



HAL
open science

Gene regulation and chromatin organization: relevance of cohesin mutations to human disease

Erwan Watrin, Frank J Kaiser, Kerstin S Wendt

► To cite this version:

Erwan Watrin, Frank J Kaiser, Kerstin S Wendt. Gene regulation and chromatin organization: relevance of cohesin mutations to human disease. *Current Opinion in Genetics and Development*, 2016, 37, pp.59–66. 10.1016/j.gde.2015.12.004 . hal-01263104

HAL Id: hal-01263104

<https://univ-rennes.hal.science/hal-01263104>

Submitted on 1 Feb 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**Gene regulation and chromatin organization:
Relevance of cohesin mutations to human disease**

Erwan Watrin^{1,2}, Frank J. Kaiser³ and Kerstin S. Wendt^{4,*}

¹ Centre National de la Recherche Scientifique, UMR 6290, Rennes, France.

² Institut de Génétique et Développement de Rennes, Université de Rennes 1, SFR BIOSIT, Rennes, France.

³ Sektion für Funktionelle Genetik am Institut für Humangenetik Lübeck, Universität zu Lübeck, Lübeck, Germany

⁴ Department of Cell Biology, Erasmus MC, Rotterdam, The Netherlands

* Corresponding author: Kerstin S. Wendt, k.wendt@erasmusmc.nl

Department of Cell Biology, Erasmus MC, Wytemaweg 80, 3015CN Rotterdam, The Netherlands. Phone: +31-10-7044007

Word count: 2250 words (without abstract, references and figure legends)

Accepted manuscript

Abstract (128 words)

Consistent with the diverse roles of the cohesin complex in chromosome biology, mutations in genes encoding cohesin and its regulators are found in different types of cancer and in developmental disorders such as Cornelia de Lange Syndrome. It is so far considered that the defects caused by these mutations result from altered function of cohesin in regulating gene expression during development. Chromatin conformation analyses have established the importance of cohesin for the architecture of developmental gene clusters and *in-vivo* studies in mouse and zebrafish demonstrated how cohesin defects lead to gene misregulation and to malformations similar to the related human syndromes. Here we present our current knowledge on cohesin's involvement in gene expression, highlighting molecular and mechanistic consequences of pathogenic mutations in the Cornelia de Lange syndrome.

Introduction

The cohesin complex (Figure 1) was originally discovered as protein complex required for sister chromatid cohesion (SCC), a function which is conserved in all eukaryotes [1•]. Within the past eight years cohesin has emerged as an essential actor in virtually all aspects of chromosome biology, including chromosome segregation, maintenance of genome stability, regulation of gene expression, chromatin structure and 3D genome organisation.

Cohesin's ability to take part in these distinct pathways arises from its unique ability to create topological links between two segments of the chromatin fibre. The topological link between two sister chromatids confers sister chromatid cohesion (Figure 2A). The same interaction is established at DNA double strand break sites and hold the sister strand in place as repair template (Figure 2B). Interactions between two regions of the same chromatid forms chromatin loops that built up the overall 3D topology of the chromatin fibre and mediate communication between genes and their regulatory elements through highly organised chromatin structure (Figures 2C-D).

Although the underlying molecular mechanisms may appear similar, these distinct functions of cohesin differ in important aspects, and the regulation of cohesin's "looping function" is not understood at all.

Mutations in genes encoding cohesin subunits or its regulators (sometimes collectively referred to as the cohesion apparatus) have been shown to cause various human disorders, in particular cancers (for review [2]) and developmental syndromes that are discussed below.

Recent developments in chromatin conformation capture techniques have allowed describing chromatin folding at different scales, mapping binding sites of important factors (including cohesin) and correlating site occupancy by cognate factors to chromatin organisation and gene activity.

In this review we will summarize the current mechanistic insights as to how defects in cohesin and its regulators relate to pathology, with a particular focus on how altered chromatin architecture may participate in developmental disorders.

Cohesin and cohesion-related functions

Cohesin mediates sister chromatid cohesion from DNA replication in S-phase until chromosomes segregate during mitosis (reviewed in [3])(Figure 3). In interphase SCC participates in maintaining genome stability by assisting recombinational repair of DNA double strand-breaks [2,4]. In mitosis and meiosis cohesion ensures correct distribution of chromosomes and, therefore faithful transmission of genome from one generation of cells or individual to the next (reviewed in [5]). The last twenty years of

research have uncovered numerous cohesin functional partners and regulators that act all along the cell cycle (Figures 1 and 3) and include cohesin modifying (Esco1 and Esco2, Aurora B and Plk1) and de-modifying (PP2A, HDAC8) enzymes, chromatin loader (NIPBL-MAU2 complex) as well as cohesive and anti-cohesive factors (SA1 and 2, Sororin, Pds5A and B, Wapl, Sgo1L) that regulate and directly bind to cohesin complexes in a cell cycle regulated manner (Figure 3, for a detailed review on cohesin regulation see [6]).

Beside roles related to its cohesive property, cohesin is also essential for proper functioning of the DNA damage checkpoint pathway [4], a surveillance mechanism that halts cell cycle progression until all DNA damage has been repaired. More recently, cohesin and its regulators have been involved in centrosome biology. Recent findings suggested that cohesin hold mother and daughter centrioles and pericentriolar material through mechanisms and regulations that resemble that of chromatid cohesion [7,8] although underlying molecular details are still elusive [9,10].

In addition to processes ensuring genome distribution and stability, cohesin also participates in other aspects of chromosome biology, which strikingly all impact on gene expression.

Cohesin's role for the regulation of gene expression

Along the vertebrate genome cohesin localizes in interphase to pericentric heterochromatin [11] and to discrete sites at chromosome arms [12-14]. The number of sites varies from about 10,000 to 60,000 depending on cell type, antibodies, technology platform and analysis pipeline [12,15-17]. In all studies however, about 50 % of the cohesin sites localize in intergenic region, 37% in introns and exons and about 7% within 5kb upstream of promoters [12]. Notably, around 60 % to 80 % of the cohesin binding sites are co-occupied by the chromatin insulator CTCF [12-14,16,17].

