Disorders of Transcriptional Regulation: An Emerging Category of Multiple Malformation Syndromes

Kosuke Izumi

Division of Human Genetics, The Children's Hospital of Philadelphia, Philadelphia, Pa., USA

Abstract

Some genetic disorders caused by mutations in genes encoding components of the transcriptional machinery as well as proteins involved in epigenetic modification of the genome share many overlapping features, such as facial dysmorphisms, growth problems and developmental delay/intellectual disability. As a basis for some shared phenotypic characteristics in these syndromes, a similar transcriptome disturbance, characterized by global transcriptional dysregulation, is believed to play a major role. In this review article, a general overview of gene transcription is provided, and the current knowledge of the mechanisms underlying some disorders of transcriptional regulation, such as Rubinstein-Taybi, Coffin-Siris, Cornelia de Lange, and CHOPS syndromes, are discussed.

© 2016 S. Karger AG, Basel

Human multiple malformation syndromes can serve as useful models to deepen the understanding of embryogenesis, since such malformation syndromes arise as a consequence of abnormal embryonic development. Human embryogenesis is intricately regulated by the spatiotemporal expression of many genes, and dysregulation of gene transcription is a known cause of many developmental syndromes. Hence, novel insights into human embryogenesis can be obtained from multiple malformation syndrome research.

The recent introduction of next-generation sequencing (NGS) technology has accelerated the pace of multiple malformation syndrome studies. NGS technology provides several advantages including high-throughput sequencing capability and precise quantification of nucleotides. By taking advantage of technology, our understanding of multiple malformation syndromes has been deepened, and the significant contribution of transcriptional disturbance in the pathogenesis of these syndromes has been revealed. In this review article, multiple malformation syndromes characterized by primary disruption of transcriptional regulation and categorized as 'disorders of transcriptional regulation' are discussed.

An Overview of Gene Transcription

The human genome contains ~20,000 protein-coding genes [Ezkurdia et al., 2014]. According to the central dogma, genes are typically transcribed into mRNAs and then translated into proteins. The gene transcription
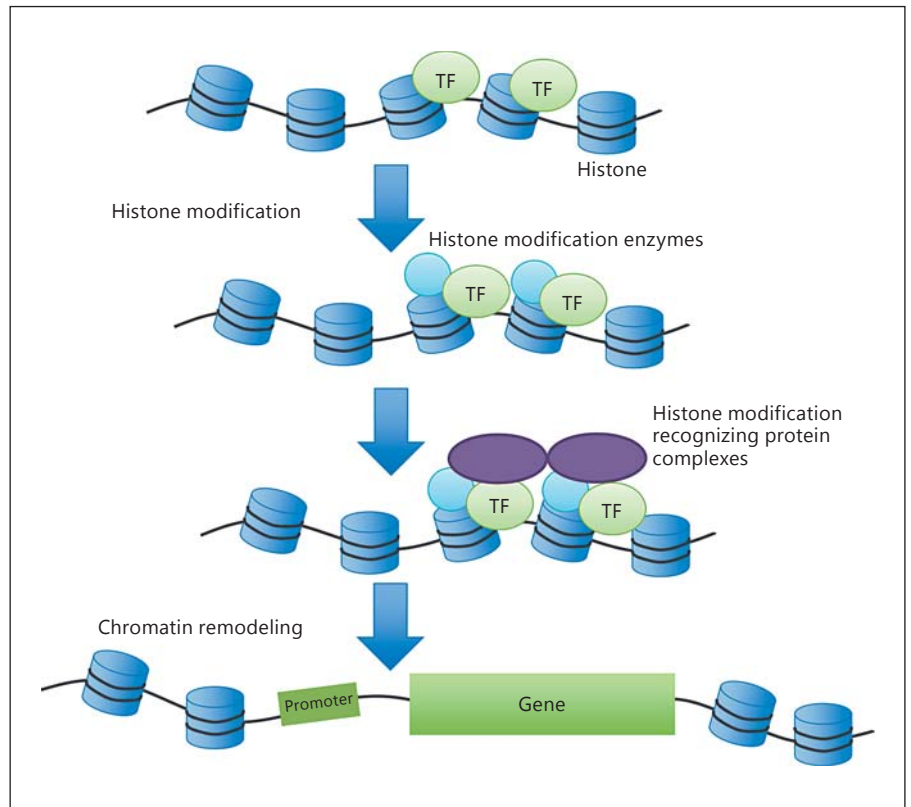
Kosuke Izumi, MD, PhD
Division of Human Genetics, The Children's Hospital of Philadelphia
3615 Civic Center Boulevard
Philadelphia, PA 19104 (USA)
E-Mail izumik1@email.chop.edu

Fig. 1. Schematic illustration of histone modification and chromatin remodeling. TF = DNA-binding transcription factor.

process is initiated by the binding of DNA-binding transcription factors to enhancer and promoter regions. Such transcription factors facilitate the binding of transcriptional co-activators such as histone modification enzymes and chromatin remodelers, and then the transcriptional machinery initiates gene transcription to produce mRNA.

The human genome is composed of over 3 billion bp. To pack such enormous stretches of genomic DNA into a small cellular nucleus, DNA is wrapped around histones. The histone core is an octamer composed of H2A, H2B, H3, and H4. Altogether, DNA and histones form a nucleosome. During gene transcription, since gene transcriptional machinery needs access to the genomic DNA, the nucleosome structure can become a physical barrier. Therefore, the unwrapping of the DNA from the histone proteins is required for active gene transcription. This unwrapping process is controlled by histone modification and chromatin remodeling (fig. 1).

Histone Modification and Chromatin Remodeling
The N-terminal domain of histones undergoes posttranslational modifications such as acetylation and meth-

ylation [Swygert and Peterson, 2014]. These modifications affect the binding of chromatin remodeler proteins. For example, acetyl-lysine modifications of histones can be recognized by a bromodomain, which is a vital region of many DNA-associated proteins including the SWI/SNF family of chromatin remodeling enzymes. Likewise, methyl-lysine modifications of histones can be recognized by plant homeodomain finger-containing and chromodomain-containing chromatin remodeler proteins [Swygert and Peterson, 2014]. By affecting the attraction of these chromatin remodelers, these so-called 'histone codes' play important roles in gene transcription regulation. Not only histone modifications, but also histone variants, DNA sequence and structure, RNA molecules, and transcription factors influence the genomic binding of chromatin remodelers [Längst and Manelyte, 2015]. Histone remodelers mobilize nucleosomes using ATP-dependent enzymatic reactions and unwrap the genomic DNA from the histones [Längst and Manelyte, 2015]. Such chromatin remodeling occurs both at regulatory elements, such as enhancers, and promoter regions. As a result, transcriptional machinery is enabled to initiate gene transcription.
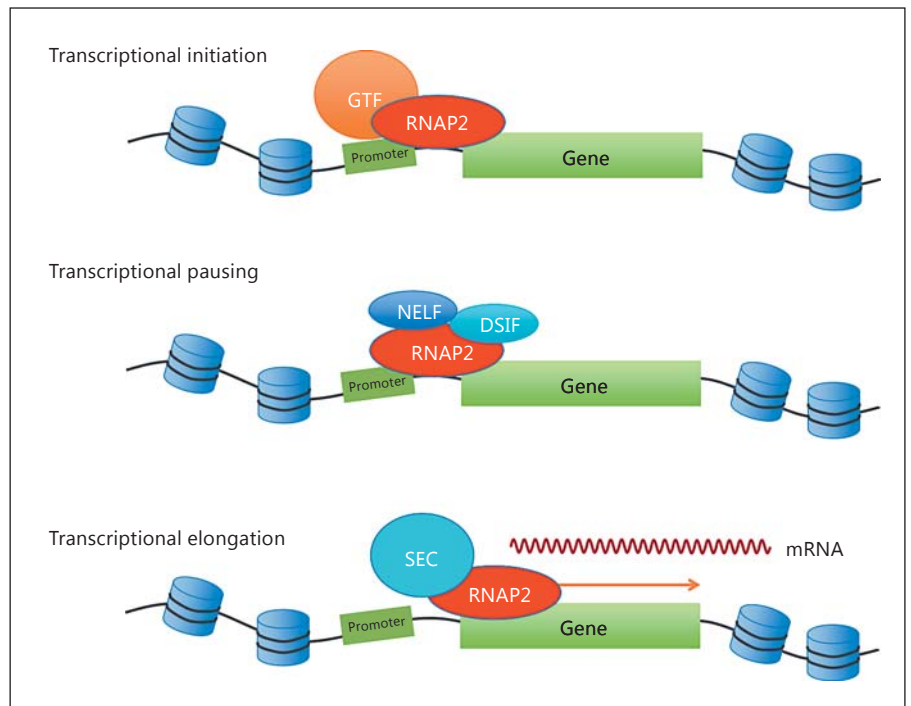
Fig. 2. Schematic illustration of gene transcription. GTF = General transcription factor complexes.

