Genome-wide screening of copy number variants in children born small for gestational age reveals several candidate genes involved in growth pathways

Ana P M Canton, Sílvia S Costa¹, Tatiane C Rodrigues¹, Debora R Bertola^{1,2},
Alexsandra C Malaquias, Fernanda A Correa³, Ivo J P Arnhold³, Carla Rosenberg¹
and Alexander A L Jorge

Unidade de Endocrinologia Genética, Laboratório de Endocrinologia Celular e Molecular LIM/25, Disciplina de Endocrinologia da Faculdade de Medicina da Universidade de São Paulo, Av. Dr. Arnaldo, 455 5º Andar Sala 5340, CEP 01246-903 São Paulo, Brazil, ¹Departamento de Genética e Biologia Evolutiva, Instituto de Biociências da Universidade de São Paulo, 05508-900 São Paulo, Brazil, ²Unidade de Genética, Instituto da Criança, Faculdade de Medicina da Universidade de São Paulo, 05403-000 São Paulo, Brazil and ³Unidade de Endocrinologia do Desenvolvimento, Laboratório de Hormônios e Genética Molecular LIM/42 do Hospital das Clínicas, Disciplina de Endocrinologia da Faculdade de Medicina da Universidade de São Paulo, 05403-900 São Paulo, Brazil

Correspondence
should be addressed
to A A L Jorge
Email
alexj@usp.br

Abstract

Background: The etiology of prenatal-onset short stature with postnatal persistence is heterogeneous. Submicroscopic chromosomal imbalances, known as copy number variants (CNVs), may play a role in growth disorders.

Objective: To analyze the CNVs present in a group of patients born small for gestational age (SGA) without a known cause.

Patients and methods: A total of 51 patients with prenatal and postnatal growth retardation associated with dysmorphic features and/or developmental delay, but without criteria for the diagnosis of known syndromes, were selected. Array-based comparative genomic hybridization was performed using DNA obtained from all patients. The pathogenicity of CNVs was assessed by considering the following criteria: inheritance; gene content; overlap with genomic coordinates for a known genomic imbalance syndrome; and overlap with CNVs previously identified in other patients with prenatal-onset short stature.

Results: In 17 of the 51 patients, 18 CNVs were identified. None of these imbalances has been reported in healthy individuals. Nine CNVs, found in eight patients (16%), were categorized as pathogenic or probably pathogenic. Deletions found in three patients overlapped with known microdeletion syndromes (4q, 10q26, and 22q11.2). These imbalances are *de novo*, gene rich and affect several candidate genes or genomic regions that may be involved in the mechanisms of growth regulation.

Conclusion: Pathogenic CNVs in the selected patients born SGA were common (at least 16%), showing that rare CNVs are probably among the genetic causes of short stature in SGA patients and revealing genomic regions possibly implicated in this condition.

European Journal of
Endocrinology
(2014) 171, 253–262

Introduction

Children born small for gestational age (SGA) are defined as those with birth weight and/or birth length at least 2 s.d. below the population mean for gestational age (1). The causes of prenatal growth impairment are often unclear

and involve maternal, placental, and/or fetal factors (1). Fetal factors include several genetic causes commonly characterized by persistent short stature, dysmorphic features, and developmental disorders (2). These genetic

syndromes of prenatal-onset growth impairment often encompass complex clinical disorders of difficult diagnosis (3).

Array-based genomic copy number analysis has recently become a research tool and a clinical genetic test in the diagnostic work-up of several clinical settings (4). One of the most used genomic copy number analysis types is array-based comparative genomic hybridization (aCGH), a technique used for genome-wide screening of submicroscopic segmental genomic gains and losses. This method enables a much higher resolution assay (greater than tenfold) of genomic imbalances than conventional cytogenetic methods (4, 5). The imbalances characterized by aCGH are called copy number variants (CNVs). A CNV is an imbalance of a genomic sequence that alters the diploid status of a particular locus in the human genome; essentially, it comprises deletions and duplications.

Currently, aCGH is indicated as a first-tier clinical diagnostic test in patients with unexplained developmental delay/intellectual disability, autism spectrum disorders, and multiple congenital anomalies (5). It is also becoming a powerful tool in the discovery of disease genes, identification of new syndromes, and delineation of known deletion/duplication syndromes, known as genomic disorders. In the field of endocrinology, genome-wide analysis of CNVs has been evaluated as a tool to identify genetic causes in several clinical conditions, including premature ovarian failure (6), severe early-onset obesity (7), congenital hypothyroidism and thyroid dysgenesis (8), ambiguous genitalia (9), skeletal defects with growth impairment (10), and Silver–Russell syndrome (11). Two recent studies have demonstrated the presence of causative and possible pathogenic CNVs in short-stature children (12, 13). However, in both studies, children were selected regardless of compromised birth weight and/or length.

