Pié J, Gil-Rodríguez MC, Ciero M, López-Viñas E, Ribate MP, Arnedo M, Deardorff MA, Puisac B, Legarreta J, de Karam JC, Rubio E, Bueno I, Baldellou A, Calvo MªT, Casals N, Olivares JL, Losada A, Hegardt FG, Krantz ID, Gómez-Puertas P, Ramos FJ. 2010. Mutations and variants in the cohesion factor genes *NIPBL*, *SMC1A*, and *SMC3* in a cohort of 30 unrelated patients with Cornelia de Lange syndrome. *Am J Med Genet Part A* 152A:924-929.

Additional supporting information may be found in the online version of this article.

<**http://www3.interscience.wiley.com/journal/123334138/suppinfo**>

Fig. 1: Sequence analysis of NIPBL mutations. A. Sequence features of human NIPBL and positions of amino acid residues mutated in Cornelia de Lange Syndrome (CdLS) in this cohort. Features include: MAU-2 interaction domain (residues 1-139), predicted coiled-coil sequence (residues 637-657), tandem repeats of undecapeptide PETPKQK(G/S)(E/D)(G/S)R (699-764), nuclear localization signal (NLS, 1108-1124), HEAT-repeat region (spanning residues 1750-2350) and HDAC1 and HDAC3 interaction domain (residues 1838-2000). Two major isoforms of NIPBL, A and B, are differentiated by the presence or absence of C-terminal residues 2698-2804. Positions of mutated residues, described in the text, are indicated by red dots. Black dot indicates the position of a mutation previously reported by Schoumans et al. [2007]. B. Multiple sequence alignment of NIPBL to several organisms (Homo sapiens: NIPBL_HUMAN; Rattus norvegicus: NIPBL_RAT; Gallus gallus: NIPBL_CHICK; Danio rerio: NIPBL_DANRE; Drosophila melanogaster: NIPB_DROME; Arabidopsis thaliana: Q9LF28_ARATH; Saccharomyces cerevisiae: SCC2_YEAST) surrounding position of V1441 and F1442 residues. Increased conservation of residues is indicated by darker shading. C. Multiple sequence alignment of the NIPBL segment located around the N1897 residue. Secondary structure assignment of the HEAT domain is included. D. Alignment of sequences homologous to NIPBL around G2081 and S2090I residues. E. Same analysis, performed in the vicinity of residue L2150. Black dot indicates the position of a mutation previously reported by Schoumans et al. [2007] and white dot indicates the position of residue I1510 of Nipped-B (Drosophila melanogaster).

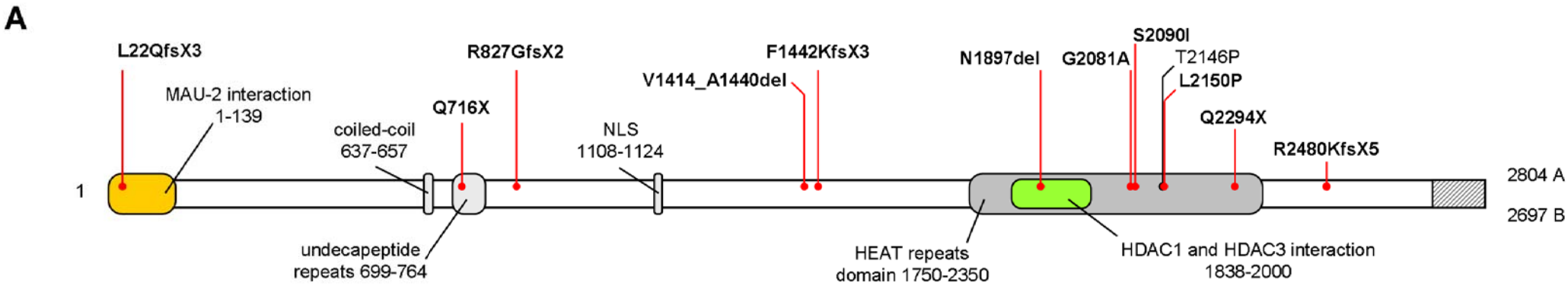
Fig. 2: Alternative splicing of NIPBL mutants. Agarose gel electrophoresis of RT-PCR product demonstrates that in patient C5, the mutation p.F1442KfsX3 causes aberrant splicing resulting in two bands: one with normal size (390bp) containing both the normal allele and one with the exonic mutation p.V1441L, and other band with a shorter size of 289bp, demonstrating a skipping of exon 20. In patient C10 the mutation p.L22QfsX3 results in aberrantly spliced transcript with a normal band (582bp) and another with a size of 416bp, skipping exon 3. In patient C28 the mutation p.V1414_A1440del results in an aberrantly spliced transcript with a normal band (390bp) and another of 309bp, skipping exon 19. WT=wild-type allele or control patient.

