Published in final edited form as:
Hum Mutat. 2010 January ; 31(1): 5–10. doi:10.1002/humu.21129.

Spectrum and Consequences of *SMC1A* Mutations: The Unexpected Involvement of a Core Component of Cohesin in Human Disease

Linda Mannini[1], **Jinglan Liu**[2], **Ian D. Krantz**[2,3], and **Antonio Musio**[1,4,*]

[1] Istituto di Tecnologie Biomediche, Consiglio Nazionale delle Ricerche, Pisa, Italy

[2] Division of Human Genetics, The Children's Hospital of Philadelphia, PA 19104, USA

[3] The University of Pennsylvania School of Medicine, Philadelphia, PA 19104, USA

[4] Istituto Toscano Tumori, Florence, Italy

Abstract

SMC1A encodes a structural component of the cohesin complex, which isnecessary for sister chromatid cohesion. In addition to its canonical role, cohesin has been shown to be involved in gene expression regulation and maintenance of genome stability. Recently, it has been demonstrated that mutations in the *SMC1A* gene are responsible for Cornelia de Lange syndrome (CdLS). CdLS is a genetically heterogeneous multisystem developmental disorder with variable expressivity, typically characterized by consistent facial dysmorphia, upper extremity malformations, hirsutism, cardiac defects, growth and cognitive retardation, gastrointestinal abnormalities and other systemic involvement. *SMC1A* mutations have also been identified in colorectal cancers. So far a total of 26 different mutations of the *SMC1A* gene have been reported. All mutations reported to date are either missense or small in frame deletions that maintain the open reading frame and presumably result in a protein with residual function. The mutations involve all domains of the protein but appear to cluster in key functional loci. At the functional level, elucidation of the effects that specific *SMC1A* mutations have on cohesin activity will be necessary to understand the etiopathology of CdLS and its possible involvement in tumorigenesis. In this review, we summarize the current knowledge of *SMC1A* mutations.

Keywords

Cohesin; *SMC1A*; Cornelia de Lange Syndrome; Cancer

INTRODUCTION

The finding that Cornelia de Lange Syndrome (CdLS; MIM#s 122470, 300590, 610759) is caused by mutations in a ubiquitous protein complex, and its regulators, which plays a fundamental physiological role during the life of every mammalian cell was surprising to both clinicians involved in the management of this serious disorder and for scientists investigating the role of the cohesin complex. Cohesin's canonical role is in holding sister chromatids together from the time of replication in S phase until their separation in anaphase to ensure proper chromosome segregation during mitosis. The cohesin core complex

*Correspondence to: Antonio Musio, Istituto di Tecnologie Biomediche, Consiglio Nazionale delle Ricerche, Pisa, Italy, Tel: +39 0503152776, Fax: +39 0503153973, antonio.musio@itb.cnr.it.

Supporting Information for this preprint is available from the *Human Mutation* editorial office upon request (humu@wiley.com)

consists of two proteins called SMC1A (for Structural Maintenance of Chromosomes 1A, also known as SMC1L1; MIM# 300040) and SMC3, and two non-SMC proteins known as Scc1/Rad21 and Scc3/SA. SMC proteins are large polypeptides containing 1000–1300 amino acids with a molecular mass ranging from 110 through 170 KDa. They are highly conserved evolutionarily being a critical group of proteins that regulate the structural and functional organization of chromosomes. SMC proteins consist of five distinct motifs, namely N- and C-terminal domains, two coiled-coil domains and the hinge domain. The N- and C-termini, each of about 100–150 amino acids, are highly conserved among SMC proteins. The N-terminal domain contains a nucleoside triphosphate (NTP) binding motif (Walker A box or P-loop), and has been shown to bind ATP. The C-terminal domain contains a DA box, known as Walker B, and binds to DNA. N- and C-terminal domains show homology to proteins from the superfamily of P-loop containing NTPases. Between the N- and C-terminal domains are two long coiled-coil motifs, each about 300–400 amino acids whereas the “hinge” domain of about 150 amino acids joins the two coiled-coil domains. It is fundamental for homo- and hetero-dimerization. The two SMC subunits assume a rod-shaped conformation folding back upon themselves through interaction of the antiparallel coiled-coil domains, and associate with each other through the “hinge” domain positioned at one end of the rod and an ATP-binding “head domain” at the other. The hinge domain is thought to contribute to the flexibility of SMC proteins by allowing the two-armed structure to open or close. The cohesin core complex interacts with several other proteins contributing to its function including NIPBL (homolog of fungal Scc2 and *Drosophila* Nipped-B), required for the loading of cohesin onto chromatin, ESCO2, a human homolog of yeast Eco1, required for the establishment of physical bridges between sisters during S phase, and PDS5A and PDS5B which have been demonstrated to modulate the dynamic association of cohesin with chromatin [reviewed in Hirano, 2006; Peters et al., 2008].