In the present study, we performed aCGH in a group of 51 children born SGA with persistent short stature without a recognized cause of prenatal and postnatal growth impairment. Our results indicate that CNVs contribute to the genetic etiology of prenatal-onset short stature and reveal novel potential candidate genes and/or loci related to this condition.

Patients and methods

Patients

This study was approved by the Local Ethics Committee, and the patients and/or guardians gave their written

informed consent. Candidates for aCGH were selected from a total of 137 children followed at our outpatient clinic due to prenatal-onset short stature using the following criteria: i) birth weight and/or length s.d. ≤ -2.0 ; ii) persistent short stature after early infancy (height s.d. ≤ -2.0 at the age of 4 years or above); and iii) the presence of dysmorphic features, developmental delay, and/or intellectual disability, but without criteria for the diagnosis of a recognized syndrome, including Silver–Russell syndrome (14). The exclusion criteria for selection were as follows: i) placental and maternal known factors of intrauterine growth retardation, including placental structural or perfusional factors and maternal infections, substance use/abuse, or medical conditions and ii) known causes of short stature, including endocrinological diseases, skeletal dysplasias, hepatic diseases, systemic diseases, and established genetic diseases. All patients had normal G-banded karyotyping. A total of 51 patients, who fulfilled all inclusion criteria and no exclusion criteria, were selected for the molecular genetic study. Among these patients, 11 were investigated for demethylation of the paternal imprinting center region one on chromosome 11p15 and 14 were investigated for insulin-like growth factor 1 (*IGF1*) and IGF1 receptor (*IGF1R*) mutations (15, 16). In all the selected patients, these molecular genetic tests proved to be normal.

Array-based comparative genomic hybridization

Genomic DNA was extracted from peripheral blood leucocytes of all patients and 28 relatives (including 14 maternal samples, 11 paternal samples, and three other first-degree relatives' samples) using standard procedures. aCGH was performed in a whole-genome customized 60K oligonucleotide platform (SurePrint G3 Human CGH Microarray Kit, 8×60K, Agilent Technologies, Inc., Santa Clara, CA, USA). The procedure was conducted according to the standard protocol of the manufacturer (17). Briefly, all standard quality assessments were implemented during DNA sample and array preparation. Microarray scanned images were processed using the Software Feature Extraction, version 10.7.3.1 (Agilent Technologies, Inc.). The Genomic Workbench 6.9 Software (Agilent Technologies, Inc.) was applied for calling CNVs using the Aberration Detection Method 2 analysis algorithm, with a sensitivity threshold of 6.7. To consider an aberration of a genomic segment as a duplication or a deletion, the log2 ratio of cyanine 3:cyanine 5 intensities was used encompassing at least three probes. Ratios >0.3 were considered duplications, whereas ratios <-0.3 were considered deletions.

All hybridizations were gender-matched and processed in reverse-labeling duplicates, which means that all patients were evaluated twice for confirmation. CNVs not detected in both experiments were disregarded.

Copy number segment analysis

Detected CNVs > 50 kb (5) were compared with CNV data from oligoarray studies documented in the Database of Genomic Variants (DGV) (18) and an independent control cohort of 400 healthy individuals with the same ethnic background, excluding common copy number polymorphisms (those with a frequency > 1%). The assessment of CNV pathogenicity was made by considering the following four criteria based on previous studies: i) inheritance and familial segregation with the prenatal and postnatal growth impairment phenotype; ii) overlap with genomic coordinates for a known genomic imbalance syndrome; iii) overlap with CNVs defined in other patients, in particular, in those presenting prenatal-onset short stature reported in the Database of Chromosomal Imbalance and Phenotype in Humans using Ensembl Resources (DECIPHER) (19); and iv) gene content of coding genes associated with impaired developmental process and/or impaired cell growth pathways and/or growth impairment in animal models (4, 5). According to

this strategy, CNVs were categorized as pathogenic if they fulfilled all these criteria. CNVs were categorized as probably pathogenic if they fulfilled all these criteria, except the overlap with genomic coordinates for a known genomic imbalance syndrome. CNVs were categorized as probably benign mainly if they did not segregate with the prenatal and postnatal growth impairment phenotype in the patients' families. Finally, CNVs were categorized as variants of uncertain clinical significance (VUS) if they were unable to be classified in one of the three categories mentioned above (Fig. 1).

