Cover Page

The handle http://hdl.handle.net/1887/21013 holds various files of this Leiden University dissertation.

Author: Duyvenvoorde, Hermine van
Title: Genetic causes of growth disorders
Issue Date: 2013-06-25

7

Copy number variants in patients with short stature

Hermine A. van Duyvenvoorde[1,2,3], Julian C. Lui[4], Sarina G. Kant[3], Wilma Oostdijk[1], Antoinet C.J. Gijsbers[3], Mariëtte J.V. Hoffer[3], Marcel Karperien[5], Marie-José E. Walenkamp[6], Cees Noordam[7], Paul G. Voorhoeve[8], Verónica Mericq[9], Alberto M. Pereira[2], Hedi L. Claahsen-van de Grinten[7], Sandy A. van Gool[1], Martijn H. Breuning[3], Monique Losekoot[3], Jeffrey Baron[4], Claudia A.L. Ruivenkamp[3] and Jan-Maarten Wit[1]

[1] Department of Pediatrics, [2] Department of Endocrinology and Metabolic Diseases,
[3] Department of Clinical Genetics, Leiden University Medical Center, Leiden, The Netherlands
[4] Section on Growth and Development, National Institutes of Health, Bethesda, USA
[5] Department of Developmental BioEngineering, MIRA Institute for Biomedical Technology and Technical Medicine, University of Twente, Enschede, The Netherlands
[6] Department of Pediatrics, VU University Medical Center, Amsterdam, The Netherlands
[7] Department of Pediatrics, Radboud University Nijmegen Medical Center, Nijmegen, The Netherlands
[8] Department of Pediatrics, Canisius-Wilhelmina Hospital, Nijmegen, The Netherlands
[9] Institute of Maternal and Child Disease, University of Chile, Santiago, Chile

Submitted

Abstract

Height is a highly heritable and classic polygenic trait. Recent Genome-Wide Association studies (GWAS) have revealed that at least 180 genetic variants influence adult height. However, these variants explain only about 10% of the phenotypic variation in height. Genetic analysis of short individuals can lead to the discovery of novel rare gene defects with a large effect on growth.

In an effort to identify novel genes associated with short stature, genome-wide analysis for copy number variants (CNVs), using Single Nucleotide Polymorphism arrays, in 162 patients (149 families) with short stature was performed. Segregation analysis was performed if possible, and genes in CNVs were compared with information from GWAS, gene expression in rodents' growth plates, and published information.

CNVs were detected in 40 families. In six families a known cause of short stature was found (*SHOX* deletion or duplication, *IGF1R* deletion), in two combined with a *de novo* potentially pathogenic CNV. Thirty-three families had one or more potentially pathogenic CNVs (n = 40). In 24 of these families segregation analysis could be performed, identifying 3 *de novo* CNVs and 9 CNVs segregating with short stature. Four were located near loci associated with height in GWAS (*ADAMTS17*, *TULP4*, *PRKG2/BMP3* and *PAPPA*).

Besides six CNVs known to be causative for short stature, 40 CNVs with possible pathogenicity were identified. Segregation studies and bioinformatics analysis suggested various potential candidate genes.

Introduction

Height is a highly heritable and classic polygenic trait. In order to discover genes involved in growth regulation, there are basically two approaches. The first approach is to carry out genome-wide association studies (GWAS) for common variants in large populations of individuals. This has led to the discovery of at least 180 loci associated with adult height. However, the contribution of each locus is small, each locus contains various genes, and cumulative loci only explain about 10% of the phenotypic variation [1]. Alternatively, when using all Single Nucleotide Polymorphisms (SNPs) identified in a GWAS approach as predictors simultaneously, up to 40% of the variance in height can be explained [2]. The second approach is to perform genetic studies in patients with extremely short or tall height, and search for causative variants [3]. With this approach one can either test for gene defects that were previously described or that appear plausible based on observations in knockout mice (candidate gene approach), or perform a genome-wide analysis for copy number variants (CNVs) or whole exome sequencing (WES) for mutations. The candidate gene approach has led to the detection of a substantial number of genes that are involved in monogenic defects associated with short or tall stature, such as *IGF1*, *STAT5B*, *IGFALS*, and *IGF1R* [4-10], but obviously does not result in finding novel genes involved in growth regulation.

In two previous papers from our group [11,12] we have described the results of a candidate gene approach in children with short stature, either associated with a low birth size (small for gestational age, SGA) [13] or with a normal birth size (idiopathic short stature, ISS) [14]. In the present paper we describe the results of a genome-wide analysis for CNVs using SNP arrays in short children, in an effort to identify novel gene variants associated with short stature.

Subjects and Methods

Patients

We studied 191 patients from 173 unrelated families with short stature of unknown origin, either born with a normal birth size or born small for gestational age (SGA). DNA was sent to our laboratory for analysis because of short stature between 2008 and 2011. Twenty-nine were excluded from the present analysis: eight because of a height standard deviation score (SDS) > –2.0, fifteen because of insufficient or low quality DNA or no parental consent, and six cases belonging to one family were separately described with a heterozygous *IGF1* mutation and an additional 435.7 Kb deletion (arr 3q26.1(162,681,814–163,117,547)×1) [6]. This resulted in an analyzable group of 162 patients from 149 families. Height standard deviation score (SDS) was calculated for Dutch population references [15], except for one patient (I.6/II.2) for whom the reference for children of Turkish ethnicity was

used [16]. With consent of the Medical Ethical Committee of the Leiden University Medical Center, clinical data were collected and anonymized for all patients.

SNP arrays

In 103 cases the Affymetrix GeneChip Human Mapping 262K *NspI* or 238K *StyI* arrays (Affymetrix, Santa Clara, CA, USA) was used, containing 262,262 and 238,304 25–mer oligonucleotides, respectively, with an average spacing of approximately 12 kb per array. An amount of 250 ng DNA was processed according to the manufacturer's protocol. Detection of SNP copy number was performed using copy number analyzer for GeneChip (CNAG) version 2.0 [17].

In 54 cases the Illumina HumanHap300 BeadChip (Illumina Inc., San Diego, CA, USA) was used, containing 317,000 TagSNPs, with an average spacing of approximately 9 kb, and in 5 cases the Illumina HumanCNV370 BeadChip (Illumina Inc., Eindhoven, The Netherlands), containing 317,000 TagSNPs and 52,000 non-polymorphic markers for specifically targeting nearly 14,000 known CNVs. This array has an average spacing of approximately 7.7 kb. A total of 750 ng DNA was processed according to the manufacturer's protocol. SNP copy number (log R ratio) and B-allele frequency were assessed using Beadstudio Data Analysis Software Version 3.2 (Illumina Inc., Eindhoven, The Netherlands).

