2016

Similarities and Differences in the Cohesinopathies of Roberts Syndrome and Cornelia de Lange Syndrome

Rachel Sternberg
Lehigh University

Similarities and Differences in the Cohesinopathies of Roberts Syndrome and Cornelia de Lange Syndrome

Rachel Sternberg
BIOS 262
Prof. Robert Skibbens
Writing Intensive Paper
May 1, 2015

Table of Contents Page

Abstract

Cohesins were first identified based on their role in chromosome segregation. The products of chromosome replication must be identified from S phase until anaphase onset in mitosis to ensure high fidelity chromosome segregation. Cohesin complexes tether together sister chromatids to maintain identity over time. More recently, cohesins were found to play numerous roles in transcription regulation and other forms of DNA metabolism, including the firing of clustered DNA replication forks, insulator function, and replication fork restart. Genetic mapping that revealed cohesin mutations are responsible for severe developmental defects spurred intense research regarding which role of cohesins was crucial to proper development.

Cohesinopathies are a group of developmental disorders caused by disruption in cohesin function. Cornelia de Lange Syndrome (CdLS) and Roberts Syndrome (RBS) are two examples of cohesinopathies that exhibit a similar suite of phenotypes but are thought to arise through very different cellular mechanisms. This review investigates the similarities and differences between these two disorders to formulate a hypothesis unifying the relationship of these disorders as a possible continuum of manifestations caused by the same underlying mechanism.

I. Introduction

Developmental maladies are defined by clinician observations. Such clinically-defined disorders are often based on just a few affected individuals, and they are made by physicians greatly separated both geographically and temporally. For syndromes exhibiting similar or overlapping manifestations, efforts meant to differentiate them complicated studies meant to help define them. Of particular interest were clinical diagnoses of numerous syndromes associated with mental and physical growth deficiencies and limb malformations that were later collectively

called cohesinopathies. These developmental disorders are caused by the disruption in the normal function of a protein complex called cohesin. The following investigation is focused on two cohesinopathies, Roberts Syndrome (RBS) and Cornelia de Lange Syndrome (CdLS). First, both disorders are described clinically. This is followed by a review of their cellular phenotypes, genotoxicities, and genotypes. Next, cohesin structure and function are presented, followed by a discussion that explores which cohesin functions may be impaired in RBS and CdLS. Lastly, a hypothesis on the interrelationship of these disorders is proposed.

II. Clinical Manifestations in Patients

II.A. Roberts Syndrome

RBS is a rare disorder, affecting approximately 150 individuals. As listed in Table 1, alternate names for the disorder include Appelt-Gerken-Lenz Syndrome, Hypomelia-Hypoichosis-Facial Hemangioma Syndrome, Pseudothalidomide Syndrome, SC Phocomelia Syndrome, and Tetraphocomelia-Cleft Palate Syndrome (Gen. Home Ref. 2015). The range and severity of RBS manifestations is now well documented (Vega 2005, 2010; Shule 2005). Affected individuals have pre- and post-natal growth retardation, reduced mental capacity, and physical malformations primarily of the skull, face, and limbs (Table 2). These malformations include bilateral symmetric limb reduction which is worse for arms than legs, cleft lip and palate, microcephaly, beak nose, micrognathia, and unusually wide-set eyes (hypertelorism) with protruding eyeballs (exophthalmos) and corneal clouding. The effects range widely in severity; however, the extent of facial malformations correlate with the extent of limb defects. Severely affected individuals have high mortality rate shortly before or after birth, while individuals mildly affected can live longer.

Until recently, SC Phocomelia Syndrome was considered a disorder separate from RBS. In part, distinguishing between these disorders was based on the relatively mild phenotypes manifested in SC Phocomelia compared to those observed in RBS. In contrast, the range of phenotypes were nearly indistinguishable. An R:S rating scale for clinical severity was developed in an attempt to help clinicians distinguish between these two difficult-to-discern disorders (reviewed by Schüle 2005). To improve the consistency of diagnoses between clinicians, the rating system was later modified to consider six criteria—1)survival, 2)growth retardation, 3)phocomelia of the arms, 4)phocomelia of the legs, 5)palatal abnormalities, and 6)craniofacial abnormalities (reviewed by Schüle 2005). Ultimately, the R:S system became defunct when SC Phocomelia Syndrome and RBS were identified as the same disorder through genetic mapping studies (reviewed by Schüle 2005). The weight of evidence included that: 1) severely and mildly affected persons occurred in the same sibship, 2) both RBS and SC Phocomelia Syndrome display Heterochromatin Repulsion, which is the puffing or repulsion of the constitutive heterochromatin, 3) somatic cell complementation studies which showed cells positive for Heterochromatin Repulsion from patients with RBS and SC Phocomelia Syndrome both failed to complement each other, 4) cells from both syndromes exhibit mitotic cell premature centromere separation, and 5) mutations in the gene *ESCO2* (see below) on chromosome 8 accounted for all the affected individuals studied. Thus, RBS and SC Phocomelia Syndrome arise from mutations within the same gene that results in a broad range in severity of effects. Understanding the mechanism through which identical mutations can give rise to a wide range in severity of effects is likely to provide important new insights regarding development regulation through genetics and environmental cues.

II.B. Cornelia de Lange Syndrome

CdLS frequency of occurrence is 1:10,000 (reviewed by Krantz 2004). The disorder is also known as the Brachmann-De Lange Syndrome (Table 1). Like RBS, people with CdLS have growth retardation, reduced mental capacity, limb abnormalities, and distinctive craniofacial malformations. Some CdLS craniofacial abnormalities are different from those in RBS. For instance, CdLS patients have ptosis, a beak-like nose, and no midfacial haemangioma. Moreover, CdLS appears to affect more systems during development than RBS: heart defects, impaired vision, gastrointestinal dysfunction, excessive hairiness in both sexes, hirsutism in women, and underdeveloped genitalia all appear pronounced in CdLS (Krantz 2004). Reduced body fat, hearing loss, autism, and self-inflicting injuries also frequently occur (Kawauchi 2009). Intriguingly, only upper limb defects are reported in CdLS individuals while both upper and lower limbs are affected in RBS (Table 2) (Krantz 2004). The significance of these distinctions is unknown, but could be simply an example of the range of phenotypes or penetrance caused by gene mutation. For instance, it is possible that lower limb malformations are very slight in comparison with those of the upper arms and underreported. Another possibility is that individuals included in CdLS studies were only mildly affected, such that lower limb malformations were absent. The differences observed between CdLS- and RBS- affected individuals also may be due to the broad range in severity of effects. Alternatively, the cohesinopathies may have different underlying mechanisms, which will be discussed further.

III. Cellular Phenotypes Differentiate RBS from CdLS

III.A. RBS Cellular Morphology

The cellular phenotypes of RBS and CdLS are different. RBS cells often contain micronuclei, exhibit increased frequency of aneuploidy, and contain nuclei that adopt an unusual morphology (Tomkins 1984). RBS cells also display Heterochromatin Repulsion in mitotic spreads, which is the puffing or repulsion of the constitutive heterochromatin (Van Den Berg 1993). Schüle (2005) claims that Heterochromatin Repulsion is exclusively pathognomonic for RBS [and SC Phocomelia] and, therefore, it can be used as a diagnostic tool to differentiate RBS from other diseases causing phocomelia and pre-and post-natal delayed development.

III.B. RBS Genotoxic Sensitivities

Micronuclei (like those observed in RBS cells) can form from acentric DNA fragments, or simply from failed mitosis in which lagging chromosomes never get to the pole to form their own nucleus. These fragments often are the result of double strand DNA breaks or cellular DNA damage. Cells from RBS and CdLS exposed to different agents that affect DNA provide evidence supporting mitotic dysfunction, but also suggest that RBS and CdLS arise through different mechanisms. Such in vitro studies are critical to predict whether patients might exhibit different sensitivities to genotoxic agents.

Exposure of Epstein-Barr virus (EBV)-immortalized lymphoblastoid cell lines derived from two persons with RBS to gamma radiation, known to cause strand breakage, increased micronucleation in the treated cells (Van Den Berg 1993). Similarly, exposure of these RBS cells to mytomycin C, which causes DNA adducts and crosslinking by alkylation, increased micronucleation. RBS cells were four times more sensitive to gamma radiation and five times more sensitive to mitomycin C than control cells. The exhibited genotoxic hypersensitive response of RBS cells support defects in DNA repair.

III.C. CdLS Cellular Morphology

In contrast with the RBS cellular phenotype, most studies find that cells from individuals with CdLS typically undergo normal mitosis, and do not have exhibit either premature chromatid separation or Heterochromatin Repulsion (reviewed by Skibbens 2013). Fluorescent *in situ* hybridization (FISH) revealed that cells altered to mimic CdLS exhibit altered chromosome architecture (Nolen 2013). For instance, chromosome architecture in lymphoblastoid cells derived from severe CdLS patients exhibit increased nuclear size. The investigators eliminated abnormal cell cycle as the cause of the increased size by demonstrating a normal cell cycle distribution using fluorescence-activated cell sorting analysis (FACS). These findings suggest that the increased nuclear size is due to chromatin decompaction—large scale unfolding of higher-order chromatin structure (Nolen 2013). Testing of different regions of chromatin using FISH indicated that the extent of decompaction is proportional to gene density (and highest density of CTCF, which is a regulator of transcription and chromatin architecture) and cohesin binding sites. In addition, cells derived from individuals with more severe phenotype have greater visible chromatin decompaction (Nolen 2013). Thus, the clinical phenotypes of CdLS individuals may be caused by dysfunction in the regulation of gene expression.

III.D. CdLS Genotoxic Sensitivities

Like the hypersensitive response of RBS cells, cells derived from CdLS individuals exhibit increased sensitivity to mitomycin C, based on the endpoints of increased cell death, decreased proliferation, and increased chromosomal aberrations (Vrouwe 2007). However, CdLS cells were exposed to a different amount of mitomycin C than that used in the RBS study,

making direct comparison difficult (Vrouwe 2007). Exposure to x-rays and gamma radiation are two ways of causing DNA breakage. X-ray irradiation of CdLS cells increased chromosomal aberrations, but only when the exposure occurred during G2-phase of the cell cycle. In contrast, RBS cells exposed to gamma radiation exhibited increasing chromosomal aberrations in exposure during any part of the cell cycle. The investigators suggest that the mutated CdLS cell sensitivity occurs when double strand DNA breakage requires establishment of sister chromatid cohesion for recombinational repair (Vrouwe 2007). Unfortunately, it is difficult to directly compare effects obtained by X-rays to those obtained by gamma radiation.