It has been proposed that cohesin associates with chromosomes by trapping DNA within its ring. Cleavage of the Scc1/Rad21 protein at the onset of anaphase releases cohesin from chromosomes allowing chromatid segregation [reviewed in Shintomi and Hirano, 2007]. Recently, an alternative mechanism for cohesin opening that requires the transient dissociation of the SMC1A and SMC3 hinge domains has been suggested indicating that hinges may not just be dimerization domains [Gruber et al., 2006].

In addition to its canonical role in chromosomal cohesion and segregation, several observations suggest that cohesin may have additional roles. First, studies in *Drosophila* demonstrate that cohesin is involved in the regulation of gene expression. This idea arises from the finding that the Nipped-B protein facilitates gene expression regulated by distal enhancers [Misulovin et al., 2008; Rollins et al., 2004]. In addition, in mammals cohesin was found to regulate gene expression through functional interaction with the insulator protein CCCTC binding factor (CTCF), a zinc finger DNA-binding protein that is conserved from invertebrates to man [Parelho et al., 2008; Rubio et al., 2008; Stedman et al., 2008; Wendt et al., 2008]. Second, cohesin factors are involved in DNA repair and genome stability. Indeed, mammalian SMC1A and SMC3 are part of the RC-1 complex, which promotes the repair of DNA gaps and deletions [Jessberger et al., 1993; 1996]. Third, SMC1A participates in a signal-transduction pathway that elicits a checkpoint response to DNA damage [Kim et al., 2002; Yazdi et al., 2002; Musio et al., 2005]. These observations show that cohesin is a versatile complex with diverse functional roles independent of chromatid cohesion. The evolving realizations that cohesin participates in a growing assortment of chromosome and chromatin-related processes suggests that we are just beginning to understand cohesin’s roles in the cell with additional activities and functions undoubtedly remaining to be discovered. Given the critical cellular processes now known to be associated with cohesin it is not surprising that human disorders have recently been found

to be caused by perturbation of cohesin structural components and regulators, including NIPBL, SMC1A, SMC3 and PDS5B (CdLS) and ESCO2 (Roberts syndrome (RBS)). *SMC1A* has also been recently associated with colorectal cancer.

CdLS is a clinically heterogeneous developmental disorder characterized by facial dysmorphia, upper extremity malformations, growth and cognitive retardation, grastrointestinal abnormalities, and other systemic involvement. The prevalence of CdLS is estimated to be 1:10,000–1:40,000 and most cases appear to be sporadic. Pedigree analysis demonstrated dominant inheritance with both maternal and paternal transmission. The clinical features seen in individuals with the classic form of CdLS are striking and easily recognizable, however there is marked variability and milder phenotypes have been described. This phenotypic variability has suggested that different genes could be responsible for CdLS. Two reports showed that mutations in *NIPBL* cause CdLS [Krantz et al., 2004; Tonkin et al., 2004]. These mutations included missense, splice site, nonsense and frameshift. Severe protein truncating mutations creating a “null” allele, are compatible with the hypothesis that the pathogenesis of the disease is due to haploinsufficiency of the NIPBL protein, although a dominant-negative effect of some missense mutations cannot be completely ruled out. A genotype-phenotype correlation has been described, with the severe truncating mutations more likely resulting a more severe clinical phenotype [Gillis et al., 2004]. Heterozygous mutations in *NIPBL* are present in about 60% of CdLS patients [Dorsett and Krantz, 2009]. However, the lack of mutations in about 40% of the patients suggested that CdLS was heterogeneous, a hypothesis also supported by the marked variability of the clinical picture [Gillis et al., 2004]. Mutations in the *SMC1A*, *SMC3* and *PDS5B* genes were subsequently found [Musio et al., 2006; Deardorff et al., 2007; Zhang et al., 2009]. *SMC1A* is responsible for 5% of CdLS cases whereas so far mutations in *SMC3* and *PDS5B* have been identified only in single probands suggesting that they account for a very small portion of CdLS.

As reported above, the *SMC1A* gene is also involved in the maintenance of genome stability. Therefore it was not surprising to find that mutations in this gene have been recently detected in human colorectal cancers [Barber et al., 2008].

In the United States, colorectal cancer is the third mostcommon cancer diagnosed in the population and the second leadingcause of death from cancer. In 2008, it was estimated that 148,810people would be diagnosed with colorectal cancer and 49,960 will diefrom this disease [Jemal et al., 2008]. Colorectal cancer progresses through a series of clinical and histopathologic stages ranging from dysplastic crypts through small benign tumors to malignant cancers. Chromosomal instability is a hallmark of colorectal cancer. This progression is the result of a series of genetic changes that involve activation of oncogenes and inactivation of tumor suppressor genes. Though the molecular basis for chromosomal instability is just beginning to be explored, it has been suggested that chromosomal instability is an early event in tumorigenesis, initiating colorectal cancer. The finding that *SMC1A*, in addition to other cohesin genes, namely *NIPBL*, *SMC3*, *SCC3*, is mutated in colorectal cancer [Barber et al., 2008] may provide a partial explanation for the aneuploidy and enhanced rate of loss of heterozygosity in these tumor cells.