Evaluation of CNVs

Deletions of at least five adjacent SNPs and a minimum region of 150 kb and duplications of at least seven adjacent SNPs and a minimum region of 200 kb were evaluated [18], except for 3 families in which a prominent, but smaller duplication than 200 kb (although consisting of ≥ 10 adjacent SNP probes) was observed. The CNVs were classified into four groups: I, known pathogenic CNVs (known microdeletion or microduplication syndromes); II, potentially pathogenic CNVs, not described in the Database of Genomic Variants (DGV; The Centre for Applied Genomics, The Hospital for Sick Children, Toronto, Canada, http://projects.tcag.ca/variation/); III, CNVs not described in the DGV, but not containing any protein-coding genes; and IV, known polymorphic CNVs described in the DGV or observed in our in-house reference set, whereby at least three individuals must have been reported with the same rearrangement. Type IV CNVs were not further evaluated. All type II CNVs were assessed with Ensembl (Wellcome Trust Genome Campus, Hinxton, Cambridge, UK, http://www.ensembl.org: Ensembl release 63 – June 2011) and the DECIPHER database (Wellcome Trust Genome Campus, Hinxton, Cambridge, UK) for gene and microRNA (miRNA) content and similar cases, respectively. If DNA from the parents was available, segregation analysis was performed by SNP array.

The type I CNVs were confirmed with multiplex ligation-dependent probe amplification

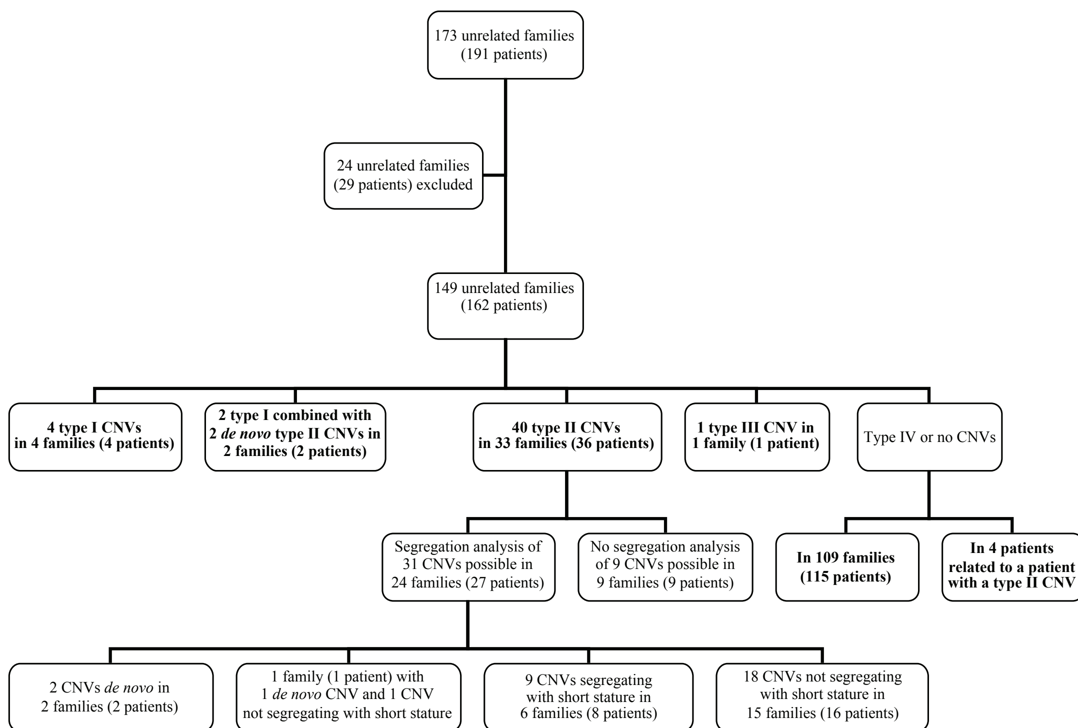

Figure 1 Organization Chart Organization chart illustrating the identified CNVs. The 149 unrelated families (162 patients) divided in the different subcategories are depicted in bold. A total of 49 CNVs were found in 40 families (43 patients).

(MLPA), using Salsa MLPA P018 probemix for *SHOX* and P217 for *IGF1R* analysis (MRC Holland, Amsterdam, The Netherlands). Amplification products were identified and quantified by capillary electrophoresis on an ABI 3130 genetic analyzer (Applied Biosystems, Nieuwerkerk aan de IJssel, The Netherlands). Fragment analysis was performed using GeneMarker (SoftGenetics, State College, USA). Thresholds for deletions and duplications were set at 0.75 and 1.25 respectively [19].

Bioinformatics approach

We checked for all CNVs whether they were located in one of the chromosomal regions associated with height in GWAS [1]. For genes in deleted or duplicated regions in cases with *de novo* CNVs, we used three additional approaches. First, the rodent homologues were checked for three criteria: 1) higher expression in 1 week old mouse growth plate than in 1 week old mouse lung, kidney, and heart; 2) spatial regulation: significant difference between zones in the 1 week old rat growth plate; and 3) temporal regulation: significant difference between 3 and 12 weeks of age in the rat growth plate using previously established mRNA expression profiles [20,21]. Second, associations were investigated for mouse growth plate-related phenotypes. Third, associations with human growth plate-related phenotypes were investigated. For details, see Lui *et al.* [21].

Results

CNVs

An organization chart illustrating the identified CNVs is shown in figure 1. In the 162 patients belonging to 149 unrelated families, a total of 49 CNVs were found in 40 families (43 patients).

In six families (4.0%, 6 patients) a type I CNV was observed and in two of them an additional *de novo* type II CNV. Table 1 shows the clinical and genetic findings of these 6 patients, including 2 microdeletions (I.1 and I.2) and 2 microduplications (I.3 and I.4) containing *SHOX*, and two terminal 15q deletions containing *IGF1R* (I.5/II.1/mi.3 and I.6/II.2). All these CNVs were confirmed with MLPA.

One or more type II CNVs (n = 40) were found in 33 unrelated families (22.1%, 36 patients). Five of these potentially pathogenic CNVs contained besides protein-coding genes also miRNAs (Table 2). In 24 families (27 patients) segregation analysis could be performed, which led to a total of 5 *de novo* CNVs (Table 3) and 9 CNVs segregating with a height below –1.5 SDS of a carrier family member (Table 4). For 19 CNVs the lack of segregation with short stature makes a causative role of the CNV unlikely (Supplementary Table 1). In 9 patients (9 CNVs) no information on segregation could be obtained (Supplementary Table 2). In two non-related patients (cases II.24 and II.25) a similar CNV (a deletion containing *DCAF12L2*, alias *WDR40C*) in the X-chromosome was identified, but both children inherited the deletion from a normal parent.

In one family (0.7%, 1 patient) a type III CNV was found encompassing a 192.3 Kb deletion of chromosome 13 (arr 13q31.1(86,733,645–86,925,974)×1). The girl (case III.1) was born SGA, had poor food intake and severe postnatal growth failure (length –8.2 SDS at 2.5 years). Screening for *IGF1* and the *IGF1R* for mutations or deletions was negative. The function of this region is unknown.

No potential pathogenic CNVs (only type IV or no CNVs) were found in 109 families (73.2 %, 119 patients).

Bioinformatics approach

Five CNVs encountered in our study are close to the loci associated with height in GWAS [1]. Four of these CNVs were *de novo* or segregating with short stature, including loci close to *ADAMTS17* (case II.5), *PRKG2/BMP3* (cases II.11 and II.13), *PAPPA* (cases II.11 and II.13) and *TULP4* (case II.7). However, none of the deletions included genes tightly linked ($r^2 < 0.5$) to a GWAS SNP implicated in human height variations. The fifth CNV is close to the *MKL2* locus (case II.37/mi.4) but did not segregate with short stature (Supplementary Table 1).