III.E. Discussion of RBS and CdLS Cellular Phenotypes

The cellular phenotypes of these two cohesinopathies are notably different. RBS cells exhibit Heterochromatin Repulsion and the results of mitotic errors. Neither occur (or are prominent) in CdLS cells. On the other hand, CdLS cells have altered chromosome architecture, exhiniting increased nuclear size due to chromatin decompaction. Thus far, as to genotoxic sensitivities, both the RBS cell and CdLS cells revealed increased sensitivities to genotoxic agents. However, the inability to make direct comparisons between cell-types genotoxic sensitivities preclude determining whether the two cohesinopathies arise through different mechanisms. A future study needs to expose RBS and CdLS cells concurrently to the same agents and at the same quantities to better compare the responses by the different cell mutations.

IV. Genetic Basis of RBS and CdLS

IV.A. RBS Genotype- Identified Mutated Gene

To delve into the genetic basis of RBS, a genome-wide search using homozygosity mapping was used to locate a potentially affected locus on chromosome 8 (Vega 2005). Using a candidate gene approach, the transcript *LOC157570* identified eight different mutations of the gene *ESCO2* in the 18 tested individuals from different families with RBS. One was a missense mutation, another was a nonsense mutation, and six were frameshift mutations. A later study of 49 cases with RBS identified 26 different mutations in *ESCO2* (Vega 2010). Most of these mutations introduce a premature stop codon in the *ESCO2* transcript by splicing, frameshift, and nonsense mutations, but all the *ESCO2* mutations result in a missing or nonfunctional protein, rendering *ESCO2* null (Schüle 2005). *ESCO2* gene mutations are autosomal recessive.

Genotype-phenotype analysis using Fisher exact test and McNemar test indicate no correlation between the severity of clinical findings and the type of *ESCO2* mutation (Vega 2010). Another study of five families affected with RBS also indicates that genotype does not predict phenotype: neither the type of mutation or location in the mutated *ESCO2* gene correlated with phenotype (Schüle 2005). Although no correlation has been established between genotype and phenotype, the timing in *ESCO2* expression during development correlated with the malformations observed in the brain, face, limb, kidney, and gonads using in situ hybridization on human embryos (Vega 2010). Regardless of the specific mutation, the gene *ESCO2* is mutated in 100% of the RBS cases (Vega 2005). As reviewed by Schüle (2005), RBC cellular characteristics are not found in heterozygous cells, and they are complemented by the Chinese hamster genome in somatic-cell hybrids.

ESCO2 is the abbreviated name for Establishment of Sister Chromatid Cohesion 1 homolog 2. The gene codes for a protein with a putative N-acetyltransferase role. *ESCO2* is also called *EFO2* (Establishment Factor Ortholog 2). Characterization of the *ESCO2* gene, using

5' and 3' RACE and RT-PCR, found a cDNA of 3,348 nucleotides with an open reading frame of 1,806 nucleotides (Vega 2005). The coding sequence consists of 11 exons spanning 30.3 kb, a start codon in exon 2, and a stop codon in exon 11. The gene's synthesized polypeptide, ESCO2 protein, consists of 601 amino acids with a molecular weight of 68.3 kDa (Vega 2005).

IV.B. Evolutionary Conservation of *ESCO2* and RBS Models

Vertebrate *ESCO2* orthologs exhibit >92%, 77%, 57%, and 41% similarity to the primate, rodent, chicken, and fish, respectively (Vega 2005). The C-terminal portion of *ESCO2* is a human homolog of *Saccharomyces cerevisiae* Eco1p/Ctf7p (Skibbens 1999; Toth 1999). *Schizosaccharomyces pombre* Eso1p and *Drosophila melanogaster* Deco are also homologs (reviewed by Vega 2005). Sequence similarity in different species suggests similar function(s) of vital importance, and enables creation of models with which to study RBS and other cohesinopathies. Much of the research on RBS uses human mutant cell lines, mice, zebrafish, and medaka, although other models are employed.

IV.C. CdLS Genotype- Identified Mutated Gene(s)

Through genome-wide linkage exclusion analysis in 12 families with CdLS and corroborative evidence of another diagnosed CdLS individual, the gene *NIPBL* was identified on chromosome 5 (Krantz 2004). Missense mutations and frameshift mutations caused by deletions and insertions were spread throughout the *NIPBL* gene; and all the mutations are expected to result in a truncated or untranslated protein (Krantz 2004). Expression patterns by northern blot on human adult and fetal tissues, and *in situ* analyses of mice embryos at days 9.5 and 10.5 of gestation, show that NIPBL protein is widely expressed in fetal and adult tissues (Krantz 2004).

Subsequent studies tend to utilize human embryonic tissue sections to avoid differences between mouse and human orthologs. *In situ* hybridization locate *NIPBL* expression in human embryonic tissue in limbs, craniofacial regions, spinal column, and notochord. This expression pattern is largely consistent with the CdLS phenotype, although a causative correlation remains to be determined (Tonkin 2004).

Mutations in *NIPBL* are responsible for the majority of CdLS cases. In fact, more than 50% of gene expression and genome-wide cohesin binding patterns in cells from CdLS probands are due to *NIPBL* mutations (Liu 2009). In addition, 80% of the mutations in the most severe cases of individuals having CdLS involve *NIPBL* (reviewed by Nolen 2013). *NIPBL* stands for *Nipped-B*-like from the Nipped-B homolog in Drosophila (Tonkin 2004). However, mutations in two other genes on separate chromosomes also cause CdLS. Mutations in *SMC1A* cause 5% of CdLS cases. *SMC1A* is the abbreviated name for Structural Maintenance of Chromosome 1A. *SMC1A* is located on the X chromosome and is inherited as X-linked dominant. The SMC1A protein is a core subunit in the cohesin complex. A *SMC3* mutation causing CdLS is very rare. *SMC3* is the abbreviated name for Structural Maintenance of Chromosome 3. The gene is located on chromosome 10 and is inherited as an autosomal dominant. The SMC3 protein is another core subunit in the cohesin complex. Mutations in *NIPBL*, *SMC1A*, and *SMC3* only account for 65% of CdLS cases, leaving 35% of the cases currently with no known origin. Thus, CdLS is more complicated than the 100% mutation incidence in *ESCO2* causing RBS. Due to the greater frequency in which *NIPBL* mutations cause CdLS and the extent of severities, the remainder of this review will focus on *NIPBL* when discussing CdLS.

NIPBL encodes for the protein delangin, which controls the deposition of cohesin complex onto the DNA that makes up sister chromatids (Tonkin 2004). Delangin belongs to a

family of chromosomal adherins having broad roles in sister chromatid cohesion, chromosome condensation, and DNA repair (Krantz 2004). The findings of mutations in a single allele of *NIPBL* in cases with CdLS supports inheritance being an autosomal dominant (Krantz 2004). A mouse model carrying one null allele of *NIPBL* also supports haploinsufficiency being the genetic mechanism of CdLS (Kawauchi 2009).

IV.D. Evolutionary Conservation of *NIPBL* and CdLS Models

As with *ESCO2*, *NIPBL* is evolutionary conserved across different species and suggests it has a significant role in development. The *NIPBL* gene product shows homology with the mouse (>92%), rat (>88%), zebrafish (63%), and fruit fly *NIPPED-B* gene (37%) (Krantz 2004; Tonkin 2004). BLAST reveals substantial homology between *NIPBL* and *Scc2* in budding yeast *S. cerevisiae* (Krantz 2004), Nipped-B in fruit fly *Drosophila melanogaster*, XM_320088 in malaria mosquito *Anapheles gambiae*, PQN-85 in worm *Caenorhabditis elegans*, CBG0727 in worm *C. briggae*, NM_121558 in plant *Arabidopsis thaliana*, and NM_186173 in plant *Oryza sativa* (Tonkin 2004). In each case, homology resides mostly in a segment of ~1,500 amino acids in the C-terminal half of the protein (Tonkin 2004).

Among the models used to investigate CdLS are pectoral fin buds of Nipbl-deficient zebrafish larvae and *NIPBL* haploinsufficient mice in their entirety or limb buds. To create the mouse model, two mouse embryonic stem cell lines have gene-trap insertions placed in *NIPBL* (Kawauchi 2009). The insertion creates premature termination of transcription and translation of the gene protein, since truncated message lacks all but the first exon. After injecting the insertion into C57BL/6 blastocysts, male chimeras were subsequently bred against both inbred (C57BL/6) and outbred (CD-1) mice. The frequency by which progeny created from inbreeding

being heterozygous (*NIPBL*$^{+/-}$) was 5.5%. This frequency is very low since Mendelian inheritance dictates a frequency of 50% heterozygous. The chimeras outbred against CD-1 females yielded a 19% frequency with the mutant allele. Again, the frequency was lower than expected. Further outcrosses with CD-1 females produced similar ratios, suggesting that 75-80% of *NIPBL*$^{+/-}$ mice die prior to genotyping (performed at 4 weeks of age). To test for the occurrence of lethality in utero, litters were examined for the appearance of the mutant allele just prior to birth. A frequency of 41%, versus the expected 50%, confirmed that most heterozygous embryos die at or after birth. It is worth noting that 41% heterozygous frequency is still below the predicted value, suggesting that other parameters could be causing the lower frequency and confounding collected data.

The above mouse model replicates many of the pathological features of CdLS: overall smaller size, heart defects, delayed bone maturation, craniofacial changes, neurological abnormalities, hearing deficits, decreased body fat, and high mortality soon after birth (Kawauchi 2009). In addition, the model shows the same wide range in severity of the effects. However, this mouse model may be of limited value to investigate the mechanism underlying the disorder in humans. The shortened limbs, reduced number of digit, or loss of other bony elements which is observed in 30-50% of individuals with CdLS, are not present in this mouse model. The model exhibits only atrial septal defects, while CdLS individuals have atrial and ventricular septal defects. Another difference is the corneal opacities observed in the model that are atypical for CdLS. The data using *NIPBL*$^{+/-}$ mice support autosomal dominant inheritance, but the significant differences raise questions regarding the validity of the data collected beyond what is confirmed through other model systems.