This review is the first mutation update for the human *SMC1A* gene. Here, in particular, we focus on the functional effects of *SMC1A* mutations as well as crucial questions that need to be addressed in the future.

MUTATIONS AND POLYMORPHISMS IN THE *SMC1A* GENE

The *SMC1A* gene maps to Xp11.22-p11.21 in a region that escapes X inactivation and consists of 25 exons. Until now, a total of 26 mutations in the *SMC1A* gene have been

identified (Table 1, Table 2, Supp. Methods). GenBank NM_006306.2 was used as *SMC1A* sequence reference. All mutations are missense or small in-frame deletions that preserve the open reading frame of the gene. Recently, a very large genomic duplication containing *SMC1A* gene has been identified but the clinical manifestations are distinct from those observed in CdLS patients [Yan et al., 2008]. None of these changes have been found in control alleles, the EST databases or are listed in the NCBI Single Nucleotide Polymorphism database (dbSNP; www.ncbi.nlm.nih.gov). Almost all were found to be *de novo* when parental samples were available for testing, with few exceptions including the hallmark family that led to the identification of this gene as a cause of CdLS [Musio et al., 2006;Deardorff et al., 2007]. The finding that protein truncating mutations that disrupt the reading frame of the gene have not been reported suggests that such mutations are likely incompatible with life in males (who have only one copy of the gene) and may result in no phenotype in females who would still have a normal functioning copy of the gene on their other allele. The mutations are scattered along the entire gene. In fact, c.173_187del15, c.397T>G and c.421G>A map to the amino terminal P-loop NTPase domain whereas c.3364T>C and c.3367C>T map to the carboxy terminal P-loop NTPase domain. The mutations c.1478A>C, c.1486C>T and c.1487G>A affect the transition region near the hinge domain. Finally, the mutations c.587G>A, c.802_804del3, 916_918del3, c.1193G>A, c.2046_2048del3, c.2077C>G, c.2131C>T, c.2342G>T, c.2369G>A, c.2446C>G, c.2467T>C, c.2493_2495del3, c.3146G>A and c.3254A>G map to the coiled-coil domains (Table 1). A total of 58 different polymorphisms have been reported. Seven of them are in coding regions and they are all synonymous, except changes in codons 662 and 685. (Table 3). It is worthy to note that 5 out of these 7 mutations map within exon 12. Fifty-one are in non-coding regions, 47 in the introns and 4 in the untranslated regions (Supp. Table S1). Information on these polymorphisms has been derived from public databases (dbSNP; www.ncbi.nlm.nih.gov), published reports (Deardorff et al., 2007) or unpublished data.

GENOTYPE-PHENOTYPE RELATIONSHIP

In order to assess for genotype-phenotype correlations, we compared the clinical features of all *SMC1A*-mutated CdLS patients described to date [Musio et al., 2006; Deardorff et al., 2007; Borck et al., 2007; Liu et al., 2009a]. It is evident that *SMC1A*-mutated patients are characterized by less significant psychomotor and growth retardation than individuals with CdLS caused by mutations in *NIPBL*. In fact most of the affected individuals demonstrated normal birth weights and head circumferences. Furthermore they are characterized by typical but very mild facial dysmorphia, and appear to be spared the limb reduction defects or other major structural anomalies. While the facial features, hirsutism, trend towards smaller stature and head circumference and small hands and feet are consistent amongst all CdLS probands the phenotype in the *SMC1A* probands is striking in the milder expression of these traits and absence of major structural abnormalities. *SMC1A* related CdLS is more similar to *NIPBL* mutated probands who carry missense changes although even this group seems to tend towards a slightly more severe phenotype than the *SMC1A* mutated individuals overall. *SMC1A* mutations, therefore, are associated with a subset of CdLS patients characterized by a milder phenotype.

MUTATIONS IN THE *SMC1A* GENE AND COLORECTAL CANCER

Recently, it has been found that *SMC1A* is one of the five genes containing 11 somatic mutations in a panel that included 132 colorectal cancers [Barber et al., 2008]. In particular, *SMC1A* accounted for 4 out of 11 mutations (36%, Table 3). All mutations are missense: c.1186T>C, c.1300C>T, c.1680C>G and c.3556G>A. The first two mutations map in the coiled-coil domain, the third in the hinge domain and the last in carboxy terminal P-loop NTPase domain.

The remaining seven mutations found in colorectal cancers are in other cohesin genes, *NIPBL, SMC3, SCC3*. Among the 11 identified mutations, four of them are nonsense mutations and they disrupt protein function through the creation of stop codons or frameshift mutations. It is worthy to note that these four mutations involve all cohesin genes analysed, with the exception of *SMC1A*. This observation further confirms that gross gene alterations of *SMC1A* are likely not tolerated in male cells that have only a single copy and are likely to be negatively selected against. In female cells that would have a normal *SMC1A* gene on their other allele, it is not anticipated that a haploinsufficient state would be associated with a phenotype as that cell would have the equivalent of a normal male complement of *SMC1A*.

