Comprehensive genetic analysis of 57 families with clinically suspected Cornelia de Lange syndrome.

Hiromi Aoi[1,2], Takeshi Mizuguchi[1], Jose Ricardo Ceroni[3], Veronica Eun Hue Kim[3], Isabel Furquim[3], Rachel S Honjo[3], Takuma Iwaki[4], Toshifumi Suzuki[2], Futoshi Sekiguchi[1], Yuri Uchiyama[1,5], Yoshiteru Azuma[1], Kohei Hamanaka[1], Eriko Koshimizu[1], Satoko Miyatake[1,6], Satomi Mitsuhashi[1], Atsushi Takata[1], Noriko Miyake[1], Satoru Takeda[2], Atsuo Itakura[2], Debora R Bertola[3], Chong Ae Kim[3], Naomichi Matsumoto[1]

[1]Department of Human Genetics, Yokohama City University Graduate School of Medicine, Yokohama, Japan; [2]Department of Obstetrics and Gynecology, Juntendo University, Tokyo, Japan; [3]Clinical Genetics Unit, Instituto da Crianca, Hospital das Clinicas HCFMUSP, Faculdade de Medicina, Universidade de Sao Paulo, Sao Paulo, Brazil; [4]Department of Pediatrics, University Hospital, Faculty of Medicine, Kagawa University, Kagawa, Japan; [5]Department of Oncology, Yokohama City University Graduate School of Medicine, Yokohama, Japan; [6]Clinical Genetics Department, Yokohama City University Hospital, Yokohama, Japan.

These authors contributed equally: Chong Ae Kim, Naomichi Matsumoto

Corresponding author: Naomichi Matsumoto, Department of Human Genetics, Yokohama City University Graduate School of Medicine, 3-9 Fukuura, Kanazawa-ku, Yokohama 236-0004, Japan.

Tel.: +81 45787 2606, Fax: +81-45 786 5219, E-mail: naomat@yokohama-cu.ac.jp

Conflict of Interest

The authors declare no conflict of interest.

ABSTRACT

Cornelia de Lange syndrome (CdLS) is a rare multisystem disorder with specific dysmorphic features. Pathogenic genetic variants encoding cohesion complex subunits and interacting proteins (e.g., *NIPBL*, *SMC1A*, *SMC3*, *HDAC8*, and *RAD21*) are the major cause of CdLS. However, there are many clinically diagnosed cases of CdLS without pathogenic variants in these genes. To identify further genetic causes of CdLS, we performed whole exome sequencing in 57 CdLS families, systematically evaluating both single nucleotides variants (SNVs) and copy number variations (CNVs). We identified pathogenic genetic changes in 36 out of 57 (63.2 %) families, including 32 SNVs and four CNVs. Two known CdLS genes, *NIPBL* and *SMC1A*, were mutated in 23 and two cases, respectively. Among the remaining 32 individuals, four genes (*ANKRD11*, *EP300*, *KMT2A*, and *SETD5*) each harbored a pathogenic variant in a single individual. These variants are known to be involved in CdLS-like. Furthermore, pathogenic CNVs were detected in *NIPBL*, *MED13L*, and *EHMT1*, along with pathogenic SNVs in *ZMYND11*, *MED13L*, and *PHIP*. These three latter genes were involved in diseases other than CdLS and CdLS-like. Systematic clinical evaluation of all patients using a recently proposed clinical scoring system showed that *ZMYND11*, *MED13L*, and *PHIP* abnormality may cause CdLS or CdLS-like.

INTRODUCTION

Cornelia de Lange syndrome (CdLS, MIM #122470, #300590, #610759, #614701, #300882) is a rare neurodevelopmental disorder characterized by dysmorphic features, prenatal onset growth restriction, hirsutism, upper limb reduction defects (which range from subtle phalangeal abnormalities to oligodactyly), developmental delay, and intellectual disability.[1] Prevalence of CdLS has been estimated at 1/10,000 to 1/30,000 of live births.[2] In addition to these cardinal phenotypes, patients show cardiac anomalies, gastroesophageal reflux, seizures, and behavioral problems.[3] A combination of signs and symptoms define the classic CdLS phenotype, which is easily recognized from birth by experienced pediatricians and clinical geneticists. However, CdLS is a genetically heterogeneous disorder presenting with extensive phenotypic variability from mild to severe, and with different degrees of facial and limb abnormalities. In addition, CdLS clinically overlaps with several other diseases including Bohring-Optiz syndrome, CHOPS syndrome, and Fryns syndrome.[4, 5] Such heterogeneity makes it difficult to clearly distinguish CdLS from other clinically overlapping diseases. Recently, an international consensus group provided clinical criteria for CdLS.[6] This criteria uses a scoring system comprised of cardinal and suggestive features.

To date, pathogenic variants in at least 15 genes are known to cause CdLS.[7-10] In this regard, cohesin complex or its functionally related genes (e.g., nipped B-like protein [*NIPBL*], structural maintenance of chromosome 1A [*SMC1A*], *SMC3*, histone deacetylase 8 [*HDAC8*], and RAD21

cohesin complex component [*RAD21*]) have been implicated. Approximately 60% of CdLS patients harbor various *NIPBL* variants.[1] Cohesin is a multisubunit protein complex consisting of four core proteins: SMC1, SMC3, RAD21, and stromal antigen (STAG).[6] Chromatin loading of cohesion is regulated by NIPBL.[11] The cohesin complex plays a significant role in mediating sister chromatid cohesion, DNA double-strand break repair, transcriptional regulation, and chromatin organization. Abnormalities of cohesion complex and its related genes in humans are known as cohesinopathy.[12] In addition, variants in *AFF4*, *ANKRD11*, *ARID1B*, *BRD4*, *EP300*, *ESPL1*, *KMT2A*, *PDGFRB*, *SETD5*, and *TAF6* also cause a CdLS-like phenotype.[7-9, 13-15]

In this study, we investigated 57 clinically suspected CdLS individuals by whole exome sequencing (WES). Genetic findings, including single nucleotide variants (SNVs) and copy number variations (CNVs), together with clinical features obtained using recent clinical criteria are presented and discussed.

METHODS

Subjects

In this study, 57 patients were recruited from 57 families, consisting of 56 Brazilian and one Japanese patients. Most of the Brazilian patients were referred by the Brazilian Association of Cornelia de Lange Syndrome (CdLS Brazil) and had the clinical diagnosis suspected by

pediatricians and/or geneticists from all over the country based on distinctive features such as synophrys, arched eyebrows, long philtrum, upper limb abnormalities, and hirsutism. For comparison of clinical manifestations within our cohort and genotype–phenotype correlations, clinical details (including atypical symptoms) were retrospectively reviewed based on recent clinical criteria reported by Kline *et al*.[6] Clinical information was obtained from all 57 patients (Table S1). Peripheral blood leukocytes were collected from patients and their parents after obtaining informed consent. Parental samples were available except for five families (Families 6, 7, 10, 22, and 30). This study was approved by the Institutional Review Boards of Yokohama City University, Faculty of Medicine, and University of Sao Paulo, Faculty of Medicine.

Whole exome sequencing

Genomic DNA was extracted from whole blood sample using QuickGene-610L (Fujifilm, Tokyo, Japan) according to the manufacturer's protocol. Genomic DNA was sheared using a S220 Focused-ultrasonicator (Covaris, Woburn, MA, USA) and captured using the SureSelect Human All Exon V6 Kit (Agilent Technologies, Santa Clara, CA, USA). Paired-end libraries were sequenced on an Illumina HiSeq 2500 platform (Illumina, San Diego, CA, USA) with 101-bp paired-end reads. Quality-controlled reads were aligned to the human reference genome (UCSC hg19, NCBI build 37.1) using NOVOALIGN (http://www.novocraft.com/products/novoalign/). After removal of

polymerase chain reaction (PCR) duplications using Picard (http://broadinstitute.github.io/picard/), variants were called using Genome Analysis Tool Kit (GATK) (https://software.broadinstitute.org/gatk/index.php). Called variants were annotated using ANNOVAR (http://annovar.openbioinformatics.org/en/latest/). Exonic and intronic variants within 30 bp from exon–intron boundaries were examined. Synonymous variants and variants with minor allele frequencies $\geqq$0.01 in our in-house exome database of 575 Japanese individuals or control population databases (including the Exome Aggregation Consortium Browser population (ExAC) [http://exac.broadinstitute.org/] and National Heart, Lung, and Blood Institute (NHLBI) exome variant server [http://evs.gs.washington.edu/EVS/]) were removed. Missense variants were evaluated using Sorting Intolerant From Tolerant (SIFT) (http://sift.jcvi.org/), Polymorphism Phenotyping v2 (Polyphen-2) (http://genetics.bwh.harvard.edu/pph2./), and MutationTaster (http://MutationTaster.org/).

In particular, the focus was on five CdLS genes (*NIPBL*, *SMC1A*, *SMC3*, *HDAC8*, and *RAD21*) and 10 CdLS-like genes (*AFF4*, *ANKRD11*, *ARID1B*, *BRD4*, *EP300*, *ESPL1*, *KMT2A*, *PDGFRB*, *SETD5*, and *TAF6*). Candidate variants were validated by Sanger sequencing. Additionally, *de novo* occurrences were validated when parental samples were available. Parentage was confirmed by analyzing 12 microsatellite markers with Gene Mapper software v4.1.1 (Life Technologies Inc., Carlsbad, CA, USA). The WES performance is summarized in Supplementary Information (Table

S2).

Real-time reverse transcription PCR

To detect aberrant transcripts caused by splice site mutations, reverse transcription PCR (RT-PCR) was performed using total RNA extracted from patient derived lymphoblastoid cell lines. Total RNA was extracted using the RNeasy Plus Mini Kit (Qiagen, Hilden, Germany) and reverse-transcribed into cDNA using the Super Script First Strand Synthesis System (Takara, Kyoto, Japan). Resultant cDNA was used as a template for PCR. PCR amplicons were subjected to Sanger sequencing and aberrant transcripts were characterized. For RT-PCR analysis of *NIPBL*, the forward and reverse primers were: 5′-GAACACTTCAGTTGCTGCAAA-3′ and 5′-CGTTTCCTAGAGGATTCAAAAGC-3′ in Patient 15 with c.3121+1G>A, and 5′-TCATCCAGTTCAGTGTGTGC-3′ and 5′-TCTCAATGACCCTGAAGTGC-3′ in Patient 28 with c.7410+4A>G.

WES-based CNV analysis

Using WES data, CNVs were analyzed by two algorithms: the eXome Hidden Markov Model (XHMM), and a program based on relative depth of coverage ratio, developed by Nord *et al*. (Nord program).[16, 17] For genome-wide screening, XHMM data were first examined in each patient. If

109 causative CNVs were detected using XHMM, altered copy numbers of such regions were further
110 verified using the Nord program. In addition, CNVs at five CdLS genes and 10 CdLS-like genes (see
111 WES section above) were tested by the Nord program.
112
113 **Quantitative polymerase chain reaction**
114 Candidate CNVs were validated by quantitative polymerase chain reaction (qPCR). Real-time qPCR
115 was performed to examine genomic DNA copy number at *NIPBL*, *C5orf42*, *MED13L*, *SMARCA2*,
116 *FREM1*, and *EHMT1* target loci. QuantiFast SYBER Green PCR kit (Qiagen) was used for real-time
117 quantification with amplification monitored on a Rotor-Gene Q real-time PCR cycler (Qiagen).
118 Relative ratios of genomic DNA copy number were calculated using the standard curve method with
119 Rotor-Gene 6000 Series Software 1.7 (Qiagen) by normalizing with autosomal internal control loci
120 (*STXBP1* and/or *FBN1*) and also compared to an unrelated control individual. Information of all
121 primers is available on request.
122
123 **RESULTS**
124 **Flowchart of this study**
125 A flowchart of this study is shown in Figure 1. Because of the genetic and clinical heterogeneity of
126 CdLS, we directly employed WES to effectively screen pathogenic variants in patients with

clinically suspected CdLS. To detect variants in CdLS, CdLS-like, or other possible genes, all 57 patients were analyzed based on autosomal dominant (*de novo*), autosomal recessive, and X-linked modes of inheritance. Based on American College of Medical Genetics and Genomics (ACMG) guidelines[18], we identified 29 pathogenic or likely pathogenic SNVs in two CdLS genes (*NIPBL* and *SMC1A*) and four CdLS-like genes (*ANKRD11*, *EP300*, *KMT2A*, and *SETD5*) (Figure 1). WES-based CNV analysis in 28 SNV-negative patients detected pathogenic CNVs in four patients (4/57 [7.0%]), involving *NIPBL*, *MED13L*, *EHMT1*, and 9q deletion (Figure 1). The remaining 24 cases had neither pathogenic SNVs nor CNVs. Consequently, these cases were subjected to trio-based analysis, except for two cases whose parental samples were unavailable. We detected three pathogenic variants in genes associated with diseases other than CdLS and CdLS-like: *ZMYND11*, *MED13L*, and *PHIP*. Altogether, if all abnormalities were included, we identified pathogenic or likely pathogenic variants in 36 out of 57 cases (63.2%) (Figure 1). Thirty-one of 36 variants occurred *de novo*, unless biological parental samples were unavailable. One variant was inherited from a mosaic mother (Patient 53). Twenty-three of 32 pathogenic SNVs were novel (Table 1).

Pathogenic SNVs in CdLS genes

We detected 22 pathogenic SNVs in *NIPBL* (22/57 [38.6%]) and two in *SMC1A* (2/57 [3.5%]) (Table 1 and Figure 1). Among 22 *NIPBL* SNVs, 14 were novel. Meanwhile, NM_0133433.3:c.6893G>A,

145 p.Arg2298His was repeatedly detected (Patients 2 and 50). Three splice site variants in *NIPBL*
146 (NM_0133433.3:c.3121+1G>A, c.7410+4A>G, and c.5329-15A>G) were detected in Patients 15,
147 28, and 54, respectively. These variants were previously described, and only c.5329-15A>G was
148 shown to result in abnormal splicing.[19-21] The other c.3121+1G>A and c.7410+4A>G mutations
149 were never examined at cDNA level.[19-21] Therefore, by RT-PCR using cDNA derived from
150 lymphoblastoid cells, we confirmed aberrant splicing in both Patient 15, with c.3121+1G>A, and
151 Patient 28, with c.7410+4A>G (Figure S1). Regarding the two missense variants in *SMC1A*,
152 NM_006306.3:c.1152C>G, p.Lys362Asn was novel.
153
154 **Pathogenic SNVs in CdLS-like genes**
155 We also detected pathogenic variants in four CdLS-like genes: *ANKRD11* (2/57 [3.5%]), *EP300*
156 (1/57 [1.8%]), *KMT2A* (1/57 [1.8%]), and *SETD5* (1/57 [1.8%]) (Table 1 and Figure 1), whose
157 pathogenic variants are known to cause KBG syndrome (MIM #148050), Rubinstein–Taybi
158 syndrome 2 (MIM #613684), Wiedemann–Steiner syndrome (MIM #605130), and mental
159 retardation autosomal dominant 23 (MIM #615761), respectively. These disorders all share
160 overlapping clinical features with CdLS. All five variants were novel, occurring *de novo* except for
161 an *EP300* variant, which was due to unavailable parental samples. According to the ACMG
162 guideline, the *EP300* variant can be classified as likely pathogenic since it is protein length changing

mutation due to in-frame deletion (PM4), it is absent from control (including the ExAC, NHLBI, and gnomAD [https://gnomad.broadinstitute.org/]) (PM2), it is predicted to be deleterious by PROVEAN [http://provean.jcvi.org/seq_submit.php] and CADD [https://cadd.gs.washington.edu/] with a score of 23.6 and 21.1, respectively (PP3), and the phenotype of patient is considered reasonable as Cornelia de Lange syndrome-like (PP4).