In an attempt to establish a causal genotype–phenotype correlation and to identify genes involved in growth impairment of prenatal onset, the gene content of the variants was analyzed using bioinformatics tools for gene prioritization (20). Gene annotation was performed using the University of California Santa Cruz Genome Browser (UCSC Genome Browser) (21). The assessment of gene function and overlap with genomic disorders was performed using the Online Mendelian Inheritance in Man (OMIM) and the PubMed databases. The assessment of microRNA function was performed using the miRBase (22) and the miRTarBase databases (23). A search for murine knockout phenotypes with growth impairment was carried out using the Mouse Genome Informatics (MGI) (24). Finally, gene prioritization to identify the most

Figure 1

Summary of study strategy with CNV identification and classification.

promising genes among the pool of candidates was performed using the Gene Distiller strategy (25) and the evaluation of putative links to other genes and biological pathways that have more established roles in the genetic control of growth.

Statistical analyses

The quantitative variables are expressed as means $\pm$ s.d. and were analyzed using unpaired Student's *t*-test or Kruskal–Wallis test, as appropriate. A *P* value <0.05 was considered to be statistically significant. All the statistical analyses were carried out using SigmaStat, version 3.5 (Windows, version 3.5; Systat Software, Inc., Erkrath, Germany).

Results

The cohort was characterized by a male predominance (63%). Among the patients, 34% were SGA just for length, 33% were SGA just for weight, and 33% were SGA for both weight and length. During the first evaluation, they presented a chronological age of 7.6 ± 3.7 years, with a marked short stature (height s.d. = -3.2 ± 0.8), even if corrected for parental height (height s.d. minus target height s.d. = -2.2 ± 1.2) (Table 1). The following clinical findings were most prevalent in the selected patients:

Table 1 Clinical characterization of the 51 selected patients with prenatal-onset short stature. Numeric values are shown as means $\pm$ s.d.; categorical values are shown as absolute numbers and percentages.

	Selected patients
Gender (M:F)	32:19
Preterm:term birth	14:37
Birth weight s.d. ≤ -2	30 (59%)
Birth length s.d. ≤ -2	32 (63%)
Target height s.d.	-1.0 ± 1.0
Father height s.d. ≤ -2	9 (17%)
Mother height s.d. ≤ -2	6 (12%)
Age during the first evaluation	7.6 ± 3.7
Height s.d.	-3.2 ± 0.8
Developmental delay	18 (25%)
Intellectual disability	5 (9.8%)
Microcephaly ^a	19 (37%)
Heart defects	7 (14%)
Skeletal defects	21 (41%)
Dysmorphic features ^b	37 (73%)

^aHead circumference s.d. < -2 .

^bTriangular face, high-arched palate, pectus excavatum, facial and/or body asymmetry, cryptorchidism, dysplastic ears, hypertelorism, epicanthus, short neck, and strabismus.

dysmorphic features (triangular face, hypertelorism, epicanthus, strabismus, dysplastic ears, high-arched palate, short neck, pectus excavatum, facial and/or body asymmetry, and cryptorchidism) in 73%; nonspecific skeletal abnormalities in 41%; microcephaly in 37%; and developmental delay in 25%. Among the selected patients, 60% did not fulfill the established criteria to indicate aCGH, according to clinical indications in the current literature (5).

In 17 of the 51 patients screened (33%), 18 rare CNVs were identified. Nine variants were deletions (ranging from 0.17 to 15.1 Mb) and nine were duplications (ranging from 0.26 to 4.4 Mb). None of these CNVs was found in our reference set of 400 controls or was considered a copy number polymorphism in comparison with DGV studies. The aCGH of parental DNA revealed that ten patients had *de novo* variants, whereas four patients inherited variants from a healthy parent. Parental DNA samples were not available for analysis in three patients.

Four distinct CNVs, found in three patients, were classified as pathogenic (patients 1–3; Tables 2 and 3 and Supplementary Fig. 1, see section on supplementary data given at the end of this article), while five distinct CNVs, found in five patients, were classified as probably pathogenic (patients 4–8; Tables 2 and 3 and Supplementary Fig. 1). All the identified pathogenic and probably pathogenic CNVs were *de novo* and gene rich. The complete list of all genes, including microRNA genes, observed in CNV regions is given in Supplementary Table 1. Eight of them overlap with CNVs defined in other patients with prenatal-onset short stature reported in the DECIPHER (Table 2). Deletions found in patient 1 (22q11.21 deletion), patient 2 (10q26.3 deletion), and patient 3 (4q deletion) overlap with known deletion syndromes (26, 27, 28). In these CNVs categorized as pathogenic or probably pathogenic, several candidate genes were identified (including one microRNA; Table 2). Patients born SGA with pathogenic or probably pathogenic CNVs have heterogeneous phenotypes (Table 3), mainly presenting with nonfamilial short stature ($n=8/8$), microcephaly ($n=5/8$), developmental delay ($n=4/8$), and body/facial asymmetry ($n=3/8$).