Transcriptional Initiation

Protein-coding genes are transcribed by RNA polymerase II (RNAP2) (fig. 2). For transcriptional initiation, RNAP2 assembles general transcription factor complexes including TFIIA, TFIIB, TFIID, TFIIE, TFIIF, and TFIIH at gene promoter regions and forms the pre-initiation complex [Cler et al., 2009; Sainsbury et al., 2015]. Each general transcription factor complex plays a different role during transcriptional initiation. For example, TFIID, composed of 14 TAF proteins (TAF1–TAF14) is important for gene promoter binding of general transcription factor complexes [Vermeulen et al., 2007]. The TFIIH complex contains CDK7, which phosphorylates the C-terminal domain of RNAP2 [Sainsbury et al., 2015]. A 7-amino acid repeat exists in this domain. Phosphorylation of this heptad repeat controls the transitions between various transcription stages from transcriptional initiation, pausing to elongation. CDK7 of TFIIH phosphorylates the serine 5 residue of the heptad repeat during transcriptional initiation. Furthermore, the Mediator complex is also required for the pre-initiation complex assembly [Allen and Taatjes, 2015]. This Mediator complex is comprised of multiple subunits and has various roles throughout gene transcription, including the organization of chromatin architecture, transcriptional initiation, pausing, and elongation. The Mediator complex can interact with different transcriptional regulators, such as DNA-binding transcription factors and RNAP2, and regulate multiple aspects of gene transcription [Allen and Taatjes, 2015].

Transcriptional Pausing and Elongation

After transcriptional initiation, gene transcription is transiently paused by the recruitment of pausing factors such as the negative elongation factor and the DRB sensitivity-inducing factor [Jonkers and Lis, 2015]. To release the pause and transition into active transcriptional elongation, the positive transcription elongation factor b (P-TEFb), comprising cyclin T1 and CDK9, is required [Jonkers and Lis, 2015]. P-TEFb is a component of the super elongation complex (SEC). CDK9 phosphorylates the negative elongation factor and the DRB sensitivity-inducing factor as well as the serine 2 residue of the heptad repeat in RNAP2. Consequently, RNAP2 is released from the promoter-proximal pausing site, and the transcriptional elongation continues [Jonkers and Lis, 2015]. The Mediator complex interacts with SEC and regulates transcriptional elongation as well.

Enhancer-Promoter Looping

The pre-initiation complex and transcription elongation factors are present not only at gene-body regions, but

Fig. 3. a Patient with RSTS. Note downslanted palpebral fissures, broad thumb and hallux. **b** Causative genes for RSTS encode the CREB-binding protein and p300 responsible for histone acetylation.
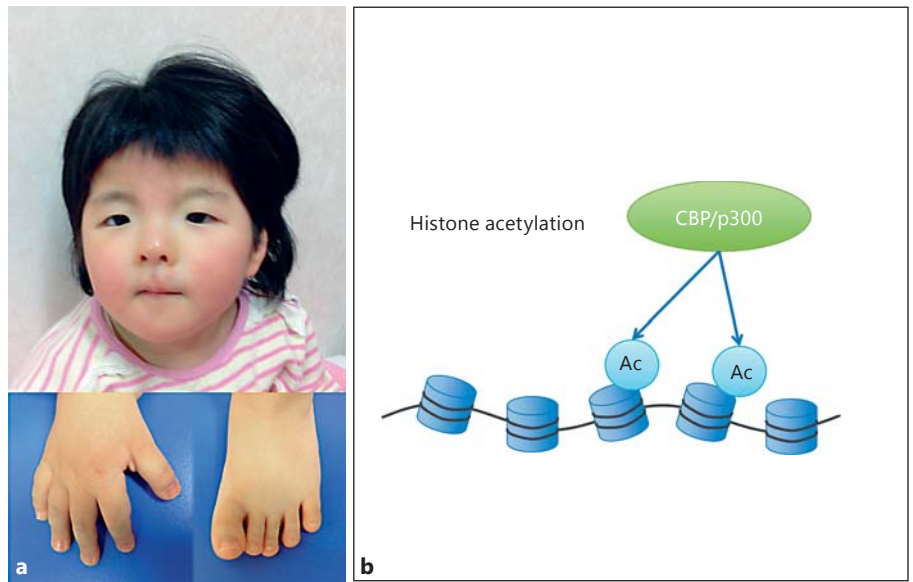
also at gene-enhancer regions [Heinz et al., 2015]. To produce transcriptional activation, the looping of the genomic DNA is required to bring enhancer elements closer to promoter regions [Heinz et al., 2015]. As a result, the amount of transcriptional machinery at gene promoter regions increases, and successful gene transcription occurs. This process is facilitated by the cohesin and Mediator complexes. The cohesin complex forms a ring structure and embraces the DNA looping to maintain enhancer-promoter interactions [Ball et al., 2014]. The roles of the Mediator complex include the organization of the genomic DNA into topological domains and mediation of DNA looping as well as in transcriptional initiation and elongation [Kagey et al., 2010; Allen and Taatjes, 2015]. Cooperative interactions among cohesin and Mediator complexes appear to play a major role in maintaining enhancer-promoter looping.

An Overview of Disorders of Transcriptional Regulation

In this review, disorders characterized by global transcriptional dysregulation are discussed. In this group of disorders, multiple malformation syndromes caused by mutations in genes encoding transcriptional machineries, chromatin remodeler, and histone modification enzymes should be included. Genetic conditions due to mutations in genes encoding DNA-binding transcription factors, such as the GATA family of transcription factors, are not included in the disorders of transcriptional regulation because the associated clinical symptoms appear to be somewhat different from those of the above-mentioned category. Interestingly, genetic disorders caused by mutations in genes encoding components of the transcriptional machinery as well as in proteins involved in epigenetic modification of the genome share many overlapping features, e.g., the Rubinstein-Taybi, Coffin-Siris, Cornelia de Lange, and CHOPS syndromes. As a basis for some shared phenotypic characteristics in these syndromes, a similar transcriptome disturbance is believed to play a major role. Therefore, the current knowledge of the mechanisms underlying these disorders will be discussed herein.