FUNCTIONAL EFFECTS OF *SMC1A* MUTATIONS

The analysis of mutated residues shows that codon 496 seems to be a mutational hot spot. All mutations that have been investigated had no effect on *SMC1A* expression with the level of SMC1A being comparable between CdLS patients and unaffected subjects [Revenkova et al., 2009; Musio et al., 2006]. The mutation p.Ile560Met in colorectal cancers and p.Glu493Ala, p.Arg496His, and p.Arg496Cys in CdLS map in, or close proximity to, the hinge domain. The hinge domain allows the dimerization of SMC1A and SMC3 proteins. It has been suggested that the loading of cohesin onto chromosomes involves the transient opening of hinge domains [Gruber et al., 2006]. Recently, we demonstrated that SMC1A mutated hinge dimers bind DNA with higher affinity than wild type proteins [Revenkova et al., 2009]. In yeast, a mutation in the hinge domain was shown to negatively affect proper cohesin localization to specific chromatin regions normally enriched with cohesin and abolished establishment of cohesion [Milutinovich et al., 2007]. It is possible that these mutations affect cohesin activity at various different levels. These mutations could alter the entire structure of the cohesin core by preventing correct protein association, affect the opening of the hinge, or prevent the entry of DNA when hinges are transiently opened.

The mutations p.Val58_Arg62del, p.Phe133Val, p.Glu141Lys, p.Phe1122Leu and p.Arg1123Trp in CdLS and p.Val1186Ile in colorectal cancer map to the amino or carboxy terminal P-loop NTPase domain. It is conceivable that they affect hinge opening or dimerization as well. It has been suggested that the hydrolysis of ATP bound to SMC1A and SMC3 heads might provide the energy needed to open the hinge. These mutations could affect any or all of the ATP processes including ATP binding, ATP hydrolysis, and SMC1A/SMC3 head domain dimerization. Several observations support this notion. In fact, the mutational analyses of Walker A and Walker B in the *Bacillus subtilis* SMC protein showed that a mutation in the Walker A motif abolishes ATP binding whereas a mutation in the Walker B motif allows ATP binding but blocks head-head engagement and ATP hydrolysis [Hirano et al., 2001].

The mutations p.Arg196His, p.Lys268del, p.Ser306del, p.Arg398Gln, p.Glu683del, p.Arg693Gly, p.Arg711Trp, p.Cys781Phe, p.Arg790Gln, p.Arg816Gly, p.Phe823Leu, p.Asp831_Gln832delinsGlu, p.Arg1049Gln, p.Tyr1085Cys in CdLS and p.Phe396Leu, and p.Arg434Trp in colorectal cancers map to the coiled-coil domain. Because SMC proteins form heterodimers, coiled-coil interactions are important for the correct folding of an SMC monomer and thus crucial for the formation of the head domain. In addition, it has been proposed that coiled-coil domains make contact with DNA and/or may be necessary for protein-protein interactions. Therefore mutations in coiled-coil domains may affect the hinge activity or disrupt the interaction with the other cohesin subunits.

Most probands found to have an *SMC1A* mutation are females [Deardorff et al., 2007] and it is expected that they have a functional wild type protein produced from their other allele as *SMC1A* escapes X inactivation [Brown et al., 1995]. This discrepancy can be explained by a

dominant negative effect of the altered protein. Mutated SMC1A proteins can assemble with native cohesin subunits to produce an inactive, or functionally restricted, cohesin complex. In theory, a single mutant can poison an entire complex, through a dominant-negative mode of action.

CLINICAL AND DIAGNOSTIC RELEVANCE

The identification of mutations in the *SMC1A* gene has provided an explanation for a significant cohort of individuals with mildly manifesting CdLS. Before molecular testing, a diagnosis of CdLS was based purely on the phenotypic features of the patients. *SMC1A* molecular testing has significantly improved the diagnostic approach to CdLS, and has allowed for more precise genetic counseling and prenatal diagnosis. It has also expanded our understanding of the clinical range seen in CdLS and has led to the suggestion that the mild extreme of the CdLS phenotype may blend in with apparent isolated mental retardation. This suggests that the organ most sensitive to cohesin disruption is the brain, given the one constant phenotypic finding in all cohesin mutant individuals is some degree of mental retardation even in the absence of other striking physical features. Hence, normal growth parameters at birth and/or the absence of major malformation typical of CdLS, in a child with mild to moderate cognitive delays and suggestive facial features might warrant the screening of *SMC1A* gene at first.