Copy number changes found in patients 9–12 were considered VUS (Supplementary Table 2, see section on supplementary data given at the end of this article). Parental DNA samples were not available for analysis in patients 9–11; patient 12 (a boy) inherited an X chromosome deletion from his healthy mother, indicating a possible X-linked inheritance pattern. The five remaining CNVs were classified as likely benign (patients 12–17;

Clinical Study	A P M Canton and others	Copy number variants in SGA children	171:2	257
----------------	-------------------------	--------------------------------------	-------	-----

Table 2 Description and functional characterization summary of CNVs categorized as pathogenic (P) or probably pathogenic (PP).

	Gain/loss	Inheritance	Locus	Genomic position ^a	Size (Mb)	No. of affected genes	DECIPHER ^b	Mouse genome database ^c	Deletion syndrome	Main candidate genes ^d
P-CNV										
1	Loss	De novo	22q11.21	Chr22: 18,894,620-21,440,656	2.5	49	11	6	22q11.2	Contiguous gene syndrome
2	Gain	De novo	10q26.2-q26.3	Chr10: 129,100,719-130,723,355	1.6	6	1	2	–	–
	Loss	De novo	10q26.3	Chr10: 130,811,542-135,404,671	4.5	34	11	2	10q26	GLRX3
3	Loss	De novo	4q28.2-q31.21	Chr4: 129,437,805-144,588,797	15.1	30	5	5	Interstitial 4q	GAB1
PP-CNV										
4	Loss	De novo	3q27.1-q27.3	Chr3: 183,902,785-187,476,327	3.5	46	4	9	–	IGF2BP2
5	Loss	De novo	20p13	Chr20: 82,892-481,772	0.4	13	1	1	–	CSNK2A1
6	Gain	De novo	14q11.2-q12	Chr14: 21,819,565-24,578,066	2.7	57	1	4	–	CHD8
7	Gain	De novo	16q24.1-q24	Chr16: 85,678,593-90,148,400	4.4	59	1	7	–	GALNS
8	Gain	De novo	Xq13.1-q13.2	ChrX: 71,553,894-73,461,083	1.9	14	–	3	–	MIR421

^aGenomic positions are given according to Human Genome Building GCRh37, HG19.^bNumber of patients reported in the DECIPHER database with overlapping CNVs and prenatal-onset short stature.^cNumber of genes involved in growth impairment in mouse model.^dMain candidate genes involved in growth impairment of prenatal onset.

Supplementary Table 2), primarily because they did not segregate with phenotype. Four of them were absent in parents who shared a growth phenotype, while three of them were inherited from a healthy parent.

On comparing the copy number segments, the mean size of pathogenic and probably pathogenic variants was found to be 4.0 Mb, greater than the mean size of VUS and benign variants (0.66 Mb; $P=0.034$). Similarly, the mean gene content of pathogenic and probably pathogenic variants (38 genes) was richer than the mean gene content of VUS and benign variants taken together (1.8 genes; $P<0.001$).

Among the 51 patients, 27 were treated with recombinant human growth hormone (rhGH, mean dose of 50 µg/kg per day), and five of them were patients carrying pathogenic or probably pathogenic CNVs. Patients born SGA with pathogenic or probably pathogenic CNVs exhibited a first-year growth response to rhGH similar to that exhibited by patients born SGA without pathogenic CNVs (Supplementary Table 3, see section on supplementary data given at the end of this article).

Discussion

Modern research on CNVs in chromosomal segments has shown the importance of these variants as a significant source of human genetic variations accounting for disease, population diversity (29), and variation in stature in the general population (30). In the present study, we investigated 51 patients with prenatal-onset short stature associated with dysmorphic features and/or developmental delay, a condition of heterogeneous etiology. In eight of them (16%), rare pathogenic or probably pathogenic CNVs were detected, indicating that submicroscopic chromosomal losses and gains may be common in this population. This frequency is similar to the frequency of pathogenic CNVs found for other known conditions tested and diagnosed with genome-wide copy number analysis techniques, including developmental delay and congenital anomalies (5).

In a recent study, underlying pathogenic CNVs have been identified in 20 of 200 patients (10%) with short stature without an apparent cause; eight of these 20 patients were born SGA (13). In another study, genome-wide analysis of CNVs of 162 patients with short stature of unknown origin revealed at least 17 (10.4%) pathogenic variants, regarding familial segregation (12). The heterogeneous phenotypes observed in eight patients carrying pathogenic and probably pathogenic CNVs in the present study are quite common

Table 3 Clinical findings of the patients with CNVs categorized as pathogenic (P) or probably pathogenic (PP).