Table 1 Type I CNVs

ID M/F	Height (SDS)	Karyotype (ISCN 2009)	Size (Mb)	Protein-coding genes[a]	Additional CNV (type)	Known gene
I.1 F	–2.9	arr Xp22.33(1–1,522,908)×1 mat	1.32	9 protein-coding genes; from *PLCXD1* to *ASMTL*	–	*SHOX*
I.2 F	–4.0	arr Xp22.33(1–2,320,027)×1 dn	2.12	13 protein-coding genes; from *PLCXD1* to *DHRSX*	–	*SHOX*
I.3 M	–2.3	46,XY,t(8;13)(q13;q12).arr Xp22.33(1–727,565)×2 mat	0.52	*PLCXD1* *GTPBP6* *PPP2R3B* *SHOX*	–	*SHOX*
I.4 M	–2.8	46,X,psu idic(Y)(q11.22) dn.arr Yp11.32p11.31(1–2,640,827)×2 dn	2.49	16 protein-coding genes; from *PLCXD1* to *XG*	–	*SHOX*
I.5/II.1/mi.3 F	–3.1	arr 15q26.2q26.3(98,374,491–102,531,392)×1 dn	4.00	23 protein-coding genes; from *ARRDC4* to *OR4F15*	Gain 15q26.1q26.2 dn (Type II)	*IGF1R*
I.6/II.2 M	–5.9	arr 15q26.3(99,131,989–102,531,392)×1 dn	3.24	21 protein-coding genes; from *IGF1R* to *OR4F15*	Gain 9p24.3p24.2 dn (Type II)	*IGF1R*

dn = *de novo*; mat = maternally inherited; pat = paternally inherited.
[a] For CNVs containing ≤ 5 protein-coding genes, all protein-coding genes are depicted. For CNVs containing ≥ 6 protein-coding genes, the number, and the first and last protein-coding gene is given.

Table 2 miRNAs

ID M/F	Height (SDS)	Karyotype (ISCN 2009)	Size (Kb)	Protein-coding genes[a]	miRNA	Additional CNV (type)
II.19/mi.1 F	–4.6	arr 7q36.3(158,183,050–158,692,049)×3	509.0	*PTPRN2* *NCAPG2* *ESYT2* *WDR60*	MIR595	–
II.32/mi.2 F	–2.8	arr 8p23.1(7,690,325–9,040,305)×3 pat, 8p23.1p22(12,242,033–13,046,661)×3 pat	8p23.1: 1,350.0 8p23.1p22: 804.6	Chr8p23.1: 10 protein-coding genes; from *DEFB104A* to *PPP1R3B* Chr8p23.1p22: *FAM86B2* *LONRF1* *KIAA1456* *DLC1*	8p23.1: MIR548I3	–
I.5/II.1/mi.3 F	–3.1	arr 15q26.1q26.2(91,199,026–98,456,575)×3 dn	7 257.6	19 protein-coding genes; from *BLM* to *SPATA8*	MIR1469	Loss 15q26.2q26.3 (type I)
II.37/mi.4 F	–2.5	arr 16p13.12p13.11(14,760,735–16,633,360)×1 pat	1 872.6	17 protein-coding genes; from *BFAR* to *NOMO3*	MIR1972-1[b] MIR484[b]	–
II.22/mi.5 M	–2.1	arr 22q11.21(21,011,217–21,928,915)×1	917.7	16 protein-coding genes; from *POM121L4P* to *UBE2L3*	MIR649[b]	–

dn = *de novo*; mat = maternally inherited; pat = paternally inherited.

[a] For CNVs containing ≤ 5 protein-coding genes, all protein-coding genes are depicted. For CNVs containing ≥ 6 protein-coding genes, the number, and the first and last protein-coding gene in the CNV are given.

[b] miRNA 484, 649, and 1972 have been predicted to bind to various isoforms of SHOX, accordingly contributing to the regulation of SHOX expression [45].

We reasoned that some of the identified CNVs might cause short stature because they contain genes that are expressed and function in the growth plate. We therefore used existing expression microarray data to identify genes that show greater expression in mouse growth plate than in soft tissues, temporal regulation in rat growth plate, or spatial regulation in rat growth plate. Within *de novo* CNVs, this approach implicated 5 genes (*Aldh1a3*, *Fam3c*, *Furin*, *Lrrk1*, and *Chsy1*), and within segregating CNVs, this implicated 7 genes (*Col14A1*, *Dscc1*, *Enpp2*, *Ezr*, *Prelid2*, *Taf2*, and *Trim32*) (Table 5). This information, in combination with other bioinformatic data, was used to formulate the arguments pro and contra an association of these genes with short stature (summarized in Tables 3 and 4). Potential candidate genes in *de novo* CNVs associated with short stature (Table 3) include *FURIN*, *DOCK8* and/or *KANK1*, *NLRP3*, *FAM3C*, *SLC13A1*, *ADAMTS17*, *ALDH1A3*, *LRRK1* and *CHSY1*. Potential candidate genes in CNVs segregating with short stature (Table 4) include *FHIT*, *PTPRG*, *TULP4*, *EZR*, *ENPP2*, *TAF2*, *COL14A1*, *DSCC1*, *LPPR1*, *ZNF675*, *C4orf22* (or *PRKG2/BMP3*), *PRELID2*, and *ASTN2* and *TRIM32* (or *PAPPA*).

For the CNVs for which insufficient information was available about segregation with short stature, the *in silico* analysis provided support for four potential candidate genes (*TBL1X*, *ROBO2*, *CHD8* and *TOX4*), as well as a candidate region (distal part of common 22q11 deletion syndrome) (Supplementary Table 2).

Table 3 *De novo* type II CNVs

ID M/F	Height (SDS)	Karyotype (ISCN 2009)	Size (Kb)	Protein-coding genes[a]	Arguments pro pathogenicity	Arguments against pathogenicity
I.5/II.1/ mi.3 F	−3.1	Type II: arr 15q26.1q26.2(91,199,026−98,456,575)×3 dn Type I: arr 15q26.2q26.3(98,374,491−102,531,392)×1 dn	7 257.6	19 protein-coding genes; from *BLM* to *SPATA8*, including MIR1469	*Furin* higher expressed in murine GP and upregulated from PZ to HZ.	*IGF1R* deletion can explain short stature [26].
I.6/II.2 M	−5.9	Type II: arr 9p24.3p24.2(1−2,612,433)×3 dn Type I: arr 15q26.3(99,131,989−102,531,392)×1 dn	2 612.4	9 protein-coding genes; from *FOXD4* to *SMARCA2*	2 short children with overlapping 9p duplication (*DOCK8* and *KANK1*, DECIPHER #256751 and #261831). Shorter than usual for IGF1R deletion [26].	*Dock8* and *Kank1* not overexpressed in murine GP.
II.3 M	−2.4	arr 1q44(246,715,197−247,652,602)×3 dn, 2q24.3(165,611,363−165,769,050)×3 pat	Chr1: 937.4 Chr2: 157.7	Chr1: 12 protein-coding genes; from *TFB2M* to *OR2B11* Chr2: *COBLL1* *SLC38A11*	Activating *NLRP3* mutations associated with short stature (NOMID). Constitutively activated *Nlrp3* in mice causes growth retardation [46].	*NLRP3* duplication described in three patients without short stature, with overlapping, smaller duplications inherited from a normal parent (DECIPHER #263423, #258032 and #253572), *Nlrp3* not overexpressed in murine GP.