Pathogenic CNVs

Using the XHMM and Nord program, we detected four pathogenic CNVs in four patients (Table 1 and Figure 1). These were confirmed by qPCR. Patient 9 has a 94-kb deletion at 5p13.2, encompassing exons 22 to 47 of *NIPBL* and the last exon of *C5orf42* (Figure S2a). Partial deletions of *NIPBL* have been reported in patients with CdLS, and *NIPBL* haploinsufficiency is apparently deleterious.[22] Patient 34 has a 4.2-Mb deletion at 12q24.1-q24.23, which contains 40 genes including the entire *MED13L* gene (Figure S2b). Patient 51 has a 14.1-Mb deletion at 9p24.3-p22.3, involving 44 genes and an adjacent 571-kb duplication at 9p22.3, altogether encompassing four genes (Figure S2c). Patient 52 has a 774-kb deletion at 9q34.3, containing 14 genes including the entire *EHMT1* gene (Figure S2d).

Variants in genes associated with diseases other than CdLS and CdLS-like

By trio-based analysis, we identified pathogenic or likely pathogenic variants in *ZMYND11*, *MED13L*, and *PHIP*. These variants are involved in other diseases, but never CdLS or CdLS-like.

A novel *ZMYND 11* frameshift variant (NM_006624.5:c.1438delG, p.Asp480Thrfs*3) was detected in Patient 53, who had typical CdLS features including left hand oligodactyly (Tables 1 and S1, and Figure 2a–e). Based on apparent double sequences implying low mutant allele peaks in the electropherogram of the mother, maternal mosaicism of this variant was examined (Figure S3). Deep sequencing of PCR products encompassing the maternal variant confirmed mosaicism (mutant/mutant+wild-type reads = 2835/27596 [10.3%)]), while Patient 53 showed heterozygosity (mutant/mutant+wild-type reads = 12514/27211 [46.0%]) (Table S3). By TA cloning of PCR products spanning the maternal variant, wild-type and mutant alleles were clearly recognized by Sanger sequencing (Figure S3), yet the mother had no CdLS-like features. *ZMYND11* has been reported as a critical gene for 10p15.3 microdeletion syndrome, including neurodevelopmental disorder, characteristic dysmorphic features, and other more frequent symptoms, such as behavioral disturbances, hypotonia, seizures, low birth weight, short stature, genitourinary malformations, and recurrent infections.[23]

A novel *MED13L* missense mutation, NM_015335.4:c.6485C>A, p.Thr2162Lys was detected in Patient 5 (Table S4). *MED13L* variants cause distinctive dysmorphic features and mental retardation with or without cardiac defects (MIM #608771), known as *MED13L* haploinsufficiency syndrome.[24]

The missense variant identified here is novel, but another variant at the same nucleotide position was previously identified, which leads to a different amino acid substitution (NM_015335.4:c.6485C>T, p.Thr2162Met).[25] Of note, we also detected a 4.2-Mb deletion involving *MED13L* in Patient 34 (Table S4 and Figure S2b). Further, a novel *PHIP* missense mutation (NM_017934.7:c.1156G>A, p.Asp386Asn) was detected in Patient 56. *PHIP* haploinsufficiency causes dysmorphic CdLS-like features, developmental delay, intellectual disability, and obesity.[26]

In the remaining 21 undetermined families, NM_025146.4:c.93C>G, p.Tyr31* in *NAA50* (encoding N-alpha-acetyltransferase 50) attracted our attention because it encodes a cohesin complex component (see Discussion). *NAA50* variants have not previously been described.

Clinical evaluation of CdLS patients using a new scoring system

Of the 57 patients with CdLS, their clinical features were re-evaluated based on the clinical scoring system reported by Kline *et al*.[6] With this scoring system, clinical features of clinically suspected CdLS are classified as cardinal (2 points each if presented) and suggestive (1 point each if presented). Clinical scores ≥ 11, 10 or 9, 8–4, and < 4 points, are classified as: classic CdLS, non-classic CdLS, sufficiently suspected to warrant molecular testing for CdLS, and insufficient indication for CdLS molecular testing, respectively. All 57 patients were classified using the above clinical scoring system (Table S1 and Figure 3). Twenty-five patients were categorized as classic

CdLS, 17 patients as non-classic CdLS, and 15 patients as sufficiently suspected to warrant molecular testing for CdLS. No patients were insufficient to indicate molecular testing. The proportion of *NIPBL* variants was 60% (15/25), 35.3% (6/17), and 13.3% (2/15) in each class, respectively. Ratios of *NIPBL* variants were compared between two of three classes, with a significant difference recognized only between classic CdLS and sufficiently suspected to warrant molecular testing for CdLS (χ^2 test, $p < 0.05$) (Figure 3). *NIPBL* variants in classic CdLS were more frequent than sufficiently suspected to warrant molecular testing for CdLS.

Interestingly, Patient 53 with a *ZMYND11* frameshift variant showed classic CdLS (15 points) with oligodactyly (Figure 2a–e). Therefore *ZMYND11* could be included as a CdLS or CdLS-like genes, although *ZMYND11* variants have not been reported in CdLS. Patients 5 (SNV) and 34 (CNV) with *MED13L* abnormality showed clinical scores of 8 and 9 points, respectively, and were consequently classified as sufficiently suspected to warrant molecular testing for CdLS and non-classic CdLS (Figure 2f–h, i–l). Patient 56 with a missense variant in *PHIP* showed CdLS-like features (6 points), including synophrys, long curly eyelashes, anteverted nostrilis, and depressed nasal bridge, although obesity was retrospectively inconsistent with CdLS (Figure 2m–p). This clinical information is summarized in Table S1.

DISCUSSION

Using WES, we identified pathogenic variants in 36 out of 57 (63.2%) patients with clinically suspected CdLS. The diagnostic yield was comparatively higher than previous studies (40–60%) as previous studies used panel or Sanger sequencing of only major CdLS genes.[27-30] Advantages of WES are clearly indicated here as CdLS and CdLS-like patients are genetically and clinically heterogeneous. Using a large clinical exome sequencing cohort, a recent genotype-driven approach of cohesinopathy also emphasized the utility of clinical exome sequencing to provide molecular diagnoses for cohesinopathies with extensive genetic and phenotypic heterogeneity, as well as to detecting mosaic variants in patients.[12] We detected no mosaicism variants in our patients, although it may be difficult to detect extremely low prevalence mosaic variants by WES.

Based on recent clinical scores,[6] *NIPBL* variants are more likely to be found in classic CdLS. Moreover, we detected a *ZMYND11* frameshift variant, NM_006624.5:c.1438delG, p.Asp480Thrfs*3 in Patient 53 with classic CdLS. ZMYND11 (also known as BS69) contains a tandem “reader” module of histone modifications, which recognizes and binds histone H3.3 trimethylated at Lys-36 (H3.3K36me3). Subsequently, this recruits histone demethylases, histone deacetylases, and the SWI/SNF chromatin-remodeling complex to reset chromatin to a relatively repressive state and prevent further transcription.[31, 32] Except for *ZMYND11*, all pathogenic variants in genes for diseases other than CdLS and CdLS-like (*MED13L* and *PHIP*) were detected in patients with scores < 9.

We found two patients with *MED13L* abnormality and one patient with a *PHIP* variant. MED13 is a subunit of the cyclin-dependent kinase 8 (CDK8) module comprised of reversible association of four subunits: cyclin C, CDK8, mediator complex subunit (MED)12/MED12L, and MED13/MED13L. The module binds the mediator complex to regulate its activity. The mediator complex bridges between gene-specific activators bound to regulatory elements and general transcription machinery comprising RNA polymerase II and general transcription factors.[33, 34] PHIP is a H3K4 methylation-binding protein that interacts with chromatin modifications associated with promoters and transcriptional cis-regulatory elements.[35] Interestingly, ZMYND11, MED13L, and PHIP are all core components of transcriptional regulatory pathways. Recently, CdLS and CdLS-like disorders were reported not only as cohesinopathies but also as "transcriptomopathies". [15] Actually, *AFF4*, *ANKRD11*, *ARID1B*, *BRD4*, *EP300*, *KMT2A*, *SETD5*, and *TAF6* have been found in patients with several clinical features overlapping with CdLS, and are related to epigenetic modification, chromatin remodeling, and transcriptional regulation pathway [8, 15, 36, 37] (Table S5). Interactive networks of 18 genes associated with CdLS and CdLS-like features were analyzed using GeneMANIA (https://genemania.org/), which covers physical interactions, pathways, and shared protein domains (Figure 4). As expected, genes encoding cohesion complex and its regulatory factors (*NIBPL*, *SMC1A*, *SMC3*, *HDAC8*, *RAD21*, and *ESPL1*) strongly interact with each other. *ZMYND11* and *PHIP* share protein domains with other genes encoding histone modification factors and transcriptional regulation factors. *MED13L* shares a

common pathway with *EP300*, and is involved in regulation of RNA polymerase Ⅱ. HIF1A is a hypoxia inducible factor subunit that induces recruitment of CDK8-mediator complex and p300 (encoding EP300) for histone acetyltransferase to stimulate RNA polymerase II elongation.[38] These functional links in three genes (*ZMYND11*, *PHIP*, and *MED13L*) may be related to CdLS-like features. Patient 51 has a 14.1-Mb deletion at 9p24.3-p22.3 (involving 44 RefSeq genes) adjacent to a 571-kb duplication at 9p22.3, containing four genes (Figure S2c). Critical genes of 9p deletion syndrome include *DMRT* (*DMRT1*, *DMRT2*, and *DMRT3* cluster) for gonadal dysgenesis from complete sex reversal to milder phenotypes in 46,XY patients,[39] *FREM1* for craniosynostosis including trigonocephaly,[40] and *DOCK8*, *KANK1*, *SLC1A1*, and *GLDC* for developmental delay and neurological disorders.[41] Trigonocephaly is one of the major features of 9p deletion syndrome, but absent in our patient. Trigonocephaly was previously mapped to a critical 4.7-Mb region at 9p22.2-p23, including *FREM1* and *CER1*.[42] Interestingly, our patient has a duplication of this critical region, and instead of trigonocephaly, exhibited delayed closure of the anterior fontanelle at 3 years of age Thus, it is conceivable that *FREM1* and/or *CER1* are potentially dosage sensitive genes related to cranial bone development and closure. In addition, *SMARCA2* was included in the deletion region. *SMARCA2* is a known causative gene for Nicolaides–Baraitser syndrome (MIM #601358), which shares several CdLS features.[43] To date, 78 variants are registered in the Human Gene Mutation Database (HGMD) V.2019.1, but no truncating variants. *SMARCA2* variants are predicted to act in a dominant-negative or

289 gain-of-function manner rather than haploinsufficiency. Indeed, it has been suggested that *SMARCA2* might
290 not be a critical gene for 9p deletion syndrome.
291 Patient 52 has a 773.8-kb deletion at 9q34.3, which contains 14 genes including *EHMT1*. Intragenic
292 *EHMT1* variants or submicroscopic 9q34.3 deletion causes Kleefstra syndrome with distinct facial
293 features, hypotonia, developmental delay, and intellectual disability.[44] *EHMT1* encodes a histone H3
294 Lys-9 methyltransferase and is consequently involved in chromatin remodeling.[45] Similar to patients
295 with CdLS, our patient showed dysmorphic features, including synophrys, long curly eyelashes, and
296 depressed nasal bridge, but no limb abnormalities (Figure 2u–w). Clinical score was 5 points,
297 suggesting that the patient is likely compatible with Kleefstra syndrome rather than CdLS. Nonetheless,
298 it is sometimes difficult to clearly differentiate these two disorders.
299 In 21 undetermined families, a *de novo* nonsense variant (NM_025146.4:c.93C>G, p.Tyr319*) was
300 detected in *NAA50* in Patient 19 with classic CdLS features (12 points). The variant was confirmed by
301 Sanger sequencing. This variant was not registered in control population databases (ExAC and
302 gnomAD). According to ExAC, probability of loss-of-function intolerance (pLI) score of 0.88 suggest
303 intolerance to loss-of-function variant. To date, no variants are registered in HGMD V.2019.1. *NAA50*
304 encodes a N-terminal acetyltransferase required for chromosome segregation during mitosis. It has
305 been reported that NAA50 is required for sister chromatid cohesion during Drosophila wing
306 development, and most likely regulates correct interaction between the cohesin subunits, RAD21 and

SMC3.[46] These findings support that *NAA50* truncation variants may cause the candidate variants of CdLS. Further studies of *NAA50* variants in patients with CdLS are necessary.

In conclusion, we have achieved a high genetic diagnosis rate of 63.2% by WES in patients with clinically diagnosed CdLS. Moreover, we have newly detected *ZMYND11*, *MED13L*, and *PHIP* variants potentially linked to CdLS or CdLS-like through abnormality of transcriptional regulation together with *NAA50* variant.

Acknowledgments

We would like to thank the patients and their families, and especially the Brazilian Association of Cornelia de Lange Syndrome (CdLS Brazil) for participating in this study. We also thank N. Watanabe, T. Miyama, M. Sato, S. Sugimoto, and K. Takabe for their technical assistance. This work was supported by Japan Agency for Medical Research and Development (AMED) under grant numbers JP18ek0109280, JP18dm0107090, JP18ek0109301, JP18ek0109348, and JP18kk0205001; by Japan Society for the Promotion of Science (JSPS) KAKENHI under grant numbers JP17H01539, JP16H05357, JP16H06254, JP17K16132, JP17K10080, JP17K15630, and JP17H06994; by the Ministry of Health, Labour, and Welfare; and by the Takeda Science Foundation. We thank Sarah Williams, PhD, from Edanz Group (www.edanzediting.com) for editing a draft of this manuscript.

Supplementary information is available at Journal of human genetics's website.

