deregulation of overlapping sets of genes.

Our newly identified variants contribute to a better understanding of the correlation between genotype and phenotype in the presence of mutations in the known-CdLS genes. Notwithstanding, different pehotypes have been observed in patients carrying the same DNA alteration, hence suggesting that environmental factors may play an important role in the delineation of the observed clinical features. Additionally, the identification of mutations in chromatin-associated factors responsible for syndromes different from CdLS indicate the existence of a broad pleiotropy that should be taken into account while assessing the clinical and molecular diagnosis.

Furthermore, we investigated the molecular mechanisms underlying the syndrome in the presence of missense substitutions or small in-frame deletions in *SMC1A*, a X-linked gene that localizes in a region of the X-chromosome that partially escapes X-inactivation in humans. Our expression analysis revealed that the transcript is expressed at higher levels in females as compared to males, and that there are no differences in the expression of the SMC1A protein between healthy and affected females. In addition, pyrosequencing analysis showed that CdLS female patients harboring mutations in *SMC1A* tend to express the wild type allele at higher levels as compared to the mutant allele. All together, these data suggests that the pathogenesis of the syndrome in the presence of mutations affecting *SMC1A* which do not disrupt the reading frame might be linked to a dominant negative effect exerted by the mutant protein on the wild type.

INTRODUCTION

CORNELIA DE LANGE SYNDROME

Cornelia de Lange syndrome (CdLS, OMIM #122470, 300590, 610759, 300882, and 614701) is a multisystem developmental disorder characterized by a dominant mode of inheritance. In the majority of cases it occurs *de novo*; though, also some familial cases have been described.

HISTORY

In 1933 the Dutch pediatrician Cornelia de Lange recognized and described for the first time in the *"Archives de Medicine des Enfants"* some comparable clinical features observed in two unrelated patients (de Lange, 1933).

The first patient (Fig.1a) was a 17-months old girl hospitalized at the Emma Children's Hospital of Amsterdam in 1932 because of bronchopneumonia. The child was born at 40 weeks of gestation with a birth weight of 1250 g; she presented with growth retardation, microcephaly, remarkably unusual facial features and feeding difficulties that persisted during the first year of life.

The second patient (Fig.1b), hospitalized few weeks after, shared the growth impairment and the digestive disorder with the first girl. Nevertheless, the most striking resembling features were related to the *facies*.

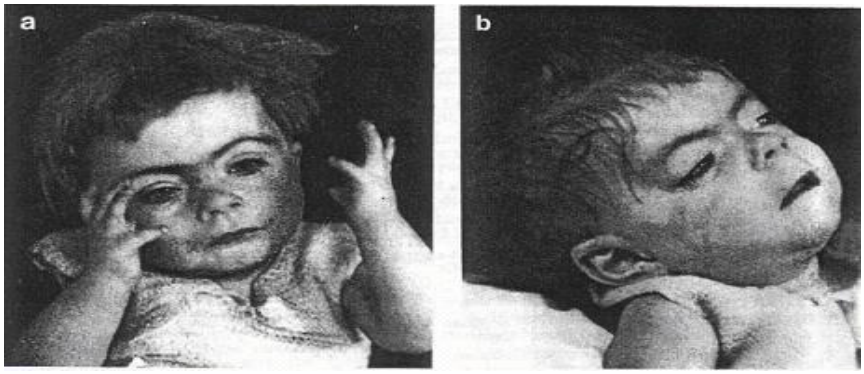


Figure 1. First two patients with the "Typus Degenerativus Amstelodamensis"described from doctor de Lange in 1933

For both cases, doctor de Lange made a highly detailed description of the clinical features and assigned to this phenotype the definition of "*Typus Degenerativus Amstelodamensis*" (de Lange, 1933).

During the following years a discreet number of patients with an analogous clinical presentation was described in the literature. Then, in 1964, Opitz and colleagues associated for

the first time the phenotype described by doctor de Lange in 1933 with the one described by the German doctor Winfried Robert Clemens Brachmann in a paper of 1916 titled *"Ein Fall von symmetrischer Monodaktylie durch Ulnadefekt, mit symmetrischer Flughaurbildung in den Ellenbogen sowie anderen Abnormalitaten"*, that likely represents the first description of the so called "*Typus Degenerativus Amstelodamensis*" syndrome. Because of this, Opitz and colleagues named the new clinical entity Brachmann-de Lange syndrome (Brachmann, 1916; Opitz et al., 1964). Today, the syndrome is mainly known and named as Cornelia de Lange syndrome (CdLS). Since 1964, many cases with similar features have been described, and a remarkable clinical heterogeneity of the syndrome has emerged.

CLINICAL FEATURES

Cornelia de Lange syndrome (CdLS) is characterized by distinct clinical features that involve different organs with variable extent; nevertheless, the muscoloskeletal, gastrointestinal, and nervous systems represent the most frequently affected systems (Kline et al., 2007).

The incidence of CdLS in the population has been discussed for several years because of the difficulties in the formulation of the diagnosis, that is mainly based on clinical observations. The first assessment of the frequency was performed in 1967 by Pearce and Pitt, who estimated the prevalence to be around 1:100.000 (Pearce and Pitt, 1967). Currently, public scientific databases (OMIM, *GeneReviews* and *GeneticsHomeReference*) and more recent papers attest the incidence between 1:10.000 and 1:40.000 live births (Kline et al., 2007; Barisic et al., 2008).

To date, minimal diagnostic criteria have been established based on the consensus of the Scientific Advisory Committee of the World CdLS Federation (SAC) and the Clinical Advisory Board of the CdLS Foundation USA (CAB). A CdLS diagnosis requires the presence of specific facial features in combination with the involvement of at least two other systems responsible for growth, development or behavior (Kline et al., 2007).

A scoring system (Table 1) is also available to determine the severity of the phenotype. This system assigns a score to each of the following categories: birth weight, age of sitting independently, age of walking independently, age of the first word, presence or absence of upper limbs malformations and presence or absence of hearing loss. Scores higher than 22 are associated with a severe phenotype, while scores between 15 and 22 indicate a moderate

presentation of CdLS; lastly, patients with a score lower than 15 are affected by a mild form of the syndrome (Table 1; Kline et al., 2007).

Parameter	1 point	3 points	5 points
Birth weight	Above 2,500 g	2,000–2,500 g	Below 2,000 g
Sitting alone[a]	Before 9 months	9–20 months	Over 20 months
Walking alone[a]	Before 18 months	18–42 months	Over 42 months
Saying first word[a]	Before 24 months	24–48 months	Over 48 months
Upper limb malformation	No defect	Partial defect (more than two digits)	Severe defect (less than two digits)
Number of other major malformations	0–1	2–3	More than 3
Hearing loss	Absent	Mild	Moderate–Severe

Scoring: >22 points, severely involved; 15–22 points, moderately involved; <15 points, mildly involved.
[a]Ages based on Kline et al. [1993b] using 25th–75th centiles of completing milestones.

Table 1. Scoring system that allows the classification of CdLS into three different categories, namely severe, moderate and mild (from Kline et al., 2007)

Prenatal diagnosis and auxological data

Despite the incidence of premature births is higher than in the normal population (30%), relevant obstetric complications rarely occur during CdLS pregnancies. As the majority of CdLS cases are *de novo*, the recurrence risk for asymptomatic parents has been estimated to be approximately 1.5% (Jackson et al., 1993). Nevertheless, in the presence of more than one affected child within one family, the existence of germline mosaicism should be considered (Slavin et al., 2012; Mariani et al., 2013).

High-resolution ultrasound examination appears to be the best technique for a prenatal diagnosis of the syndrome. Intra-uterine growth retardation (IUGR) and limb malformations are the most discernible features, but reported ultrasound findings also include microcephaly and diaphragmatic hernia (Sekimoto et al., 2000; Huang and Porto, 2002). In some cases it was also possible to notice micrognathia with prominent maxilla and a depressed nasal bridge with anteverted nares, features that define the unique facial profile of CdLS (Clark et al., 2012).

Maternal serum screening might also be performed: the levels of the Pregnancy-Associated Plasma Protein-A (PAPP-A) have shown to be reduced during the first and second trimester in pregnancies affected by CdLS; nevertheless, this biomarker is not specific for CdLS (Westergaard et al., 1983; Aitken et al., 1999; Arbuzova et al., 2003). Lastly, an increased fetal nuchal translucency during the first trimester has also been observed (Sekimoto et al., 2000; Huang and Porto, 2002).

Auxological data are highly peculiar for CdLS; for this reason, specific charts are used to define the growth parameters of patients, who are usually characterized by pre- and postnatal growth retardation (Kline et al., 2007). In point of fact, the body parameters at birth are

frequently under the 10th centile, with an average weight of 2.28 kg, an average length of 45.5 cm and an average head circumference of 30.9 cm. The growth curve runs parallel to the one of the healthy population but tends to decrease under the 5th centile during early childhood (Kline et al., 1993a; Kline et al., 1993b). During adulthood, the average weight and height are 47.6 kg and 156 cm for men, and 30.5 kg and 131 cm for women. The head circumference can reach the average value of 49 cm, consistent with a significant microcephaly (Kline et al., 1993a; Kline et al., 1993b).

Facial dysmorphisms

The facial features are the most recognizable findings in CdLS (Fig.2). Patients usually present with a short neck and a low anterior and posterior hairline. Eyebrows are typically arched, hypertrophic and fused above the bridge of the nose (synophrys). Eyelashes are thick, long and curly. Narrow palpebral fissures and ptosis are also reported. The midface is usually flattened and characterized by the presence of a short nose presenting with a depressed bridge and anteverted nares. Beneath, the philtrum is frequently smooth and long. Exemplary oral features include thin lips with down-turned commissures, downslanting superior lip, a high and arched palate, small and widely spaced teeth and micrognathia (Jackson et al., 1993).

Figure 2. Typical facies observed in four patients with CdLS. These patients present with a variable degree of severity and are placed from the mildest to the most severe. Though, the striking facial features typical for CdLS are discernible in all of them (adapted from Liu and Krantz, 2009)

Low-set dysplastic ears are also observed with a high frequency. The ear canal is often stenotic or narrow, thus causing predisposition to otitis media and sinusitis (Jackson et al., 1993).

The facial features tend to evolve during aging: the face usually lengthens and becomes coarser. In some cases, the facial appearance also becomes less typical with age. For this reason, patients with a mild presentation of the syndrome would have been difficult to diagnose as CdLS if the clinical evaluation would have occurred in adulthood (Kline et al., 2007). Premature aging has also been reported (Kline et al., 2007).

Muscoloskeletal system

The muscoloskeletal system is the district with the highest number of alterations: almost all patients present with upper limbs anomalies that range from the presence of small hands to the presence of severe reductions or malformations (Fig.3). Typical findings are disproportional shortening of the first metacarpal, proximally placed thumbs, brachydactyly, clinodactyly of the fifth finger, radial head dislocation, radioulnar synostosis and incomplete elbow extension (Jackson et al., 1993).

Figure 3. Examples of upper limbs anomalies observed in CdLS patients, placed from the mildest to the most severe: almost all patients present with small hands, but in some cases absence of digits or of the entire forearms are also observed (adapted from Liu and Krantz, 2009)

Lower extremities are usually less affected; the feet are small in almost all cases, and syndactyly of the second and third toes are frequently observed (Jackson et al., 1993).

Other orthopedic complications include scoliosis, development of bunions, tight Achilles tendons, cervical malformations, *pectum excavatum* and hip dislocation or dysplasia, that occurs in 10-15% of patients and normally requires surgery during childhood. Osteoporosis may appear earlier than expected (Jackson et al., 1993).

Gastrointestinal system

Feeding problems are typical during infancy and young childhood, and can be caused by micrognathia, cleft palate and decreased muscle tone in the oral area. In addition, gastroesophageal reflux disease (GERD) is observed in more than 90% of patients and often requires surgical intervention (Luzzani et al., 2003).

Malrotation, diaphragmatic hernia, increased risk for volvulus formation and intestinal obstruction and pyloric stenosis have also been frequently reported (Masumoto et al., 2001).

Cardiac system

Congenital heart defects are observed in about 25% of CdLS patients. Ventriculo-septal defects or atrial-septal defects are the most commonly represented, but tetralogy of Fallot, pulmonic stenosis and hypoplastic left heart syndrome are also frequent (Jackson et al., 1993; Metha and Ambalavanan, 1997; Tsukahara et al., 1998; Kline et al., 2007).

Genitourinary system

Renal dysfunctions and malformations can be seen in approximately 40% of CdLS patients and include pelvic dilatation, vesiculoureteral reflux and renal dysplasia. Male patients frequently present hypoplastic genitalia and cryptorchidism; other common findings are hypospadias and micropenis. On the other hand, small *labia majora* and abnormally formed uteri can be observed in females. Fertility is decreased in severely affected individuals (Jackson et al., 1993).

Other malformations

Ophthalmologically, almost all patients present peripapillary pigmentation. Other common findings are myopia, blepharitis, ptosis and mild forms of microcornea. Nystagmus, cataract, glaucoma and nasolacrimal duct obstruction have been reported with a lower frequency (Jackson et al., 1993).

Vestibular and auditory anomalies include both conductive and sensorineural hearing loss together with recurrent otitis media and sinusitis (Marchisio et al., 2014).

Neurologically, hypoplasia of the cerebellar vermis, enlarged ventricles particularly at the basal cisterns and atrophy of the white matter are frequently observed (Kline et al., 2007). Neurofibrillary tangles inside neurons, myelination defects and gyral structural abnormalities were also reported during autopsies (Yamaguchi K and Ishitobi F, 1999). Patients also present with high pain tolerance, which might be linked to peripheral neuropathy (Kline et al., 2007). Seizures are the main neuropathological manifestation. Though, no specific EEG patterns have been described and the seizures are usually well managed with standard medical treatments (Verrotti et al., 2013). The gait of those patients who achieve walking is often wide-based (Kline et al., 2007).

Patients often present with sleep disturbances, namely frequent awakenings, reduced need of sleeping with an average sleeping time of 2-4 hours per night, and ability to remain awake for a longer time compared to the normal population (Berney et al., 1999).

The presence of diaphragmatic hernia is infrequent but important to mention because of the severe clinical implications (Masumoto et al., 2001).

Generalized hypertrichosis is also one of the main manifestations of CdLS, particularly regarding the face, back and extremities; *cutis marmorata* is also another frequent finding. Small nipples and umbilicus have also been observed (Jackson et al., 1993).

Puberty usually begins at ages 12-13 for girls and 13-14 for boys, slightly later if compared with the average onset of the healthy population (Basile et al., 2007). Menstrual cycles are observed in 76% of female patients, even though frequently irregular. Breast development is reported for 78% of the patients. Fertility is unaffected for those patients with a mild presentation of the syndrome (Kline et al., 2007).

Obesity in association with type II diabetes has also been described (Kline et al., 2007).

Cognitive impairment

Cognitive impairment is always observed with an average IQ around 50. Nevertheless, the severity of the disability is highly variable: both borderline cases characterized by mild learning disabilities and severe impairments have in fact been reported. Learning continues throughout life without evidence of regression, but patients often need supervised living (Jackson et al., 1993).

Developmental delay is also a main feature of CdLS and speech acquisition is usually more severely affected than the motor development. Perceptual organization and visual-spatial memory are more preserved (Kline et al., 2007). Based on the period of language acquisition, patients are divided into four categories:

- Talkers (3-4%): the psychomotor development is almost normal and they spontaneously start talking
- Late talkers (35-40%): they sit independently after 18 months and walk independently after 30 months. They formulate the first words between four and eight years of age. They normally present with attention deficits

- Limited talkers (20-25%): they show a very slow psychomotor development, and pronounce the first words between seven and ten years of age; the ability to combine words appears after 10 years of age
- NonTalkers (25-30%): they usually present with a very severe phenotype, including limb anomalies, hearing loss, autistic features and inability to walk before five years of age. They are unable to speak.

Behavioral problems

Almost all patients show behavioral issues that are usually caused or aggravated by physical complications. The following problems have been frequently reported: obsessive-compulsive disorders, self-injurious behavior, short attention span, attention deficit disorder with or without hyperactivity, depression, extreme shyness and autistic features (Basile et al., 2007). Many of these behavioral issues are thought to be secondary to frustration due to the inability to communicate or to the presence of GERD (Basile et al., 2007).

Social and environmental interactions are achieved at variable degrees (Berney et al., 1999).

Prognosis

The prognosis is usually good in the absence of major malformations and life expectancy is estimated to be 10-20 years shorter compared to the healthy population, depending on the number and severity of the clinical signs (Coppus, 2013).

The most common causes of death are mainly connected to gastrointestinal complications like diaphragmatic hernia, aspiration pneumonia complicated by GERD and volvulus at older ages (Schrier et al., 2011).

MOLECULAR BASIS: THE COHESIN COMPLEX

CdLS is caused by mutations in different subunits or regulators of the cohesin complex; therefore, it falls within the category of the so called cohesinopathies (Krantz et al., 2004; Tonkin et al., 2004; Musio et al., 2006; Deardorff et al., 2007; Deardorff et al., 2012a; Deardorff et al., 2012b).

Cohesin is a chromatin-associated multisubunit protein complex characterized by a high level of conservation during evolution. Cohesin was originally identified for its role in mediating sister chromatid cohesion during both meiosis and mitosis. Nevertheless, several studies have demonstrated an important role of the complex in different cellular processes including DNA repair, chromatin remodeling, regulation of gene expression and long-range interactions between distant genomic regions (Uhlmann, 2008; Sjögren and Ström, 2010; Cuylen and Haering, 2011; Dorsett, 2011; Feeney et al., 2012; Sofueva and Hadjur, 2012). Consequently, cohesin plays a very important role in the maintenance of genomic stability.

Interestingly, cell lines of CdLS patients do not display cohesion defects. This finding suggests that the etiopathology of CdLS is probably not linked to the disruption of cohesion between sister chromatids but to the inability of the cohesin complex to mediate more dosage-sensitive cellular functions (Dorsett and Krantz, 2009).

ARCHITECTURE OF THE COHESIN COMPLEX

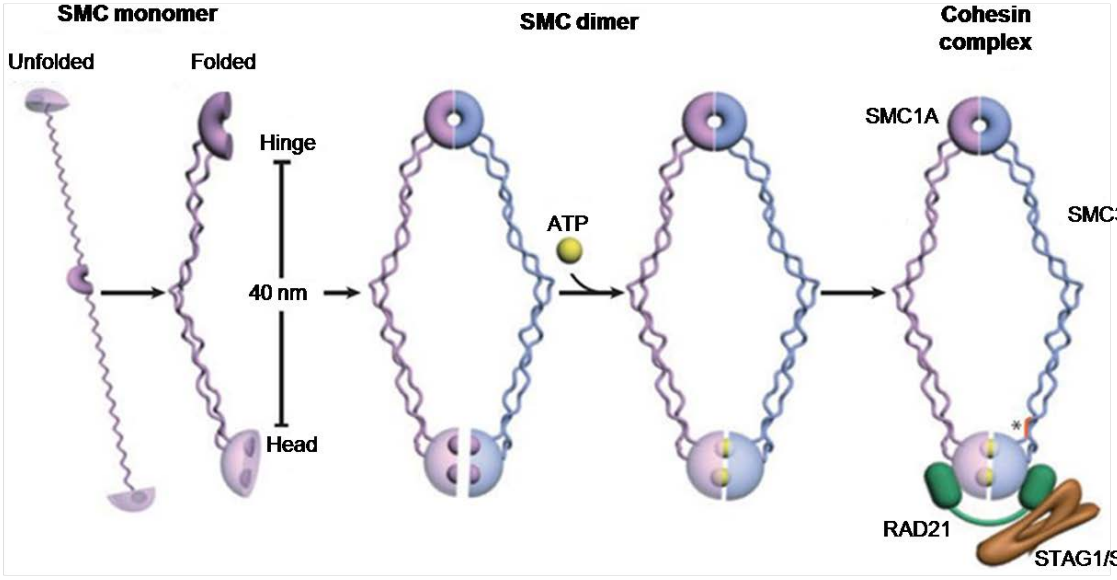


Figure 4. Architecture of the SMC proteins and of the cohesin complex (adapted from Onn et al., 2008)

The cohesin complex consists of four core subunits: SMC1A, SMC3, the α-kleisin protein RAD21 and the HEAT-repeat containing proteins STAG1 or STAG2 (Fig.4; Guacci et al., 1997; Michaelis et al., 1997; Losada et al., 1998; Nasmyth and Haering, 2005).

SMC1A (OMIM #300040, Xp11.22) and SMC3 (OMIM #606062, 10q25.2) are members of the Structural Maintenance of Chromosome (SMC) family, a very large family of ATPases characterized by a peculiar structure: one ATPase domain is localized at both the N- and C-termini of the proteins and the two domains are then connected to each other through a long α-helix. During the folding process, the α-helix bends on itself at the central hinge domain, thus forming 40 nm of anti-parallel coiled-coil structure. The bending of the α-helix brings the ATPases at the N- and C-termini close to each other, generating a globular ATPase head domain (Haering et al., 2002). By interacting through their respective hinge domains, SMC1A and SMC3 constitute the heterodimer that represents the core of the cohesin complex. The heads of both proteins are physically connected to RAD21, thus determining the generation of a tripartite ring structure. RAD21 (OMIM #606462, 8q24.11) belongs to the α-kleisin family; kleisin is a word that comes from the Greek language and means closure. In fact, RAD21 determines the closure of the cohesin ring by bridging the head domains of the SMC proteins. Particularly, the N-terminus of RAD21 binds to the head domain of SMC3, whereas its C-terminus interacts with the ATPase domain of SMC1A (Haering et al., 2002). Hence, SMC1A and SMC3 bind to each other directly through their respective hinge domains and indirectly through RAD21 at their head domains. The association of the SMC proteins with RAD21 requires ATP: the binding of ATP determines the closure of the ring while its hydrolysis is needed for the dissociation of the complex (Arumugam et al., 2003; Weitzer et al., 2003).

The fourth subunit of the cohesin complex, known as Stromalin Antigen (STAG), interacts with RAD21 and contains HEAT repeats that are important for protein-protein interactions (Neuwald and Hirano, 2000; Gruber et al., 2003). The functional role of this last subunit and of its HEAT repeats is still poorly understood. Vertebrates present two different isoforms of the STAG subunit, but the cohesin complex can contain only one subunit per time. STAG1-containing cohesin seems to be responsible for transcriptional regulation and replication of telomeres, whereas complexes containing the STAG2 subunit are mainly involved in sister chromatid cohesion (Losada et al., 2000).

It is worth to be noted that the meiotic cohesin complex differs from the mitotic one: in meiosis, the RAD21 subunit is substituted by the paralog REC8, while the STAG1/2 subunits are replaced by the protein STAG3.

COHESIN-ASSOCIATED PROTEINS

In addition to the core subunits, many other cohesin-associated proteins have been identified that modulate cohesin activity and function (Table 2).

The heterodimeric protein complex formed by NIPBL and MAU2 is the main regulator of cohesin and is responsible for the the loading of cohesin onto DNA (Krantz et al., 2004; Tonkin et al., 2004).

NIPBL (OMIM #608667, 5p13.2) is the homologous of the *Drosophila melanogaster Nipped-B* gene and of the *Saccharomyces cerevisiae Scc2* gene (Sister Chromatid Cohesion). The protein encoded by *NIPBL*, named Delangin, belongs to an evolutionary conserved family of proteins called adherins (Krantz et al., 2004; Tonkin et al., 2004). The human gene consists of 47 exons, while the coding sequence extends from exon 2 to exon 47 and can produce two main isoforms: the longer delangin-A (2804 aa) and the shorter delangin-B (2697 aa), that does not include exon 47 and results from the translation of an additional fragment of exon 46 (Krantz et al., 2004; Tonkin et al., 2004).

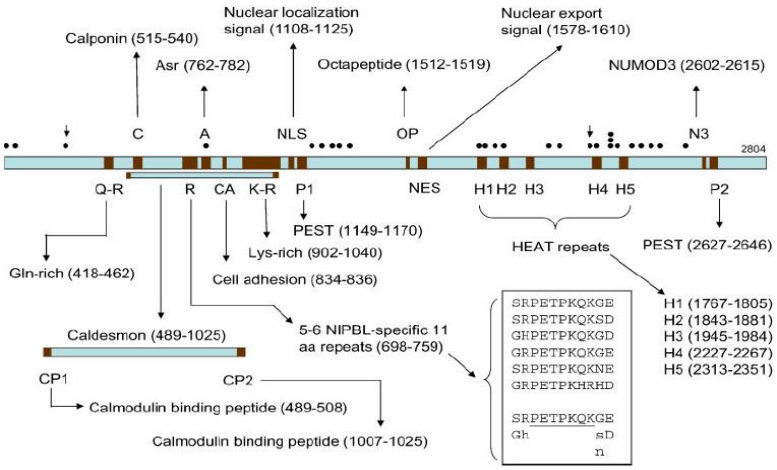


Figure 5. Structural organization of the NIPBL protein. The different domains of NIPBL and their localization are shown together with the amino acid sequences of the NIPBL specific repeats (from Yan et al., 2006)

The amino acid sequence of NIPBL indicates different functional domains (Fig.5). From the N-terminus to the C-terminus, the following domains are predicted: one caldesmon domain, one calponin domain, two calmodulin-binding motifs, one nuclear localization signal, one nuclear export signal, five HEAT repeats and one DNA binding domain (Yan et al., 2006). The caldesmon, calponin and calmodulin domains are so called for their analogy with the Ca_2-dependent proteins CALD1, CNN1 and calmodulin, that are involved in the contractility of the smooth muscle (Winder and Walsh, 1993). The HEAT repeats have instead been proven to play an important role in the establishment of cohesion, since they have been identified in other

cohesin-associated proteins. These repeats consist of sequences of 34-43 aa that are organized in clusters and are known for their role in mediating protein-protein interactions (Neuwald and Hirano, 2000).

Northern Blot analyses have shown that the *NIPBL* transcript is ubiquitously expressed. Though, the level of transcription is tissue-specific. Wheras high levels of *NIPBL* expression are observed in cardiac and skeletal muscles, the expression appear to be very low in brain, lungs and intestine (Krantz et al., 2004; Tonkin et al., 2004).

The adherin NIPBL is involved in various cellular processes including gene expression regulation, sister chromatid cohesion and DNA repair. It was also shown to be functionally linked to chromatin remodeling complexes through the direct interaction with the histones deacetylases (Jahnke et al., 2008). Its role in the regulation of gene expression has been well documented in *Drosophila melanogaster*. Here, the homolog *Nipped-B* activates the transcription of the *cut*, *ultrabithorax* and Notch signaling by mediating long-range interactions between distant enhancers and promoters (Rollins et al., 1999). Furthermore, NIPBL was reported to be involved in the loading of the cohesin complex onto chromatin by stimulating the ATP hydrolysis that determines the opening of the cohesin ring (Dorsett, 2004).

The second component of the cohesin-loader complex MAU2 is characterized by the presence of TRP domains (Tetratricopeptide repeats) that are responsible for protein-protein interactions. Besides mediating the cohesin loading onto DNA in association with NIPBL, little is known about MAU2 functions (Watrin et al., 2006).

PDS5 is also evolutionary conserved and contains HEAT repeats, similarly to other cohesin-associated proteins (Neuwald and Hirano, 2000). Its function is to promote the generation of a cohesive state of the ring through the direct interaction with the structural elements SMC1A, SMC3 and RAD21 and the regulators ESCO1, sororin and WAPL (Tanaka et al., 2001; Dorsett et al., 2005; Losada et al., 2005; Losada and Hirano, 2005). Vertebrates present two different isoforms of *PDS5*, *PDS5A* and *PDS5B*, both involved in the regulation of the cohesin' cycle. The depletion of *PDS5* does not affect the binding of cohesin to the DNA, but prevents the establishment of cohesion (Tanaka et al., 2001; Losada and Hirano, 2005).

Cohesin also interacts with a protein called WAPL (Wings-apart like) through RAD21 and STAG1/STAG2 (Kueng et al., 2006). The main role of WAPL is to mediate the removal of cohesin from chromosomes during prophase, thus playing an antagonist function compared to PDS5. In WAPL-depleted cells cohesin does not dissociate from chromosomes arms; as a consequence, metaphase chromosomes show low resolution of the sister chromatids (Kueng et

al., 2006). Conversely, the overexpression of WAPL was shown to cause premature separation of sister chromatids, thus driving tumorigenesis in mice (Oikawa et al., 2004).

Sororin is another cohesin-associated protein that has been found only in vertebrate cells. As PDS5, Sororin is dispensable for the association of the cohesin complex with the DNA, but is needed for the establishment of cohesion. As point of fact, human cells that were depleted of Sororin display severe cohesion defects, despite the overall amount of chromatin-bound cohesin remains unchanged (Schmitz et al., 2007). Additionally, it has also been suggested that Sororin protects cohesin from a precocious dissociation from the DNA, possibly by inhibiting the activity of WAPL. This is in agreement with the finding that Sororin becomes dispensable for sister chromatid cohesion in the absence of WAPL (Nishiyama et al., 2010). Consequently, it can be assumed that Sororin plays a role both in cohesion establishment and maintenance (Diaz-Martinez et al., 2007).

ESCO1/2, the human hortologues of the *Saccharomyces cerevisiae* Ctf7/Eco1, are also required for the establishment of a cohesive state (Ivanov et al., 2002). These genes encode for acetyltransferases whose human substrate is SMC3. It has been proved that depletion of ESCO1 or ESCO2 in different cell lines is responsible for massive cohesion defects (Unal et al., 2008; Zhang et al., 2008). Hence, ESCO1/2 are essential for the establishment of cohesion. As a proof, mutations in *ESCO2* result in Roberts syndrome, a genetic disorder associated to a peculiar cytological phenotype characterized by heterochromatin repulsion of sister chromatids (OMIM #268300; Vega et al., 2005).

After the separation of sister chromatids, SMC3 needs to be deacetylated in order to allow the proper dissolution of pro-cohesive elements and the recycling of cohesin for the next cycle. HDAC8 (OMIM #300269, Xq13.1) was recently identified as the vertebrate deacetylase for SMC3 (Deardorff et al., 2012b). Thus, the role of HDAC8 is indispensable for the establishment of a new functional cohesin complex.

S. cerevisiae	Human	*Drosophila*	*S. pombe*	Function
Smc1	Smc1a	Smc1	Psm1	Cohesin subunit
	Smc1b			Cohesin subunit (meiosis)
Smc3	Smc3	Smc3	Psm3	Cohesin subunit
Mcd1/Scc1	Rad21/Scc1	Rad21	Rad21	Cohesin subunit (α-kleisin)
Rec8	Rad21L/Rec8	C(2)M	Rec8	Cohesin subunit (meiotic α-kleisin)
Scc3/Irr1	SA1	SA	Psc3	Cohesin subunit
	SA2	SA2		
	STAG3		Rec11	Cohesin subunit (meiosis)
Pds5	Pds5a	Pds5	Pds5	Cohesion maintenance
	Pds5b			
Wpl1/Rad61	Wap1	Wap1	Wap1	Cohesin dissociation
Sororin	CDCA5			Cohesion establishment/maintenance
Scc2	Nipbl/Scc2	*Nipped-B*	Mis4	Cohesin loading
Scc4	Mau2/Scc4	Scc4	Ssl3	Cohesin loading
Eco1/Ctf7	Esco1	Deco	Eso1	Cohesion establishment
	Esco2	San		
Esp1	Espl1	Sse	Cut1	Separase
		Thr		
Pds1	Pttg1	Pim	Cut2	Securin
Hos1	HDAC8	?	?	Smc3 deacetylase
Sgo1	Sgo1	MEI-S332	Sgo1	Protection of centromeric cohesin
	Sgo2		Sgo2	

Table 2. Cohesin structural elements and regulatory proteins in different organisms (from Metha et al., 2013)

MODELS OF INTERACTION OF COHESIN WITH DNA

Numerous models have been proposed in order to explain how the cohesin complex is able to hold sister chromatids together (Losada, 2007). These models can be classified into four categories: the embrace model, the two rings model, the bracelet model and the rod model (Fig.6).