II.4 M	−3.5	arr 7q31.31q31.32(119,770,125–123,600,606)×1 dn	3 830.5	21 protein-coding genes; from *KCND2* to *SPAM1*	*Fam3c* higher expressed in murine GP and downregulated from RZ to PZ. Homozygous mutations in *Slc13a1* in sheep and mice cause dwarfism [27,28].	Only 2 out of 9 patients with bigger overlapping deletions reported with short stature (DECIPHER).
II.5 M	−3.5	arr 15q26.3(101,003,122–102,374,592)×1 dn	1 371.5	13 protein-coding genes; from *CERS3* to *OR4F15*	*ADAMTS17* associated with height in human and dog (GWAS) [1,29]. Short child with overlapping 15q deletion (DECIPHER #251400). Mutations cause chondrodysplasia [30-33]. Associated with fibrillin-1 function [31,33]. *Aldh1a3* and *Lrrk1* higher expressed in murine GP; *Chsy1* highly expressed in HZ and downregulated with age.	Deletion is located 244 Kb downstream of the *ADAMTS17* locus.

dn = *de novo*; mat = maternally inherited; pat = paternally inherited.
GP = Growth Plate; RZ = Resting zone; PZ = Proliferative zone; HZ = Hypertrophic zone.
[a] For CNVs containing ≤ 5 protein-coding genes, all protein-coding genes are depicted. For CNVs containing ≥ 6 protein-coding genes, the number, and the first and last protein-coding gene in the CNV are given.

Table 4 Type II CNVs segregating with short stature

ID M/F	Height (SDS)	Karyotype (ISCN 2009)	Size (Kb)	Protein-coding genes[a]	Arguments pro pathogenicity	Arguments against pathogenicity
II.6 M	–2.0	arr 3p14.2(59,235,764–61,832,828)×3 pat	2 597.1	*FHIT* *PTPRG*	*FHIT* acts as a repressor of beta-catenin transcriptional activity [34]. *PTPRG* possibly inhibits cell growth [35]. Height father –1.8 SDS.	Not overexpressed in murine GP.
II.7 M	–2.9	arr 6q25.3(159,026,380–159,929,652)×3 mat	903.3	8 protein-coding genes; from *TMEM181* to *FNDC1*	*TULP4* associated with height (GWAS) [1]. *Ezr* downregulated with age in murine GP.	Duplication is located 94 Kb downstream of the *TULP4* locus. Height mother –1.5 SDS.
II.8 M	–2.9	arr 8q24.12(120,463,609–121,849,380)×3 mat	1 385.8	8 protein-coding genes; from *ENPP2* to *SNTB1*	*ENPP2* encodes for a lysophospholipase D, producing lysophosphatidic acid involved in cell proliferation [36]. *Enpp2* highly expressed in murine kidney and GP, and highly upregulated from PZ to HZ. *Taf2* upregulated from PZ to HZ. *Col14a1* downregulated with age in murine GP and upregulated from RZ to PZ. *Dscc1* higher expressed in murine GP and downregulated from PZ to HZ.	Height mother –1.6 SDS.

II.9 F	−2.5	arr 9q31.1(103,493,752–104,059,876)×1 mat	566.1	*LPPR1*(*PRG3*)	*Prg1* knockout mice are small [37]. Maternal height −3.3 SDS. Height non-carrier maternal relative −0.1 SDS.	Height carrier maternal sister −1.7 SDS. Height non-carrier maternal relative −2.3 SDS. *LPPR1* not overexpressed in murine GP.
II.10 M	−3.7	arr 19p12(23,661,801–24,041,650)×1 pat	379.8	*ZNF675* *ZNF681* *RPSAP58*	*ZNF675* possibly involved in osteoclast differentiation [38]. Paternal height −2.6 SDS.	Not overexpressed in murine GP.
II.11[b] F	−3.0	arr 1p31.1(72,546,864–72,940,272)×1, 4q21.21(81,301,396–81,775,934)×1, 5q32(145,109,219–145,250,730)×3, 9q33.1(119,411,013–119,601,157)×1	Chr1: 393.4 Chr4: 494.4 Chr5: 141.5 Chr9: 190.1	Chr1: *NEGR1* Chr4: *C4orf22* Chr5: *PRELID2* *GRXCR2* Chr9: *ASTN2* *TRIM32*	*PRKG2/BMP3* (4q21.21) and *PAPPA* (9q33.1) associated with height (GWAS) [1]. *Prelid2* higher expressed in murine GP, downregulated with age and from PZ to HZ. *Trim32* downregulated from PZ to HZ.	4q21.21 Deletion is located 176 Kb upstream of the *PRKG2/BMP3* locus. 9q33.1 Deletion is located 246 Kb downstream of the *PAPPA* locus.
II.12[b] M	−2.0	arr 5q32(145,109,219–145,250,730)×3 mat	Chr5: 141.5	Chr5: *PRELID2* *GRXCR2*		
II.13[b] M	−1.5	arr 4q21.21(81,301,396–81,775,934)×1 mat, 5q32(145,109,219–145,250,730)×3 mat, 9q33.1(119,411,013–119,601,157)×1 mat	Chr4: 494.4 Chr5: 141.5 Chr9: 190.1	Chr4: *C4orf22* Chr5: *PRELID2* *GRXCR2* Chr9: *ASTN2* *TRIM32*		

dn = *de novo*; mat = maternally inherited; pat = paternally inherited.
GP = Growth Plate; RZ = Resting zone; PZ = Proliferative zone; HZ = Hypertrophic zone.
[a] For CNVs containing ≤ 5 protein-coding genes, all protein-coding genes are depicted. For CNVs containing ≥ 6 protein-coding genes, the number, and the first and last protein-coding gene in the CNV are given.
[b] Family; mother and 2 sons.

Table 5 Bioinformatic approach (mouse GP vs soft tissues expression, and spatial and temporal regulation of gene expression in the rat GP)