Patient	Gender	Age (years)	Height s.d.	Target height s.d.	GA (weeks)	Birth weight s.d.	Birth length s.d.	Clinical findings
P-CNV								
1	F	5.8	−2.7	−1.5	39	−2.5	−1.4	Spina bifida without communication, clinodactyly of the fifth finger, and seizures
2	F	10.0	−3.7	−0.8	39	−2.3	−3.0	Microcephaly, cleft palate, deafness, squint, body asymmetry, pectus carinatum, triangular face, learning difficulties, and strabismus
3	M	5.7	−2.7	0.3	38	−2.3	−2.6	Microcephaly, developmental delay, and dysplastic ears
PP-CNV								
4	F	4.8	−3.0	−0.1	38	−1.8	−2.9	Microcephaly, developmental delay, hyperactivity, and facial asymmetry
5	F	6.3	−2.9	−1.6	40	−2.1	−3.3	Epicanthus, short neck, marked lumbar lordosis, and muscular pseudohypertrophy
6	F	1.5	−2.6	−1.3	38	−2.0	−3.8	Microcephaly
7	F	19.0	−2.5	−0.1	39	−1.7	−3.5	Asplenia, corpus callosum dysgenesis, pulmonic stenosis, atrial septal defect, developmental delay, and scoliosis
8	M	12.4	−3.3	−0.4	31	−2.1	−3.6	Microcephaly, strabismus, developmental delay, ventricular septal defect, renal agenesis, cryptorchidism, body asymmetry, and vertebral fusion

GA, gestational age.

among syndromic individuals carrying chromosomal aberrations (29). Our results draw attention to the fact that pathogenic and probably pathogenic variants detected occurred in different chromosomal regions, which could explain the clinical variability of the SGA patients investigated in the present study. The same genetic variability was observed in the pathogenic and probably pathogenic CNVs in the two studies mentioned above (12, 13).

CNVs categorized as pathogenic or probably pathogenic in the present study fulfill some established deleterious criteria. All of them are *de novo*, are gene rich, are of significant size, and involve genes with functions in cell cycle, developmental process, and/or cell growth pathways (Supplementary Table 2) (5). Moreover, deletions found in three patients overlapped with known deletion syndromes. Therefore, aCGH allowed clinical genetic diagnosis and elucidation of the genetic basis of prenatal-onset short stature in eight patients carrying pathogenic or probably pathogenic CNVs in the present study, not suspected before copy number segment analysis.

The 22q11.21 deletion found in patient 1 overlapped with the genomic coordinates of the 22q11.2 deletion syndrome (or DiGeorge syndrome; OMIM #188400). This is one of the most common genomic alterations in humans, with an estimated frequency of 1:2000–1:4000 live births and a markedly variable phenotypic expression, leading to difficult diagnosis based on clinical grounds (26). Short stature during childhood has been described in

several clinical series, with frequencies ranging from 10 to 60% and ~26% of these short children being born SGA (31). Among the 49 coding genes deleted in patient 1, 24 have functions in cell cycle, developmental process, and/or cell growth pathways. However, no specific gene included in this deletion was found to be associated with the short stature phenotype. Nevertheless, it is worth pointing out that 22q11.2 deletion was the only pathogenic CNV identified in common in the present study as well as in other recent studies of CNVs in patients with short stature of unknown origin (14, 15).

The terminal 10q26.3 deletion found in patient 2 overlapped with the genomic coordinates of the 10q26 deletion syndrome (OMIM #609625). This is a rare contiguous gene deletion syndrome with more than 15 cases being reported (27). They all share distinct clinical features also found in our patient, mainly short stature of prenatal onset, microcephaly, deafness, strabismus, body asymmetry, and learning difficulties. Analysis of this variant revealed glutaredoxin 3 (*GLRX3*, OMIM #612754) to be the main candidate gene involved in the growth impairment of prenatal onset. This gene encodes an oxidoreductase enzyme (member of the glutaredoxin family), which plays a role in cellular growth and may inhibit apoptosis (32). *Glrx3*-knockout mice have been found to exhibit morphological defects with growth impairment (33).

The interstitial 4q28.2–31.2 deletion found in patient 3 overlapped with the 4q deletion syndrome, a rare

chromosomal disorder with an estimated incidence of 1:100 000 and nearly 150 cases being reported in the literature. Interstitial and terminal deletions of the long arm of chromosome 4 have been reported with several clinical conditions affecting multiple organs and systems (28). Pre- or postnatal-onset short stature has been reported in ~50% of the cases, but the critical region responsible for growth impairment remains unclear. Nevertheless, in our analysis, we identified the growth factor receptor-bound protein 2-associated binding protein 1 (*GAB1*, OMIM #604439) to be the main candidate gene. The protein encoded by this gene plays a role in a broad range of growth factor and cytokine signaling pathways, including the insulin receptor signaling pathway (34). *GAB1* and 22 other coding genes included in the deletion found in patient three are part of the Stature Quantitative Trait Locus 12 (STQTL12, OMIM #612224), a region mapped as underlying an association with height trait based on the measurement of gene expression (35).