The embrace model (Fig.6a), also known as one ring model or topological model, seems to be the most plausible one; based on this theory, the tripartite structure formed by SMC1A, SMC3 and RAD21 embraces the two DNA duplexes, that remain therefore entrapped into the cohesin ring following the closure of the ring determined by the binding of ATP. In the two ring model (Fig.6b), one single cohesin ring encircles one sister chromatid; the two chromatids are then hold together by direct or indirect interaction of the two cohesin complexes. Based on the bracelet model (Fig.6c), SMC heterodimers form multimeric filaments that entrap sister chromatids inside the oligomeric bracelet. The multiple heterodimers are connected to each other through RAD21.

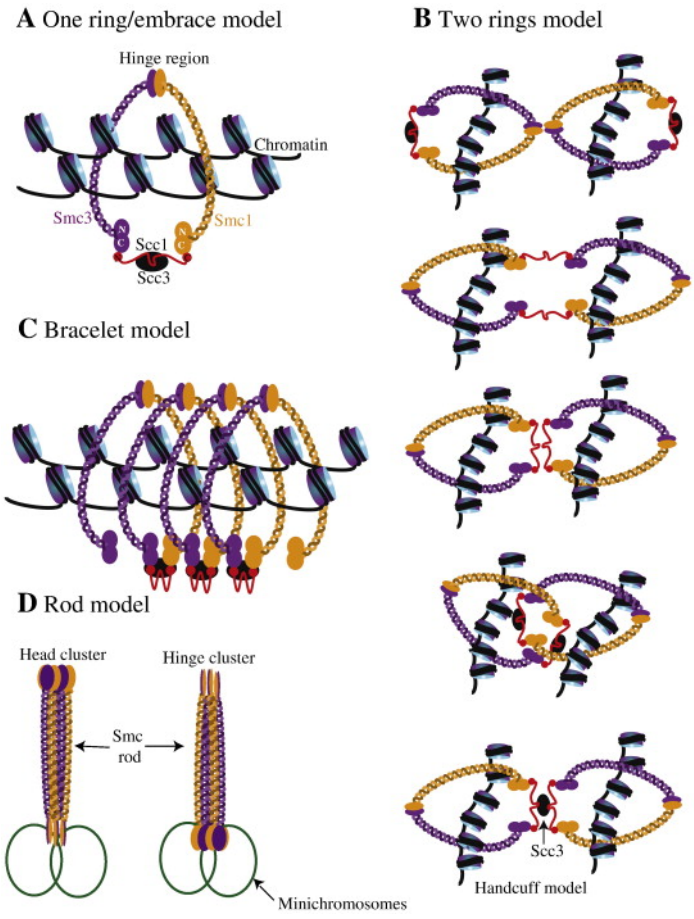


Figure 6. Models proposed to explain how the cohesin complex is able to mediate cohesion of sister chromatids during mitosis (from Mehta et al., 2012)

Lastly, for the fourth model (Fig.6d), a rod shaped structure of cohesin have been proposed. Transmission electron microscopy experiments in budding yeast have shown that cohesin present a rod shaped structure and that sister minichromosomes are able to interact at one end of the rod. Moreover, the width of the rod indicates that multiple cohesin complexes interact with each other along their arms in order to mediate cohesion.

CELL CYCLE REGULATION BY COHESIN

The activity of the cohesin complex is tightly regulated throughout the cell cycle in order to allow the proper dissociation of sister chromatids (Fig.7). Interactions with accessory proteins and post-translational modifications are responsible for this regulation.

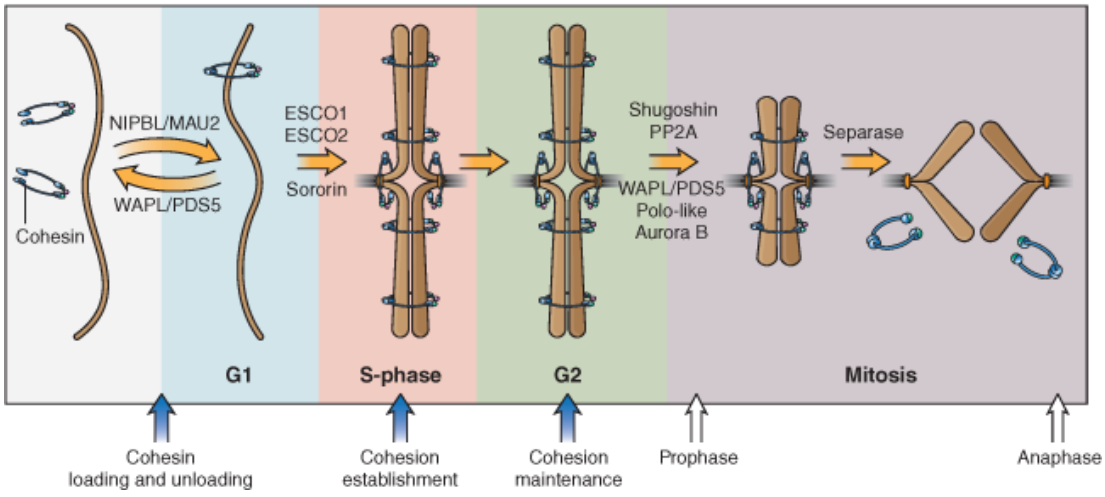


Figure 7. Schematic representation of the regulation of the cohesin complex throughout the cell cycle (from Zakari et al., 2015)

The loading of cohesin onto chromatin

In order to hold sister chromatids together, cohesin must bind to the DNA before its replication. In vertebrates, the loading of cohesin onto the DNA is performed during the telophase of the previous cell cycle by the protein complex NIPBL/MAU2 (Michaelis et al., 1997; Losada et al., 2000; Haering et al., 2004). No sequence specificity has been observed for the loading of cohesin onto chromosomes: its initial positions and its permanent locations vary among different organisms (Michaelis et al., 1997; Losada et al., 2000; Haering et al., 2004).

Once cohesin is bound to the DNA, the cohesin loader NIPBL/MAU2 becomes dispensable (Ciosk et al., 2000). This suggests that this complex is responsible for the physical association between cohesin and chromosomes, but it does not have any role in the assembly of the cohesin complex starting from the single subunits or in the establishment and maintenance of cohesion.

Therefore, it can be speculated that a preassembled cohesin ring gets loaded onto chromatin. For this to be possible, the cohesin complex should be opened in order to be able to embrace sister chromatids. In this context, evidence suggests that chromosomes enter the cohesin ring through its hinge domain (Gruber et al., 2006). The energy required for this opening is given by the mechanism through which the hinge bends towards the ATPase heads; in this way the head domains become able to provide the energy for the opening of the ring through the hydrolysis of ATP (McIntyre et al., 2007).

Establishment of cohesion

The association of the cohesin complex with chromosomes is not enough to keep sister chromatids together. For this purpose, DNA-associated cohesin must achieve a "pairing competent state", a process known as establishment of cohesion. For this attainment, the key player are the conserved proteins ESCO1/ESCO2, that are responsible for the acetylation of SMC3 (Ivanov et al., 2002). Cell cycle dependent studies in yeasts and humans demonstrated that SMC3 is not acetylated during the G1 phase, but its level of acetylation rises during S phase. Accordingly, the essential function of ESCO1/ESCO2 in generating cohesion is confined to the S phase. Hence, these acetyltransferases are only required for the establishment of a cohesive state concomitant to replication, but are unable to mediate the maintaining of cohesion. Nevertheless, the "pairing competent state" should be preserved until the beginning of anaphase. In this context, PDS5 plays a very important role because it ensures cohesion throughout the G2 phase and during the transition between G2 and M. In point of fact, as previously mentioned, PDS5 interacts with other cohesin-associated proteins like Sororin and WAPL. It has been demonstrated that the association of PDS5 with Sororin maintains a cohesive state, while a non-cohesive state is promoted when PDS5 interacts with WAPL (Sutani et al., 2009; Nishiyama et al., 2010). In other words, Sororin and WAPL compete for the binding with PDS5 in order to mediate the maintenance or the disruption of cohesion, respectively.

The removal of cohesin from chromatin

The dissociation of cohesin from the DNA is mediated by two different pathways based on the position of the complex along chromosomes: cohesins localized on chromosomes arms are removed during prophase whereas cohesins bound to pericentromeric and centromeric regions are preserved until the onset of anaphase (Waizenegger et al., 2000). The dissociation of cohesin during prophase is mediated by the two mitotic kinases Polo-like kinase 1 (Plk1) and Aurora B, which are responsible for the phosphorylation of RAD21 and of the STAG subunits, thus triggering the opening of the cohesin ring. During this phase, centromeric cohesins are protected from the activity of the kinases by the protein shugoshin and the phosphatase PP2A: Shugoshin directly interacts with PP2A and recruits it at the centromeres, thus neutralizing the activity of the kinases (Giménez-Abián et al., 2004; McNairn e Gerton, 2008). It is worth to be noted that this prophase pathway determines the dissociation of cohesin from chromosomes without cleaving the single subunits, thus sparing the complex from the destruction. As a result, cells exit mitosis with almost an unchanged pool of cohesin complexes available for a new loading immediately after the generation of the nuclear envelope. This is of high physiological significance because it has been demonstrated that cohesin plays an important role in gene expression regulation during the G1 phase, and so before the establishment of cohesion (Wendt et al., 2008).

The dissociation of cohesin during anaphase and the complete dissolution of the complex are mediated by three elements: securin, separase and the Anaphase-Promoting Complex (APC). Before the onset of anaphase, the proteolytic activity of the separase is inhibited because of its physical association with securin. Once the sister kinetochores becomes bioriented on mitotic spindles, the anaphase checkpoint is inactivated, thus determining the activation of APC. Activated-APC mediates the ubiquitin-dependent degradation of the securin. The separase is now free to cleave RAD21: the cohesin complex dissociates from the DNA and sister chromatids can now reach the opposite poles of the cell (Ciosk et al., 1998; Uhlmann et al, 1999).

The deacetylation of SMC3 performed by HDAC8 is then essential in both prophase and anaphase pathways in order to allow the proper dissolution of the cohesin complex and the recycling of the subunits for the next cell cycle (Deardorff et al., 2012b).

COHESIN´S FUNCTIONS

In the last decade, extensive research on the function and structure of cohesins has revealed that the complex is involved in many other cellular processes apart from mediating sister chromatid cohesion. These additional functions include transcriptional regulation, chromatin remodeling, DNA double-strand breaks (DSB) repair, chromosomes condensation and morphogenesis and centrosome-related activities (Fig.8; Uhlmann, 2008; Sjögren and Ström, 2010; Cuylen and Hearing, 2011; Dorsett, 2011; Feeney et al., 2012; Sofueva and Hadjur, 2012). It is worth to be noted that sister chromatid cohesion per se might be directly involved in these non-canonical roles of the cohesin complex.

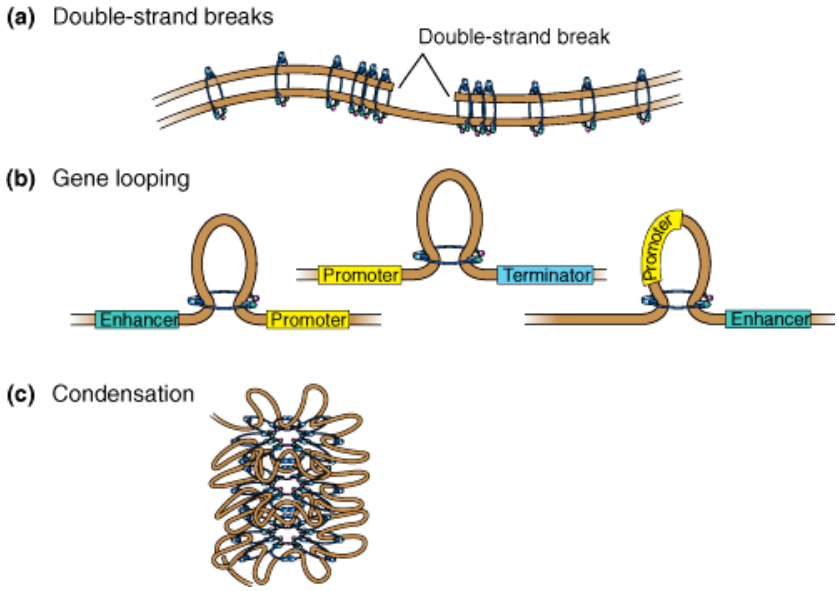


Figure 8. Schematic representation of the non-canonical functions of cohesins. (a) DSB repair is mediated by cohesin binding to the break site. In response to a DSB, cohesion is also re-enforced genome-wide. (b) Cohesin regulates gene expression through gene looping. (c) Cohesin promotes chromosome condensation (from Zakari et al., 2015)

Transcriptional regulation

The earliest evidence of the role of cohesin as a key regulator of transcription came with studies on *Saccharomyces cerevisiae* and *Drosophila melanogaster* (Donze et al., 1999; Rollins et al., 1999).

In budding yeast cohesin was found to play an important role at the silent mating type locus, where it acts as boundary element in order to restrict the spreading of transcriptional silencing (Donze et al., 1999). In fission yeast the cohesin complex was instead shown to regulate transcriptional termination (Gullerova and Proudfoot, 2008).

In *Drosophila melanogaster*, both *Nipped-B* and cohesin were found to be involved in gene expression regulation. Specifically, Nipped-B activates the transcription of the homeobox genes *Cut* and *Ultrabithorax* by mediating the long-range interaction of the promoters with the respective enhancers (Rollins et al., 1999). On the other hand, in postmitotic mushroom body γ-neurons, cohesin is required for the transcriptional activation of the Ecdyson steiroid hormone Receptor (EcR), a factor indispensable for axon pruning (Schuldiner et al., 2008). Recent studies in *Drosophila melanogaster* have also proved that cohesin selectively binds to genes on which the RNA polymerase II is paused just downstream of the transcription start site. Therefore, it is presumable that cohesin is involved in the regulation of the transition of the polymerase between a paused state and an active transcription elongation state (Fay et al., 2011).

In the following years, the role of *NIPBL* and cohesin in the transcriptional regulation was demonstrated also in other model organisms. In addition, even translation has been proved to be influenced by cohesin, thought indirectly. In point of fact, cohesin was reported to augment the translational capacity of cells by increasing the level of transcription of rRNA in budding yeast and humans (Bose et al., 2012).

Nevertheless, the mechanisms by which cohesin mediates the regulation of gene expression are still not fully understood and vary among different organisms. For instance, in *Drosophila melanogaster* cohesin accumulates together with Nipped-B at the transcriptional start sites of active genes (Misulovin et al., 2008). Instead, in humans, cohesin co-localizes with both NIPBL and the CCCTC-binding factor (CTCF), a well-known insulator (Wendt et al., 2008).

The way cohesin fulfills its role in the regulation of gene expression is linked with the mechanism through which it mediates sister chromatid cohesion. As previously mentioned, cohesin is able to hold sister chromatids together by entrapping them in a ring-like structure. In the same way, the cohesin ring might be able to bring together non-contiguous genomic regions through the generation of DNA loops. Long-range interactions between distant genomic regions obtained thanks to DNA looping might therefore be the key event through which cohesin is able to influence gene expression (Rollins et al., 1999). Notably, cohesin can induce both transcriptional activation or repression depending on which elements are brought in close proximity and on whether an enhancer or a silencer is sequestered into the loop.

Chromatin remodeling

Cohesin was also shown to interact with chromatin remodeling complexes. In particular, ChIP experiments showed a direct interaction between the cohesin subunit RAD21 and the ATPase SNF2h of a chromatin remodeling complex. This cooperation is needed to modify histone tails in order to regulate the accessibility of the DNA (Hakimi et al., 2002). In addition, NIPBL was found to interact with the deacetylases HDAC1 and HDAC3. The minimal region of interaction in NIPBL is confined to 163 aa at the C-terminal of the protein (1838-2000). Mutations in this region determine an impairment of the ability of NIPBL to bind HDAC1 and HDAC3 (Jahnke et al., 2008). It might be assumed that the previously reported interactions are needed to promote the binding of cohesin to the target DNA following the modification of the chromatin structure.

DNA repair

Studies on *Saccharomyces pombe* provided the first evidence for the involvement of cohesin in DNA repair, and so in the maintenance of genomic stability. In fact, in fission yeast, cohesin mutants are sensitive to γ-irradiation due to a defect in the repairing of DSBs (Birkenbihl and Subramani, 1992). Since then, the role of cohesin in DSBs repair was confirmed in different organisms (Sjörgen and Nasmyth, 2001; Musio et al., 2005; Cortés-Ledesma and Aguilera, 2006).

At the beginning, the repair of DSBs was thought to depend on the cohesin complexes already bound to the DNA during the S phase. Conversely, it has been demonstrated that, in the presence of DNA damage, cohesin can be loaded onto DNA and generate cohesion even in the post-replicative phase as a result of the increased activity of the cohesin loader and of the acetylases ESCO1/2 (Ström et al., 2004). It is worth to be noted that this damage-induced cohesion is generated globally on all chromosomes and not only at the site of damage (Unal et al., 2007). In this context, the job of cohesin is to bring sister chromatids together, so that the wild type chromatid can be used as template for homologous recombination in order to mediate the repair of the strand carrying the DSB.

In budding yeast, epigenetic marks were analyzed to understand the targeting of the cohesin complex at the damage site. It has been demonstrated that the main epigenetic mark for this purpose is the phosphorylation of histone H2AX by the checkpoint kinases Mec1 and Tel1 at least 60 kb from the DSB site (Unal et al., 2004). In addition, also the MRX complex

(Mre11/Rad50/Xrs2) is able to promote the assembly of cohesin at breakage sites (Fig.9; Unal et al., 2004).

In humans, cohesin has also been found to play an important role for the activation of the intra S phase and G2-M checkpoints following the generation of a DSB (Kim et al., 2002; Yazdi et al., 2002; Luo et al., 2008; Watrin and Peters, 2009).

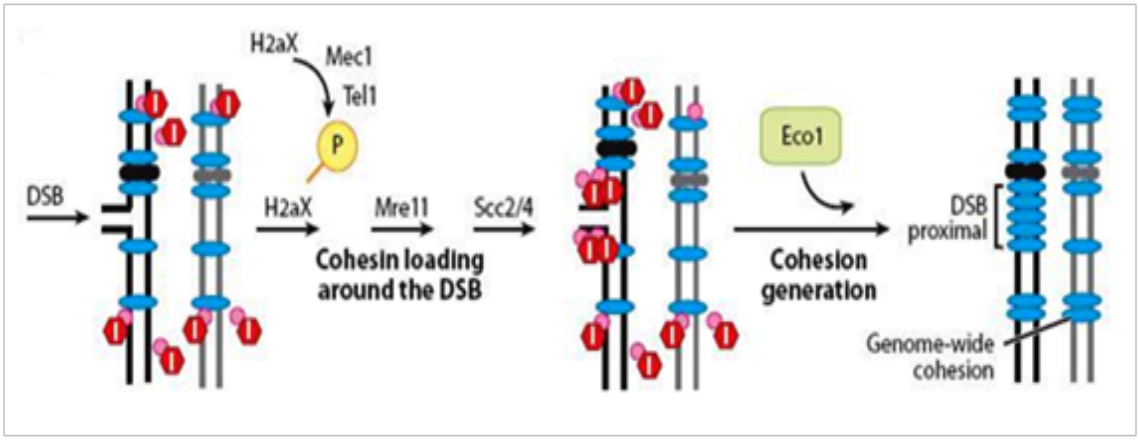


Figure 9. Schematic representation of DNA repair mediated by cohesin in yeast: upon a DSB, cohesin loading is directed to the region of the damage following the phosphorylation of H2AX and binding of MRX to the break site.The inhibition of Eco1 is then removed, thus allowing the generation of a cohesive-state (from Onn et al., 2008)

Interestingly, it has been shown in budding yeast that a reduction of more than 80% of the cohesin' activity is associated to normal cohesion, but leads to defects in DNA repair and chromosome condensation (Ellermeier and Smith, 2005). These data underlines the importance of cohesin stoichiometry for the regulation of cellular processes different from sister chromatid cohesion, that hence appear to be highly dosage-sensitive. Accordingly, a small amount of weakly bound cohesin seems to be sufficient to mediate sister chromatid cohesion whereas non-canonical functions depend upon high levels of strongly bound cohesin. This hypothesis is further confirmed by the absence of cohesion defects in cell lines of CdLS patients (Castronovo et al. 2009; Dorsett and Krantz, 2009).

Chromosomes condensation

Chromosome morphogenesis is an essential step for the correct distribution of the genetic material during mitosis and depend on DNA condensation. Chromosome compaction is mainly mediated by condensins, a highly conserved cohesin-like complex containing SMC proteins (Lavoie et al., 2002).

In fission yeast and budding yeast changes in the level of cohesin were found to be responsible for a hypo- or hyper-condensation of chromosomes, possibly due to altered targeting of condensins to the chromosomes (Hartman et al., 2000; Losada et al., 2002). These results suggest that cohesin might be involved in the control of the localization of condensins on chromosomes. This hypothesis was proved in 2004, when a study in budding yeast demonstrated that chromosome compaction occurs in two steps, and that the first step depends on cohesin (Lavoie et al., 2004). A further confirmation of this theory was given by Chromatin Immunoprecipitation (ChIP) experiments on budding yeast, which proved that occupancy sites of condensins overlap with those of the cohesin loader complex, thus implying a co-localization of condensins with cohesin (D'Ambrosio et al., 2008). Nevertheless, the mechanisms of interaction between condensin and cohesin are still poorly understood.

Centrosome-related functions

Several studies have demonstrated a role of cohesin in the duplication of the microtubule-organizing centers, also known as centrosomes. The first evidence came in 2006, when it was shown that the separase is needed not only for the cleavage of the cohesin ring, but also for centriole disengagement and centrosome duplication (Tsou and Stearns, 2006). In the following years cohesin subunits were also identified at the spindle pole and centrosomes (Wong and Blobel, 2008; Kong et al., 2009). Moreover, depletion of RAD21 in HeLa cells was associated with a premature separation of paired centrioles, and also shugoshin was shown to be required for the protection of the cohesion between centrioles (Wang et al., 2008; Nakamura et al., 2009). Lastly, the premature activation of the separase or the depletion of the shugoshin were found to cause both premature sister chromatid separation and premature centriole disengagement (Schockel et al., 2011). These data suggest that the chromosome and centrosome cycles might be coordinated with each other as a result of the dual role played by cohesin (Schockel et al., 2011).

THE CdLS-GENES: STATE OF ART

To date, mutations in five different cohesin-related genes have been described as a genetic cause of CdLS, namely *NIPBL*, *SMC1A*, *SMC3*, *RAD21* and *HDAC8* (Krantz et al., 2004; Tonkin et al., 2004; Musio et al., 2006; Deardorff et al., 2007; Deardorff et al., 2012a; Deardorff et al., 2012b).

NIPBL

The cohesin loader *NIPBL* is the "major gene" of the syndrome: mutations can be identified in more than half of CdLS cases (Borck et al., 2004; Gillis et al., 2004; Selicorni et al., 2007).

To date, more than 300 variants affecting *NIPBL* have been identified in CdLS patients. Most of the mutations arise *de novo*, though also a small number of familial cases has been described. All types of variants have been reported so far, including out-of-frame deletions or insertions, missense, nonsense, and splicing mutations. Large scale genomic alterations have also been identified, but with a lower frequency. Among point mutations, out-of-frame deletions resulting in premature termination codons are the most frequent (28%), followed by missense substitutions (21%), nonsense and splicing mutations (17% each), duplications (14%), insertions (1,5%) and indels (1,5%) (Mannini et al., 2013).

Mutations in *NIPBL* are scattered along the entire gene without apparent hotspots. Nevertheless, a large number of variants has been identified in exons 10 and 47 because of their length (Oliveira et al., 2010). Out-of-frame deletions, nonsense and splicing mutations, all presumably leading to a truncated and nonfunctional NIPBL protein, are associated with a more typical and severe phenotype characterized by profound developmental and cognitive delay with absence of speech, severe pre- and postnatal growth retardation, typical facial dysmorphisms and high frequency of limb reductions or malformations. Conversely, missense substitutions are generally associated with a milder phenotype; in point of fact, limb anomalies are rarely observed and cognitive impairment, growth retardation and developmental delay are less severe (Fig.10 A-H) (Selicorni et al. 2007). Patients with large scale genomic rearrangements present instead with a phenotype that correlates with the size of the region and with the number of exons involved: patients with deletions spanning beyond *NIPBL* sequence or involving large portions of the gene are associated with severe growth and cognitive delay and with upper limb malformations. Contrariwise, deletions of one or few exons frequently result in a milder phenotype (Russo et al., 2012; Gervasini et al., 2013; Mannini et al., 2013).

The severity of the phenotype also depends on the functional domain involved. Indeed, variants affecting the HEAT repeats were shown to be associated with a more severe phenotype, putatively based on the modified ability of NIPBL to interact with other proteins (Mannini et al., 2013). Nevertheless, CdLS unrelated patients sharing the same mutation might present with different clinical features. This observation points out that, beside the effects of the mutations themselves, other genetic or environmental factors may play an important role in determining the phenotypcal expression (Mannini et al., 2013).

Cell lines of CdLS patients carrying mutations in *NIPBL* show a reduction of the *NIPBL* transcript of 30%, thus suggesting that the pathogenic mechanism is likely to be linked to haploinsufficiency (Dorsett and Krantz, 2009).

Furthermore, mosaicism was recently proved to play a very important role in association with *NIPBL*: more than 20% of patients who were reported to be mutation-negative on DNA extracted from blood were then found to carry a mutation in *NIPBL* on DNA isolated from buccal mucosa or fibroblasts (Husiman et al., 2012; Baquero-Montoya et al., 2014). No significant differences were identified between the phenotype of patients with somatic mosaicism o with a constitutional mutations, thus excluding the possibility to discern in advance the two categories of patients (Huisman et al., 2013). Three mechanisms have been proposed in order to explain the onset of mosaicism, namely somatic mutation after fertilization, negative selection of mutant lymphocytes or loss of the mutation due to reversion. Though, negative selection seems to be the most likely explanation, since no remarkable phenotypical differences were observed between patients with or without mosaicism, and since mutation reversion is mainly associated to skin disorders (Huisman et al., 2013).

SMC1A

Mutations in *SMC1A* have been identified in about 5% of CdLS patients. So far, 44 CdLS individuals harboring 29 different mutations have been described (Borck et al., 2007; Deardorff et al., 2007, Liu et al., 2009; Limongelli et al., 2010; Mannini et al., 2010; Pié et al., 2010; Hoppman-Chaney et al., 2011).

Mutations were identified throughout the entire protein with the exception of the highly conserved hinge domain. Moreover, only missense mutations or in frame-deletions have been reported in CdLS patients with mutations in *SMC1A*. The observation that nonsense mutations

or out-of-frame deletions have never been reported points out that they might lead to a different phenotype or they are likely not tolerated (Mannini et al., 2013).

In agreement with the X-linked dominant transmission, females present with a very broad spectrum of phenotypical features ranging from very mild to very severe. Conversely, males show a more homogeneous and severe clinical presentation. In general, patients harboring mutations in *SMC1A* are characterized by a milder phenotype compared to those who carry mutations in *NIPBL*: in point of fact *SMC1A*-patients present with a tendency towards normal auxological parameters, a mild to moderate cognitive impairment and less remarkable facial features (Fig.10 M-N; Mannini et al., 2013).

Interestingly, *SMC1A* localizes at Xp11.22, a region of the X-chromosome that partially escapes X-inactivation in humans. The rate of escaping of the allele on the inactive X-chromosome ranges from 15% to 30% (Carrell and Willard, 2005). It is worth to be noted that healthy females were shown to express *SMC1A* at higher levels compared to healthy males (Liu et al., 2009). Moreover, patients with mutations in *SMC1A* were shown to express equal levels of transcript and protein compared to controls (Revenkova et al., 2009). Furthermore, mutations in *SMC1A* do not prevent the incorporation of the mutant protein into the cohesin complex and SMC1A-mutated heterodimers show a higher affinity for chromatin than normal dimers (Revenkova et al., 2009; Gimigliano et al., 2012). All together, these findings suggest that a dominant negative effect might be the pathogenic mechanism associated to mutations in *SMC1A*.

SMC3

To date, only one mutation in *SMC3* has been described in a single CdLS patient (Fig.10 Q-U). The mutation is a small in-frame deletion of three nucleotides (c.1464-1466delAGA) that results in the deletion of a single amino acid (p.E488del) and localizes in the coiled-coil region of the SMC3 protein (Deardorff et al., 2007). The patient showed pre- and postnatal growth retardation and moderate cognitive and developmental delay. He presented with GERD, but did not show any other major malformation (Deardorff et al., 2007). The low frequency of mutations in *SMC3* indicates that this gene might play important roles beyond cohesion (Mannini et al., 2013). In agreement with this hypothesis, SMC3 acetylation was found to control the fork processivity in human cells (Zhang et al., 2008).

RAD21

Mutations in *RAD21* were identified in eight CdLS patients (Deardorff et al., 2012a; Minor et al., 2014). Cell lines expressing mutant-RAD21 showed significantly lower cell survival and elevated levels of chromosomal structural rearrangements compared to normal cells after irradiation (Deardorff et al., 2012a).

Mutations in this gene are responsible for a human cohesinopathy that overlaps with CdLS and is mainly characterized by short stature, minor skeletal anomalies and facial features including synophrys, micrognathia and brachydactyly (Fig.10 O-P). Interestingly, individuals with mutations in *RAD21* present with a significantly milder cognitive impairment (Deardorff et al., 2012a).

HDAC8

To date, mutations in *HDAC8* have been reported in about 5% of patients (Deardorff et al., 2012b; Kaiser et al., 2014; Feng et al., 2014; Yuan et al., 2015). All types of mutations have been identified so far; nevertheless, no correlations between the mutation and the clinical phenotype could be deduced. Similarly to *NIPBL* haploinsufficiency, mutations in *HDAC8* were found to cause only a modest reduction of the levels of the transcript (Deardorff et al., 2012b).

Patients with mutations in *HDAC8* tend to present with a phenotype that partially overlaps with CdLS but that is also characterized by atypical features including the late closure of the anterior fontanelle, a broad or bulbous nasal tip, hypertelorism, hooding of the eyelids, mosaic skin pigmentation and teeth anomalies (Fig.10 I-L; Kaiser et al., 2014). Importantly, *HDAC8* is also an X-linked gene. Accordingly, as for *SMC1A*, males show a homogeneous and more severe clinical presentation while females, who represents the vast majority of *HDAC8*-mutated CdLS patients, present with a broad range of phenotypes also based on the level of X-inactivation. In point of fact, most of the female patients show complete skewing towards the wild type allele in blood, thus indicating a strong selection against the mutation (Deardorff et al., 2012b; Kaiser et al., 2014).