Gene	Growth Plate vs Soft Tissues (Mouse Array)						Growth Plate, 3 vs 12 wk; RZ vs PZ and PZ vs HZ at 1 wk (Rat Array)					
	GP vs Heart (FC)	p-value[a]	GP vs Kidney (FC)	p-value[b]	GP vs Lung (FC)	p-value[c]	3 vs 12 wk (FC)	p-value[d]	RZ vs PZ (FC)	p-value[e]	PZ vs HZ (FC)	p-value[f]
***De novo* CNVs**												
Aldh1a3 [(S)]	15.9	<0.001	1.3	0.2	20.7	<0.001	1.3	0.2	1.0	0.7	**32.7**	**<0.001**
Fam3c [(G,S)]	**3.2**	**<0.001**	**3.0**	**<0.001**	**1.9**	**<0.001**	1.2	0.002	**–2.2**	**<0.001**	1.2	0.2
Furin [(G,S)]	**2.2**	**<0.001**	**1.9**	**<0.001**	**2.2**	**<0.001**	–1.4	0.03	–1.2	0.008	**1.6**	**<0.001**
Lrrk1 [(S)]	3.4	<0.001	2.2	<0.001	1.4	0.002	1.2	0.2	–1.4	0.02	**1.9**	**0.001**
Chsy1 [(T,S)]	no probe in mouse array						**–2.5**	**<0.001**	–1.2	0.3	**2.0**	**0.004**
Segregating CNVs												
Col14a1 [(T,S)]	–14.1	<0.001	–10.9	<0.001	–6.3	<0.001	**–2.5**	**<0.001**	**5.1**	**0.001**	–2.4	0.02
Dscc1 [(G,S)]	**2.4**	**<0.001**	**2.3**	**<0.001**	**2.0**	**<0.001**	1.2	0.2	–1.7	0.02	**–3.2**	**<0.001**
Enpp2 [(S)]	7.4	<0.001	–1.8	<0.001	2.3	<0.001	1.1	0.6	–1.2	0.4	**27.3**	**<0.001**
Ezr [(T)]	–6.8	<0.001	–15.4	<0.001	–9.1	<0.001	**–1.6**	**0.001**	–1.4	0.1	1.5	0.07
Prelid2 [(G,T,S)]	**2.4**	**<0.001**	**4.1**	**<0.001**	**6.0**	**<0.001**	**–2.8**	**<0.001**	1.1	0.3	**–1.9**	**<0.001**
Taf2 [(G,S)]	**2.1**	**<0.001**	**1.6**	**<0.001**	**1.5**	**<0.001**	–1.2	0.1	1.0	1.0	**2.2**	**0.005**
Trim32 [(S)]	–2.2	0.007	–2.7	0.001	–1.9	0.02	1.2	0.2	-1.3	0.1	**–2.0**	**0.002**

GP = Growth Plate; FC = Fold change; RZ = Resting zone; PZ = Proliferative zone; HZ = Hypertrophic zone.
[a] P<0.0048 considered statistically significant (False Discovery Rate (FDR) <0.01).
[b] P<0.0048 considered statistically significant (FDR<0.01).
[c] P<0.0047 considered statistically significant (FDR<0.01).
[d] P<0.0042 considered statistically significant (FDR<0.05).
[e] P<0.0017 considered statistically significant (FDR<0.05).
[f] P<0.0086 considered statistically significant (FDR<0.05).
[(G)] Growth plate specific gene, defined as expression in GP vs Soft Tissue ≥ 1.5–fold and FDR<0.01 for all three soft tissues.
[(T)] Temporally regulated gene, defined as 3 vs 12 wk ≥ ±1.5–fold, FDR<0.05.
[(S)] Spatially regulated gene, defined as RZ vs PZ ≥ ±1.5–fold, FDR<0.05; and/or PZ vs HZ ≥ ±1.5–fold, FDR<0.05.
Genes fulfilling the criteria mentioned above are depicted in bold.

Discussion

Whole genome SNP array analysis in 162 patients with short stature from 149 unrelated families (Fig. 1) led to the detection of type I CNVs known to cause short stature (involving *SHOX* or *IGF1R*) in six families (in two of them combined with type II CNVs), and 40 potentially pathogenic CNVs (type II) in 33 families. Out of the total of 42 type II CNVs, five were *de novo* and nine others were associated with short stature in their families. In one severely short child a deletion without protein-coding genes was found, and in 5 CNVs 6 microRNAs were encountered.

A recent study on a genome-wide association analysis of copy-number variation and stature showed that children with short stature had a greater global burden of lower-frequency and rare deletions and a greater average CNV length than controls [22]. There were no significant associations with tall stature. These observations suggest that CNVs might contribute to genetic variation in stature in the general population. These authors also identified three preliminary candidate regions as having significant associations with stature; a duplication at 11q11 and deletions at 14q11.2 and 17q21.31. In our analysis these regions all display common CNVs, which have been often observed in our in-house database and in the DGV (type IV CNVs).

The two patients carrying a heterozygous deletion containing the *SHOX* gene had disproportionate short stature, but no Madelung deformity. Case I.1 (sitting height/height (SH/H) ratio +3.7 SDS) inherited the deletion from her mother, who also had disproportionate short stature (height –1.8 SDS, SH/H ratio +4.2 SDS). Case I.2 (SH/H ratio +3.8 SDS) carries besides a *de novo SHOX* haploinsufficiency also a heterozygous unclassified variant (UV) in the *IGFALS* gene (c.1555C>T, p.Arg519Trp) inherited from her father (height –1.1 SDS). IGFALS sequencing was performed because of a low circulating IGF-I and IGFBP-3 despite elevated GH secretion. While the referring physician had not suspected Leri-Weill syndrome, in retrospect the increased sitting height/height ratio would have been sufficient reason to directly test for *SHOX* defects. The two patients in whom a duplication of the *SHOX* gene including surrounding genes was observed (*de novo* and inherited via a normal statured parent, respectively), had a sitting height/height ratio of approximately +1.9 SDS. We and others have recently reported that a phenotype similar to Leri-Weill syndrome (including short stature) can be associated with *SHOX* duplication [11,23,24].

In two patients a heterozygous deletion on chromosome 15 containing the *IGF1R* gene was identified, a well-established cause of short stature [11,25,26]. In both patients an additional *de novo* CNV was present (Table 3). In case I.5/II.1/mi.3 this was a duplication in 15q26.1q26.2 (located upstream of the deleted area). Although this patient's growth failure is similar to that of other patients with IGF1R defects [26], duplication of *FURIN* may play an additional

chapter 7

role. In case I.6/II.2, considerably shorter than usual for *IGF1R* deletions [26], the terminal 15q deletion was combined with a terminal 9p24.3p24.2 duplication, suggesting the presence of an unbalanced reciprocal translocation. We suspect that one of the parents is a carrier of a balanced 9;15 translocation, but unfortunately parental chromosomes were not available for testing. The presence of two patients in the DECIPHER database with a similar 9q duplication and short stature suggests that there may be an association between the genes *DOCK8* and *KANK1*, and stature.

Bioinformatics analysis of the three other cases with *de novo* type II CNVs led to several candidate genes (Table 3). In case II.3 a duplication of *NLRP3* may be associated with short stature. The CNV in case II.4 (who has besides short stature also mental retardation, behavioral problems, strabismus, and various dysmorphic features) suggests that *FAM3C* and *SLC13A1* deletions may be associated with short stature, particularly because of the expression data of *Fam3c* in the murine growth plate and the dwarfism and skeletal deformities in Texel sheep and mice with loss of function of *Slc13a1* [27,28].