References

1. Nizon, M., Henry, M., Michot, C., Baumann, C., Bazin, A., Bessieres, B. et al. A series of 38 novel germline and somatic mutations of NIPBL in Cornelia de Lange syndrome. Clinical Genetics. 2016;89:584–89.
2. Mannini, L., Cucco, F., Quarantotti, V., Krantz, I.D., Musio, A. Mutation spectrum and genotype-phenotype correlation in Cornelia de Lange syndrome. Human Mutation. 2013;34:1589–96.
3. Bhuiyan, Z.A., Klein, M., Hammond, P., van Haeringen, A., Mannens, M.M., Van Berckelaer-Onnes, I. et al. Genotype-phenotype correlations of 39 patients with Cornelia De Lange syndrome: the Dutch experience. Journal of Medical Genetics. 2006;43:568–75.
4. Bedoukian, E., Copenheaver, D., Bale, S., Deardorff, M. Bohring-Opitz syndrome caused by an ASXL1 mutation inherited from a germline mosaic mother. American Journal of Medical Genetics. Part A. 2018;176:1249–52.
5. McInerney-Leo, A.M., Harris, J.E., Gattas, M., Peach, E.E., Sinnott, S., Dudding-Byth, T. et al. Fryns Syndrome Associated with Recessive Mutations in PIGN in two Separate Families. Human Mutation. 2016;37:695–702.
6. Kline, A.D., Moss, J.F., Selicorni, A., Bisgaard, A.M., Deardorff, M.A., Gillett, P.M. et al. Diagnosis and management of Cornelia de Lange syndrome: first international consensus statement. Nature Reviews Genetics. 2018;19:649–66.
7. Parenti, I., Teresa-Rodrigo, M.E., Pozojevic, J., Ruiz Gil, S., Bader, I., Braunholz, D. et al. Mutations in chromatin regulators functionally link Cornelia de Lange syndrome and clinically overlapping phenotypes. Human Genetics. 2017;136:307–20.
8. Olley, G., Ansari, M., Bengani, H., Grimes, G.R., Rhodes, J., von Kriegsheim, A. et al. BRD4 interacts with NIPBL and BRD4 is mutated in a Cornelia de Lange-like syndrome. Nature Genetics. 2018;50:329–32.
9. Ansari, M., Poke, G., Ferry, Q., Williamson, K., Aldridge, R., Meynert, A.M. et al. Genetic heterogeneity in Cornelia de Lange syndrome (CdLS) and CdLS-like phenotypes with observed and predicted levels of mosaicism. Journal of Medical Genetics. 2014;51:659–68.
10. Yavarna, T., Al-Dewik, N., Al-Mureikhi, M., Ali, R., Al-Mesaifri, F., Mahmoud, L. et al. High diagnostic yield of clinical exome sequencing in Middle Eastern patients with Mendelian disorders. Human Genetics. 2015;134:967–80.
11. Muto, A., Schilling, T.F. Zebrafish as a Model to Study Cohesin and

Cohesinopathies. In: Methods in Molecular Biology. Clifton, N.J; 2017. p. 177–196.
12. Yuan, B., Neira, J., Pehlivan, D., Santiago-Sim, T., Song, X., Rosenfeld, J. et al. Clinical exome sequencing reveals locus heterogeneity and phenotypic variability of cohesinopathies. Genetics in Medicine: Official Journal of the American College of Medical Genetics. 2019;3:663-75.
13. Izumi, K., Nakato, R., Zhang, Z., Edmondson, A.C., Noon, S., Dulik, M.C. et al. Germline gain-of-function mutations in AFF4 cause a developmental syndrome functionally linking the super elongation complex and cohesin. Nature Genetics. 2015;47:338–44.
14. Woods, S.A., Robinson, H.B., Kohler, L.J., 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. American Journal of Medical Genetics. Part A. 2014;164a:251–58.
15. Yuan, B., Pehlivan, D., Karaca, E., Patel, N., Charng, W.L., Gambin, T. et al. Global transcriptional disturbances underlie Cornelia de Lange syndrome and related phenotypes. The Journal of Clinical Investigation. 2015;125:636–51.
16. Fromer, M., Moran, J.L., Chambert, K., Banks, E., Bergen, S.E., Ruderfer, D.M. et al. Discovery and statistical genotyping of copy-number variation from whole-exome sequencing depth. American Journal of Human Genetics. 2012;91:597–607.
17. Nord, A.S., Lee, M., King, M.C., Walsh, T. Accurate and exact CNV identification from targeted high-throughput sequence data. BMC Genomics. 2011;12:184.
18. Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J. et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genetics in medicine : official journal of the American College of Medical Genetics. 2015;17:405-24.
19. Selicorni, A., Russo, S., Gervasini, C., Castronovo, P., Milani, D., Cavalleri, F. et al. Clinical score of 62 Italian patients with Cornelia de Lange syndrome and correlations with the presence and type of NIPBL mutation. Clinical Genetics. 2007;72:98–108.
20. Gillis, L.A., McCallum, J., Kaur, M., DeScipio, C., Yaeger, D., Mariani, A. et al. NIPBL mutational analysis in 120 individuals with Cornelia de Lange syndrome and evaluation of genotype-phenotype correlations. American Journal of Human Genetics. 2004;75:610–23.
21. Teresa-Rodrigo, M.E., Eckhold, J., Puisac, B., Pozojevic, J., Parenti, I., Baquero-Montoya, C. et al. Identification and Functional Characterization of Two Intronic

NIPBL Mutations in Two Patients with Cornelia de Lange Syndrome. BioMed Research International. 2016;8742939.

22. Cheng, Y.W., Tan, C.A., Minor, A., Arndt, K., Wysinger, L., Grange, D.K. et al. Copy number analysis of NIPBL in a cohort of 510 patients reveals rare copy number variants and a mosaic deletion. Molecular Genetics & Genomic Medicine. 2014;2:115–23.

23. Tumiene, B., Ciuladaite, Z., Preiksaitiene, E., Mameniskiene, R., Utkus, A., Kucinskas, V. Phenotype comparison confirms ZMYND11 as a critical gene for 10p15.3 microdeletion syndrome. Journal of Applied Genetics. 2017;58:467–74.

24. Smol, T., Petit, F., Piton, A., Keren, B., Sanlaville, D., Afenjar, A. et al. MED13L-related intellectual disability: involvement of missense variants and delineation of the phenotype. Neurogenetics. 2018;19:93–103.

25. Bowling, K.M., Thompson, M.L., Amaral, M.D., Finnila, C.R., Hiatt, S.M., Engel, K.L. et al. Genomic diagnosis for children with intellectual disability and/or developmental delay. Genome Medicine. 2017;9:43.

26. Jansen, S., Hoischen, A., Coe, B.P., Carvill, G.L., Van Esch, H., Bosch, D.G.M. et al. A genotype-first approach identifies an intellectual disability-overweight syndrome caused by PHIP haploinsufficiency. European Journal of Human Genetics. 2018;26:54–63.

27. Mariani, M., Decimi, V., Bettini, L.R., Maitz, S., Gervasini, C., Masciadri, M. et al. Adolescents and adults affected by Cornelia de Lange syndrome: A report of 73 Italian patients. American Journal of Medical Genetics. Part C, Seminars in Medical Genetics. 2016;172:206–13.

28. Marchisio, P., Selicorni, A., Bianchini, S., Milani, D., Baggi, E., Cerutti, M. et al. Audiological findings, genotype and clinical severity score in Cornelia de Lange syndrome. International Journal of Pediatric Otorhinolaryngology. 2014;78:1045–48.

29. Hei, M., Gao, X., Wu, L. Clinical and genetic study of 20 patients from China with Cornelia de Lange syndrome. BMC Pediatrics. 2018;18:64.

30. Mehta, D., Vergano, S.A., Deardorff, M., Aggarwal, S., Barot, A., Johnson, D.M. et al. Characterization of limb differences in children with Cornelia de Lange Syndrome. American journal of medical genetics. Part C, Seminars in Medical Genetics. 2016;172:155–62.

31. Wen, H., Li, Y., Li, H., Shi, X. ZMYND11: an H3.3-specific reader of H3K36me3. Cell Cycle. 2014;13:2153–54.

32. Wen, H., Li, Y., Xi, Y., Jiang, S., Stratton, S., Peng, D. et al. ZMYND11 links histone

H3.3K36me3 to transcription elongation and tumour suppression. Nature. 2014:508:263–68.

33. Utami, K.H., Winata, C.L., Hillmer, A.M., Aksoy, I., Long, H.T., Liany, H. et al. Impaired development of neural-crest cell-derived organs and intellectual disability caused by MED13L haploinsufficiency. Human Mutation. 2014;35:1311–20.

34. Miao, Y.L., Gambini, A., Zhang, Y., Padilla-Banks, E., Jefferson, W.N., Bernhardt, M.L. et al. Mediator complex component MED13 regulates zygotic genome activation and is required for postimplantation development in the mouse. Biology of Reproduction. 2018;98:449–64.

35. Morgan, M.A.J., Rickels, R.A., Collings, C.K., He, X., Cao, K., Herz, H.M. et al. A cryptic Tudor domain links BRWD2/PHIP to COMPASS-mediated histone H3K4 methylation. Genes & Development. 2017;31:2003–14.

36. Izumi, K. Disorders of Transcriptional Regulation: An Emerging Category of Multiple Malformation Syndromes. Molecular Syndromology. 2016;7:262–73.

37. Ka, M., Kim, W.Y. ANKRD11 associated with intellectual disability and autism regulates dendrite differentiation via the BDNF/TrkB signaling pathway. Neurobiology of Disease. 2018;111:138–52.

38. Galbraith, M.D., Allen, M.A., Bensard, C.L., Wang, X., Schwinn, M.K., Qin, B. et al. HIF1A employs CDK8-mediator to stimulate RNAPII elongation in response to hypoxia. Cell. 2013;**153**:1327-39.

39. Onesimo, R., Orteschi, D., Scalzone, M., Rossodivita, A., Nanni, L., Zannoni, G.F. et al. Chromosome 9p deletion syndrome and sex reversal: novel findings and redefinition of the critically deleted regions. American Journal of Medical Genetics. Part A. 2012;158a:2266–71.

40. Vissers, L.E., Cox, T.C., Maga, A.M., Short, K.M., Wiradjaja, F., Janssen, I.M. et al. Heterozygous mutations of FREM1 are associated with an increased risk of isolated metopic craniosynostosis in humans and mice. PLoS Genetics. 2011;7:e1002278.

41. Recalcati, M.P., Bellini, M., Norsa, L., Ballarati, L., Caselli, R., Russo, S. et al. Complex rearrangement involving 9p deletion and duplication in a syndromic patient: genotype/phenotype correlation and review of the literature. Gene. 2012;502:40–5.

42. Kawara, H., Yamamoto, T., Harada, N., Yoshiura, K., Niikawa, N., Nishimura, A. et al. Narrowing candidate region for monosomy 9p syndrome to a 4.7-Mb segment at 9p22.2-p23. American Journal of Medical Genetics. Part A. 2006;140:373–77.

43. Sousa, S.B., Hennekam, R.C. Phenotype and genotype in Nicolaides-Baraitser syndrome. American Journal of Medical Genetics. Part C, Seminars in Medical

Genetics. 2014;166c:302–14.
44. Willemsen, M.H., Vulto-van Silfhout, A.T., Nillesen, W.M., Wissink-Lindhout, W.M., van Bokhoven, H., Philip, N. et al. Update on Kleefstra Syndrome. Molecular Syndromology. 2012;2:202–12.
45. Iacono, G., Dubos, A., Meziane, H., Benevento, M., Habibi, E., Mandoli, A. et al. Increased H3K9 methylation and impaired expression of Protocadherins are associated with the cognitive dysfunctions of the Kleefstra syndrome. Nucleic Acids Research. 2018;46:4950–65.
46. Ribeiro, A.L., Silva, R.D., Foyn, H., Tiago, M.N., Rathore, O.S., Arnesen, T. et al. Naa50/San-dependent N-terminal acetylation of Scc1 is potentially important for sister chromatid cohesion. Scientific Reports. 2016;6:39118.
47. Gillis, L.A., McCallum, J., Kaur, M., DeScipio, C., Yaeger, D., Mariani, A. et al. NIPBL mutational analysis in 120 individuals with Cornelia de Lange syndrome and evaluation of genotype-phenotype correlations. American journal of human genetics. 2004;75:610-23.
48. Oliveira, J., Dias, C., Redeker, E., Costa, E., Silva, J., Reis Lima, M. et al. Development of NIPBL locus-specific database using LOVD: from novel mutations to further genotype-phenotype correlations in Cornelia de Lange Syndrome. Human mutation. 2010;31:1216-22.
49. Deardorff, M.A., Kaur, M., Yaeger, D., Rampuria, A., Korolev, S., Pie, J. et al. Mutations in cohesin complex members SMC3 and SMC1A cause a mild variant of cornelia de Lange syndrome with predominant mental retardation. American journal of human genetics. 2007;80:485-94.

Titles and legends to figures

Figure 1. **Flowchart of this study.**

All the 57 patients with clinically suspected CdLS were analyzed by whole exome sequencing (WES). Twenty-nine patients had pathogenic single nucleotide variants (SNVs) in two CdLS genes (*NIPBL* and *SMC1A*) and four CdLS-like genes (*ANKRD11*, *EP300*, *KMT2A*, and *SETD5*). WES-based copy number variation (CNV) analysis in patients with no causative SNVs identified pathogenic CNVs in four patients. The remaining 24 patients with neither pathogenic SNVs nor CNVs were subjected to trio-based analysis, except for two cases whose parental samples were unavailable. Three causative variants were identified in *ZMYND11*, *MED13L*, and *PHIP*. Diagnostic yield was 63.2 % (36/57) when all 32 SNVs (32/57 [56.1%]) and four CNVs (4/57 [7.0%]) were included. A novel candidate variant was detected in *NAA50*.

Figure 2. Clinical photographs of individuals with *ZMYND11, MED13L,* and *PHIP* abnormalities. (a–e) Photos of Patient 53 with a *ZMYND11* frameshift mutation. (a, b) Facial features include microcephaly, synophrys, highly arched eyebrows, long curly eyelashes, low set ears, anteverted nasal nostrilis, long philtrum, thin upper lip, downturned corners of the mouth, and micrognathia. (c) Note left hand oligodactyly (only one finger). (d, e) Right hand and bilateral feet. Right hand shows abnormal palmer crease. Feet show no abnormalities. (f–h) Facial photos of

Patient 5 with a *MED13L* missense mutation at (f) 3 months and (g) 18 years. (h) Broad forehead, synophrys, long curly eyelashes, low set ears, anteverted nasal bridge, and full cheeks are seen at 23 years. (i–l) Clinical features of Patient 34 with a 4.2-Mb deletion involving *MED13L*. (i) Note synophrys, arched eyebrows, upslanting palpebral fissures, long curly eyelashes, low set ears, anteverted nasal bridge, and bulbous nasal tip. (j) Hirsutism in the back. (k, l) Bilateral hands and feet. Hands show clinodactyly of the fifth finger. Feet show no abnormalities. (m–p) Photos of Patient 56 with a *PHIP* missense mutation. (m) Facial features include macrocephaly, synophrys, long curly eyelashes, anteverted nasal nostrilis, depressed nasal bridge, and short neck. (n) Full whole body view with obesity at 11 years (weight, 82.5 kg [> 95 percentile]; height, 157.5 cm [> 95 percentile]; occipital frontal circumference, 58 cm [> 98 percentile]). (o, p) Hands and feet were normal. (q–t) Photos of Patient 51 with a 4.1-Mb deletion at 9p24.3-p22.3 adjacent to a 571-kb duplication at 9p22.3. (q, r) Facial features include synophrys, upslanting palpebral fissures, anteverted nostrilis, and long philtrum. (s, t) Hands were normal. (u–w) Phenotype of Patient 52 with a 773.8-kb deletion at 9q34.3. (u) Facial features include synophrys, long curly eyelashes, and depressed nasal bridge. (v, w) Hands and feet were normal.

Figure 3. Classification of 57 CdLS patients by clinical score. All patients were classified based on clinical score. Scores of ≥ 11, 10 or 9, 8–4, and < 4 enabled categorization of four classes: classic

CdLS, non-classic CdLS, sufficiently suspected to warrant molecular testing for CdLS indicated, and insufficient to indicate molecular testing for CdLS. The 57 patients were classified to classic CdLS (*N* = 25), non-classic CdLS (*N* = 17), and sufficiently suspected to warrant molecular testing for CdLS indicated (*N* = 15). The number of individuals with variants are indicated in rows of genes.

Figure 4. Schematic presentation of interacting networks of mutated genes in CdLS and CdLS-like. Interactive gene networks of mutated genes with CdLS and CdLS-like features. Three networks are highlighted using GeneMANIA (https://genemania.org/), based on physical interactions (red line), connecting pathways (blue line), and shared protein domains (green line).

