

# *NIPBL* rearrangements in Cornelia de Lange syndrome: evidence for replicative mechanism and genotype–phenotype correlation

Davut Pehlivan, MD<sup>1</sup>, Melanie Hullings, BSE<sup>2</sup>, Claudia M.B. Carvalho, PhD<sup>1</sup>, Claudia G. Gonzaga-Jauregui, BS<sup>1</sup>, Elizabeth Loy, BS<sup>2</sup>, Laird G. Jackson, MD<sup>3</sup>, Ian D. Krantz, MD<sup>2,4</sup>, Matthew A. Deardorff, MD, PhD<sup>2,4</sup> and James R. Lupski, MD, PhD<sup>1,5,6</sup>

**Purpose:** Cornelia de Lange syndrome (CdLS) is a multisystem congenital anomaly disorder characterized by mental retardation, limb abnormalities, distinctive facial features, and hirsutism. Mutations in three genes involved in sister chromatid cohesion, *NIPBL*, *SMC1A*, and *SMC3*, account for ~55% of CdLS cases. The molecular etiology of a significant fraction of CdLS cases remains unknown. We hypothesized that large genomic rearrangements of cohesin complex subunit genes may play a role in the molecular etiology of this disorder.

**Methods:** Custom high-resolution oligonucleotide array comparative genomic hybridization analyses interrogating candidate cohesin genes and breakpoint junction sequencing of identified genomic variants were performed.

**Results:** Of the 162 patients with CdLS, for whom mutations in known CdLS genes were previously negative by sequencing,

deletions containing *NIPBL* exons were observed in 7 subjects (~5%). Breakpoint sequences in five patients implicated microhomology-mediated replicative mechanisms—such as serial replication slippage and fork stalling and template switching/microhomology-mediated break-induced replication—as a potential predominant contributor to these copy number variations. Most deletions are predicted to result in haploinsufficiency due to heterozygous loss-of-function mutations; such mutations may result in a more severe CdLS phenotype.

**Conclusion:** Our findings suggest a potential clinical utility to testing for copy number variations involving *NIPBL* when clinically diagnosed CdLS cases are mutation-negative by DNA-sequencing studies.

*Genet Med* 2012;14(3):313–322

**Key Words:** aCGH; CdLS; CNV; genomic rearrangement; *NIPBL*

## INTRODUCTION

Cornelia de Lange syndrome (CdLS; OMIM# 122470, 300590, and 610759) is a multiple congenital anomaly disorder, characterized by distinctive facial features, intellectual disability/developmental delay, hirsutism, and limb abnormalities.<sup>1</sup> Growth failure, which typically manifests in the second trimester, occurs proportionally.<sup>2</sup> Hirsutism is commonly observed especially on the face, neck, back, and extremities. Less frequently associated clinical findings include cardiac septal defects, gastrointestinal malformations/dysfunction, genitourinary malformations, ocular findings, and hearing problems.<sup>1,3,4</sup> Despite differences in clinical severity from patient to patient, the distinctive recognizable facies have provided the most differentiating feature in establishing the diagnosis. The prevalence of CdLS is estimated to be 1:10,000 live births but the incidence may be underestimated due to lack of recognition of milder cases.<sup>1</sup>

Thus far, three cohesin complex subunit genes—*NIPBL*, *SMC1A*, and *SMC3*—have been described in the molecular etiology of CdLS.<sup>5–8</sup> Heterozygous mutations in *NIPBL*,

which encodes a key regulatory protein of the cohesin complex that functions in sister chromatid cohesion and transcriptional regulation, are responsible for ~50% of all CdLS cases. Subsequently, the involvement of two additional core cohesin subunit genes, *SMC1A* and *SMC3*, were documented, but point mutations in these genes account for only 5% of all CdLS cases.<sup>9,10</sup> Hence, the molecular etiology of ~45% of CdLS cases remains unknown. Therefore, other cohesin complex subunit genes, and/or large genomic rearrangements—that is, copy number variation (CNV)—of known genes not detected by DNA-sequencing methods may have a role in the molecular etiology of CdLS.

The majority of disease-causing mutations in *NIPBL* are point mutations or single-nucleotide variants (SNVs); however, case reports of large genomic rearrangements have been published rarely.<sup>11–14</sup> Genotype–phenotype correlation studies revealed that patients who are mutation-negative tend to have a milder phenotype than patients who are mutation-positive, and patients with missense mutations have a milder phenotype

<sup>1</sup>Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, Texas, USA; <sup>2</sup>Division of Human Genetics, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA; <sup>3</sup>Department of Obstetrics and Gynecology, Drexel University School of Medicine, Philadelphia, Pennsylvania, USA; <sup>4</sup>University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania, USA; <sup>5</sup>Department of Pediatrics, Baylor College of Medicine, Houston, Texas, USA; <sup>6</sup>Texas Children's Hospital, Houston, Texas, USA. Correspondence: James R. Lupski (jlupski@bcm.edu)