Rubinstein-Taybi Syndrome: A Disorder of Histone Modifications

Rubinstein-Taybi syndrome (RSTS; MIM 180849 and 613684) is characterized by facial dysmorphisms, including downslanting palpebral fissures, broad thumbs and halluces, developmental delay/intellectual disability, and short stature (fig. 3a) [Hennekam, 2006]. RSTS is caused by heterozygous mutations in the *CBP* and *EP300* genes [Roelfsema et al., 2005]. In about 55% of the RSTS cases, mutations are found in *CBP*, whereas mutations in *EP300* have been identified only in up to 8% of the cases [Roelfsema et al., 2005; Hennekam, 2006; Negri et al., 2015]. Mutations in *CBP* cause the most typical RSTS phenotype, and although *EP300* mutations can be found in typical RSTS cases, a wider phenotypic spectrum is seen in

individuals with *EP300* mutations [Roelfsema et al., 2005; Bartholdi et al., 2007; Negri et al., 2015]. Given the genomic deletions of *CBP* or *EP300* causing the typical phenotype of RSTS, haploinsufficiency of either *CBP* or *EP300* is suggested as the disease mechanism of RSTS [Roelfsema et al., 2005; Bartholdi et al., 2007; Tsai et al., 2011; Negri et al., 2015, 2016; Rusconi et al., 2015]. These genes encode transcriptional coactivators with lysine acetyltransferase activity (fig. 3b) [Kasper et al., 2014].

Pathogenesis of Rubinstein-Taybi Syndrome

CBP encodes the CREB-binding protein, and *EP300* encodes the p300 protein. These proteins have high sequence similarity, including a bromodomain and a histone acetyltransferase domain. The acetylation of histone lysines is the hallmark of a transcriptionally active chromatin region. CREB-binding protein and p300 are the primary acetyltransferases to mark H3K18 and H3K27 acetylation [Kasper et al., 2014]. H3K27 acetylation is known as an active enhancer and promoter histone mark. Consistent with the concept that RSTS is caused by defective histone acetylation, lymphoblastoid cell lines obtained from individuals with RSTS showed a reduction of histone H2B and H2A acetylation [Lopez-Atalaya et al., 2012]. Therefore, the molecular mechanism of RSTS appears to be a global transcriptional disturbance due to an altered genome-wide histone modification. Recently, in an individual with features of Cornelia de Lange syndrome (CdLS), an *EP300* mutation was identified, suggesting a similar transcriptome disturbance between RSTS and CdLS [Woods et al., 2014].

Coffin-Siris Syndrome: A Disorder of Chromatin Remodeling

Coffin-Siris syndrome (CSS; MIM 135900) is characterized by facial dysmorphisms, such as thick eyebrows, periorbital fullness, a wide mouth with full lips, and coarse facial features as well as fifth fingernail hypoplasia, absence of terminal phalanges, hypertrichosis, developmental delay/intellectual disability, and short stature [Vergano and Deardorff, 2014]. CSS is caused by heterozygous germline mutations in genes encoding a component of BAF (BRG1/hBRM-associated factors) complex, which belongs to the SWI/SNF chromatin remodeler complex family [Santen et al., 2012; Tsurusaki et al., 2012]. In individuals with CSS, mutations have been found in *ARID1A* (encoding BAF250a), *ARID1B* (encoding BAF250b), *SMARCA4* (encoding BRG1), *SMARCB1* (encoding BAF47), and *SMARCE1* (encoding BAF57). Mutations in *SMARCA2* (encoding BRM) have been re-

ported in Nicolaides-Baraitser syndrome (NCBRS), which has phenotypic overlap with CSS (see below) [Santen et al., 2012; Tsurusaki et al., 2012; Van Houdt et al., 2012]. All of the encoded proteins are components of the BAF complex (fig. 4a). BRM and BRG1 possess an ATPase domain, and only one of these proteins exists per BAF complex. Likewise, BAF250a and BAF250b are mutually exclusive in the BAF complex [Kadoch and Crabtree, 2015].

Mutations in one of these genes can be found in about 60–70% of the patients with CSS [Santen et al., 2013; Wieczorek et al., 2013]. *ARID1B* mutations are found most frequently and are associated with relatively milder phenotypes compared to CSS due to mutations in other genes [Santen et al., 2013; Wieczorek et al., 2013; Tsurusaki et al., 2014]. Since loss-of-function mutations in *ARID1B* and *ARID1A* were found in individuals with CSS, the disease mechanism appears to be haploinsufficiency of *ARID1A* and *ARID1B* [Santen et al., 2013; Wieczorek et al., 2013; Tsurusaki et al., 2014]. On the other hand, mutations in *SMARCA4*, *SMARCB1,* and *SMARCE1* were mainly missense or in-frame deletions [Santen et al., 2013; Tsurusaki et al., 2014].

NCBRS (MIM 601358) is also caused by mutations in the BAF complex protein-encoding *SMARCA2* gene [Van Houdt et al., 2012]. The clinical phenotype of NCBRS highly overlaps with that of CSS and is characterized by facial dysmorphisms, sparse hair, short stature, microcephaly, and intellectual disability [Van Houdt et al., 2012]. As a distinguishing feature, NCBRS patients have interphalangeal joint swelling, and they do not show hypoplasia of digits, which is typically associated with CSS. The main *SMARCA2* mutations identified in NCBRS were missense; therefore, the molecular mechanism resulting from *SMARCA2* mutation may not be a simple loss of function [Van Houdt et al., 2012]. In addition to *SMARCA2*, *ARID1B* mutations were found to be associated with NCBRS clinical features [Wieczorek et al., 2013]. Given the clinical similarities and molecular overlap between CSS and NCBRS, a similar molecular pathogenesis of CSS and NCBRS might be considered.

Pathogenesis of Coffin-Siris Syndrome and Nicolaides-Baraitser Syndrome

Since SWI/SNF complex remodelers alter chromatin conformation by their ATPase enzymatic activity, mutations in BAF complex protein-coding genes are expected to cause transcriptional disturbance due to modifications in the chromatin accessibility. A previous study, using chromatin immunoprecipitation followed by massive
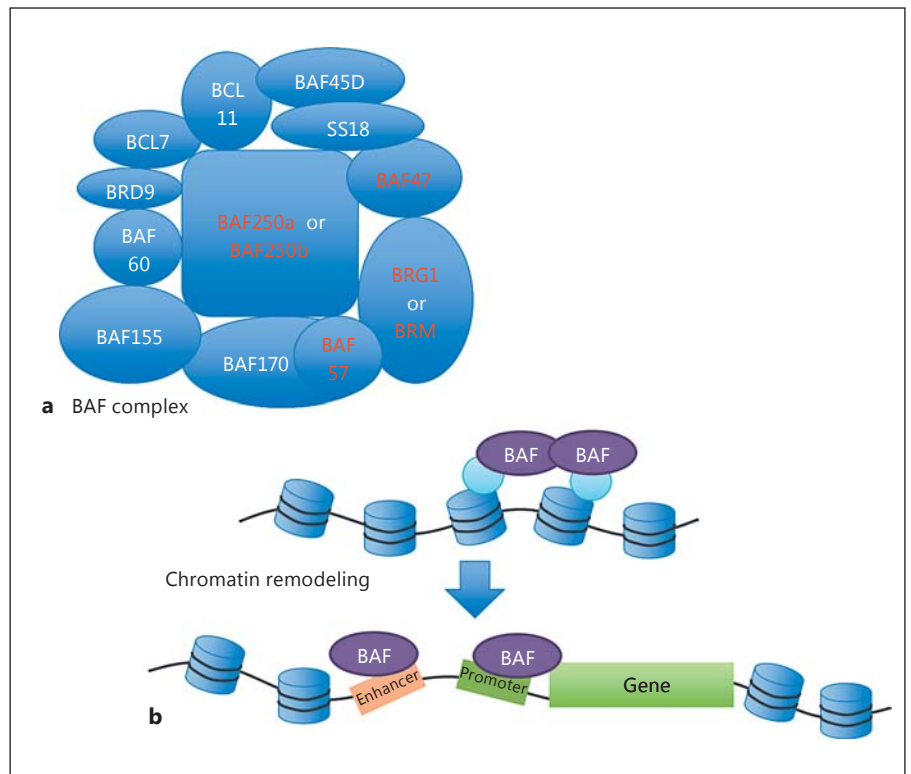
Fig. 4. a Schematic illustration of the BAF complex, mutated in CSS. The BAF complex components whose mutations were found in CSS and NCBRS are indicated by red characters. **b** The BAF complex is responsible for chromatin remodeling.