Another model system is pectoral fin buds of Nipbl-deficient zebrafish larvae. Transcriptome analysis of *NIPBL* haploinsufficient mouse limb buds show many changes in gene expression similar to the zebrafish pectoral fin bud (Muto 2014). Both the mutant pectoral fin buds and mouse limb buds show significant decrease in size (Muto 2014). For instance, at 72 hours post fertilization, the zebrafish larvae pectoral fin had a 40% decrease in size. Evidence separate from an overall developmental delay is based on partial rescue by injection of exogenous Nipbl mRNA. In addition, decreased limb bud growth occurs in the absence of cell death, suggesting a cumulative effect of slower rates of cell division as the cause of the reduced size (Muto 2014).

While the data may mirror what occurs in CdLS, the zebrafish pectoral fin may not be a good model. These investigators identify the zebrafish pectoral fin as a homolog of the mammalian forelimb; however, this fin is more analogous to a human hand or wrist. Also, the fin bones develop by intramembranous ossification in contrast to human long bones developing by endochondral ossification (personal communication Dr. M. Katherine Iovine 2/24/2015). Although not ideal, these models continue to provide useful information which can validated by human cells carrying CdLS mutations.

V. Mechanisms Posited to Produce Developmental Defects

V.A. Cohesin Structure and Function in Sister Chromatid Cohesion

Cohesin is a protein complex consisting of four subunits. The protein complex is evolutionary conserved, suggesting vital functional significance, and is documented for many species including *Saccharomyces cerevisiae*, *Schizosaccharomyces pombre*, *Drosophila melanogaster*, *Xenopus laevis*, and *Homo sapiens*. In *S. cerevisiae*, the subunits are called Smc1,

Smc3, Mcd1/Scc1, and Scc3/Irr1. Smc1 and Smc3 are the structural maintenance of chromosome units. They are rod-shaped cohesin subunits. Each SMC subunit has ATPase activity in the nucleotide-binding domain (NBD) at one end, and a dimerization domain on the other end. The latter domain creates the V-shaped hinge to produce the Smc1-Smc3 heterodimer. Scc1 bridges the ATPase heads of the SMC subunits to form a ring-shaped structure, and Scc1 recruits Scc3. The human homologs for Smc1, Smc3, Scc1, and Scc3 are SMC1, SMC3, RAD21, and STAG, respectively (Table 3). Humans also contain the protein SORORIN, which is required to maintain cohesion, although lower species do not possess SORORIN.

Additional proteins associate with the cohesin complex, while the remaining cohesion factors play a non-structural role. For instance, the loading factors Scc2 and Scc4 in budding yeast (NIPBL and MAU2, respectively in humans) are needed during G1 phase of the cell cycle for the loading of cohesin onto DNA. Eco1 in yeast (ESCO1 and ESCO2 in humans) functions in the establishment of cohesion between sister chromatids during S phase. Pds5 is necessary to maintain cohesion. It binds cohesins, balancing cohesin establishment with cohesin dissociation. Rad61 in budding yeast (WAPL in humans) functions in cohesin dissociation. Hos1 (HDAC8 in humans) deacetylates the Smc3 subunit and recycles cohesin (Borges 2010). Each of these non-structural, cohesin-associated proteins, therefore, have multiple roles in chromosome structure, cell division, repair of DNA damage, and gene transcription.

Mapping the gene mutations through which RBS and CdLS arise provided important insights regarding the mechanism involved in these developmental maladies. For instance, *ESCO2*, *NIPBL*, *RAD21*, *SMC1A*, *SMC3* and *HDAC8* all function in a singular pathway to tether together DNA molecules (Table 3). The complicating issue as to why it is difficult to assign

mechanism to disease is that this tethering participates in numerous distinct activities. For instance, *SMC1*, *SMC3* and *RAD21* form a protein complex termed cohesin. This complex is implicated in several functions: chromosome segregation, DNA replication, DNA repair, and regulating gene expression both in proliferating and post-mitotic cells. Cohesins were first identified for their role in sister chromatid pairing. Cohesins tether together the products of chromosome replication, identifying them as sisters until the onset of anaphase. The cohesin one-ring model suggests a mechanism through which the cohesin ring structure acts to physically entrap paired sister DNAs and hold them together during metaphase of mitosis and meiosis. Breakage in the cohesin ring can cause early separation of sister chromatids (Tóth 1999). Cohesin mutations result in aneuploidy—a common attribute of cancer cells. Cohesins also can tether together DNA elements along the same chromosome. In this way, cohesins can generate large DNA loops that allow for the close positioning of promoters and enhancers that is required for transcription of many genes.

Cohesinopathies develop from mutations affecting cohesin structure or associated proteins responsible for the establishment, loading, or maintenance/ removal of cohesin. Table 3 lists key cohesin subunits and cohesin regulators, along with their functions in several species. The following sections offer a discussion on cohesin mechanisms implicated in RBS and CdLS from which I formulate a hypothesis as to the interrelationship of these disorders and possible unifying underlying mechanism.

V.B. Impairment of Establishment of Sister Chromatid Cohesin by RBS

Establishment of cohesion is a process in which chromatin-associated cohesin becomes cohesion-“able” during S phase. It is an essential step since sister chromatid cohesion will not

occur simply in the presence of functional cohesin subunit loading proteins and DNA replication (reviewed by Vega 2005). The discovery of Eco1/Ctf7 in yeast, and subsequent characterization of the human homolog ESCO1 and ESCO2, resulted in the first model of establishment activities as cohesin activation. Eco1 function is essential only during S phase, when the protein acetylates lysine residues in the Smc3 subunit of cohesin. Eco1 mutants exhibit precocious sister separation. This cohesion defect is rescued by elevated levels of DNA replication factors such as PCNA. Eco1 also associates with DNA replication factors (including PCNA). Eco1 acetylates only cohesins that are chromatin-bound. Cohesin most likely undergoes a conformational change due to the acetylation before it can physically hold the sister chromatids together. Cells with Eco1 mutation arrested prior to the beginning of anaphase show separated sister chromatids (Skibbens 1999). The Eco1 links the establishment of sister chromatid adhesion to the DNA replication machinery (Skibbens 1999).

ESCO2 mutations create the observed RBS cellular phenotype: premature separation of centromeres, lagging chromosomes, aneuploidy, and micronuclei due to mis-segregation of chromosomes, decreased cell proliferation due to mitotic arrest or delay, and sensitivity to DNA-damaging agents (Vega 2005). Similar phenotypes were observed in zebrafish– for instance, esco2 knockdown results in an increase in both TUNEL staining which is used to detect DNA fragmentation resulting from apoptotic signaling cascades, and Caspase activation which is a family of cysteine protease enzymes that are the executioners of apoptosis (Mönnich 2011). Thus, the consensus of investigations with human mutant cell lines, mice, zebrafish, and medaka is that the multi-system wide-range in severity of individuals with RBS may be due to the increased apoptosis and fewer progenitor cells due to the errors generated in mitosis caused by *ESCO2* mutations.

The chromosome region(s) affected by the mutation varies with the species. Yeast Eco1 mutants affect cohesion both in the centromeres and chromosome arms, while fly Deco mutants disrupt cohesion only at the centromeres. Human *ESCO2* mutants disrupt cohesion specifically at heterochromatic C-banding regions around centromeres and the distal portion of the long arm of the Y chromosome (reviewed by Vega 2005). Differential regulation is proposed of arm versus centromere cohesion in humans since most cohesin in already removed from chromosome arms prior to the removal of cohesin at centromeres (Giménez-Abián 2004). Differential regulation may mean that another influence is behind the range of effects seen in RBS patients; the ESCO2 protein may have more than one function, or the singular function may initiate a cascade of overlapping effects.

V.C. Transcriptional Basis for CdLS Impairment?

Cohesin loading is necessary for the association of the cohesin complex with chromosomes during S-phase. The genes Scc2 and Scc4 are responsible for cohesin loading in budding yeast. *NIPBL* and *MAU2* are the human homologs of Scc2 and Scc4, respectively. The homology suggests that *NIPBL* and *MAU2* perform analogous roles in cohesin loading, however, CdLS cells do not show HR nor exhibit elevated levels of precocious sister chromatid separation exhibited by most cohesin mutations. Moreover, cultured *NIPBL*$^{+/-}$ mouse embryo fibroblasts, *NIPBL*$^{+/-}$ embryonic stem cells, or adult B-lymphocytes failed to produce aneuploidy or cohesion defects (Kawauchi 2009). These results suggest that the multi-system disorders of CdLS with a mutation in *NIPBL* is caused by a mechanism distinct from that of chromosome mis-segregation (cohesion defects). But how can mutations in Scc2, which loads cohesins, not have a segregation defect? One possibility is that, in the heterozygous state, sufficient Scc2 protein is generated to

function in cohesin loading, but insufficient quantity necessary to prevent CdLS via the protein's use in a separate mechanism.

In the absence of cohesin defects that might lead to progenitor cell death, numerous studies instead point to alterations in transcription as the basis of CdLS. In general, initiation of transcription is a major control point in gene expression. A particular cell state is created or maintained by gene expression programs, and these are controlled by transcription factors. The transcription factors bind to control elements called enhancers. The location of an enhancer may be far from where it binds to the core promoter elements. The enhancer-bound transcription factors *in vitro* bind coactivators, which subsequently bind the transcription initiation apparatus during active transcription. For this to occur, the physical bending of DNA into a loop excludes/displaces the intervening DNA such that distal DNA elements come into close proximity. Thus, DNA loops most likely form during active transcription by selective gene activation. Chromosome conformation capture studies show that some cohesin dependent enhancers are brought close to the promoter during active transcription (reviewed by Kagey 2010).

The most widely accepted model is that *NIPBL* may influence long-range chromosomal regulatory interactions which could cause collective or synergistic effects—that cohesin regulates development by controlling the expression of many other genes. Supporting this hypothesis, analysis of the RNA in the $NIPBL^{+/-}$ mouse model indicated a decrease of approximately 30% in NIPBL transcript levels, which subsequently altered protein expression in several hundreds of genes (Kawauchi 2009). Differential expression was detected in 978 genes of the $NIPBL^{+/-}$ embryonic brain itself. The findings reflect the extreme sensitivity of development to small changes in *NIPBL* activity.