Fig. 3: A. Schematic representation of the SMC1A/SMC3 heterodimer in the Cohesin complex and the locations of SMC1A mutations in coiled-coil structure. Coiled-coil arms connect the hinge domain to the head domain. Position of mutated residues in patients with Cornelia de Lange Syndrome, described in the text, are indicated by red dots. An altered residue, which was previously reported by Deardorff et al. [2007], is indicated by a black dot. B. Multiple sequence alignment of several proteins homologous to SMC1A in the area surrounding K268 residue. Represented sequences are: Homo sapiens (SMC1A_HUMAN), Rattus norvegicus (SMC1A_RAT), Gallus gallus (Q8AWB7_CHICK), Danio rerio (Q6DRM9_DANRE), Drosophila melanogaster (NIPB_DROME), Arabidopsis thaliana (Q9LF28_ARATH), Saccharomyces cerevisiae (SMC1_YEAST), and Methanococcus jannaschii (SMC_METJA). C. Same analysis, performed in the vicinity of residue R711. Heptad signature corresponding to the coiled-coil structure of the protein segment is indicated as "coiled_coil". Residues are colored as above.

Table I: Clinical and molecular features of 14 patients with CdLS with mutation in NIPBL or SMC1A.

Table II: Novel NIPBL, SMC1A and SMC3 polymorphisms and variants of unknown significance identified.

Table III: Previously reported NIPBL, SMC1A and SMC3 polymorphisms.



B

```
NIPBL_HUMAN   1430 LQLCAIKLVTAVFS--RYEKHRQLILEEIFTSLARLPTSKR 1468
NIPBL_RAT     1424 LQLCAIKLVTAVFS--RYEKHRQLILEEIFTSLARLPTSKR 1462
NIPBL_DANRE   1444 LQLCAIKLVTAVFS--RYEKHRQLILEEIFTSLARLPTSKR 1482
NIPBL_CHICK   1406 LQLCAIKLVTAVFS--RYEKHRQLILEEIFTSLARLPTSKR 1444
SCC2_YEAST     391 VKRISSDILVSLFG--SFDQQRGFIIEELLSHIEKLPTKRI  429
NIPB_DROME     719 LQFVCLELVTTIFRKERYDKIRNSILGDILTSIDRLPSSKK  759
Q9LF28_ARATH   438 LHYSWLTQGPVVGR--------------------KLPSSKR  458
```