Figure 1

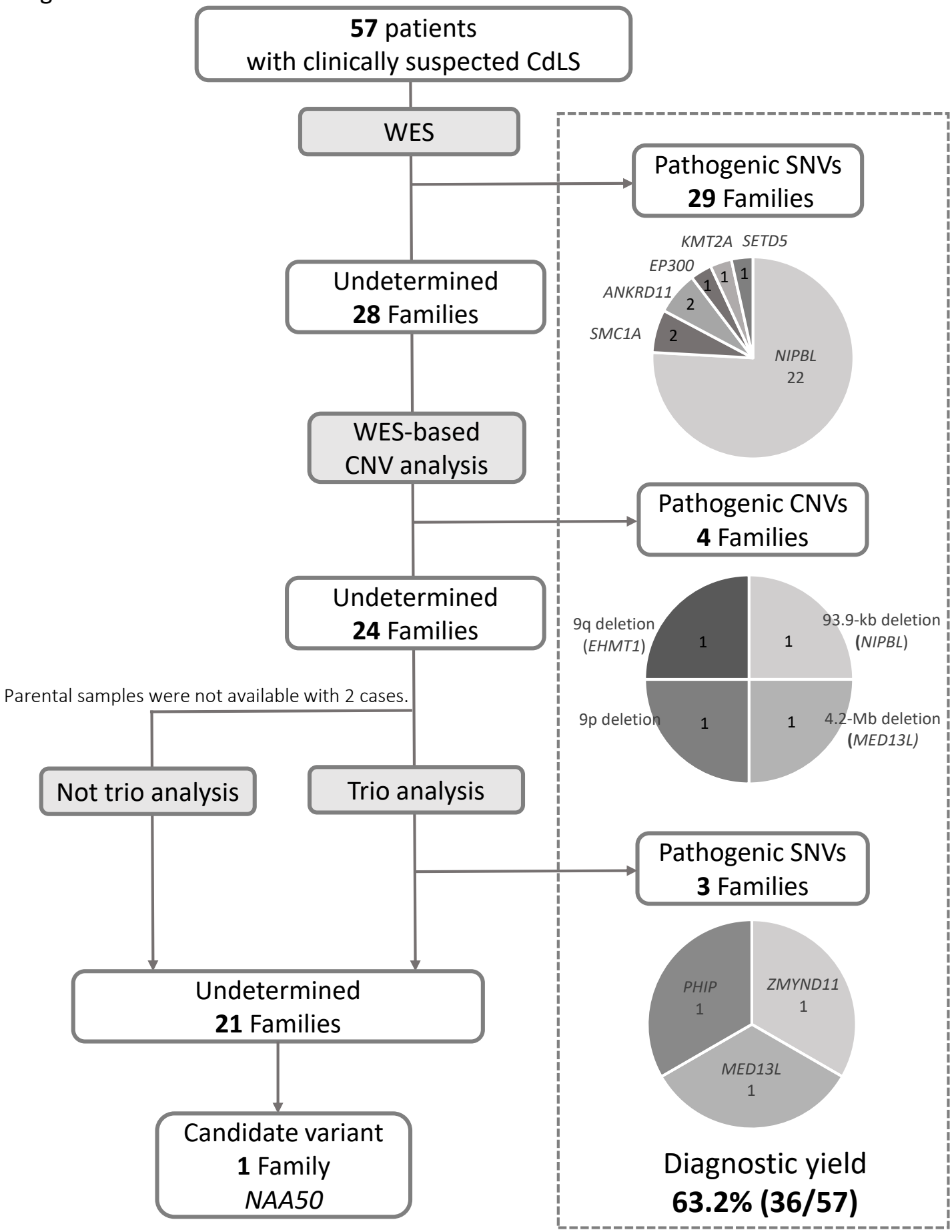

57 patients
with clinically suspected CdLS
WES
Pathogenic SNVs
29 Families
KMT2A
SETD5
EP300
ANKRD11
SMC1A
NIPBL
22
Undetermined
28 Families
WES-based
CNV analysis
Pathogenic CNVs
4 Families
9q deletion
(EHMT1)
93.9-kb deletion
(NIPBL)
9p deletion
4.2-Mb deletion
(MED13L)
Undetermined
24 Families
Parental samples were not available with 2 cases.
Not trio analysis
Trio analysis
Pathogenic SNVs
3 Families
PHIP
ZMYND11
MED13L
Undetermined
21 Families
Candidate variant
1 Family
NAA50
Diagnostic yield
63.2% (36/57)

Figure 2

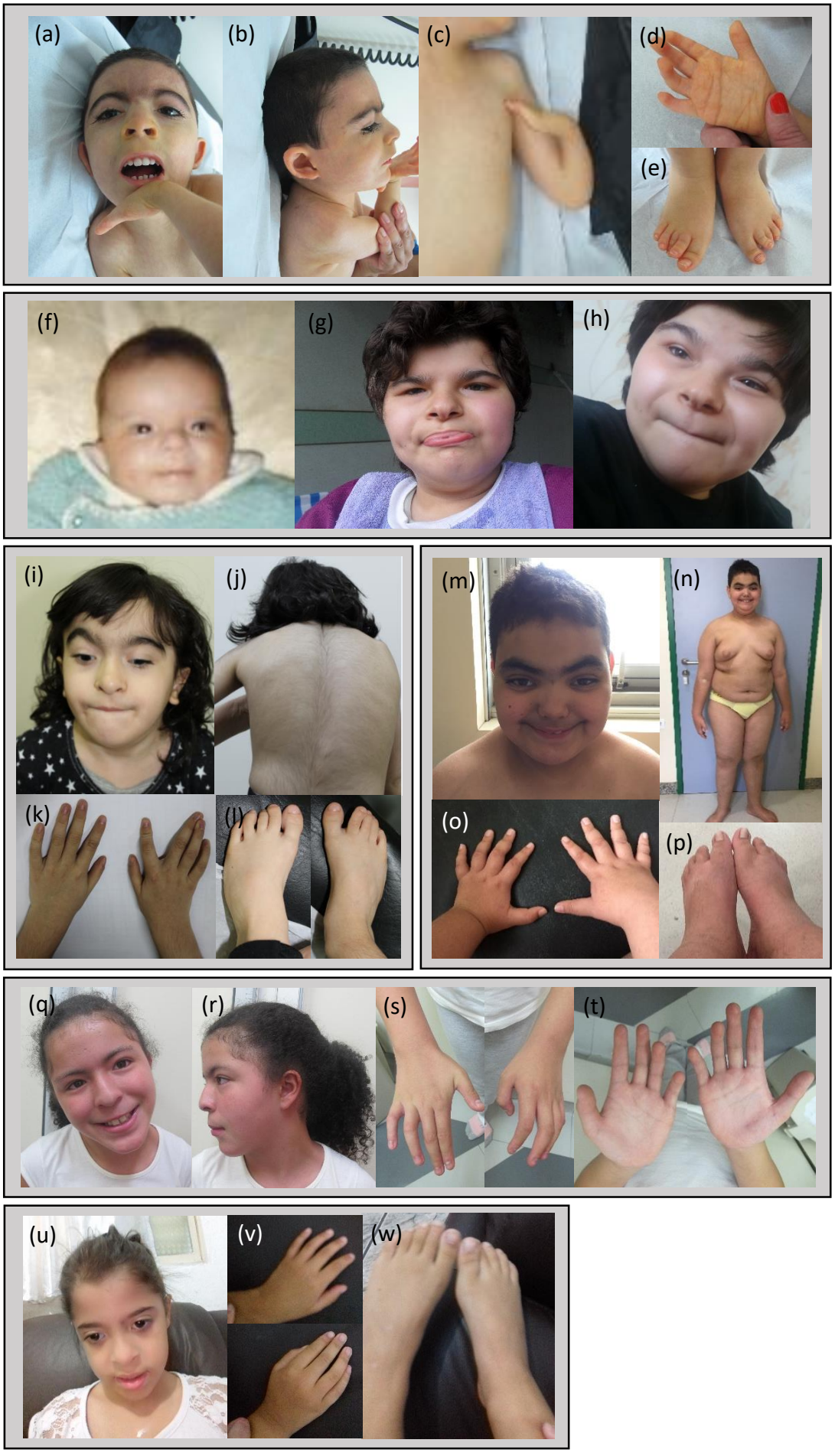
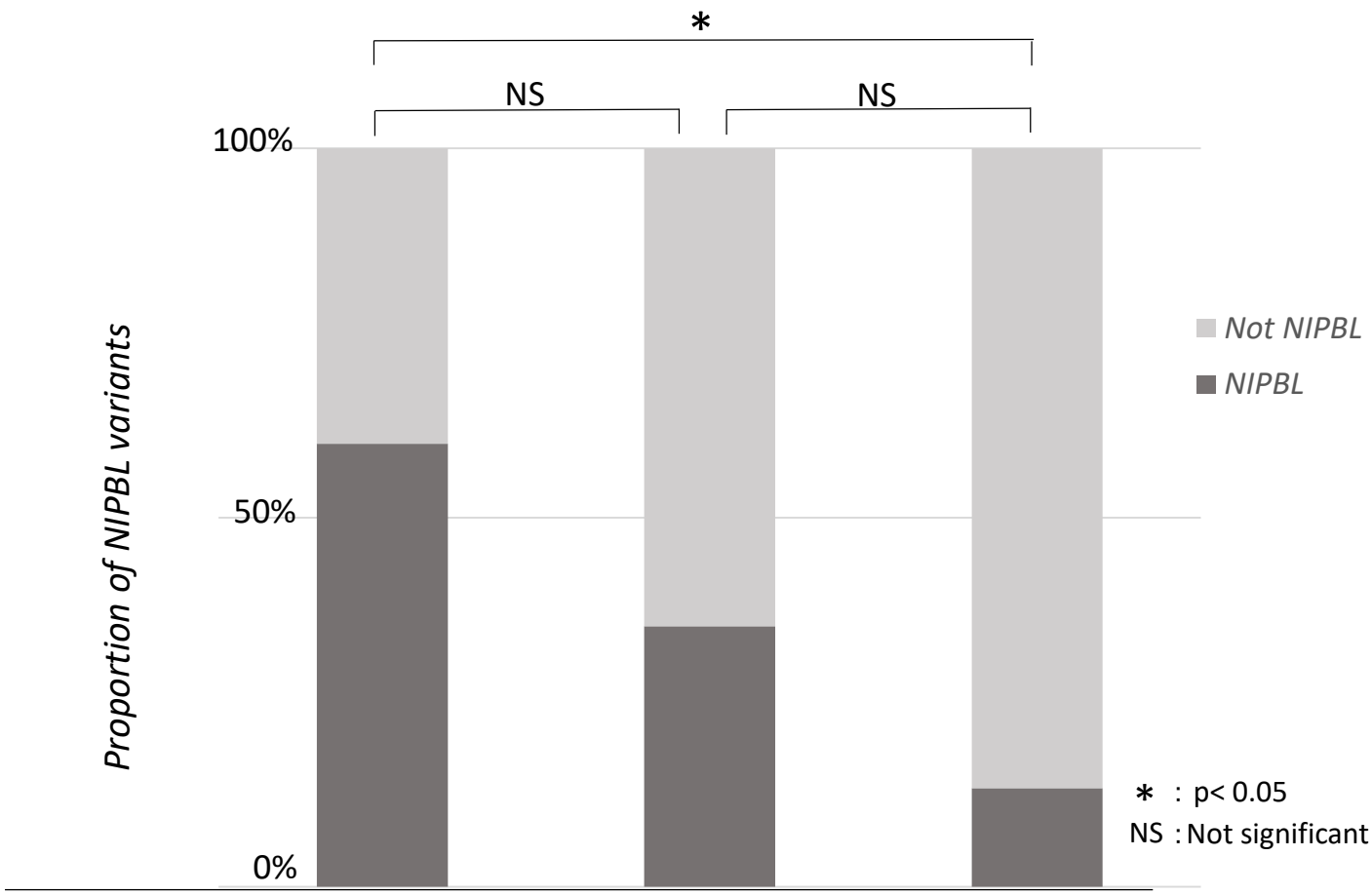
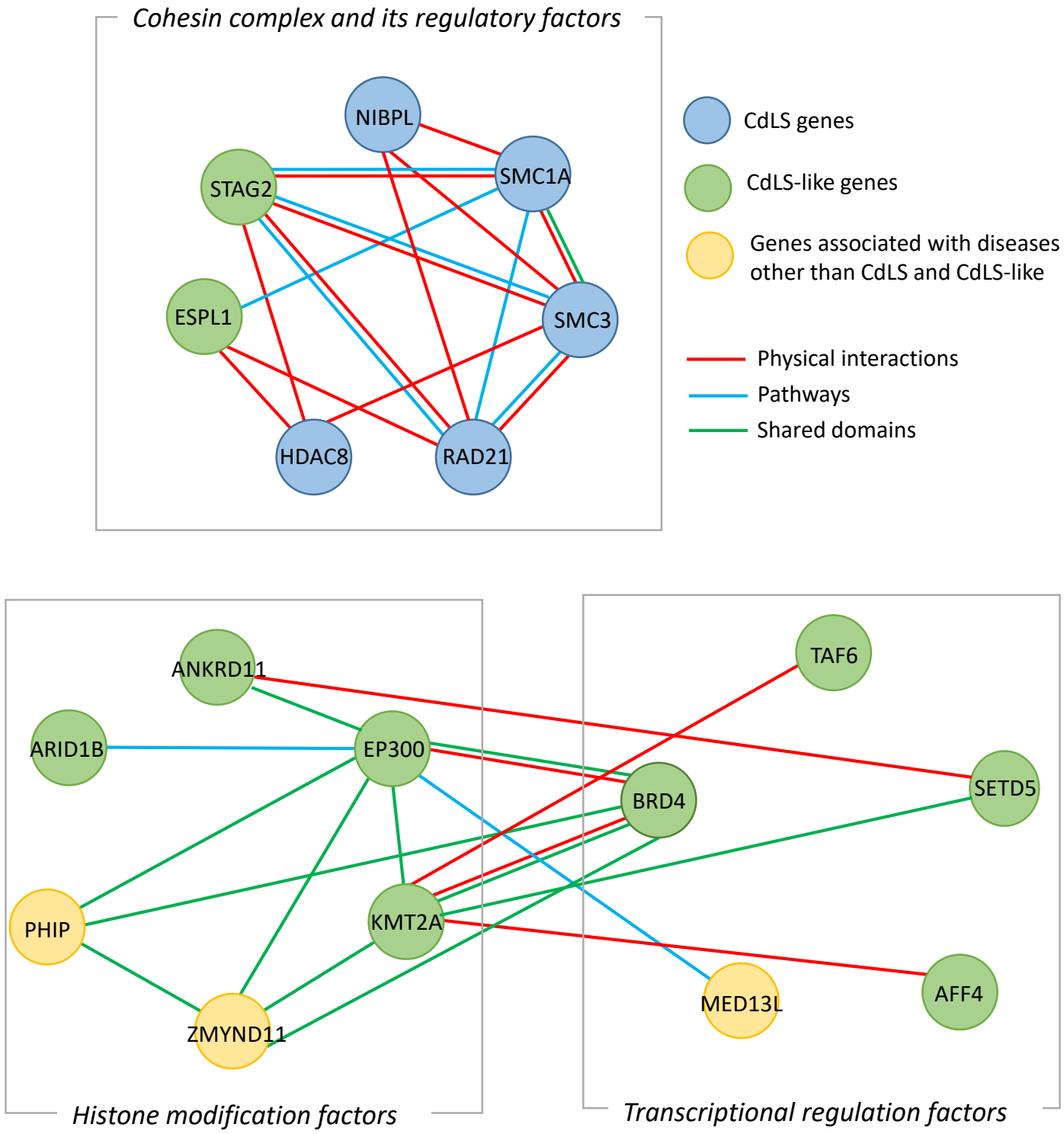


Figure 3



	Classic CdLS (Score ≧ 11)	Non-classic CdLS (Score 10 or 9)	Molecular test (Score 8-4)
NIPBL	14 (SNV) 1 (CNV)	6	2
SMC1A	1	0	1
ANKRD11	0	1	1
EP300	0	1	0
KMT2A	0	0	1
SETD5	0	1	0
ZMYND11	1	0	0
MED13L	0	1 (CNV)	1
PHIP	0	0	1
CNV	0	0	2
Undetermined	8	7	6
Total	25	17	15