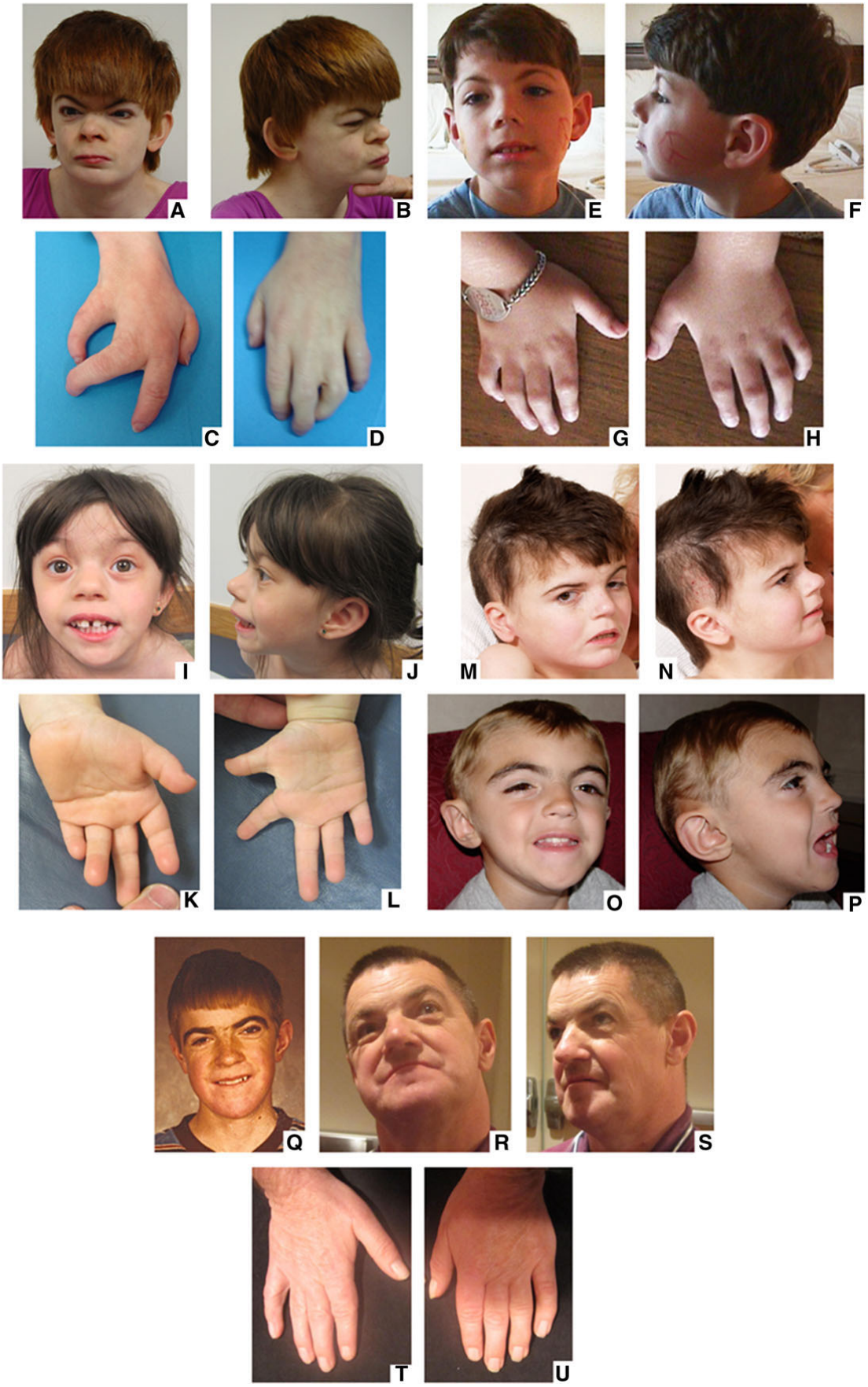


Figure 10. Phenotype of patients carrying mutations in the five cohesin-associated genes responsible for CdLS. (A-D) Female patient carrying a truncating mutation in NIPBL. (E-H) Male patient harboring a missense mutation in NIPBL. (I-L) Female patient with an in-frame deletion in HDAC8. (M-N) Female patient with a missense mutation in SMC1A. (O-P) Male patient with a deletion of RAD21. (Q-U) Male patient carrying the in-frame deletion in SMC3 (from Mannini et al., 2013).

MISSING HERITABILITY

The five genes *NIPBL*, *SMC1A*, *SMC3*, *RAD21*, and *HDAC8*, also taking into account the role of mosaicism, are responsible for about 70% of CdLS cases. Thus, the molecular cause underlying the syndrome is still not known in about 30% of CdLS patients.

A lot of effort has been invested in order to try to cover this missing heritability. It can be speculated that the remaining cases could be explained by the presence of undetectable mutations in the known CdLS genes as well as of mutations affecting genes with a function similar to the one of cohesin. The employment of high-throughput techniques such as next generation sequencing or CGH-array could help to shed light on this phenomenon.

DIFFERENTIAL DIAGNOSIS

Part of the missing heritability might also be linked to sometimes miscarried diagnosis due to phenotypical overlap between CdLS and other syndromes. The main syndromes that represent a differential diagnosis to CdLS are the following.

KBG SYNDROME

KBG syndrome (OMIM #148050) is a rare genetic disorder characterized by postnatal growth retardation, developmental delay, cognitive impairment, skeletal anomalies, brachycephaly and a typical round *facies* displaying broad and arched eyebrows, hypertelorism, a broad or bulbous nasal tip and a long philtrum (Hermann et al., 1975; Ockeloen et al., 2014). The hallmark of the syndrome is the presence of dental anomalies, a feature that is instead not common for CdLS (Ockeloen et al., 2014).

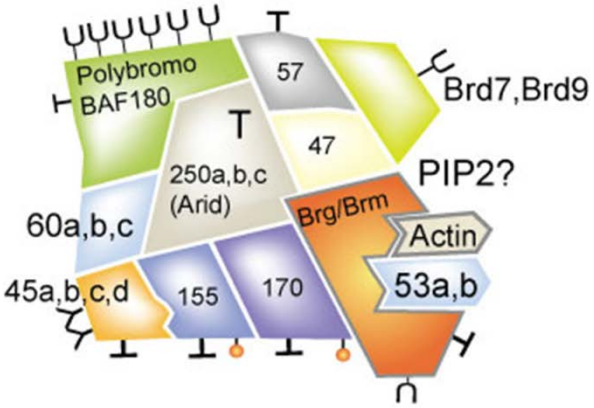
KBG syndrome is caused by loss-of-function mutations in the *ANKRD11* gene (OMIM #611192, 16q24.3), encoding for a transcriptional repressor also known as ankyrin repeat-containing protein 11 (Sirmaci et al., 2011). ANKRD11 contains one activation domain that is able to stimulate transcription, two repression domains and five ankyrin repeats needed to mediate protein-protein interactions (Zhang et al., 2007). The protein is able to inhibit transcriptional activation by interacting with the deacetylases HDAC3, HDAC4 and HDAC5 and by recruiting them to target genes of nuclear receptors, thus inhibiting their transcription (Sirmaci et al., 2011). ANKRD11 was also found to localize within the nuclei of neurons and to

accumulate in distinct inclusions after their depolarization. This data suggests that ANKRD11 also plays a role in neural plasticity (Sirmaci et al., 2011).

COFFIN-SIRIS SYNDROME

Coffin-Siris syndrome (CSS, OMIM #135900) is a rare multisystem developmental disorder characterized by a coarse *facies* and typical facial features including thick and arched eyebrows, long eyelashes, a broad nasal bridge and nasal tip, a wide mouth and thick and everted lips (Schrier et al., 2012). CSS patients also present with short stature, cognitive impairment, psychomotor delay, feeding difficulties and hypoplasia or aplasia of the fifth fingernail or digit. Additionally, they can display hirsutism, microcephaly and sparse scalp hair (Schrier et al., 2012).

Mutations in different subunits of the SWItch/Sucrose Non-Fermentable (SWI/SNF) complex were found to be associated with the syndrome (Santen et al., 2012; Tsurusaki et al., 2012).



Mouse	Human
BAF250a,b,c	ARID1A,B,C
BAF 57	SMARCE1
BAF47	SMARCB1
BAF155	SMARCC1
BAF170	SMARCC2
Brg1	SMARCA4
Brm	SMARCA2

Figure 11. Schematic representation of the SWI/SNF complex, that is highly conserved during evolution. The names of the main subunitsare indicated for different organisms (mouse and humans)

The SWI/SNF complex (Fig.11) is a family of highly conserved chromatin remodelers. Though, the subunits of the complex can vary among different organisms. In addition, in humans, the numerous subunits are encoded by more than one gene, so that different combinatorial assemblies of the complex are possible. The various combinations are then tightly regulated during cell development and differentiation (Kosho et al., 2014). Typically, a human SWI/SNF complex contains a single ATPase, represented by SMARCA2 (OMIM #600014, 9p24.3) or SMARCA4 (OMIM #603254, 19p13.2), some core subunits including SMARCB1 (OMIM #601607, 22q11.23), SMARCC1 and SMARCC2 and some accessory subunits, the most important of which are ARID1A (OMIM #603024, 1p36.11), ARID1B

(OMIM #614556, 6q25.3) and SMARCE1 (OMIM #603111, 17q21.2). To date, CSS-causing mutations have been identified in *ARID1B, ARID1A*, *SMARCA2*, *SMARCB1* and *SMARCA4*, thus indicating genetic heterogeneity (Santen et al., 2012; Tsurusaki et al., 2012). In addition, mutations in the SWI/SNF-associated protein PHF6 (OMIM #300414, Xq26.2) have also been proven to cause CSS (Wieczorek et al., 2013). *ARID1B* represents the major gene of the syndrome and loss-of-function mutations can be identified in more than 70% of CSS cases (Wieczorek et al., 2013). All the *ARID1B*-mutations so far identified are truncating and are predicted to determine haploinsufficiency (Wieczorek et al., 2013).

As cohesin, the SWI/SNF complex is involved in various cellular processes including ATP-dependent chromatin remodeling, gene expression regulation, cell cycle regulation and differentiation. Accordingly, mutations in subunits of the SWI/SNF complex have been identified in different types of cancer (Wieczorek et al., 2013).

WIEDEMANN-STEINER SYNDROME

Wiedemann-Steiner syndrome (WDSTS, OMIM # 605130) is a heterogeneous disorder also known as “hairy elbows syndrome” due to a peculiar and frequently observed excessive hair growth around the proximal third of the forearm and the distal third of the upper arm. The syndrome is also characterized by short stature, intellectual disability and distinctive facial features including thick eyebrows, long eyelashes, vertically narrowed and downslanted palpebral fissures, broad nasal bridge and nasal tip (Wiedemann et al., 1989; Steiner and Marques, 2000).

Mutations in the *KMT2A* gene (OMIM #159555, 11q23.3) are responsible for WDSTS (Jones et al., 2012). KMT2A is a lysine-specific methyltransferase that plays an essential role in the regulation of gene expression during hematopoiesis and early development. The protein is highly conserved and consists in multiple functional domains, the most important of which is the SET domain, which is responsible for the enzymatic activity (Jones et al., 2012).

Multiple chromosomal translocations involving *KMT2A* are described in association with acute lymphoid leukemias and acute myeloid leukemias (Nilson et al., 1996; Taki et al., 1996).

RUBINSTEIN-TAYBI SYNDROME

Rubinstein-Taybi syndrome (RSTS, OMIM #180849) is characterized by postnatal growth retardation, intellectual disability, broad and angulated thumbs and halluces and typical facial features that include downslanted palpebral fissures, a beaked nose with the columella extending below the nares, grimacing smile, highly arched palate and Talon cusps (Rubinstein and Taybi, 1963; Hennekam, 2006).

To date, mutations in two different genes have been associated with RSTS. The major gene of the syndrome is *CREBBP* (OMIM #600140, 16p13.3). *CREBBP* mutations are responsible for more than 50% of RSTS cases (Bartsch et al., 2005; Bentivegna et al., 2006; Spena et al., 2014). The second gene *EP300* (OMIM #602700, 22q13.2) accounts for about 8% of RTS cases (Roelfsema et al., 2005; Negri et al., 2015). Both CREBBP and EP300 are transcriptional co-activators involved in the regulation of the expression of various signaling pathways. The two genes encode for homologous proteins highly conserved during evolution that play an important role as acetyltransferases (Roelfsema et al., 2005). Interestingly, a patient with a clinical diagnosis of CdLS was found to carry a mutation in the *EP300* gene in 2013 (Fig.12; Woods et al., 2013).

Figure 12. Facial appearance of the CdLS patient carrying a mutation in the EP300 gene. To be noted the presence of thick and arched eyebrows with synophrys, long eyelashes, depressed nasal bridge with anteverted nares and smooth philtrum (from Woods et al., 2013)

The patient, a Caucasian male born to healthy parents, presented with typical CdLS facial features including microbrachycephaly, thick and arched eyebrows with synophrys, ptosis, long eyelashes, depressed nasal bridge with anteverted nares, smooth and long philtrum, and thin lips. He also displayed growth retardation, cognitive impairment, developmental delay, hirsutism, respiratory distress, supraventricular tachycardia, hypospadias, cryptorchidism,

increased bifrontal subarachnoid spaces and a severe, progressive and rigid scoliosis. He died at the age of 5 years 3 months because of severe bronchopneumonia (Woods et al., 2013).

The finding of a mutation in *EP300* in a patient with a clinical diagnosis of CdLS is of particular interest and highlights the already suggested phenotypical overlap between CdLS and RSTS (Gorlin et al., 2001). A functional link between *EP300* and the cohesin-associated genes has also been established. In point of fact, both EP300 and HDAC8 are key regulators of the transcriptional activity of p53 (Yan et al., 2013). In addition, EP300 and NIPBL have been shown to localize at putative enhancer regions in different cell types and were recently ranked as top enhancer signatures (Chen et al., 2012). It cannot be excluded that *EP300* and the cohesin-associated genes are involved in the regulation of other common pathways, thus explaining the phenotypical overlap between the two syndromes.

AIM OF THE PRESENT WORK

CdLS is a heterogeneous disease characterized by clinical and genetic heterogeneity. Mutations in five genes have been so far identified as a genetic cause of CdLS, namely *NIPBL*, *SMC1A*, *SMC3*, *RAD21* and *HDAC8*. Nevertheless, the impact each gene exerts on the syndrome has yet to be defined. In addition, mutations in the five genes are all together responsible for about 70% of CdLS cases. Hence, a molecular diagnosis is still not available for about 30% of CdLS patients.

The present work originates from an international cooperation that involves different countries and is divided into the following section:

- Definition of the mutational frequency of each CdLS-associated gene with particular attention to the role played by mosaicism. The assessment of the genotype-phenotype correlation is also of utmost importance for the definition of the most appropriate molecular testing in the presence of a diagnosis of CdLS

- Investigation of the peculiar pattern of expression of *SMC1A* in healthy controls and in the CdLS population. In point of fact, this gene localizes on a region of the X chromosome that partially escapes X-inactivation. Therefore, the study of the overall and allele-specific expression of this gene could shed light on the pathogenic mechanisms underlying the syndrome in the presence of mutations in *SMC1A*

- Identification of new causative genes by the use of next generation sequencing techniques. In this context, it can be speculated that genes encoding for proteins functionally associated with cohesin or with chromatin-related activity might be responsible for CdLS-overlapping phenotypes. Both targeted gene panel or genome-wide approaches can be used for this purpose

MATERIALS AND METHODS

PATIENTS

The patients herein described were recruited thanks to a large international cooperation that includes Italy, Spain, Germany, Sweden, Poland and the United States.

All procedures involving human participants were performed in accordance with the ethical standards of the respective national research committees. Informed consent was obtained from all individuals participants included in this study. An additional consent was obtained for the publication of photographic material comprising patients.

All children underwent complete physical and dysmorphologic evaluation by their referring clinical geneticists, who assigned a diagnosis of CdLS to most of the patients. Patients with an uncertain diagnosis but presenting CdLS-overlapping features were also included in this study.

Patients' DNA was analyzed through different techniques including Sanger sequencing and next generation sequencing approaches in order to identify the causative mutations.

CELL CULTURE ESTABLISHMENT

Cell cultures were generated from lymphoblasts or fibroblasts.

Lymphoblastoid cell lines (LCLs) were established in collaboration with the Galliera Genetic Bank of the Galliera Hospital through Epstein-Barr virus transformation. LCLs were cultured in RPMI 1640 medium supplemented with 20% Fetal Bovine Serum (FBS) and 1% Antibiotics. Cells were grown at 37°C with 5% CO_2.

Skin biopsies were used to establish fibroblasts cell lines, which were grown in DMEM supplemented with 10% FBS and 1% Antibiotics at 37°C with 5% CO_2.

DNA ISOLATION

Genomic DNA was isolated from blood, LCLs, fibroblasts, buccal mucosa and urine.

DNA ISOLATION FROM BLOOD, LCLs AND FIBROBLASTS

DNA extraction from blood, LCLs and fibroblasts was performed with the QIAamp® DNA blood Mini Kit (Qiagen, Hilden, Germany), which enables rapid and efficient purification of high quality genomic DNA from a diverse variety of sample materials for a broad range of downstream applications. The isolation was performed according to the manufacturer's instruction starting from 200 µl of whole blood or $5x10^6$ cells.

DNA ISOLATION FROM URINE

Genomic DNA was isolated from 50 ml of urine.

Cells were sedimented by centrifuging at 5000 rpm for 20 minutes at 4°C. The pellet was resuspended in 1 ml TNE buffer (10 mM Tris pH 8, 1 mM EDTA pH 8, 100 mM NaCl) and transferred to a 1,5 ml Eppendorf tube. Cells were then centrifuged again at 3000 rpm for 10 minutes at 4°C and the pellet resuspended in 400 µl TNE plus (10 mM Tris pH 8, 1 mM EDTA pH 8, 100 mM NaCl, 1% SDS, 100 µg/ml Proteinase K). The resulting solution was incubated overnight in a thermo-shaker at 56°C and 500 rpm.

The day after, 200 µl Phenol and 200 µl Clorophorm were added to the resuspension, which was then vortexed and incubated at room temperature for 10 minutes. After the incubation, a centrifugation at room temperature at 10000 rpm for 10 minutes was performed in order to let the different phases separate.

The DNA contained in the supernatant was precipitated with 40 µl of 3M Sodium Acetate and 1 ml of 100% Isopropanol (centrifugation at 14000 for 30 minutes at 4°C) and subsequently washed with 500 µl of 70% Ethanol (centrifugation at 14000 rpm for 5 minutes at 4°C).

The purified DNA was dried in a speed vacuum and resuspended in 30 µl of distilled water.

DNA ISOLATION FROM BUCCAL MUCOSA

Isolation of genomic DNA from buccal mucosa was performed with the MasterAmp™ Buccal Swab Kit (Epicentre, Cambridge, UK) according to the manufacturer's instructions.

The resulting DNA was precipitated with 1/10 volume of 3M Sodium Acetate and two volumes of Ethanol 100% by centrifuging at 14000 rpm for 30 minutes at 4°C. After discarding the supernatant, the DNA was washed with 500 µl of 70% Ethanol and centrifuged at 14000 rpm for 5 minutes at 4°C. The purified DNA was then dried in a speed vacuum and resuspended in 20 µl of distilled water.

SEQUENCING APPROACHES

Different techniques were employed in order to identify the causative mutations of each patient. Specifically, patients were analyzed through Sanger sequencing or next generation sequencing approaches targeted to the exome or to a cluster of selected gene.

SANGER SEQUENCING

The genes of interest were amplified with primers flanking the intron-exon boundaries with the GoTaq® Flexi DNA Polymerase (Promega, Madison, WI, USA) with the following protocol:

H_2O	13,875 µl		
Buffer 5x	5 µl	95°C x 5'	
dNTPs 2 mM	2,5 µl	95°C x 30''	x 35 cycles
$MgCl_2$ 25 mM	1,5 µl	Ta°C x 30''	
Primer Forward (25 µM)	0,5 µl	72°C x 30''	
Primer Reverse (25 µM)	0,5 µl	72°C x 5'	
Taq (5 U/µl)	0,125 µl	4°C ∞	
DNA (50 ng/µl)	1 µl		

PCR products were sequenced with the Big Dye terminator v3.1 Sequencing Kit following the manufacturer's protocol and run on the ABI PRISM 3130xl Genetic Analyzer (Applied Biosystems, Foster City, CA, USA). Electropherograms were analyzed with the ChromasPro Software version 1.7.6 (Technelysium Pty Ltd, Tewantin QLD, Australia).

TARGETED GENE PANEL

Targeted next generation sequencing was based on the design of a custom-made gene panel for the Personal Genome Machine (Ion Torrent PGM) with the AmpliSeq Designer Tool (Life Technologies, Darmstadt, Germany). The designed panel amplifies 75 kb of selected genomic sequences including the five known CdLS-associated genes, with an estimate coverage of >95%. After quantification with the Qubit® 2.0 Fluorometer (Life Technologies), libraries were prepared using 20 ng of gDNA as starting template with the Ion AmpliSeq Library Kit (Life Technologies) according to the manufacturer's instructions.

IonSphere particles were loaded with the clonally amplified templates through the Ion PGM Template OT2 200 Kit; the loaded template positive particles were then sequenced on the Ion 316v2 TM Chip with the Ion PGM Sequencing 200 Kit v2 (Life Technologies). Given the high frequency of mosaic mutations in CdLS, the filtering process was adjusted to a low stringency protocol that excluded only those variants that were represented in less than 5% of the total reads.

SNaPshot ASSAY

SNaPshot assays (SNaPshot® Multiplex Kit, Life Technologies) were used to confirm low percentage variants emerged from the targeted gene panel. In fact, this primer-extension based assay allows the detection of DNA sequence alterations with a very high sensitivity by the use of fluorescent-labeled ddNTPs.

The region of interest was amplified with the GoTaq® Flexi DNA Polymerase (Promega) as previously described. The PCR product was purified with ExoSAP-IT® (Affymetrix, Santa Clara, CA, USA) following the manufacturer's instructions.The SNaPshot reaction was carried out as follows:

H_2O	1 µl		
SNaPshot Multiplex Ready Reaction Mix	5 µl	96°C x 10'' 50°C x 5'' 60°C x 30''	x 25 cycles
Sequencing primer (2 µM)	1 µl		
Purified PCR	3 µl	4°C ∞	

SNaPshot samples were run on the ABI PRISM 3130xl Genetic Analyzer (Applied Biosystems) and evaluated with the GeneScan® 4.0 Analysis Software (Life Technologies).

EXOME SEQUENCING

DNAs were assayed for quality and quantity with the Qubit® 2.0 Fluorometer (Life Technologies).

The TruSeq Exome Enrichment Preparation Kit (Illumina, San Diego, CA, USA) was used for exome library preparation following the manufacturer'sprotocol. Briefly, 2 µg of gDNA were fragmented by Covaris E220; subsequently, end repair, A-tailing, adapter/barcode ligation and clonal amplification were performed.

The pooled libraries, at a final concentration of 8 pMol, were then sequenced on the Illumina's HiSeq 2500 system (Illumina).

For bioinformatic analysis, read tags were aligned with the human reference genome (hg19) using bwa 0.6.1 (Li and Durbin, 2010). After removing duplicate and off-target reads, GATK (DePristo et al., 2011) was used to perform a joint call of SNPs and Indels. Variants were filtered after Variant Quality Score Recalibration and annotated to dbSNP using SnpSift. Functional impact of each variant was predicted using snpEff (Cingolani et al., 2012). Variants that were predicted to impact the coding sequences and with a frequency in the population lower than 2% were retained. Resulting variants were filtered according to a dominant/de novo inheritance model.

CGH-ARRAY

Peripheral blood DNA samples were tested with the qChip Post oligonucleotide microarray (Quantitative Genomic Medicine Laboratories, Barcelona, Spain). This CGH-array comprises 60,000 probes mainly located in pericentromeric and subtelomeric regions, as well as regions associated to syndromes caused by recurrent genomic alterations. The overall median probes spacing is 35 Kb and the resolution is of 100-125Kb.

After hybridization of the DNA of the patients and of the reference DNA, slides were scanned and analyzed for relative gain or loss of fluorescent signals. Genomic region analyses were performed according to the human reference sequence build 36.1 (hg18) with the Genomic Workbench software.

EXPRESSION STUDIES

The level of expression of *SMC1A* was assessed in seven female patients carrying missense mutations. The X-inactivation status of each patient was first evaluated in order to discriminate between a random or skewed X-inactivation. Subsequently, Real Time PCR and semiquantitative Western Blot were performed in order to assess the level of expression of the *SMC1A* transcript and of the SMC1A protein, respectively. Lastly, a pyrosequencing assay was carried out in order compare the expression of the wild type and mutant alleles.

X-INACTIVATION ANALYSIS

Detection of the X-inactivation status was performed at the HUMARA (Human Androgen Receptor gene) and DXS6673E loci. For each sample, 250 ng of gDNA extracted from blood or LCLs were digested with the appropriate enzymes.

Digestion of the HUMARA locus was performed with 2.5 U of the methylation sensitive restriction enzymes HhaI and HpaII (New England Biolabs, Ipswich, Massachusetts, USA), while the digestion of the DXS6673E locus was performed with 5 U of HhaI and RsaI (New England Biolabs). In both cases, gDNA was digested overnight at 37°C, and enzymes were inactivated for 20 minutes at 80°C. Additionally, a no enzymes control was set up, and a male control was additionally digested in order to verify the completeness of enzymatic cleavage. Digested and undigested samples were subsequently PCR-amplified; the HUMARA locus was amplified with the Fw 5'-FAM-TCCAGAATCTGTTCCAGAGCGTGC-3' and Rev 5'-GCTGTGAAGGTTGCTGTTCCTCAT-3' primers, while the DXS6673E locus was amplified with the following primers: Fw 5'-FAM-ATGCTAAGGACCATCCAGGA-3' and Rev 5'-GGAGTTTTCCTCCCTCACCA-3'.

PCR products were separated on the ABI PRISM 3130xl Genetic Analyzer (Applied Biosystems) and the GeneScan® 4.0 Analysis Software (Life Technologies) was used to determine peak position and area intensity of each allele. The level X-inactivation was calculated by dividing the ratio of the allele peak areas in the digested sample by the ratio of the allele peak areas in the undigested sample.

RNA EXTRACTION AND cDNA SYNTHESIS

RNA extraction was performed from blood or LCLs. RNA isolation from blood was performed using the Tempus Spin RNA Isolation Reagent Kit (Applied Biosystems) according to manufacturer's instructions. Extraction from LCLs was instead performed with the TRI reagent RNA Isolation Reagent (Sigma-Aldrich, St. Louis, MO, USA). Subsequent treatment with DNase I (RNase-free, New England Biolabs) was carried out on all RNA samples in order to avoid genomic DNA contaminations.

The High Capacity cDNA Reverse Transcription Kit (Applied Biosystem) was used to retro-transcribe 250 ng of RNA with random examers. cDNA synthesis was performed in two independent experiments for each sample.

REAL TIME PCR

The expression level of the *SMC1A* transcript was assessed by the use of the Real Time PCR assay. The investigation was run on the OneStep Real Time PCR System Instrument (Applied Biosystems). The TaqMan gene expression assay (assay ID: Hs01091953-m1; Applied Biosystems), spanning the junction between exons 19 and 20 of *SMC1A*, was used to carry out the Real Time PCR reaction. Based on efficiency and stability experiments, the *RPLP0* gene was selected as endogenous normalizer and amplified with the TaqMan gene expression assay ID:4333761F (Applied Biosystems).

Relative gene expression was determined using the ΔΔCt method, as previously described (Livak and Schmittgen, 2001).

PROTEIN EXTRACTION AND WESTERN BLOT

LCLs were lysed using lysis buffer (150 mM NaCl, 50 mM Tris pH 7.5, 1% NP-40, 0.25% deoxycholic acid) and a protease inhibitor cocktail (Roche, Mannheim, Germany). The suspension was incubated on ice for 30 minutes with occasional inversion to obtain complete lysis. The lysate was centrifuged at 14000 rpm for 25 minutes at 4°C. The supernatant (whole cell lysate) was stored at -20°C. Protein concentration was determined through the BCA Protein Assay Kit (Pierce, Rockford, IL, USA) according to the manufacturer's protocol with the ND-1000 Spectrophotometer (Thermo Fisher Scientific Inc., NanoDrop Products).

For each sample, 20 µg of proteins were supplemented with Laemmli buffer (60 mM Tris-Cl pH 6.8, 2% SDS, 10% glycerol, 5% β-mercaptoethanol, 0.01% bromophenol blue) and denatured at 95°C for 5 minutes. Proteins were separated by 8% SDS polyacrylamide gels and then transferred by electroblotting to a polyvinylidene fluoride (PVDF) membrane (Roche). Subsequently, membranes were washed twice in PBS-T (100 mM NaCl, 80 mM Na_2HPO_4, 20 mM NaH_2PO_4, 0.3% Tween20) and non-specific binding was blocked by incubating the membranes in 5% skimmed milk in PBS-T for 1 hour at room temperature. Membranes were incubated overnight at 4°C in agitation with the respective antibodies, namely the rabbit anti-SMC1A antibody (1:2000) (ab21583; Abcam, Cambridge, UK) and the mouse anti-GAPDH antibody (1:12500) (ab8245; Abcam), both diluted in PBS-T. After four washing steps in PBS-T, the membranes were incubated at room temperature for one hour in agitation with the HRP-conjugated secondary antibodies diluted in PBS-T: the goat anti-rabbit IgG (1:10.000) (sc-2004; Santa Cruz Biotechnology, Santa Cruz, CA, USA) and the goat anti-mouse IgG (1:25000) (sc-2005; Santa Cruz Biotechnology) were used for SMC1A and GAPDH detection, respectively. After four washes in PBS-T and two washes in PBS, bound antibodies were detected using enhanced chemiluminescence (Westar ηC; Cyangene, Bologna, Italy).

Blot images were acquired with GboxChemi XT4 system (Syngene, Cambridge, UK) and semiquantitative analysis of the SMC1A protein expression was performed using the Gene Tools Gel Analysis software (Syngene) after normalization with the expression of the GAPDH protein.

PYROSEQUENCING ASSAY

Pyrosequencing is a quantitative technique based on the activity of the luciferase enzyme.

Briefly, pyrosequencing consists of five main steps (Fig.13):

1. The sequence of interest, after being amplified, is incubated with the four different enzymes DNA polymerase, ATP sulfurylase, luciferase and apyrase and the two substrates Adenosine 5'-phosphosulfate (APS) and luciferin.
2. The four dNTPs are added one by one. Every time a dNTP is added, two different scenarios might occur: in the first case, the dNTP is degraded by the apyrase because there is no complementarity with the PCR-template. In the second scenario, the dNTP is complementary to the template, and the DNA polymerase incorporates it to the newly

forming strand, thus determining the release of one molecule of inorganic pyrophosphate (PPi) for each dNTP that has been added.

3. After the release of the PPi and in the presence of the substrate APS, the ATP sulfurylase is able to catalyze the generation of one molecule of adenosine triphosphate (ATP) for each molecule of PPi that has been released. The so formed ATP is then used by the luciferase to convert its substrate luciferin into oxyluciferin with emission of light. The amount of produced light is directly proportional to the amount of produced ATP, and so to the amount of PPi that was released, and consequently to the number of dNTPs that were incorporated to the newly forming strand. The emitted light is then captured by a Charged Coupled Device (CCD) and converted into a single peak in the program generated by the analysis software. The height of the peak is directly proportional to the number of incorporated dNTPs.
4. The apyrase catalyzes the degradation of the dNTPs which were not incorporated into the new DNA strand.
5. Another dNTP is added, and the cycle starts over.

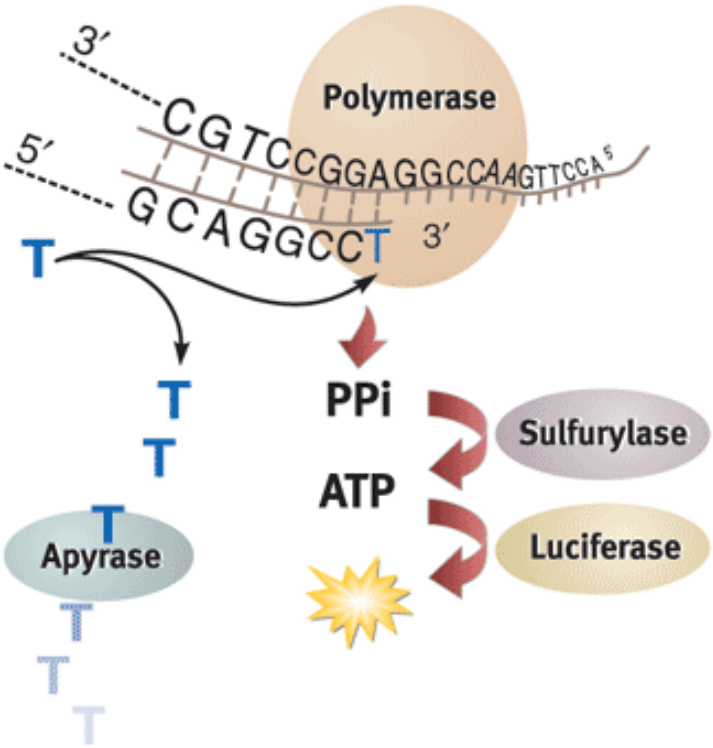


Figure 13. Diagram of operation of pyrosequencing

For the run, the target region was amplified in a 50 µl reaction using the GoTaq Flexi DNA polymerase (Promega) and a specific couple of primers, one of which biotinylated. The analysis was performed with the Pyro Gold Reagent Kit (Qiagen). Briefly, 40 µl of biotinylated PCR product were mixed for 15 minutes at 1400 rpm with 40 µl of a Binding Mix made up of 37 µl of Binding Buffer (10 mM Tris-HCL; 2 M NaCl; 1 mM EDTA; 0.1 % TweenTM 20; pH 7,6) and 3 µl of Streptavidin-Sepharose Beads. The PCR product bound to the beads was then denatured and washed with a vacuum pump. The resulting single strand biotinylated PCR product bound to the beads was hybridized with the sequencing primer at a final concentration of 2 µM in a 96 wells PSQ Plate (Qiagen) and incubated at 80°C for 2 minutes. The plate was then transferred to the Pyro MarkID instrument (Qiagen). The PyroMark Q96 ID Software 2.5 (Qiagen) was used for the analysis.