C

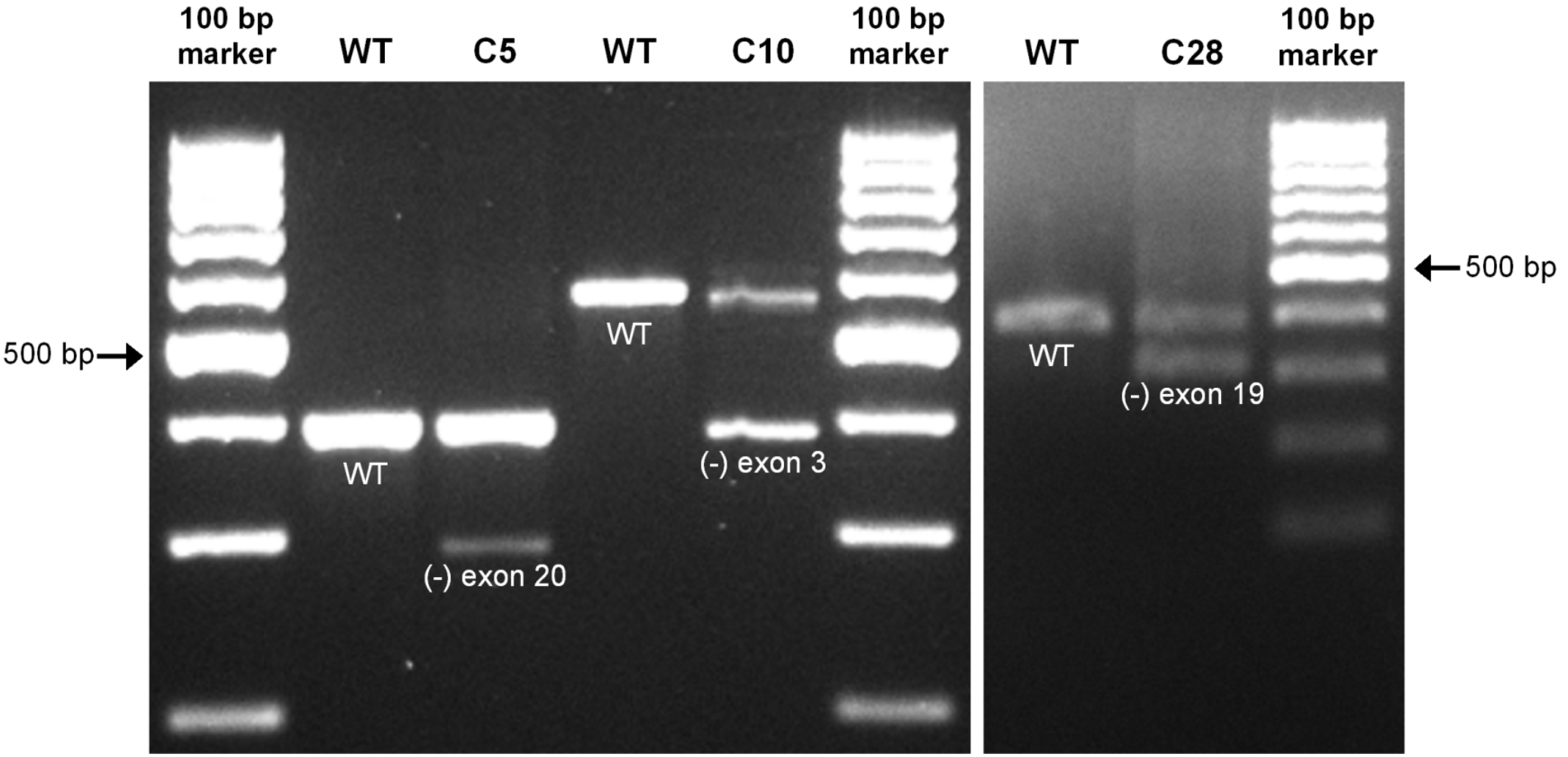
```
NIPBL_HUMAN   1888 MCVKMIRRVNDEEG-IKKLVNETFQKLWF 1915
NIPBL_RAT     1882 MCVKMIRRVNDEEG-IKKLVNETFQKLWF 1909
NIPBL_DANRE   1904 MCVKMIRRVNDEEG-IKKLVNETFQKLWF 1931
NIPBL_CHICK   1864 MCVKMIRRVNDEEG-IKKLVNETFQKLWF 1891
SCC2_YEAST     810 VIARILMKIEDEEDNIIDMARLILLNRWI  838
NIPB_DROME    1245 ICVKMIRRVHDEEG-IQKLVTEVFMKMWF 1272
Q9LF28_ARATH   892 ACAEILSRISDDESSVQDLVCKTFYEFWF  920
```