The reduction in hundreds of gene transcript and protein expression caused by *NIPBL*$^{+/-}$ is similar to the decrease causing CdLS in people. However, since a 50% decrease was expected if the mutated allele of the *NIPBL* gene is nonfunctional, the *NIPBL* gene may be autoregulatory. An up-regulation of the wild-type copy most likely occurs to compensate for the mutant *NIPBL* copy and permits sufficient quantity of delangin to be synthesized to function in chromatid cohesion during cell replication. The amount of delangin may be insufficient, however, for interphase functions related to gene regulation during development, and this would account for the phenotype observed in CdLS patients.

All 16 mutant cell lines from CdLS individuals have a unique profile of dysregulated gene expression which could prove useful for diagnosis of the disorder (Liu 2009). Specifically, *NFATC2*, *PAPSS2*, and *ZNF608* were identified as biomarkers for CdLS. These genes also represent affected pathways that may contribute to CdLS phenotypes. *NFATC2*, for example, is involved in multiple signaling pathways during development of muscle, cartilage, and axons. *NFATC2* dysregulation was subsequently validated in a larger human study (Liu 2009). The extent that *NFATC2*, *PAPSS2*, and *ZNF608* dysregulation correlate with CdLS phenotype severity, however, remains undefined.

Numerous lines of evidence support the models that NIPBL regulates transcription and that transcriptional dysregulation results in CdLS phenotypes. *Cebpb* and *EBF1*, which encode transcriptional factors key to the differentiation of adipocytes, are down-regulated in *NIPBL*$^{+/-}$ mouse embryo fibroblasts (Kawauchi 2009). This regulation may explain the decreased body fat observed in the mouse model and CdLS individuals. Knockdown of *NIPBL* expression in zebrafish fin bud causes subsequent dysregulation in the expression of important early limb development genes, including *fgfs*, *shha*, and *hand2* (Muto 2014). Much of this work was

duplicated in mouse embryonic limb buds, but not every aspect was tested in both models. For example, reduced zone of polarizing activity (ZPA) expression is shown in *NIPBL*$^{+/-}$ mouse limb buds, but there is no data for the pectoral fin. Figure 2 summarizes the cascade of effects in mouse limb bud development investigated by Muto *et al* (2014). Thus, a *NIPBL* mutation could explain the limb defects observed in people with CdLS. In addition, *NIPBL* haploinsufficiency in mice is detrimental to the DNA looping that controls the selective expression of erythrocyte beta-globin (reviewed by Muto 2014). These data suggest that *NIPBL* has more than one role in development since no single target, when mis-regulated, can explain everything.

Knockdown of *NIPBL* expression in both the mouse limb bud and zebrafish pectoral fin bud models also causes dysregulation expression of multiple *hox* genes, or homeotic genes, which control the body plan of an embryo along its anterior to posterior axis (Muto 2014). The changes in expression of *hox* genes in fin and limb bud correlate with its position in the genome. This effect may explain the range of severity of the effects observed in individuals with CdLS.

Another series of investigations revealed that the *Drosophila melanogaster NIPPED-B* gene regulates Notch receptor signaling and the homeobox genes *cut* and *Ultrabithorax (Ubx)*, which facilitates enhancer-promoter communication (reviewed by Krantz 2004). Mutations in fly *cut* causes leg and wing abnormalities, and the mouse homolog *Cutl2* (or *Cux2*) causes anomalies in the branchial arch and limb bud (reviewed by Tonkin 2014). *Ubx* suppresses limb formation in the fly abdomen by repressing the gene Distalless (*Dll*) which is essential for distal limb development. The Dlx family of mammalian *Dll* homologs are associated with limb, neural, and other developmental processes (reviewed by Tonkin 2014). *NIPBL* mutation may mistakenly activate DLX genes and cause the limb malformations observed in CdLS individuals (Tonkin

2014). The assumption of a DLX role for vertebrate langins functioning in CdLS needs to be substantiated before accepting the proposed as true because other mechanisms are plausible.

V.D. Proposed Mechanisms of How Cohesin May Impact Transcription

Data suggest that cohesin can function as a transcription factor via the location and number of its binding sites. In control lymphoblastoid cells, cohesin preferentially binds to the starter (promoter) regions of actively expressed genes where there are an increased number of binding sites, as compared to non-transcribed genes. In contrast, human CdLS cells have a general loss of cohesin binding sites, and a greater loss at transcription start sites. The result is an 80% cohesin binding preference towards intergenic regions (Liu 2009). The dysregulation in *NIPBL* mutation could result from decreased cohesin binding to starter regions, preventing gene expression.

An alternative mechanism for how cohesin might impact transcription accounts for the observation that the majority of genes do not carry known cohesin binding sites (Liu *et al* 2009). Thus, cohesin might act like a boundary/insulator protein in non-coding DNA regions. This boundary model is supported by findings that NIPBL or cohesin binds CTCF. CTCF (also known as transcriptional repressor CTCF, 11-zinc finger protein, and CCCTC-binding factor) is a vertebrate insulator capable of blocking enhancers from interacting with promoters or preventing the spread of epigenetic signals. Thus, cohesin interactions with the transcriptional regulator CTCF may be critical regulators of chromatin architecture.

Relationships between NIPBL, Mediator, and cohesin in transcription are demonstrated in laboratory experiments. Mediator is a large complex composed of a core which interacts with the gene specific transcriptional regulators and RNA polymerase II, and a Cdk8 submodule that

regulates transcription positively or negatively (reviewed by Muto 2014). The use of a small hairpin RNA library links NIPBL, cohesins, and transcription regulators needed for the maintenance of murine embryonic stem cells (Kagey 2010). The application of Hoechst to the stem cells stained DNA and enabled quantifiable changes in the pixel density of Oct4, a known master regulator of the pluripotent state. In addition to Oct4, most known regulators of the murine embryonic stem cell state were identified, including Sox2 and Nanog. Importantly, additional components were found which implicate them as important for the maintenance of the stem cell state. The additional components include many subunits of the Mediator complex, cohesin complex, and the cohesin loading factor *NIPBL*. The results suggest that murine embryonic stem cell pluripotency is sensitive to changes in *NIPBL*, Mediator, and cohesin components.

If Mediator and cohesins mirror Oct4 levels in stem cells, perhaps they also regulate the stem cell fate? Indeed, a knock down of Mediator, cohesin, or *NIPBL* all produced the same effects as knocking down Oct4—a loss of embryonic stem cell state. This loss of state is defined by a decrease in Oct4 protein, change in cell colony morphology, decreased mRNAs specifying transcription factors linked to pluripotency, and increased mRNAs specifying transcription factors for development, although the relationship between colony morphology and pluripotency isn't discussed (Kagey 2010). The results suggest a very important possibility—that Mediator, cohesin, and *NIPBL* have an inter-related and essential role in gene expression.

To further investigate how Mediator functions, chromatin immunoprecipitation coupled by massively parallel DNA sequencing (ChIP-Seq) was employed. Mediator was found to occupy the enhancers and the promoter regions of more than 60% of actively transcribed genes in murine embryonic stem cells (Kagey 2010). Mediator is a transcriptional coactivator that

plays a key role in transmitting regulatory signals from gene-specific activators and repressors to RNA polymerase II. Cohesin and Mediator co-localize on active genes. A population of cohesin is associated with the enhancer and core promoter sites occupied by Mediator in many active promoters. Whether interactions between Mediator and cohesin are direct or mediated by DNA remains unclear – an issue that will require new studies that include DNAse treatment. In addition, Mediator pulls down only a very small fraction of cohesin in this study, making physiological relevance of this interaction questionable. Despite these limitations, the combination of results suggest that when transcription activators bind Mediator, the complex undergoes a conformational change and then binds cohesin and *NIPBL* (Kagey 2010). In this model, cohesin connects the enhancer to the core promoter site and RNA polymerase, while *NIPBL* associated with Mediator-cohesin complexes is the factor that loads cohesin at promoters. This activity occurs independently of CTCF (reviewed by Nolen 2013). A schematic of the interaction of Mediator, *NIPBL*, and cohesin in transcription is shown in Figure 1.

Is the interaction of *NIPBL* and Mediator conserved? The zebrafish pectoral fin bud and mouse limb bud models show that *NIPBL* and Mediator cooperatively regulate gene expression to control limb development (Muto 2014). Deficiency in Mediator subunit Med12 in the zebrafish pectoral fin larvae cause morphology and gene expression changes similar to *NIPBL* deficiency, and incomplete deficiency of both *NIPBL* and Med12 together create a strong synergistic effect (Muto 2014). These results support both *NIPBL* and Mediator acting together in a common pathway. They regulate the expression of *hox* genes in developing limb and fin buds by modulating the interaction of promoters with remote enhancers. This is demonstrated in 3-D fluorescent in situ hybridization with the zebrafish *hoxda* cluster (Muto 2014). The model that NIPBL regulation is gene specific is undermined by the lack of evidence that any genes

remain unaffected by NIPBL knockdown. For instance, while the effects on *fgf24* gene expressions upon reduced NIPBL levels is modest, these are clearly reduced – in contrast to the authors' claims that there is no difference.

Is cohesin-dependent DNA looping required only for differentiation, or does it also maintain the differentiation state? The difference is an important one given that DNA loops may be transient or long-lived and thus persist during DNA replication and transcription. For instance, DNA looping is observed between enhancers and promoters co-occupied by Mediator and cohesin in murine embryonic stem cells and mouse embryonic fibroblasts (Kagey 2010). Mediator and cohesin co-occupy different promoters in different cells, enabling cell type-specific DNA loops linked to the selective gene expression program of each cell (Kagey 2010). Thus, the specific DNA loops are associated with gene activation or suppression as a consequence of the different promoters that Mediator and cohesin co-occupy. Problems with the ability to retain looping, however, might result in the inability to maintain the state of a cell type. The investigators propose that loss of gene transcription control may be the primary cause of CdLS and other cohesinopathies that cause a diverse range of developmental defects.