Case II.5, with a terminal *de novo* 15q deletion located 1.5 Mb downstream of *IGF1R* and 244 Kb downstream of the *ADAMTS17* locus on the reverse strand, had a normal birth size, but showed proportionate progressive growth failure (SH/H ratio +1.58 SDS) with a normal head circumference. Clinical characteristics included slight frontal bossing of the skull, a high pitched voice and slight abdominal adiposity and delayed bone age. GH secretion and circulating IGF-I were normal, but IGFBP-3 was low (–2 SDS). Several arguments are in favor of a role of *ADAMTS17* in growth regulation (for summary, see Table 3), including: 1) significant association with height in population GWAS [1]; 2) a short child with a similar terminal deletion in the DECIPHER database; 3) significant association with size in a GWAS in the domestic dog [29]; 4) human mutations in *ADAMTS17* causing the acromelic chondrodysplasia Weill-Marchesani-like syndrome (OMIM #277600 and #608328) [30-33]; and 5) association of members of the ADAMTSL/ADAMTS family with the modulation of fibrillin-1 function [31,33]. Unfortunately, expression of the rodent homologue of *ADAMTS17* could not be investigated, because the gene was not represented on the microarrays used. Besides *ADAMTS17*, this deletion contains three other genes, *ALDH1A3*, *LRRK1* and *CHSY1*, that might be implicated in short stature.

Nine CNVs in 6 families (5 families with one index patient each, and one family consisting of a mother and her 2 sons) segregated with a height of less than –1.5 SDS of a carrier family member (Table 4). The 3p duplication that case II.6 (height –2.0 SDS) inherited from his father (–1.8 SDS) contains *FHIT* and the first part of *PTPRG*. Both genes are considered tumor suppressors [34,35]. The 6q duplication that case II.7 inherited from his mother is located nearby (97 Kb downstream) a locus (*TULP4)* associated with height [1]. One of the duplicated genes (*ENPP2*) in case II.8 encodes for a lysophospholipase D, producing lysophosphatidic

acid (LPA) inducing cell proliferation [36]. The mouse homologs of *TAF2*, *COL14A1* and *DSCC1* are differentially expressed in the growth plate. In case II.9, the 9q deletion containing part of *LPPR1* (also known as *PRG3*) did not fully segregate with short stature in the family, but the observation that *Prg1* knockout mice are smaller compared to wild type littermates [37] suggests a role for this gene in height regulation. The 19p deletion that case II.10 inherited from his father includes *ZNF675*, associated with osteoclast differentiation [38]. Out of the four CNVs in cases II.11, 12 and 13 (the short members of one family), *C4orf22*, *ASTN2* and *TRIM32* are located close to loci (374 KB upstream *PRKG2/BMP3* and 289 Kb downstream *PAPPA*, respectively) associated with height [1], suggesting that the 4q and/or 9q deletion are associated with stature.

Four out of nine patients in whom no segregation analysis could be performed (Supplementary Table 2) carry a CNV suggestive for an association with short stature. One of the genes in the duplication of case II.14 is *TBL1X* (alias *TBL1*), encoding for transducin beta-like protein 1 (TBL1). TBL1 and its highly related family member *TBLR1* are required for Wnt-beta-catenin-mediated transcription [39]. Case II.17, described previously [12], carries a duplication of 3p12.3 containing part of *ROBO2*, as well as his younger brother (height –4.3), but his mother (height –3.6 SDS) does not carry the variant, while no DNA is available from the father. The encoded protein is a receptor for SLIT2 and probably SLIT1, which are thought to function in axon guidance and cell migration [40]. Case II.21 was born SGA, and at 1.2 years her length was –3.7 SDS and head circumference –3.1 SDS. Further clinical characteristics include clinodactily, a protruded tongue and delayed bone age. The mother does not carry this duplication, and DNA from her father is not available. A search in the DECIPHER database revealed 2 patients with (partially) overlapping duplications, one of whom was short (patient #258583) and one was not (#258497). Out of the 6 genes outside the overlapping region with patient #258497 *CHD8* and *TOX4* appear potential candidate genes [41,42]. Case II.22/mi.5 has a 22q deletion containing only the distal part of the common 22q11 deletion syndrome (Velocardiofacial/DiGeorge syndrome). His mother does not carry the duplication, and DNA from the father is not available. In 8 patients in the DECIPHER database with overlapping deletions short stature was observed. The common deleted region contains *PI4KA*, *SERPIND1*, *SNAP29*, *CRKL*, *AIFM3*, *LZTR1*, *THAP7*, and *P2RX6*.

Although non-coding DNA can play an important regulatory role [43,44], no supportive evidence could be obtained on a possible role of novel type III CNVs. Similarly, none of the 6 miRNAs (Table 2) identified in the type II CNVs could be directly linked to short stature, due to lack of segregation with short stature (data not shown), although miRNA 484, 649, and 1972 have been predicted to bind to various isoforms of SHOX, and contribute to the regulation of SHOX expression [45].

In conclusion, whole genome SNP array analysis in this exploratory study on 162 patients

with short stature belonging to 149 unrelated families identified 6 CNVs in 6 families (4%) for which the association with short stature is virtually certain, and 40 CNVs in 33 families (22.1%) with possible pathogenicity. Several of the deleted or duplicated genes may be considered as potential candidate genes for growth disorders, including four genes associated with height in the genome-wide association studies (*ADAMTS17*, *PRKG2/BMP3*, *PAPPA*, *TULP4*). Future studies are needed to support the role of these and other genes in longitudinal growth regulation.

Acknowledgments
We are grateful to the patients and their parents for participating in this study, and to the physicians for referring the patients.

chapter

7

Reference List

1 Lango Allen H, Estrada K, Lettre G *et al*: Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* 2010; **467**: 832-838.

2 Yang J, Benyamin B, McEvoy BP *et al*: Common SNPs explain a large proportion of the heritability for human height. *Nat Genet* 2010; **42**: 565-569.

3 Cirulli ET, Goldstein DB: Uncovering the roles of rare variants in common disease through whole-genome sequencing. *Nat Rev Genet* 2010; **11**: 415-425.

4 Kant SG, Kriek M, Walenkamp MJ *et al*: Tall stature and duplication of the insulin-like growth factor I receptor gene. *Eur J Med Genet* 2007; **50**: 1-10.

5 van Duyvenvoorde HA, Kempers MJ, Twickler TB *et al*: Homozygous and heterozygous expression of a novel mutation of the acid-labile subunit. *Eur J Endocrinol* 2008; **159**: 113-120.

6 van Duyvenvoorde HA, van Setten PA, Walenkamp MJ *et al*: Short stature associated with a novel heterozygous mutation in the insulin-like growth factor 1 gene. *J Clin Endocrinol Metab* 2010; **95**: E363-E367.

7 Vidarsdottir S, Walenkamp MJ, Pereira AM *et al*: Clinical and biochemical characteristics of a male patient with a novel homozygous STAT5b mutation. *J Clin Endocrinol Metab* 2006; **91**: 3482-3485.

8 Walenkamp MJ, Karperien M, Pereira AM *et al*: Homozygous and heterozygous expression of a novel insulin-like growth factor-I mutation. *J Clin Endocrinol Metab* 2005; **90**: 2855-2864.

9 Walenkamp MJ, van der Kamp HJ, Pereira AM *et al*: A variable degree of intrauterine and postnatal growth retardation in a family with a missense mutation in the insulin-like growth factor I receptor. *J Clin Endocrinol Metab* 2006; **91**: 3062-3070.

10 Woods KA, Camacho-Hubner C, Savage MO, Clark AJ: Intrauterine growth retardation and postnatal growth failure associated with deletion of the insulin-like growth factor I gene. *N Engl J Med* 1996; **335**: 1363-1367.