parallel sequencing, demonstrated a combinatorial regulation of gene expression achieved by distinct BAF complexes with different components [Raab et al., 2015]. Although *ARID1A*-encoded BAF250a and *ARID1B*-encoded BAF250b proteins do not coexist in a single complex, they bind to similar genomic locations. BAF250a and BAF250b both bind close to transcriptional start sites of genes, as well as gene enhancer regions, suggesting the importance of the BAF complex in remodeling the promoter and enhancer regions of the chromatin structure (fig. 4b) [Raab et al., 2015]. However, the mechanism by which the BAF complex chromatin remodeler controls gene transcription remains to be fully understood. Given the role of the BAF complex in transcriptional regulation, global transcriptional disturbance probably underlies the molecular mechanism of CSS and NCBRS.

Cornelia de Lange Syndrome: A Disorder of the Cohesin Complex

CdLS (MIM 122470) is characterized by craniofacial dysmorphisms, such as synophrys, high-arched eyebrows and a thin downturned upper lip, hirsutism, intellectual disability, microcephaly, growth retardation, limb anomalies, such as micromelia, phocomelia and oligodactyly as well as several other systemic abnormalities (fig. 5a) [Liu

and Krantz, 2009]. CdLS is caused by germline mutations in genes encoding the structural and regulatory components of the cohesin complex. Mutations have been found in *NIPBL, HDAC8, SMC1A,* and *SMC3* in CdLS [Krantz et al., 2004; Tonkin et al., 2004; Musio et al., 2006; Deardorff et al., 2007, 2012a]. A CdLS-like phenotype was also observed in individuals with *RAD21* mutations [Deardorff et al., 2012b]. The cohesin complex is composed of *SMC1A, SMC3, RAD21,* and *STAG1* or *STAG2* (fig. 5b). The cohesin complex has a ring-like structure and 'embraces' the genomic DNA. *NIPBL* encodes a protein that controls the cohesin complex genome loading. The SMC3 component of the cohesin complex becomes acetylated once loaded onto the genome and is subsequently deacetylated by HDAC8 for cohesin protein recycling.

NIPBL mutations can be found in ~50% of the individuals with CdLS, and in general, more severely (or 'classically') affected individuals are much more likely to harbor a mutation in *NIPBL*, with loss-of-function mutations causing a severe phenotype [Mannini et al., 2013]. Mutations in *SMC1A, SMC3,* and *HDAC8* can be found in a small percentage of the individuals with CdLS. In individuals with *NIPBL* mutations, the disease mechanism is due to haploinsufficiency of *NIPBL* [Pehlivan et al., 2012; Mannini et al., 2013]. Although both *HDAC8* and
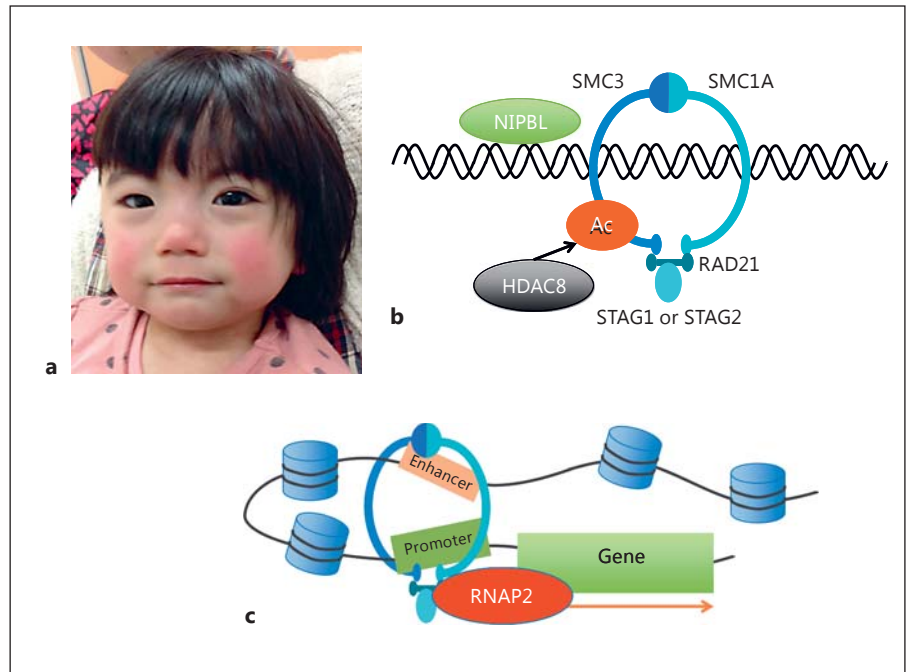
Fig. 5. a Patient with CdLS. Note arched eyebrow and long philtrum. **b** Schematic illustration of the cohesin complex. **c** The Cohesin complex promotes enhancer-promoter interaction and controls gene transcription via interacting with RNAP2. Ac = Acetylation site of the SMC3 protein.

SMC1A are located in the X chromosome, affected females with heterozygous *HDAC8* or *SMC1A* mutations have been reported [Deardorff et al., 2007; Kaiser et al., 2014]. Most mutations in *HDAC8* are missense and de novo. In general, males with *HDAC8* mutations tend to show more severe manifestations compared to females. Although heterozygous females with *HDAC8* mutations can be asymptomatic, affected heterozygous females were also reported, depending on the skewing level of X inactivation [Kaiser et al., 2014]. *HDAC8* mutations produce loss of function in the acetyltransferase activity [Deardorff et al., 2012a; Kaiser et al., 2014]. *SMC1A* escapes X inactivation and shows a biallelic expression pattern. Interestingly, female CdLS probands with an *SMC1A* mutation tend to show a more severe phenotype compared to males with *SMC1A* mutations [Deardorff et al., 2007]. *SMC1A* and *SMC3* mutations found in individuals with CdLS were missense mutations or in-frame deletions [Deardorff et al., 2007; Gil-Rodríguez et al., 2015]. Recently, female probands with heterozygous loss-of-function mutations in *SMC1A* were found in association with seizures and intellectual disability without a clear CdLS phenotype [Goldstein et al., 2015; Lebrun et al., 2015; Jansen et al., 2016]. Therefore, the disease mechanism of *SMC1A* and *SMC3* mutations in CdLS is unlikely to be a mere loss-of-function effect, and a dominant-negative effect is suspected.