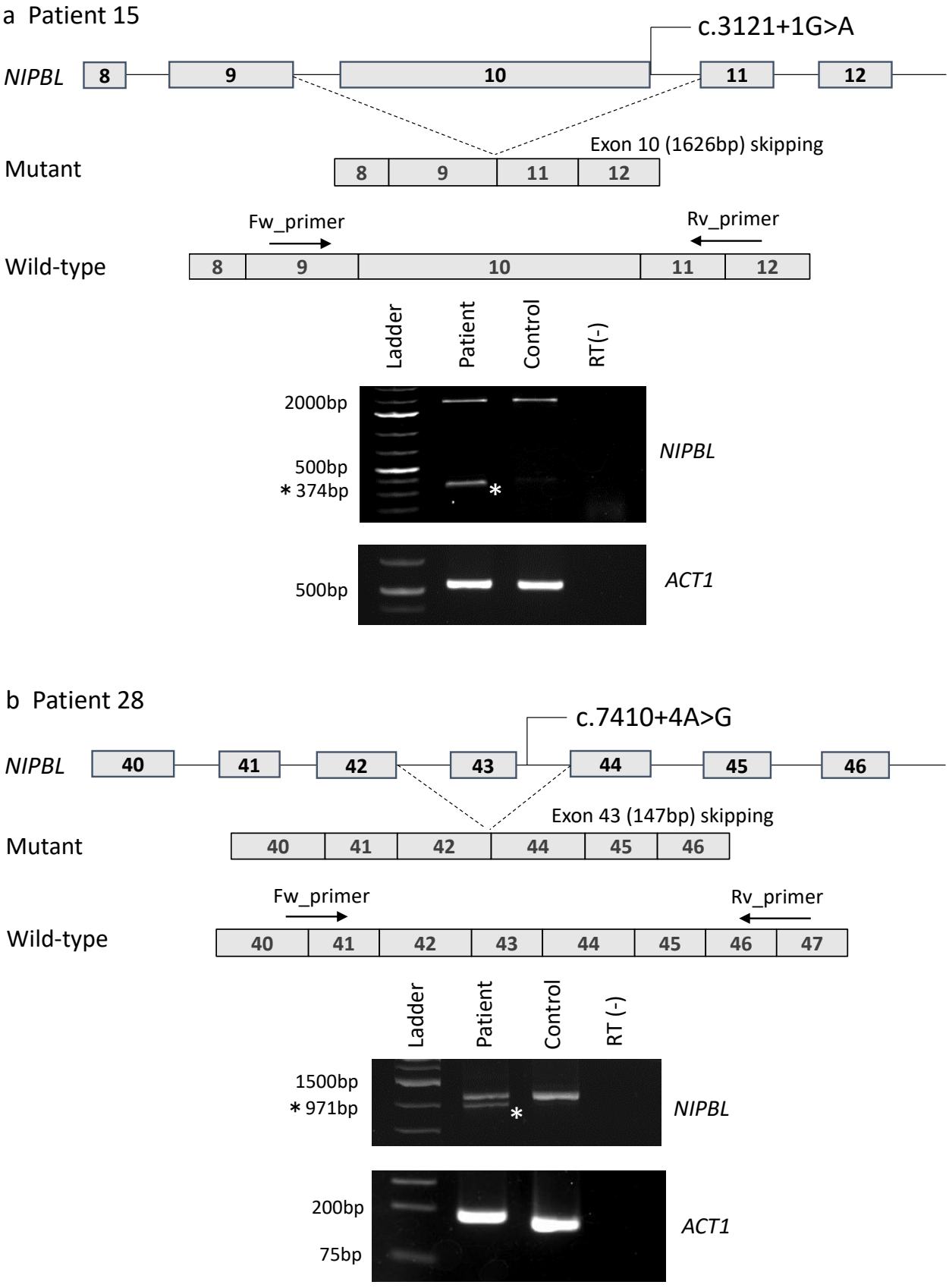
Figure 4


Cohesin complex and its regulatory factors
NIBPL
STAG2
SMC1A
ESPL1
SMC3
HDAC8
RAD21
CdLS genes
CdLS-like genes
Genes associated with diseases other than CdLS and CdLS-like
Physical interactions
Pathways
Shared domains
ANKRD11
ARID1B
EP300
BRD4
PHIP
KMT2A
ZMYND11
Histone modification factors
TAF6
SETD5
MED13L
AFF4
Transcriptional regulation factors

Gene	HGMD accession	Patient	Variant type	Mutation (hg 19)	Protein	Inheritance	Prediction scores			Control database			Novelty
							SIFT	Polyphen2	Mutation taster	ESP6500	ExAC	575 in-house	
NIPBL	NM_0133433.3	2	missense	c.6893G>A	p.Arg2298His	*de novo*	D	D	D	-	-	0	Reported [47]
		3	nonsense	c.6179dup	p.His2060Glnfs*4	*de novo*	-	-	-	-	-	0	Novel
		4	missense	c.7699T>G	p.Tyr2567Asp	*de novo*	D	D	D	-	-	0	Novel
		7	missense	c.5595G>T	p.Arg1865Ser	unavailable	D	D	D	-	-	0	Novel
		8	missense	c.6620T>C	p.Met2207Thr	*de novo*	D	D	D	-	-	0	Novel
		9	CNV	93.9-Kb deletion	-	not confirmed	-	-	D	-	-	-	Novel
		10	frameshift	c.5174delA	p.Lys1725Serfs*17	unavailable	-	-	D	-	-	0	Novel
		13	frameshift	c.2479_2480delAG	p.Arg827Glyfs*2	*de novo*	-	-	D	-	-	0	Reported [47]
		15	splicing	c.3121+1G>A	-	*de novo*	-	-	-	-	-	0	Reported [19]
		17	frameshift	c.1903_1904insA	p.Ser635Tyrfs*3	*de novo*	-	-	D	-	-	0	Novel
		25	frameshift	c.5030_5031del	p.Ile1677Serfs*21	*de novo*	-	-	D	-	-	0	Novel
		28	splicing	c.7410+4A>G	-	*de novo*	-	-	-	-	-	0	Reported [20]
		31	in-frame deletion	c.6653_6655del	p.Asn2218del	*de novo*	-	-	D	-	-	0	Novel
		36	nonsense	c.5509C>T	p.Arg1837*	*de novo*	-	-	D	-	-	0	Novel
		38	nonsense	c.826C>T	p.Gln276*	*de novo*	-	-	D	-	-	0	Reported [48]
		39	nonsense	c.190C>T	p.Gln64*	*de novo*	-	-	D	-	-	0	Reported [9]
		41	missense	c.6343G>T	p.Gln2115Cys	*de novo*	D	D	D	-	-	0	Novel
		45	missense	c.6027G>C	p.Leu2009Phe	*de novo*	D	D	D	-	-	0	Novel
		48	frameshift	c.8325_8326delinsT	p.Lys2775Asnfs*4	*de novo*	-	-	D	-	-	0	Novel
		49	missense	c.6448C>G	p.Leu2150Val	*de novo*	T	D	D	-	-	0	Novel
		50	missense	c.6893G>A	p.Arg2298His	*de novo*	D	D	D	-	-	0	Reported [47]
		54	splicing	c.5329-15A>G	-	*de novo*	-	-	-	-	-	0	Reported [21]
		55	missense	c.7079G>T	p.Gly2360Val	*de novo*	D	D	D	-	-	0	Novel
SMC1A	NM_006306.3	11	missense	c.1152C>G	p.Lys362Asn	*de novo*	D	D	D	-	-	0	Novel
		42	missense	c.1487G>A	p.Arg496His	*de novo*	D	D	D	-	-	0	Reported [49]
ANKRD11	NM_013275.5	21	frameshift	c.3255_3256del	p.Lys1086Glufs*15	*de novo*	-	-	D	-	-	0	Novel
		43	nonsense	c.5434C>T	p.Gln1812*	*de novo*	-	-	D	-	-	0	Novel
EP300	NM_001429.3	6	in-frame deletion	c.7014_7028del	p.His2338_Pro2342del	unavailable	-	-	P	-	-	0	Novel
KMT2A	NM_001197104.1	27	nonsense	c.3592C>T	p.Gln1198*	*de novo*	-	-	D	-	-	0	Novel
SETD5	NM_001080517.2	1	nonsense	c.1852C>T	p.Arg618*	*de novo*	-	-	-	-	-	0	Novel
MED13L	NM_015335.4	5	missense	c.6485C>A	p.Thr2162Lys	*de novo*	D	D	D	-	-	0	Novel
		34	CNV	4.2-Mb deletion	-	*de novo*	-	-	-	-	-	-	Novel
ZMYND11	NM_006624.5	53	frameshift	c.1438delG	p.Asp480Thrfs*3	maternal (mosaic)	-	-	D	-	-	0	Novel
PHIP	NM_017934.7	56	missense	c.1156G>A	p.Asp386Asn	*de novo*	D	D	D	-	-	0	Novel
-		51	CNV	9p 14.1-Mb del 571-kb dup	-	*de novo*	-	-	-	-	-	-	Novel
EHMT1	NM_24757.4	52	CNV	9q 773.8-kb del	-	*de novo*	-	-	-	-	-	-	Novel
NAA50	NM_025146.4	19	nonsense	c.93C>G	p.Tyr31*	*de novo*	-	-	-	-	-	-	Novel

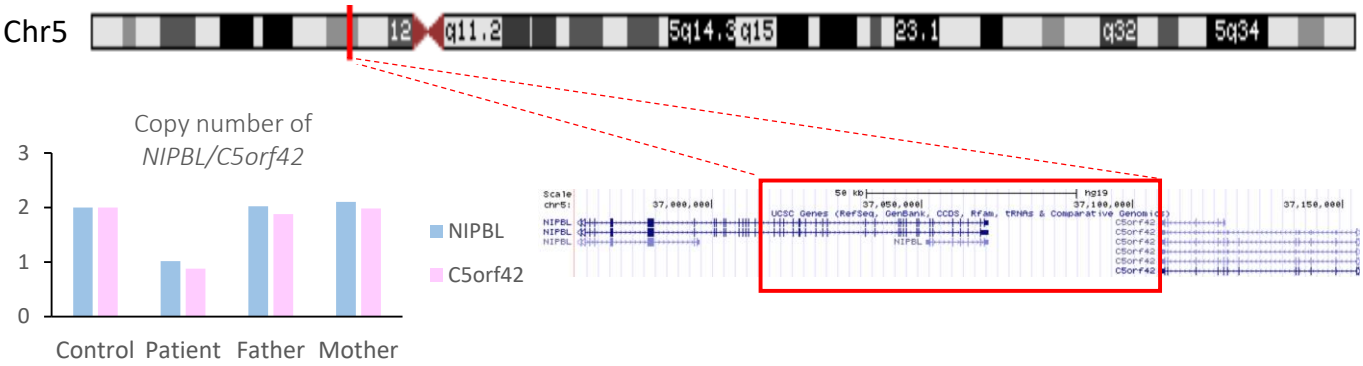
Table 1. Pathogenic variants were identified in this study.

SIFT, Sorting Intolerant From Tolerant (http://sift.bii.a-star.edu.sg/); PolyPhen-2, PolymorphismPhenotyping v2 (http://genetics.bwh.harvard.edu/pph2/); MutationTaster (http://www.mutationtaster.org/); ESP6500, National Heart, Lung, and Blood Institute (NHLBI) Exome Sequencing Project (ESP) Exome Variant Server (http://evs.gs.washington.edu/EVS/); ExAc browser (http://exac.broadinstitute.org/); 575 in-house, in-house 575 Japanese control exome dataset

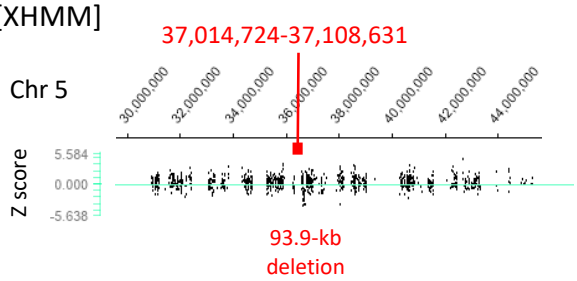


Supplementary Figure S1. Abnormal *NIPBL* transcripts generated by splice site variants. Abnormal transcripts were confirmed in (a) Patient 15 with c.3121+1G>A in *NIPBL* and (b) Patient 28 with c.7410+4A>G in *NIPBL*. The upper schematic shows a partial gene structure with the splicing variant. The middle schematic shows the mutant and wild-type cDNA. The lower image shows wild-type and mutant cDNA products in electrophoresed gels. The asterisk (*) indicates aberrant transcripts of 374 bp, generated by exon 10 skipping (a), and 971 bp, generated by exon 43 skipping (b). *ACT1* was used as the control.

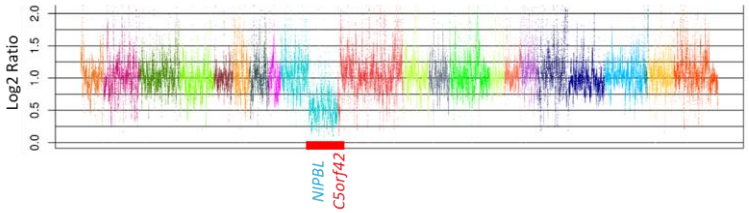
a Patient 9



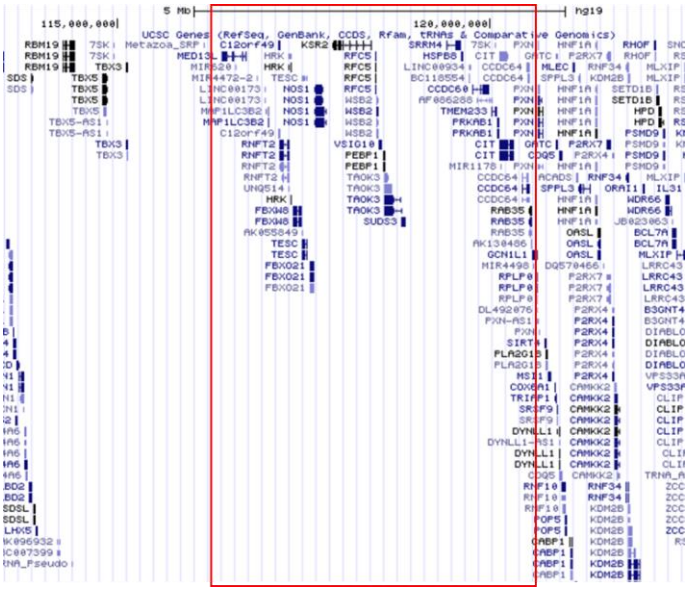
[XHMM]

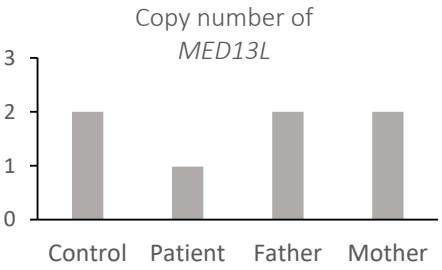


[Nord]