CONCLUSION AND FUTURE PROSPECTS

The discovery of SMC proteins [Strunnikov et al., 1993] and cohesin [Michaelis et al., 1997] provided the framework from which the puzzle of how the genome is segregated amongst daughter cells during meiosis and mitosis has begun to be elucidated. Recent studies have shed new light on the mechanistic action of the cohesin complex and has also expanded its known repertoire of cellular functions. The discovery that mutations in the *SMC1A* gene is responsible for CdLS [Musio et al., 2006] and the ability to clinically screen for alterations in this gene in CdLS probands and individuals with subtle features of CdLS disease has increased our understanding of the phenotypic spectrum associated with this disease disorder. In addition, the finding that *SMC1A* mutations have also been detected in colorectal cancers [Barber et al., 2008] coupled with its known involvement in genome stability [Kim et al., 2005; Musio et al., 2003; 2005] and gene expression regulation [Peric-Hupkes and van Steensel, 2008; Liu et al., 2009b] has provided some clues to the functions of *SMC1A*. A number of basic questions remain to be answered, however.

SMC1A and cancer

Cancer cells can variably display gross chromosome alterations and aneuploidy. There is ample experimental evidence that variation in nuclear DNA content and chromosomal aneuploidies may occur at early stages of tumorigenesis and are invariably observed in carcinomas. According to this view, by gaining or losing entire chromosomes, aneuploidy could simulate both overexpression and loss of specific loci relevant to cell growth, thus providing the substrate for selection of more aggressively growing cells. Our knowledge of the molecular mechanisms underlying chromosomal instability is extremely limited. The finding that colorectal cancers have mutations in the *SMC1A* gene and other cohesin related genes (*NIPBL, SMC3, SCC3*), involved in sister chromatid cohesion suggests that cohesin genes are mutational targets whose disruption leads to chromosomal instability [Barber et al., 2008]. Though this hypothesis is very intriguing, mutational screening of *SMC1A* in a larger sample of diverse tumors is needed to determine if *SMC1A* mutations, or cohesin disruption in general, are common events in cancer, particularly in those with significant chromosomal instability and aneuploidy, or whether they are restricted to colorectal cancers. How early during cancer development *SMC1A* mutations occur is another critical question.

To gain new insights into this process, the *SMC1A* mutational screening in precancerous lesions will be also necessary. *SMC1A* mutations can contribute to chromosomal instability by two different ways. The first one is directly linked to the canonical role of cohesin. In fact, mutations can affect correct chromosome segregation leading to chromosome imbalance with chromosome loss or gain. The second one comes from the recent findings that cohesin is involved in gene expression regulation with *SMC1A* mutations altering the expression of proto-oncogenes or tumor suppresor genes leading to tumorigenesis. It is worthy to note that some CdLS patients developed metaplasia of the esophagus, Barrett esophagus and esophageal adenocarcinoma, suggesting a genetic predisposition to esophageal alterations, although this is complicated by the chronic and often under-ascertained diagnosis of gastroesophageal reflux which is thought to be near universal in CdLS [Kline et al., 2007].

SMC1A and CdLS

Since the identified mutations do not affect SMC1A synthesis [Musio et al., 2006; Revenkova et al., 2009], we speculate that the presence of a mutated SMC1A subunit is able to alter correct cohesin complex function in a dominant-negative manner. Recently, we demonstrated that several mutations in *SMC1A* gene causing CdLS affect the association of SMC hinge dimers with DNA [Revenkova et al., 2009] suggesting a potential role of SMC proteins in remodeling chromatin architecture. Cohesin binds DNA with residence half-lives ranging from several minutes to a few hours [Kueng et al., 2006; Gerlich et al., 2006]. What is the effect of mutations mapping in protein domains other than the hinge domain on the affinity of cohesin for DNA? Has cohesin stability, and consequently its turn over, changed in the presence of mutated subunits? These are critical questions that need to be addressed. An increasing number of manuscripts report on the varying ways that cohesin acts to regulate gene expression in human and vertebrate cells. Four groups mapped cohesin-binding sites in mammalian cells, and found that most of CTCF-binding sites bind cohesion [reviewed in Gause et al., 2008]. CTCF is well known for its role at transcriptional insulators. Cohesin could stabilize long-range chromosomal interactions mediated by CTCF. It has been suggested that CTCF participates in the formation of chromosome loops that facilitate some enhancer promoter interactions and prevent others [Parelho et al., 2008; Rubio et al., 2008; Stedman et al., 2008; Wendt et al., 2008]. *SMC1A* mutations can affect the dynamic association between the cohesin complex and chromatin, leading to gene expression alteration. To investigate this possibility, genome wide studies and the identification of cohesin binding sites by ChIP in cells from CdLS probands with *SMC1A* mutations will also be necessary. The availability of an *Smc1a* mouse model would offer new approaches that may help in understanding these roles of cohesin and provide insights into CdLS pathogenesis.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work is partially funded by grants from Consiglio Nazionale delle Ricerche (CNR), Fondazione Cassa di Risparmio delle Province Lombarde (CARIPLO), Istituto Superiore di Sanità, to A.M. I.D.K. is supported by PO1 HD052860 and R21 HD050538, NICHD.