Cohesin sites without CTCF (cohesin/no CTCF) tend to be enriched at transcription start sites and are found to colocalize with cell-type specific transcription factors like estrogen receptor in MCF7 cells, or HNFA4 and CEBPA in HepG2 cells [16]. In mouse ES cells cohesin/no CTCF sites correlate with binding sites of the cohesin loader NIPBL and the mediator complex [17]. Also the transcription factor KIF4A was shown to be necessary to recruit cohesin to interacting regions [18].

Cohesin/CTCF sites are involved in long-range interactions (Figure 2 and 4) of the chromatin fibre, thereby mediating contacts between promoter and distal enhancers and gene activation at a number of developmentally important loci, for example the apolipoprotein gene cluster [19], the imprinted gene cluster H19/IFG2 that controls embryonic growth [20], the IFNG locus [21], the β -globin locus [22][23]

and the protocadherin locus [24]. Overall, cohesin/CTCF sites appear to mediate constitutive chromatin interactions [25•].

Cohesin/no CTCF sites also mediate long range interactions and transcriptional activation, likely depending on the co-binding of specific additional transcription factors. In mouse ES cells cohesin and the mediator complex mediate promoter-enhancer interactions of important pluripotency genes (OCT4, NANOG) [17,26,27] and are, therefore, important for reprogramming and maintenance of the pluripotency state.

Cohesin-mediated looping between promoter and enhancer mostly occurs in a tissue-specific fashion, although interactions between promoters and distal regions can also be conserved between tissues, the tissue-specific gene activity being then regulated by different mechanism [28]. Besides promoter-enhancer interactions, cohesin is also involved in larger scale chromosomal rearrangements, like T-cell receptor rearrangement [29], V(D)J recombination and the genomic architecture of the IgH locus [30].

Genome-wide chromatin interaction maps of the fly, mouse and human genomes [31,32••] provided by the Hi-C technique [33••] show an organisation of the chromatin fibre into discrete domains, termed topological domains (TADs) [32••]. Noteworthy, both cohesin and CTCF are enriched at boundaries between TADs, suggesting a role of these proteins for defining the TAD. Subsequent cohesin inactivation studies in cycling cells and postmitotic cells have established the importance of cohesin [15,34,35] and CTCF [15•] for the chromatin interactions forming the TADs, possibly by assembling clusters of chromatin loops (Figure 2E). Interestingly, the TAD boundaries seem not to depend on cohesin since only the depletion of CTCF but not cohesin depletion leads to increased interactions between neighbouring domains that indicates weakening of the TAD boundaries, [15•]. Genes misregulated after cohesin depletion were correlated with cohesin binding sites and long-range interactions, revealing that these genes are more likely to have lost the contact to their respective enhancer than other genes [15•].

These studies have also stressed how essential cohesin-mediated chromatin conformations are for developmental gene clusters to allow timely access to cell-type specific enhancers to drive development-specific gene expression patterns. Examples are the *HOX* genes that encode master transcription factors responsible for patterning the anterior-posterior body axis during development. The best studied *HOX* clusters, *HOXA* and *HOXD*, are expressed in the developing limb and reside in a different TAD than their respective enhancer and the TAD boundary changes during development to provide the characteristic “proximo-distal patterning” of *HOX* gene expression [36]. Inactivation of cohesin in HEK293T cells disturbs this 3D-organization and leads to *HOX* gene misregulation [15,37]. The

relevance of these regulatory connections *in vivo* has been addressed in vertebrate models. Depletion of the cohesin loading complex NIPBL in the zebrafish pectoral fin (forelimb) and in the mouse limb buds gives rise to limb defects, a common condition in individuals with CdLS. In the zebrafish model dysregulated expression of early limb development genes, including multiple *hox* genes, was observed and a change of chromatin interactions at the HOXD cluster confirmed by DNA-FISH. Similar genes are affected in the NIPBL-haploinsufficient mice [38]. A detailed description of the phenotypes observed in the different available mouse models is provided by Sing and Gerton [39].

Cohesin-related pathologies

In increasing numbers of cancer types mutations in the cohesin complex or its regulators are found. For example, mutations in the cohesin complex subunits STAG2, RAD21, SMC1 and SMC3 were observed in different types of leukaemia [40,41] and mutations of the cohesin subunit STAG2 have been linked to bladder cancer [42,43] and colorectal cancer [44]. In addition, changes in protein levels of cohesion apparatus components were linked to cancer, for instance increased RAD21 levels have been associated with poor prognosis in breast cancers [45]. Only very recently also individual genomic cohesin binding sites came into focus since they were found to be frequently mutated in cancers [46]. It is still very unclear how the cohesin complex might be implicated in the different cancers. Defects in the cohesin complex might mainly lead to genome instability caused by defective DNA repair, as well as to chromosome instability and aneuploidy through chromosome missegregation. However, patients suffering from CdLS carry mutations in several cohesin subunits and regulators but do not show increased genomic instability. It was also noted that cancers with cohesin complex mutations do not necessarily present aberrant karyotypes [40,47].