PROTEIN-PROTEIN INTERACTION STUDIES

Protein-protein interaction studies were carried out in order to verify putative interactions of the subunits of the cohesin complex with other transcription factors.

CLONING

PCR based cloning was used to place DNA fragments encoding for the gene of interest into the backbone of the vectors needed for the yeast two hybrid assay.

PCRs were carried out using primers containing the desired restriction sites at the 5' position and with the high fidelity polymerase TaKaRa primeSTAR GXL DNA Polymerase (Clontech, Saint-Germain-en-Laye, France) with the following protocol:

H_2O	30 µl		
Buffer 5x	10 µl	98°C x 1'	
dNTPs 10 mM	4 µl	98°C x 10''	x 35 cycles
Primer Forward (10 mM)	1,5 µl	Ta°C x 15''	
Primer Reverse (10 mM)	1,5 µl	68°C x 1'x kb	
Taq (5 U/µl)	1 µl	68°C x 5'	
cDNA	2 µl	4°C ∞	

In order to produce PCR fragments suitable for a TA-cloning, 0,25 µl of Taq Poymerase (MP Biomedicals, Santa Ana, CA, USA) were added to the PCR reaction after the run, and the tubes were then incubated at 72°C for 20 minutes.

PCR fragment were run on 1% agarose gel and the specific band corresponding to the fragment of interest was cut out. Gel extraction was subsequently performed with the QIAquick Gel Extraction Kit (Qiagen) according to the manufacturer's instruction.

The purified fragment was ligated with the shuttle vector pGEM®-T Easy (Promega) by the use of the T4 DNA Ligase (Promega). *E. coli* DH5α cells were then chemically transformed with the recombinant DNA construct and plated onto LB agar plates containing 100 µg/ml of ampicillin, in order to specifically select the transformed bacterial cells. Transformed bacteria containing the plasmid of interest were grown overnight at 37°C at 225 rpm in 3 ml of LB medium supplemented with 3 µl of ampicillin. Plasmids were then extracted using the QIAGEN Plasmid Mini Kit (Qiagen) following the manufacturer's instructions, digested with the appropriate restriction enzymes and run on 1% agarose gel in order to check the restriction profile. Clones presenting the right profile were sequenced in order to verify the absence of mutations. Wild type plasmids were then subcloned into the expression vectors of the yeast two hybrid assay, namely pGADT7 and pGBKT7 (Clontech). Transformed bacteria containing the plasmid of interest were grown overnight at 37°C at 225 rpm in 50 ml of LB medium supplemented with 50µl of antibiotic. Plasmids were then extracted using the QIAGEN Plasmid Midi Kit (Qiagen) following the manufacturer's instructions.

YEAST TWO HYBRID ASSAY

The yeast two hybrid assay is used to detect protein-protein interactions (Fig.14). The technique is highly sensitive and is able to identify also weak and transient interactions. This assay is based on the properties of the yeast transcriptional activator GAL4, which can be divided in two domains: one DNA binding domain (BD) responsible for the interaction of the protein with the DNA, and one activation domain (AD) responsible for the transcriptional activation of the target gene. In the yeast two hybrid assay, a gene of interest (bait gene) is fused to the GAL4 BD, while a putative interactor (prey gene) is expressed as a fusion to the GAL4 AD. If the proteins encoded by the bait and prey genes interact, the BD and AD domains of the GAL4 are brought into proximity, thus activating the transcription of four reporter genes in the AH109 yeast strain, namely *ADE2*, *HIS3*, *MEL1* and *LacZ*, which confer nutritional advantage to the transformed yeasts.

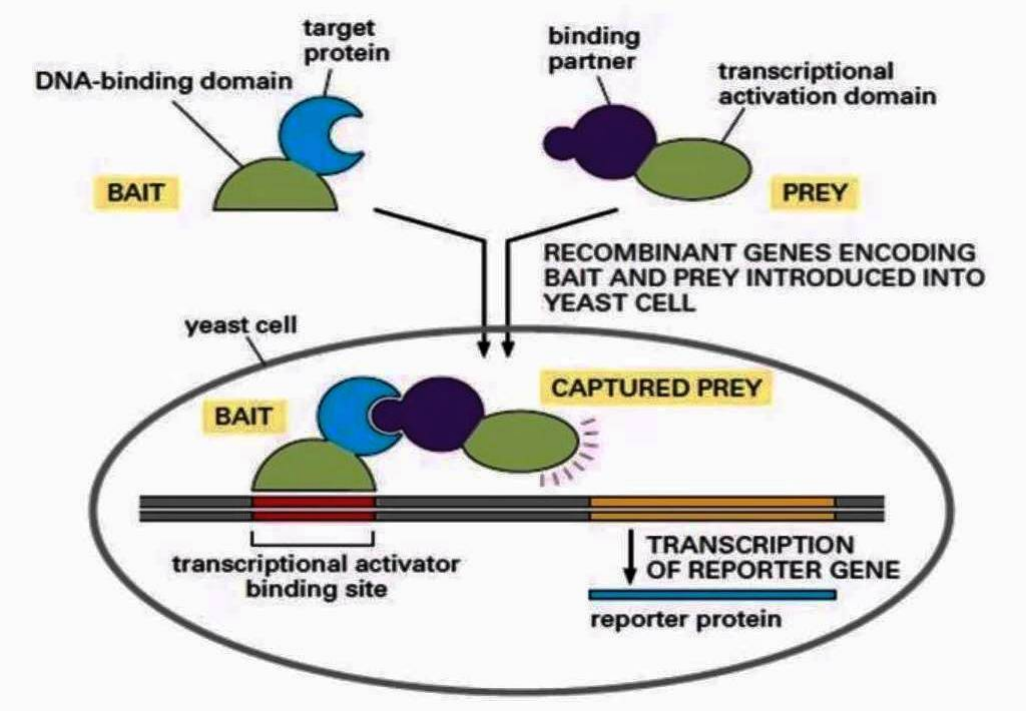


Figure 14. Schematic representation of the yeast two hybrid assay

Transformation of the yeast AH109 strain was carried out with the lithium acetate method (Gietz and Woods, 2002).

Yeasts growing on an agar plate were inoculated in 5 ml of YPD medium (Yeast Extract-Peptone-Dextrose) and grown for 2-3 days at 30°C and 250 rpm (pre-culture). The day of the transformation, the entire pre-culture was poured into 50 ml of YPD medium and grown at 30°C at 250 rpm for 5-6 hours. Once the desired density was reached, yeasts were centrifuged at 3000 rpm for 5 minutes and then washed with 20 ml of sterile water. Two other washing steps were subsequently performed: one in 1 ml 1x TE Buffer (Tris/EDTA) and one in 1 ml of 1x TE/LiAc (TE/Lithium Acetate). The resulting pellet was resuspended in 250-300 µl TE/LiAc. The transformation mix was prepared as follows: 50 µl of yeasts suspension, 5 µl of salmon sperm, 300 µl of PEG 40 (Polyethylenglycol) and 1 µg of each midi preparation (bait and prey plasmids). After vortexing, the transformation mix was incubated at 30°C for 30 minutes. Subsequent heat-shock was carried out at 42°C for 15 minutes, followed by two minutes on ice. The entire transformation mix was then plated on -2 SD Plates. Three or four days after the transformation, -2 and -4 star plates were prepared and incubated at 30°C.

EXPRESSION CONTROLS

Immunoblotting was performed in order to verify the expression of the protein of interests.

Liquid cultures were established by inoculating yeasts in 3 ml of -2 SD medium and grown for 2-3 days at 30°C at 250 rpm. The liquid culture was then centrifuged for 5 minutes at 3000 rpm. After discarding the supernatant, the yeast pellet was freezed with liquid nitrogen and subsequently heat-shocked at 42°C. The freezing-thawing step was repeated five times in order to determine the lysis of the cell wall. The resulting lysed yeasts were resuspended in 50 µl of Laemmli Buffer and denatured at 95°C for 5 minutes.

Proteins were separated on 10% SDS polyacrylamide gels and then transferred by electroblotting to a polyvinylidene fluoride (PVDF) membrane (Roche). Subsequently, membranes were incubated with the blocking solution (4% milk in PBS) for 1 hour at room temperature, in agitation. Membranes were incubated overnight at 4°C in agitation with the respective antibodies, namely the rabbit anti-GAL4 (DNA binding domain) antibody (06-262; Merck Millipore, Billerica, MA, USA) and the rabbit anti-Gal4 (Activation Domain) antibody (G9293; Sigma-Aldrich), both diluted in 1:1000 in 4% milk. After four washing steps in PBS, membranes were incubated at room temperature for 90 minutes in agitation with the HRP-conjugated secondary antibody goat anti-rabbit IgG (sc-2004; Santa Cruz Biotechnology) diluted in 4% milk. After four more washes in PBS, bound antibodies were detected using enhanced chemiluminescence (Westar ηC; Cyangene). Blot images were acquired with GboxChemi XT4 system (Syngene).

Co-IMMUNOPRECIPITATION

Co-immunoprecipitation was performed to confirm the protein-protein interactions observed with the yeast two hybrid assay. 70 millions of HeLa cells were used for the analysis. Cells were harvested and washed with 1 ml of PBS. The lysis was performed with 500 µl of RIPA buffer (50 mM HEPES, 1 mM EDTA, 1% NP40, 0,5 M LiCl) and 2 µl of protein inhibitor cocktail (Roche).

Briefly, 100 µl of beads (Affi-Prep® Protein A Matrix, Bio-Rad, Hercules, CA, USA) were transferred to a 1,5 ml Eppendorf tube and pelleted at 2500 rpm for 2 minutes at 4°C. The pellet was then washed twice with 1 ml of PBS/BSA (5 mg/ml) and resuspended again in 1 ml of PBS/BSA by gently inverting. 10 µl of the anti-NIPBL antibody (A301-779A, Bethyl Laboratories Inc, Montgomery, TX, USA) or anti-SMC3 antibody (A300-060A, Bethyl

Laboratories Inc) were added to the resuspension, which was then incubated for 5-6 hours in agitation at 4°C in order to immobilize the antibody on the beads. After the incubation, the complexes beads-antibody were centrifuged at 2500 rpm for 2 minutes at 4°C and resuspended in the HeLa cell lysate. A overnight incubation in agitation at 4°C was performed to let the protein of interest and its interactors bind the antibody immobilized on the beads. After centrifugation, the supernatant (unbound fraction) was transferred to a new tube, while the beads pellet was washed twice with 1 ml of RIPA buffer. The pellet was then resuspended in 50 µl of Laemmli buffer and denatured at 95°C for 5 minutes (Immuno-precipitated fraction). 100 µl of unbound fraction were mixed with 100 µl of Laemmli buffer and denatured at 95°C for 5 minutes. Immunoblotting was performed on both fractions using the following antibodies for the detection of the proteins: rabbit anti-NIPBL antibody (A302-778A, Bethyl Laboratories Inc), mouse anti-SMC3 antibody (A300-060A, Bethyl Laboratories Inc) and mouse anti-SNF5 antibody (ab58209, Abcam).

RESULTS

SEQUENCING ANALYSIS OF THE KNOWN CdLS-GENES

Sequencing analysis was carried out on an internationally assembled cohort consisting of more than 200 patients with a diagnosis of CdLS or with overlapping phenotypes resembling features characteristic for CdLS.

Sequencing analyses were performed by Sanger sequencing and targeted gene panel in order to identify the causative mutation underlying the syndrome. When possible, the analysis was performed on different tissues to allow the detection of mosaic mutations.

NIPBL

By our investigations we could identify 109 distinct mutations in the *NIPBL* gene in 131 individuals. The identified mutations comprise 37 missense substitutions (34%), 28 out-of-frame deletions (25%), 12 nonsense mutations (11%), 25 splicing variants (23%), 3 in frame deletions/insertions (3%) and 4 deletions of one or multiple exons (4%). The alignment of the NIPBL protein sequence of six different organisms showed a high level of conservation of the mutant residues during evolution. Recurrence was not observed for any of the identified variants.

Mutations localize throughout the entire gene without mutational hotspots. Among the 109 mutations, 26 reside in exons coding functional domains. The most affected domains are the large caldesmon domain (11 mutations) and the HEAT repeats (9 mutations). In addition, sequence variants were also observed in the QR-rich region (1 mutation), in the nuclear export signal (1 mutation), in the DNA-binding domain (1 mutation) and in a region enriched in prolin, glutamic acid, serine and threonine (PEST, 3 mutations).

Somatic mosaicism was identified in nine out of 131 patients (7%). The mosaic variants include three missense substitutions, two out-of-frame deletions, two splicing and two nonsense mutations. Patients carrying these variants present with variable phenotypes.

Out of the nine mosaic mutations, three appear of particular interest because they were not detectable by Sanger sequencing neither using DNA extracted from blood nor using DNA extracted from buccal mucosa. Solely through high-coverage targeted next generation sequencing we were able to detect a very low amount of mutant reads on buccal mucosa DNA

(published as Braunholz et al., 2015). All mutations were confirmed by SNaPshot analysis and Sanger sequencing of DNA extracted from fibroblasts.

The three patients carrying the low grade mosaic variants are all characterized by intellectual disability, a low weight and microcephaly (Fig.15). Patient A carries the nonsense mutation c.4094T>A, p.L1365* affecting exon 18 and represented in 17% of 64 reads. This patient, a 2-year-old girl, displayed typical CdLS facial features including synophrys, long eyelashes, long philtrum and thin upper lip. In addition, she presented with strabismus, feeding problems, GERD, restriction of elbow movements and transverse palmar crease. Patient B is a 19-year-old girl harboring a missense substitution in exon 23 (c.4751T>G, p.L1584R; 13% of 93 reads). She suffered from GERD and displayed a very characteristic CdLS *facies* with the following features: short neck, brachycephaly, low anterior hairline, depressed nasal bridge, long and smooth philtrum and thin lips. Lastly, patient C, a 6-year-old girl, carries a variant in the splice site of intron 27 that determines the skipping of the entire exon 27, thus resulting in a frame-shift deletion of 103 bp (c.5328+1G>C, p.M1743Sfs*35; 21% of 267 reads). This patient presented with mild facial features and other major complications including a severe motor and verbal developmental delay and a brain tumor detected at the age of nine months. The presence of hemihypoplasia was the only feature suggestive for a mosaic mutation.

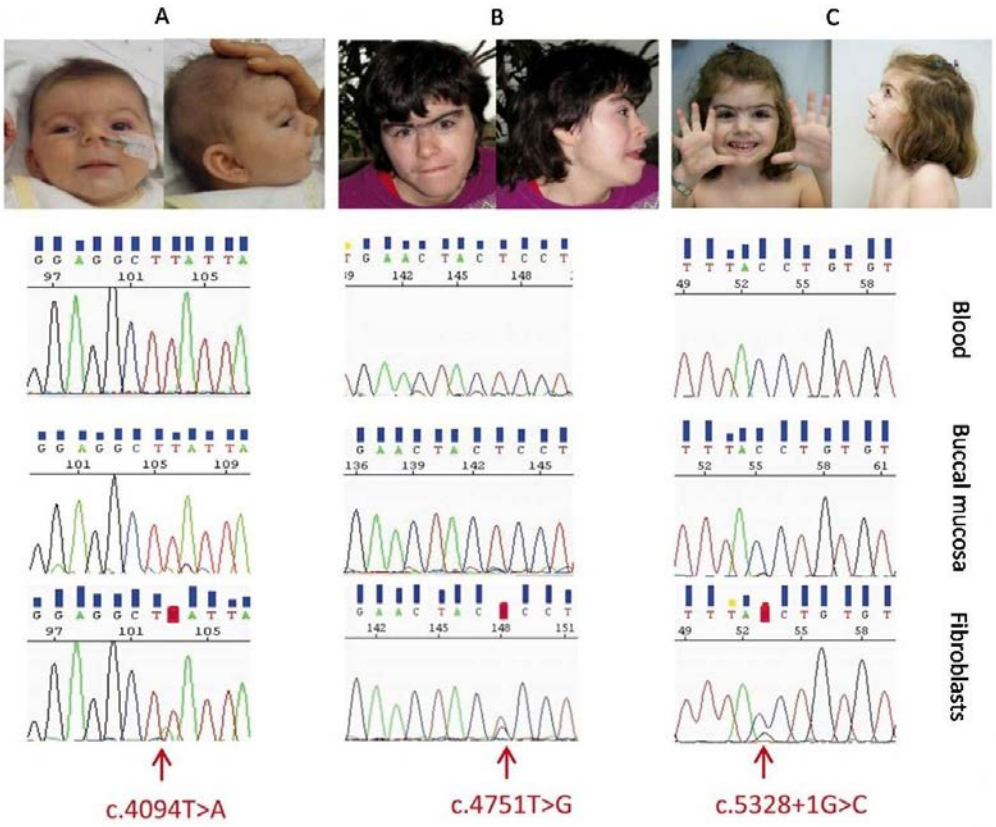


Figure 15. Photographs and electropherograms of the three CdLS patients carrying mosaic mutations in NIPBL that were detectable by Sanger sequencing only on DNA extracted from fibroblasts (from Braunholz et al., 2015)

SMC1A

(Published as Gervasini et al., 2013)

Mutations in *SMC1A* were identified in eight patients with different ethnic origins including Europe, North Africa and South America. The eight patients comprise one male and seven females with ages ranging from 3 to 15 years (Fig.16).

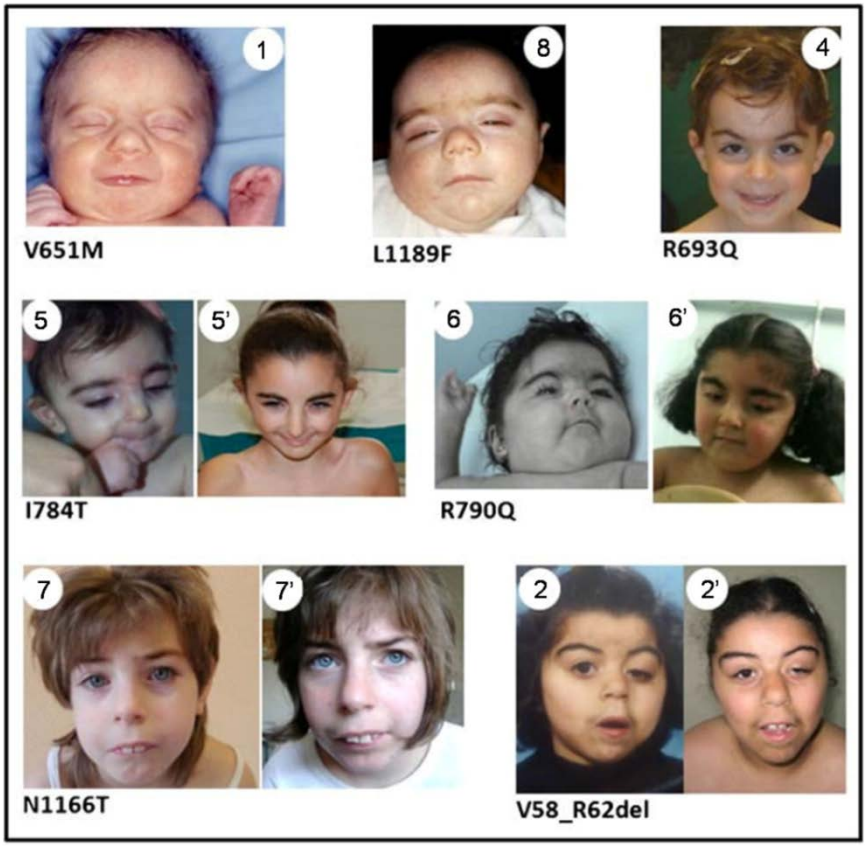


Figure 16. Facial appearance of seven out of the eight CdLS patients harboring mutations in the SMC1A gene. For four patients, pictures were taken at two different ages. For each patient, the respective mutation and the identification number are indicated below the photographs (from Gervasini et al., 2013)

All patients displayed the characteristic facial features of CdLS, including thick and arched eyebrows, long eyelashes, anteverted nares, a long philtrum and thin lips with downturned commissures. Their clinical presentation ranges from mild to severe, with patients 1 and 8 being the most severely affected. Intra-uterine growth retardation was reported for all patients, while postnatal growth appears to be less affected. No major limb malformations were observed. The reported clinical findings included GERD, feeding disorder, pulmonary stenosis, mild brain anomalies, supernumerary ribs, aortic coarctation, epilepsy and sensorineural hypoacusis. A severe developmental delay is the most striking feature of this cohort: sitting independently and walking independently were achieved at a median age of 1 year 4 months and 2 years 8 months, respectively; patient 1 was still not able to walk autonomously at the age

of 6 years. A moderate to severe cognitive impairment was observed in the entire cohort. Speech acquisition appears to be even more delayed, with patients 1, 5 and 6 being still unable to speak at the age of 6, 11 and 4 years, respectively. The remaining five patients were able to pronounce the first words at a median age of 3 years 6 months. The association of more than two words occurred at a median age of 4 years 6 months.

The identified mutations comprise six missense substitution and one in frame-deletion, namely p.V58_R62del, p.R398G, p.V651M, p.R693Q, p.I784T, p.T790Q, p.N1166T and p.L1189F. The substitutions p.R398G and p.R693Q affect residues that were previously described in the literature, though with a different amino acid exchange (Liu et al., 2009). All variants were proven to arise *de novo* in the patients and none of them was observed in more than 400 X chromosomes of healthy individuals.

The relative positions of each mutation according the localization of the structural domains and on the SMC1A/SMC3 heterodimer are depicted in Fig.17a and Fig.17b, respectively. The mutations localize throughout the entire protein with the exception of the hinge domain. It is worth to be noted that the substitutions p.N1166T and p.L1189F affect the C-terminal ATPase cassette, and p.L1189F is the first mutation residing in the functional motif H-loop. All variants were predicted to be disease causing by different *in silico* bioinformatic tools (Polyphen-2, SIFT and Mutation Taster). In addition, the alignment of multiple regions of the SMC1A protein across nine species highlighted a high level of conservation of the affected residues and of the surrounding amino acids.

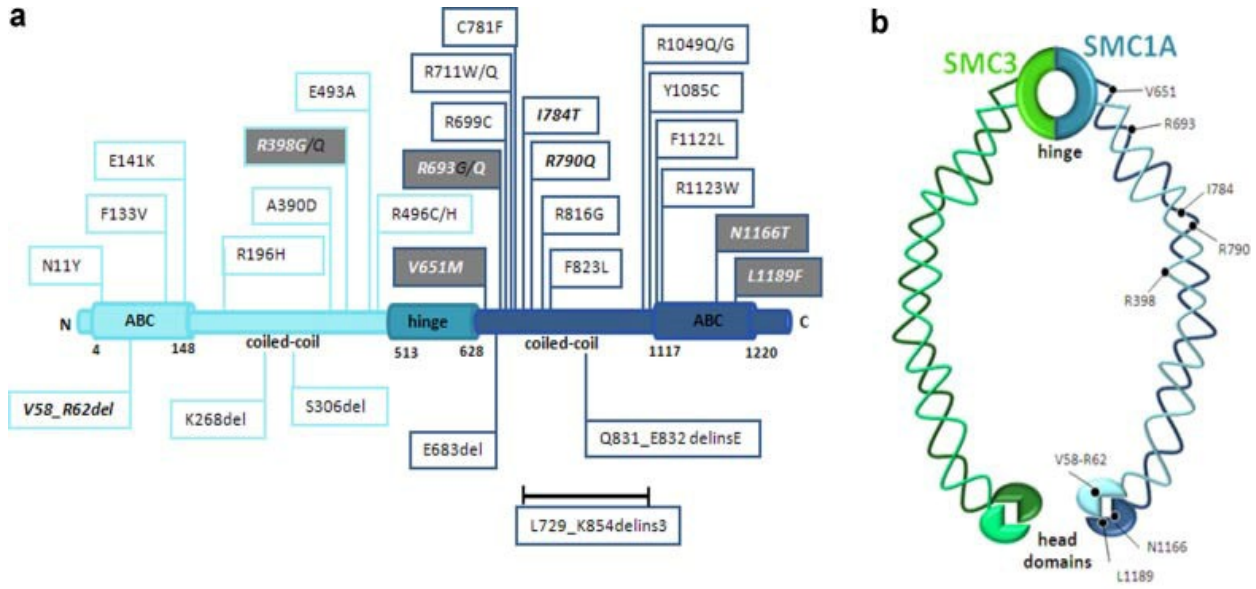


Figure 17. (a) Schematic representation of the SMC1A protein and of its domains. The position of novel and reported mutations are indicated. (b) Schematic representation of the core heterodimer of the cohesin complex with relative position of the described mutations (from Gervasini et al., 2013)

SMC3

(Published as Gil-Rodrìguez et al., 2015)

Figure 18. Phenotype of the patients harboring mutations in the SMC3 gene so far described in the literature, including the one reported in Deardorff et al., 2007 (P6). For each individual, photographs show a frontal view, a lateral view, hands and feet. The identification number and the relative mutation are also indicated above each picture (from Gil-Rodrìguez et al., 2015)

Sequencing analysis of patients with CdLS or CdLS-overlapping features led to the identification of 15 additional mutations in *SMC3* in 15 unrelated individuals.

The phenotypical appearance of these patients is shown in Fig.18. The most characteristic clinical feature in this group is a variable level of cognitive impairment. Considerable behavioral problems were not observed, and many patients of this cohort were described as having friendly personalities. Major limb malformations were also not reported. Typical CdLS facial features were observed with the following frequency: low anterior hairline (50%), brachycephaly (73%), synophrys (73%), arched eyebrows (93%), ptosis (27%), long eyelashes (94%), depressed nasal bridge (47%), anteverted nares (57%), long philtrum (67%), thin upper lip (81%), downturned commissures (60%), dental anomalies (38%), high palate (45%) and micrognathia (40%). Interestingly, despite often long, the philtrum of these individuals was typically not smooth, and only one patient presented with cleft palate.

The *SMC3* mutations so far identified are listed in Table 3 and comprise nine missense substitutions, one nonsense mutation and five in-frame deletions, one of which (p.E488del) observed in two unrelated patients. The SMC3 protein sequence of nine different organisms was aligned in order to verify the degree of conservation of the mutant residues. This analysis showed that the conservation can be traced back to yeast for all the amino acids of interest. In addition, all the variants were defined as disease causing by three different *in silico* bioinformatic tools (Polyhen-2, SIFT and Mutation Taster).

ID	DNA variant	Protein change	De novo	Exon	Protein domain
P1	c.193T>C	p.F47L	na	4	Head
P2	c.703_705del (mosaic)	p.T235del	+	9	Coiled-coil
P3	c.707G>C	p.R236P	+	9	Coiled-coil
P4	c.859_861dup	p.E287dup	na	11	Coiled-coil
P5	c.1200_1202delGTC	p.K400_S401delinsN	na	13	Coiled-coil
P6	c.1464_1466delAGA	p.E488del	+	15	Coiled-coil
P7	c.1464_1466delAGA	p.E488del	+	15	Coiled-coil
P8	c.1462G>A	p.E488K	+	15	Coiled-coil
P9	c.1561C>T	p.R521*	na	16	Hinge
P10	c.1964G>A	p.G655D	+	19	Hinge
P11	c.1997G>C	p.G666A	+	19	Hinge
P12	c.2494_2499del	p.L832_N833del	+	22	Coiled-coil
P13	c.2515C>T	p.R839C	na	22	Coiled-coil
P14	c.2750A>C	p.H917P	+	24	Coiled-coil
P15	c.3439C>G	p.Q1147E	+	27	Head
P16	c.3644C>T	p.T1215I	na	29	Head

Table 3. List of all mutations so far identified in SMC3, including the one reported in Deardorff et al., 2007 (P6). For each mutation, the DNA and protein change are indicated. The localization of the mutations with regards to the exons and to the functional/structural domains of the protein are also designated. (na=not assessed)

Mutations are scattered across the entire protein without apparent hotspots (Fig.19). Interestingly, three mutations affect the highly conserved hinge domain, responsible for the interaction with the specular domain of the SMC1A protein. Some of the affected residues appear to play an important role in the tertiary structure of the protein. In point of fact, the substitution of G655D structurally destabilize the core of the hinge domain. In addition, the deletions of T235, R236, R839 and H917, all residing in the coiled-coil domain, are responsible for the displacement of the two antiparallel helices.

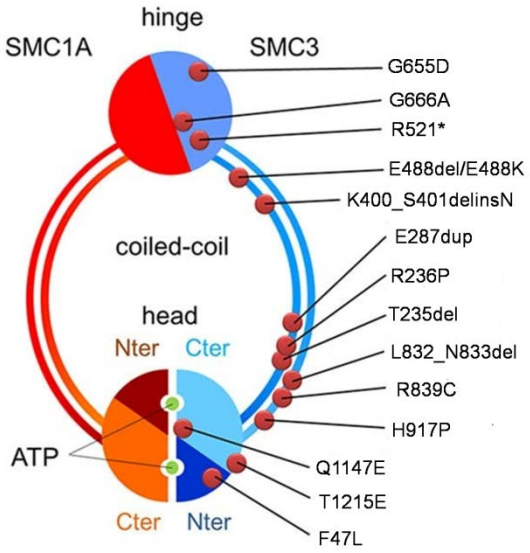


Figure 19. Schematic representation of the core heterodimer of the cohesin complex with relative position of the described SMC3 mutations according to the localization of the functional domains of the protein (from Gil-Rodrìguez et al., 2015)

Besides, Q1147 localizes in the functional D-loop motif, close to interaction surface between SMC1A and SMC3 and to the gamma-phosphate of ATP. Therefore, the substitution of this polar residue with the negatively charged glutamate amino acid might impair the interaction between SMC1A and SMC3 as well as the ATPase activity (Fig.20).

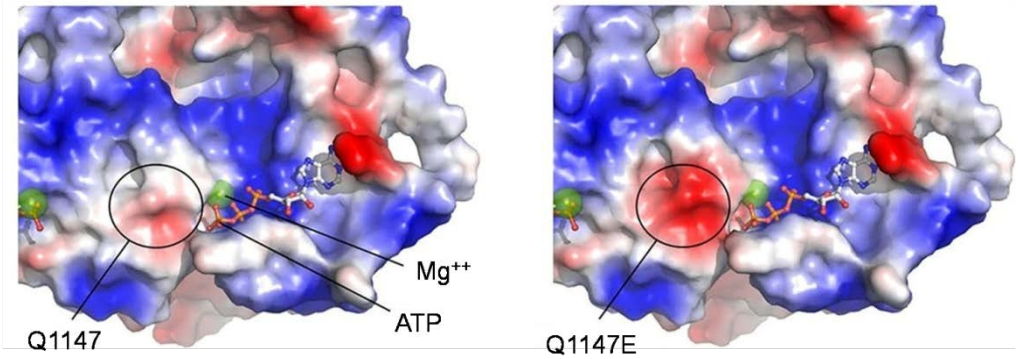


Figure 20. Representation of the predicted tridimensional structure of the SMC3 protein in proximity of the ATPase active centre in the presence of a wild type (left) or mutant (right) residue at position 1147. The interaction surface between SMC3 and SMC1A has been colored with regard to the electrostatic characteristics (blue=positive; red=negative; white=neutral). The negatively charged patch that appears in the presence of the mutation close to the gamma phosphate of ATP is highlighted (from Gil-Rodrìguez et al., 2015)

HDAC8

Ten distinct mutations in the *HDAC8* gene were identified in 11 patients, namely eight females and three males (Fig.21). Patients presented with a very broad range of phenotypical features that are summarized in Table 4 and include microcephaly (90%), brachycephaly (36%), late closure of the anterior fontanelle (36%), low anterior hairline (45%), short neck (90%), synophrys (90%), arched eyebrows (90%), long eyelashes (45%), hypertelorism (90%), hooding of the eyelids (45%), myopia (54%), astigmatism (45%), strabismus (36%), depressed nasal bridge (45%), anteverted nares (54%), broad/bulbous nasal tip (81%), long philtrum (36%), smooth philtrum (54%), thin upper lip (72%), downturned commissures (54%), micro/retrognathia (90%), widely spaced teeth (45%), highly arched palate (45%), small hands (72%), small feet (36%), clinodactyly (45%), proximally set thumbs (63%), short fifth finger (54%), hirsutism (63%), pigmentation defects of the skin (36%), gastroesophageal reflux (36%), feeding problems in infancy (45%), cryptorchidism (100% of males) and hearing problems (63%). Growth retardation, intellectual disability and motor and verbal developmental delay were reported in all patients.