V.E. *NIPBL* Cohesin-Independent Role in Development

Is NIPBL's role in transcription dysregulation anchored in cohesin function? Just this past year, investigators discovered that *NIPBL* may also have a cohesin-independent role in development—functioning at promoter sites as a transcription cofactor (Zuin 2014). Using mitotic HeLa cells fixed with paraformaldehyde and immunostained with antibodies specific for NIPBL, CTCF, RAD21, and SA2/STAG2, NIPBL and CTCF were found to precede cohesin binding to DNA with NIPBL associating at an early stage of the mitotic exit. Conversely,

cohesion appearance onto chromatin occurred during or after nuclear envelope reformation. The mitotic stage was determined using the marker Lamin B-EGFP which correlates with the reassembly of the nuclear envelop in the HeLa cells. Based on this timing, it is proposed that cohesin is recruited to NIPBL-bound chromatin. However, additional studies are required to document that NIPBL functions in transcription separate from cohesin function, since cohesin binds DNA long after transcription profile is put into play.

There is evidence that supports the model that NIPBL acts independent of cohesin during transcription regulation. In HB2 cells, *NIPBL* localizes in somatic cells independently of cohesin (Zuin 2014). This function of *NIPBL* initially seems like a contradiction to the previously mentioned co-localization of *NIPBL* and cohesin at enhancers and core promoter regions of transcriptionally active genes (which are also bound by Mediator in mouse embryonic stem cells) (Kagey 2010). However, Zuin *et al* (2014) found the presence of two different *NIPBL* binding sites based on the use of different *NIPBL* antibodies in ChIP sequencing protocols on HB2 cells and human lymphoblastoid cells. The significance of cell type used is unclear. However, NIPBL#1 antibodies used by Zuin (2014) bind to highly enriched "major sites" that localize at promoters and do not overlap with cohesin. In contrast, NIPBL#6 antibodies used by Kagey (2010) and NIPBL#1 bind to low-enriched "minor sites" which overlap with cohesin binding sites. Approximately 80% of *NIPBL* "major binding sites" localize to nucleosome-free promoter areas which are less than 1000 base pairs from transcription start sites, in contrast with 10% of cohesin and CTCF sites localizing to promoters (Zuin 2014). Analysis of RNA sequencing indicate that more than 98% of the *NIPBL*-bound genes at major binding sites are actively transcribed. Analysis of major binding profiles and heat maps on human lymphoblastoid cells identified five transcription factors on *NIPBL* sites: NFYA/NFYB, SP1, PBX3, C-FOS, and

IRF3 (Zuin 2014). Unfortunately, the studies stopped short of testing whether mis-regulation of these genes phenocopy birth defects seen in CdLS, or whether upregulating these genes in *NIPBL* knockdowns rescues CdLS phenotypes. These genes are linked to several different cell functions including control of the cell cycle, cell death, gene expression, control of cellular growth and multiplication, and RNA post-translational modification. The continued presence of *NIPBL* and *MAU2* bound to protein coding genes was found necessary to maintain the levels of gene expression. Investigators concluded that altered gene expression when *NIPBL* is depleted is not due to decreased cohesin binding or cohesin's function in transcription (Zuin 2014).

Also independent from cohesin, *NIPBL* haploinsufficiency downregulates the canonical Wnt pathway in zebrafish embryos and CdLS patients fibroblasts (Pistocchi 2013). The canonical Wnt pathway is a signal transduction pathway involved in the regulation of gene transcription during embryonic development, controlling processes that include cell proliferation, cell fate or differentiation, cell migration, and body axis. While investigating the effects of *nipblb* loss of function in the developing hindbrain in zebrafish embryos, *NIPBL* expression was found localized in the central nervous system at 24 hours post fertilization (hpf). Injected with *nipblb*-morpholino, 50% (100/200) of the embryos were microphthalmic and microcephalic at 24 hpf, and Western blot analysis determined Nipbl protein level reduced by 39% compared to controls (Pistocchi 2013). Increased apoptosis, with no alteration in proliferation, was observed in the central nervous system of *nipblb*-knocked down embryos in the areas where *nipblb* is normally expressed. Based on expression of AP identity markers hoxb2a, pax2a, and krox20, and the marker *atoh1* of hindbrain dorsal progenitors and ventricle opening, central nervous system axis and anatomical patterning may be unaffected by *nipblb*- morpholino at 24 hpf. Also unaffected was neuronal differentiation of the dopaminergic cell population. However, 80%

(96/120) had mis-regulation of *wnt1* expression, with the severity of phenotype correlating with the extent of hindbrain fusion. *axin2* expression, a key downstream component of the wnt pathway, was similarly mis-regulated (Pistocchi 2013).

One direct target of the canonical Wnt pathway, *Cyclin D1* or *CCND1*, was downregulated in *NIPBL* haploinsufficiency. Fibroblasts from patients with CdLS had almost 50% decrease in *CCND1* (Pistocchi 2013). *Ccnd1* levels in *nipblb*- morpholino injected zebrafish embryos showed decreases by q-PCR, western blot analysis, and whole-mount *in situ* hybridization (WISH). Also observed were decreases in the active form of beta-catenin in the CdLS fibroblast and zebrafish models. The phenotype of the *nipblb*- morpholino injected zebrafish embryos could be rescued by chemically inducing the Wnt pathway with lithium chloride. Results included restored morphology of cephalic structures, *wnt1* distribution pattern similar to control embryos, reduction in elevated apoptosis levels, increased active form of beta-catenin, and increased *cnd1* levels.

Therefore, the evidence suggests that *NIPBL* plays multiple roles in developmental regulation by loading cohesin onto the chromatin, indirectly alter gene expression via long range chromosomal interactions in conjunction with cohesin and with or without CTCF. These long range interactions maintain the architecture in higher-order chromatin structure, directly affecting genes when *NIPBL* binds to promoters, and impacting the canonical Wnt pathway. It is presently unknown how many gene expressions are primarily due to direct *NIPBL* transcription, or how many are secondary effects caused by gene dysregulation.

V.F. Impact of *ESCO2* on Transcription Regulation of Genes Downstream

Since *NIPBL* regulates transcription in development, might ESCO2 also have this function? A microarray analysis of Esco2-depleted zebrafish embryos was used to examine the genes regulated downstream of Esco2 (Mönnich 2011). Instead of enrichment for transcriptional regulators responsible for the generation of CdLS, the RNAs differentially expressed in *esco2* morphants were enriched for RNAs associated with the cell cycle—cell proliferation and apoptosis—instead of any particular functional category or developmental gene mis-regulation. The data support high levels of cell proliferation defects and cell death created by the disruption of the cell cycle as the primary cause of RBS features, in contrast to the alteration of developmental pathways in CdLS. Therefore, Mönnich *et al* (2011) conclude that RBS and CdLS are distinct diseases, and that they are grouped as cohesinopathies because both have mutations in genes in proteins responsible for sister chromatic cohesion.

V.G. Cohesinopathy vs. Trancriptomopathy

Identified as a single group of disorders permits speculation that all cohesinopathies have a common underlying mechanism. Recently, Yuan *et al* (2015) proposed transcriptional disturbances as the common underlying mechanism unifying several developmental maladies, including CdLS. Claiming that the syndromes exhibit similar or overlapping CdLS-like manifestations in multiple systems and broad range of severity, the investigators would prefer to call the collection of syndromes "transcriptomopathies," each syndrome caused by mutations in different genes related to transcription; CdLS should not be called a cohesinopathy but a transcriptomopathy. The eight syndromes reviewed include Wiedemann-Steiner syndrome (WDSTS; Hairy Elbows Syndrome), Floating-Harbor syndrome, and Rubinstein-Taybi syndrome, in addition to CdLS. Roberts Syndrome was not included in the study. The

justification for reviewing these eight syndromes together, according to the researchers, are the overlapping phenotypes. However, when examining the table of characteristics, the commonality isn't as obvious. For example, excessive hair growth of the upper and lower arm around the elbow area is a prominent WDSTS clinical feature, while general excessive hairiness is common in CdLS in both sexes, but not in all CdLS individuals. Another reason for the review of these syndromes collectively are the mutations that the investigators observed.

CdLS-like phenotypes are caused by mutations in different genes, while WDSTS is caused by mutation in lysine-specific methyltransferase 2A (KMT2A) on the X-chromosome. KMT2A is thought to be a broad regulator of gene transcription. Whole-exon sequencing (WES) for two brothers diagnosed with WDSTS identified a hemizygous missense mutation in *SMC1A* (Yuan 2015). It is unclear in the publication if the two brothers also have a KMT2A mutation, but *SMC1A* mutation is known to cause CdLS. Since syndromes and diagnoses are first identified through observations made by clinicians, and due to the broad spectrum and varied intensity of effects for each of these syndromes, it is possible that instead of a common mechanism unifying WDSTS and CdLS, an alternative explanation may be that the brothers were incorrectly diagnosed. Similarly, one of 32 Turkish patients diagnosed with CdLS and having clinical features of CdLS (and not WDSTS) had a novel heterozygous nonsense KMT2A mutation, detected by WES using DNA samples extracted from the peripheral blood. In addition, one patient with characteristics of both CdLS and WDSTS had a mutation in SMC1A, while another patient with characteristics of both CdLS and WDSTS had a mutation in SMC3 (Yuan 2015). Thus, it is unclear as to whether mis-diagnoses is prevalent in the diagnoses of these syndromes, or if there is a transcriptional commonality causing them. Also, since SMC1A

and SMC3 contributed to some phenotypes resembling WDSTS, perhaps the molecular mechanism by which WDSTS is generated involves cohesin genes and proteins.

Yuan *et al* (21015) also identify a population in Saudi Arabia with CdLS-like phenotype that would be the first report of an autosomal recessive inheritance responsible for the disease. This population has a variant in *TAF6* gene, coding for Transcription initiation factor TFIID subunit 6 which may function in activation or promoter recognition to facilitate the initiation of transcription by RNA polymerase II. The phenotype requires homozygous recessive mutation in *TAF6*. But should this population be associated with CdLS? Since RBS has autosomal recessive inheritance, the question is really what defines CdLS-like? Is a group of syndromes being made to fit a hypothesis?