Submitted 12 April 2011; accepted 9 September 2011; advance online publication 5 January 2012. doi:10.1038/gim.2011.13

than those with truncating mutations, which suggests that *NIPBL* is a dosage-sensitive gene.<sup>15</sup> Furthermore, Gause et al.<sup>16</sup> have shown that Nipped-B regulates cohesin chromosome binding in a dosage-sensitive manner in *Drosophila* salivary glands. Because of the rarity of CNV mutational events identified, genotype–phenotype correlations of large deletion CNVs involving *NIPBL* have not been extensively studied. Moreover, we showed recently that copy number gains involving *NIPBL* convey a different clinical phenotype than CdLS, further supporting the dosage sensitivity of *NIPBL*.<sup>13</sup> *NIPBL* duplication cases have common facial dysmorphic features including frontal bossing, broad nasal root, low-set ears, short philtrum, and high arched palate; however, most of these features do not overlap with the distinctive CdLS facial gestalt. Furthermore, some clinical findings such as overweight body habitus, short philtrum, and long fingers are completely opposite to what is observed with the CdLS phenotype.

CNVs are increased or decreased number of copies of a genomic segment, that is, deviations from the normal diploid state, which like SNV may either represent benign variations or result in a disease phenotype.<sup>17</sup> After application of genome-wide analysis tools such as comparative genomic hybridization (CGH), the human genome has been found to contain a high degree of CNVs comprising ~12% of the human genome in varying sizes from kilobases to megabases.<sup>18</sup> Duplication CNVs in the etiology of genetic diseases have long been known,<sup>19–21</sup> and more continue to be identified with recent advances in genome-wide scanning technologies.<sup>22</sup>

Three major mechanisms have been proposed for the formation of human genomic disorder-associated CNVs, including nonallelic homologous recombination, nonhomologous end joining, and the DNA replication-based mechanisms of fork stalling and template switching (FoSTeS)/microhomology-mediated break-induced replication (MMBIR).<sup>17</sup> Nonallelic homologous recombination is the predominant mechanism for the formation of recurrent genomic rearrangements by using low copy repeats as a substrate for recombination. In nonhomologous end joining, breaks in double strands occur, and then both broken DNA ends are bridged. The product of repair often contains additional nucleotides at the junction, leaving a “molecular scar.” FoSTeS/MMBIR is a recently described replication-based mechanism of DNA repair that utilizes nucleotide microhomology at the breakpoint junctions to prime DNA replication of a template switch; it has been found to be associated with the formation of nonrecurrent and complex rearrangements.<sup>23–25</sup>

To investigate a potential role for large genomic rearrangements in the etiology of CdLS, and to further delineate potential genotype/phenotype correlations of CNVs in the CdLS phenotype, we designed an Agilent 8×60K custom array interrogating all cohesin complex subunit genes and report herein our findings at the *NIPBL* locus. Furthermore, by defining and characterizing the breakpoint junctions of these genomic rearrangements, we sought to gain insights into underlying molecular mechanisms.

## MATERIALS AND METHODS

### Subjects

All patients were offered enrollment into the study after procuring informed consent at the Children’s Hospital of Philadelphia. Consent for publishing subject photographs was also independently procured. All patients were examined by clinical dysmorphologists experienced with the CdLS phenotype. To date, 162 CdLS cases, in which SNV mutations in *NIPBL*, *SMC1A*, and *SMC3* were excluded by targeted sequencing methods, and genome-wide single-nucleotide polymorphism arrays, were investigated. This study was approved by the institutional review boards of both Baylor College of Medicine and the Children’s Hospital of Philadelphia.

### Array CGH (aCGH)

We designed an Agilent 8×60K custom microarray with oligonucleotides interrogating cohesin complex subunit genes using the Agilent eArray website (<http://earray.chem.agilent.com/earray>). We included 46 genes and their flanking 50 kb of up- and downstream regions with an average genomic resolution of ~1 probe/200 bp. Genes related to cohesin structure and function were selected after a bibliographic search, including the OMIM and PubMed websites.

Experiments for digestion, labeling, purification of the labeled product, hybridization with gender-matched male (NA10851) or female (NA15510) control DNAs (obtained from Coriell Cell Repositories; <http://ccr.coriell.org>), washing, and scanning were conducted per the manufacturer’s protocol and previously described methods.<sup>26</sup> Computational analyses including data extraction, background subtraction, and normalization were done by using Agilent Feature Extraction Software 10\_7\_3\_1 (Agilent Technologies, Santa Clara, CA). These data were subsequently imported into array CGH analytics software (Genomic Workbench Standard Edition 5.0.14; Agilent Technologies). The genomic copy number was defined by analysis of the normalized log<sub>2</sub> (Cy5/Cy3) ratio average of the CGH signal. Regions that reached a threshold of at least 0.6 were considered gains consistent with duplication, and thresholds of at least –1.0 were considered significant losses consistent with deletion.