References

Barber TD, McManus K, Yuen KW, Reis M, Parmigiani G, Shen D, Barrett I, Nouhi Y, Spencer F, Markowitz S, Velculescu VE, Kinzler KW, Vogelstein B, Lengauer C, Hieter P. Chromatid

cohesion defects may underlie chromosome instability in human colorectal cancers. Proc Natl Acad Sci U S A. 2008; 105:3443–3448. [PubMed: 18299561]

Borck G, Zarhrate M, Bonnefont JP, Munnich A, Cormier-Daire V, Colleaux L. Incidence and clinical features of X-linked Cornelia de Lange syndrome due to SMC1L1 mutations. Hum Mutat. 2007; 28:205–206. [PubMed: 17221863]

Brown CJ, Miller AP, Carrel L, Rupert JL, Davies KE, Willard HF. The DXS423E gene in Xp11.21 escapes X chromosome inactivation. Hum Mol Genet. 1995; 4:251–255. [PubMed: 7757075]

Deardorff MA, Kaur M, Yaeger D, Rampuria A, Korolev S, Pie J, Gil-Rodriguez C, Arnedo M, Loyes B, Kline A, Wilson M, Lillquist K, Siu V, Ramos FJ, Musio A, Jackson LS, Dorsett D, Krantz I. Mutations in cohesin complex members SMC3 and SMC1A cause a mild variant of Cornelia de Lange syndrome with predominant mental retardation. Am J Hum Genet. 2007; 80:485–494. [PubMed: 17273969]

Dorsett D, Krantz ID. On the molecular etiology of Cornelia de Lange syndrome. Ann N Y Acad Sci. 2009; 1151:22–37. [PubMed: 19154515]

Gause M, Schaaf CA, Dorsett D. Cohesin and CTCF: cooperating to control chromosome conformation? Bioessays. 2008; 30:715–718. [PubMed: 18623068]

Gerlich D, Koch B, Dupeux F, Peters JM, Ellenberg J. Live-cell imaging reveals a stable cohesin-chromatin interaction after but not before DNA replication. Curr Biol. 2006; 16:1571–1578. [PubMed: 16890534]

Gillis LA, McCallum J, Kaur M, DeScipio C, Yaeger D, Mariani A, Kline AD, Li HH, Devoto M, Jackson LG, Krantz ID. NIPBL mutational analysis in 120 individuals with Cornelia de Lange syndrome and evaluation of genotype-phenotype correlations. Am J Hum Genet. 2004; 75:610–623. [PubMed: 15318302]

Gruber S, Arumugam P, Katou Y, Kuglitsch D, Helmhart W, Shirahige K, Nasmyth K. Evidence that loading of cohesin onto chromosomes involves opening of its SMC hinge. Cell. 2006; 127:523–537. [PubMed: 17081975]

Hirano T. The heart of the chromosome: SMC proteins in action. Nat Rev Mol Cell Biol. 2006; 7:311–322. [PubMed: 16633335]

Hirano M, Anderson DE, Erickson HP, Hirano T. Bimodal activation of SMC ATPase by intra- and inter-molecular interactions. EMBO J. 2001; 20:3238–3250. [PubMed: 11406600]

Jemal A, Siegel R, Ward E, Hao Y, Xu J, Murray T, Thun MJ. Cancer statistics. CA Cancer J Clin. 2008; 58:71–96. [PubMed: 18287387]

Jessberger R, Podust V, Hubscher U, Berg P. A mammalian protein complex that repairs double-strand breaks and deletion by recombination. J Biol Chem. 1993; 268:15070–15079. [PubMed: 8392064]

Jessberger R, Riwar B, Baechtold H, Akhmedov AT. SMC proteins constitute two subunits of the mammalian recombination complex RC-1. Embo J. 1996; 15:4061–4068. [PubMed: 8670910]

Kim ST, Xu B, Kastan MB. Involvement of the cohesin protein, Smc1, in Atm-dependent and independent responses to DNA damage. Genes Dev. 2002; 16:560–570. [PubMed: 11877376]

Kline AD, Grados M, Sponseller P, Levy HP, Blagowidow N, Schoedel C, Rampolla J, Clemens DK, Krantz I, Kimball A, Pichard C, Tuchman D. Natural history of aging in Cornelia de Lange syndrome. Am J Med Genet. 2007; 145C:248–260. [PubMed: 17640042]

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. Cornelia de Lange syndrome is caused by mutations in NIPBL, the human homolog of Drosophila melanogaster Nipped-B. Nat Genet. 2004; 36:631–635. [PubMed: 15146186]

Kueng S, Hegemann B, Peters BH, Lipp JJ, Schleiffer A, Mechtler K, Peters JM. Wapl controls the dynamic association of cohesin with chromatin. Cell. 2006; 127:955–967. [PubMed: 17113138]

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. SMC1A expression and mechanism of pathogenicity in probands with X-linked Cornelia de Lange syndrome. Hum Mutat. 2009a (in press).

Liu J, Zhang Z, Bando M, Itoh T, Deardorff MA, Clark D, Kaur M, Tandy S, Kondoh T, Rappaport E, Spinner NB, Vega H, Jackson LG, Shirahige K, Krantz ID. Transcriptiona dysregulation in NIPBL and cohesin mutant human cells. PLoS Biol. 2009b; 7(5):e 1000119.