Defects directly related to cohesin's sister chromatid cohesion function occur in germ cells, which relies on a meiosis-specific cohesin complex for homologous recombination and synapsis (reviewed in [48]). Human oocytes have a prolonged arrest in meiosis I (up to 50 years) but not spermatocytes. Of importance, women fertility decreases with age and observations in mouse models and human oocytes allowed to correlate reduction of cohesin complexes bound to meiotic chromosomes, and consequent cohesion loss, with infertility [49] (see review [50]). The current model proposes that gradual loss of cohesion during the prolonged oocyte arrest results in chromosome missegregation, aneuploidy and ultimately oocyte death. One syndrome that was suggested to be linked to this mechanism is Down syndrome (DS), also named Trisomy 21, characterized by a triploid chromosome 21. A number of syndromic genetic disorders also originate from mutations in cohesion apparatus genes. The most

prominent member of this group of developmental disorders, termed cohesinopathies, is the Cornelia de Lange syndrome (CdLS). Mutations in genes encoding structural components or associated regulators of cohesin were identified as the genetic cause of CdLS (see table 1). CdLS is a multi-system malformation syndrome with a broad spectrum of clinical involvement including characteristic facial features, malformations of the upper-extremities, small stature, growth and cognitive delay as well as musculoskeletal, cardiac and gastrointestinal abnormalities. Interestingly, cell lines derived from CdLS patients do not show obvious cohesion defects [51] but an increased sensitivity for DNA damage [52]. However, CdLS patients do not have an increased risk for cancer. Expression profiling of lymphoblastoid cells obtained from CdLS patients with different mutations revealed differential expression of about 1501 non-redundant genes and allowed identifying a panel of marker genes that also correlate with the severity of the phenotype [53 •]. For a part of the genes a direct correlation with cohesin binding was shown, for the other genes it is unclear whether the promoter-enhancer communication is disturbed or whether they are downstream targets.

Studies using zebrafish and mouse as models revealed several pathways crucial for embryonic development, besides the aforementioned *HOX* genes, that might be affected by cohesin insufficiency. In NIPBL heterozygous knockout mice the protocadherin genes (*PCDH*), clusters of genes that are expressed in an exclusive manner to provide neuronal identity, were found to be misregulated [54 •]. Subsequent studies showed that cohesin-mediated DNA looping is involved in *PCDH* promoter choice [24,55,56], indicating how cohesin insufficiencies might affect the developing brain.

Nipbl depletion in zebrafish produces a spectrum of specific heart and gut/visceral organ defects, reminiscent to the defects observed in CdLS. Analysis of gene expression changes revealed the misregulation of genes that are involved in gene regulatory networks controlling endodermal differentiation and left-right patterning [57 •].

Altogether, these studies demonstrate that the effects of cohesin deficiencies in the developing embryo are highly complex. Multiple pathways are affected, either by a direct involvement of cohesin in the regulation of the gene or as downstream target.

Quite unexpectedly, and although all CdLS patient-derived cells reported so far display hypersensitivity to genotoxic stress, CdLS patients do not exhibit any cancer susceptibility caused by an altered response to DNA damage. This intriguing observation may be accounted for, at least partly, by a very protective patient environment, but has no satisfactory molecular explanation so far. Rare cases of cancer in CdLS

patients have been attributed to the consequences of malformations, like gastroesophageal reflux, but not to the genetic defect [60].

An important issue concerns how mutations causative for tumorigenesis and for CdLS differ in their physiopathological outputs. Cohesin mutations found in either disease do not cluster in distinct, functionally relevant protein domains that would allow correlating particular mutations with given altered cohesin functions and thus with CdLS or tumorigenesis. This observation first suggests that cellular and developmental context, i.e. adult and terminally differentiated *versus* developing embryonic tissues, plays a key role in cellular consequences of given mutations. Second, it also indicates that acquisition of cohesin mutations might not initiate but contribute to tumorigenesis. Obviously much work will be required to decipher how individual cohesin mutations affect its many functions and how these alterations can account for human diseases.

Conclusion

Our concept of how cohesin regulates genes, in particular in context with disease, needs to evolve from single genes to gene regulatory networks and include the fact that cohesin insufficiency can affect individual genes via multiple mechanisms.

The challenge in the field will be to evaluate observations made under drastic knockdown conditions in the respective model systems for the relevance in the patient with a much milder depletion of the respective proteins.

Acknowledgements

Research in our labs is financed in part by The German Federal Ministry of Education and Research (BMBF) (to FJK), the French National Research Agency (ANR)(to EW) and the Netherlands organisation for health research and development (ZonMw) (to KSW) under the frame of E-Rare-2 (TARGET-CdLS) the ERA-Net for Research on Rare Diseases.

References:

1. • Michaelis C, Ciosk R, Nasmyth K: **Cohesins: chromosomal proteins that prevent premature separation of sister chromatids.** *Cell* 1997, **91**:35-45.
Discovery of the cohesin complex. First description of cohesin in S. cerevisiae.