Pathogenesis of Cornelia de Lange Syndrome

The cohesin complex was originally identified with a critical role in preventing premature sister chromatid separation during mitosis [Ball et al., 2014]. Therefore, when *NIPBL* was identified as a causative gene for CdLS, it was originally thought that sister chromatid separation defects might underlie the disease mechanism of CdLS. However, cell lines obtained from patients with CdLS did not show an increased frequency of premature sister chromatid separation, suggesting that disruption of cohesin's canonical role was unlikely to be the underlying mechanism causing the CdLS phenotype [Kaur et al., 2005; Castronovo et al., 2009]. Instead, animal model studies, including drosophila, zebrafish, and mice as well as patient-derived cell line studies demonstrated a clear role for cohesin in transcriptional regulation [Kawauchi et al., 2009; Liu et al., 2009; Muto et al., 2014; Wu et al., 2015].

There are 2 proposed mechanisms regarding the transcriptional regulatory mechanism of cohesin (fig. 5c) [Ball et al., 2014; Izumi et al., 2015]. The first suggests that the cohesin complex facilitates gene promoter-enhancer interactions, whereas the second indicates a direct role in transcriptional regulation by interacting with RNAP2. By using its ring structure, cohesin is known to promote DNA looping, allowing physical interactions between distal regulatory elements and promoter regions [Ball et
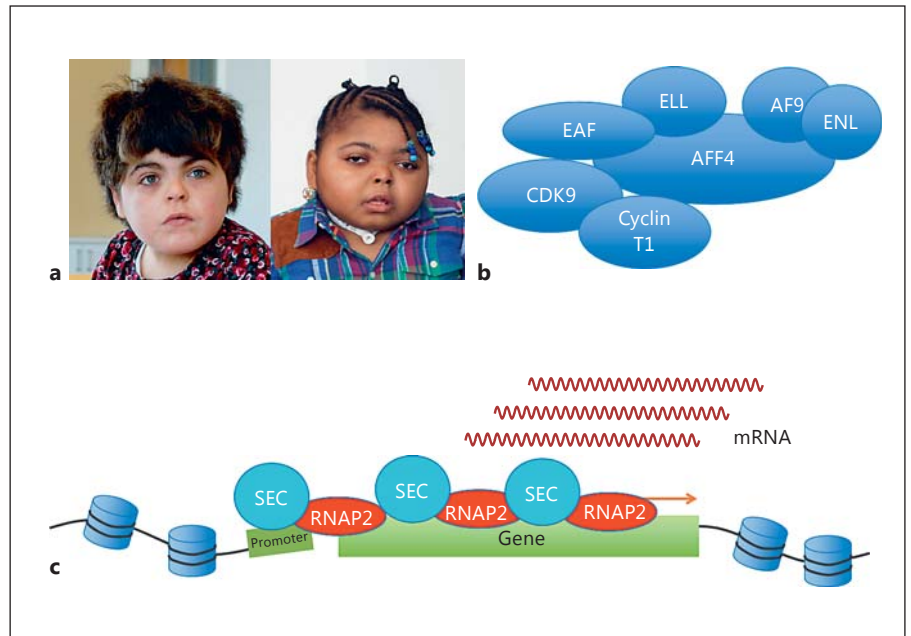
Fig. 6. a Patients with CHOPS syndrome. Note round faces and arched eyebrows. **b** Schematic illustration of SEC. **c** Disease mechanism of CHOPS syndrome. SEC is responsible for transcriptional elongation. In the CHOPS syndrome, accumulation of SEC causes transcriptional overactivation by excessive transcriptional elongation.

al., 2014]. A direct transcriptional regulatory role of cohesin is supported by the observation that RNAP2, as well as the SEC, interacts with cohesin, specifically STAG1-cohesin [Izumi et al., 2015]. Furthermore, cohesin-binding sites were enriched within the promoter regions of CdLS dysregulated genes, suggesting its role in transcriptional initiation or elongation control [Liu et al., 2009].

NIPBL forms a heterodimer complex with MAU2, and functions as a cohesin loader. In yeast, RSC chromatin remodeler complex, yeast ortholog of BAF complex, was shown to play a major role in cohesin loaders (SCC2/SCC4) genome binding, and cohesin loaders and RSC remodeler cooperatively maintain nucleosome-free regions [Lopez-Serra et al., 2014]. Furthermore, transcriptional similarities were established between the cohesin component inactivation and the RSC complex inactivation in the yeast model [Lopez-Serra et al., 2014]. Therefore, the derangements of nucleosome positioning might cause transcriptional disturbances not only in CSS but also in CdLS.

Recently, mutations in *TAF1* and *TAF6* were shown to be associated with a CdLS-like phenotype [O'Rawe et al., 2015; Yuan et al., 2015]. *TAF1* and *TAF6* encode components of the TFIID complex, which plays a major role in transcriptional initiation. Therefore, disruption of transcriptional initiation or transition from transcriptional initiation to the elongation phase may be the steps affected in CdLS.

CHOPS Syndrome: A Disorder of Transcriptional Elongation

CHOPS syndrome (MIM 616368) is a multisystem developmental disorder characterized by coarse facial features, cognitive impairment, heart defects, obesity, pulmonary symptoms, skeletal defects, and short stature (fig. 6a) [Izumi et al., 2015]. CHOPS syndrome probands were originally suspected to have CdLS, although they lacked several typical features of CdLS, such as microcephaly. CHOPS syndrome is caused by gain-of-function mutations in *AFF4*, which encodes a key component of the SEC (fig. 6b) [Izumi et al., 2015].

Pathogenesis of CHOPS Syndrome
SEC includes P-TEFb, which plays important roles in RNAP2 pausing release. Other components of the SEC are AFF1, AFF4, AF9, ENL, and ELL protein families. AFF4 is a scaffolding protein located at the center of the SEC, which does not exist if AFF4 is knocked down [Lin et al., 2010]. All the mutations found in CHOPS syndrome probands were missense and were located within the ALF homology domain of the AFF4 protein [Izumi et al., 2015]. The AFF4 protein is degraded by ubiquitination-dependent proteasomal degradation. For AFF4 to be transported to the proteasome, AFF4 is ubiquitinated by E3 ubiquitin ligase, SIAH1. SIAH1 binds to AFF4 by recognizing the SIAH1-binding motif located within the