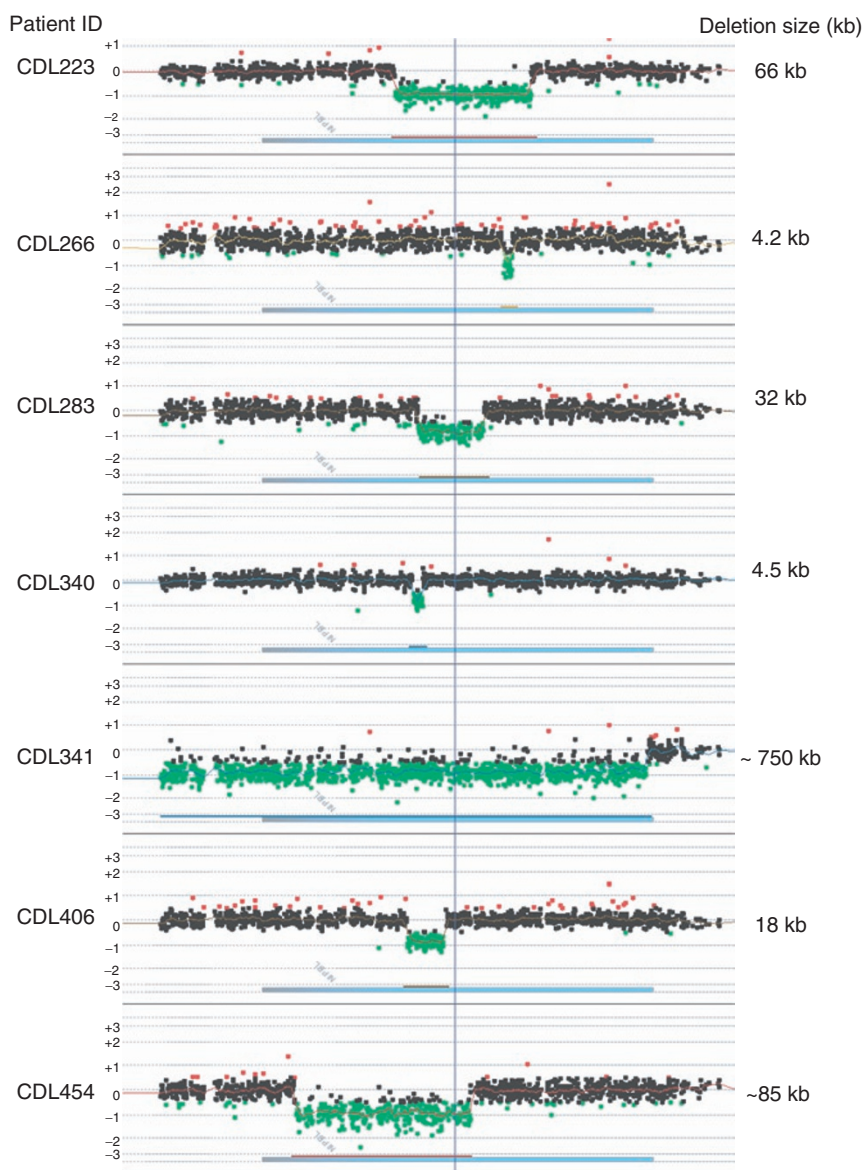
### Breakpoint analysis

To detect the breakpoint junctions of the large genomic rearrangements, primers were designed at the apparent boundaries of each segment based on aCGH analysis and the genomic coordinates of interrogating probes demarcating transitions from normal copy to apparent deletion. Both long-range and conventional PCR methods were conducted for each primer pair. Long-range PCR was performed as previously described.<sup>27</sup> Standard PCR was carried out in 12 µl of reaction mixture with 0.52 pmol/µl of each primer, 50 ng of genomic DNA, 10× PCR buffer, 0.2 mmol/l of each deoxynucleotide triphosphate, and 0.6 U of HotStar *Taq* DNA polymerase (Qiagen, Valencia, CA). The initial denaturation step at 95°C for 15 min was followed by 40 cycles of denaturation at 94°C for 30 s, annealing at 60°C for 30 s, and an extension at 72°C