2. Losada A: **Cohesin in cancer: chromosome segregation and beyond.** *Nat Rev Cancer* 2014, **14**:389-393.
3. Peters JM, Tedeschi A, Schmitz J: **The cohesin complex and its roles in chromosome biology.** *Genes Dev* 2008, **22**:3089-3114.
4. Watrin E, Peters JM: **The cohesin complex is required for the DNA damage-induced G2/M checkpoint in mammalian cells.** *Embo J* 2009, **28**:2625-2635.
5. Jeppsson K, Kanno T, Shirahige K, Sjogren C: **The maintenance of chromosome structure: positioning and functioning of SMC complexes.** *Nat Rev Mol Cell Biol* 2014, **15**:601-614.
6. Peters JM, Nishiyama T: **Sister chromatid cohesion.** *Cold Spring Harb Perspect Biol* 2012, **4**.
7. Wang X, Yang Y, Duan Q, Jiang N, Huang Y, Darzynkiewicz Z, Dai W: **sSgo1, a major splice variant of Sgo1, functions in centriole cohesion where it is regulated by Plk1.** *Dev Cell* 2008, **14**:331-341.
8. Schockel L, Mockel M, Mayer B, Boos D, Stemmann O: **Cleavage of cohesin rings coordinates the separation of centrioles and chromatids.** *Nat Cell Biol* 2011, **13**:966-972.
9. Oliveira RA, Nasmyth K: **Cohesin cleavage is insufficient for centriole disengagement in Drosophila.** *Curr Biol* 2013, **23**:R601-603.
10. Cabral G, Sans SS, Cowan CR, Dammermann A: **Multiple mechanisms contribute to centriole separation in C. elegans.** *Curr Biol* 2013, **23**:1380-1387.
11. Bernard P, Maure JF, Partridge JF, Genier S, Javerzat JP, Allshire RC: **Requirement of heterochromatin for cohesion at centromeres.** *Science* 2001, **294**:2539-2542.
12. Wendt KS, Yoshida K, Itoh T, Bando M, Koch B, Schirghuber E, Tsutsumi S, Nagae G, Ishihara K, Mishiro T, et al.: **Cohesin mediates transcriptional insulation by CCCTC-binding factor.** *Nature* 2008, **451**:796-801.
13. Rubio ED, Reiss DJ, Welch PL, Distchech 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.
14. Parelho V, Hadjur S, Spivakov M, Leleu M, Sauer S, Gregson HC, Jarmuz A, Canzonetta C, Webster Z, Nesterova T, et al.: **Cohesins functionally associate with CTCF on mammalian chromosome arms.** *Cell* 2008, **132**:422-433.
15. • Zuin J, Dixon JR, van der Reijden MI, Ye Z, Kolovos P, Brouwer RW, van de Corput MP, van de Werken HJ, Knoch TA, van IWF, et al.: **Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells.** *Proc Natl Acad Sci U S A* 2014, **111**:996-1001.
The effect of proteolytic destruction of cohesin and CTCF depletion on the topological domains is studied using Hi-C and correlated with changes of the transcriptome, revealing significant differences between both factors.
16. Schmidt D, Schwalie PC, Ross-Innes CS, Hurtado A, Brown GD, Carroll JS, Flicek P, Odom DT: **A CTCF-independent role for cohesin in tissue-specific transcription.** *Genome Res* 2010, **20**:578-588.
17. Kagey MH, Newman JJ, Bilodeau S, Zhan Y, Orlando DA, van Berkum NL, Ebmeier CC, Goossens J, Rahl PB, Levine SS, et al.: **Mediator and cohesin connect gene expression and chromatin architecture.** *Nature* 2010, **467**:430-435.
18. Wei Z, Gao F, Kim S, Yang H, Lyu J, An W, Wang K, Lu W: **Klf4 organizes long-range chromosomal interactions with the oct4 locus in reprogramming and pluripotency.** *Cell Stem Cell* 2013, **13**:36-47.
19. Mishiro T, Ishihara K, Hino S, Tsutsumi S, Aburatani H, Shirahige K, Kinoshita Y, Nakao M: **Architectural roles of multiple chromatin insulators at the human apolipoprotein gene cluster.** *Embo J* 2009, **28**:1234-1245.
20. Nativio R, Wendt KS, Ito Y, Huddleston JE, Uribe-Lewis S, Woodfine K, Krueger C, Reik W, Peters JM, Murrell A: **Cohesin is required for higher-order chromatin conformation at the imprinted IGF2-H19 locus.** *PLoS Genet* 2009, **5**:e1000739.

21. Hadjur S, Williams LM, Ryan NK, Cobb BS, Sexton T, Fraser P, Fisher AG, Merckenschlager M: **Cohesins form chromosomal cis-interactions at the developmentally regulated IFNG locus.** *Nature* 2009, **460**:410-413.
22. Hou C, Dale R, Dean A: **Cell type specificity of chromatin organization mediated by CTCF and cohesin.** *Proc Natl Acad Sci U S A* 2010, **107**:3651-3656.
23. Chien R, Zeng WH, Kawauchi S, Bender MA, Santos R, Gregson HC, Schmiesing JA, Newkirk DA, Kong XD, Ball AR, et al.: **Cohesin Mediates Chromatin Interactions That Regulate Mammalian beta-globin Expression.** *Journal of Biological Chemistry* 2011, **286**:17870-17878.
24. • Guo Y, Monahan K, Wu H, Gertz J, Varley KE, Li W, Myers RM, Maniatis T, Wu Q: **CTCF/cohesin-mediated DNA looping is required for protocadherin alpha promoter choice.** *Proc Natl Acad Sci U S A* 2012, **109**:21081-21086.
This publication describes the role of cohesin and CTCF for protocadherin promoter choice, hinting how cohesin complex defects can impair neuronal development.
25. • Phillips-Cremins JE, Sauria ME, Sanyal A, Gerasimova TI, Lajoie BR, Bell JS, Ong CT, Hookway TA, Guo C, Sun Y, et al.: **Architectural protein subclasses shape 3D organization of genomes during lineage commitment.** *Cell* 2013, **153**:1281-1295.
High resolution interaction maps of seven genomic regions show that distinct combinations of CTCF, mediator and cohesin that anchor different length scales of chromatin interactions. Cohesin and CTCF anchor long-range constitutive interactions.
26. Nitzsche A, Paszkowski-Rogacz M, Matarese F, Janssen-Megens EM, Hubner NC, Schulz H, de Vries I, Ding L, Huebner N, Mann M, et al.: **RAD21 cooperates with pluripotency transcription factors in the maintenance of embryonic stem cell identity.** *PLoS ONE* 2011, **6**:e19470.
27. Zhang H, Jiao W, Sun L, Fan J, Chen M, Wang H, Xu X, Shen A, Li T, Niu B, et al.: **Intrachromosomal looping is required for activation of endogenous pluripotency genes during reprogramming.** *Cell Stem Cell* 2013, **13**:30-35.
28. DeMare LE, Leng J, Cotney J, Reilly SK, Yin J, Sarro R, Noonan JP: **The genomic landscape of cohesin-associated chromatin interactions.** *Genome Res* 2013, **23**:1224-1234.
29. Seitan VC, Hao B, Tachibana-Konwalski K, Lavagnoli T, Mira-Bontenbal H, Brown KE, Teng G, Carroll T, Terry A, Horan K, et al.: **A role for cohesin in T-cell-receptor rearrangement and thymocyte differentiation.** *Nature* 2011, **476**:467-471.
30. Degner SC, Verma-Gaur J, Wong TP, Bossen C, Iverson GM, Torkamani A, Vettermann C, Lin YC, Ju ZL, Schulz D, et al.: **CCCTC-binding factor (CTCF) and cohesin influence the genomic architecture of the Igh locus and antisense transcription in pro-B cells.** *Proceedings of the National Academy of Sciences of the United States of America* 2011, **108**:9566-9571.
31. Sexton T, Yaffe E, Kenigsberg E, Bantignies F, Leblanc B, Hoichman M, Parrinello H, Tanay A, Cavalli G: **Three-dimensional folding and functional organization principles of the Drosophila genome.** *Cell* 2012, **148**:458-472.
32. •• Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B: **Topological domains in mammalian genomes identified by analysis of chromatin interactions.** *Nature* 2012, **485**:376-380.
Genome-wide chromatin interaction maps determined by Hi-C reveal genome reveal topological domains as building blocks for genome architecture. Position of these domains is invariant between different cell types and in part even conserved between mouse and human. Boundaries of the domains are enriched for CTCF, housekeeping genes, transfer RNAs and short interspersed element (SINE) retro-transposons.
33. •• Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, et al.: **Comprehensive mapping of long-range interactions reveals folding principles of the human genome.** *Science* 2009, **326**:289-293.