Table 1. Examples of disorders of transcriptional regulation

Syndrome	Causative genes	Function of encoded proteins	Clinical features
Rubinstein-Taybi syndrome	*CBP*, *EP300*	histone modification enzymes	facial dysmorphisms (downslanted palpebral fissures, prominent beaked nose), broad thumbs and halluces, short stature, and developmental delay
Wiedemann-Steiner syndrome	*KMT2A*	histone modification enzymes	facial dysmorphisms (hypertelorism, downslanted palpebral fissures), hypertrichosis, and developmental delay
Kabuki syndrome	*KMT2D*, *KDM6A*	histone modification enzymes	facial dysmorphisms (long palpebral fissures, eversion of lateral third of lower eyelids), skeletal anomalies, persistence of fetal fingertip pads, developmental delay, and growth deficiency
Say-Barber-Biesecker/ Young-Simpson syndrome	*KAT6B*	histone modification enzymes	facial dysmorphisms (blepharophimosis), developmental delay, hypotonia, joint stiffness, cryptorchidism, hypothyroidism, and patellar hypoplasia/agenesis
KBG syndrome	*ANKRD11*	regulate histone acetylation	facial dysmorphisms (anteverted nares), macrodontia of the upper central incisors, short stature, skeletal anomalies, and developmental delay
Coffin-Siris syndrome	*ARID1A*, *ARID1B*, *SMARCA4*, *SMARCB1*, *SMARCE1*	chromatin remodelers	facial dysmorphisms (bushy eyebrows, large mouth), aplasia/hypoplasia of the distal phalanx or nail, developmental delay, hypotonia, hirutism, and sparse scalp hair
Nicolaides-Baraitser syndrome	*SMARCA2*	chromatin remodeler	facial dysmorphisms/coarse facial features, sparse scalp hair, interphalangeal joint swelling, microcephaly, seizures, and developmental delay
CHARGE syndrome	*CHD7*	chromatin remodeler	facial dysmorphisms (facial asymmetry, cup-shaped ears), coloboma, heart defects, choanal atresia, retarded growth and development, genital abnormalities, and ear anomalies
FG syndrome	*MED12*	mediator protein	facial dysmorphisms (prominent forehead), developmental delay, macrocephaly, hypotonia, corpus callosum abnormalities, and broad thumbs/ halluces
Cornelia de Lange syndrome	*NIPBL*, *SMC1A*, *SMC3*, *HDAC8*	cohesin complex and regulatory factors	facial dysmorphisms (synophrys, long philtrum), developmental delay, growth retardation, hirsutism, and upper limb abnormalities
CHOPS syndrome	*AFF4*	transcriptional elongation factor	facial dysmorphisms (round face), developmental delay, heart defect, obesity, pulmonary abnormalities, short stature, and skeletal anomalies

ALF domain [Oliver et al., 2004]. In CHOPS syndrome, missense mutations hamper the binding of SIAH1, causing a less efficient proteasomal degradation of AFF4. As a result, the AFF4 protein accumulates in the cells of CHOPS syndrome probands [Izumi et al., 2015].

The effect of the accumulation of AFF4 proteins is transcriptional overactivation of genes, whose expression is controlled by transcriptional elongation, such as *MYC* (fig. 6c). Chromatin immunoprecipitation followed by massive parallel sequencing demonstrated accumulation

of the AFF4 protein around transcription start sites in genes in which expression is activated in CHOPS syndrome [Izumi et al., 2015]. Furthermore, the elongating form of RNAP2 (RNAP2 Ser2ph) was accumulated around gene bodies of overexpressed genes in CHOPS syndrome. Combining these observations, the transcriptional disturbance of CHOPS syndrome results from an overactivation of the transcriptional elongation system.

Transcriptome comparison between CHOPS syndrome and CdLS showed similar transcriptome profiles [Izumi et al., 2015]. Thus, disruption of molecular interactions of cohesin, SEC, and RNAP2 would lead to similar transcriptional disturbances by altering transcriptional elongation. Since CHOPS syndrome probands have clinical features that overlap those seen in CdLS, a part of the CdLS disease mechanism could be explained by excessive transcriptional elongation, similar to that seen in CHOPS syndrome.

Other Disorders of Transcriptional Regulation

In recent years, an increasing number of genetic disorders, due to mutations in genes encoding histone modifiers and chromatin remodeler proteins, have been identified (table 1). These disorders include: CHARGE syndrome due to *CHD7* mutations, Wiedemann-Steiner syndrome due to *KMT2A* mutations, Kabuki syndrome due to *KMT2D* and *KDM6A* mutations, and Say-Barber-Biesecker/Young-Simpson syndrome due to *KAT6B* mutations [Vissers et al., 2004; Ng et al., 2010; Clayton-Smith et al., 2011; Jones et al., 2012; Lederer et al., 2012]. CHD7 is a chromatin remodeler protein with a chromodomain histone-modification recognition motif. *KMT2A*, *KMT2D*, *KDM6A*, and *KAT6B* encode histone modification enzymes. KBG syndrome may also be categorized in this group of disorders because the causative gene for KBG syndrome, *ANKRD11*, encodes Ankyrin repeat domain 11 protein, known to regulate histone acetylation via functioning with HDAC3 [Sirmaci et al., 2011; Gallagher et al., 2015]. Mutations in genes encoding the mediator complex proteins have also been reported. *MED23* and *MED25* mutations were found in individuals with intellectual disability, and *MED12* mutations have been found in individuals with FG syndrome and X-linked Ohdo syndrome [Risheg et al., 2007; Hashimoto et al., 2011; Vulto-van Silfhout et al., 2013; Basel-Vanagaite et al., 2015; Figueiredo et al., 2015]. There are clinical features that allow astute clinicians to distinguish between typical cases of these syndromes. However, in their atyp-ical presentations, a clinical distinction may not be feasible. The use of exome sequencing in clinical practice has allowed for an expansion of the phenotypic spectrum of these Mendelian disorders. We recently reported a case of an RSTS patient who had overlapping features with CHARGE syndrome [Masuda et al., 2015]. Likewise, mutations in the *ANKRD11* gene, whose mutations cause KBG syndrome, are identified in CdLS-overlapping phenotype patients [Parenti et al., 2016]. As illustrated by these cases with overlapping phenotypes among various disorders of transcriptional regulation, unbiased genome screening, by either exome or genome sequencing, may obscure the clinical distinction among disorders of transcriptional regulation.

Conclusion

Similarities between the clinical phenotypes of diverse syndromic diagnoses facilitate the identification of common molecular mechanisms linking these diverse diagnoses on a pathogenetic level. Such insights have contributed significantly not only to deepening our understanding of the etiologies of these malformation syndromes, but also in shedding light on the mechanism of fundamental cellular processes such as gene transcription. Due to the increasing use of NGS-based technologies in the clinical and research fields, novel gene discovery will continue to provide insights into the molecular controls of transcription and will further our understanding of a growing class of diagnoses – the disorders of transcriptional regulation. Elucidation of the mechanisms of transcriptional abnormalities underlying disorders of transcriptional regulation will contribute to the advancement of both clinical genetics and molecular cell biology.

Disclosure Statement

The author has no conflict of interest to declare.

References

Allen BL, Taatjes DJ: The Mediator complex: a central integrator of transcription. Nat Rev Mol Cell Biol 16:155–166 (2015).

Ball AR Jr, Chen YY, Yokomori K: Mechanisms of cohesin-mediated gene regulation and lessons learned from cohesinopathies. Biochim Biophys Acta 1839:191–202 (2014).

Bartholdi D, Roelfsema JH, Papadia F, Breuning MH, Niedrist D, et al: Genetic heterogeneity in Rubinstein-Taybi syndrome: delineation of the phenotype of the first patients carrying mutations in *EP300*. J Med Genet 44:327–333 (2007).

Basel-Vanagaite L, Smirin-Yosef P, Essakow JL, Tzur S, Lagovsky I, et al: Homozygous *MED25* mutation implicated in eye-intellectual disability syndrome. Hum Genet 134:577–587 (2015).

Castronovo P, Gervasini C, Cereda A, Masciadri M, Milani D, et al: Premature chromatid separation is not a useful diagnostic marker for Cornelia de Lange syndrome. Chromosome Res 17:763–771 (2009).

Clayton-Smith J, O'Sullivan J, Daly S, Bhaskar S, Day R, et al: Whole-exome-sequencing identifies mutations in histone acetyltransferase gene *KAT6B* in individuals with the Say-Barber-Biesecker variant of Ohdo syndrome. Am J Hum Genet 89:675–681 (2011).