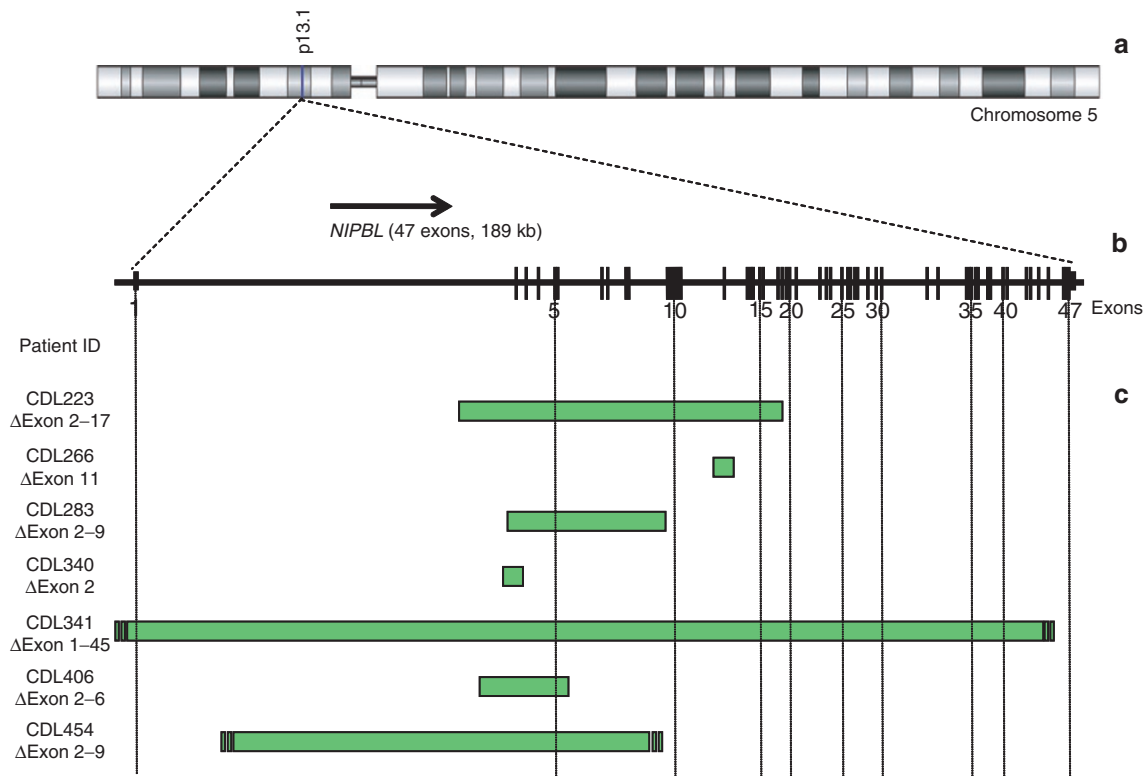
for 1 min. A final extension step at 72°C for 7 min was added. Amplification products were electrophoresed on 0.8–1% agarose gels. PCR products were purified using ExoSAP-IT (Affymetrix, Santa Clara, CA) and analyzed by standard Sanger di-deoxy nucleotide sequencing (DNA Sequencing Core Facility at Baylor College of Medicine, Houston, TX). We successfully amplified the breakpoints of five of seven patients by using different combinations of primers (223-F1: 5'-TTG TTCTGGCAGTCTGTAGTATGG-3' and 223-R1: 5'-TTAA TGGCACACAACACTGTAGTTCAC-3' for patient CDL223; 266-F1: 5'-CAGCGTTCACTTTTGGAGGATGATA-3' and 266-R2: 5'-CCTTCAACATTTTCCCCTAACCTTC-3' for

patient CDL266; 283-F2: 5'-TGTCAGTCATTCACCAAAGG AAAGT-3' and 283-R2: 5'-TCTGCCAATATACCAAACAGG AAA-3' for patient CDL283; 340-F1: 5'-CATGGCAAAAAGTA AGATGCAGAAGA-3' and 340-R1: 5'-CCAAAGAAAAGTA TGCCATCCTCTC-3' for patient CDL340; and 406-F2: 5'-CCTTGTGAGATGAGTATGCTTTTCC-3' and 406-R2: 5'-GTGTGTTATTTCTCCTATCAGACAGT-3' for patient CDL406).

PCR products of CDL223 and CDL406 were further sequenced by primer walking using the following primers: 223-FW1: 5'-GGATTCAAAACTAAGCAATT-3' and 223-RW1: 5'-GAATTAAGAGAACACAATTT-3' for patient



**Figure 1** Array comparative genomic hybridization (aCGH) results displaying the *NIPBL* gene region. Results for aCGH analyses in each of the seven Cornelia de Lange syndrome (CdLS) cases are shown. Individual dots represent interrogating oligonucleotide probes: black dot represents normal copy number as compared with a gender-matched control, red dots represent copy number gain, and green dots represent copy number losses as compared with a gender-matched control. Horizontal blue bar represents the *NIPBL* gene. Numbers on the y axis show the log<sub>2</sub> ratio of the hybridization signal of patient versus control. The approximate size in kb of the deletion is shown at right.



**Figure 2 Schematic view of *NIPBL* displaying exonic deletions in seven patients.** (a) Chromosome 5 karyogram with G bands indicated (top). The location of *NIPBL* is demarcated with a blue vertical line at 5p13.1. (b) Graphic view of 47 exons (vertical black bars) of *NIPBL*; size and orientation of the gene above the exons. (c) Solid green bars represent genomic regions deleted with approximate sizes. Vertical dotted lines track exons on deleted regions. Patients' code and deleted exons ( $\Delta$ ) are given at the left. The graphical normalized data for each patient was obtained by inputting the most distal and proximal oligonucleotide genomic probe coordinates into the custom track at the University of California, Santa Cruz website, <http://genome.ucsc.edu/cgi-bin/hgGateway>. Narrow green vertical bars depict uncertainty for proximal and distal ends of the deleted regions in patients CDL341 and CDL454, for which breakpoint junctions were not determined. Note the 77-kb length of intron 1; thus intron 1 harbors the distal breakpoint in five of seven cases.

CDL223, and 406-RW1: 5'-ATTTGAGAATGTCTACTCAC-3' for patient CDL406.