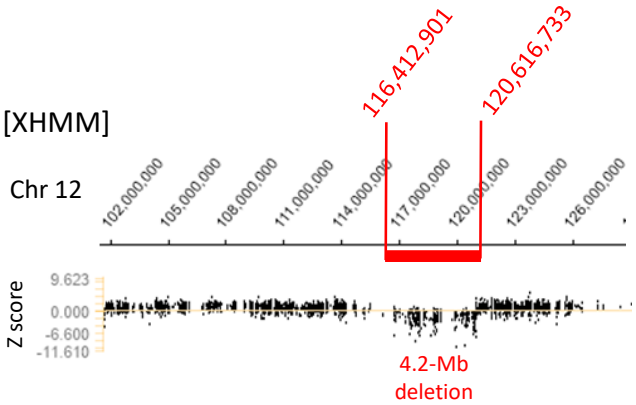


b Patient 34

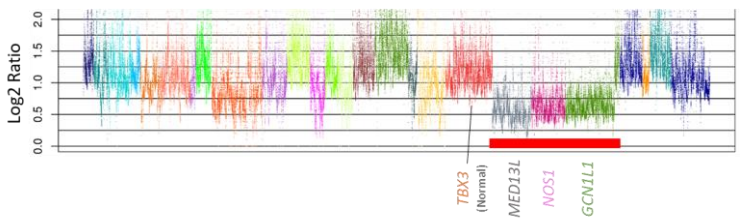




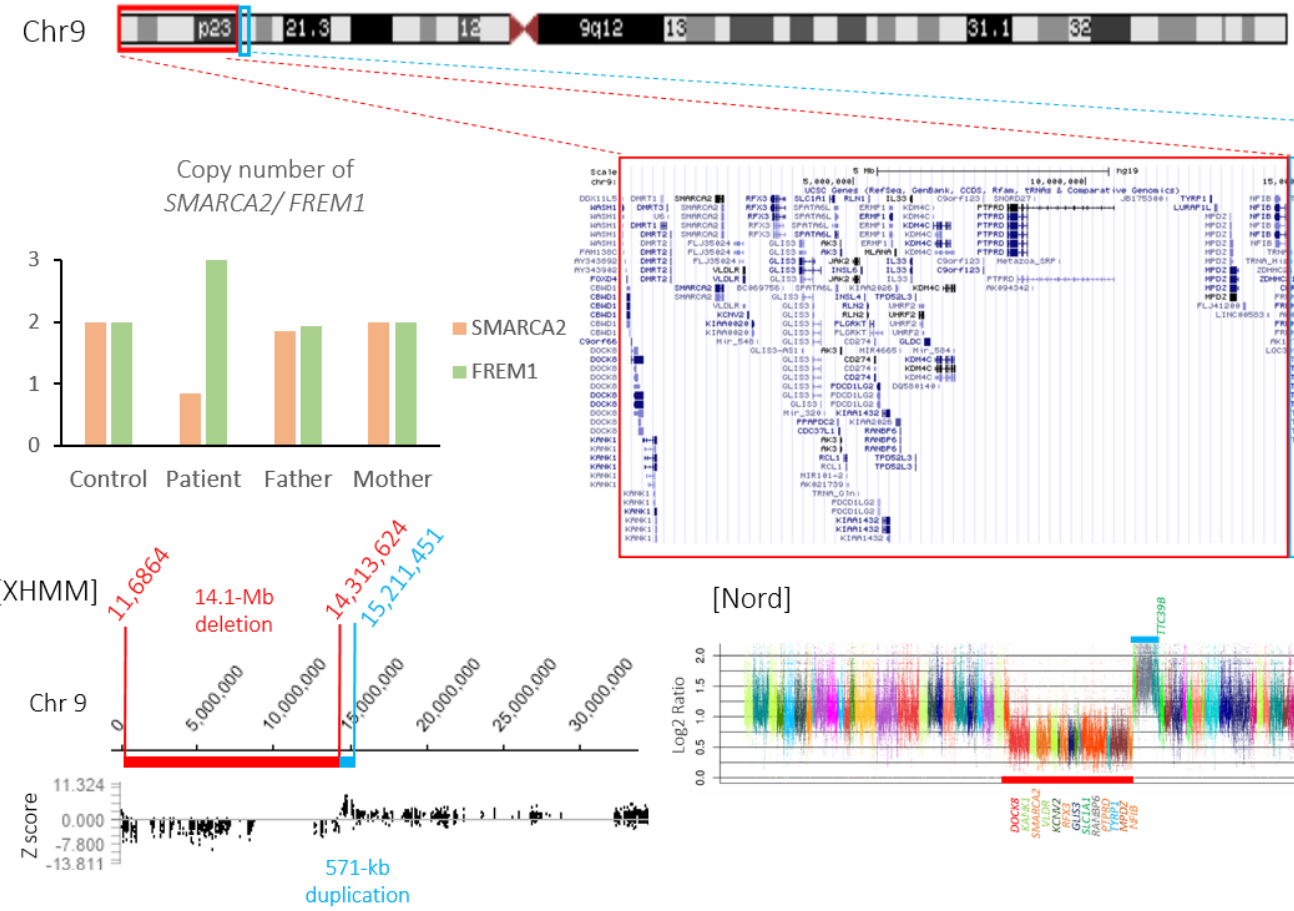
[XHMM]



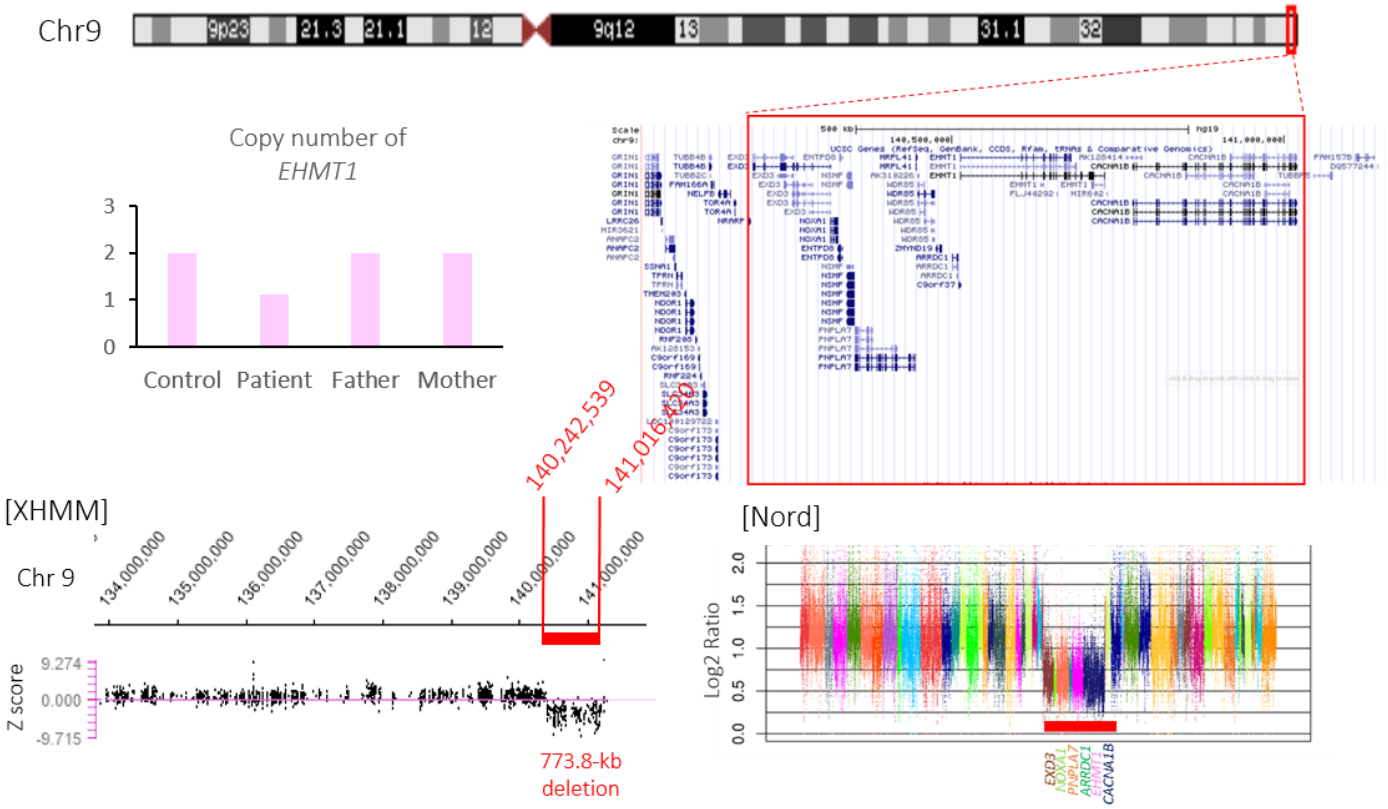
[Nord]



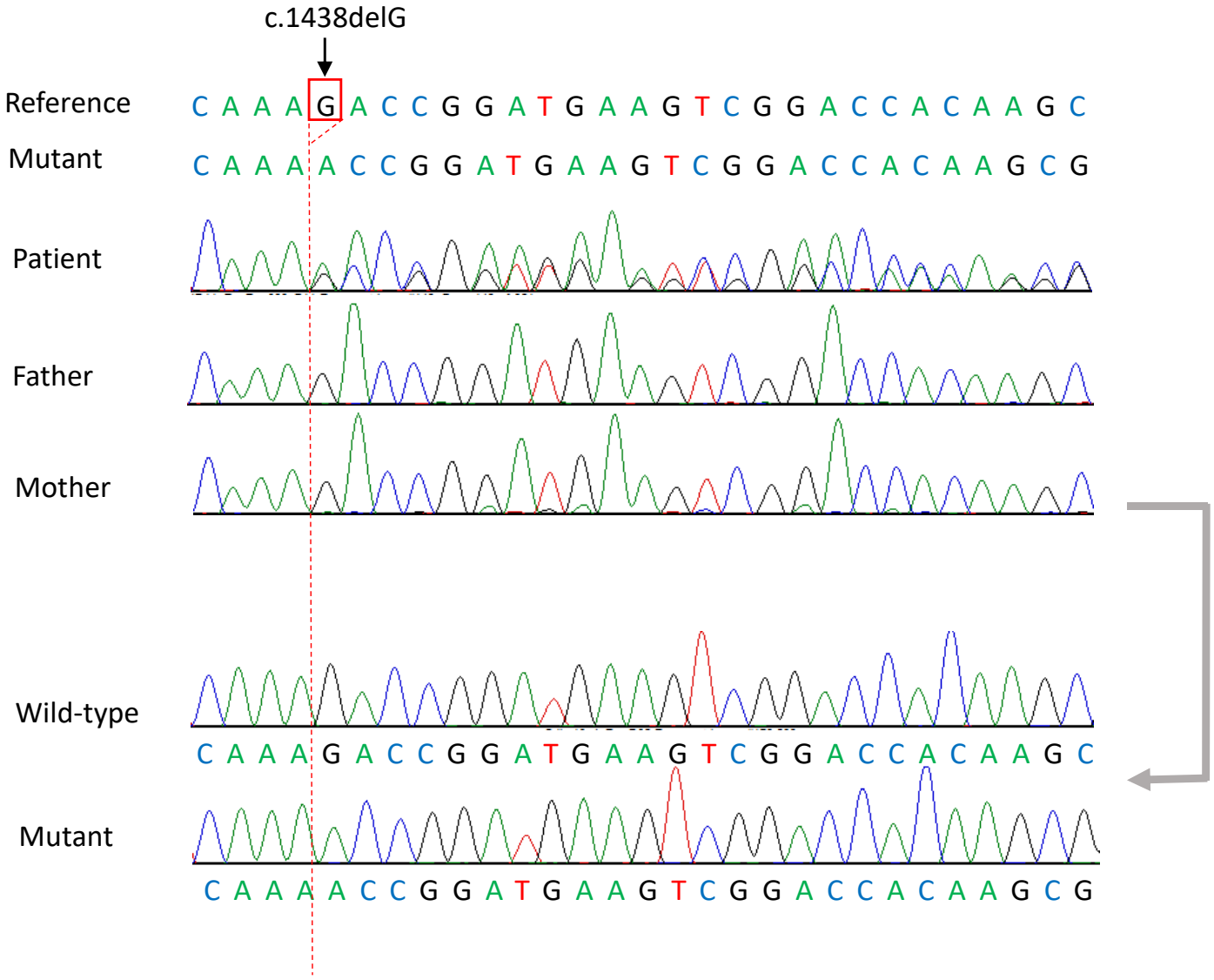
c Patient 51



d Patient 52



Supplementary Figure S2. Four patients with pathogenic CNVs detected using WES data. Red squares/bars indicates deletions, and blue squares/bars indicates duplications identified using XHMM and Nord programs. Copy number variations (CNVs) were confirmed by qPCR. (a) Patient 9 had a *de novo* 93.9-kb deletion at 5p13.2 corresponding to exons 22 to 47of *NIPBL* and the last exon of *C5orf42*. *NIPBL* and *C5orf 42* deletion were confirmed by qPCR. (b) Patient 34 possessed a *de novo* 4.2-Mb deletion involving 40 RefSeq genes including *MED13L* at 12q24.1-q24.23. *MED13L* deletion was confirmed by qPCR. (c) Patient 51 had a *de novo* 14.1-Mb deletion at 9p24.3-p22.3 involving 44 RefSeq genes adjacent to a 571-kb duplication at 9p22.3 encompassing four genes. *SMARCA2* deletion and *FREM1* duplication were confirmed by qPCR. (d) Patient 52 showed a *de novo* 773.8-kb deletion at 9q34.3 involving 14 RefSeq genes including *EHMT1*. *EHMT1* deletion was confirmed by qPCR.



Supplementary Figure S3. A frameshift *ZMYND11* variant in Patient 53 and his mosaic mother with low level mosaicism. Heterozygous c.1438delG was detected in electropherograms of Patient 53. Double sequences were detected in the maternal electropherogram.By cloning the PCR product in the mother, wild-type and mutant alleles were clearly recognized by Sanger sequencing.

Supplementary Table S1. Pathogenic variants and clinical features in all 57 patients. Please see the separate file. Pathogenic variants and clinical features of CdLS are summarized.