D

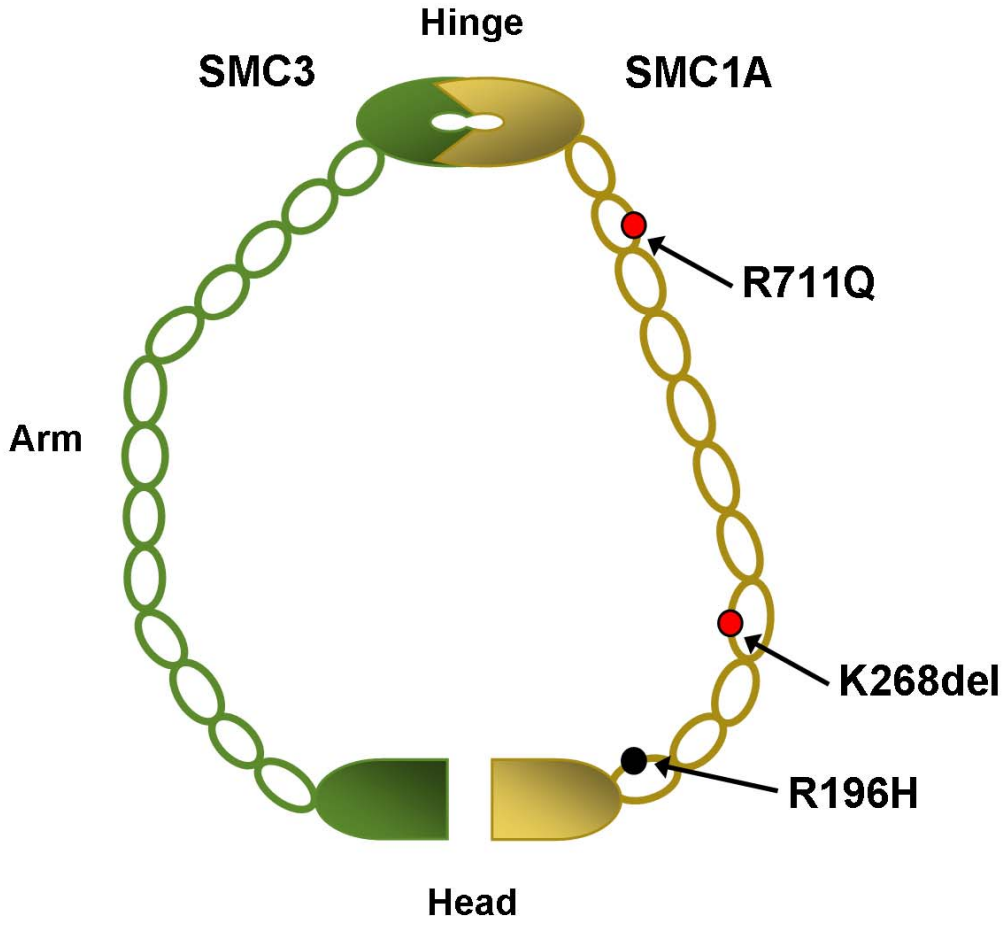
```
NIPBL_HUMAN   2065 FLATIEEDLMKLIIKY-GMTVVQHCVSCLGAVVNKV 2099
NIPBL_RAT     2059 FLATIEEDLMKLIIKY-GMTVVQHCVSCLGAVVNKV 2093
NIPBL_DANRE   2081 FLTTIEEDLMKLIIKY-GMTVVQHCVSCLGAVVNRV 2115
NIPBL_CHICK   2041 FLATIEEDLMKLIIKY-GMTVVQHCVSCLGSVVNKV 2075
SCC2_YEAST     994 LETTLLSRLPKMNVRE--IDEAMPLIWSVATHRHDT 1027
NIPB_DROME    1429 FLASLEEHLMLLVVSR-NQAEVTSCVSCLGALVNKI 1463
Q9LF28_ARATH  1090 VTEDLEQDLKHMIVRHSFLTVVHACVSKLAGKG--- 1122
```