## RESULTS

### High-resolution genomic analyses of the *NIPBL* gene region in CdLS

To investigate a potential role for large genomic rearrangements as an etiology for CdLS, we performed high-resolution genome-wide gene-targeted CGH array analyses. In total, 7 of 162 unrelated patients were found to harbor *NIPBL* deletions ranging in size from 4.2 to 750 kb (Figure 1). As anticipated, parental studies revealed that deletions occurred *de novo* in all CdLS cases for whom parental samples were available (4 of 7, both parents; 2 of 7, only one parent). Patients CDL266 and CDL340 each had a relatively small single-exon dropout mutation; deletions of exon 11 (4.2 kb in size) and exon 2 (4.5 kb in size), respectively. Except for patient CDL341, all deletions were intragenic exonic deletions of varying sizes. Patient CDL341 has a deletion spanning ~750 kb, that comprised almost the entire *NIPBL* gene and the 5'-flanking *SLC1A3*; we could not detect the breakpoint for this patient using our aCGH assay, as our aCGH probe coverage did not allow us to

easily assess whether the deletion includes the distal *RANBP3L* gene or not. The deletions observed in patients CDL283 and CDL454 encompassed exons 2–9; although the deletion in each patient is different in size: 32 and 85 kb, respectively. Two patients, CDL223 and CDL406, were found to have multiexon intragenic *NIPBL* deletions including exon 2–17 deleting 66 kb and exon 2–6 deleting 18 kb, respectively (Figure 2). Of note, intron 1, the longest intron, which constitutes 77 kb of the 189 kb *NIPBL* locus, harbors one end of the breakpoints in 5 of 7 patients.

### Breakpoint junction sequence analyses

We performed breakpoint sequencing to further fine-map the deletions and potentially infer rearrangement mechanisms. The PCR amplification of breakpoint junctions was achieved in five of seven patients (CDL223, CDL266, CDL283, CDL340, and CDL406; Figure 3). In four of five breakpoints, the distal end of the deletion harbors a repetitive sequence such as a microsatellite, mammalian interspersed repeat element, or a member of the long interspersed element family. Breakpoints from each of these four patients showed 1–5 bp of shared microhomology between proximal and



**Table 1** Clinical features of patients with deletions in *NIPBL*

Patient info							
Patient ID	CDL223	CDL266	CDL283	CDL340	CDL341	CDL406	CDL454
Sex	F	M	F	M	M	M	M
CdLS info							
Clinical severity	Severe	Mild	Moderate	Mild	Severe	Mild	Mild
<i>NIPBL</i> deletion	Exons 2–17	Exon 11	Exons 2–9	Exon 2	Exons 1–45	Exons 2–6	Exons 2–9
Deletion size	66 kb	4.2 kb	32 kb	5.2 kb	750 kb	18 kb	85 kb
Family testing							
Mother	N/A	–	–	–	–	–	–
Father	N/A	–	N/A	–	–	N/A	–
Head							
Microcephaly	+	+	+	+	+	+	+
Low anterior hairline	+	+	+	+	+	+	–
Other	Mild bitemporal narrowing			Low-set ears			
Eyes							
Arched eyebrows	+	+	+	–	+		+
Synophrys	High arched		–	–	+	+	+
Long eyelashes	+	+	+	+	+	+	+
Ptosis	+	–	–	–	+	–	–
Myopia	+	–				–	–
Nasolacrimal duct stenosis	+		–	+		–	+
Other		Horizontal and vertical nystagmus					
Nose							
Depressed nasal bridge	+	+	+	+	+	+	+
Anteverted nostrils	+	+	+	+	+	+	+
Long/flat philtrum	+	+	+	+	+	+	+
Mouth							
Thin upper lip	+	+	+	+	+	+	+
Downturned corners of mouth	+	+	+	+	+	+	+
Palate—high arch	+	–		–		–	–
Palate—cleft		–	+	–		–	–
Teeth (small, widely spaced)	+	–	+	–		–	+
Micrognathia	+	–	+	+	+	+	+
Other							
Skin							
Cutis marmorata	+	–	–	+	+	+	
Hirsutism	Face, arms, back, lateral thighs		+	–	+	–	
Hands							
Small hands	+	+	+	+	nr	+	+
Proximally set thumbs	R	+	+	+	nr	+	+
Clinodactyly 5th finger	nr	+	+	+	nr	+	+

No notation indicates that information is unavailable.

CdLS, Cornelia de Lange syndrome; CNS, central nervous system; HC, head circumference; N/A, not applicable; nr, not relevant; R, right; reg., regular.