11 Caliebe J, Broekman S, Boogaard M *et al*: IGF1, IGF1R and SHOX mutation analysis in short children born small for gestational age and short children with normal birth size (idiopathic short stature). *Horm Res Paediatr* 2012; **77**: 250-260.

12 Wit JM, van Duyvenvoorde HA, Scheltinga SA *et al*: Genetic analysis of short children with apparent growth hormone insensitivity. *Horm Res Paediatr* 2012; **77**: 320-333.

13 Clayton PE, Cianfarani S, Czernichow P, Johannsson G, Rapaport R, Rogol A: Management of the child born small for gestational age through to adulthood: a consensus statement of the International Societies of Pediatric Endocrinology and the Growth Hormone Research Society. *J Clin Endocrinol Metab* 2007; **92**: 804-810.

14 Wit JM, Clayton PE, Rogol AD, Savage MO, Saenger PH, Cohen P: Idiopathic short stature: definition, epidemiology, and diagnostic evaluation. *Growth Horm IGF Res* 2008; **18**: 89-110.

15 Fredriks AM, van Buuren S, Burgmeijer RJ *et al*: Continuing positive secular growth change in The Netherlands 1955-1997. *Pediatr Res* 2000; **47**: 316-323.

16 Fredriks AM, van Buuren S, Jeurissen SE, Dekker FW, Verloove-Vanhorick SP, Wit JM: Height, weight, body mass index and pubertal development reference values for children of Turkish origin in the Netherlands. *Eur J Pediatr* 2003; **162**: 788-793.

17 Nannya Y, Sanada M, Nakazaki K *et al*: A robust algorithm for copy number detection using high-density oligonucleotide single nucleotide polymorphism genotyping arrays. *Cancer Res* 2005; **65**: 6071-6079.

18 Hehir-Kwa JY, Egmont-Petersen M, Janssen IM, Smeets D, van Kessel AG, Veltman JA: Genome-wide copy number profiling on high-density bacterial artificial chromosomes, single-nucleotide polymorphisms, and oligonucleotide microarrays: a platform comparison based on statistical power analysis. *DNA Res* 2007; **14**: 1-11.

19 Lalic T, Vossen RH, Coffa J *et al*: Deletion and duplication screening in the DMD gene using MLPA. *Eur J Hum Genet* 2005; **13**: 1231-1234.

20 Lui JC, Andrade AC, Forcinito P *et al*: Spatial and temporal regulation of gene expression in the mammalian growth plate. *Bone* 2010; **46**: 1380-1390.

21 Lui JC, Nilsson O, Chan Y *et al*: Synthesizing genome-wide association studies and expression microarray reveals novel genes that act in the human growth plate to modulate height. *Hum Mol Genet* 2012; **21**: 5193-5201.

22 Dauber A, Yu Y, Turchin MC *et al*: Genome-wide association of copy-number variation reveals an association between short stature and the presence of low-frequency genomic deletions. *Am J Hum Genet* 2011; **89**: 751-759.

23 Benito-Sanz S, Barroso E, Heine-Suner D *et al*: Clinical and molecular evaluation of SHOX/PAR1 duplications in Leri-Weill dyschondrosteosis (LWD) and idiopathic short stature (ISS). *J Clin Endocrinol Metab* 2011; **96**: E404-E412.

24 Iughetti L, Capone L, Elsedfy H *et al*: Unexpected phenotype in a boy with trisomy of the SHOX gene. *J Pediatr Endocrinol Metab* 2010; **23**: 159-169.

25 Ester WA, van Duyvenvoorde HA, de Wit CC *et al*: Two short children born small for gestational age with insulin-like growth factor 1 receptor haploinsufficiency illustrate the heterogeneity of its phenotype. *J Clin Endocrinol Metab* 2009; **94**: 4717-4727.

26 Klammt J, Kiess W, Pfaffle R: IGF1R mutations as cause of SGA. *Best Pract Res Clin Endocrinol Metab* 2011; **25**: 191-206.

27 Zhao X, Onteru SK, Piripi S *et al*: In a shake of a lamb's tail: using genomics to unravel a cause of chondrodysplasia in Texel sheep. *Anim Genet* 2012; **43 Suppl 1**: 9-18.

28 Markovich D: Slc13a1 and Slc26a1 KO models reveal physiological roles of anion transporters. *Physiology (Bethesda)* 2012; **27**: 7-14.

29 Hoopes BC, Rimbault M, Liebers D, Ostrander EA, Sutter NB: The insulin-like growth factor 1 receptor (IGF1R) contributes to reduced size in dogs. *Mamm Genome* 2012.

30 Khan AO, Aldahmesh MA, Al-Ghadeer H, Mohamed JY, Alkuraya FS: Familial spherophakia with short stature caused by a novel homozygous ADAMTS17 mutation. *Ophthalmic Genet* 2012; **33**: 235-239.

31 Le Goff C, Cormier-Daire V: The ADAMTS(L) family and human genetic disorders. *Hum Mol Genet* 2011; **20**: R163-R167.

32 Morales J, Al-Sharif L, Khalil DS *et al*: Homozygous mutations in ADAMTS10 and ADAMTS17 cause lenticular myopia, ectopia lentis, glaucoma, spherophakia, and short stature. *Am J Hum Genet* 2009; **85**: 558-568.

33 Sengle G, Tsutsui K, Keene DR *et al*: Microenvironmental regulation by fibrillin-1. *PLoS Genet* 2012; **8**: e1002425.

34 Weiske J, Albring KF, Huber O: The tumor suppressor Fhit acts as a repressor of beta-catenin transcriptional activity. *Proc Natl Acad Sci U S A* 2007; **104**: 20344-20349.

35 Wang Z, Shen D, Parsons DW *et al*: Mutational analysis of the tyrosine phosphatome in colorectal cancers. *Science* 2004; **304**: 1164-1166.

36 van Meeteren LA, Moolenaar WH: Regulation and biological activities of the autotaxin-LPA axis. *Prog Lipid Res* 2007; **46**: 145-160.

37 Trimbuch T, Beed P, Vogt J *et al*: Synaptic PRG-1 modulates excitatory transmission via lipid phosphate-mediated signaling. *Cell* 2009; **138**: 1222-1235.

38 Shin JN, Kim I, Lee JS, Koh GY, Lee ZH, Kim HH: A novel zinc finger protein that inhibits osteoclastogenesis and the function of tumor necrosis factor receptor-associated factor 6. *J Biol Chem* 2002; **277**: 8346-8353.

39 Li J, Wang CY: TBL1-TBLR1 and beta-catenin recruit each other to Wnt target-gene promoter for transcription activation and oncogenesis. *Nat Cell Biol* 2008; **10**: 160-169.

40 Yue Y, Grossmann B, Galetzka D, Zechner U, Haaf T: Isolation and differential expression of two isoforms of the ROBO2/Robo2 axon guidance receptor gene in humans and mice. *Genomics* 2006; **88**: 772-778.

41 Thompson BA, Tremblay V, Lin G, Bochar DA: CHD8 is an ATP-dependent chromatin remodeling factor that regulates beta-catenin target genes. *Mol Cell Biol* 2008; **28**: 3894-3904.

42 Lee JH, You J, Dobrota E, Skalnik DG: Identification and characterization of a novel human PP1 phosphatase complex. *J Biol Chem* 2010; **285**: 24466-24476.