Michaelis C, Ciosk R, Nasmyth K. Cohesins: chromosomal proteins that prevent premature separation of sister chromatids. Cell. 1997; 91:35–45. [PubMed: 9335333]

Milutinovich M, Unal E, Ward C, Skibbens RV, Koshland D. A multi-step pathway for the establishment of sister chromatid cohesion. PLoS Genet. 2007; 3:e12. [PubMed: 17238288]

Misulovin Z, Schwartz YB, Li XY, Kahn TG, Gause M, Macarthur S, Fay JC, Eisen MB, Pirrotta V, Biggin MD, Dorsett D. Association of cohesin and Nipped-B with transcriptionally active regions of the Drosophila melanogaster genome. Chromosoma. 2008; 117:89–102. [PubMed: 17965872]

Musio A, Montagna C, Zambroni D, Indino E, Barbieri O, Citti L, Villa A, Ried T, Vezzoni P. Inhibition of BUB1 results in genomic instability and anchorage-independent growth of normal human fibroblasts. Cancer Res. 2003; 63:2855–2863. [PubMed: 12782591]

Musio A, Montagna C, Mariani T, Tilenni M, Focarelli ML, Brait L, Indino E, Benedetti PA, Chessa L, Albertini A, Ried T, Vezzoni P. Smc1 involvement in fragile site expression. Hum Mol Genet. 2005; 14:525–533. [PubMed: 15640246]

Musio A, Selicorni A, Focarelli ML, Gervasini C, Milani D, Russo S, Vezzoni P, Larizza L. X-linked Cornelia de Lange syndrome owing to SMC1L1 mutations. Nat Genet. 2006; 38:520–530.

Parelho V, Hadjur S, Spivakov M, Leleu M, Sauer S, Gregson HC, Jarmuz A, Canzonetta C, Webster Z, Nesterova T, Cobb BS, Yokomori K, Dillon N, Aragon L, Fisher AG, Merkenschlager M. Cohesins functionally associate with CTCF on mammalian chromosome arms. Cell. 2008; 132:422–433. [PubMed: 18237772]

Peric-Hupkes D, van Steensel B. Linking cohesin to gene regulation. Cell. 2008; 132:925–928. [PubMed: 18358805]

Peters JM, Tedeschi A, Schmitz J. The cohesin complex and its roles in chromosome biology. Genes Dev. 2008; 22:3089–3114. [PubMed: 19056890]

Revenkova E, Focarelli ML, Susani L, Paulis M, Bassi MT, Mannini L, Frattini A, Delia D, Krantz I, Vezzoni P, Jessberger R, Musio A. Cornelia de Lange syndrome mutations in SMC1A or SMC3 affect binding to DNA. Hum Mol Genet. 2009; 18:418–427. [PubMed: 18996922]

Rollins RA, Korom M, Aulner N, Martens A, Dorsett D. Drosophila nipped-B protein supports sister chromatid cohesion and opposes the stromalin/Scc3 cohesion factor to facilitate long-range activation of the cut gene. Mol Cell Biol. 2004; 24:3100–3111. [PubMed: 15060134]

Rubio ED, Reiss DJ, Welcsh PL, Disteche CM, Filippova GN, Baliga NS, Aebersold R, Ranish JA, Krumm A. CTCF physically links cohesin to chromatin. Proc Natl Acad Sci U S A. 2008; 105:8309–8314. [PubMed: 18550811]

Shintomi K, Hirano T. How are cohesin rings opened and closed? Trends Biochem Sci. 2007; 32:154–157. [PubMed: 17320400]

Stedman W, Kang H, Lin S, Kissil JL, Bartolomei MS, Lieberman PM. Cohesins localize with CTCF at the KSHV latency control region and at cellular *c-myc* and *H19/Igf2* insulators. EMBO J. 2008; 27:654–666. [PubMed: 18219272]

Strunnikov AV, Larionov VL, Koshland D. SMC1: an essential yeast gene encoding a putative head-rod-tail protein is required for nuclear division and defines a new ubiquitous protein family. J Cell Biol. 1993; 123:1635–48. [PubMed: 8276886]

Tonkin ET, Wang TJ, Lisgo S, Bamshad M, Strachan T. NIPBL, encoding a homolog of fungal Scc2-type sister chromatid cohesin proteins and fly Nipped-B, is mutated in Cornelia de Lange syndrome. Nat Genet. 2004; 36:636–341. [PubMed: 15146185]

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. Cohesin mediates transcriptional insulation by CCCTC-binding factor. Nature. 2008; 451:796–801. [PubMed: 18235444]

Yan J, Zhang F, Brundage E, Scheuerle A, Lanpher B, Erickson RP, Powis Z, Robinson HB, Trapane PL, Stachiw-Hietpas D, Keppler-Noreuil KM, Lalani SR, Sahoo T, Chinault AC, Patel A, Cheung SW, Lupski JR. Genomic duplication resulting in increased copy number of genes encoding the

sister chromatid cohesion complex conveys clinical consequences distinct from Cornelia de Lange. J Med Genet. 2008 (in press).