**Table 1** Continued on next page.

**Table 1 Continued**

**Patient info**

Single palmar crease	+	–	+	+	nr	+	Right
Other	3 Digits on left arm, right 5th digit tip abnormal				Oligodactyly left arm, truncation of right arm		
<b>Feet</b>							
Small feet			+	+	+		+
Syndactyly of toes	+		–				+
<b>Arms</b>							
Restriction of elbow movements	+	–	–		+		–
Other skeletal	Minor scoliosis, hip click						
Cardiac defects	–	–	–	–		–	–
Genitourinary defects	–		Hypospadias		–	–	Inguinal hernia
<b>Gastroenterology</b>							
Gastroesophageal reflux	+	Possible	+	–			+
Feeding problems in infancy	+	+	+	–			+
<b>Otolaryngology</b>							
Hearing loss	+	–	–	+		–	+
<b>CNS</b>							
CNS anomalies	–	–	–	–			–
Seizures	+		–	Febrile	Severe		–
<b>Cognitive</b>							
Mental retardation	+		+			At 11 years, reads at first-grade level	Mild
Behavior, personality	Nonverbal		Alert	Autistic-like			Pleasant
<b>Development</b>							
Verbal development	None	Delayed (spoke at 4.5 years)	Delayed (no words at 17 months)	Delayed (4 words at 4 years)	Severe delay	Mild delay	First words, 3 years
Motor development	Nonambulatory, rolls	Walking at 20 months	Delayed	Delayed	Delayed	Mild delay	Walked at 2 years
Other	Asthma, pneumonia, ringworm	Bunions, horizontal nystagmus	Bilateral inguinal hernia, wheezing	Affected brother, asthma		Chronic poor weight gain	
<b>Measurements</b>							
Gestational age (weeks)	37	40	27	37		40	32
<b>Birth</b>							
Birth weight (g)	2,240	2,950	1,190	2,900		2,700	1,490
Percentile weight (reg. chart)	20%	40%	<3	55%		15%	25%
Length at birth (cm)	48	46		44.5		47	14 in
Percentile height (reg. chart)	50%	5%		15%		15%	<10%
HC at birth		33					
Percentile HC (reg. chart)		25%					
APGAR score (1 min/5 min)							

No notation indicates that information is unavailable.

CdLS, Cornelia de Lange syndrome; CNS, central nervous system; HC, head circumference; N/A, not applicable; nr, not relevant; R, right; reg., regular.

**Table 1 Continued on next page.**

Table 1 Continued

## Patient info

Later evaluations						
Age at evaluation	3 years	15.7 years	17 months	4 years	11 years	11 years
Weight (kg)	7.4		7.38	12.3	20.3	17.7
Percentile weight (reg. chart)	<5%		<3%	<3%	<<1%	<5%
Percentile weight (CdLS chart)	5–10%		90%	90%	50%	30%
Height (cm)	77.9		71.5	92.5	122.5	112.5
Percentile height (reg. chart)	<5%	<5%	<3%	<3%	<<1%	<5%
Percentile height (CdLS chart)	25%		90%	90%	60%	40%
HC (cm)	41		42.3	47	47.5	44.2
Percentile HC (reg. chart)			<3%	<3%	<<1%	<3%
Percentile HC (CdLS chart)			45%	90%	90%	10%

No notation indicates that information is unavailable.

CdLS, Cornelia de Lange syndrome; CNS, central nervous system; HC, head circumference; N/A, not applicable; nr, not relevant; R, right; reg., regular.



**Figure 4 Patient photographs.** Frontal view of facies and extremity pictures of patients with limb abnormality.

of cases. Because studies in *Drosophila*<sup>16</sup>, mouse,<sup>28</sup> and human<sup>29</sup> showed that *NIPBL* is a dosage-sensitive gene, we hypothesized that large genomic deletions may contribute to a fraction of the remaining molecularly unidentified cases. Previously, two

groups have reported a low frequency of large genomic rearrangements in CdLS.<sup>11,12</sup> Ratajska et al.<sup>12</sup> studied 11 *NIPBL*/*SMC1A* mutation-negative cases and found one deletion spanning 62.7 kb and encompassing exons 35–47 of the *NIPBL* gene. Bhuiyan et al.<sup>11</sup> analyzed 50 CdLS probands negative for *NIPBL* mutation and found a single 5.2-kb deletion encompassing exons 41–42 of *NIPBL*.

This study systematically assessed genomic rearrangements in a large cohort of point mutation-negative CdLS cases. Furthermore, the genomic span and breakpoint junctions of multiple *NIPBL* deletions have been comprehensively examined. Our data show that large intragenic deletions of *NIPBL* can account for ~5% of mutation-negative CdLS cases. Thus, we suggest screening for large genomic deletions of *NIPBL* in SNV mutation-negative CdLS cases.

We identified microhomology at four of the five sequenced breakpoints that ranged from 1 to 5 bp, consistent with a possible replicative mechanism such as FoSTeS/MMBIR. Patient CDL266 has a 47-bp insertion and no microhomology at the breakpoint junction. We initially evaluated this case as a non-homologous end joining event; however, detailed examination of the inserted segment revealed that another replicative mechanism, serial replication slippage, potentially underlies this complex rearrangement (see **Supplementary Figure S1** online). Serial replication slippage has been proposed to explain complex rearrangements, especially those potentially occurring within 100-bp intervals of the replication fork.<sup>30</sup>

In reviewing the two reported CNV-associated CdLS cases, a breakpoint junction has been studied in detail in only one patient. In the case reported by Ratajska et al.,<sup>12</sup> the proximal and distal deletion breakpoints mapped within two different long interspersed element 1 from distinct families. They did not observe microhomology at the breakpoint junction, but a 15-bp insertion was evaluated as reflecting a potential nonhomologous end joining event as the mechanism for the deletion. We re-analyzed the case reported by Ratajska et al.<sup>12</sup> and now suggest that serial replication slippage may be the responsible



mechanism for that deletion; as viewed in this perspective, apparent 15-bp insertion actually consist of smaller insertions flanked by microhomology (see **Supplementary Figure S2** online). The breakpoint junction of the case reported by Bhuiyan *et al.*,<sup>11</sup> was not studied in detail.