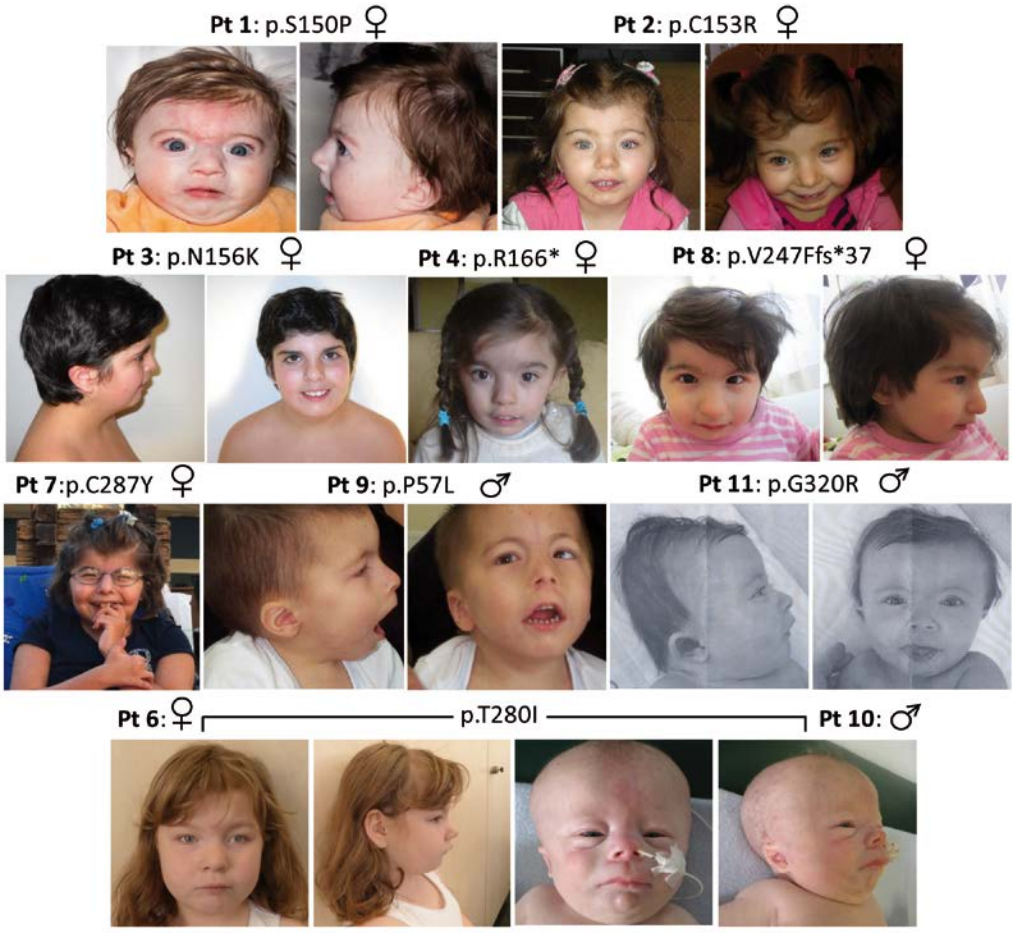


Figure 21. Facial features of the patients harboring mutations in the HDAC8 gene. For each patient, the mutation and the gender are indicated

		Females									Males				General frequency
Patient number		1	2	3	4	5	6	7	8	Relative frequency	9	10	11	Relative frequency	
X-inactivation		99/1	100/0	17/83	100/0	100/0	100/0	100/0	100/0		-	-	-		
Head	Microcephaly	+	+	+	+	+	+	+	-	87%	+	+	+	100%	90%
	Brachycephaly	+	+	+	-	-	-	-	-	37%	-	-	+	33%	36%
	Late closure of the anterior fontanelle	+	+	na	-	-	-	-	+	37%	-	+	-	33%	36%
Eyes	Arched eyebrows	+	+	+	+	-	+	+	+	87%	+	+	+	100%	90%
	Synophrys	+	+	+	+	+	-	+	+	87%	+	+	+	100%	90%
	Hypertelorism	+	+	+	+	+	+	+	+	100%	-	+	+	67%	90%
	Telecanthus	-	+	-	+	-	-	-	-	25%	-	-	-	0%	18%
	Myopia/Astigmatism	+	+	+	-	+	-	+	+	75%	+	-	-	33%	63%
	Strabismus	-	-	-	-	+	-	+	+	37%	+	-	-	33%	36%
	Hooding of lids	-	-	-	-	-	-	+	+	25%	+	+	-	67%	36%
	Lacrimal duct obstruction	-	-	-	-	-	na	-	-	0%	-	na	+	33%	9%
Nose	Depressed nasal bridge	+	+	-	+	-	-	-	-	37%	-	+	+	67%	45%
	Anteverted nares	+	+	-	+	-	+	-	-	50%	-	+	+	67%	54%
	Long philtrum	+	-	-	+	-	-	-	-	25%	-	+	+	67%	36%
	Smooth philtrum	+	+	-	-	-	-	+	+	50%	+	-	+	67%	54%
	Broad nasal tip	+	+	+	+	+	+	+	+	100%	+	-	-	33%	81%
Mouth	Thin lips	+	+	-	-	-	+	+	+	62%	+	+	+	100%	72%
	Downturned commissures	+	+	-	+	+	-	-	-	50%	-	+	+	67%	54%
	Micro/Retrognathia	+	+	+	+	+	+	-	+	87%	+	+	+	100%	90%
Ears	Dysplastic ears	+	-	-	+	-	-	-	+	37%	-	-	-	0%	27%
	Hearing problems	+	-	-	+	+	-	+	+	62%	+	+	-	67%	63%
Hands	Small hands	-	+	+	+	+	-	+	-	62%	+	+	+	100%	72%
	Proximally set thumbs	+	+	+	-	+	-	+	-	62%	+	-	+	67%	63%
	Clinodactyly 5th finger	-	+	+	+	-	-	-	-	37%	+	+	-	67%	45%
Feet	Small feet	-	+	+	-	na	-	na	-	25%	na	+	+	67%	36%
	Syndactyly of toes	-	-	+	-	-	-	-	-	12%	-	-	-	0%	9%
Other	Cardiac defects	+	-	-	+	+	+	-	-	50%	-	+	+	67%	54%
	Genitourinary defects	-	+	+	-	+	-	-	-	37%	+	+	+	100%	54%
	GER	-	-	+	-	+	-	+	-	37%	+	-	-	33%	36%
	Feeding problems	+	+	+	-	-	-	+	-	50%	-	+	-	33%	45%
	Seizures	-	-	+	-	-	-	-	-	12%	-	+	+	67%	27%
Development	Growth retardation	+	+	+	+	+	+	+	+	100%	+	+	+	100%	100%
	Intellectual disability	+	+	+	+	+	+	+	+	100%	+	+	+	100%	100%
	Psychomotor delay	+	+	+	+	+	+	+	+	100%	+	+	+	100%	100%
	Behavioral problems	na	+	+	-	+	-	-	-	37%	na	na	-	0%	27%

Table 4. Clinical features of the eleven patients with mutations in the HDAC8 gene

The three male patients, who are hemizygous for the mutation, displayed very severe multi-organ complications including cryptorchidism, vesicoureteral reflux, pulmonary stenosis, hypoacusis, inguinal hernia, corpus callosum agenesis and cerebral atrophy. In addition, they all showed marked growth retardation and prominent motor and verbal developmental delay. Patient 8 even died of pneumonia at the age of two years. Conversely, the eight female patients presented with a broad range of clinical features and a variable degree of severity of the phenotype. Accordingly, patient 2 presented with the mildest clinical presentation, while patient 7 showed the most severe features among females. The observed broad range of phenotypes could be partially explained by the X-inactivation status: in point of fact, all females except patient 3 were found to carry a completely skewed X-inactivation towards the wild type allele on blood DNA.

The identified variants include one nonsense mutation (p.R166*), one out-of-frame deletion (p.F207Nfs*3), one splicing mutation (c.910+1G>A; p.V247Ffs*37) and seven missense substitutions (p.S150P, p.C153R, p.N156K, p.T280I, p.P257L, p.C287Y, and p.G320R) (Fig.22). Eight mutations are novel, while two affect amino acids (C153 and G320) that were

already found to be mutated in the literature (Deardorff et al., 2012b; Kaiser et al., 2014). None of these variants was identified in over a hundred of healthy X chromosomes. All mutations were proven to affect highly conserved residues and were predicted to be pathogenic by different *in silico* bioinformatic tools (PolyPhen-2, SIFT and Mutation Taster).

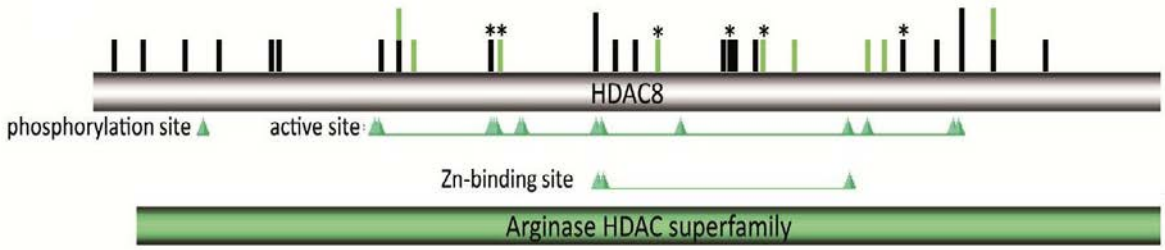


Figure 22. Schematic representation of HDAC8 and of its functional domains, with relative position of already described (black bars) and novel (green bars) mutations. Stars indicate out-of-frame deletions or nonsense mutations, that lead to an early truncation of the protein

Interestingly, the p.T280I missense exchange is shared by two siblings, a boy and a girl (patients 6 and 10). Next generation sequencing approaches highlighted the presence of the same mutation in a low number of reads on the DNA of the mother (18 of 200 reads). Hence, pyrosequencing was performed in order to investigate the suspected mosaicism. The analysis was carried out on blood DNA of the affected daughter (patient 6) as a control and on DNA extracted from three different tissues of the mother, namely blood, buccal mucosa and urine (Fig.23). This analysis pointed out the presence of tissue-dependent level of heterozygosity in the mother: while the daughter displayed an equal amount of wild type and mutant alleles, the mutation in the mother was totally absent on buccal mucosa, barely detectable on blood (2%) and was represented at a rate of 19% on DNA extracted from urine. In addition, X-inactivation analysis on blood DNA showed a complete skewing in the daughter (100/0) and a random selection of the inactive X-chromosome in the mother (40/60).

The mother, who presents with short stature and appears cognitively normal, displayed facial features reminiscent of CdLS during childhood; anyhow, these features were not discernible anymore during adulthood. Both children showed instead CdLS features including intellectual disability, growth retardation, patent foramen ovale, microcephaly, hypertelorism, arched eyebrows, anteverted nares, micro/retrognathia and thin lips. In addition, the male sibling (patient 10) presented with short neck, hooding of the lids, synophrys, ptosis, a depressed nasal bridge, long philtrum, downturned commissures, small hands and feet, clinodactyly of the fifth finger, cystic renal degeneration, pulmonary stenosis, hypospadias, cryptorchidism,

sensorineural hearing loss, feeding problems and seizures. He died at the age of two years because of pneumonia.

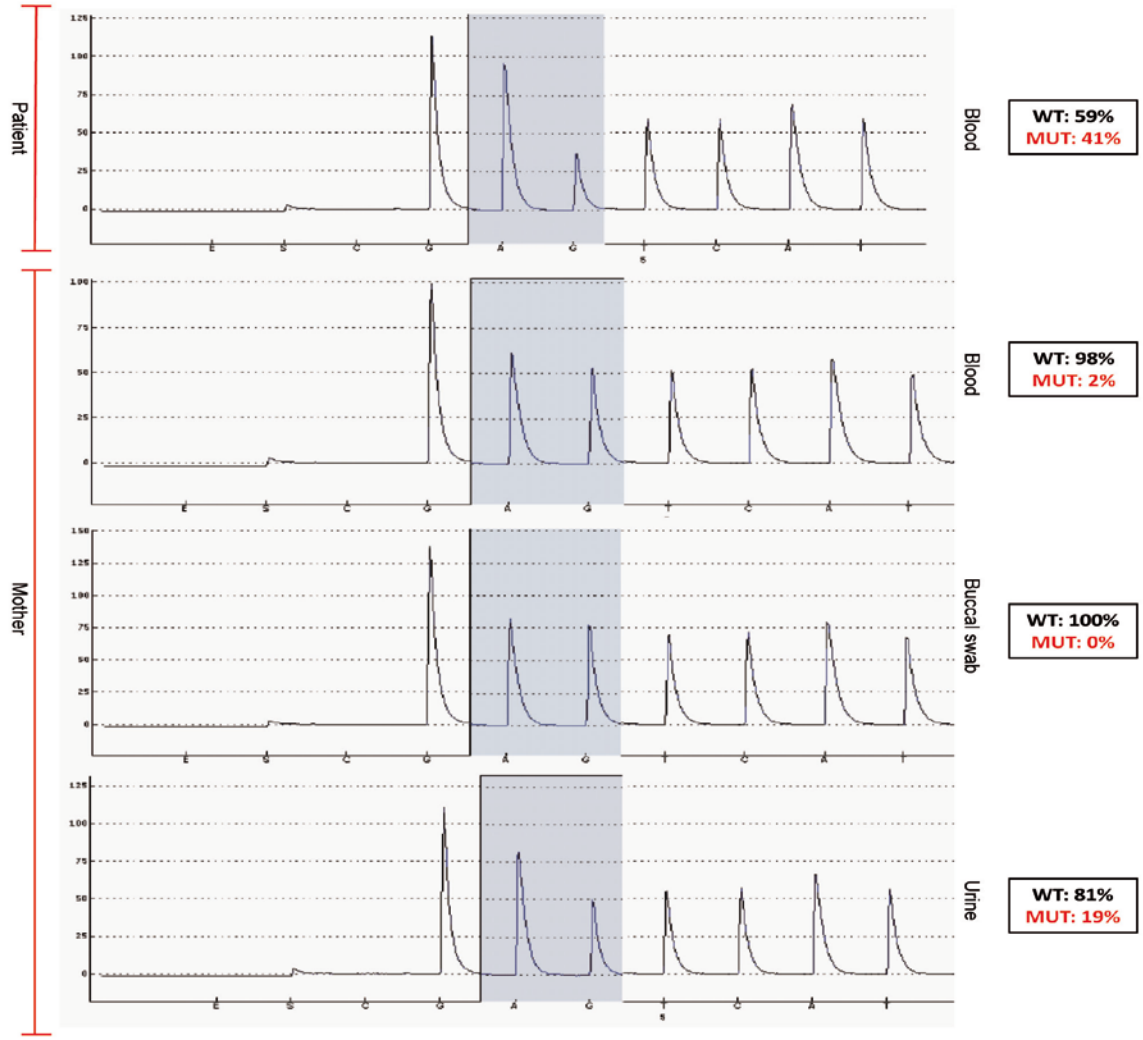


Figure 23. Pyrosequencing analysis of the mother of the siblings harboring the T280I variant on different tissues. The mutation is not detectable on buccal swab DNA (0%), barely detectable on blood DNA (2%) and detectable at a rate of 19% on DNA extracted from urine

OVERALL AND ALLELE-SPECIFIC EXPRESSION OF THE *SMC1A* GENE

(Published as Parenti et al., 2014)

SMC1A maps at Xp11.22, a region of the X chromosome that partially escapes X-inactivation in humans. Genes that localize in this region are characterized by a variable level of expression among females and between females and males. In particular, females were reported to express *SMC1A* at a higher level than males, but the ratio between the two genders vary among different studies (Craig et al., 2004; Johnston et al., 2008; Liu et al., 2009). Therefore, we established a Real Time PCR assay in order to assess the expression level of *SMC1A* in 17 healthy controls, namely ten females and seven males. In this way we were able to prove that *SMC1A* expression in females is approximately 50% higher than in males (***P=0,0001; one tail Mann-Whitney U test) (Fig.24a). Subsequently, the same investigation was carried out at a protein level by mean of semi-quantitative western blot on protein lysates of five LCLs from females and four LCLs from males. Our results indicate that SMC1A expression in females is approximately 44% higher than in males, as shown in Fig.24b and 24c. (**P=0,0079; one tail Mann-Whitney U test). Hence, the examination of both *SMC1A* transcript and protein concordantly showed a higher expression in females as compared to males.

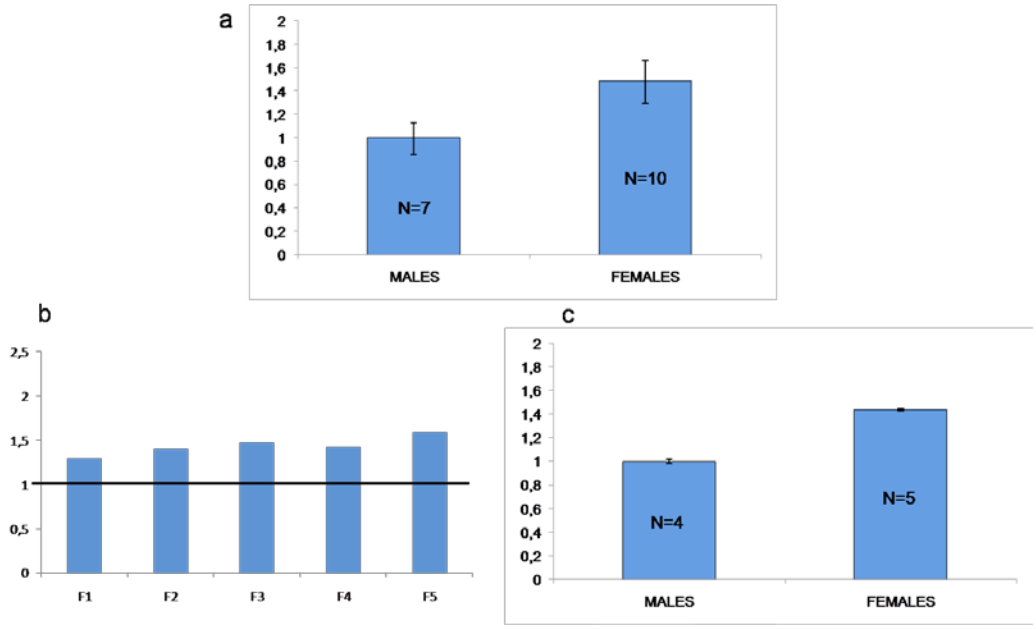


Figure 24. SMC1A expression in the healthy population. (a) Graphical representation of the transcript analysis in seven males and ten females, indicating a 50% higher expression of the gene in females as compared to males. (b-c) Graphical representation of the protein analysis in four males and five females (b) The histograms indicate the relative expression of SMC1A in the female controls in comparison with the average level of expression in the males (set as 1 and represented by the horizontal black line) (c) Average expression of the protein in the two groups of controls. Females are shown to express SMC1A 44% more than males (from Parenti et al., 2014)

We then investigated the SMC1A protein in the CdLS population in order to assess whether the presence of mutations could impair the ability of the cell to synthesize the protein or the stability of the protein itself. The analysis was performed on five out of the seven CdLS female patients identified through the mutational analysis and on two previously described patients (Musio et al., 2006; Limongelli et al., 2010). The seven patients, herein indicated as patient 1-7, respectively carry the following mutations in *SMC1A*: p.Q831_D832delinsE, p.L1189F, p.N1166T, p.I784T (shared by patient 4 and 5), p.V58_R62del and p.R693Q. Semi-quantitative western blot analyses (Fig.25a) could show that all patients express SMC1A at comparable levels with few variability among samples (Fig.25b). No differences among the distinct mutations were observed. Subsequently, we compared the SMC1A expression of the CdLS population with the one of the five female controls. This analysis indicated that there are no significant differences in the protein level between the two groups (P=0,5; one tail Mann-Whitney U test) (Fig.25c).

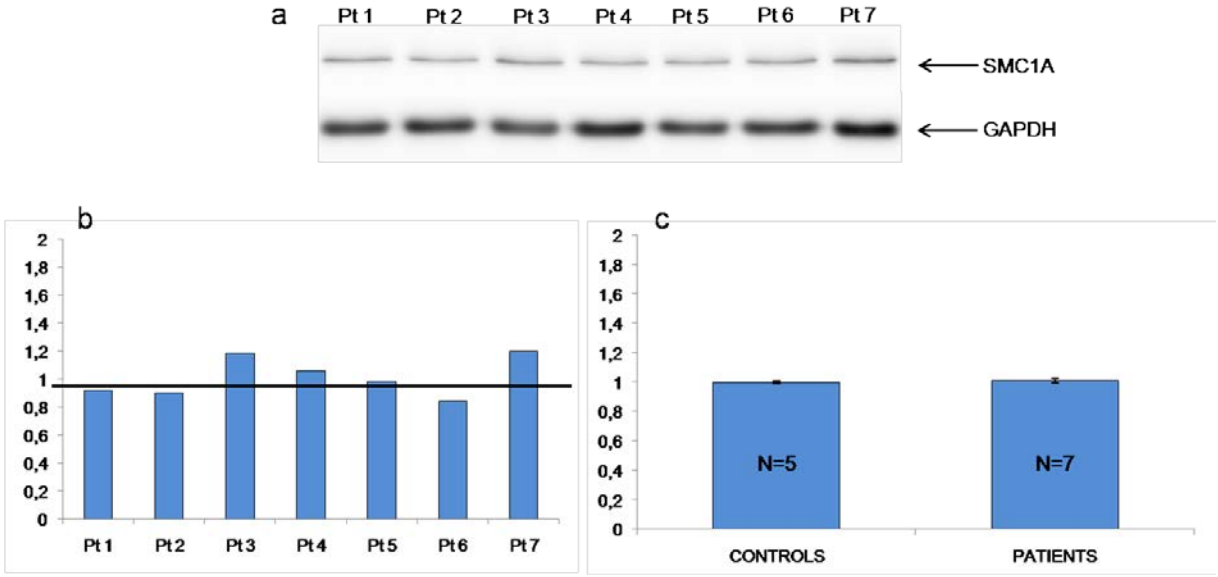


Figure 25. Western blot analysis of the SMC1A protein in a group of CdLS female patients (a) Representative blot showing the SMC1A protein (150 kDa) and the normalizer GAPDH (36 kDa) (b) The histograms indicate the relative expression of the protein in the CdLS patients in comparison with the average level of expression of the five female controls (set as 1 and represented by the horizontal black line) (c) Average expression of SMC1A in controls and patients. No significant difference is observed (from Parenti et al., 2014)

Lastly, we investigated the expression of the wild type and mutant alleles in the seven female patients with mutations in SMC1A in order to identify a possible differential expression of the two alleles. The X-inactivation status was first checked for each patient at the HUMARA or DSX6673E loci. A completely skewed X-inactivation was identified in patients 3, 4, 5 and 6. The remaining patients presented with a borderline or random X-inactivation. The allele-

specific expression analysis was carried out by means of pyrosequencing on 15 female controls and on six out of seven CdLS patients (patient 6, carrying the p.V58_R62del, was excluded from this analysis due to the inability to establish an effective pyrosequencing assay). The SNP rs1264011 was used to discriminate the two alleles of the female controls; this SNP is characterized by a high level of heterozigosity in different populations and is localized in the 5'UTR of the gene. Patients were analyzed with assays specific for each mutation. When possible, the analysis was carried out on RNA extracted from both blood and LCLs in order to verify whether the allele expression ratios might be affected by the type of starting material. In this context, no differences in the wild type/mutant ratio were observed between the two pools of RNA (Fig.26a). The pyrosequencing assays showed that the expression ratio of the two alleles of the SNP in the healthy population was 1:1. Conversely, the average ratio between wild type and mutant allele in the CdLS population was 2:1 (Fig.26b). In point of fact, in all patients, more than 60% of the total transcript carried the wild type allele, and the greatest difference between the wild type and mutant allele was observed in patient 1, carrying the in-frame deletion p.Q831_D832delinsE, who expressed 73% of wild type allele and 27% of mutant allele.

a

	Allele 1 / Allele 2 PB	Allele 1 / Allele 2 LCL
Pt 1		73/27
Pt 2		71/29
Pt 3	66/34	66/34
Pt 4	69/31	62/38
Pt 5	66/34	65/35
Pt 7	65/35	61/39

b

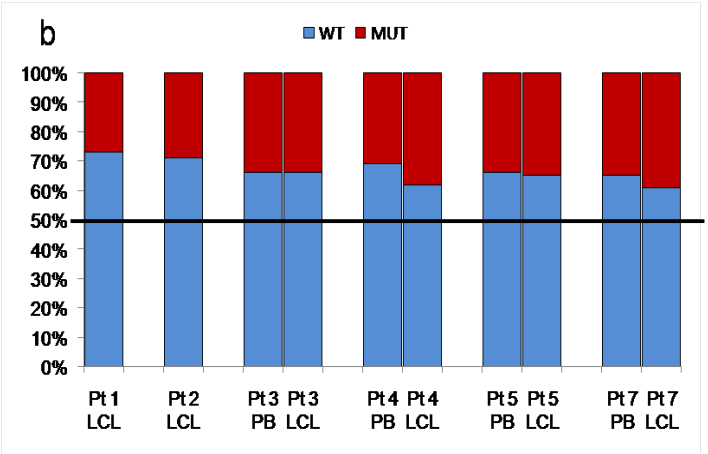


Figure 26. Results of the pyrosequencing assay (a) Level of expression of the wild type (allele 1) and mutant (allele 2) alleles of the SMC1A gene in six CdLS patients. Results from both peripheral blood (PB) and LCLs are shown. (b) Graphical representation of the pyrosequencing results. The level of expression of the wild type allele is indicated in blue, while the mutant allele is indicated in red. The 1:1 ratio between the two alleles observed in the control group is graphically represented as a black line (from Parenti et al., 2014)

IDENTIFICATION OF NEW CAUSATIVE GENES

Mutations in the cohesin-related genes *NIPBL*, *SMC1A*, *SMC3*, *RAD21*, and *HDAC8* are found in about 70% of CdLS patients. Hence, additional causative genes might explain the remaining 30% of cases. In this context, genes functionally related to cohesin or genes responsible for clinical entities partially overlapping with CdLS might be able to cover this missing heritability. For this reason, we analyzed our cohort of patients with CdLS or CdLS-overlapping features by different next generation sequencing techniques including exome sequencing, targeted gene panel and CGH-array.

IDENTIFICATION OF MUTATIONS IN THE *ANKRD11* GENE IN TWO PATIENTS WITH A TENTATIVE CLINICAL DIAGNOSIS OF CdLS

(Published as Parenti et al., 2015)

Exome sequencing led us to the identification of mutations in the *ANKRD11* gene in two patients with a tentative diagnosis of CdLS. As previously mentioned, mutations in *ANKRD11* are responsible for KBG syndrome, a rare genetic disorder characterized by growth retardation and skeletal anomalies, cognitive impairment, developmental delay and characteristic facial features partially overlapping with CdLS, in particular concerning the ocular and peri-nasal regions (Ockeloen et al., 2014).

Patient 1 (Fig.27a-h) is a German female born to healthy and non-consanguineous parents at 37 weeks of gestation. The Apgar scores at 1 and 5 minutes were 9 and 10, respectively. Birth weight was 2350 g (3^{rd} centile), birth length was 46 cm (3^{rd} centile) and the Occipital Frontal Circumference (OFC) was 31.5 cm (3^{rd} centile). The patient presented with general muscular hypertonia and persistent feeding difficulties. At the age of four months, weight was 5600 g (10^{th} centile), height was 59 cm (10^{th} centile) and the OFC was 38.3 cm ($<3^{rd}$ centile). At this age some dysmorphic features typical of CdLS were observed, including microcephaly, brachycephaly, short neck, open anterior fontanelle, broad eyebrows with synophrys, downslanting palpebral fissures, depressed nasal bridge, anteverted nares, long philtrum, thin upper and lower lips with downturned commissures and micrognathia. She also displayed proximally set thumbs and a mild brachydactyly. At the age of the last clinical evaluation (4 years) the patient still needed assisted feeding. She was able to sit independently at the age of 16 months and to walk independently at the age of 20 months, but was still unable to speak,

thus indicating a marked speech delay. She also showed intellectual disability and a self-injurious and aggressive behavior associated with autistic features.

Figure 27. Clinical features of the patients harboring mutations in the ANKRD11 gene. (a-h) Patient 1: (a-c) Facial gestalt at the age of four months and four years, respectively. (d) Teeth conformation at the age of four years, showing small and widely spaced teeth. (e-h) Pictures of the limbs, revealing clinodactyly of the fifth finger, shortened first metacarpal and proximally set thumb. (i-k) Patient 2: (i, j) Facial gestalt at the age of 8 and 15 years, respectively. (k) Left hand of the patient, characterized by proximally set thumbs and clinodactyly of the fifth finger (from Parenti et al., 2015)

Patient 2 (Fig.26i-k) is an Italian male born to healthy and non-consanguineous parents at 39 weeks of gestation. Birth weight was 3150 g (25th centile), lenght was 53 cm (90th centile) and the OFC was 33 cm (3rd centile). He presented with vesicoureteral reflux, kidney dysplasia, GERD and generalized epilepsy. At the age of four years he was able to speak in sentences, walk independently and to manage himself in daily life. He showed a moderate intellectual disability and an aggressive behavior towards himself and towards others. At the age of 15

years, weight was 41.2 kg (3rd centile), height was 161 cm (10th centile) and the OFC was 54 cm (3rd centile). The following facial dysmorphisms were observed: low anterior hairline, synophrys, broad and arched eyebrows, long eyelashes, high nasal bridge with a broad nasal tip, smooth philtrum, thin lips, large mouth, mild micrognathia and macrodontia of the superior central incisors. Lastly, he presented with hirsutism, astigmatism and myopia, clinodactyly of the fifth finger and proximally set thumbs.

Both patients received a clinical diagnosis of CdLS during early childhood and were therefore analyzed for mutations in the five CdLS genes. The high coverage targeted gene panel could not detect any pathogenic variant on DNA extracted from blood, buccal mucosa and fibroblasts. Hence, the two patients were selected for a exome sequencing analysis.

Four *de-novo* variants were identified in patient 1: a missense mutation in *ACAP3* (ArfGAP with coiled-coil, Ankyrin repeat and PH domains 3), a possible splicing variant in *BMPR2* (Bone Morphogenetic Protein Receptor 2), a missense mutation in *ERC2* (ELKS/RAB6-interacting/CAST family member 2) and a nonsense mutation in *ANKRD11*. Different *in silico* bioinformatic tools predicted the variants in the *ACAP3*, *BMPR2* and *ERC2* genes to be not pathogenic. In addition, no functional overlap between these genes and the cohesin complex was identified. Therefore, we focused on the nonsense mutation in *ANKRD11*, which appeared as an ideal candidate because mutations in this gene were recently reported in three patients with clinical features reminiscent of CdLS (Ansari et al., 2014). The mutation in *ANKRD11*, namely c.5483G>T; p.S1828*, was found in 836 of 2664 sequencing reads (31%), hence raising the suspicion of mosaicism.