Patient	Sequenced base(bp)	Mean depth	Covered regions (%)		
			> 5 reads	> 10 reads	> 20 reads
1	2,717,491,258	81.19	98.1	97.6	95.9
2	2,345,160,622	70.06	98.1	97.5	94.9
3	2,410,547,560	72.02	98.1	97.5	95
4	2,193,943,963	65.55	98.1	97.3	94.3
5	2,134,650,052	63.77	98	97.2	93.4
6	2,278,525,905	68.07	98	97.3	94.7
7	3,327,143,568	99.4	98.3	98	96.9
8	2,353,682,334	70.32	98.1	97.5	95
9	1,956,621,164	58.46	97.9	97	92.7
10	2,467,364,092	73.71	97.9	97.2	94.1
11	2,163,878,827	64.65	97.9	96.9	92.7
12	2,386,979,976	71.31	97.9	97.2	94.2
13	2,709,790,390	80.96	98.1	97.5	95.2
14	2,271,433,984	67.86	97.9	97.1	93.6
15	2,497,750,149	74.62	98	97.3	94.5
16	2,401,689,900	71.75	98	97.3	94
17	3,068,833,794	91.68	98.2	97.7	96
18	2,811,745,691	84	98	97.5	95.4
19	2,244,406,012	67.05	98	97.2	93.3
20	2,515,145,707	75.14	98	97.3	94.8
21	2,304,367,691	68.84	98	97.4	94.5
22	2,071,779,476	61.9	98	97.2	93.3
23	2,077,593,792	62.07	98	97.1	93
24	2,584,558,871	77.22	98	97.5	95.3
25	2,214,779,528	66.17	97.9	96.9	92.2
26	3,250,938,863	97.12	98.1	97.7	96
27	2,588,522,835	77.33	98.1	97.4	94.8
28	2,617,702,774	78.21	97.9	97.1	94.4
29	3,294,551,016	98.43	98	97.5	95.8
30	2,217,645,104	66.25	97.8	96.1	88.2
31	2,037,671,095	60.88	97.8	96.5	90.7
32	2,627,629,431	78.5	98	97.3	94.3
33	1,924,415,550	57.49	97.9	96.8	91.2
34	2,743,248,459	81.96	97.8	96.7	92.2
35	3,250,384,892	97.11	98.1	97.7	96.1
36	2,232,039,428	66.68	98.1	97.3	93.5
37	2,722,699,721	81.34	98	97.5	95
38	2,480,268,680	74.1	98	97.3	94.4
39	3,425,858,008	102.35	98.2	97.7	96.2
40	2,127,989,922	63.57	98	97	92.4
41	2,140,157,321	63.94	98	96.9	92.1
42	2,970,702,586	88.75	98.1	97.6	95.8
43	3,052,462,950	91.19	98.1	97.6	95.6
44	2,956,719,611	88.33	98.1	97.6	95.7
45	2,708,263,123	80.91	98.1	97.5	95.2
46	2,760,250,251	82.46	98	97.5	95.5
47	2,438,746,246	72.86	98	97.3	94.3
48	2,902,940,061	86.73	98.1	97.6	95.6
49	3,442,610,124	102.85	98	97.5	95.5
50	3,625,730,620	108.32	98.1	97.6	95.8
51	2,395,445,229	71.57	98	97.3	94.5
52	2,221,757,172	66.38	97.9	97	93.3
53	2,226,712,912	66.52	98	97.1	93.4
54	2,934,068,828	87.66	98.3	97.9	95.4
55	2,759,645,091	82.45	98.3	97.8	94.9
56	3,109,737,458	92.91	98.1	97.6	95.6
57	2,900,351,453	86.65	98.1	97.4	94.5

Supplementary Table S2. WES performance. Whole exome sequencing (WES) was performed in all 57 patients with suspected CdLS. Mean read depth of protein coding regions ranged from 57.49× to 108.32×, with an average 94.3% of target bases sequenced by 20 or more reads.

		total	G (WT)	Deletion (Mut)	Other substitusions			
					A	C	T	N
Patient	read count (%)	27211	14629 (53.8)	12514 (46.0)	30 (0.1)	24 (0.1)	13 (0.0)	1 (0.0)
Mother	read count (%)	27596	24718 (89.6)	2835 (10.3)	20 (0.1)	14 (0.1)	9 (0.0)	0 (0.0)
Father	read count (%)	29548	29500 (99.8)	0	30 (0.1)	6 (0.0)	10 (0.0)	2 (0.0)
Control	read count (%)	45574	45507 (99.9)	1 (0.0)	32 (0.1)	13 (0.0)	19 (0.0)	2 (0.0)

Supplementary Table S3. Mutant read frequencies of *ZMYND11*. Deep sequencing indicated a mutant allele read frequency of 10.3% in the mother and 46.0% in Patient 53.

Sex		female	female
age		24y	6y
Clinical score		8	9
Classification		molecular test	non-classic
MED13L alteration		c.6485C>A, p.(Thr2162Lys)	4.2-Mb deletion
Inheritance		*de novo*	*de novo*
Dysmorphic features			
MED13L features	broad/prominent forehead	+	-
	bitemporal narrowing	+	+
	horizontal eyebrows	-	-
	upslanting palpebral fissures	-	+
	long down palpebral fissures	-	-
	enlargement of the palpebral fissures	-	-
	everted lower eyelids	-	-
	full cheeks	+	-
	bulbous nasal tip	-	+
	deep philtrum	-	-
	wide open mouth	-	-
	protruding tongue	-	-
	cupid-bow upper lip	-	-
CdLS features	microcephaly	+	-
	synophrys	+	+
	hyghly arched eyebrows	-	+
	long curly eyelashes	+	+
	anteverted nostrilis	+	+
	long philtrum	-	-
	downturned corners of the mouth	-	-
	short neck	-	+
Upper limb abnormalities		+	+
		5th finger clinodactyly	5th finger clinodactyly
Congenital heart defects		-	+
		bradycardia	ASD
Hypotonia		+	+
Age for independent walking		not able to walk	60 months
Speech delay		+	+
		only vocals ; "ma" "pa"	only a few sounds
Epilepsy		+	-
		from4-6y	
Developmental delay		+	+
Intellectual disability		+	+
Brain MRI findings		NA	NA
Others		recurrent otitis	-
Death or alive		death	alive

Abbreviations: ASD, atrial septal defects; NA, Non available

Supplementary Table S4. Comparison of clinical features of *MED13L* haploinsufficiency syndrome and CdLS in two patients with *MED13L* abnormality. Pathogenic variants were identified in *MED13L* in two patients: c.6485C>A, p.Thr2162Lys in Patient 5, and a 4.2-Mb deletion involving *MED13L* in Patient 34. Both patients shared several clinical features of *MED13L* haploinsufficiency syndrome as well as CdLS.

	Gene	*Gene-phenotype relationships [OMIM]*	*Function of proteins*
Cohesin complex and its regulation factors	*NIPBL*	Cornelia de Lange syndrome 1	cohesin loading to the genome
	SMC1A	Cornelia de Lange syndrome 2	cohesin ring components
	SMC3	Cornelia de Lange syndrome 3	cohesin ring components
	HDAC8	Cornelia de Lange syndrome 5	SMC3 is deacetylated to be reused the cohesin by follwing cycle
	RAD21	Cornelia de Lange syndrome 4	cohesin ring components
	STAG2	Neurodevelopmental disorder, X-linked, with craniofacial abnormalities	cohesin ring components
	ESPL1	-	cleave RAD21 of the centromeric cohesin and open the cohesin ring
Chromatin modification factors	*ANKRD11*	KBG syndrome	transcription inhibition by interacting with histone deacetylases (HDACs) and histone molecules
	ARID1B	Coffin-Siris syndrome 1	components of BAF complex, bind close to transcriptional start sites and responsible for chromatine remodeling
	EP300	Rubinstein-Taybi syndrome 2	encoding acetyltransferase to mark H3K18 and H3K27 acetylation
	KMT2A	Wiedemann-Steiner syndrome	encoding methiltransferase to mark H3K4 methylation
	ZMYND11	Mental retardation, autosomal dominant 30	the SWI/SNF chromatin-remodeling complex to reset the chromatin to a repressive state to prevent further transcription
	PHIP	Developmental delay, intellectual disability, obesity, and dysmorphic features	H3K4 methylation-binding protein
Transcriptional regulation factors	*AFF4*	CHOPS syndrome	components of the super elongeation complex (SEC), RNAP2 pausing release
	BRD4	-	binding super-enhansers with NIPBL and co-regulate developmental gene expression
	MED13L	Mental retardation and distinctive facial features with or without cardiac defects	crucial link between transcription factors,coactivators, and the main mediator complex
	SETD5	Mental retardation, autosomal dominant 23	interacts with the PAF1 co-transcriptional complex
	TAF6	Alazami-Yuan syndrome	components of TFIID, binding promoter with RANP2 and transcriptional intiation

Supplementary Table S5. Function of genes associated with CdLS and CdLS-like. A total of 18 genes were classified to three functional categories: cohesin complex and its regulation, chromatin modification, and transcriptional regulation. Three genes (*ZMYND11*, *MED13L*, and *PHIP*) associated with diseases other than CdLS and CdLS-like were categorized to chromatin modification factors or transcriptional regulation factors. ZMYND11 was categorized to chromatin modification factors, of which the SWI/SNF chromatin-remodeling complex resets chromatin to a repressive state to prevent further transcription. MED13L was categorized to transcriptional regulation factors showing a crucial link to EP300 (as a histone modification factor). PHIP was categorized to chromatin modification factors (as an H3K4 methylation-binding protein).

Supplementary Table S1. Pathogenic variants and clinical features in all 57 patients.

Patient	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Sex	female	male	male	male	female	female	male	male	female	female	female	female	male	female
Age (at point of study entry)	59y	8y	1y 1m	3y 10m	24y	7827	21y	17y	4y	8y	2y 6m	5y 9m	12y	3y
Clinical score	10	13	14	12	8	9	11	11	12	12	4	13	9	9
Classification	non-classic	classic	classic	classic	molecular testing	non-classic	classic	classic	classic	classic	molecular testing	classic	non-classic	non-classic
Gene	*SETD5*	*NIPBL*	*NIPBL*	*NIPBL*	*MED13L*	*EP300*	*NIPBL*	*NIPBL*	*NIPBL*	*NIPBL*	*SMC1A*	undetermined	*NIPBL*	undetermined
Variant type	nonsense	missense	nonsense	missense	missense	frameshift	missense	missense	CNV	frameshift	missense	-	frameshift	-
Mutation (hg19)	c.1852C>T	c.6893G>A	c.6179dup	c.7699T>G	c.6485C>A	c.7014_7028del	c.5595G>T	c.6620T>C	93.9-kb deletion	c.5174delA	c.1152C>G	-	c.2479_2480del	-
Protein	p.Arg618*	p.Arg2298His	p.His2060Glnfs*4	p.Tyr2567Asp	p.Thr2162Lys	p.His2338_Pro2342del	p.Arg1865Ser	p.Met2207Thr	-	p.Lys1725Serfs*17	p.Lys362Asn	-	p.Arg827Glyfs*2	-
Dysmorphic features														
synophrys	+	+	+	-	+	+	+	+	+	+	-	+	+	+
highly arched eyebrows	+	+	+	+	-	+	+	+	-	+	-	+	+	+
long curly eyelashes	+	+	+	+	+	+	+	+	+	+	-	+	+	+
ptosis	-	-	-	-	-	-	-	-	+	-	+	-	-	-
cleft lip	-	-	-	-	-	-	-	-	-	-	-	-	-	-
cleft palate	-	-	-	-	-	-	-	-	-	-	-	-	+	-
microcephaly	-	+	+	+	+	+	-	+	+	+	-	+	+	+
anteverted nostrils	+	+	+	+	+	+	-	+	-	+	-	+	+	+
depressed nasal bridge	+	-	+	-	-	-	+	-	+	-	-	-	+	+
long philtrum	+	+	+	+	-	-	+	+	-	+	-	+	-	+
thin upper lip	+	+	+	+	-	+	-	-	+	+	-	+	+	-
downturned corners of the mouth	+	+	+	+	-	+	-	-	-	+	-	+	+	-
micrognathia	-	-	-	+	+	-	-	-	-	+	-	+	-	+
short neck	+	+	-	-	-	+	+	+	+	+	-	+	-	+
high palate	-	-	-	-	-	+	+	-	+	-	-	+	-	-
widely spaced or absent teeth	-	-	-	-	NA	-	-	-	-	-	-	-	-	+
others					low set ears									
Growth														
weight below 5th percentile for age	+	-	+	+	-	-	-	-	+	+	+	+	-	-
height or length below 5th percentile for age	-	+	+	+	+	+	+	+	+	+	-	+	-	-
prenatal growth retardation	-	+	+	+	-	NA	+	-	+	+	+	+	+	-
others														
Development														
intellectual disability	+	-	NA	+	+	+	+	-	+	+	-	+	+	+
developmental delay or mental retardation	+	-	+	+	+	+	+	+	+	+	+	+	+	+
Behavior														
attention deficit disorder	+	+	NA	+	+	+	-	+	NA	+	+	+	+	+
anxiety	+	+	NA	+	+	+	+	+	NA	+	+	+	-	+
aggression	-	-	NA	+	+	+	+	+	NA	-	-	-	-	+
self-injurious behavior	-	+	NA	-	-	+	+	-	NA	+	+	+	+	+
autistic behavior	+	-	NA	-	-	+	+	-	NA	+	-	-	-	+
Limb abnormalities														
absence of forearms	-	-	-	-	NA	-	-	-	+	-	-	-	-	-
small hands and/or feet	-	-	-	-	NA	-	-	-	-	+	-	-	+	-
oligodactyly	-	-	-	-	NA	-	-	-	-	+	-	-	+	-
5th finger clinodactyly	-	+	+	-	+	+	+	+	-	-	-	-	-	-
abnormal palmar crease	-	+	+	-	NA	+	+	+	-	+	-	-	+	-
syndactyly	-	-	+	+	NA	+	-	+	-	+	-	+	-	-
phocomelia	-	-	-	-	NA	-	-	-	-	-	-	-	-	-
limited elbow extension	-	-	-	-	NA	-	+	-	-	-	-	-	-	-
proximally placed thumbs	-	-	+	-	NA	+	+	-	-	-	-	-	-	-
others					club feet			unilateral absence of hand						

Neurosensory–Skin														
ptosis	-	-	-	-	-	-	-	-	-	-	+	-	+	-
myopia	-	-	NA	-	+	+	-	-	-	+	+	-	-	+
deafness or hearing loss	-	-	NA	-	-	+	-	-	+	+	-	-	+	-
seizures	-	-	-	+	+	-	-	+	+	+	-	-	+	+
hirsutism, generalized	-	+	+	-	+	+	+	+	-	+	+	+	+	+
others									-					
Genitourinary														
cryptorchidism	-	+	+	-	-	-	+	-	-	-	-	-	+	-
hypoplastic (small) genitalia	-	-	+	+	-	-	-	-	-	-	-	-	+	+
renal abnormalities	-	-	-	-	-	-	-	-	-	-	-	-	-	-
others														
Cardiovascular														
ventricular septal defects	+	-	-	-	-	NA	-	-	-	-	-	-	-	-
atrial septal defects	-	-	-	-	-	NA	-	-	-	-	-	-	-	-
pulmonic ste-sis	-	-	+	-	-	NA	-	+	-	-	-	-	-	-
tetralogy of Fallot	-	-	-	-	-	NA	-	-	+	-	-	-	-	-
hypoplastic left heart	-	-	-	-	-	NA	-	-	-	-	-	-	-	-
bicuspid aortic valve	-	-	-	-	-	NA	-	-	-	-	-	-	-	-
others														
Others														
language delay	+	-	+	-	+	+	+	-	+	-	+	+	+	+
diaphragmatic hernia	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gastroesophageal reflux	-	+	-	+	+	-	+	+	+	+	+	+	+	+
others														