The clinical features of patients with genomic deletions that wholly or partially alter the *NIPBL* locus are comparable to the distinctive facial features, cognitive delay, and growth deficiencies associated with classical CdLS. However, a broad range of clinical severity that is manifest in their degree of cognitive, growth, and structural involvement can be observed with *NIPBL* deletions. Largely, this correlates with the size of deletion and the number of exons involved, with more exon involvement manifesting as more severe features. However, there are some notable findings that may explain some of the molecular basis for the differences between patients (**Table 1**). For example, the most severe patients include patient CDL223, who demonstrates a severe, classic, *NIPBL* phenotype due to deletion of a large part of the open reading frame of *NIPBL* (exons 2–17). In addition, patient CDL341 demonstrates severe cognitive and growth delays and has a deletion of *NIPBL* exons 1–45 and upstream genes.

In contrast, patient CDL266, with a milder form of the disorder, is a higher-functioning patient with a deletion of only exon 11 that is predicted to result in an in-frame deletion. Also, patient CDL340 has less cognitive and growth involvement with a deletion of only exon 2. It is possible that an in-frame ATG at c.334 in exon 4 provides an alternative start codon.

Consistent with this model, three patients have deletions that include exons 5–8. Each of these patients demonstrates a phenotype that is intermediate to those above; however, their resultant features vary significantly despite deletions at the same three exons. Patient CDL406 is mildly affected with typical CdLS facies and a deletion of exons 2–6. It is possible that the ATG at c.748 in exon 7 could serve as an alternative start codon. Patient CDL283 is a moderately severe Caucasian patient with typical facies with a deletion of exons 2–9. Because exon 2 contains the primary start codon, it is possible that an alternative start codon at c.1618, p.540 in exon 10 could be used, resulting in a truncated protein. In comparison, patient CDL454, an African-American patient with a deletion of the same exons 2–9, demonstrates typical CdLS facies and growth delays, although he appears cognitively less affected than patient CDL283 despite having a deletion that is twice as large. Although these latter two patients could share the same alternative start codon, there are clearly modifying factors that influence the cognitive outcomes, either within the *NIPBL* locus or at other genomic sites.

To fully understand the multiple potential causes of CdLS and to be able to predict the growth and cognitive outcomes for patients with *NIPBL* mutations, much more work will need to be done to clarify the type of mutations and their resultant effect on the genome. Nevertheless, this study identifies additional underlying causes of CdLS, and provides insight into the molecular bases of these genomic rearrangements.

## SUPPLEMENTARY MATERIAL

Supplementary material is linked to the online version of the paper at <http://www.nature.com/gim>

## ACKNOWLEDGMENTS

We thank the patients with Cornelia de Lange syndrome and their families who participated in this study as well as the referring physicians and colleagues who have contributed samples and clinical information. We are indebted to the continued support of the US and International Cornelia de Lange syndrome foundations. This work was supported by National Institutes of Health Grants K08HD055488 (National Institute of Child Health and Human Development (NICHD); M.A.D.), P01 HD052860 (NICHD; I.D.K.), research grants from the US CdLS Foundation, institutional funds from the Children's Hospital of Philadelphia, and March of Dimes grant 6-FY09-504 (J.R.L.).

## DISCLOSURE

J.R.L. is a paid consultant for Athena Diagnostics and Ion Torrent Systems, has stock ownership in 23andMe, and is a co-inventor on multiple US and European patents related to molecular diagnostics for inherited neuropathies, eye diseases, and bacterial genomic fingerprinting. The Department of Molecular and Human Genetics at Baylor College of Medicine derives revenue from the chromosomal microarray analysis offered in the Medical Genetics Laboratory (<http://www.bcm.edu/geneticlabs>). The other authors declare no conflict of interest.