Four variants were also identified in patient 2. Nevertheless, the missense variants in the *ZW10* (centromere/kinetocore protein zw10 homologue), *SMC1B* (Structural Maintenance of Chromosome protein 1B) and *ASPM* (Abnormal Spindle-like Microcephaly-associated protein) genes were all found to be inherited. Therefore, we focused on a four base-pairs deletion in *ANKRD11*, resulting in a premature stop codon. This deletion, c.2297_2300delAGAA; p.K766Rfs*10, was found in 13 of 60 sequencing reads (22%), thus suggesting a possible mosaic state also for this patient.

Both *ANKRD11* variants were confirmed by Sanger sequencing on DNA extracted from different tissues (Fig.28).

Electropherograms of patient 1 showed an unequal distribution of the wild type and mutant alleles on DNA extracted from blood and fibroblasts, further supporting the presence of

mosaicism. Conversely, no significant differences between the two alleles were observed in three different tissues of patient 2.

Pyrosequencing was performed in order to determine the role exerted by the putative mosaicism in the two patients (Fig.28). This assay could confirm the presence of mosaicism in patient 1, while the mutation in patient 2 was excluded to be mosaic. Particularly, the two alleles appear to be equally represented on all tissues of patient 2 and on fibroblasts DNA of patient 1, while a 70:30 ratio between the wild type and mutant allele was observed on blood DNA of this latter patient.

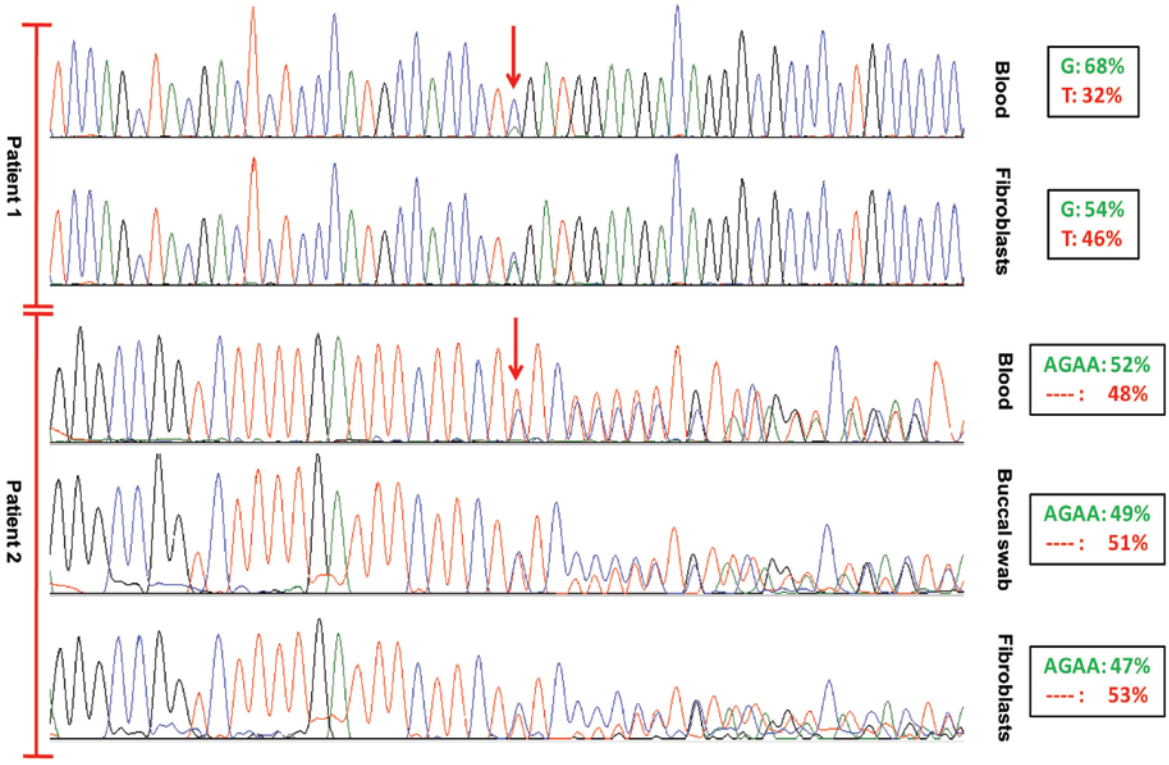


Figure 28. Examination of the putative mosaicism by means of Sanger sequencing and pyrosequencing. Multiple tissues were analyzed for both patients. On the left, electropherograms of patient 1 suggests an unequal distribution of the wild type and mutant alleles, while no differences are observed in the electropherograms of patient 2. On the right, the results of the pyrosequencing assays are showed, with the wild type allele indicated in green and the mutant allele indicated in red (adapted from Parenti et al., 2015)

COHESINOPATHIES ARE BRANCHING OUT: CLINICAL AND MOLECULAR CORRELATION BETWEEN CdLS AND CSS

As previously mentioned, several syndromes represent a differential diagnosis to CdLS. For the establishment of our gene panel we focused on the phenotypical overlap between CdLS and CSS and on the functional overlap between the protein complexes responsible for these clinical entities. Thus, our gene panel comprises the following 16 genes: the five known CdLS genes, five genes functionally related to cohesins (*STAG1*, *STAG2*, *STAG3*, *HDAC1* and *HDAC3*) and six subunits of the SWI/SNF complex including *SMARCA2*, *SMARCA4*, *SMARCB1*, *SMARCE1*, *ARID1A* and *ARID1B*.

Targeted gene panel and exome sequencing led us to the identification of mutations in different subunits of the SWI/SNF complex in four patients with a clinical diagnosis of CdLS.

Patient 1 (Fig.29a,b) is a German female who carries a missense mutation in *SMARCB1*, namely c.998A>G; p.K333R. Patient 2 (Fig.29c), a boy from the United States, harbors the out-of-frame deletion in *ARID1A* c.4682_4779del; p.V1561Afs*9. Patient 3 (Fig.29d) is a German boy carrying a splicing variant in *ARID1B*, namely c.3136-2A>G. Patient 4 (Fig.29e,f), a Spanish boy, was found to carry a gross deletion on chromosome six by means of exome sequencing. The deletion was subsequently confirmed and defined through CGH-array, which showed a loss in copy number at the 6q25.3q26 region spanning ~4.1 Mb[arr 6q25.3q26(157,377,717-161,484,554)x1 hg18]. The deletion comprises 34 genes, including the entire *ARID1B*.

Correspondingly, we were also able to find a missense substitution in the *NIPBL* gene in a patient with a clinical diagnosis of CSS (patient 5, Fig.29g,h). The mutation (c.6886A>G; p.S2296G) affects the HEAT repeats, indispensable for protein-protein interactions. This patient is of particular interest because she was described as CSS unsolved case in a paper of 2013 (Wieczorek et al., 2013). In fact, she presented with clinical findings consistent with a diagnosis of CSS, namely a coarse face, low frontal hairline, synophrys, thick eyebrows, a broad nose with upturned nasal tip, a large mouth, hypoplasia of distal phalanges and fingernails, short metacarpals and metatarsals, delayed bone age, scoliosis and hirsutism. She also displayed severe intellectual disability, psychomotor delay, an autistic behavior, hearing loss and seizures.

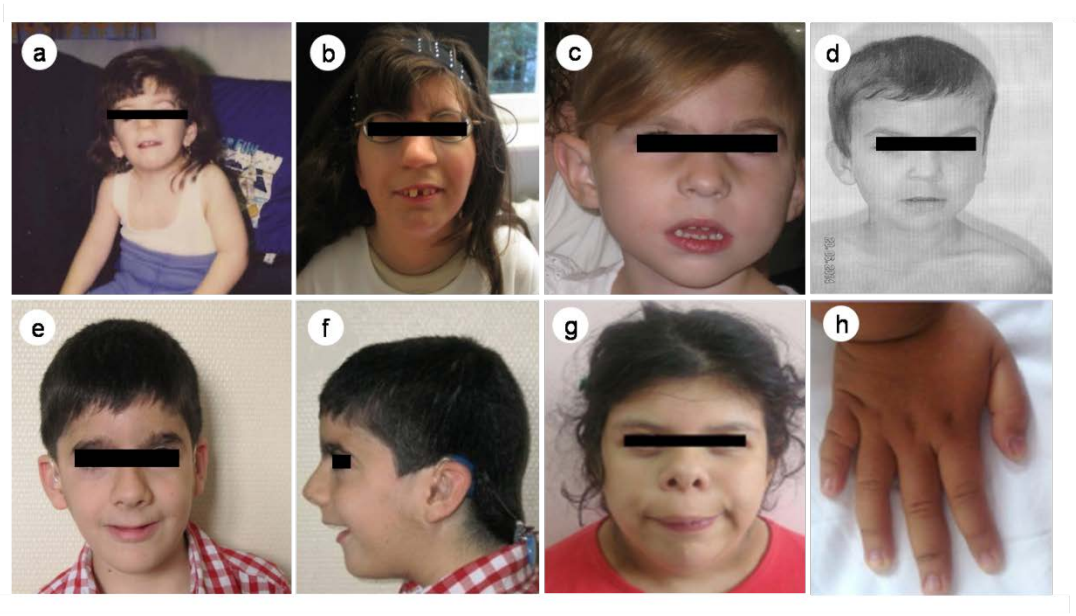


Figure 29. Phenotype of the patients. (a,b) Pictures taken at different ages of patient 1, carrying the missense mutation in SMARCB1. (c) Facial features of patient 2, harboring the out-of-frame deletion in ARID1A. (d) Facies of patient 3, carrying the splicing mutation in ARID1B. (e,f) Frontal and lateral view of patient 4, presenting with the deletion of chromosome 6 encopassing the entire ARID1B gene. (g,h) Pictures of the CSS patient (patient 5) carrying a missense mutation in NIPBL. To be noted the presence of hypoplasia of the distal phalanges and of the fingernails, clinical findings characteristic of CSS

Importantly, a recent paper could demonstrate a direct interaction between the cohesin loader NIPBL and the SWI/SNF complex (Lopez-Serra et al., 2014). Particularly, in yeast, NIPBL was found to be recruited by the SWI/SNF complex to broad nucleosome-free regions. It is therefore tempting to speculate that NIPBL and the SWI/SNF complex cooperate in order to mediate the transcriptional regulation of multiple genes (Lopez-Serra et al., 2014). Based on this hypothesis, we checked for a putative direct interaction between cohesin-related proteins and subunits of the SWI/SNF complex by a yeast two hybrid approach (Y2H). Hence, we generated Y2H plasmids containing cohesin-related proteins as a BD-bait and SWI/SNF subunits as a prey in a AD-fusion-construct. In detail, the following expression constructs were generated: NIPBL F1 (aa 1-879), NIPBL F2 (aa 728-1555), NIPBL F3 (aa 1501-2470), NIPBL F4 (aa 2121-2804), RAD21 full length, SMARCB1 full length, SMARCA5 F1 (aa 1-668), SMARCA5 F2 (aa 558-1052), ARID1B F1 (aa 1-668), ARID1B F2 (aa 562-1273), ARID1B F3 (aa 1188-1818), ARID1B F4 (aa 1720-2249), ARID1A F1 (aa 1-708), ARID1A F2 (aa 602-1297), ARID1A F3 (aa 1196-1871), ARID1A F4 (aa 1755-2285), SMARCA2 F1 (aa 1-701) and SMARCA2 F2 (aa 601-1590). Because of the peculiarity of their protein structures, SMC1A and SMC3 expression constructs were not generated.

Through this technique we were able to detect an interaction of NIPBL fragment F1 (aa 1-879) with SMARCB1 (Fig.30).

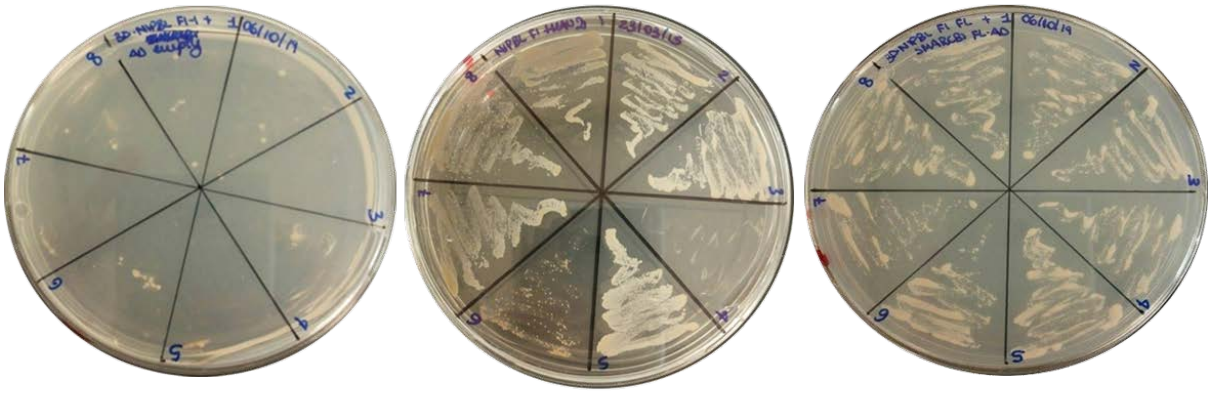


Figure 30. Yeast two hybrid –4 plates (a) Negative control NIPBL F1-BD + AD-empty (b) Positive control NIPBL F1-BD + MAU2-AD (c) Positive growth observed for those yeasts which were transformed with NIPBL F1-BD and SMARCB1-AD, indicating the existence of an interaction between the two proteins

We subsequently performed a Co-Immunoprecipitation (Co-IP) assay in HeLa cells in order to confirm the interaction between NIPBL and SMARCB1. By means of Co-IP we were able not only to confirm this interaction, but also to prove the existence of an interaction between SMC3 and SMARCB1.

We then assessed whether the identified interactions take place before or after the binding of the proteins to the DNA. Therefore, we separated the chromatin-bound and the soluble proteins of the cells and performed Co-IP on both protein fractions (Fig.31). By this we could demonstrate that the interaction of SMARCB1 with SMC3 as well as NIPBL takes place on both fractions, thus suggesting that cohesin and the SWI/SNF complex interact before binding to the DNA.

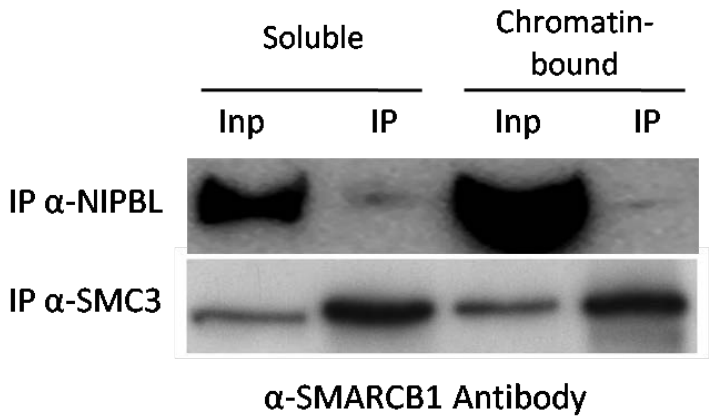


Figure 31. Results of the Co-IP experiments performed on two cellular fractions of HeLa cells: soluble proteins (left) and chromatin-bound proteins (right). A positive SMARCB1 signal is detected on the IP fraction of the α-NIPBL Co-IP as well as on the IP fraction of the α-SMC3 Co-IP, thus indicating that the interaction of SMARCB1 with the cohesin-related proteins takes place in both protein fractions. Inp (Input) represents the whole cell extract which was loaded as a positive control

Subsequently, we created smaller fragments of NIPBL and SMARCB1 in order to narrow down the protein domain responsible for the interaction. The new NIPBL fragments (Fig.32) include NIPBL F1-1 (aa 1-193), which comprises the domain responsible for the interaction of NIPBL with MAU2 (aa 1-38) and NIPBL F1-2, which does not contain any particular functional or structural domain. Similarly, SMARCB1 F2 (aa 157-404) contains the functional domain SNF5, while no functional or structural domains are localized in SMARCB1 F1 (aa 1-193). In addition, all the mutations so far reported in the literature in association with CSS (p.K363N, p.K364del, p.R377H) as well the mutation identified in patient 1 of our cohort (p.K324R) localize within the second fragment of SMARCB1. The Y2H assay performed on the smaller fragments showed that the interaction involves the first fragment of NIPBL and the second fragment of SMARCB1.

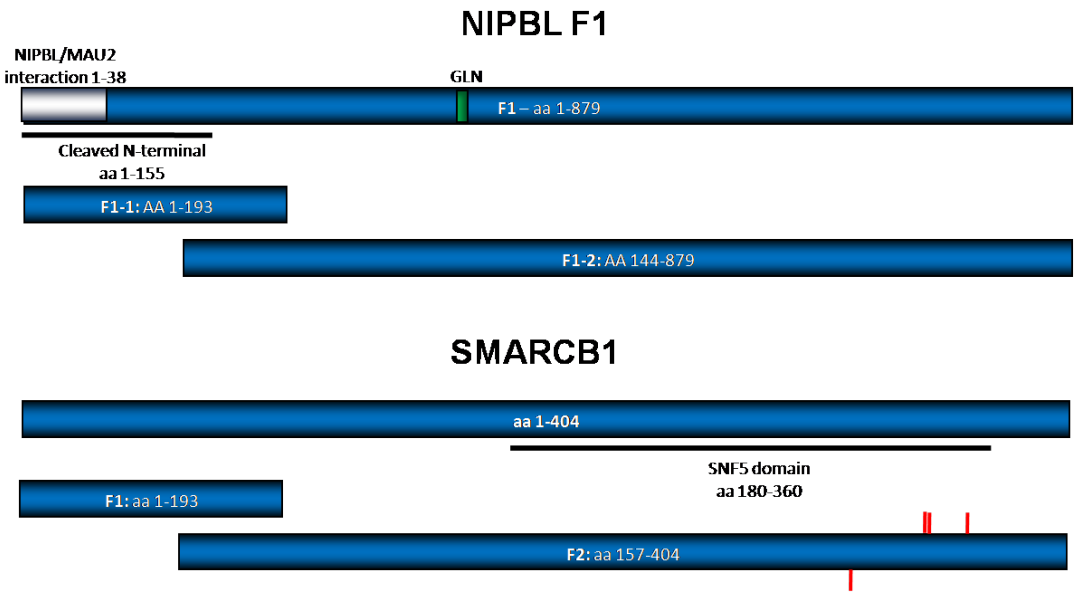


Figure 32. Schematic representation of the protein fragments created for the yeast two hybrid assay in order to narrow down the domains responsible for the interaction between NIPBL and SMARCB1. The first 879 aa of NIPBL were subdivided into two fragment: the first fragment (aa 1-193) comprises the domain of NIPBL responsible for the interaction with MAU2, whereas no functional or protein-protein interaction domains localize within the second fragment (aa 144-879). The full-length SMARCB1 was also subdivided in two fragments: the first fragment (aa 1-193) does not contain any relevant domain, while the second fragment (aa 157-404) contains the functional domain of the protein (SNF5 domain). In addition, all the mutations so far identified (red bars) localize in this fragment. Our experiments showed that the interaction between the two proteins is mediated by the first fragment of NIPBL and the second fragment of SMARCB1.

DISCUSSION

MUTATIONAL FREQUENCIES AND GENOTYPE-PHENOTYPE CORRELATION

CdLS is characterized by a wide clinical expressivity: the reported facial features as well as the number and severity of complications affecting other organs or systems are highly variable among patients (Kline et al., 2007). The broad range of phenotypes is partially accounted for by the genetic heterogeneity that underlies the syndrome. Up to date, mutations in five different genes have been reported in association with CdLS. These five genes comprise the structural components of the cohesin complex *SMC1A*, *SMC3* and *RAD21* and the regulators *NIPBL* and *HDAC8* (Krantz et al., 2004; Tonkin et al., 2004; Musio et al., 2006; Deardorff et al., 2007; Deardorff et al., 2012a; Deardorff et al., 2012b).

Thanks to a large and international cooperation we were able to assemble an extensive cohort of patients with CdLS or CdLS-overlapping phenotypes. By means of different techniques, patients were analyzed for the presence of mutations in the five CdLS-associated genes. This analysis led us to the identification of 109 mutations in *NIPBL*, 8 mutations in *SMC1A*, 15 mutations in *SMC3* and 10 mutations in *HDAC8*. Mutations in *RAD21* were not identified in any of the investigated patients. All the observed alterations affect highly conserved amino acids, and were reported to be disease causing by different bioinformatic tools (PolyPhen-2, SIFT and Mutation Taster). Our newly identified variants raise the total number of mutations affecting *NIPBL* to 470, thus confirming the cohesin loader as the major gene of the syndrome, being responsible for more than half of CdLS cases (Fig.33) (http://www.hgmd.cf.ac.uk/ac/gene.php?gene=NIPBL). Out-of-frame deletions or insertions together with missense substitutions are the most frequent observed alteration. As previously reported, patients harboring mutations in *NIPBL* tend to present with a more canonical phenotype characterized by marked growth retardation and developmental delay and by typical facial dysmorphisms. Limb malformations are also observed with a high frequency, in particular in association with out-of-frame alterations (Selicorni et al., 2007). Importantly, nine patients were found to carry mutations in a mosaic condition. The identified mosaic alterations comprise two splicing variants, two out-of-frame deletions, three missense substitutions and two nonsense mutations. Hence, no correlation can be established between the onset of mosaicism and the type of alteration. Interestingly, the nine patients carrying the mosaic variants present with a broad range of clinical features. Classic facial dysmorphisms and severe limbs reductions were also observed in this group of patients, thus indicating that the mosaic condition is not necessarily correlated to a mitigation of the effect of the mutation. The

assessment of whether mosaicism arises from a somatic mutation after fertilization or from negative selection of the cells harboring the mutation is currently an issue of utmost importance. All together, our findings strengthen the need for the integration of highly sensitive sequencing technologies into routine molecular work up as well as the importance to investigate DNA samples derived from more than one tissue, given that mutations are frequently not detectable on blood DNA.

All together, the remaining four genes *SMC1A*, *SMC3*, *HDAC8* and *RAD21* account for about 10-15% of CdLS cases (Fig.33). Our work raises the total number of mutations in these genes to 49, 16, 38 and 8, respectively (Borck et al., 2007; Deardorff et al., 2007, Liu et al., 2009; Limongelli et al., 2010; Mannini et al., 2010; Pié et al., 2010; Hoppman-Chaney et al., 2011; Deardorff et al., 2012a; Deardorff et al., 2012b; Gervasini et al., 2013; Feng et al., 2014; Kaiser et al., 2014; Gil-Rodriguez et al., 2015; Lebrun et al., 2015; Goldstein et al., 2015; Yuan et al., 2015). Albeit all kinds of variants have been reported for *HDAC8*, the vast majority of mutations in *SMC1A* or *SMC3* do not disrupt the reading frame of the protein, thus suggesting that the pathogenic mechanism associated with alterations of the core subunits of the cohesin complex may act through a dominant negative effect. Nevertheless, very recently, one alteration of the splicing and two 4-bp deletions, all resulting in a premature stop, have been identified in *SMC1A* in three female patients presenting with epilepsy and with few or absent CdLS features (Goldstein et al., 2015; Lebrun et al., 2015). These findings might indicate that haploinsufficiency of *SMC1A* might lead to a phenotype different from CdLS, whereas a dominant negative pathomechanism results in CdLS or CdLS-overlapping phenotypes.

The 34 patients herein described with mutations affecting *SMC1A*, *SMC3* and *HDAC8* present with a milder or less typical phenotype: several facial features commonly observed in typical CdLS are in fact absent or infrequent in this group of patients. For instance, synophrys is often absent or subtle, the nasal tip is often bulbous or broad and the nasal bridge is less frequently depressed. Nevertheless, these patients present with a broad range of phenotypes, for which a variable degree of cognitive impairment and a severe verbal and motor developmental delay are the only unifying features. The clinical heterogeneity is even wider in the presence of mutations affecting the X-linked genes *SMC1A* and *HDAC8*. Male patients with mutations in these genes usually display a highly recognizable, more homogeneous and severe CdLS phenotype. Conversely, females present with a broad range of clinical features and a variable degree of severity of the syndrome. In this context, the wide clinical expressivity observed in females might also be correlated with the extent of the skewing of the X-inactivation towards the

normal allele. Consistent with this possibility, a skewed X-inactivation was observed on blood DNA of almost all female patients harboring mutations in *SMC1A* and *HDAC8*.

Interestingly, clinical comparisons between patients sharing the same mutations revealed that the displayed clinical signs are not identical. For instance, the two male patients harboring the single amino acid deletion p.E488del in *SMC3* presented with a similar craniofacial appearance during infancy, but showed a different progression of the clinical signs with aging (Deardorff et al, 2007; Gil-Rodrìguez et al., 2015). The two patients also displayed a different degree of cognitive impairment and developmental delay, with one patient being unable to speak at the age of 23 years and with limited motor development and the other patient working as an adult. Similarly, the clinical comparison between two female patients carrying the variant p.V58_R62del in *SMC1A* indicated a similar limb involvement and the presence of GERD and myopia in both patients (Deardorff et al., 2007; Gervasini et al., 2013). Nevertheless, distinguishable clinical features were also observed: the patient described by Deardorff and colleagues displayed pulmonary stenosis, hearing loss and a borderline cognitive impairment whereas our patient presented with ptosis, highly arched palate, teeth malformations and a moderate degree of intellectual disability. Likewise, despite the existence of similar craniofacial features in the presence of the p.G320R substitution in *HDAC8*, the male patient of our cohort presented with pulmonary stenosis and vesicoureteral reflux, whereas the patient described in Deardorff et al., 2012b displayed hearing loss and seizures. These clinical comparisons reveal that the genetic background given by the combination of polymorphic and non polymorphic variants as well as the influence of environmental factors might contribute to the definition of the clinical phenotype, thus making the genotype-phenotype correlation complicated and difficult to interpret. All together, these findings suggest that the range of clinical features caused by alterations of the cohesin complex might be significantly broader than previously appreciated.

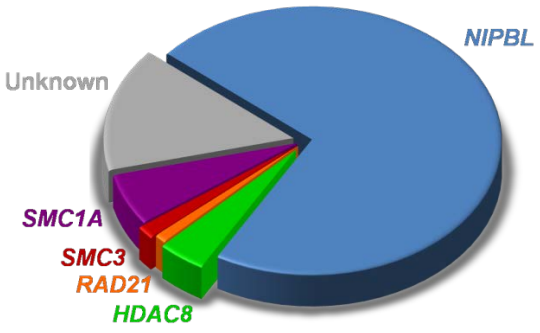


Figure 33. Graphical representation of the contribution exerted by each of the identified CdLS-genes. Up to date, 470 mutations have been observed in NIPBL, 45 in SMC1A, 16 in SMC3, 38 in HDAC8 and 8 in RAD21. Nevertheless, it is still not possible to identify the molecular cause of the syndrome in a large group of CdLS patients

OVERALL AND ALLELE-SPECIFIC EXPRESSION OF THE *SMC1A* GENE

Several studies have contributed to define the genetic cause of CdLS (Krantz et al., 2004; Tonkin et al., 2004; Musio et al., 2006; Deardorff et al., 2007; Deardorff et al., 2012a; Deardorff et al., 2012b). Nevertheless, little is known on the pathogenic mechanism underlying the syndrome. In case of heterozygous mutations affecting the level of the synthetized *NIPBL* gene product, upregulation of the wild type allele as an attempt to compensate for the lower expression of the mutant allele was demonstrated in both mouse and humans cells (Dorsett and Krantz, 2009; Kawauchi et al., 2009).Compensatory upregulation of the wild type allele is even more relevant for the X-linked gene *SMC1A*, which partially escapes X inactivation in humans and show a variable expression of the allele lying on the inactive X chromosome (Carrell and Willard, 2005). The portion of the allele that escapes X inactivation further contributes to the wide variability of features observed in female patients with *SMC1A* mutations. Therefore, the analysis of the expression of *SMC1A* is highly relevant for the clinical evaluation of female patients.

Previous studies in heterozygous females could demonstrate the presence of the mutant transcript and of normal levels of the SMC1A protein in patients (Musio et al., 2006; Liu et al., 2009; Revenkova et al., 2009). In addition, Gimigliano and colleagues demonstrated the incorporation of the mutant SMC1A into the cohesin complex (Gimigliano et al., 2012). Therefore, in order to ascertain the pathogenic mechanism associated with SMC1A, we analyzed the overall and allele-specific expression of the gene in the healthy population and in female patients with X-linked CdLS, carrying *SMC1A* mutations. So far, different studies reported a variable but always higher expression of SMC1A in females as compared to males (Craig et al., 2004; Johnston et al., 2008; Liu et al., 2009). In line with the previous data, our Real Time PCR experiments could demonstrate that the level of expression of the *SMC1A* transcript in healthy female controls is 50% higher than in healthy males. In agreement with the results of the transcriptional analysis, our western blot experiments demonstrated for the first time that the level of the SMC1A protein in females is 44% higher than in males. The same analysis, when extended to our cohort of female patients, revealed that there are no significant differences between healthy females and CdLS patients in the total amount of the SMC1A protein, thus consolidating the hypothesis that the pathogenic mechanism associated to missense mutations in *SMC1A* in female patients is not linked to haploinsufficiency, but to a probable dominant negative effect.

Albeit the overall amount of SMC1A remains unchanged, a previous study reported a different level of expression of the wild type and mutant alleles in a single patient (Liu et al., 2009). For this reason, we carried out a pyrosequencing assay on a group of six *SMC1A*-mutated females and 15 *SMC1A*-heterozygous controls in order to assess the expression ratio between the wild type and mutant alleles. This analysis revealed that all female patients expressed the two alleles differentially. In point of fact, patients always expressed the wild type allele at higher levels as compared to the mutant allele, with an average ratio between wild type and mutant allele of 2:1. Our findings partially justify the existence of patients with mutations in *SMC1A* displaying very mild or borderline phenotypes. For this reason, the assessment of the expression ratio between the wild type and the mutant allele becomes of utmost importance for the comprehension of the wide clinical expressivity in females.

Taken into account the dominant negative effect enacted by the altered SMC1A protein, it might be possible that level of expression of the mutant allele may act as modulator of the CdLS phenotype. In agreement with this hypothesis, patient 1, who presented with a borderline phenotype, displayed the lowest expression level of the mutant *SMC1A* allele. Anyhow, a clear correlation between the expression levels of SMC1A and the severity of the phenotype could be established only by extending the analysis to a larger cohort, and especially to those borderline cases that are frequently overlooked except for the identification of the affected progeny.