In functional characterization of the 3.5 Mb deletion found in patient 4, we identified IGF2 mRNA-binding protein 2 (*IGF2BP2*, OMIM #608289) to be the most promising gene with a putative role in the genetic control of growth. This gene was also deleted in four other patients reported in the DECIPHER with prenatal-onset short stature and CNVs overlapping the deletion of our patient (subject identification: 1495, 256 994, 257 537, and 257 773). *IGF2BP2* is a mRNA-binding protein involved in *IGF2* RNA localization, stability, and translation, being necessary for embryonic growth and development (36). Recently, the genome-wide association study Genetic Investigation of Anthropocentric Traits Consortium (GIANT) has identified the *IGF2BP2* locus to affect human height (rs720390) (37).

In bioinformatics analysis of the 20p13 deletion found in patient 5, we identified casein kinase II alpha 1 subunit isoform b (*CSNK2A1*, OMIM #115440) as the main candidate gene to be involved in the growth phenotype. This gene encodes the catalytic subunit of a serine/threonine protein kinase that phosphorylates acidic proteins, regulating cellular processes, such as cell cycle progression, apoptosis, and transcription (38). *Csnk2a1*-knockout mice have defects in limb development with growth impairment (39).

The remaining variants classified as probably pathogenic in the present study (patients 6–8) are copy number gains. Until recently, the interpretation of this type of CNV was considered a challenge in the literature. However, current studies and guidelines indicate its clinical relevance (40). In bioinformatics analysis of the

duplication found in patient 6, we identified 14 genes with functions in cell cycle, developmental process, and/or cell growth pathways. Among them, the most promising gene to be involved in growth impairment appeared to be chromodomain helicase DNA-binding protein 8 (*CHD8*, OMIM #610528). This gene encodes a DNA helicase with functions in transcription repression by remodeling chromatin structure. It binds to β -catenin and negatively regulates the Wnt signaling pathway, which plays a role in early development and morphogenesis (41).

In the analysis of the 4.4 Mb duplication found in patient 7, we identified three genes with loci that have recently been characterized as affecting human height by genome-wide association studies: galactosamine-6-sulfate sulfatase (*GALNS*, OMIM #612222, rs8052560) (37), cytosolic thiouridylase subunit 2 homolog (*CTU2*, rs8052560) (37), and spermatogenesis-associated 33 (*C16ORF55* (*SPATA33*), rs154663) (42). *GALNS* encodes a lysosomal enzyme involved in the catabolism of keratan and chondroitin sulfate, the deficiency of which leads to mucopolysaccharidosis type IVA (43). *CTU2* and *SPATA33* encode proteins without a known function yet. In addition to establishing a causal genotype–phenotype correlation in our functional characterization, we found that the genomic region duplicated in patient seven is part of the STQTL22 (OMIM #613547).

The Xq13.1–13.2 duplication found in patient 8 is a rare condition and partially overlaps with duplications found in 12 other male patients reported in the literature (44). All these patients have been described to have growth impairment. In our bioinformatics analysis, we identified microRNA 421 (*MIR421*) as the main candidate gene to be involved in the growth phenotype. Recently, it has been shown that miR-421 levels are inversely correlated with that of mothers against DPP homolog 4 gene (*SMAD4*, OMIM #600993), functioning as a negative regulator of the expression of this target gene (45). *SMAD4* encodes a protein with an essential role in the signal transduction of the bone morphogenetic pathway (BMP) and the transforming growth factor (TGF) β pathway. Also recently, heterozygous missense *SMAD4* mutations have been described as the causative agents of Myhre syndrome (OMIM #139210) in patients fulfilling diagnostic criteria for the syndrome (46). Myhre syndrome is a rare developmental disorder, and interestingly its phenotype resembles some important clinical findings of patient 8 in the present study: prenatal-onset short stature, microcephaly, facial dysmorphisms, skeletal anomalies, developmental delay, congenital heart defects, and cryptorchidism. Thus, we hypothesize that the

duplication found in patient 8 could have been a gain-of-function effect of *MIR421*, which could have enhanced its negative regulation of *SMAD4* expression, leading to a clinical picture quite similar to that of Myhre syndrome.