E

```
NIPBL_HUMAN   2142 RSLFTVGALCRHFDFDLEDFKGNS----KVNIKDKVL 2174
NIPBL_RAT     2136 RSLFTVGALCRHFDFDLEDFKGNS----KVNIKDKVL 2168
NIPBL_DANRE   2158 RSLFTVGALCRHFDFDQEEFKGSN----KVVIKDKVL 2190
NIPBL_CHICK   2118 RSLFTVGALCRHFDFDHEDFKGNS----KVNIKDKVL 2150
SCC2_YEAST    1062 QRLIYLSTGFARFCFPKPSNDKIAFLQEGETLYEHIT 1098
NIPB_DROME    1505 RSLFTIGILMRYFDFKSPIALGETNDGLPVSICEDVF 1541
Q9LF28_ARATH  1150 RSLFCLGLLIRHGNSLISTSGGKN------FNLSGCL 1180
```


100 bp marker
WT
C5
WT
C10
100 bp marker
WT
C28
100 bp marker
500 bp
500 bp
WT
(-) exon 20
WT
(-) exon 3
WT
(-) exon 19

A



B

coiled_coil		abcdefgabcdefgabcdefg	
SMC1A_HUMAN	256	MDKVEDELKEKKKELGKMMRE	276
SMC1A_RAT	256	MDKVEDELKEKKKELGKMMRE	276
Q6DRM9_DANRE	256	MDRVEEELKDKKKELGRMMRD	276
Q8AWB7_CHICK	256	MDRVEDELKDRKKELGKMMRE	276
SMC1_YEAST	270	INNEMKSLQRSKSSFVKESAV	290
Q9VCD8_DROME	279	KEAADEILREKKKDAGKITRD	299
Q6Q1P4_ARATH	266	LEKFEREAGKRKVEQAKYLKE	286
SMC_METJA	266	VREIDVEIENLKLRLNNIINE	286

C

coiled_coil		cdefgabcdefgabcdefg	
SMC1A_HUMAN	700	QVQSQAHGLQMRLKYSQSD	718
SMC1A_RAT	700	QVQSQAHGLQMRLKYSQSD	718
Q6DRM9_DANRE	700	QVQSQAHGLQMRLKYSQSD	718
Q8AWB7_CHICK	700	QVQSQAHGLQMRLKYSQSD	718
SMC1_YEAST	712	EVENSVSLLNSDIANLRTQ	730
Q9VCD8_DROME	726	TVESQIKGLENRLKYSMVD	744
Q6Q1P4_ARATH	705	EISGKISGLEKKIQYAEIE	723
SMC_METJA	708	RSSAKKMEIENTLEIIKKN	726

Supplementary Table I. Clinical and molecular features of 14 patients with CdLS with mutation in *NIPBL* or *SMC1A*.