Landmark paper introducing for the first time the Hi-C technology as tool to study genome architecture and reveals the organization of the genome in Mb sized domains.

34. Seitan VC, Faure AJ, Zhan Y, McCord RP, Lajoie BR, Ing-Simmons E, Lenhard B, Giorgetti L, Heard E, Fisher AG, et al.: **Cohesin-based chromatin interactions enable regulated gene expression within preexisting architectural compartments.** *Genome Res* 2013, **23**:2066-2077.
35. Sofueva S, Yaffe E, Chan WC, Georgopoulou D, Rudan MV, Mira-Bontenbal H, Pollard SM, Schroth GP, Tanay A, Hadjur S: **Cohesin-mediated interactions organize chromosomal domain architecture.** *Embo Journal* 2013, **32**:3119-3129.
36. Lonfat N, Duboule D: **Structure, function and evolution of topologically associating domains (TADs) at HOX loci.** *FEBS Lett* 2015.
37. Kim YJ, Cecchini KR, Kim TH: **Conserved, developmentally regulated mechanism couples chromosomal looping and heterochromatin barrier activity at the homeobox gene A locus.** *Proc Natl Acad Sci U S A* 2011, **108**:7391-7396.
38. Muto A, Ikeda S, Lopez-Burks ME, Kikuchi Y, Calof AL, Lander AD, Schilling TF: **Nipbl and mediator cooperatively regulate gene expression to control limb development.** *PLoS Genet* 2014, **10**:e1004671.
39. Singh VP, Gerton JL: **Cohesin and human disease: lessons from mouse models.** *Curr Opin Cell Biol* 2015, **37**:9-17.
40. Kon A, Shih LY, Minamino M, Sanada M, Shiraishi Y, Nagata Y, Yoshida K, Okuno Y, Bando M, Nakato R, et al.: **Recurrent mutations in multiple components of the cohesin complex in myeloid neoplasms.** *Nat Genet* 2013, **45**:1232-1237.
41. Dolnik A, Engelmann JC, Scharfenberger-Schmeer M, Mauch J, Kelkenberg-Schade S, Haldemann B, Fries T, Kronke J, Kuhn MW, Paschka P, et al.: **Commonly altered genomic regions in acute myeloid leukemia are enriched for somatic mutations involved in chromatin remodeling and splicing.** *Blood* 2012, **120**:e83-92.
42. Solomon DA, Kim JS, Bondaruk J, Shariat SF, Wang ZF, Elkahloun AG, Ozawa T, Gerard J, Zhuang D, Zhang S, et al.: **Frequent truncating mutations of STAG2 in bladder cancer.** *Nat Genet* 2013, **45**:1428-1430.
43. Balbas-Martinez C, Sagrera A, Carrillo-de-Santa-Pau E, Earl J, Marquez M, Vazquez M, Lapi E, Castro-Giner F, Beltran S, Bayes M, et al.: **Recurrent inactivation of STAG2 in bladder cancer is not associated with aneuploidy.** *Nat Genet* 2013, **45**:1464-1469.
44. Cucco F, Servadio A, Gatti V, Bianchi P, Mannini L, Prodosmo A, De Vitis E, Basso G, Friuli A, Laghi L, et al.: **Mutant cohesin drives chromosomal instability in early colorectal adenomas.** *Hum Mol Genet* 2014, **23**:6773-6778.
45. Yan M, Xu H, Waddell N, Shield-Artin K, Haviv I, kConFab a, McKay MJ, Fox SB: **Enhanced RAD21 cohesin expression confers poor prognosis in BRCA2 and BRCA1, but not BRCA1 familial breast cancers.** *Breast Cancer Res* 2012, **14**:R69.
46. Katainen R, Dave K, Pitkanen E, Palin K, Kivioja T, Valimaki N, Gylfe AE, Ristolainen H, Hanninen UA, Cajuso T, et al.: **CTCF/cohesin-binding sites are frequently mutated in cancer.** *Nat Genet* 2015, **47**:818-821.
47. Kihara R, Nagata Y, Kiyoi H, Kato T, Yamamoto E, Suzuki K, Chen F, Asou N, Ohtake S, Miyawaki S, et al.: **Comprehensive analysis of genetic alterations and their prognostic impacts in adult acute myeloid leukemia patients.** *Leukemia* 2014, **28**:1586-1595.
48. McNicoll F, Stevense M, Jessberger R: **Cohesin in gametogenesis.** *Curr Top Dev Biol* 2013, **102**:1-34.
49. Le Quesne Stabej P, Williams HJ, James C, Tekman M, Stanescu HC, Kleta R, Ocaka L, Lescai F, Storr HL, Bitner-Glindzicz M, et al.: **STAG3 truncating variant as the cause of primary ovarian insufficiency.** *Eur J Hum Genet* 2015.