It is important to note that the CNVs considered as pathogenic or probably pathogenic in the present study probably represent an underestimate: CNVs present in patients and inherited from a clinically normal parent may still contribute to the abnormal phenotype. There are several well-documented cases in the literature, including the 1q21 deletions associated with TAR syndrome (47) or the 16p12 deletions associated with autism (48) and Mullerian aplasia (49). Additionally, it is important to be aware that the aCGH technique does not detect uniparental disomy, a known cause for children being born SGA (14).

In conclusion, in the present study, we found CNVs in distinct chromosomal regions with several candidate genes that may be involved in the mechanisms of growth regulation and/or in the regulatory pathways of intra-uterine development. It is important to note that these candidate genes identified are hypotheses based on bioinformatics analyses and that further studies are necessary to clarify their involvement in growth pathways. Moreover, we found at least 16% pathogenic or probably pathogenic CNVs in a group of SGA patients with persistent short stature associated with dysmorphic features and/or developmental delay, indicating that rare CNVs are probably among the genetic causes of this condition. Our results improve the evidence of the importance of genome-wide copy number analysis as a clinical genetic test to clarify diagnoses in patients with growth impairment.

Supplementary data

This is linked to the online version of the paper at <http://dx.doi.org/10.1530/EJE-14-0232>.

Declaration of interest

The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

Funding

This study was supported by grants 2013/03236-5 (to A A L Jorge) and 2009/00898-1 (to C Rosenberg) from the Sao Paulo Research Foundation (FAPESP) and grants 307922/2013-8 (to I J P Arnhold), 473285/2011-8, and 304678/2012-0 (to A A L Jorge) from the National Council for Scientific and Technological Development (CNPq).

References

- 1 Lee PA, Chernausk SD, Hokken-Koelega AC & Czernichow P. International Small for Gestational Age Advisory Board consensus development conference statement: management of short children born small for gestational age, April 24–October 1, 2001. *Pediatrics* 2003 **111** 1253–1261. (doi:10.1542/peds.111.6.1253)
- 2 Wit JM, Kiess W & Mullis P. Genetic evaluation of short stature. *Best Practice & Research. Clinical Endocrinology & Metabolism* 2011 **25** 1–17. (doi:10.1016/j.beem.2010.06.007)
- 3 Renes JS, Willemssen RH, Wagner A, Finken MJ & Hokken-Koelega AC. Bloom syndrome in short children born small for gestational age: a challenging diagnosis. *Journal of Clinical Endocrinology and Metabolism* 2013 **98** 3932–3938. (doi:10.1210/jc.2013-2491)
- 4 Lee C, Iafrate AJ & Brothman AR. Copy number variations and clinical cytogenetic diagnosis of constitutional disorders. *Nature Genetics* 2007 **39** S48–S54. (doi:10.1038/ng2092)
- 5 Miller DT, Adam MP, Aradhya S, Biesecker LG, Brothman AR, Carter NP, Church DM, Crolla JA, Eichler EE, Epstein CJ *et al.* Consensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. *American Journal of Human Genetics* 2010 **86** 749–764. (doi:10.1016/j.ajhg.2010.04.006)
- 6 Aboura A, Dupas C, Tachdjian G, Portnoi MF, Bourcigaux N, Dewailly D, Frydman R, Fauser B, Ronci-Chaix N, Donadille B *et al.* Array comparative genomic hybridization profiling analysis reveals deoxyribonucleic acid copy number variations associated with premature ovarian failure. *Journal of Clinical Endocrinology and Metabolism* 2009 **94** 4540–4546. (doi:10.1210/jc.2009-0186)
- 7 Bochukova EG, Huang N, Keogh J, Henning E, Purmann C, Blaszczyk K, Saeed S, Hamilton-Shield J, Clayton-Smith J, O'Rahilly S *et al.* Large, rare chromosomal deletions associated with severe early-onset obesity. *Nature* 2010 **463** 666–670. (doi:10.1038/nature08689)
- 8 Thorwarth A, Mueller I, Biebermann H, Ropers HH, Grueters A, Krude H & Ullmann R. Screening chromosomal aberrations by array comparative genomic hybridization in 80 patients with congenital hypothyroidism and thyroid dysgenesis. *Journal of Clinical Endocrinology and Metabolism* 2010 **95** 3446–3452. (doi:10.1210/jc.2009-2195)
- 9 Ledig S, Hiort O, Scherer G, Hoffmann M, Wolff G, Morlot S, Kuechler A & Wieacker P. Array-CGH analysis in patients with syndromic and non-syndromic XY gonadal dysgenesis: evaluation of array CGH as diagnostic tool and search for new candidate loci. *Human Reproduction* 2010 **25** 2637–2646. (doi:10.1093/humrep/deq167)
- 10 de Pontual L, Yao E, Callier P, Faivre L, Drouin V, Cariou S, Van Haeringen A, Genevieve D, Goldenberg A, Oufadem M *et al.* Germline deletion of the miR-17 approximately 92 cluster causes skeletal and growth defects in humans. *Nature Genetics* 2011 **43** 1026–1030. (doi:10.1038/ng.915)
- 11 Spengler S, Begemann M, Ortiz Bruchle N, Baudis M, Denecke B, Kroisel PM, Oehl-Jaschkowitz B, Schulze B, Raabe-Meyer G, Spaich C *et al.* Molecular karyotyping as a relevant diagnostic tool in children with growth retardation with Silver–Russell features. *Journal of Pediatrics* 2012 **161** 933–942. (doi:10.1016/j.jpeds.2012.04.045)
- 12 van Duyvenvoorde HA, Lui JC, Kant SG, Oostdijk W, Gijbbers AC, Hoffer MJ, Karperien M, Walenkamp MJ, Noordam C, Voorhoeve PG *et al.* Copy number variants in patients with short stature. *European Journal of Human Genetics* 2014 **22** 602–609. (doi:10.1038/ejhg.2013.203)
- 13 Zahnleiter D, Uebe S, Ekici AB, Hoyer J, Wiesener A, Wiczorek D, Kunstmann E, Reis A, Doerr HG, Rauch A *et al.* Rare copy number variants are a common cause of short stature. *PLoS Genetics* 2013 **9** e1003365. (doi:10.1371/journal.pgen.1003365)
- 14 Dias RP, Nightingale P, Hardy C, Kirby G, Tee L, Price S, Macdonald F, Barrett TG & Maher ER. Comparison of the clinical scoring systems in Silver–Russell syndrome and development of modified diagnostic