Patient	C2	C5	C8	C10	C13	C14	C18✞
Gene mutated	*SMC1A*	*NIPBL*	*NIPBL*	*NIPBL*	*SMC1A*	*NIPBL*	*NIPBL*
Exon	4	20	35	Intron 3	5	10	37
cDNA mutation*	c.587G→A[a]	c.4321G→T	c.6242G→C	c.230+1G→A	c.802_804delAAG	c.2146C→T	c.6449T→C
Effect on mRNA/protein	p.R196H	Skipping of exon 20, p.F1442KfsX3 p.V1441L	p.G2081A	Skipping of exon 3, p.L22QfsX3	p.K268del	p.Q716X	p.L2150P
Type of mutation	Missense	Splice site/Missense	Missense	Splice site	In-frame deletion	Nonsense	Missense
Gender	M	F	M	M	F	F	F
Year of birth	2001	2005	2001	2004	1988	2003	2003
Birth weight (g)	2.770	1.700	2.650	1.660	1.850	1.650	1.980
Length at birth (cm)	46	40	47	42	43	40	43,5
OFC at birth (cm)	32	29	N/A	27	29	24	29
APGAR score	9/10	8/9	9/10	9/9	5/10	N/A	8/10
IUGR	-	+	-	+	+	+	+
Postnatal growth retardation	+	+	+	+	-	+	+
Limb malformations	Brachydactyly, clinodactyly	Small hands	Syndactyly	Brachydactyly, feet syndactyly	Short fingers, *cubitus valgus*, flat feet	Clinodactyly, feet syndactyly, small hands	Hypomelia
Psychomotor delay	+	+	+	+	+	+	+
Mental retardation	+	+	+	+	+	N/A	N/A
Microcephaly	+	+	+	-	+	+	+
Hirsutism	+	-	+	+	-	+	+
Cardiovascular abnormality	ASD-OS	-	VSD	Heart murmur	-	Pulmonic stenosis	-
Gastroesophageal reflux	-	-	-	-	+	-	+
ENT-Hearing	-	-	Hearing loss	-	-	Hearing loss	Hearing loss
Genitourinay problems	-	-	Testicular cyst	Bilateral cryptorchidism	Polycystic ovary	-	Anteriorly placed anus
Craniofacial malformations	Arched palate	+	+	+	Cleft palate	+	Cleft palate
CNS alterations	-	Cortical-subcortical atrophy	-	-	-	-	-
Feeding problems in infancy	-	-	-	-	-	-	+
Seizures	-	+	-	-	-	-	-
Other findings	Hyperextensible joints, hyperactivity	-	Hyperactivity	-	Hyperandrogenism, insulin resistance	Atelectasia (deceased)	-

Patient	C20	C21	C25	C26	C28	C29	C30
Gene mutated	*NIPBL*	*NIPBL*	*NIPBL*	*NIPBL*	*NIPBL*	*NIPBL*	*SMC1A*
Exon	10	30	40	44	Intron 19	33	13
cDNA mutation*	c.2479_2480delAG[b]	c.5689_5691delAAT	c.6880C→T	c.7438_7439 delAG[C]	c.4320+5G→C	c.6269G→T	c.2132G→A
Effect on mRNA/protein	p.R827GfsX2	p.N1897del	p.Q2294X	p.R2480KfsX5	Skipping of exon 19 p. V1414_A1440 del	p.S2090I	p.R711Q
Type of mutation	Frameshift	In-frame deletion	Nonsense	Frameshift	Splice site	Missense	Missense
Gender	F	M	F	F	M	F	M
Year of birth	1998	1997	2000	2000	2008	2007	2005
Birth weight (g)	1.150	2.855	1.750	1.800	1.425	1.530	2.940
Length at birth (cm)	35,5	48	39,5	42	42	40	48
OFC at birth (cm)	25,5	31	30,5	N/A	27,7	28	33
APGAR score (1´/5´)	4/8	9/10	5/7	N/A	8/8	5/10	9/10
IUGR	+	-	+	+	+	+	+
Postnatal growth retardation	+	-	+	+	+	+	+
Limb malformations	Bilateral hypoplasia, feet syndactyly	Clinodactyly	5th finger (distally placed (bilateral)	Monodactyly (left hand), brachydactyly (right hand).	Oligodactyly, syndactyly, brachydactyly, clinodactyly	Camptodactyly, brachyclinodactyly, proximally placed thumbs	Feet syndactyly
Psychomotor delay	+	Language onset delay	+	+	+	+	+
Mental retardation	+	+	+	+	N/A	N/A	+
Microcephaly	+	+	+	-	+	+	+
Hirsutism	-	+	-	+	-	+	-
Cardiovascular abnormalities	ASD, persistent foramen ovale. Heart murmur, cardiomegaly	Heart murmur	Heart murmur	Heart murmur	-	-	Heart murmur
Gastroesophageal reflux	-	+	-	+	-	+	+ (severe)
ENT-Hearing	Hearing loss	Adenoid hypertrophy	-	Malformed internal auditory structures	Hearing loss, external auditory canal stenosis	Hearing loss	Hearing loss
Genitourinay problems	-	Unilateral cryptorchidism, renal cyst, hydrocele	-	Horseshoe kidney	Cryptorchidism	Hypoplasia labia minor	Pyelectasis, cryptorchidism
Craniofacial malformations	+	+	+	Arched palate	Cleft palate	Cleft palate	Arched palate
CNS alterations	Hyperechogeneicity of periventricular white substance	-	Hypotonia (mild)	-	-	Marked peritrigonal hyperechoecogene icity	-
Feeding problems in infancy	+	-	-	+	-	+	-
Seizures	-	-	-	+	-	-	+
Other findings	-	-	Pulmonary stenosis, limited elbow movements	Limited elbow movements	Limited elbow movements, lacrimonasal duct obstruction	Flat angiomata in neck	Umbilical hernia