## REFERENCES

1. Jackson L, Kline AD, Barr MA, Koch S. de Lange syndrome: a clinical review of 310 individuals. *Am J Med Genet* 1993;47:940–946.
2. Kline AD, Barr M, Jackson LG. Growth manifestations in the Brachmann-de Lange syndrome. *Am J Med Genet* 1993;47:1042–1049.
3. de Lange C. Sur un type nouveau de degeneration (typus Amstelodamensis). *Arch Med Enfants* 1933;36:713–719.
4. Allanson JE, Hennekam RC, Ireland M. De Lange syndrome: subjective and objective comparison of the classical and mild phenotypes. *J Med Genet* 1997;34:645–650.
5. Krantz ID, McCallum J, DeScipio C, et al. Cornelia de Lange syndrome is caused by mutations in *NIPBL*, the human homolog of *Drosophila melanogaster* Nipped-B. *Nat Genet* 2004;36:631–635.
6. 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* 2004;36:636–641.
7. Deardorff MA, Kaur M, Yaeger D, 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* 2007;80:485–494.
8. Musio A, Selicorni A, Focarelli ML, et al. X-linked Cornelia de Lange syndrome owing to *SMC1L1* mutations. *Nat Genet* 2006;38:528–530.
9. Yan J, Saifi GM, Wierzbica TH, et al. Mutational and genotype-phenotype correlation analyses in 28 Polish patients with Cornelia de Lange syndrome. *Am J Med Genet A* 2006;140:1531–1541.
10. Dorsett D, Krantz ID. On the molecular etiology of Cornelia de Lange syndrome. *Ann NY Acad Sci* 2009;1151:22–37.
11. Bhuiyan ZA, Stewart H, Redeker EJ, Mannens MM, Hennekam RC. Large genomic rearrangements in *NIPBL* are infrequent in Cornelia de Lange syndrome. *Eur J Hum Genet* 2007;15:505–508.
12. Ratajska M, Wierzbica J, Pehlivan D, et al. Cornelia de Lange syndrome case due to genomic rearrangements including *NIPBL*. *Eur J Med Genet* 2010;53:378–382.

13. Yan J, Zhang F, Brundage E, et al. Genomic duplication resulting in increased copy number of genes encoding the sister chromatid cohesion complex conveys clinical consequences distinct from Cornelia de Lange. *J Med Genet* 2009;46:626–634.
14. Hulinsky R, Byrne JL, Lowichik A, Viskochil DH. Fetus with interstitial del(5)(p13.1p14.2) diagnosed postnatally with Cornelia de Lange syndrome. *Am J Med Genet A* 2005;137A:336–338.
15. Gillis LA, McCallum J, Kaur M, et al. NIPBL mutational analysis in 120 individuals with Cornelia de Lange syndrome and evaluation of genotype-phenotype correlations. *Am J Hum Genet* 2004;75:610–623.
16. Gause M, Misulovin Z, Bilyeu A, Dorsett D. Dosage-sensitive regulation of cohesin chromosome binding and dynamics by Nipped-B, Pds5, and Wapl. *Mol Cell Biol* 2010;30:4940–4951.
17. Zhang F, Gu W, Hurler ME, Lupski JR. Copy number variation in human health, disease, and evolution. *Annu Rev Genomics Hum Genet* 2009;10:451–481.
18. Redon R, Ishikawa S, Fitch KR, et al. Global variation in copy number in the human genome. *Nature* 2006;444:444–454.
19. Lupski JR, de Oca-Luna RM, Slaugenhaupt S, et al. DNA duplication associated with Charcot-Marie-Tooth disease type 1A. *Cell* 1991;66:219–232.
20. Inoue K, Osaka H, Sugiyama N, et al. A duplicated PLP gene causing Pelizaeus-Merzbacher disease detected by comparative multiplex PCR. *Am J Hum Genet* 1996;59:32–39.
21. Lupski JR. Genomic disorders ten years on. *Genome Med* 2009;1:42.
22. Stankiewicz P, Lupski JR. Structural variation in the human genome and its role in disease. *Annu Rev Med* 2010;61:437–455.
23. Lee JA, Carvalho CM, Lupski JR. A DNA replication mechanism for generating nonrecurrent rearrangements associated with genomic disorders. *Cell* 2007;131:1235–1247.
24. Hastings PJ, Ira G, Lupski JR. A microhomology-mediated break-induced replication model for the origin of human copy number variation. *PLoS Genet* 2009;5:e1000327.
25. Hastings PJ, Lupski JR, Rosenberg SM, Ira G. Mechanisms of change in gene copy number. *Nat Rev Genet* 2009;10:551–564.
26. Shinawi M, Liu P, Kang SH, et al. Recurrent reciprocal 16p11.2 rearrangements associated with global developmental delay, behavioural problems, dysmorphism, epilepsy, and abnormal head size. *J Med Genet* 2010;47:332–341.
27. Carvalho CM, Zhang F, Liu P, et al. Complex rearrangements in patients with duplications of MECP2 can occur by fork stalling and template switching. *Hum Mol Genet* 2009;18:2188–2203.
28. Kawachi S, Calof AL, Santos R, et al. Multiple organ system defects and transcriptional dysregulation in the Nipbl(+/-) mouse, a model of Cornelia de Lange Syndrome. *PLoS Genet* 2009;5:e1000650.
29. Liu J, Zhang Z, Bando M, et al. Genome-wide DNA methylation analysis in cohesin mutant human cell lines. *Nucleic Acids Res* 2010;38:5657–5671.
30. Chen JM, Chuzhanova N, Stenson PD, Férec C, Cooper DN. Meta-analysis of gross insertions causing human genetic disease: novel mutational mechanisms and the role of replication slippage. *Hum Mutat* 2005;25:207–221.