Cler E, Papai G, Schultz P, Davidson I: Recent advances in understanding the structure and function of general transcription factor TFIID. Cell Mol Life Sci 66:2123–2134 (2009).

Deardorff MA, Kaur M, Yaeger D, Rampuria A, Korolev S, et al: Mutations in cohesin complex members SMC3 and SMC1A cause a mild variant of Cornelia de Lange syndrome with predominant mental retardation. Am J Hum Genet 80:485–494 (2007).

Deardorff MA, Bando M, Nakato R, Watrin E, Itoh T, et al: *HDAC8* mutations in Cornelia de Lange syndrome affect the cohesin acetylation cycle. Nature 489:313–317 (2012a).

Deardorff MA, Wilde JJ, Albrecht M, Dickinson E, Tennstedt S, et al: *RAD21* mutations cause a human cohesinopathy. Am J Hum Genet 90:1014–1027 (2012b).

Ezkurdia I, Juan D, Rodriguez JM, Frankish A, Diekhans M, et al: Multiple evidence strands suggest that there may be as few as 19,000 human protein-coding genes. Hum Mol Genet 23:5866–5878 (2014).

Figueiredo T, Melo US, Pessoa ALS, Nobrega PR, Kitajima JP, et al: Homozygous missense mutation in *MED25* segregates with syndromic intellectual disability in a large consanguineous family. J Med Genet 52:123–127 (2015).

Gallagher D, Voronova A, Zander MA, Cancino GI, Bramall A, et al: Ankrd11 is a chromatin regulator involved in autism that is essential for neural development. Dev Cell 32:31–42 (2015).

Gil-Rodríguez MC, Deardorff MA, Ansari M, Tan CA, Parenti I, et al: De novo heterozygous mutations in *SMC3* cause a range of Cornelia

de Lange syndrome-overlapping phenotypes. Hum Mutat 36:454–462 (2015).

Goldstein JHR, Tim-Aroon T, Shieh J, Merrill M, Deeb KK, et al: Novel *SMC1A* frameshift mutations in children with developmental delay and epilepsy. Eur J Med Genet 58:562–568 (2015).

Hashimoto S, Boissel S, Zarhrate M, Rio M, Munnich A, et al: *MED23* mutation links intellectual disability to dysregulation of immediate early gene expression. Science 333:1161–1163 (2011).

Heinz S, Romanoski CE, Benner C, Glass CK: The selection and function of cell type-specific enhancers. Nat Rev Mol Cell Biol 16:144–154 (2015).

Hennekam RCM: Rubinstein-Taybi syndrome. Eur J Hum Genet 14:981–985 (2006).

Izumi K, Nakato R, Zhang Z, Edmondson AC, Noon S, et al: Germline gain-of-function mutations in *AFF4* cause a developmental syndrome functionally linking the super elongation complex and cohesin. Nat Genet 47:338–344 (2015).

Jansen S, Kleefstra T, Willemsen MH, de Vries P, Pfundt R, et al: De novo loss-of-function mutations in X-linked *SMC1A* cause severe ID and therapy resistant epilepsy in females: expanding the phenotypic spectrum. Clin Genet (2016), E-pub ahead of print.

Jones WD, Dafou D, McEntagart M, Woollard WJ, Elmslie FV, et al: De novo mutations in *MLL* cause Wiedemann-Steiner syndrome. Am J Hum Genet 91:358–364 (2012).

Jonkers I, Lis JT: Getting up to speed with transcription elongation by RNA polymerase II. Nat Rev Mol Cell Biol 16:167–177 (2015).

Kadoch C, Crabtree GR: Mammalian SWI/SNF chromatin remodeling complexes and cancer: mechanistic insights gained from human genomics. Sci Adv 1:e1500447 (2015).

Kagey MH, Newman JJ, Bilodeau S, Zhan Y, Orlando DA, et al: Mediator and cohesin connect gene expression and chromatin architecture. Nature 467:430–435 (2010).

Kaiser FJ, Ansari M, Braunholz D, Concepción Gil-Rodríguez M, Decroos C, et al: Loss-of-function *HDAC8* mutations cause a phenotypic spectrum of Cornelia de Lange syndrome-like features, ocular hypertelorism, large fontanelle and X-linked inheritance. Hum Mol Genet 23:2888–2900 (2014).

Kasper LH, Qu C, Obenauer JC, McGoldrick DJ, Brindle PK: Genome-wide and single-cell analyses reveal a context dependent relationship between CBP recruitment and gene expression. Nucleic Acids Res 42:11363–11382 (2014).

Kaur M, DeScipio C, McCallum J, Yaeger D, Devoto M, et al: Precocious sister chromatid separation (PSCS) in Cornelia de Lange syndrome. Am J Med Genet A 138:27–31 (2005).

Kawauchi S, Calof AL, Santos R, Lopez-Burks ME, Young CM, et al: Multiple organ system defects and transcriptional dysregulation in

the *Nipbl*$^{+/-}$ mouse, a model of Cornelia de Lange syndrome. PLoS Genet 5:e1000650 (2009).

Krantz ID, McCallum J, DeScipio C, Kaur M, Gillis LA, et al: Cornelia de Lange syndrome is caused by mutations in *NIPBL*, the human homolog of *Drosophila melanogaster Nipped-B*. Nat Genet 36:631–635 (2004).

Längst G, Manelyte L: Chromatin remodelers: from function to dysfunction. Genes (Basel) 6:299–324 (2015).

Lebrun N, Lebon S, Jeannet PY, Jacquemont S, Billuart P, Bienvenu T: Early-onset encephalopathy with epilepsy associated with a novel splice site mutation in *SMC1A*. Am J Med Genet A 167:3076–3081 (2015).

Lederer D, Grisart B, Digilio MC, Benoit V, Crespin M, et al: Deletion of KDM6A, a histone demethylase interacting with MLL2, in three patients with Kabuki syndrome. Am J Hum Genet 90:119–124 (2012).

Lin C, Smith ER, Takahashi H, Lai KC, Martin-Brown S, et al: AFF4, a component of the ELL/P-TEFb elongation complex and a shared subunit of MLL chimeras, can link transcription elongation to leukemia. Mol Cell 37:429–437 (2010).

Liu J, Krantz ID: Cornelia de Lange syndrome, cohesin, and beyond. Clin Genet 76:303–314 (2009).

Liu J, Zhang Z, Bando M, Itoh T, Deardorff MA, et al: Transcriptional dysregulation in *NIPBL* and cohesin mutant human cells. PLoS Biol 7:e1000119 (2009).

Lopez-Atalaya JP, Gervasini C, Mottadelli F, Spena S, Piccione M, et al: Histone acetylation deficits in lymphoblastoid cell lines from patients with Rubinstein-Taybi syndrome. J Med Genet 49:66–74 (2012).

Lopez-Serra L, Kelly G, Patel H, Stewart A, Uhlmann F: The Scc2-Scc4 complex acts in sister chromatid cohesion and transcriptional regulation by maintaining nucleosome-free regions. Nat Genet 46:1147–1151 (2014).

Mannini L, Cucco F, Quarantotti V, Krantz ID, Musio A: Mutation spectrum and genotype-phenotype correlation in Cornelia de Lange syndrome. Hum Mutat 34:1589–1596 (2013).

Masuda K, Akiyama K, Arakawa M, Nishi E, Kitazawa N, et al: Exome sequencing identification of *EP300* mutation in a proband with coloboma and imperforate anus: possible expansion of the phenotypic spectrum of Rubinstein-Taybi syndrome. Mol Syndromol 6:99–103 (2015).