Abbreviations: y, years; m, months; NA, not available

Patient	15	16	17	18	19	20	21	22	23	24	25	26	27	28
Sex	female	male	male	female	male	female	female	male	male	female	male	male	male	female
Age (at point of study entry)	24y	26y	2y	3y	4y	10y	3y 10m	25y	18y	6y	16y 2m	9y	41y	40y
Clinical score	11	4	12	8	12	6	8	7	12	10	13	9	7	11
Classification	classic	molecular testing	classic	molecular testing	classic	molecular testing	molecular testing	molecular testing	classic	non-classic	classic	non-classic	molecular testing	classic
Gene	*NIPBL*	undetermined	*NIPBL*	undetermined	*NAA50*	undetermined	*ANKRD11*	undetermined	undetermined	undetermined	*NIPBL*	undetermined	*KMT2A*	*NIPBL*
Variant type	splicing	-	frameshift	-	nonsense	nonsense	frameshift	-	-	-	frameshift	-	nonsense	splicing
Mutation (hg19)	c.3121+1G>A	-	c.1903_1904insA	-	c.93C>G	-	c.3255_3256del	-	-	-	c.5030_5031del	-	c.3592C>T	c.7410+4A>G
Protein	-	-	p.Ser635Tyrfs*3	-	p.Tyr31*	-	p.Lys1086Glufs*15	-	-	-	p.Ile1677Serfs*21	-	p.Gln1198*	-
Dysmorphic features														
synophrys	+	+	+	+	+	+	+	+	+	+	+	+	+	+
highly arched eyebrows	+	-	+	+	+	-	-	-	+	+	+	+	+	+
long curly eyelashes	+	-	+	+	+	-	-	+	+	+	+	+	+	-
ptosis	-	-	-	+	-	-	-	-	-	-	-	-	-	-
cleft lip	-	-	-	-	-	-	-	-	-	-	-	-	-	-
cleft palate	-	+	-	-	+	-	-	-	-	-	-	-	-	-
microcephaly	+	-	+	+	+	-	+	+	+	-	+	-	-	+
anteverted nostrils	+	-	+	-	+	-	+	-	+	+	+	+	-	-
depressed nasal bridge	-	-	-	+	+	-	+	-	-	+	+	+	-	-
long philtrum	-	-	+	-	+	-	-	+	+	+	+	+	-	+
thin upper lip	+	-	+	-	+	+	+	-	+	+	+	+	+	+
downturned corners of the mouth	+	-	+	-	+	+	+	-	+	-	+	+	+	+
micrognathia	-	-	-	-	-	-	-	-	+	-	-	-	-	-
short neck	+	-	+	-	-	+	-	-	-	+	-	+	+	+
high palate	+	-	-	-	-	-	-	-	+	-	-	-	-	-
widely spaced or absent teeth	-	-	-	+	-	-	-	+	+	-	-	+	-	-
others														
Growth														
weight below 5th percentile for age	+	-	-	+	+	-	-	+	+	+	+	-	-	-
height or length below 5th percentile for age	+	-	-	+	+	-	-	+	+	+	+	-	+	+
prenatal growth retardation	+	-	+	-	+	-	-	-	+	-	+	-	+	+
others														
Development														
intellectual disability	+	+	-	+	-	+	-	+	-	+	+	-	+	-
developmental delay or mental retardation	+	+	+	+	+	+	+	+	+	+	+	+	+	-
Behavior														
attention deficit disorder	+	+	-	+	+	+	+	-	-	+	+	+	+	+
anxiety	+	+	+	-	+	+	-	-	-	+	+	+	+	-
aggression	+	+	-	-	+	+	-	-	-	+	+	-	-	+
self-injurious behavior	+	+	-	+	+	+	-	+	-	+	+	-	-	-
autistic behavior	+	+	-	-	-	-	-	+	-	+	-	-	-	-
Limb abnormalities														
absence of forearms	-	-	-	-	-	-	-	-	-	-	-	-	-	-
small hands and/or feet	-	-	+	-	+	+	-	-	+	-	+	-	-	+
oligodactyly	-	-	-	-	+	-	-	-	-	-	+	-	-	-
5th finger clinodactyly	+	-	-	-	-	-	-	-	-	+	-	-	-	+
abnormal palmar crease	-	-	+	+	-	-	-	-	+	-	+	-	-	-
syndactyly	-	+	+	-	+	-	-	-	-	-	-	-	-	+
phocomelia	-	-	-	-	-	-	-	-	-	-	-	-	-	-
limited elbow extension	-	+	+	-	-	-	-	-	+	-	-	-	-	+
proximally placed thumbs	+	-	+	-	-	-	-	-	-	-	-	-	+	-
others														

Neurosensory–Skin														
ptosis	-	-	-	+	-	-	-	-	-	-	-	-	-	-
myopia	-	+	+	+	+	+	-	-	+	+	+	+	-	-
deafness or hearing loss	+	-	+	-	-	-	+	+	-	-	-	+	-	-
seizures	-	+	+	-	-	+	-	-	-	+	+	-	-	-
hirsutism, generalized	-	+	+	+	-	+	-	-	-	+	+	-	+	-
others														
Genitourinary														
cryptorchidism	-	-	+	-	+	-	-	-	-	-	+	-	-	-
hypoplastic (small) genitalia	-	-	+	-	+	-	-	-	-	-	+	-	-	-
renal abnormalities	-	-	-	-	-	-	-	-	+	-	-	-	-	-
others													hypospadias	
Cardiovascular														
ventricular septal defects	-	-	-	-	-	-	-	-	-	-	+	-	-	-
atrial septal defects	-	-	-	-	-	-	-	-	-	-	+	-	-	-
pulmonic ste-sis	-	-	-	-	-	-	-	-	-	-	-	-	-	-
tetralogy of Fallot	-	-	-	-	-	-	-	-	-	-	-	-	-	-
hypoplastic left heart	-	-	-	-	-	-	-	-	-	-	-	-	-	-
bicuspid aortic valve	-	-	-	+	-	-	-	-	-	-	-	-	-	-
others									aneurism between atrial septal					
Others														
language delay	+	+	+	+	+	+	+	+	-	+	+	-	+	-
diaphragmatic hernia	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gastroesophageal reflux	+	+	+	+	+	-	+	+	+	+	+	+	+	-
others														

Patient	29	30	31	32	33	34	35	36	37	38	39	40	41	42
Sex	female	male	male	male	male	female	female	male	female	female	male	male	male	female
Age (at point of study entry)	25y	43y	1y 1m	16y	13y	6y	9y	26y	3y	14y	3y 9m	5y 10m	NA	15y
Clinical score	9	12	13	7	10	9	9	13	11	13	11	12	10	13
Classification	non-classic	classic	classic	molecular testing	non-classic	non-classic	non-classic	classic	classic	classic	classic	classic	non-classic	classic
Gene	undetermined	undetermined	*NIPBL*	undetermined	undetermined	*MED13L*	undetermined	*NIPBL*	undetermined	*NIPBL*	*NIPBL*	undetermined	*NIPBL*	*SMC1A*
Variant type	-	-	frameshift	-	-	CNV	-	nonsense	-	nonsense	nonsense	-	missense	missense
Mutation (hg19)	-	-	c.6653_6655del	-	-	4.2-Mb deletion	-	c.5509C>T	-	c.826C>T	c.190C>T	-	c.6343G>T	c.1487G>A
Protein	-	-	p.Asn2218del	-	-	-	-	p.Arg1837*	-	p.Gln276*	p.Gln64*	-	p.Gln2115Cys	p.Arg496His
Dysmorphic features														
synophrys	+	+	+	+	+	+	+	+	+	+	+	+	+	+
highly arched eyebrows	+	+	+	+	+	+	+	+	+	+	+	+	+	-
long curly eyelashes	-	+	+	-	-	+	-	+	+	+	+	+	+	+
ptosis	-	-	-	-	-	-	-	-	-	-	-	-	-	-
cleft lip	-	-	-	-	-	-	-	-	-	-	-	-	-	-
cleft palate	-	-	-	-	-	-	+	-	-	-	-	+	-	-
microcephaly	+	-	+	-	+	-	-	+	+	+	+	+	-	+
anteverted nostrils	-	-	+	-	-	+	-	+	-	+	+	+	+	-
depressed nasal bridge	-	+	+	-	-	-	+	+	-	+	-	-	-	+
long philtrum	-	+	+	-	+	-	-	+	+	+	+	+	-	+
thin upper lip	+	+	+	+	+	-	-	+	+	+	-	+	+	+
downturned corners of the mouth	+	+	+	+	+	-	-	+	+	+	-	+	-	+
micrognathia	-	-	+	-	+	-	-	+	-	+	-	-	-	-
short neck	+	+	-	+	-	+	-	+	+	+	+	-	-	-
high palate	-	+	-	+	-	-	+	-	-	-	+	-	+	-
widely spaced or absent teeth	-	-	-	-	-	-	+	-	-	+	-	-	-	-
others						low set ears								
Growth														
weight below 5th percentile for age	+	-	+	+	+	-	+	+	+	+	+	+	+	+
height or length below 5th percentile for age	+	-	-	+	+	+	+	+	+	+	+	+	+	+
prenatal growth retardation	+	+	+	-	-	+	+	+	+	+	+	-	+	-
others														
Development														
intellectual disability	+	+	NA	+	+		+	+	+	+	+	+	+	+
developmental delay or mental retardation	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Behavior														
attention deficit disorder	+	+	NA	+	+	+	-	+	+	+	+	+	-	+
anxiety	+	+	NA	+	-	+	-	+	+	+	-	+	-	+
aggression	-	-	NA	-	-	+	+	+	+	-	+	-	-	-
self-injurious behavior	+	-	NA	-	+	-	+	+	-	+	+	-	-	+
autistic behavior	-	+	NA	+	+	-	-	+	-	-	+	-	-	-
Limb abnormalities														
absence of forearms	-	-	-	-	-	-	-	-	-	-	-	-	-	-
small hands and/or feet	-	-	-	-	-	-	-	+	-	-	-	-	-	-
oligodactyly	+	-	-	-	-	-	+	-	-	-	-	-	-	-
5th finger clinodactyly	-	+	+	-	+	+	-	-	-	-	-	-	-	+
abnormal palmar crease	-	-	+	+	-	+	+	+	-	+	-	-	-	-
syndactyly	-	+	-	-	-	-	+	+	-	+	-	-	+	-
phocomelia	-	-	-	-	-	-	-	-	-	-	-	-	-	-
limited elbow extension	+	+	-	-	+	-	-	+	-	+	-	-	-	+
proximally placed thumbs	-	-	+	-	+	-	+	+	-	+	-	-	-	-
others		asymmetry									camptodactyly			

Neurosensory–Skin														
ptosis	-	-	-	-	-	-	-	-	-	-	-	-	-	-
myopia	-	-	-	+	-	-	-	-	-	-	-	-	-	-
deafness or hearing loss	-	-	-	+	-	-	-	-	+	-	+	-	-	+
seizures	+	-	+	+	-	-	-	-	+	-	-	+	-	-
hirsutism, generalized	+	+	-	+	-	+	-	+	+	+	+	-	+	+
others							coloboma		white lesions	lacrimal glands				
Genitourinary														
cryptorchidism	-	-	+	+	-	-	-	+	-	-	-	+	-	-
hypoplastic (small) genitalia	-	+	+	-	-	-	-	+	+	-	-	-	-	+
renal abnormalities	-	+	-	+		-	-	-	-	+	-	-	-	-
others					hydrocele									
Cardiovascular														
ventricular septal defects	-	-	-	-	-	-	-	+	-	-	-	-	-	-
atrial septal defects	-	-	-	-	-	+	-	+	+	-	-	+	-	-
pulmonic ste-sis	-	-	-	-	-	-	-	+	-	-	+	-	-	-
tetralogy of Fallot	-	-	-	-	-	-	-	-	-	-	-	-	-	-
hypoplastic left heart	-	-	-	-	-	-	-	-	-	-	-	-	-	-
bicuspid aortic valve	-	-	-	-	-	-	-	-	-	-	-	-	-	-
others														
Others														
language delay	+	-	+	+	+	-	+	+	+	+	+	-	+	+
diaphragmatic hernia	-	-	-	+	-	-	-	-	-	-	-	-	-	-
gastroesophageal reflux	+	+	-	-	+	+	+	+	+	+	-	+	+	+
others					scoliosis					small cerebellum				scoliosis

Patient	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57
Sex	male	male	male	female	female	male	female	female	female	female	male	male	female	male	male
Age (at point of study entry)	11y	4y	8m	6m	16y	18y	8y	11y	8y	5y	5y 10m	7y	4y 6m	11y	6y
Clinical score	9	14	10	14	5	9	10	6	7	5	15	6	10	6	9
Classification	non-classic	classic	non-classic	classic	nolecular testin	non-classic	non-classic	molecular testing	molecular testing	molecular testing	classic	molecular testing	non-classic	molecular testing	non-classic
Gene	*ANKRD11*	undetermined	*NIPBL*	undetermined	undetermined	*NIPBL*	*NIPBL*	*NIPBL*	-	*EHMT1*	*ZMYND11*	*NIPBL*	*NIPBL*	*PHIP*	undetermined
Variant type	nonsense	-	missense	-	-	frameshift	missense	missense	CNV	CNV	frameshift	splicing	missense	missense	-
Mutation (hg19)	c.5434C>T	-	c.6027G>C	-	-	c.8325_8326delinsT	c.6448C>G	c.6893G>A	9p 14.1-Mb deletion, 571-kb duplication	9q 773.8-Kb deletion	c.1438delG	c.5329-15A>G	c.7079G>T	c.1156G>A	-
Protein	p.Gln1812*	-	p.Leu2009Phe	-	-	p.Lys2775Asnfs*4	p.Leu2150Val	p.Arg2298His	-	-	p.Asp480Thrfs*3	-	p.Glu2360Val	p.Asp386Asn	-
Dysmorphic features															
synophrys	+	+	+	+	-	+	+	+	+	+	+	+	+	+	+
highly arched eyebrows	-	+	+	+	+	+	+	+	-	-	+	+	+	-	+
long curly eyelashes	+	+	+	+	+	+	+	-	-	+	+	+	+	+	+
ptosis	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
cleft lip	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
cleft palate	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
microcephaly	+	+	+	+	+	-	-	-	-	-	+	-	+	-	+
anteverted nostrils	-	+	+	+	-	-	+	+	+	-	+	+	-	+	+
depressed nasal bridge	+	+	+	+	-	+	-	-	-	+	-	+	-	+	-
long philtrum	+	+	-	+	-	+	-	-	+	-	+	-	+	-	-
thin upper lip	+	+	-	+	-	-	+	-	-	-	+	-	+	-	+
downturned corners of the mouth	+	+	+	+	-	-	+	-	-	-	+	-	+	-	-
micrognathia	-	+	-	-	-	-	+	-	-	-	+	-	+	-	-
short neck	-	-	+	-	-	+	-	-	-	-	-	-	-	+	-
high palate	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
widely spaced or absent teeth	+	-	-	-	-	+	+	-	-	-	-	-	-	-	-
others														macrocephaly	
Growth															
weight below 5th percentile for age	-	+	+	+	+	-	-	-	-	-	+	+	+	-	-
height or length below 5th percentile for age	-	+	+	+	+	+	-	-	-	-	-	+	+	-	-
prenatal growth retardation	+	+	+	+	+	-	+	-	-	-	+	-	+	+	-
others														obesity	
Development															
intellectual disability	-	+	NA	NA	+	-	+	+	+	-	+	-	+	+	+
developmental delay or mental retardation	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+
Behavior															
attention deficit disorder	+	+	NA	NA	+	+	+	+	+	+	-	+	+	+	+
anxiety	+	+	NA	NA	+	+	+	+	+	+	+	-	+	+	+
aggression	-	+	NA	NA	-	-	+	+	+	+	-	-	+	-	+
self-injurious behavior	-	-	NA	NA	-	-	+	-	+	+	+	-	+	+	-
autistic behavior	-	-	NA	NA	+	-	-	+	+	+	-	-	-	-	-
Limb abnormalities															
absence of forearms	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-
small hands and/or feet	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-
oligodactyly	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-
5th finger clinodactyly	-	+	-	+	-	-	+	-	-	-	-	-	-	-	-
abnormal palmar crease	-	-	+	-	-	-	+	-	-	-	+	-	-	-	-
syndactyly	-	-	+	+	-	+	+	-	-	-	-	-	-	-	-
phocomelia	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
limited elbow extension	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-
proximally placed thumbs	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-
others		camptodactyly					sternum	asymmetry							

Neurosensory–Skin																
ptosis	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
myopia	-	-	-	-	-	+	-	-	-	-	+	-	-	-	-	
deafness or hearing loss	-	+	-	-	-	-	-	-	+	+	-	-	-	-	+	
seizures	+	-	-	-	+	+	-	-	-	-	-	-	-	-	+	
hirsutism, generalized	-	+	-	+	+	+	+	+	-	-	+	+	+	-	+	
others																
Genitourinary																
cryptorchidism	+	-	+	-	-	+	-	-	-	-	+	-	-	-	-	
hypoplastic (small) genitalia	-	-	-	-	+	+	NA	-	-	-	+	-	-	-	+	
renal abnormalities	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
others																
Cardiovascular																
ventricular septal defects	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	
atrial septal defects	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
pulmonic ste-sis	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	
tetralogy of Fallot	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
hypoplastic left heart	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
bicuspid aortic valve	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	
others			aortic coarctation									aortic reflux				
Others																
language delay	-	+	NA	+	+	+	+	+	+	+	+	-	+	+	+	
diaphragmatic hernia	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
gastroesophageal reflux	-	+	+	+	+	-	+	+	-	+	+	-	-	-	+	
others									inguinal hernia							