Numbering is based on *SMC1A* and *NIPBL* cDNA sequences (RefSeq numbers NM_006306 and NM_133433, respectively), starting from the first nucleotide of the ORF. Nomenclature is according to den Dunnen and Antonarakis and to the Human Genome Variation Society Mutation Nomenclature Recommendations. In all these cases the parents were unaffected except C18 (adopted child).

a. Mutation previously reported by Borck et al. [2007] or Deardorff et al.[2007].

b. Mutation previously reported by Gillis et al. [2004], Kaur et al. [2005], Bhuiyan et al. [2006] or Selicorni et al. [2007].

c. Mutation previously reported by Yan et al. [2006].

(+) Present; (-) Not present; N/A: Not available; OFC: Occipito-Frontal Circumference; IUGR: Intrauterine Growth Retardation; ENT: Ear, Nose and Throat; CNS: Central Nervous System; ASD-OS: Atrial Septal Defect-*Ostium Secundum*; VSD: Ventricular Septal Defect.

Supplementary Tables II. Novel *NIPBL*, *SMC1A* and *SMC3* polymorphisms and variants of unknown significance identified.

Gene	Nucleotide change	Location	db SNP (frequency)	Carrier status of other family members	Control alleles identified	Estimated allele frequency
NIPBL:						
	c.3305-85delT	Intron 11		f	-	-
	c.4421+7A→G[#]	Intron 20		f	0/100	0.00
	c.4561-85C→T	Intron 21		f	-	-
	c.5011-62T→C	Intron 25	*rs16903455*	-	-	0.052
	c.5575-168A→T	Intron 29	*rs3100685*	-	-	0.024
	c.6108+7A→T[#]	Intron 34		*de novo*	0/100	0.00
	c.7263+153A→T	Intron 42	*rs300059*	-	-	0.152
SMC1A:						
	c.2197-5T→C[#]	Intron 13	*rs2297104*	f,m	0/100	0.00
	c.*14C→T[#]	3'UTR		m	0/98	0.00
SMC3:						
	c.92-193G→A	Intron 2	*rs7083749*	f,m	100/100	1.00
	c.92-128_127insGTT	Intron 2	*rs10658641*	f,m	96/100	0.96
	c.804+55C→G	Intron 10	-	-	6/100	0.06
	c.1092-64_62delATT	Intron 12	-	*	1/100	0.01
	c.1305+136A→T[#]	Intron 13	-	m	0/100	0.00
	c.1305+166_167insTC	Intron 13	-	f	11/100	0.11
	c.1306-159A→G	Intron 13	*rs2419572*	-	-	0.57
	c.3105+83G→T	Intron 25	*rs2039874*	f,m	94/98	0.96

Numbering is based on cDNA sequences for *SMC1A*, *SMC3* and *NIPBL* (RefSeq accession numbers: NM_006306, NM_005445 and NM_133433 respectively).
f= father, m= mother, -= no data, *= no parents available (adopted child), [#]= variants of unknown significance.

Supplementary Tables III. Previously reported *NIPBL*, *SMC1A* and *SMC3* polymorphisms.