43 Shabalina SA, Spiridonov NA: The mammalian transcriptome and the function of non-coding DNA sequences. *Genome Biol* 2004; **5**: 105.

44 Stamoulis C: Estimation of correlations between copy-number variants in non-coding DNA. *Conf Proc IEEE Eng Med Biol Soc* 2011; **2011**: 5563-5566.

45 Durand C, Roeth R, Dweep H *et al*: Alternative splicing and nonsense-mediated RNA decay contribute to the regulation of SHOX expression. *PLoS ONE* 2011; **6**: e18115.

46 Bonar SL, Brydges SD, Mueller JL *et al*: Constitutively activated NLRP3 inflammasome causes inflammation and abnormal skeletal development in mice. *PLoS ONE* 2012; **7**: e35979.

chapter

7

Supplemental Tables

Supplementary Table 1 Type II CNVs with lack of segregation with short stature

ID M/F	Height (SDS)	Karyotype (ISCN 2009)	Size (Kb)	Protein-coding genes[a]	miRNA
II.23 M	–2.1	arr Xq22.3(105,062,645–105,739,894)×2	677.2 Kb	*NRK* *SERPINA7* *MUM1L1*	-
II.24 F	–2.7	arr Xq25(125,001,006–125,318,345)×1 pat	317.3	*DCAF12L2* (*WDR40C*)	-
II.25 M	–4.5	arr Xq25(125,263,911–125,536,426)×0 mat	272.5	*DCAF12L2*	-
II.26 M	–2.4	arr 1q25.1(173,834,684–174,140,227)×1 mat	305.5	*ZBTB37* *SERPINC1* *RC3H1* *RABGAP1L*	-
II.27 F	–8.5	arr 1q43q44(243,546,954–243,821,364)×1 mat	274.4	*SDCCAG8* *AKT3*	-
II.3 M	–2.4	arr 1q44(246,715,197–247,652,602)×3 dn, 2q24.3(165,611,363–165,769,050)×3 pat	Chr1: 937.4 Chr2: 157.7	Chr1: 12 protein-coding genes; from *TFB2M* to *OR2B11* Chr2: *COBLL1* *SLC38A11*	-
II.28 F	–4.7	arr 2p25.3(1,101,473–1,742,700)×3 pat	641.2	*SNTG2* *TPO* *PXDN*	-
II.29 F	–3.2	arr 2q31.1(169,703,120–170,064,498)×3 mat, 2q31.1(172,454,934–172,657,695)×3 mat	Gain1: 361.4 Gain2: 202.8	Gain1: 6 protein-coding genes; from *NOSTRIN* to *LRP2* Gain2: *DYNC1I2* *SLC25A12*	-
II.30 M	–1.4	arr 5q33.1q33.2(152,375,974–153,460,042)×3 mat, 7p14.1(40,298,879–40,528,146)×1 mat	Chr5: 1,084.1 Chr7: 229.3	Chr5: *GRIA1* *FAM114A2* *MFAP3* Chr7: *C7orf10*	-
II.31 F	–2.6	arr 7q34q35(142,491,575–143,864,670)×3 mat	1,373.1	29 protein-coding genes; from *EPHB6* to *OR2A14*	-

II.32/mi.2 F	–2.8	arr 8p23.1(7,690,325–9,040,305)×3 pat, 8p23.1p22(12,242,033–13,046,661)×3 pat	8p23.1: 1,350.0 8p23.1p22: 804.6	Chr8p23.1: 10 protein-coding genes; from *DEFB104A* to *PPP1R3B* Chr8p23.1p22: *FAM86B2* *LONRF1* *KIAA1456* *DLC1*	8p23.1: MIR548I3
II.33 F	–2.9	arr 9q31.3(111,555,994–111,711,514)×1 mat	155.5	*ACTL7B* *ACTL7A* *IKBKAP* *C9orf6* *CTNNAL1*	-
II.34 F	–3.0	arr 9q34.13q34.2(134,789,097–136,484,291)×3 mat	1,695.2	31 protein-coding genes; from *MED27* to *FAM163B*	-
II.35[b] F	–2.1	arr 15q24.2q24.3(76,314,543–76,727,022)×3 mat	412.5	*NRG4* *C15orf27* *ETFA* *ISL2* *SCAPER*	-
II.36[b] F	–3.4	arr 15q24.2q24.3(76,314,543–76,727,022)×3 mat	412.5	*NRG4* *C15orf27* *ETFA* *ISL2* *SCAPER*	-
II.37/ mi.4 F	–2.5	arr 16p13.12p13.11(14,760,735–16,633,360)×1 pat	1,872.6	17 protein-coding genes; from *BFAR* to *NOMO3*	MIR1972-1 MIR484
II.38 F	–2.5	arr 16q12.1(49,088,824–49,615,386)×3 mat	526.6	*CBLN1* *C16orf78,* *ZNF423*	-

[a] For CNVs containing ≤ 5 protein-coding genes, all protein-coding genes are depicted. For CNVs containing ≥ 6 protein-coding genes, the number, and the first and last protein-coding gene in the CNV are given. [b] Family; 2 sisters.

Supplementary Table 2 Type II CNVs with no information about segregation

ID M/F	Height (SDS)	Karyotype (ISCN 2009)	Size (Kb)	Protein-coding genes[a]	miRNA	Arguments pro pathogenicity
II.14 F	–4.1	arr Xp22.2(9,594,546–10,046,186)×3	451.6	*TBL1X* *GPR143* *SHROOM2* *WWC3*	–	*TBL1X* required for WNT signalling
II.15 M	–4.7	arr 1q24.2(169,296,745–169,474,353)×3	177.6	*NME7* *BLZF1* *C1orf114* *SLC19A2*	–	–
II.16 F	–2.5	arr 3p24.3(21,756,554–22,231,872)×3	475.3	*ZNF385D*	–	–
II.17 M	–5.2	arr 3p12.3(76,153,037–77,581,256)×3	1,428.2	*ROBO2*	–	Brother carrying same CNV short (–4.3 SDS); function in axon guidance and cell migration
II.18 M	–4.6	arr 7q21.11(80,071,944–80,768,261)×3	696.3	*CD36* *GNAT3* *SEMA3C*	–	–
II.19/mi.1 F	–4.6	arr 7q36.3(158,183,050–158,692,049)×3	509.0	*PTPRN2* *NCAPG2* *ESYT2* *WDR60*	MIR595	–
II.20 F	–4.4	arr 9p23(9,503,942–10,018,120)×3	514.2	*PTPRD*	–	–
II.21 F	–3.7	arr 14q11.2(21,530,059–22,022,116)×3	492.1	13 protein-coding genes; from *NDRG2* to *SALL2*	–	*CHD8* negative regulator of Wnt. *TOX4* involved in cell cycle progression
II.22/mi.5 M	–2.1	arr 22q11.21(21,011,217–21,928,915)×1	917.7	16 protein-coding genes; from *POM121L4P* to *UBE2L3*	MIR649	Distal part of common 22q11 deletion syndrome

[a] For CNVs containing ≤ 5 protein-coding genes, all protein-coding genes are depicted. For CNVs containing ≥ 6 protein-coding genes, the number, and the first and last protein-coding gene in the CNV are given.

chapter
7 S