BEYOND COHESINOPATHIES: MUTATIONS IN CHROMATIN-ASSOCIATED FACTORS AS GENETIC CAUSE OF CdLS-OVERLAPPING PHENOTYPES

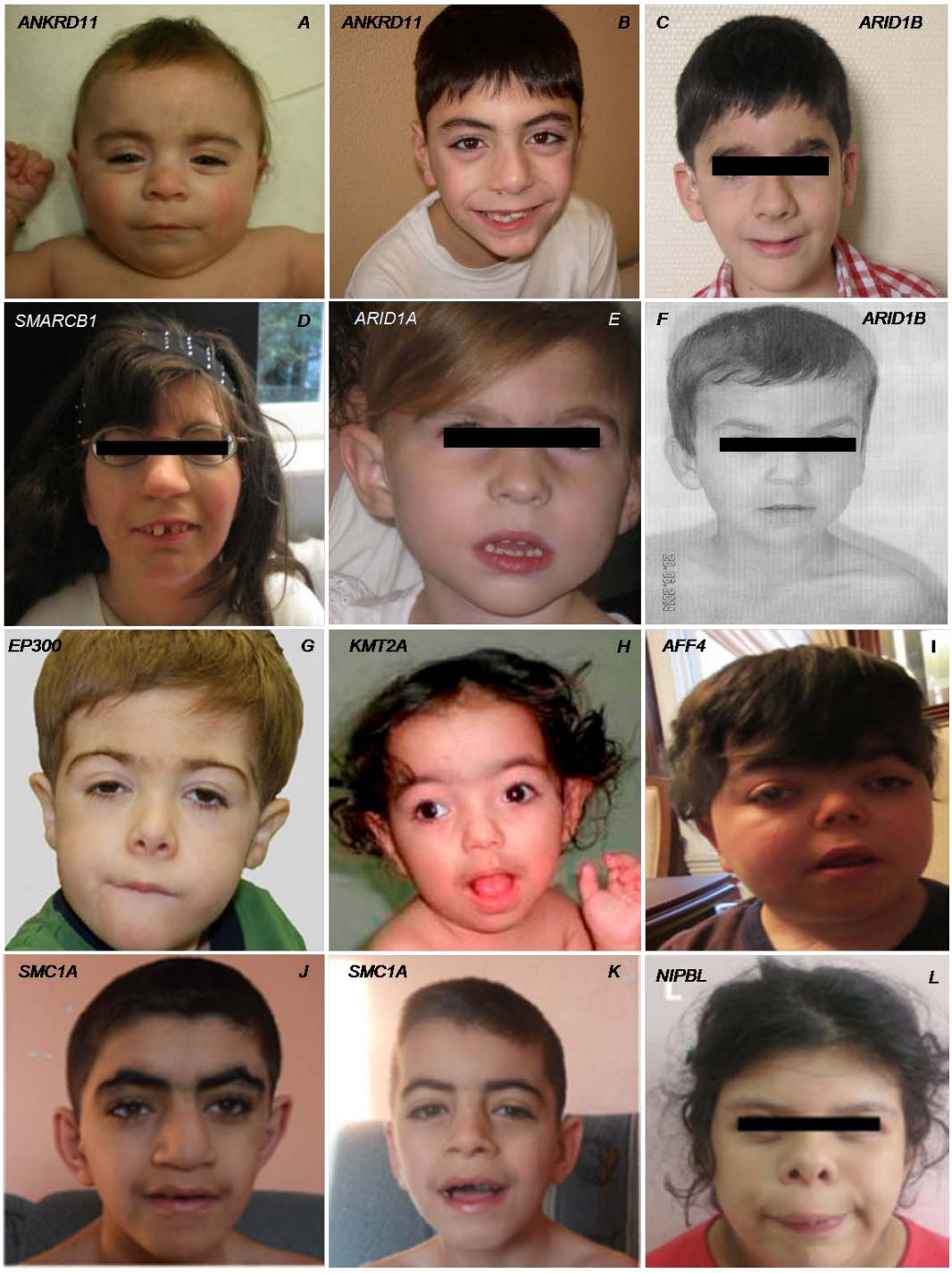


Figure 34. (A-I) Facial features of patients with CdLS or CdLS-overlapping features harboring mutations in transcription factors different from cohesins. (J-L) Phenotypical appearance of patients affected by multisystem developmental disorders different from CdLS, who were found to carry mutations in CdLS-related genes

The employment of high-throughput sequencing techniques has helped to shed light on part of the missing heritability in CdLS: mosaicism has been proven to be responsible for a high percentage of CdLS cases previously found to be negative for the presence of mutations in the five CdLS genes investigating DNA extracted from blood. The role of mosaicism is of particular relevance in association with *NIPBL*, but few mosaic cases have also been reported for *SMC1A*, *SMC3* and *HDAC8* (Huisman et al., 2012; Ansari et al., 2014; Baquero-Montoya et al., 2014; Braunholz et al., 2015). Despite the significant role played by mosaicism, the molecular cause for more than 20% of CdLS patients is still unknown. In this context, genes functionally related to cohesin or genes responsible for clinical entities partially overlapping with CdLS might both be able to cover this missing heritability.

Exome sequencing has recently led to the identification of mutations in chromatin-associated factors different from cohesin in patients with CdLS or CdLS-overlapping phenotypes (Fig.34A-I). These transcriptional regulators, which are responsible for the onset of other multisystem developmental disorders, include *EP300*, *AFF4* and *KMT2A* (Woods et al., 2014; Izumi et al., 2015; Yuan et al., 2015). In addition, next generation sequencing techniques targeted to the exome or to a selected set of genes allowed us to identify mutations in our cohort of patients in *ANKRD11* (Parenti et al., 2015) and in different subunits of the SWI/SNF complex, namely *SMARCB1*, *ARID1A* and *ARID1B*. EP300 is a transcriptional coactivator which is known to be responsible for Rubinstein-Taybi syndrome (Roelfsema et al., 2005). *KMT2A* is instead associated to Wiedemann-Steiner syndrome and encodes for a transcription factor that acts as a lysine-specific methyltransferase (Jones et al., 2012). The *ANKRD11* gene codes for the ankyrin repeat-containing protein 11, a repressor of the transcription of nuclear receptors' target genes and mutations in this gene are associated with KBG syndrome (Sirmaci et al., 2011). Mutations in different subunits of the SWI/SNF complex, responsible for chromatin remodeling, were instead associated to Coffin-Siris syndrome (Santen et al., 2012). Lastly, mutations in *AFF4*, a core component of the super elongation complex, were found to be responsible for a new syndrome that phenotypically overlaps with CdLS named CHOPS (Izumi et al., 2015). Correspondingly, mutations in the cohesin gene *SMC1A* were recently reported in two patients with a clinical diagnosis of Wiedemann-Steiner syndrome (Fig.34J-K) (Yuan et al., 2015). Similarly, we report herein of a mutation in the cohesin loader *NIPBL* in a patient with Coffin-Siris syndrome (Fig.34L).

Interestingly, although the aforementioned disorders are *de facto* distinct syndromes, they all share some phenotypical features such as a variable degree of intellectual disability, a general developmental delay, growth impairment and some craniofacial features including

brachycephaly and thick and arched eyebrows. Despite the presence of some relatively unique features for each syndrome, the observed phenotypical overlap suggests the existence of a common pathophysiology among these disorders. In this context, it is worth to be noted that the genes responsible for these syndromes are all involved in the regulation of gene expression. In addition, our yeast two hybrid experiments could prove a physical interaction between the SWI/SNF subunit SMARCB1 and the cohesin genes NIPBL and SMC3, whereas Lopez-Serra and colleagues demonstrated that the cohesin loader NIPBL and the SWI/SNF complex cooperate in order to maintain nucleosome depletion at target regions of the genome in yeast (Lopez-Serra et al., 2014). It is therefore tempting to speculate that mutations in the previously mentioned genes might be responsible for the alteration of the expression of functionally interconnected sets of genes, thus giving a molecular explanation for the similarities observed among the different syndromes. For this reason,Yuan and colleagues recently conceived the term “transcriptomopathies” to describe those syndromes sharing core clinical features as a consequence of global disturbances in the transcriptional regulation of the cell (Fig.35).

In conclusion, mutations in different transcription factors can result in clinical pictures that are difficult to distinguish from each other, thus indicating a wide pleiotropy that should be taken into account while addressing the clinical diagnosis. The extensive clinical and functional data reported in here indicate the existence of limitations of the current “state of the art” for the molecular diagnosis of CdLS and CdLS-overlapping phenotypes. The clinical heterogeneity of the syndrome and the numerous differential clinical diagnoses strongly suggest the integration of next generation sequencing technologies into the routine molecular diagnostics in order to enable a higher sensitivity and the concomitant analysis of multiple sets of genes.

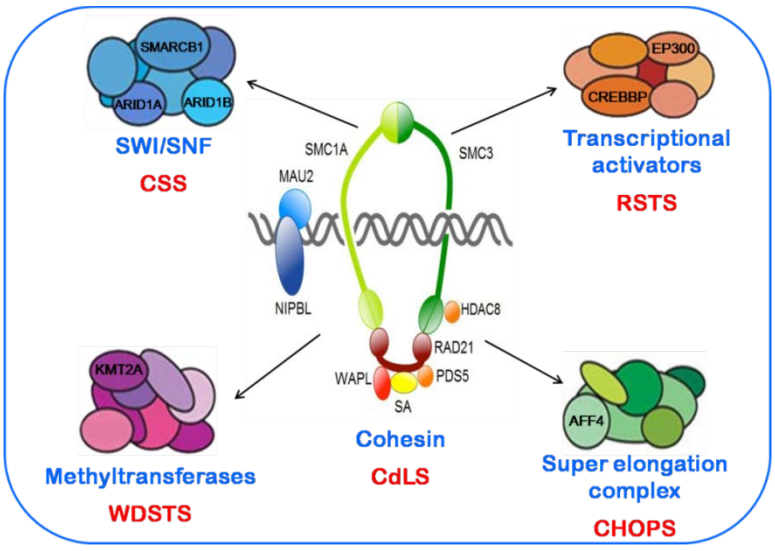


Figure 35. Mutations in different chromatin-associated factors are responsible for the so called “transcriptomopathies”, that is syndromes caused by alterations of interconnected sets of genes that share a variable degree of cognitive impairment and a general developmental delay

PERSPECTIVES

An extensive clinical and molecular heterogeneity characterizes CdLS. The identification of new genes responsible for overlapping phenotypes and the recognition of the important role played by mosaicism has helped to shed light on the molecular basis of the syndrome. Nevertheless, further advances are needed to fully understand the causative molecular mechanisms. In this context, additional experiments will be addressed to the resolution of the following issues:

- Definition of the mechanisms responsible for the onset of mosaicism
- Analysis of the expression mechanisms responsible for the onset of different phenotypes in the presence of the same DNA alteration
- Definition of the pathogenic mechanisms underlying the syndrome in the presence of out-of-frame deletions in *SMC1A*
- Identification of new causative genes
- Characterization of the interaction of SMARCB1 with NIPBL, with particular attention to the following aspects:
    - Does the presence of mutations disrupt the interaction between the two proteins?
    - Does MAU2 influence the interaction between NIPBL and SMARCB1?

REFERENCES

Aitken DA, Ireland M, Berry E, Crossley JA, Macri JN, Burn J, Connor JM (1999). Second-trimester pregnancy associated plasma protein-A levels are reduced in Cornelia de Lange syndrome pregnancies. *Prenat Diagn* 19(8):706-10

Ansari M, Poke G, Ferry Q, Williamson K, Aldridge R, Meynert AM, Bengani H, Chan CY, Kayserili H, Avci S, Hennekam RC, Lampe AK, Redeker E, Homfray T, Ross A, Falkenberg Smeland M, Mansour S, Parker MJ, Cook JA, Splitt M, Fisher RB, Fryer A, Magee AC, Wilkie A, Barnicoat A, Brady AF, Cooper NS, Mercer C, Deshpande C, Bennett CP, Pilz DT, Ruddy D, Cilliers D, Johnson DS, Josifova D, Rosser E, Thompson EM, Wakeling E, Kinning E, Stewart F, Flinter F, Girisha KM, Cox H, Firth HV, Kingston H, Wee JS, Hurst JA, Clayton-Smith J, Tolmie J, Vogt J, Tatton-Brown K, Chandler K, Prescott K, Wilson L, Behnam M, McEntagart M, Davidson R, Lynch SA, Sisodiya S, Mehta SG, McKee SA, Mohammed S, Holden S, Park SM, Holder SE, Harrison V, McConnell V, Lam WK, Green AJ, Donnai D, Bitner-Glindzicz M, Donnelly DE, Nellåker C, Taylor MS, FitzPatrick DR (2014). Genetic heterogeneity in Cornelia de Lange syndrome (CdLS) and CdLS-like phenotypes with observed and predicted levels of mosaicism. J Med Genet. 2014 Oct;51(10):659-68

Arbuzova S, Nikolenko M, Krantz D, Hallahan T, Macri J (2003). Low first-trimester pregnancy-associated plasma protein-A and Cornelia de Lange syndrome. *Prenat Diagn* 23(10):864

Arumugam P, Gruber S, Tanaka K, Haering CH, Mechtler K, Nasmyth K (2003). ATP hydrolysis is required for cohesin's association with chromosomes. *Curr. Biol* (13) 1941–1953

Baquero-Montoya C, Gil-Rodríguez MC, Braunholz D, Teresa-Rodrigo ME, Obieglo C, Gener B, Schwarzmayr T, Strom TM, Gómez-Puertas P, Puisac B, Gillessen-Kaesbach G, Musio A, Ramos FJ, Kaiser FJ, Pié J (2014). Somatic mosaicism in a Cornelia de Lange syndrome patient with NIPBL mutation identified by different next generation sequencing approaches. *Clin Genet* 86(6):595-7

Barisic I, Tokic V, Loane M, Bianchi F, Calzolari E, Garne E, Wellesley D, Dolk H, EUROCAT Working Group (2008). Descriptive Epidemiology of Cornelia de Lange Syndrome in Europe. *Am J Med Genet A* 146A(1):51-9

Bartsch O, Schmidt S, Richter M, Morlot S, Seemanová E, Wiebe G, Rasi S (2005). DNA sequencing of CREBBP demonstrates mutations in 56% of patients with Rubinstein-Taybi syndrome (RSTS) and in another patient with incomplete RSTS. *Hum Genet* 117(5):485-93

Basile E, Villa L, Selicorni A, Molteni M (2007). The behavioural phenotype of Cornelia de Lange Syndrome: a study of 56 individuals. *J Intellect Disabil Res* 51(9):671-81

Bentivegna A, Milani D, Gervasini C, Castronovo P, Mottadelli F, Manzini S, Colapietro P, Giordano L, Atzeri F, Divizia MT, Uzielli ML, Neri G, Bedeschi MF, Faravelli F, Selicorni A, Larizza L (2006). Rubinstein-Taybi Syndrome: spectrum of CREBBP mutations in Italian patients. *BMC Med Genet* 7:77

Berney TP, Ireland M, Burn J (1999). Behavioural phenotype of Cornelia de Lange syndrome. *Arch Dis Child* 81(4):333-6

Birkenbihl RP, Subramani S (1992). Cloning and characterization of rad21 an essential gene of Schizosaccharomyces pombe involved in DNA double-strand-break repair. *Nucleic Acids Res* 25;20(24):6605-11

Borck G, Redon R, Sanlaville D, Rio M, Prieur M, Lyonnet S, Vekemans M, Carter NP, Munnich A, Colleaux L, Cormier-Daire V (2004). NIPBL mutations and genetic heterogeneity in Cornelia de Lange syndrome. *J Med Genet* 41(12):e128

Borck G, Zarhrate M, Bonnefont JP, Munnich A, Cormier-Daire V, Colleaux L (2007). Incidence and clinical features of X-linked Cornelia de Lange syndrome due to SMC1L1 mutations. *Hum Mutat* 28(2):205-6

Bose T, Lee KK, Lu S, Xu B, Harris B, Slaughter B, Unruh J, Garrett A, McDowell W, Box A, Li H, Peak A, Ramachandran S, Seidel C, Gerton JL (2012). Cohesin proteins promote ribosomal RNA production and protein translation in yeast and human cells. *PLoS Genet* 8(6):e1002749

Brachmann E (1916). Ein Fall von symmetrischer Monodaktylie durch Ulnadefekt, mit symmetrischer Flughaurbildung in den Ellenbogen sowie anderen Abnormalitaten. *Jahrbuch für Kinderheilkunde und physische Erziehung* 84:225-235

Braunholz D, Obieglo C, Parenti I, Pozojevic J, Eckhold J, Reiz B, Braenne I, Wendt KS, Watrin E, Vodopiutz J, Rieder H, Gillessen-Kaesbach G, Kaiser FJ (2015). Hidden mutations in cornelia de lange syndrome limitations of sanger sequencing in molecular diagnostics. *Hum Mutat* 36(1):26-9

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

Castronovo P, Gervasini C, Cereda A, Masciadri M, Milani D, Russo S, Selicorni A, Larizza L (2009). Premature chromatid separation is not a useful diagnostic marker for Cornelia de Lange syndrome. *Chromosome Res* 17(6):763-71

Chen CY, Morris Q, Mitchell JA (2012). Enhancer identification in mouse embryonic stem cells using integrative modeling of chromatin and genomic features. *BMC Genomics* 13:152

Cingolani P, Platts A, Wang le L, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly (Austin)* 6(2):80-92

Ciosk R, Zachariae W, Michaelis C, Shevchenko A, Mann M, Nasmyth K (1998). An ESP1/PDS1 complex regulates loss of sister chromatid cohesion at the metaphase to anaphase transition in yeast. *Cell* 93(6):1067-76

Ciosk R, Shirayama M, Shevchenko A, Tanaka T, Toth A, Shevchenko A, Nasmyth K (2000). Cohesin's binding to chromosomes depends on a separate complex consisting of Scc2 and Scc4 proteins. *Mol Cell* 5:1–20

Clark DM, Sherer I, Deardorff MA, Byrne JL, Loomes KM, Nowaczyk MJ, Jackson LG, Krantz ID (2012). Identification of a prenatal profile of Cornelia de Lange syndrome (CdLS): a review of 53 CdLS pregnancies. *Am J Med Genet A* 158A: 1848-56

Coppus AM (2013). People with intellectual disability: what do we know about adulthood and life expectancy? *Dev Disabil Res Rev* 18: 6-16

Cortés-Ledesma F, Aguilera A (2006). Double-strand breaks arising by replication through a nick are repaired by cohesin-dependent sister-chromatid exchange. *EMBO Rep* 7(9):919-26

Craig IW, Mill J, Craig GM, Loat C, Schalkwyk LC (2004). Application of microarrays to the analysis of the inactivation status of human X-linked genes expressedin lymphocytes. *Eur J Hum Genet* 12:639-46

Cuylen S, Haering CH (2011). Deciphering condensin action during chromosome segregation. *Trends Cell Biol* 21 552–559

D'Ambrosio C, Schmidt CK, Katou Y, Kelly G, Itoh T, Shirahige K, Uhlmann F (2008). Identification of cis-acting sites for condensin loading onto budding yeast chromosomes. *Genes Dev* 22(16):2215-27

Deardorff MA, Kaur M, Yaeger D, Rampuria A, Korolev S, Pie J, Gil-Rodríguez C, Arnedo M, Loeys B, Kline AD, Wilson M, Lillquist K, Siu V, Ramos FJ, Musio A, Jackson LS, Dorsett D, Krantz ID (2007). 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* 80(3):485-94

Deardorff MA, Wilde JJ, Albrecht M, Dickinson E, Tennstedt S, Braunholz D, Mönnich M, Yan Y, Xu W, Gil-Rodríguez MC, Clark D, Hakonarson H, Halbach S, Michelis LD, Rampuria A, Rossier E, Spranger S, Van Maldergem L, Lynch SA, Gillessen-Kaesbach G, Lüdecke HJ, Ramsay RG, McKay MJ, Krantz ID, Xu H, Horsfield JA, Kaiser FJ (2012a). RAD21 Mutations Cause a Human Cohesinopathy. *Am J Hum Genet* 90(6):1014-27

Deardorff MA, Bando M, Nakato R, Watrin E, Itoh T, Minamino M, Saitoh K, Komata M, Katou Y, Clark D, Cole KE, De Baere E, Decroos C, Di Donato N, Ernst S, Francey LJ, Gyftodimou Y, Hirashima K, Hullings M, Ishikawa Y, Jaulin C, Kaur M, Kiyono T, Lombardi PM, Magnaghi-Jaulin L, Mortier GR, Nozaki N, Petersen MB, Seimiya H, Siu VM, Suzuki Y, Takagaki K, Wilde JJ, Willems PJ, Prigent C, Gillessen-Kaesbach G, Christianson DW, Kaiser FJ, Jackson LG, Hirota T, Krantz ID, Shirahige K (2012b). HDAC8 mutations in Cornelia de Lange syndrome affect the cohesin acetylation cycle. *Nature* 489(7415):313-7

de Lange C 1933. Sur un type nouveau de dégénération (typus Amstelodamensis). *Arch Med Enfants* 36 :713-7193

DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M, McKenna A, Fennell TJ, Kernytsky AM, Sivachenko AY, Cibulskis K, Gabriel SB, Altshuler D, Daly MJ (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* 43(5):491-8

Díaz-Martínez LA, Giménez-Abián JF, Clarke DJ (2008). Chromosome cohesion - rings, knots, orcs and fellowship. *J Cell Sci* 1;121(Pt 13):2107-14

Donze D, Adams CR, Rine J, Kamakaka RT (1999). The boundaries of the silenced HMR domain in Saccharomyces cerevisiae. *Genes Dev.* 13(6):698–708

Dorsett D (2004). Adherin: key to the cohesin ring and cornelia de Lange syndrome. *Curr Biol* 14:R834-6

Dorsett D, Krantz ID (2009). On the molecular etiology of Cornelia de Lange syndrome. *Ann N Y Acad Sci* 1151:22-37

Dorsett D (2011). Cohesin: genomic insights into controlling gene transcription and development. *Curr Opin Genet Dev* (2):199-206

Ellermeier C, Smith GR (2005). Cohesins are required for meiotic DNA breakage and recombination in Schizosaccharomyces pombe. Proc. *Natl. Acad Sci* USA 102, 10952–10957

Fay A, Misulovin Z, Li J, Schaaf CA, Gause M, Gilmour DS, Dorsett D (2011). Cohesin selectively binds and regulates genes with paused RNA polymerase. *Curr. Biol* 21, 1624–1634

Feeney AJ, Verma-Gaur J (2012). CTCF-cohesin complex: architect of chromatin structure regulates V(D)J rearrangement. *Cell Res* 22, 280–282

Feng L, Zhou D, Zhang Z, Liu Y, Yang Y (2014). Exome sequencing identifies a *de novo* mutation in HDAC8 associated with Cornelia de Lange syndrome. *J Hum Genet* 59(9):536-9

Gervasini C, Parenti I, Picinelli C, Azzollini J, Masciadri M, Cereda A, Selicorni A, Russo S, Finelli P, Larizza L (2013). Molecular characterization of a mosaic NIPBL deletion in a Cornelia de Lange patient with severe phenotype. *Eur J Med Genet* 56(3):138-43

Gervasini C, Russo S, Cereda A, Parenti I, Masciadri M, Azzollini J, Melis D, Aravena T, Doray B, Ferrarini A, Garavelli L, Selicorni A, Larizza L (2013). Cornelia de Lange individuals with new and recurrent SMC1A mutations enhance delineation of mutation repertoire and phenotypic spectrum. *Am J Med Genet A* 161A(11):2909-19

Gietz RD, Woods RA (2002). Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Methods Enzymol* 350:87-96

Gillis LA, McCallum J, Kaur M, DeScipio C, Yaeger D, Mariani A, Kline AD, Li HH, Devoto M, Jackson LG, Krantz ID (2004). NIPBL mutational analysis in 120 individuals with Cornelia de Lange syndrome and evaluation of genotype-phenotype correlations. *Am J Hum Genet* 75(4):610-23

Gil-Rodríguez MC, Deardorff MA, Ansari M, Tan CA, Parenti I, Baquero-Montoya C, Ousager LB, Puisac B, Hernández-Marcos M, Teresa-Rodrigo ME, Marcos-Alcalde I, Wesselink JJ, Lusa-Bernal S, Bijlsma EK, Braunholz D, Bueno-Martinez I, Clark D, Cooper NS, Curry CJ, Fisher R, Fryer A, Ganesh J, Gervasini C, Gillessen-Kaesbach G, Guo Y, Hakonarson H, Hopkin RJ, Kaur M, Keating BJ, Kibaek M, Kinning E, Kleefstra T, Kline AD, Kuchinskaya E, Larizza L, Li YR, Liu X, Mariani M, Picker JD, Pié Á, Pozojevic J, Queralt E, Richer J, Roeder E, Sinha A, Scott RH, So J, Wusik KA, Wilson L, Zhang J, Gómez-Puertas P, Casale CH, Ström L, Selicorni A, Ramos FJ, Jackson LG, Krantz ID, Das S, Hennekam RC, Kaiser FJ, FitzPatrick DR, Pié J (2015). De novo heterozygous mutations in SMC3 cause a range of Cornelia de Lange syndrome-overlapping phenotypes. *Hum Mutat* 36(4):454-62

Giménez-Abián JF, Sumara I, Hirota T, Hauf S, Gerlich D, de la Torre C, Ellenberg J, Peters JM (2004). Regulation of sister chromatid cohesion between chromosome arms. *Curr Biol* 14(13):1187-93

Gimigliano A, Mannini L, Bianchi L, Puglia M, Deardorff MA, Menga S, Krantz ID, Musio A, Bini L (2012). Proteomic profile identifies dysregulated pathways in Cornelia de Lange syndrome cells with distinct mutations in SMC1A and SMC3 genes. *J Proteome Res* 11(12):6111-23

Goldstein JH, Tim-Aroon T, Shieh J, Merrill M, Deeb KK, Zhang S, Bass NE, Bedoyan JK (2015). Novel SMC1A frameshift mutations in children with developmental delay and epilepsy. *Eur J Med Genet* pii: S1769-7212(15)30019-7. doi: 10.1016/j.ejmg.2015.09.007. [Epub ahead of print]

Gorlin RJ, Cohen M, Hennekam RCM (2001). Syndromes of the head and neck. *Ney York, NY: Oxford University Press. pp.* 382–387

Gruber S, Haering CH, Nasmyth K (2003). Chromosomal cohesin forms a ring. *Cell* 112(6):765-77

Gruber S, Arumugam P, Katou Y, Kuglitsch D, Helmhart W, Shirahige K, Nasmyth K (2006) Evidence that loading of cohesin onto chromosomes involves opening of its SMC hinge. *Cell* 127 523–537

Guacci V, Koshland D, Strunnikov A (1997). A direct link between sister chromatid cohesion and chromosome condensation revealed through the analysis of MCD1 in S. cerevisiae. *Cell* 91(1):47-57

Gullerova M, Proudfoot NJ (2008). Cohesin complex promotes transcriptional termination between convergent genes in S. pombe. *Cell* 132 983–995.

Haering CH, Löwe J, Hochwagen A, Nasmyth K (2002). Molecular architecture of SMC proteins and the yeast cohesin complex. *Mol Cell* 9(4):773-88

Haering CH, Schoffnegger D, Nishino T, Helmhart W, Nasmyth K, Löwe J (2004). Structure and stability of cohesin's Smc1-kleisin interaction. *Mol Cell* 15(6):951-64

Hakimi MA, Bochar DA, Schmiesing JA, Dong Y, Barak OG, Speicher DW, Yokomori K, Shiekhattar R (2002). A chromatin remodelling complex that loads cohesin onto human chromosomes. *Nature* 418(6901):994-8

Hartman T, Stead K, Koshland D, Guacci V (2000). Pds5p is an essential chromosomal protein required for both sister chromatid cohesion and condensation in Saccharomyces cerevisiae. *J Cell Biol* 151(3):613-26

Hennekam RCM (2006). Rubinstein-Taybi syndrome. *Eur J Hum Genet* 14(9):981-5

Herrmann J, Pallister PD, Tiddy W, Opitz JM (1975). The KBG syndrome-a syndrome of short stature, characteristic facies, mental retardation, macrodontia, and skeletal anomalies. *Birth Defects Orig Artic Ser* 11: 7–18

Hoppman-Chaney N, Jang JS, Jen J, Babovic-Vuksanovic D, Hodge JC (2011). In-frame multi-exon deletion of SMC1A in a severely affected female with Cornelia de Lange Syndrome. Am J Med Genet A 158A(1):193-8

Huang WH, Porto M (2002). Abnormal first-trimester fetal nuchal translucency and Cornelia De Lange syndrome. *Obstet Gynecol* 99:956-8

Huisman SA, Redeker EJ, Maas SM, Mannens MM, Hennekam RCM (2013). High rate of mosaicism in individuals with Cornelia de Lange syndrome. *J Med Genet* 50(5):339-44

Ivanov D, Schleiffer A, Eisenhaber F, Mechtler K, Haering CH, Nasmyth K (2002). Eco1 is a novel acetyltransferase that can acetylate proteins involved in cohesion. *Curr Biol* 12(4):323-8

Izumi K, Nakato R, Zhang Z, Edmondson AC, Noon S, Dulik MC, Rajagopalan R, Venditti CP, Gripp K, Samanich J, Zackai EH, Deardorff MA, Clark D, Allen JL, Dorsett D, Misulovin Z, Komata M, Bando M, Kaur M, Katou Y, Shirahige K, Krantz ID (2015). Germline gain-of-function mutations in AFF4 cause a developmental syndrome functionally linking the super elongation complex and cohesin. *Nat Genet*. 2015 Apr;47(4):338-44

Jackson L, Kline AD, Barr MA, Koch S (1993). De Lange syndrome: a clinical review of 310 individuals. *Am J Med Genet* 47(7):940-6

Jahnke P, Xu W, Wülling M, Albrecht M, Gabriel H, Gillessen-Kaesbach G, Kaiser FJ (2008). The Cohesin loading factor NIPBL recruits histone deacetylases to mediate local chromatin modifications. *Nucleic Acids Res* 36(20):6450-8

Johnston CM, Lovell FL, Leongamornlert DA, Stranger BE, Dermitzakis ET, Ross MT (2008). Large-scale population study of human cell lines indicates that dosage compensation is virtually complete. *PLoS Genet* 4:e9; PMID:18208332

Jones WD, Dafou D, McEntagart M, Woollard WJ, Elmslie FV, Holder-Espinasse M, Irving M, Saggar AK, Smithson S, Trembath RC, Deshpande C, Simpson MA (2012). De novo mutations in MLL cause Wiedemann-Steiner syndrome. *Am J Hum Genet* 91(2):358-64

Kaiser FJ, Ansari M, Braunholz D, Concepción Gil-Rodríguez M, Decroos C, Wilde JJ, Fincher CT, Kaur M, Bando M, Amor DJ, Atwal PS, Bahlo M, Bowman CM, Bradley JJ, Brunner HG, Clark D, Del Campo M, Di Donato N, Diakumis P, Dubbs H, Dyment DA, Eckhold J, Ernst S, Ferreira JC, Francey LJ, Gehlken U, Guillén-Navarro E, Gyftodimou Y, Hall BD, Hennekam R, Hudgins L, Hullings M, Hunter JM, Yntema H, Innes AM, Kline AD, Krumina Z, Lee H, Leppig K, Lynch SA, Mallozzi MB, Mannini L, McKee S, Mehta SG, Micule I; Care4Rare Canada Consortium, Mohammed S, Moran E, Mortier GR, Moser JA, Noon SE, Nozaki N, Nunes L, Pappas JG, Penney LS, Pérez-Aytés A, Petersen MB, Puisac B, Revencu N, Roeder E, Saitta S, Scheuerle AE, Schindeler KL, Siu VM, Stark Z, Strom SP, Thiese H, Vater I, Willems P, Williamson K, Wilson LC; University of Washington Center for Mendelian Genomics, Hakonarson H, Quintero-Rivera F, Wierzba J, Musio A, Gillessen-Kaesbach G, Ramos FJ, Jackson LG, Shirahige K, Pié J, Christianson DW, Krantz ID, Fitzpatrick DR, Deardorff MA (2014). 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* 23(11):2888-900

Kawauchi S, Calof AL, Santos R, Lopez-Burks ME, Young CM, Hoang MP, Chua A, Lao T, Lechner MS, Daniel JA, Nussenzweig A, Kitzes L, Yokomori K, Hallgrimsson B, Lander AD (2009). Multiple organ system defects and transcriptional dysregulation in the Nipbl(+/-) mouse, a model of Cornelia de Lange syndrome. *PLoS Genet* 5(9):e1000650

Kim ST, Xu B, Kastan MB (2002). Involvement of the cohesin protein, Smc1, in Atm-dependent and independent responses to DNA damage. *Genes Dev* 16(5):560-70

Kline AD, Barr M, Jackson LG (1993a). Growth manifestations in the Brachmann-de Lange syndrome. *Am J Med Genet* 47(7):1042-9

Kline AD, Stanley C, Belevich J, Brodsky K, Barr M, Jackson LG (1993b). Developmental data on individuals with the Brachmann-de Lange syndrome. *Am J Med Genet* 47(7):1053-8

Kline AD, Krantz ID, Sommer A, Kliewer M, Jackson LG, FitzPatrick DR, Levin AV, Selicorni A (2007). Cornelia de Lange syndrome: clinical review, diagnostic and scoring systems, and anticipatory guidance. *Am J Med Genet A* 143A(12):1287-96

Kong X, Ball Jr, Sonoda E, Feng J, Takeda S, Fukagawa T, Yen TJ, Yokomori K (2009). Cohesin associates with spindle poles in a mitosis-specific manner and functions in spindle assembly in vertebrate cells. *Mol Biol Cell* 20, 1289–1301

Kosho T, Miyake N, Carey JC (2014). Coffin-Siris syndrome and related disorders involving components of the BAF (mSWI/SNF) complex: historical review and recent advances using next generation sequencing. *Am J Med Genet C Semin Med Genet* 166C(3):241-51

Krantz ID, McCallum J, DeScipio C, Kaur M, Gillis LA, Yaeger D, Jukofsky L, Wasserman N, Bottani A, Morris CA, Nowaczyk MJ, Toriello H, Bamshad MJ, Carey JC, Rappaport E, Kawauchi S, Lander AD, Calof AL, Li HH, Devoto M, Jackson LG (2004). Cornelia de Lange syndrome is caused by mutations in NIPBL, the human homolog of Drosophila melanogaster Nipped-B. *Nat Genet* 36(6):631-5

Kueng S, Hegemann B, Peters BH, Lipp JJ, Schleiffer A, Mechtler K, Peters JM (2006). Wapl controls the dynamic association of cohesin with chromatin. *Cell* 127(5): 955–967

Lavoie BD, Hogan E, Koshland D (2002). In vivo dissection of the chromosome condensation machinery: reversibility of condensation distinguishes contributions of condensin and cohesin. *J Cell Biol* 156, 805–815

Lavoie BD, Hogan E, Koshland D (2004) In vivo requirements for rDNA chromosome condensation reveal two cell-cycle-regulated pathways for mitotic chromosome folding. *Genes Dev* (18) 76–87

Lebrun N, Lebon S, Jeannet PY, Jacquemont S, Billuart P, Bienvenu T (2015). Early-onset encephalopathy with epilepsy associated with a novel splice site mutation in SMC1A. *Am J Med Genet A* doi: 10.1002/ajmg.a.37364. [Epub ahead of print]

Li H, Durbin R (2010) Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 26: 589–595

Limongelli G, Russo S, Digilio MC, Masciadri M, Pacileo G, Fratta F, Martone F, Maddaloni V, D'Alessandro R, Calabro P, Russo MG, Calabro R, Larizza L (2010). Hypertrophic Cardiomyopathy in a Girl With Cornelia de Lange Syndrome Due to Mutation in SMC1A. *Am J Med Genet Part A* 152A:2127-2129

Liu J, Krantz ID (2009). Cornelia de Lange syndrome, cohesin, and beyond. *Clin Genet* 76(4):303-14

Liu J, Feldman R, Zhang Z, Deardorff MA, Haverfield EV, Kaur M, Li JR, Clark D, Kline AD, Waggoner DJ, Das S, Jackson LG, Krantz ID (2009). SMC1A expression and mechanism of pathogenicity in probands with X-linked Cornelia de Lange syndrome. *Hum Mutat* 30(11):1535-42

Livak KJ, Schmittgen TD (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Δ Δ C(T)) Method. *Methods* 25:402-8

Lopez-Serra L, Kelly G, Patel H, Stewart A, Uhlmann F (2014). The Scc2-Scc4 complex acts in sister chromatid cohesion and transcriptional regulation by maintaining nucleosome-free regions. Nat Genet 46(10):1147-51

Losada A, Hirano M, Hirano T (1998). Identification of Xenopus SMC protein complexes required for sister chromatid cohesion. *Genes Dev* 12(13):1986-97

Losada A, Yokochi T, Kobayashi R, Hirano T (2000). Identification and characterization of SA/Scc3p subunits in the Xenopus and human cohesin complexes. *J Cell Biol* 150(3):405-16

Losada A, Hirano M, Hirano T (2002). Cohesin release is required for sister chromatid resolution, but not for condensin-mediated compaction, at the onset of mitosis. *Genes Dev* (16) 3004–3016

Losada A, Hirano T (2005). Dynamic molecular linkers of the genome: the first decade of SMC proteins. *Genes Dev* 19(11):1269-87

Losada A (2007). Cohesin regulation: fashionable ways to wear a ring. *Chromosoma* 116(4):321-9

Luo H, Li Y, Mu JJ, Zhang J, Tonaka T, Hamamori Y, Jung SY, Wang Y, Qin J (2008). Regulation of intra S phase checkpoint by ionizing radiation (IR)-dependent and IR-independent phosphorylation of SMC3. *J Biol Chem* 283, 19176–19183.