Gene	Nucleotide Change	Location	db SNP	Reference
NIPBL:				
	c.3575-17A→G	Intron 13		Borck et al. [2004]
	c.3855+52A→G	Intron 16	*rs62654860*	Krantz et al. [2004], Gillis et al. [2004], Selicorni et al. [2007]
	c.4239+53T→C	Intron 18	*rs159753*	Gillis et al. [2004], Selicorni et al. [2007]
	c.4239+152C→G	Intron 18	*rs41270323*	Gillis et al. [2004]
	c.4321-35T→C	Intron 19		Gillis et al. [2004]
	c.4560+77A→G	Intron 21	*rs35011787*	Gillis et al. [2004]
	c.4560+108delT	Intron 21		Gillis et al. [2004]
	c.4561-9T→A	Intron 21		Gillis et al. [2004]
	c.4777-108delA	Intron 23		Gillis et al. [2004]
	c.4921-59G→A	Intron 24	*rs300060*	Gillis et al. [2004]
	c.5575-193T→C	Intron 29		Gillis et al. [2004]
	c.5710-78G→A	Intron 30		Gillis et al. [2004]
	c.5862+74delTT	Intron 32		Krantz et al. [2004], Gillis et al. [2004]
	c.5863-52delT	Intron 32		Gillis et al. [2004]
	c.5863-30delAT	Intron 32	*rs10554564*	Gillis et al. [2004]
	c.5863-12delAT	Intron 32	*rs10587827*	Krantz et al. [2004], Gillis et al. [2004]
	c.5874C→T (p.=)	Exon 33	*rs61748200*	Krantz et al. [2004], Gillis et al. [2004]
	c.6955-9delT	Intron 40		Gillis et al. [2004]
	c.7860+39G→A	Intron 45		Gillis et al. [2004]
	c.*282_285delACAA	3'UTR		Gillis et al. [2004]
SMC1A:				
	c.-19C→T	5'UTR	*rs1264011*	Deardorff et al. [2007]
	c.1338-32C→A	Intron 8	*rs1264008*	Deardorff et al. [2007]
SMC3:				
	c.-99C→A	5'UTR		Deardorff et al. [2007]
	c.15+89_90insA	Intron 1		Deardorff et al. [2007]
	c.91+67C→G	Intron 2	*rs4917577*	Deardorff et al. [2007]
	c.350+21T→A	Intron 6	*rs11195194*	Deardorff et al. [2007]
	c.350+30T→G	Intron 6	*rs7914351*	Deardorff et al. [2007]
	c.351-9T→C	Intron 6		Deardorff et al. [2007]
	c.547+92A→G	Intron 8	*rs7911129*	Deardorff et al. [2007]
	c.548-45A→C	Intron 8	*rs2275570*	Deardorff et al. [2007]
	c.548-4_3insTT	Intron 8		Deardorff et al. [2007]
	c.724-206_201delTTGTAG	Intron 9		Deardorff et al. [2007]
	c.724-5_6insT	Intron 9	*rs11380915*	Deardorff et al. [2007]
	c.805-26A→G	Intron 10	*rs11815960*	Deardorff et al. [2007]
	c.970-8G→A	Intron 11	*rs11195199*	Deardorff et al. [2007]
	c.1092-18T→C	Intron 12	*rs11195200*	Deardorff et al. [2007]
	c.1306-81A→G	Intron 13		Deardorff et al. [2007]
	c.1365T→C (p.=)	Exon 14		Deardorff et al. [2007]
	c.1410-48T→C	Intron 14	*rs3737293*	Deardorff et al. [2007]
	c.2116+23G→A	Intron 19	*rs7075340*	Deardorff et al. [2007]
	c.2644+48A→G	Intron 23	*rs11195213*	Deardorff et al. [2007]
	c.3039A→G (p.=)	Exon 25	*rs2419565*	Deardorff et al. [2007]
	c.3582+51G→A	Intron 28		Deardorff et al. [2007]

Numbering is based on cDNA sequences for *SMC1A*, *SMC3* and *NIPBL* (RefSeq accession numbers: NM_006306, NM_005445 and NM_133433 respectively).