VI. Final Thoughts- Discussion of Cohesinopathy Commonality

RBS, CdLS, and the other cohesinopathies share some common features but appear to be clinically distinct. Multiple systems are affected, albeit different in their number of systems, and they manifest themselves in a broad range of severity. To target several systems plus vary the extent to which targets are affected suggests to me simultaneous mechanisms: one direct mechanism that disrupts cohesion function, and the second indirect mechanism via transcription that can cause developmental gene dysfunction or cell cycle disruptions. Figure 3 describes my proposed interpretation of the dual mechanism in *NIPBL* mutation, which would occur in a simultaneous gradation. As with the case of *NIPBL*, a mutation in this gene prevents the loading of cohesin onto chromatin, and regulates development by controlling the expression of multiple other genes. A dual function exists for *ESCO2*, but the affected downstream genes are associated with the cell cycle. The same studies on binding to promoters, indirectly altering gene expression via long range chromosomal interactions in conjunction with cohesin and with or

without CTCF, and maintaining the architecture in higher-order chromatin structure should be performed with *ESCO2*. The dual functions of both *NIPBL* and *ESCO2* in cohesion function and transcription suggests a unifying mechanism that can explain how all cohesinopathies develop while still creating the distinct characteristics of each disorder. Assume a continuum of manifestations exists. At one end is the disruption of cohesion during S phase, causing increased errors in mitosis leading to apoptosis and aberrant development due to fewer progenitor cells. The other end is transcriptional control, either of the cell cycle or the dysregulated expression of multiple developmental genes during G_0 and G_1, which also causes aberrant development. A mutation in *ESCO2* is more of a mitotic error causing fewer progenitor cells, and *NIPBL* is more gene transcriptional dysregulation. The continuum has a central area with overlapping or common effects. Other cohesinopathies may fall somewhere in between on this continuum and would depend on the genes transcribed downstream from the mutation. Acetylation is necessary in both processes. Lysine residues in the Smc3 subunit of cohesin are acetylated before cohesin can physically hold together sister chromatids. Acetylation also tends to open up the coiled chromatin to enable transcription. It is important to note that the cohesinopathies are created by the loss of control in gene expression, while cancer can arise from improper chromosome segregation or an inability to repair DNA damage. Thus, dysfunction of cohesion causing a mitotic error can develop into a cohesinopathy, but it can also develop cancer if the abnormal cell survives and does not undergo DNA repair.

The continuum itself can be the cell cycle and the role cohesin plays in the different parts of the cycle. For example, the synthesis stage or S stage is DNA replication. Therefore, mutation in *ESCO2* prevents the establishment of cohesin with the result of mitotic errors due to failed sister chromatid cohesion. A *NIPBL* mutation disrupts normal growth and preparation for

DNA synthesis by dysregulation of gene transcription. Figure 4 shows my hypothesis of the links between the cell cycle phases, cohesin activities, *NIPBL* and *ESCO2* which can unify all cohesinopathies but still reflect their distinct differences.

Table 1.

Overview of Roberts Syndrome and Cornelia de Lange Syndrome		
	Roberts Syndrome (RBS)	**Cornelia de Lange Syndrome (CdLS)**
Alternate Names	Appelt-Gerken-Lenz Syndrome; Hypomelia-Hypotrichosis-Facial Hemangioma Syndrome; Pseudothalidomide Syndrome; RBS; Roberts-SC Phocomelia Syndrome; SC Phocomelia Syndrome; SC Pseudothalidomide Syndrome; SC Syndrome; Tetraphocomelia-Cleft Palate Syndrome	BDLS; Brachmann-De Lange Syndrome; CDLS; De Lange Syndrome
Frequency of Occurrence	~150 cases	1 in 10,000
Inheritance	Homozygous recessive	Autosomal dominant: *NIPBL*, *SMC3*; X-link dominant: *SMC1A*
Gene mutated	*ESCO2*: in 100% of cases	• *NIPBL*: present in >50% cases • *SMC1A*: 5% of cases • *SMC3*: rare 35% of cases unknown cause
Gene full name	*ESCO2* : "establishment of sister chromatid cohesion N-acetyltransferase 2"	• *NIPBL*: "Nipped-B homolog (Drosophila)." • *SMC1A*: "structural maintenance of chromosome 1A" • *SMC3*: "structural maintenance of chromosome 3"
Gene location	Chromosome 8 at position 21.1 from base pair 27,774,540 to base pair 27,812,623	• *NIPBL* on chromosome 5 at position 13.2 base pairs 36,876,758 - 37,065,823 • *SMC1A* on the X chromosome between positions 11.22 and 11.21 base pairs 53,374,148 - 53,422,727 • *SMC3* on chromosome 10 at position 25 base pairs 110,567,690 - 110,604,633
Defective protein	ESCO2 protein: functions in the establishment of cohesion between sister chromatids. Cells respond by delaying cell division and can be a signal that the cell should undergo self-destruction.	• Delangin, from *NIPBL*: controls the interaction between the cohesion complex and the DNA that makes up the sister chromatids. • SMC3 protein, from *SMC3*: part of a protein group called the cohesion complex that holds the sister chromatids together. • SMC1 protein, from *SMC1*: part of a protein group called the cohesion complex that holds the sister chromatids together.

Table 2.

Major Clinical Features in Robert's Syndrome and Cornelia de Lange Syndrome, and Some Reported Incidence Frequencies*

ROBERT'S SYNDROME (RBS)	CORNELIA DE LANGE SYNDROME (CdLS)
• Growth retardation- 100% • Mental retardation- ~60% - **No problems reported on behavioral parameters** • Craniofacial abnormalities - Head o Microcephaly- ~95% o Brachycephaly- ~63% - Eyes o **Exophthalmos**- ~60% o **Hypertelorism**- ~85% o **Corneal opacity**- ~36% - Ears o Misshapen- ~66% o **No hearing loss reported** - Nose/Mouth/Chin o **Narrowed nostrils**- ~*92%* o **Midfacial hemangioma- ~**68% o Cleft lip and palate- ~55% o Micrognathia- ~75% • **Limb abnormalities- 100%** - Phocomelia - Bilaterally, arms affected more than legs - **Predominantly all four appendages affected** • Cardiac defects- ~26% • **Enlarged phallus/clitoris**-~25-45% • Cryptorchidism- ~25-45%	• Growth retardation • Mental retardation - **Aggressive behavior, autism, self-inflicting injuries, seizures** • Craniofacial abnormalities - Head o Microcephaly o Brachycephaly - Eyes o **Ptosis** o **Myopia** o **No corneal opacity reported** - Ears o Misshapen o **Hearing loss** - Nose/Mouth/Chin o **Beak-like nose** o **No midfacial haemangioma** o Cleft lip and palate o Micrognathia • **Limb abnormalities- ~33%** - Phocomelia - Bilaterally, arms affected more than legs - **Predominantly only upper extremities affected** • Cardiac defects • **Underdeveloped male genitals** • Cryptorchidism • **Hypospadias** • **Small and abnormal kidneys** • **Unibrow (synophyrs) and overall excessive hairiness (hirsutism)** • **Gastro-esophageal dysfunction** • **Reduced body fat**

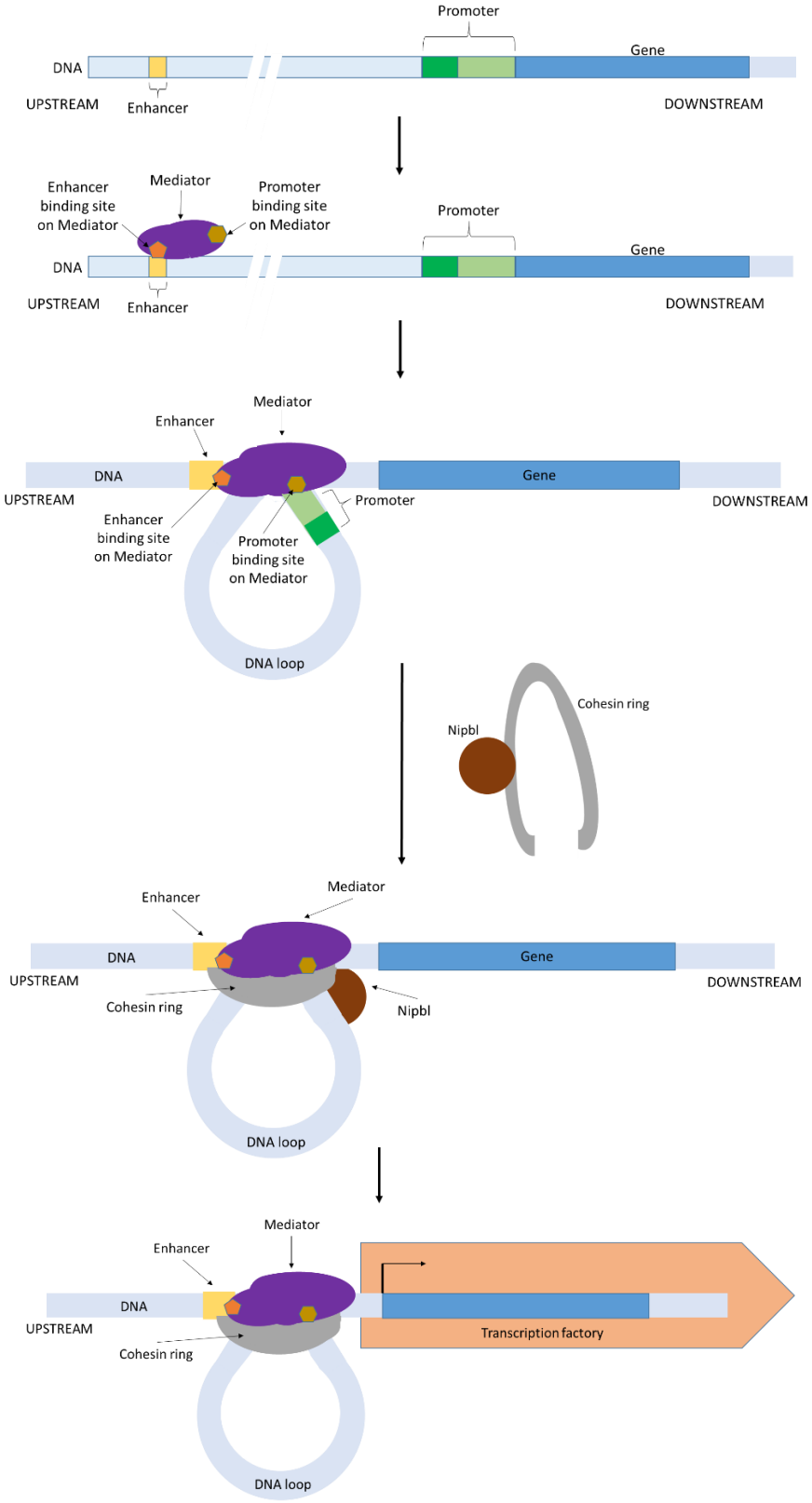
* Both RBS and CdLS individuals have growth retardation, reduced mental capacity, distinctive craniofacial malformations, and limb abnormalities; however, specific clinical details that differ in these characteristics are highlighted in bold, along with additional irregularities.