50. Jessberger R: **Deterioration without replenishment-the misery of oocyte cohesin.** *Genes & Development* 2010, **24**:2587-2591.
51. Kaur M, DeScipio C, McCallum J, Yaeger D, Devoto M, Jackson LG, Spinner NB, Krantz ID: **Precocious sister chromatid separation (PSCS) in Cornelia de Lange syndrome.** *Am J Med Genet A* 2005, **138**:27-31.
52. Vrouwe MG, Elghalbzouri-Maghrani E, Meijers M, Schouten P, Godthelp BC, Bhuiyan ZA, Redeker EJ, Mannens MM, Mullenders LH, Pastink A, et al.: **Increased DNA damage sensitivity of Cornelia de Lange syndrome cells: evidence for impaired recombinational repair.** *Hum Mol Genet* 2007, **16**:1478-1487.
53. • Liu J, Zhang Z, Bando M, Itoh T, Deardorff MA, Clark D, Kaur M, Tandy S, Kondoh T, Rappaport E, et al.: **Transcriptional dysregulation in NIPBL and cohesin mutant human cells.** *PLoS Biol* 2009, **7**:e1000119.
- This study analyzed LCL from a large cohort of CdLS patient to describe differentially expressed genes and define a panel of misregulated genes that can be used for diagnostic purposes.*
54. • Kawauchi S, Calof AL, Santos R, Lopez-Burks ME, Young CM, Hoang MP, Chua A, Lao T, Lechner MS, Daniel JA, et al.: **Multiple organ system defects and transcriptional dysregulation in the Nipbl(+/-) mouse, a model of Cornelia de Lange Syndrome.** *PLoS Genet* 2009, **5**:e1000650.
- First Nipbl heterozygous knockout mouse model demonstrating that part of the defects observed in CdLS patients can be recapitulated in a mouse model.*
55. Remeseiro S, Cuadrado A, Gomez-Lopez G, Pisano DG, Losada A: **A unique role of cohesin-SA1 in gene regulation and development.** *EMBO J* 2012, **31**:2090-2102.
56. Cuadrado A, Remeseiro S, Grana O, Pisano DG, Losada A: **The contribution of cohesin-SA1 to gene expression and chromatin architecture in two murine tissues.** *Nucleic Acids Res* 2015, **43**:3056-3067.
57. • Muto A, Calof AL, Lander AD, Schilling TF: **Multifactorial origins of heart and gut defects in nipbl-deficient zebrafish, a model of Cornelia de Lange Syndrome.** *PLoS Biol* 2011, **9**:e1001181.
- Study demonstrating the involvement of Nipbl in embryonic development of heart and gut, two organs affected in CdLS, using zebrafish as model system.*
58. • Rao SS, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES, et al.: **A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping.** *Cell* 2014, **159**:1665-1680.
- Landmark paper providing high resolution Hi-C interaction maps for a large number of human cell types.*
59. Stamatoyannopoulos JA, Snyder M, Hardison R, Ren B, Gingeras T, Gilbert DM, Groudine M, Bender M, Kaul R, Canfield T, et al.: **An encyclopedia of mouse DNA elements (Mouse ENCODE).** *Genome Biol* 2012, **13**:418.
60. Schrier SA, Sherer I, Deardorff MA, Clark D, Audette L, Gillis L, Kline AD, Ernst L, Loomes K, Krantz ID, Jackson LG: **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 Part A* 2011, **155**:3007-3024.

Table 1

Gene Symbol	Locus	Protein Name	CdLS cases
<i>NIPBL</i>	5p13.2	Nipped-B-like protein	65 %
<i>SMC1A</i>	Xp11	Structural maintenance of chromosomes 1A	5 %
<i>SMC3</i>	10q25.2	Structural maintenance of chromosomes 1A	< 1 %
<i>HDAC8</i>	Xq13.1	Histone deacetylase 8	~5 %
<i>RAD21</i>	8q24.11	Double-strand break repair protein Rad21 homolog	< 1 %

Table 1. List of genes whose mutations cause Cornelia de Lange syndrome.

Gene symbol, genomic localisation, gene product name and proportion of CdLS patients are given for the five known CdLS genes.

Figure legends

Figure 1. Cohesin complex structure.

Cohesin core complex subunit *SMC1A*, *SMC3* and *SCC1/RAD21* form a tripartite molecular ring that embraces DNA fibres (here indicated as cross section of two 11 nm fibres), thereby ensuring SCC in a topological manner. Additional regulatory proteins, i.e. *SA1/2* and *Pds5A/B*, sororin and *Wapl* bind to cohesin core complex in various combinations.

Figure 2. Models of cohesin's involvement in chromatin related processes.

(A) Cohesin embraces and holds sister chromatid together to ensure SCC.

(B) Cohesive property of cohesin assists recombinational repair of DNA double strand breaks by tethering the break site to the intact sister strand.

(C-D) Cohesin also acts as chromatin loop structural basis that serves different processes: interaction between distal enhancer and gene promoter to trigger expression (C) or physical separation (insulation) of gene regulatory elements (D).

(E) Cohesin-mediated chromatin loops cluster together, as seen in TADs, and provide the spatial organization of gene clusters.

Figure 3. Cohesin complex and sister chromatid cohesion during cell cycle.

In vertebrates, cohesin complexes are loaded onto chromatin at mitotic exit in a NIPBL/MAU2 complex (also called Kollerin) dependent manner. SMC3 acetylation by ESCO1/ESCO2 acetyltransferases during DNA replication recruits sororin that counteracts anti-establishment activity of WAPL, thereby triggering cohesion establishment. SCC is dissolved during mitosis by the complete removal of cohesin complexes from chromosomes in a two-step process: dissociation from chromosome arms in early mitosis and proteolytic cleavage of SCC1/RAD21 by separase at metaphase-to-anaphase transition.