Luzzani S, Macchini F, Valadè A, Milani D, Selicorni A (2003). Gastroesophageal reflux and Cornelia de Lange syndrome: typical and atypical symptoms. *Am J Med Genet A* 119A(3):283-7

Mannini L, Liu J, Krantz ID, Musio A (2010). Spectrum and consequences of SMC1A mutations: the unexpected involvement of a core component of cohesin in human disease. *Hum Mutat* 31(1):5-10

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

Marchisio P, Selicorni A, Bianchini S, Milani D, Baggi E, Cerutti M, Larizza L, Principi N, Esposito S (2014). Audiological findings, genotype and clinical severity score in Cornelia de Lange syndrome. *Int J Pediatr Otorhinolaryngol* 78(7):1045-8

Mariani M, Bettini LR, Cereda A, Maitz S, Gervasini C, Russo S, Masciadri M, Biondi A, Larizza L, Selicorni A (2013). Germline mosaicism in Cornelia de Lange syndrome: dilemmas and risk figures. *Am J Med Genet A* 161A(7):1825-6

Masumoto K, Izaki T, Arima T (2001). Cornelia de Lange syndrome associated with cecal volvulus: report of a case. *Acta Paediatr* 90(6):701-3

McIntyre J, Muller EGD, Weitzer S, Snydsman BE, Davis TN, Uhlmann F (2007). In vivo analysis of cohesin architecture using FRET in the budding yeast Saccharomyces cerevisiae. *EMBO J* (26) 3783–3793

McNairn AJ, Gerton JL (2008). Cohesinopathies: One ring, many obligations. *Mutat Research* 647(1-2):103–111

Mehta AV, Ambalavanan SK (1997). Occurrence of congenital heart disease in children with Brachmann-de Lange syndrome. *Am J Med Genet* 71(4):434-5

Mehta GD, Rizvi SM, Ghosh SK (2012). Cohesin: a guardian of genome integrity. *Biochim Biophys Acta* 1823(8):1324-42

Michaelis C, Ciosk R, Nasmyth K (1997). Cohesins: chromosomal proteins that prevent premature separation of sister chromatids. *Cell* 91(1):35-45

Minor A, Shinawi M, Hogue JS, Vineyard M, Hamlin DR, Tan C, Donato K, Wysinger L, Botes S, Das S, Del Gaudio D (2014). Two novel RAD21 mutations in patients with mild Cornelia de Lange syndrome-like presentation and report of the first familial case. *Gene* 537(2):279-84

Misulovin Z, Schwartz YB, Li XY, Kahn TG, Gause M, MacArthur S, Fay JC, Eisen MB, Pirrotta V, Biggin MD, Dorsett D (2008). Association of cohesin and Nipped-B with transcriptionally active regions of the Drosophila melanogaster genome. *Chromosoma* 117(1):89-102

Musio A, Montagna C, Mariani T, Tilenni M, Focarelli ML, Brait L, Indino E, Benedetti PA, Chessa L, Albertini A, Ried T, Vezzoni P (2005). SMC1 involvement in fragile site expression. *Hum Mol Genet* 14(4):525-33

Musio A, Selicorni A, Focarelli ML, Gervasini C, Milani D, Russo S, Vezzoni P, Larizza L (2006). X-linked Cornelia de Lange syndrome owing to SMC1L1 mutations. *Nat Genet* 38(5):528-30

Nakamura A, Arai H, Fujita N (2009). Centrosomal Aki1 and cohesin function in separase-regulated centriole disengagement. *J Cell Biol* 187, 607–614

Nasmyth K, Haering CH (2005). The structure and function of SMC and kleisin complexes. *Annu Rev Biochem* 74:595-648

Negri G, Milani D, Colapietro P, Forzano F, Della Monica M, Rusconi D, Consonni L, Caffi LG, Finelli P, Scarano G, Magnani C, Selicorni A, Spena S, Larizza L, Gervasini C (2015). Clinical and molecular characterization of Rubinstein-Taybi syndrome patients carrying distinct novel mutations of the EP300 gene. *Clin Genet* 87(2):148-54

Neuwald AF, Hirano T (2000). HEAT repeats associated with condensins, cohesins, and other complexes involved in chromosome-related functions. *Genome Res* 10(10):1445-52

Nilson I, Lochner K, Siegler G, Greil J, Beck JD, Fey GH, Marschalek R (1996). Exon/intron structure of the human ALL-1 (MLL) gene involved in translocations to chromosomal region 11q23 and acute leukemias. *Br J Haematol* 93(4):966-72

Nishiyama T, Ladurner R, Schmitz J, Kreidl E, Schleiffer A, Bhaskara V, Bando M, Shirahige K, Hyman AA, Mechtler K, Peters JM (2010). Sororin mediates sister chromatid cohesion by antagonizing Wapl. *Cell* 143 737–749

Ockeloen CW, Willemsen MH, de Munnik S, van Bon BW, de Leeuw N, Verrips A, Kant SG, Jones EA, Brunner HG, van Loon RL, Smeets EE, van Haelst MM, van Haaften G, Nordgren A, Malmgren H, Grigelioniene G, Vermeer S, Louro P, Ramos L, Maal TJ, van Heumen CC, Yntema HG, Carels CE, Kleefstra T (2014). Further delineation of the KBG syndrome phenotype caused by ANKRD11 aberrations. *Eur J Hum Genet* doi: 10.1038/ejhg.2014.253

Oikawa K, Ohbayashi T, Kiyono T, Nishi H, Isaka K, Umezawa A, Kuroda M, Mukai K (2004). Expression of a novel human gene, human wings apart-like (hWAPL), is associated with cervical carcinogenesis and tumor progression. *Cancer Res* 64(10):3545-9

Oliveira J, Dias C, Redeker E, Costa E, Silva J, Reis Lima M, den Dunnen JT, Santos R (2010). Development of NIPBL locus-specific database using LOVD: from novel mutations to further genotype-phenotype correlations in Cornelia de Lange Syndrome. *Hum Mutat* 31(11):1216-22

Onn I, Heidinger-Pauli JM, Guacci V, Unal E, Koshland DE (2008). Sister Chromatid Cohesion: A Simple Concept with a Complex Reality. *Annu Rev Cell Dev Biol* 24:105-129

Opitz JM, Segal AT, Lehrke R, Nadler H (1964). Brachmann/de Lange syndrome. *Lancet* 2(7367):1019

Parenti I, Rovina D, Masciadri M, Cereda A, Azzollini J, Picinelli C, Limongelli G, Finelli P, Selicorni A, Russo S, Gervasini C, Larizza L (2014). Overall and allele-specific expression of the SMC1A gene in female Cornelia de Lange syndrome patients and healthy controls. *Epigenetics* 9(7):973-9

Parenti I, Gervasini C, Pozojevic J, Graul-Neumann L, Azzollini J, Braunholz D, Watrin E, Wendt KS, Cereda A, Cittaro D, Gillessen-Kaesbach G, Lazarevic D, Mariani M, Russo S, Werner R, Krawitz P, Larizza L, Selicorni A, Kaiser FJ (2015). Broadening of cohesinopathies: exome sequencing identifies mutations in ANKRD11 in two patients with Cornelia de Lange-overlapping phenotype. *Clin Genet* 2015 Feb 4. doi: 10.1111/cge.12564

Pearce PM, Pitt DB (1967). Six cases of de Lange's syndrome; parental consanguinity in two. *Med J Aust* 1(10):502-6

Pié J, Gil-Rodríguez MC, Ciero M, López-Viñas E, Ribate MP, Arnedo M, Deardorff MA, Puisac B, Legarreta J, de Karam JC, Rubio E, Bueno I, Baldellou A, Calvo MT, Casals N, Olivares JL, Losada A, Hegardt FG, Krantz ID, Gómez-Puertas P, Ramos FJ (2010). Mutations and variants in the cohesion factor genes NIPBL, SMC1A, and SMC3 in a cohort of 30 unrelated patients with Cornelia de Lange syndrome. *Am J Med Genet* A 152A(4):924-9

Revenkova E, Focarelli ML, Susani L, Paulis M, Bassi MT, Mannini L, Frattini A, Delia D, Krantz I, Vezzoni P and others (2009). Cornelia de Lange syndrome mutations in SMC1A or SMC3 affect binding to DNA. *Hum Mol Genet* 18(3):418-27

Roelfsema JH, White SJ, Ariyürek Y, Bartholdi D, Niedrist D, Papadia F, Bacino CA, den Dunnen JT, van Ommen GJ, Breuning MH, Hennekam RC, Peters DJ (2005). Genetic heterogeneity in Rubinstein-Taybi syndrome: mutations in both the CBP and EP300 genes cause disease. *Am J Hum Genet* 76(4):572-80

Rollins RA, Morcillo P, Dorsett D (1999). Nipped-B, a Drosophila homologue of chromosomal adherins, participates in activation by remote enhancers in the cut and Ultrabithorax genes. *Genetics* 152(2):577-93

Rubinstein JH, Taybi H (1963). Broad thumbs and toes and facial abnormalities. A possible mental retardation syndrome. *AmJ Dis Child* 105:588-608

Russo S, Masciadri M, Gervasini C, Azzollini J, Cereda A, Zampino G, Haas O, Scarano G, Di Rocco M, Finelli P, Tenconi R, Selicorni A, Larizza L (2012). Intragenic and large NIPBL rearrangements revealed by MLPA in Cornelia de Lange patients. *Eur J Hum Genet* 20(7):734-41

Santen GWE, Kriek M, van Attikum H (2012). SWI/SNF complex in disorder: SWItching from malignancies to intellectual disability. *Epigenetics* 7(11):1219-24

Schmitz J, Watrin E, Lénárt P, Mechtler K, Peters JM (2007). Sororin is required for stable binding of cohesin to chromatin and for sister chromatid cohesion in interphase. *CurrBiol* 17 630–636

Schockel L, Mockel M, Mayer B, Boos D, Stemmann O (2011). Cleavage of cohesin rings coordinates the separation of centrioles and chromatids. *Nat. Cell Biol* 13, 966–972

Schrier SA, Sherer I, Deardorff MA, Clark D, Audette L, Gillis L, Kline AD, Ernst L, Loomes K, Krantz ID, Jackson LG (2011). 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* 155A: 3007-24

Schrier SA, Bodurtha JN, Burton B, Chudley AE, Chiong MA, D'avanzo MG, Lynch SA, Musio A, Nyazov DM, Sanchez-Lara PA, Shalev SA, Deardorff MA (2012). The Coffin-Siris syndrome: a proposed diagnostic approach and assessment of 15 overlapping cases. *Am J Med Genet A* 158A: 1865-76

Schuldiner O, Berdnik D, Levy JM, Wu JS, Luginbuhl D, Gontang AC, Luo L (2008). PiggyBac-based mosaic screen identifies a postmitotic function for cohesin in regulating developmental axon pruning. *Dev Cell* 14(2):227-38

Sekimoto H, Osada H, Kimura H, Kamiyama M, Arai K, Sekiya S (2000). Prenatal findings in Brachmann-de Lange syndrome. *Arch Gynecol Obstet* 263(4):182-4

Selicorni A, Russo S, Gervasini C, Castronovo P, Milani D, Cavalleri F, Bentivegna A, Masciadri M, Domi A, Divizia MT, Sforzini C, Tarantino E, Memo L, Scarano G, Larizza L (2007). Clinical score of 62 Italian patients with Cornelia de Lange syndrome and correlations with the presence and type of NIPBL mutation. *Clin Genet* 72(2):98-108

Sirmaci A, Spiliopoulos M, Brancati F, Powell E, Duman D, Abrams A, Bademci G, Agolini E, Guo S, Konuk B, Kavaz A, Blanton S, Digilio MC, Dallapiccola B, Young J, Zuchner S, Tekin M (2011). Mutations in ANKRD11 cause KBG syndrome, characterized by intellectual disability, skeletal malformations, and macrodontia. *Am J Hum Genet* 89 (2):289–294

Slavin TP, Lazebnik N, Clark DM, Vengoechea J, Cohen L, Kaur M, Konczal L, Crowe CA, Corteville JE, Nowaczyk MJ, Byrne JL, Jackson LG, Krantz ID (2012). Germline mosaicism in Cornelia de Lange syndrome. *Am J Med Genet A* 158A(6):1481-5

Sofueva S, Hadjur S (2012). Cohesin-mediated chromatin interactions-into the third dimension of gene regulation. Brief Funct. Genomics 11(3):205-16

Spena S, Milani D, Rusconi D, Negri G, Colapietro P, Elcioglu N, Bedeschi F, Pilotta A, Spaccini L, Ficcadenti A, Magnani C, Scarano G, Selicorni A, Larizza L, Gervasini C (2014). Insights into genotype-phenotype correlations from CREBBP point mutation screening in a cohort of 46 Rubinstein-Taybi syndrome patients. *Clin Genet* doi: 10.1111/cge.12537. [Epub ahead of print]

Sjögren C, Nasmyth K (2001). Sister chromatid cohesion is required for postreplicative double-strand break repair in Saccharomyces cerevisiae. *Curr Biol* 11(12):991-5

Sjögren C, Ström L (2010). S-phase and DNA damage activated establishment of sister chromatid cohesion–importance for DNA repair. *Exp Cell Res* 316, 1445–1453

Steiner CE, Marques AP (2000). Growth deficiency, mental retardation and unusual facies. *Clin Dysmorphol* 9(2):155-6

Ström L, Lindroos HB, Shirahige K, Sjögren C (2004). Postreplicative recruitment of cohesin to double-strand breaks is required for DNA repair. *Mol Cell* 16(6):1003-15

Sutani T, Kawaguchi T, Kanno R, Itoh T, Shirahige K (2009). Budding yeast Wpl1(Rad61)-Pds5 complex counteracts sister chromatid cohesion-establishing reaction. *Curr Biol* 24 492–497

Taki T, Hayashi Y, Taniwaki M, Seto M, Ueda R, Hanada R, Suzukawa K, Yokota J, Morishita K (1996). Fusion of the MLL gene with two different genes, AF-6 and AF-5alpha, by a complex translocation involving chromosomes 5, 6, 8 and 11 in infant leukemia. *Oncogene* 13(10):2121-30

Tanaka K, Hao Z, Kai M, Okayama H (2001). Establishment and maintenance of sister chromatid cohesion in fission yeast by a unique mechanism. *EMBO J* 20 5779–5790

Tonkin ET, Wang TJ, Lisgo S, Bamshad MJ, Strachan T (2004). 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* 36(6):636-41

Tsou MF, Stearns T (2006). Mechanism limiting centrosome duplication to once per cell cycle. *Nature* 442, 947–951

Tsukahara M, Okamoto N, Ohashi H, Kuwajima K, Kondo I, Sugie H, Nagai T, Naritomi K, Hasegawa T, Fukushima Y, Masuno M, Kuroki Y (1998). Brachmann-de Lange syndrome and congenital heart disease. *Am J Med Genet* 75(4):441-2

Tsurusaki Y, Okamoto N, Ohashi H, Kosho T, Imai Y, Hibi-Ko Y, Kaname T, Naritomi K, Kawame H, Wakui K, Fukushima Y, Homma T, Kato M, Hiraki Y, Yamagata T, Yano S, Mizuno S, Sakazume S, Ishii T, Nagai T, Shiina M, Ogata K, Ohta T, Niikawa N, Miyatake S, Okada I, Mizuguchi T, Doi H, Saitsu H, Miyake N, Matsumoto N (2012). Mutations affecting components of the SWI/SNF complex cause Coffin-Siris syndrome. *Nat Genet* 44(4):376-8

Uhlmann F, Lottspeich F, Nasmyth K (1999). Sister-chromatid separation at anaphase onset is promoted by cleavage of the cohesin subunit Scc1. *Nature* 400(6739):37-42

Uhlmann F (2008). Molecular biology: cohesin branches out. *Nature* 451 777–778

Unal E, Arbel-Eden A, Sattler U, Shroff R, Lichten M, Haber JE, Koshland D (2004). DNA damage response pathway uses histone modification to assemble a double-strand break-specific cohesin domain. *Mol Cell* 16(6):991-1002

Unal E, Heidinger-Pauli JM, Koshland D (2007). DNA double-strand breaks trigger genome-wide sister-chromatid cohesion through Eco1 (Ctf7). *Science* 317(5835):245-8

Vega H, Waisfisz Q, Gordillo M, Sakai N, Yanagihara I, Yamada M, van Gosliga D, Kayserili H, Xu C, Ozono K, Jabs EW, Inui K, Joenje H (2005). Roberts syndrome is caused by mutations in ESCO2, a human homolog of yeast ECO1 that is essential for the establishment of sister chromatid cohesion. Nat Genet 37(5):468-70

Verrotti A, Agostinelli S, Prezioso G, Coppola G, Capovilla G, Romeo A, Striano P, Parisi P, Grosso S, Spalice A, Foiadelli T, Curatolo P, Chiarelli F, Savasta S (2013). Epilepsy in patients with Cornelia de Lange syndrome: a clinical series. *Seizure* 22: 356-9

Yamaguchi K, Ishitobi F (1999). Brain dysgenesis in Cornelia de Lange syndrome. *Clin Neuropathol* 18(2):99-105

Yan J, Saifi GM, Wierzba TH, Withers M, Bien-Willner GA, Limon J, Stankiewicz P, Lupski JR, Wierzba J (2006). Mutational and genotype-phenotype correlation analyses in 28 Polish patients with Cornelia de Lange syndrome. *Am J Med Genet A* 140(14):1531-41

Yan W, Liu S, Xu E, Zhang J, Zhang Y, Chen X, Chen X (2013). Histone deacetylase inhibitors suppress mutant p53 transcription via histone deacetylase 8. *Oncogene* 32:599–609

Yuan B, Pehlivan D, Karaca E, Patel N, Charng WL, Gambin T, Gonzaga-Jauregui C, Sutton VR, Yesil G, Bozdogan ST, Tos T, Koparir A, Koparir E, Beck CR, Gu S, Aslan H, Yuregir OO, Al Rubeaan K, Alnaqeb D, Alshammari MJ, Bayram Y, Atik MM, Aydin H, Geckinli BB, Seven M, Ulucan H, Fenercioglu E, Ozen M, Jhangiani S, Muzny DM, Boerwinkle E, Tuysuz B, Alkuraya FS, Gibbs RA, Lupski JR (2015). Global transcriptional disturbances underlie Cornelia de Lange syndrome and related phenotypes. *J Clin Invest* 125(2):636-51

Yazdi PT, Wang Y, Zhao S, Patel N, Lee EY, Qin J (2002). SMC1 is a downstream effector in the ATM/NBS1 branch of the human S-phase checkpoint. *Genes Dev* 16(5):571-82

Waizenegger IC, Hauf S, Meinke A, Peters JM (2000). Two distinct pathways remove mammalian cohesin from chromosome arms in prophase and from centromeres in anaphase. *Cell* 103 399–410

Wang X, Yang Y, Duan Q, Jiang N, Huang Y, Darzynkiewicz Z, Dai W (2008). SSgo1, a major splice variant of Sgo1, functions in centriole cohesion where it is regulated by Plk1. *Dev Cell* 14, 331–341

Watrin E, Schleiffer A, Tanaka K, Eisenhaber F, Nasmyth K, Peters JM (2006). Human Scc4 is required for cohesin binding to chromatin, sister-chromatid cohesion, and mitotic progression. *Curr Biol* 16(9):863-74

Watrin E and Peters JM (2009). The cohesin complex is required for the DNA damage-induced G2/M checkpoint in mammalian cells. *EMBO J* 28, 1128 2625–2635

Weitzer S, Lehane C, Uhlmann F (2003). A model for ATP hydrolysis-dependent binding of cohesin to DNA, *CurrBiol* 13 1930–1940

Wendt KS, Yoshida K, Itoh T, Bando M, Koch B, Schirghuber E, Tsutsumi S, Nagae G, Ishihara K, Mishiro T, Yahata K, Imamoto F, Aburatani H, Nakao M, Imamoto N, Maeshima K, Shirahige K, Peters JM (2008). Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature* 451(7180):796-801

Westergaard JG, Sinosich MJ, Bugge M, Madsen LT, Teisner B, Grudzinskas JG (1983). Pregnancy-associated plasma protein A in the prediction of early pregnancy failure. *Am J Obstet Gynecol* 145(1):67-9

Wieczorek D, Bögershausen N, Beleggia F, Steiner-Haldenstätt S, Pohl E, Li Y, Milz E, Martin M, Thiele H, Altmüller J, Alanay Y, Kayserili H, Klein-Hitpass L, Böhringer S, Wollstein A, Albrecht B, Boduroglu K, Caliebe A, Chrzanowska K, Cogulu O, Cristofoli F, Czeschik JC, Devriendt K, Dotti MT, Elcioglu N, Gener B, Goecke TO, Krajewska-Walasek M, Guillén-Navarro E, Hayek J, Houge G, Kilic E, Simsek-Kiper PÖ, López-González V, Kuechler A, Lyonnet S, Mari F, Marozza A, Mathieu Dramard M, Mikat B, Morin G, Morice-Picard F, Ozkinay F, Rauch A, Renieri A, Tinschert S, Utine GE, Vilain C, Vivarelli R, Zweier C, Nürnberg P, Rahmann S, Vermeesch J, Lüdecke HJ, Zeschnigk M, Wollnik B (2013). A comprehensive molecular study on Coffin-Siris and Nicolaides-Baraitser syndromes identifies a broad molecular and clinical spectrum converging on altered chromatin remodeling. *Hum Mol Genet* 22(25):5121-35

Wiedemann HR, Kunze J, Grosse FR, Dibbern H (1989). A syndrome of abnormal facies, short stature, and psychomotor retardation. In: Atlas of Clinical Syndromes: A Visual Aid to Diagnosis for Clinicians and Practicing Physicians. *2nd ed. London*, United Kingdom: Wolfe Publishing Ltd; 1989:198–199

Winder SJ, Walsh MP (1993). Calponin: Thin filament-linked regulation of smooth muscle contraction. *Cell Signal* 5:677–686

Wong RW, Blobel G (2008). Cohesin subunit SMC1 associates with mitotic microtubules at the spindle pole. *Proc Natl Acad Sci* USA 105, 15441–15445

Woods SA, Robinson HB, Kohler LJ, Agamanolis D, Sterbenz G, Khalifa M (2014). Exome sequencing identifies a novel EP300 frame shift mutation in a patient with features that overlap Cornelia de Lange syndrome. *Am J Med Genet A* 164A(1):251-8

Zakari M, Yuen K, Gerton JL (2015). Etiology and pathogenesis of the cohesinopathies. *Wiley Interdiscip Rev Dev Biol* doi: 10.1002/wdev.190

Zhang A, Li CW, Chen JD (2007). Characterization of transcriptional regulatory domains of ankyrin repeat cofactor 1. *Biochem Biophys Res Commun* 358: 1034–1040

Zhang J, Shi X, Li Y, Kim BJ, Jia J, Huang Z, Yang T, Fu X, Jung SY, Wang Y, Zhang P, Kim ST, Pan X, Qin J (2008). Acetylation of Smc3 by Eco1 is required for S phase sister chromatid cohesion in both human and yeast. *Mol Cell* 31 143–151

PUBLICATIONS

Parenti I, Gervasini C, Pozojevic J, Graul-Neumann L, Azzollini J, Braunholz D, Watrin E, Wendt KS, Cereda A, Cittaro D, Gillessen-Kaesbach G, Lazarevic D, Mariani M, Russo S, Werner R, Krawitz P, Larizza L, Selicorni A, Kaiser FJ.
Broadening of cohesinopathies: exome sequencing identifies mutations in ANKRD11 in two patients with Cornelia de Lange-overlapping phenotype.
Clin Genet. 2015 Feb 4. doi: 10.1111/cge.12564. [Epub ahead of print]

Bramswig NC, Lüdecke HJ, Alanay Y, Albrecht B, Barthelmie A, Boduroglu K, Braunholz D, Caliebe A, Chrzanowska KH, Czeschik JC, Endele S, Graf E, Guillén-Navarro E, Kiper PÖ, López-González V, **Parenti I**, Pozojevic J, Utine GE, Wieland T, Kaiser FJ, Wollnik B, Strom TM, Wieczorek D.
Exome sequencing unravels unexpected differential diagnoses in individuals with the tentative diagnosis of Coffin-Siris and Nicolaides-Baraitser syndromes.
Hum Genet. 2015 Jun;134(6):553-68

Gil-Rodríguez MC, Deardorff MA, Ansari M, Tan CA, **Parenti I**, Baquero-Montoya C, Ousager LB, Puisac B, Hernández-Marcos M, Teresa-Rodrigo ME, Marcos-Alcalde I, Wesselink JJ, Lusa-Bernal S, Bijlsma EK, Braunholz D, Bueno-Martinez I, Clark D, Cooper NS, Curry CJ, Fisher R, Fryer A, Ganesh J, Gervasini C, Gillessen-Kaesbach G, Guo Y, Hakonarson H, Hopkin RJ, Kaur M, Keating BJ, Kibaek M, Kinning E, Kleefstra T, Kline AD, Kuchinskaya E, Larizza L, Li YR, Liu X, Mariani M, Picker JD, Pié Á, Pozojevic J, Queralt E, Richer J, Roeder E, Sinha A, Scott RH, So J, Wusik KA, Wilson L, Zhang J, Gómez-Puertas P, Casale CH, Ström L, Selicorni A, Ramos FJ, Jackson LG, Krantz ID, Das S, Hennekam RC, Kaiser FJ, FitzPatrick DR, Pié J.
De Novo Heterozygous Mutations in SMC3 Cause a Range of Cornelia de Lange Syndrome-Overlapping Phenotypes.
Hum Mutat. 2015 Apr;36(4):454-62

Braunholz D, Obieglo C, **Parenti I**, Pozojevic J, Eckhold J, Reiz B, Brænne I, Wendt KS, Watrin E, Vodopiutz J, Rieder H, Gillessen-Kaesbach G and Kaiser FJ.
Hidden Mutations in CdLS – Limitations of Sanger Sequencing in Molecular Diagnostics.
Hum Mutat. 2015 Jan;36(1):26-9

Azzollini J, Rovina D, Gervasini C, **Parenti I**, Fratoni A, Cubellis MV, Cerri A, Pietrogrande L, Larizza L.
Functional characterisation of a novel mutation affecting the catalytic domain of MMP2 in siblings with multicentric osteolysis, nodulosis and arthropathy.
J Hum Genet. 2014 Nov;59(11):631-7

Parenti I, Rovina D, Masciadri M, Cereda A, Azzollini A, Picinelli C, Limongelli G, Finelli P, Selicorni A, Russo S, Gervasini C and Larizza L.
Overall and allele-specific expression of the SMC1A gene in female Cornelia de Lange syndrome patients and healthy controls.
Epigenetics 2014 Jul 1;9(7):973-9

Gervasini C*, Russo S*, Cereda A, **Parenti I**, Masciadri M, Azzollini J, Melis D, Aravena T, Doray B, Ferrarini A, Garavelli L, Selicorni A, Larizza L.
Cornelia de Lange Individuals with New and Recurrent SMC1A Mutations Enhance Delineation of Mutation Repertoire and Phenotypic Spectrum.
American Journal of Medical Genetics Part A 2013 Nov;161A(11):2909-19

Gervasini C, **Parenti I**, Picinelli C, Azzollini J, Masciadri M, Cereda A, Selicorni A, Russo S, Finelli P, Larizza L.
Molecular characterization of a mosaic NIPBL deletion in a Cornelia de Lange patient with severe phenotype.
European Journal of Medical Genetics 2013 Mar;56(3):138-43