Key Subunits of Cohesin Complexes and Cohesin Regulators Implicated in Cohesinopathies in Several Species- *Saccharomyces cerevisiae*, *Schizosaccharomycies pombre*, *Drosophila melanogaster*, *Xenopus laevis*, *Dario reio*, and *Homo sapiens*

Chromosome Cohesion Regulator	*Saccharomyces cerevisiae* (yeast)	*Schizosaccharo-myces pombre* (fission yeast)	*Drosophila melanogaster* (fruit fly)	*Xenopus laevis* (frog)	*Dario reio* (zebrafish)	*Homo sapiens* (man)	Function
(SMC) Structural maintenance of chromosomes units	Smc1	Psm1	Smc1	smc1a; *smc1b* [predicted]	Smc1a; *Smc1b* [predicted]	SMC1A; *SMC1B*	Core cohesin subunit; *Core subunit (meiosis)*
	Smc3	Psm3	Cap/Smc3	smc3/cspg6 [predicted]	Smc3	SMC3	Core cohesin subunit
α-Kleisin subunit	Mcd1/Scc1	Rad21	Vtd/Rad21	rad21/mcd1/nxp1 /scc1 [predicted]	rad21a; rad21b	RAD21	Core cohesin subunit
	Rec8	*Rec8*	*C(2)M*	*rec8*	*Rec8/zgc*	*REC8*	*Core subunit (meiosis)*
Stromalin/SA unit	Scc3/IRR1	Psc3	DSA1	stag1/sa1; stag2/sa2 [predicted]	Stag1; Stag2 [predicted]	STAG1/SA1/SCC3A; STAG2/SA2/SCC3B	Cohesin subunit
Adherin/Kollerin	Scc2	Mis4	Nipped-B	nipbl/scc2/ delangin	Nipbla/Scc2a; Nipblb/Scc2b	NIPBL/SCC2/ DELANGIN	Cohesin loading
	Scc4	Ssl3	--	mau2/scc4 [predicted]	Mau2/zgc [predicted]	MAU2/SCC4	Cohesin loading
Interactors of α-Kleisin and SA	Pds5	Pds5	Pds5	pds5a	Pds5a/zgc [predicted]	PDS5A	Balances cohesin establishment/dissociation
	--	--	--	pds5b/as3/aprin [predicted]	Pds5b [predicted]	PDS5B/AS3/ APRIN	Balances cohesin establishment/dissociation
	?	?	Dmt (Dalmatian)	cdca5/sororin [predicted]	Cdca5	CDCA5/SORORIN	Balances cohesin establishment/dissociation
	Rad61/Wpl1	Wapl	Wapl	wapal	Wapl/KIAA [predicted]	WAPAL/WAPL	Balances cohesin establishment/dissociation
Cohesin acetyltransferase	Eco1/Ctf7	Eso1	Eco/Deco	esco1	Esco1 [predicted]	ESCO1	Establishment of cohesin
	--	--	San	esco2/rbs/efo2 [predicted]	Esco2	ESCO2	Establishment of cohesin
Cohesin deacetylase	Hos1	*Protein not yet identified*	*Protein not yet identified*	hdac8	Hdac8	HDAC8	Recycling of cohesin

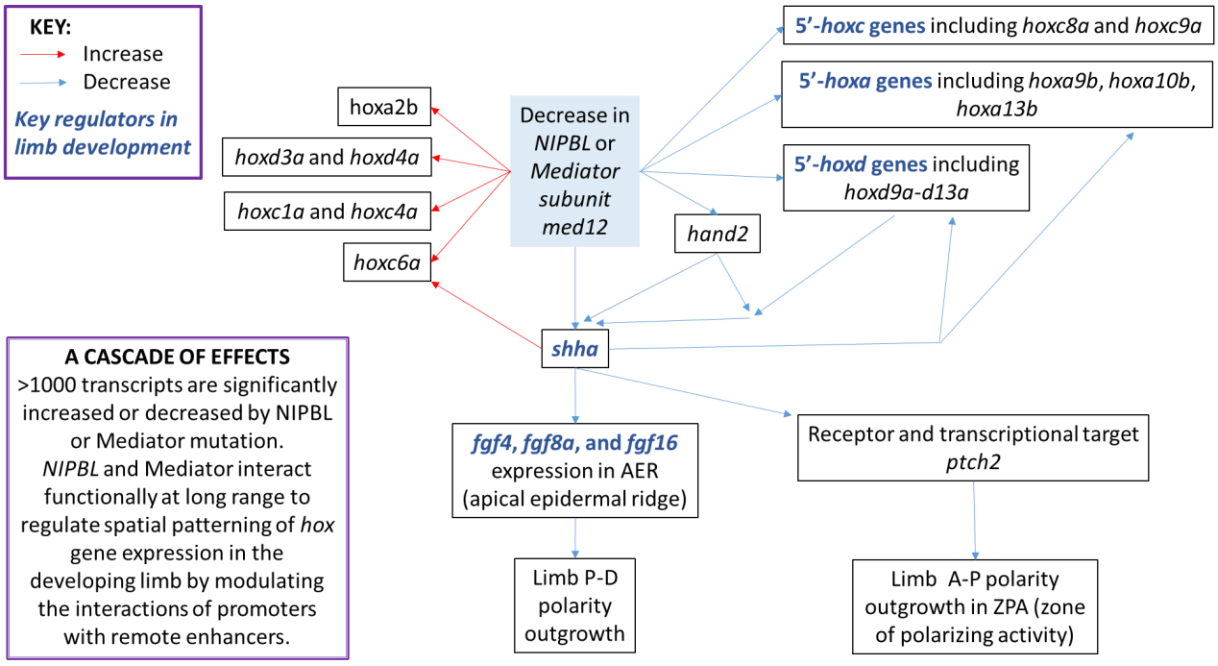
[Modified from: Horsfield, J.A., Print, C.G. & Front, M.M. Diverse developmental disorders from The One Ring: distinct molecular pathways underlie the cohesinopathies. Front. Genet., 12 September 2012 | doi: 10.3389/fgene.2012.00171]

Figure 1. Interaction of Mediator, NIPBL, and Cohesin in Transcription



Modified from: http://www.nature.com/nature/journal/v467/n7314/images/467406a-f1.2.jpg

Figure 2. Cascade of Effects in Limb Bud Development



[Data from Muto *et al*. (2014). Nipbl and Mediator Cooperatively Regulate Gene Expression to Control Limb Development. *PLOS Genetics* 10(9):1-22, e1004671]

Figure 3. Proposed Interpretation of the Dual Mechanism in NIPBL Mutation

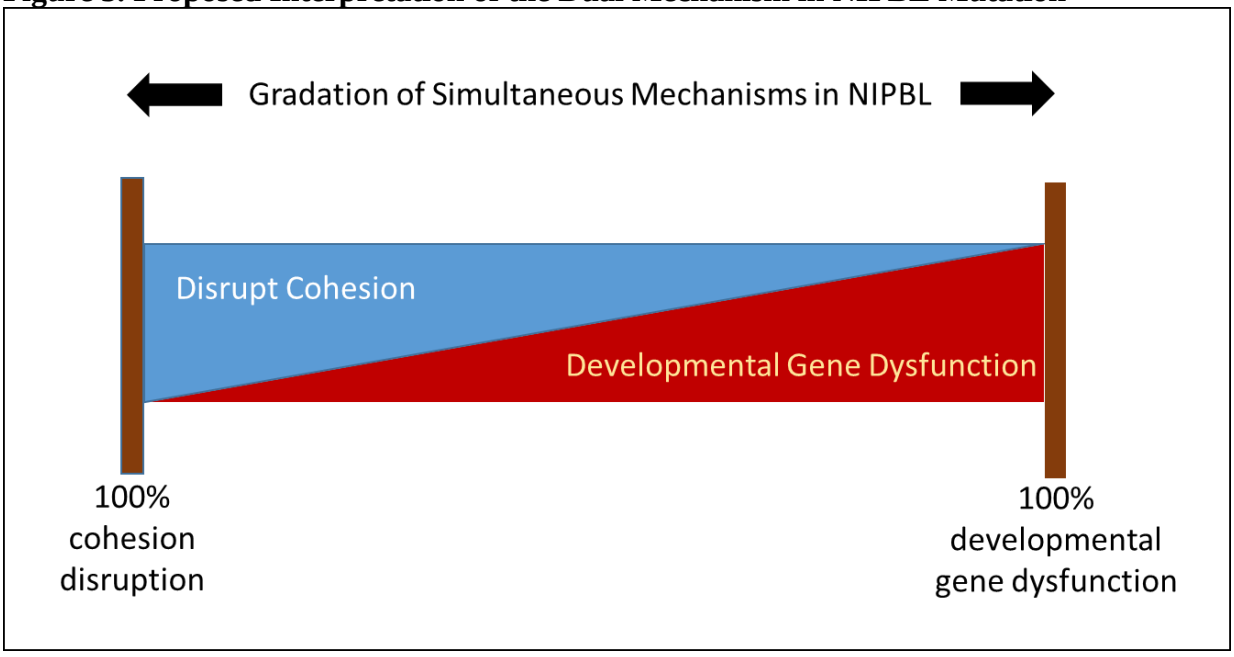
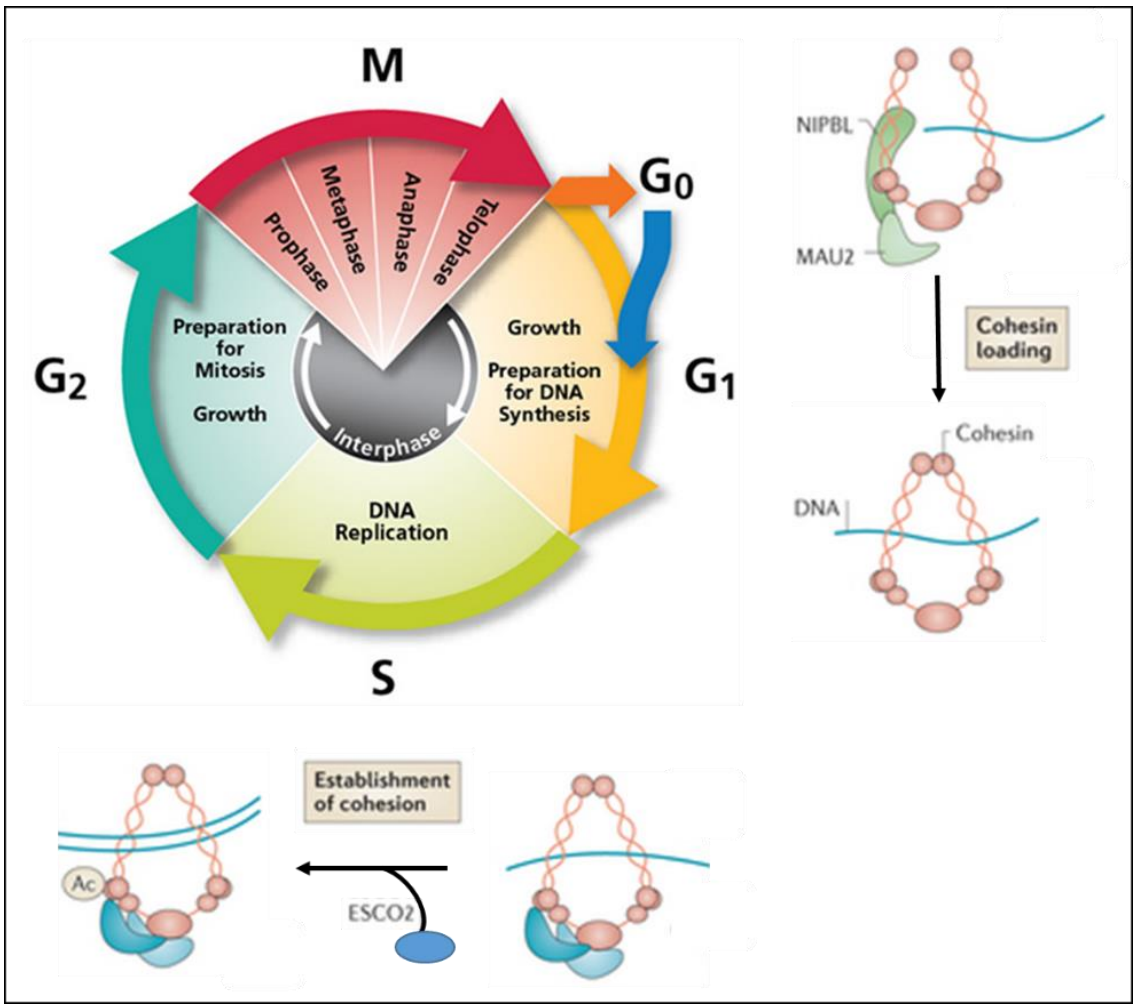