Musio A, Selicorni A, Focarelli ML, Gervasini C, Milani D, et al: 2006. X-linked Cornelia de Lange syndrome owing to *SMC1L1* mutations. Nat Genet 38:528–530 (2006).

Muto A, Ikeda S, Lopez-Burks ME, Kikuchi Y, Calof AL, et al: Nipbl and mediator cooperatively regulate gene expression to control limb development. PLoS Genet 10:e1004671 (2014).

Negri G, Milani D, Colapietro P, Forzano F, Della Monica M, et al: Clinical and molecular characterization of Rubinstein-Taybi syndrome patients carrying distinct novel mutations of the *EP300* gene. Clin Genet 87:148–154 (2015).

Negri G, Magini P, Milani D, Colapietro P, Rusconi D, et al: From whole gene deletion to point mutations of *EP300*-positive Rubinstein-Taybi patients: new insights into the mutational spectrum and peculiar clinical hallmarks. Hum Mutat 37:175–183 (2016).

Ng SB, Bigham AW, Buckingham KJ, Hannibal MC, McMillin MJ, et al: Exome sequencing identifies *MLL2* mutations as a cause of Kabuki syndrome. Nat Genet 42:790–793 (2010).

Oliver PL, Bitoun E, Clark J, Jones EL, Davies KE: Mediation of Af4 protein function in the cerebellum by Siah proteins. Proc Natl Acad Sci USA 101:14901–14906 (2004).

O'Rawe JA, Wu Y, Dörfel MJ, Rope AF, Au PYB, et al: *TAF1* variants are associated with dysmorphic features, intellectual disability, and neurological manifestations. Am J Hum Genet 97:922–932 (2015).

Parenti I, Gervasini C, Pozojevic J, Graul-Neumann L, Azzollini J, et al: Broadening of cohesinopathies: exome sequencing identifies mutations in *ANKRD11* in two patients with Cornelia de Lange-overlapping phenotype. Clin Genet 89:74–81 (2016).

Pehlivan D, Hullings M, Carvalho CMB, Gonzaga-Jauregui CG, Loy E, et al: *NIPBL* rearrangements in Cornelia de Lange syndrome: evidence for replicative mechanism and genotype-phenotype correlation. Genet Med 14:313–322 (2012).

Raab JR, Resnick S, Magnuson T: Genome-wide transcriptional regulation mediated by biochemically distinct SWI/SNF complexes. PLoS Genet 11:e1005748 (2015).

Risheg H, Graham JM Jr, Clark RD, Rogers RC, Opitz JM, et al: A recurrent mutation in *MED12* leading to R961W causes Opitz-Kaveggia syndrome. Nat Genet 39:451–453 (2007).

Roelfsema JH, White SJ, Ariyürek Y, Bartholdi D, Niedrist D, et al: Genetic heterogeneity in Rubinstein-Taybi syndrome: mutations in both the *CBP* and *EP300* genes cause disease. Am J Hum Genet 76:572–580 (2005).

Rusconi D, Negri G, Colapietro P, Picinelli C, Milani D, et al: Characterization of 14 novel deletions underlying Rubinstein-Taybi syndrome: an update of the *CREBBP* deletion repertoire. Hum Genet 134:613–626 (2015).

Sainsbury S, Bernecky C, Cramer P: Structural basis of transcription initiation by RNA polymerase II. Nat Rev Mol Cell Biol 16:129–143 (2015).

Santen GWE, Aten E, Sun Y, Almomani R, Gilissen C, et al: Mutations in SWI/SNF chromatin remodeling complex gene *ARID1B* cause Coffin-Siris syndrome. Nat Genet 44:379–380 (2012).

Santen GWE, Aten E, Vulto-van Silfhout AT, Pottinger C, van Bon BWM, et al: Coffin-Siris syndrome and the BAF complex: genotype-phenotype study in 63 patients. Hum Mutat 34:1519–1528 (2013).

Sirmaci A, Spiliopoulos M, Brancati F, Powell E, Duman D, et al: Mutations in *ANKRD11* cause KBG syndrome, characterized by intellectual disability, skeletal malformations, and macrodontia. Am J Hum Genet 89:289–294 (2011).

Swygert SG, Peterson CL: Chromatin dynamics: interplay between remodeling enzymes and histone modifications. Biochim Biophys Acta 1839:728–736 (2014).

Tonkin ET, Wang TJ, Lisgo S, Bamshad MJ, Strachan T: *NIPBL*, encoding a homolog of fungal Scc2-type sister chromatid cohesion proteins and fly Nipped-B, is mutated in Cornelia de Lange syndrome. Nat Genet 36:636–641 (2004).

Tsai ACH, Dossett CJ, Walton CS, Cramer AE, Eng PA, et al: Exon deletions of the *EP300* and *CREBBP* genes in two children with Rubinstein-Taybi syndrome detected by aCGH. Eur J Hum Genet EJHG 19:43–49 (2011).

Tsurusaki Y, Okamoto N, Ohashi H, Kosho T, Imai Y, et al: Mutations affecting components of the SWI/SNF complex cause Coffin-Siris syndrome. Nat Genet 44:376–378 (2012).

Tsurusaki Y, Okamoto N, Ohashi H, Mizuno S, Matsumoto N, et al: Coffin-Siris syndrome is a SWI/SNF complex disorder. Clin Genet 85:548–554 (2014).

Van Houdt JKJ, Nowakowska BA, Sousa SB, van Schaik BDC, Seuntjens E, et al: Heterozygous missense mutations in *SMARCA2* cause Nicolaides-Baraitser syndrome. Nat Genet 44:445–449 (2012).

Vergano SS, Deardorff MA: Clinical features, diagnostic criteria, and management of Coffin-Siris syndrome. Am J Med Genet C Semin Med Genet 166C:252–256 (2014).

Vermeulen M, Mulder KW, Denissov S, Pijnappel WW, van Schaik FMA, et al: Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. Cell 131:58–69 (2007).

Vissers LELM, van Ravenswaaij CMA, Admiraal R, Hurst JA, de Vries BBA, et al: Mutations in a new member of the chromodomain gene family cause CHARGE syndrome. Nat Genet 36:955–957 (2004).

Vulto-van Silfhout AT, de Vries BBA, van Bon BWM, Hoischen A, Ruiterkamp-Versteeg M, et al: Mutations in *MED12* cause X-linked Ohdo syndrome. Am J Hum Genet 92:401–406 (2013).

Wieczorek D, Bögershausen N, Beleggia F, Steiner-Haldenstätt S, Pohl E, et al: A comprehensive molecular study on Coffin-Siris and Nicolaides-Baraitser syndromes identifies a broad molecular and clinical spectrum converging on altered chromatin remodeling. Hum Mol Genet 22:5121–5135 (2013).

Woods SA, Robinson HB, Kohler LJ, Agamanolis D, Sterbenz G, Khalifa M: Exome sequencing identifies a novel *EP300* frame shift mutation in a patient with features that overlap Cornelia de Lange syndrome. Am J Med Genet A 164A:251–258 (2014).

Wu Y, Gause M, Xu D, Misulovin Z, Schaaf CA, et al: *Drosophila Nipped-B* mutants model Cornelia de Lange syndrome in growth and behavior. PLoS Genet 11:e1005655 (2015).

Yuan B, Pehlivan D, Karaca E, Patel N, Charng WL, et al: Global transcriptional disturbances underlie Cornelia de Lange syndrome and related phenotypes. J Clin Invest 125:636–651 (2015).