Figure 4. Long-range interactions shaping the chromatin loops around the human HOXA genes

Hi-C interaction matrices of IMR90 cells at 10kb resolution by Rao and co-workers [58 ••] were displayed using <http://promoter.bx.psu.edu/hi-c/index.html>. Below, the Refseq genes are annotated and the CHIP-sequencing profiles for cohesin (RAD21) and CTCF from IMR90 cells (available from ENCODE/SYDH [59]) are shown.

Below the Hi-C profile two grey bars indicate major chromatin loops/TADs in this region. The left loop comprises most of the HoxA genes and the right loop contains the region where the limb-specific enhancers localize. Both loops/TADs are flanked by cohesin/CTCF binding sites with a most striking clustering of sites in the HoxA gene region.

Figure 1

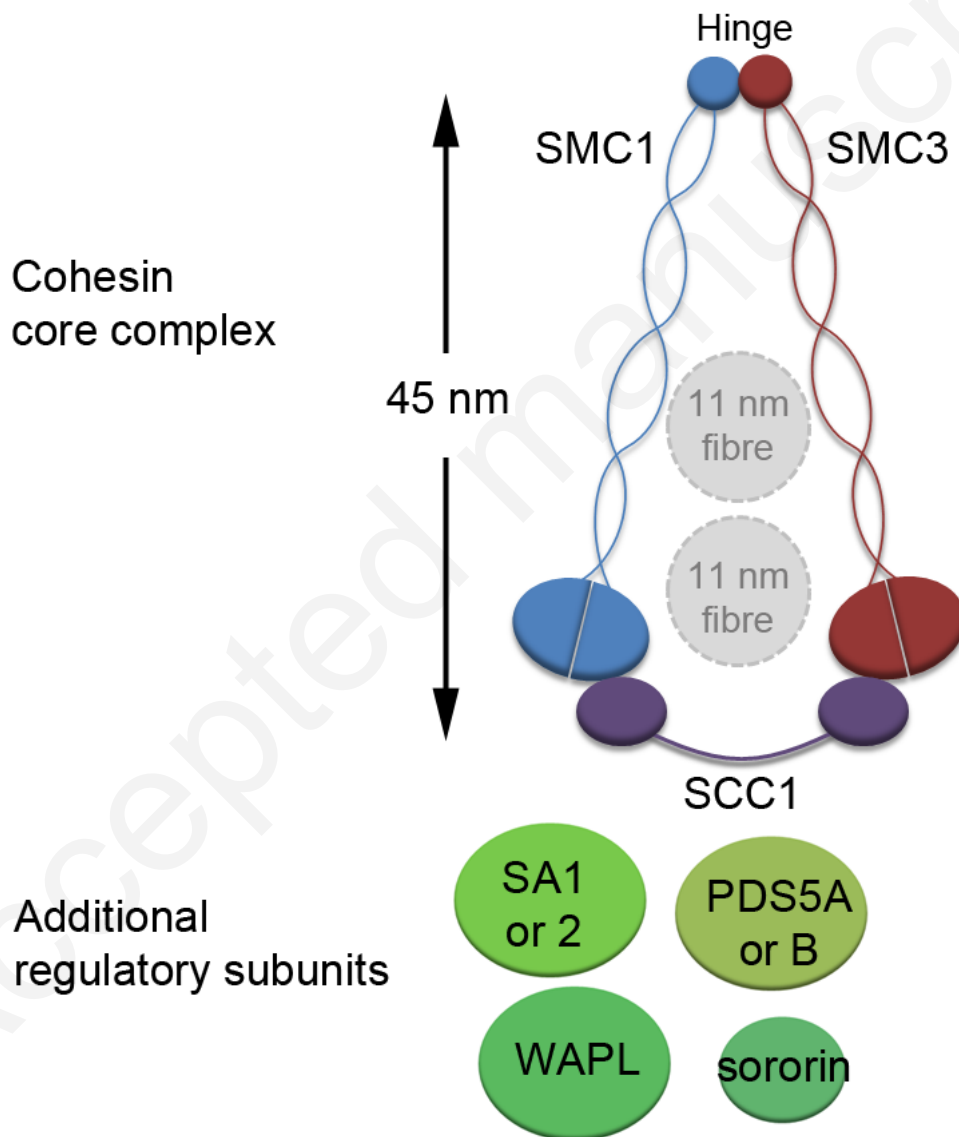


Figure 2

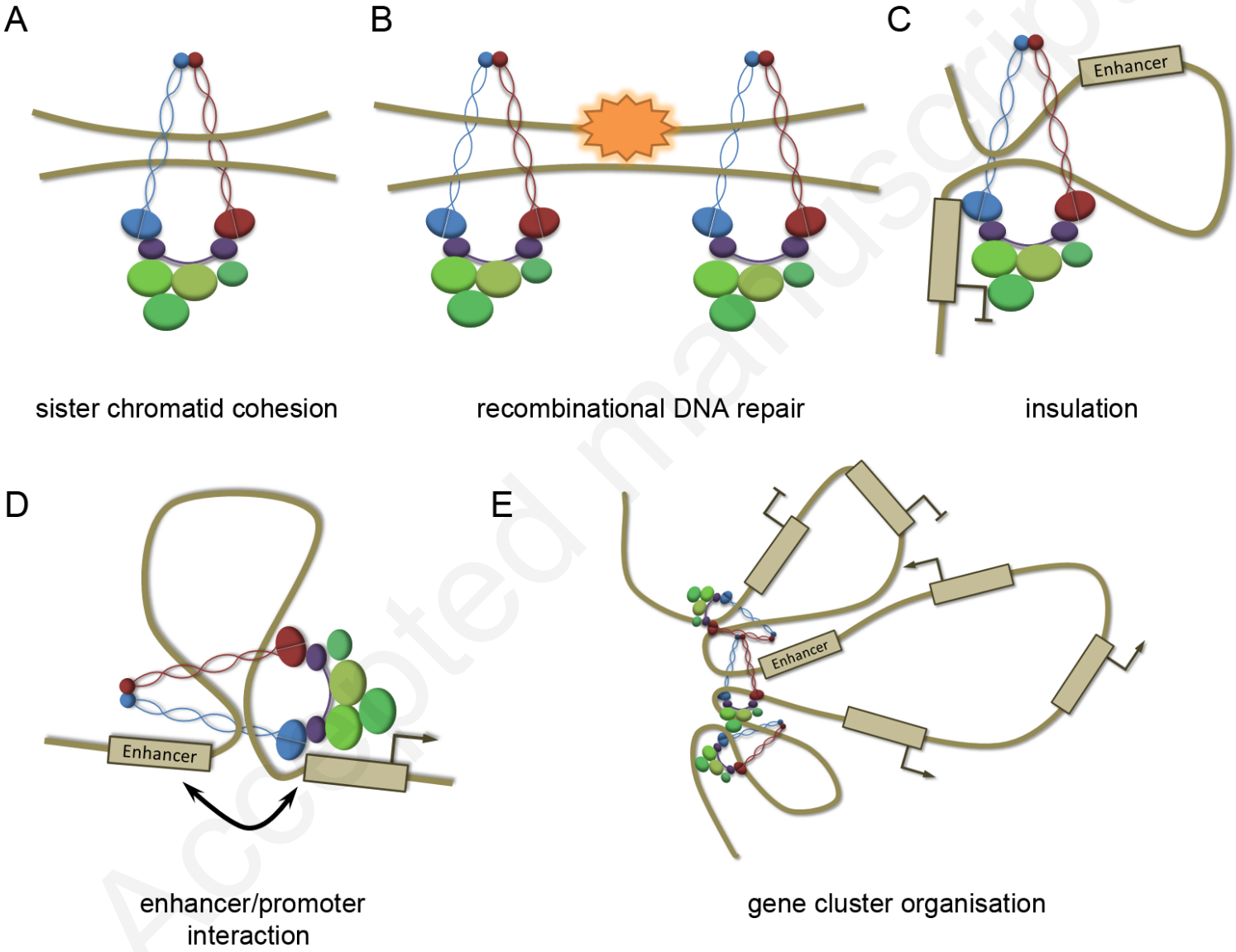


Figure 3

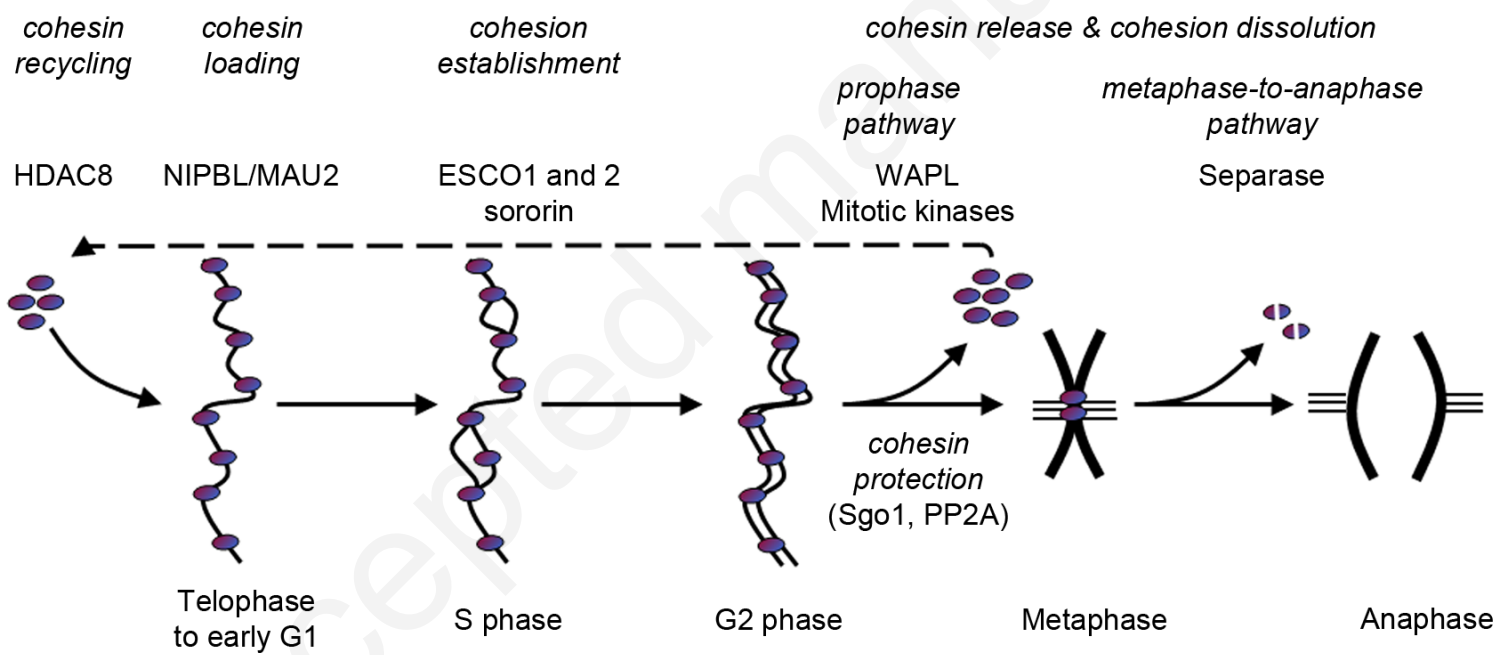


Figure 4

Hi-C:
IMR90 cells
10 kb resolution