Yazdi PT, Wang Y, Zhao S, Patel N, Lee EY, Qin JS. SMC1 is a downstream effector in the ATM/NBS1 branch of the human S-phase checkpoint. Genes Dev. 2002; 16:571–582. [PubMed: 11877377]

Zhang B, Chang J, Fu M, Huang J, Kashyap R, Salavaggione E, Jain S, Shashikant K, Deardorff MA, Uzielli ML, Dorsett D, Beebe DC, Jay PY, Heuckeroth RO, Krantz I, Milbrandt J. Dosage effects of cohesin regulatory factor PDS5 on mammalian development: implications for cohesinopathies. PLoS One. 2009; 4:e5232. [PubMed: 19412548]

TABLE 1

Mutational spectrum of the *SMC1A* gene in CdLS

Nucleotide changes*	Effect on amino acid	Protein domain	First description
c.173_187del15	p.Val58-Arg62del	Amino terminal P-loop NTPase domain	Deardorff et al. [2007]
c.397T>G	p.Phe133Val	Amino terminal P-loop NTPase domain	Deardorff et al., [2007]
c.421G>A	p.Glu141Lys	Amino terminal P-loop NTPase domain	Liu et al., [2009]
c.587G>A	p.Arg196His	Coiled-coil	Borck et al., [2007]
c.802_804del3	p.Lys268del	Coiled-coil	Liu et al., [2009]
c.916_918del3	p.Ser306del	Coiled-coil	Liu et al., [2009]
c.1193G>A	p.Arg398Gln	Coiled-coil	Liu et al., [2009]
c.1478A>C	p.Glu493Ala	Transition region	Musio et al., [2006]
c.1486C>T	p.Arg496Cys	Transition region	Deardorff et al.,[2007]
c.1487G>A	p.Arg496His	Transition region	Deardorff et al., [2007]
c.2046_2048del3	p.Glu683del	Coiled-coil	Liu et al., [2009]
c.2077C>G	p.Arg693Gly	Coiled-coil	Liu et al., [2009]
c.2131C>T	p.Arg711Trp	Coiled-coil	Deardorff et al.[2007]
c.2342G>T	p.Cys781Phe	Coiled-coil	Liu et al., [2009]
c.2369G>A	p.Arg790Gln	Coiled-coil	Deardorff et al. [2007]
c.2446C>G	p.Arg816Gly	Coiled-coil	Liu et al., [2009]
c.2467T>C	p.Phe823Leu	Coiled-coil	This report
c.2493_2495del3	p.Asp831-Gln832del	Coiled-coil	Musio et al., [2006]
c.3146G>A	p.Arg1049Gln	Coiled-coil	Liu et al., [2009]
c.3254A>G	p.Tyr1085Cys	Coiled-coil	Borck et al., [2007]
c.3364T>C	p.Phe1122Leu	Carboxy terminal P-loop NTPase domain	Deardorff et al., [2007]
c.3367C>T	p.Arg1123Trp	Carboxy terminal P-loop NTPase domain	Liu et al., [2009]

*Nucleotide numbering reflects cDNA numbering with +1 corresponding to the A of the ATG translation initiation codon in the reference sequence NM_006306.2, according to journal guidelines (www.hgvs.org/mutnomen). The initiation codon is codon 1.

TABLE 2

Mutational spectrum of the *SMC1A* gene in colorectal cancer

Nucleotide changes*	Effect on amino acid	Protein domain	First description
c.1186T>C	p.Phe396Leu	Coiled-coil	Barber et al., [2008]
c.1300C>T	p.Arg434Trp	Coiled-coil	Barber et al., [2008]
c.1680C>G	p.Ile560Met	Hinge	Barber et al., [2008]
c.3556G>A	p.Val1186Ile	Carboxy terminal P-loop NTPase domain	Barber et al., [2008]

*Nucleotide numbering reflects cDNA numbering with +1 corresponding to the A of the ATG translation initiation codon in the reference sequence NM_006306.2, according to journal guidelines (www.hgvs.org/mutnomen). The initiation codon is codon 1.

TABLE 3

Polymorphisms in the *SMC1A* gene

Exon	Nucleotide change*	Codon change	Reference
10	c.1698G>A	Glu566Glu	NCBI
12	c.1986G>C	Lys662Asn	NCBI
12	c.1986G>A	Lys662Lys	This report
12	c.1993C>G	Arg665 Gly	NCBI
12	c.2052G>A	Glu684Glu	NCBI
12	c.2053C>G	Leu685Val	NCBI
25	c.3663C>G	Thr1221Thr	NCBI

*Nucleotide numbering reflects cDNA numbering with +1 corresponding to the A of the ATG translation initiation codon in the reference sequence NM_006306.2, according to journal guidelines (www.hgvs.org/mutnomen). The initiation codon is codon 1.