Figure 4. Linking Cell Cycle Phases, Cohesin Activities, and the Cohesinopathies of *NIPBL* & *ESCO2*



Modified from: http://www.nature.com/nrc/journal/v14/n6/images_article/nrc3743-f2.jpg and http://www.bdbiosciences.com/in/wcmimages/apoptosis_analysis_cellcycle_phases_lrg.jpg

References

Borges, V., Lehane, C., Lopez-Serra, L., Flynn, H., Skehel, M., Rolef Ben-Shahar, T., and Uhlmann, F. (2010). Hos1 Deacetylates Smc3 to Close the Cohesin Acetylation Cycle. *Molecular Cell* 39:677-688.

Genetics Home Reference. 13 April 2015. U.S. National Institute of Health. 15 April 2015. <http://ghr.nlm.nih.gov/>

Giménez-Abián, J.F., Sumara, I., Hirota, T., Hauf, S., Gerlich, D., Torre, C., Ellenberg, J. and Peters, J. (2004). Regulation of Sister Chromatid Cohesion between Chromosome Arms. *Current Biology* 14(13):1187-1193.

Kagey, M.H., Newman, J.J., Bilodeau, S., Zhan, Y., Orlando, D.A., van Berkum, N.L., Ebmeier, C.C., Goossens, J., Rahl, P.B., Levine, S.S., Taatjes, D.J., Dekker, J. and Young, R.A. (2010). Mediator and cohesin connect gene expression and chromatin architecture. *Nature* 467:430-435 (23 Sept. 2010).

Kaur, M., DeScipio, C., McCallum, J., Yaeger, D., Devoto, M., Jackson, L.G., Spinner, N.B. and Krantz, I.D. (2005). Precocious Sister Chromatid Separation (PSCS) in Cornelia de Lange Syndrome. *American Journal of Medical Genetics* 138(A):27-31.

Kawauchi, S., Calof, A.L., Santos, R., Lopez-Burks, M.E., Young, C.M., Hoang, M.P., Chua, A., Lao, T., Lechner, M.S., Daniel, J.A., Nussenzweig, A., Kitzes, L., Yokomori, K., Hallgrimsson, B. and Lander, A.D. (2009). Multiple Organ System Defects and Transcriptional Dysregulation in the *Nipbl*$^{+/-}$ Mouse, a Model of Cornelia de Lange Syndrome. *PLOS Genetics* 5(9):1-17, e1000650.

Krantz, I.D., McCallum, J., DeScipio, C.,Kaur, M., Gillis, L.A. *et al*. (2004). Cornelia de Lange syndrome is caused by mutations in *NIPBL*, the human homolog of *Drosophila melanogaster Nipped-B. Nature Genetics* 36(6):631-635.

Liu, J., Zhang, Z., Bando, M., Itoh, T., Deardoff, M.A., Clark, D., Kaur, M., Tandy, S., Kondoh, T., Rappaport, E., Spinner, N.B., Vega, H., Jackson, L.G., Shirahige, K. and Krantz, I.D. (2009). Transcriptional Dysregulation in NIPBL and Cohesin Mutant Human Cells. *PLOS Biology* 7(5):1-16, e1000119.

Mönnich, M., Kuringer, Z., Print, C.G. and Horsfield, J.A. (2011). A Zebrafish Model of Roberts Syndrome Reveals that Esco2 Depletion Interferes with Development by Disrupting the Cell Cycle. *PLoS ONE* 6(5):1-12, e20051.

Muto, A., Ikeda, S., Lopez-Burks, M.E., Kikuchi, Y., Calof, A.L., Lander, A.D. and Schilling, T.F. (2014). Nipbl and Mediator Cooperatively Regulate Gene Expression to Control Limb Development. *PLOS Genetics* 10(9):1-22, e1004671.

Nolen, L.D., Boyle, S., Ansari, M., Pritchard, E. and Bickmore, W.A. (2013). Regional chromatin decompaction in Cornelia de Lange syndrome associated with NIPBL disruptions can be uncoupled from cohesin and CTCF. *Human Molecular Genetics* 22(20):4180-4193.

Pistocchi, A., Fazio, G., Cereda, A., Bettini, L.R., Messina, G., Cotelli, F., Biondi, A., Selicorni, A. and Massa, V. (2013) Cornelia de Lange Syndrome: *NIPBL* haploinsufficiency downregulates canonical Wnt pathway in zebrafish embryos and patients fibroblasts. *Cell Death and Disease* 4:1-9, e866.

Schüle, B., Oviedo, A., Johnston, K, Pai, S. and Francke, U. (2005). Inactivating Mutations in *ESCO2* Cause SC Phocomelia and Roberts Syndrome: No Phenotype-Genotype Correlation. *American Journal of Human Genetics* 77:1117-1128.

Skibbens, R.V., Colquhoun, J.M., Green, M.J., Molnar, C.A., Sin, D.N., Sullivan, B.J. and Tanzosh, E.E. (2013). Cohesinopathies of a Feather Flock Together. *PLOS Genetics* 9(12):1-6, e1004036.

Skibbens, R. V., L. B. Corson, D. Koshland and P. Hieter. (1999). Ctf7p is essential for sister chromatid cohesion and links mitotic chromosome structure to the DNA replication machinery. *Genes & Development* 13(3):307-319.

Tomkins, D.J. and Sisken, J.E. (1984). Abnormalities in the cell-division cycle in Roberts syndrome fibroblasts: a cellular basis for the phenotypic characteristics? *Am. J. Human Genetics* 36:1332-1340.

Tonkin, E.T., Wang, T.J., Lisgo, S., Bamshad, M.J. and 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. *Nature Genetics* 36(6):636-641.

Tóth, A., Ciosk, R., Uhlmann, F., Galova, M., Scheiffer, A. and Nasmyth, K. (1999). Yeast Cohesin complex requires a conserved protein, Eco1p(Ctf7), to establish cohesion between sister chromatids during DNA replication. *Genes and Development* 13:320-333.

Van Den Berg, D.J. and Francke, U. (1993). Sensitivity of Roberts syndrome cells to gamma radiation, mitomycin C, and protein synthesis inhibitors. *Somatic Cell and Molecular Genetics* 19:377-392.

Vega, H., Trainer, A.H., Gordillo, M., Crosier, M., Kayserili, H. *et al*. (2010). Phenotypic variability in 49 cases of *ESCO2* mutations, including novel missense and codon deletion in the acetyltransferase domain, correlates with *ESCO2* expression and establishes the clinical criteria for Roberts syndrome. *Journal Medical Genetics* 47:30-37.

Vega, H., Waisfisz, Q., Gordillo, M., Sakai, N., Yanagihara, I. *et al*. (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. *Nature Genetics* 37(5):468-470.

Vrouwe, M.G., Elghalbzouri-Maghrani, E., Meijers, M., Schouten, P., Godthelp, B.C., Bhuiyan, Z.A., Redeker, E.J. *et al*. (2007). Increased DNA damage sensitivity of Cornelia de Lange syndrome cells: evidence for impaired recombinational repair. *Human Molecular Genetics* 16(12):1478-1487.

Yuan, B., Pehlivan, D., Karaca, E., Patel, N., Charng, W.L. *et al*. (2015). Global transcriptional disturbances underlie Cornelia de Lange syndrome and related phenotypes. *The Journal of Clinical Investigation* 125(2):636-651.

Zuin, J., Franke, V., van Ijcken, W.F.J., van der Sloot, A., Krantz, I.D., van der Reijden, M.I.J.A., Nakato, R., Lenhard, B. and Wendt, K.S. (2014). A Cohesin-Independent Role for NIPBL at Promoters Provides Insights in CdLS. *PLOS Genetics* 10(2):1-15, e1004153.