- criteria to guide molecular genetic testing. *Journal of Medical Genetics* 2013 **50** 635–639. (doi:10.1136/jmedgenet-2013-101693)
- 15 Coutinho DC, Coletta RR, Costa EM, Pachi PR, Boguszewski MC, Damiani D, Mendonca BB, Arnhold IJ & Jorge AA. Polymorphisms identified in the upstream core polyadenylation signal of IGF1 gene exon 6 do not cause pre- and postnatal growth impairment. *Journal of Clinical Endocrinology and Metabolism* 2007 **92** 4889–4892. (doi:10.1210/jc.2007-1661)
 - 16 Leal AC, Montenegro LR, Saito RF, Ribeiro TC, Coutinho DC, Mendonca BB, Arnhold IJ & Jorge AA. Analysis of the insulin-like growth factor 1 receptor gene in children born small for gestational age: *in vitro* characterization of a novel mutation (p.Arg511Trp). *Clinical Endocrinology* 2012 **78** 558–563. (doi:10.1111/cen.12048)
 - 17 Allemeersch J, Van Vooren S, Hannes F, De Moor B, Vermeesch JR & Moreau Y. An experimental loop design for the detection of constitutional chromosomal aberrations by array CGH. *BMC Bioinformatics* 2009 **10** 380. (doi:10.1186/1471-2105-10-380)
 - 18 Iafrate AJ, Feuk L, Rivera MN, Listewnik ML, Donahoe PK, Qi Y, Scherer SW & Lee C. Detection of large-scale variation in the human genome. *Nature Genetics* 2004 **36** 949–951. (doi:10.1038/ng1416)
 - 19 Firth HV, Richards SM, Bevan AP, Clayton S, Corpas M, Rajan D, Van Vooren S, Moreau Y, Pettett RM & Carter NP. DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans using Ensembl Resources. *American Journal of Human Genetics* 2009 **84** 524–533. (doi:10.1016/j.ajhg.2009.03.010)
 - 20 Moreau Y & Tranchevent LC. Computational tools for prioritizing candidate genes: boosting disease gene discovery. *Nature Reviews. Genetics* 2012 **13** 523–536. (doi:10.1038/nrg3253)
 - 21 Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM & Haussler D. The human genome browser at UCSC. *Genome Research* 2002 **12** 996–1006. (doi:10.1101/gr.229102)
 - 22 Kozomara A & Griffiths-Jones S. miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Research* 2011 **39** D152–D157. (doi:10.1093/nar/gkq1027)
 - 23 Hsu SD, Lin FM, Wu WY, Liang C, Huang WC, Chan WL, Tsai WT, Chen GZ, Lee CJ, Chiu CM *et al.* miRTarBase: a database curates experimentally validated microRNA–target interactions. *Nucleic Acids Research* 2010 **39** D163–D169. (doi:10.1093/nar/gkq1107)
 - 24 Eppig JT, Blake JA, Bult CJ, Kadin JA & Richardson JE. The Mouse Genome Database (MGD): comprehensive resource for genetics and genomics of the laboratory mouse. *Nucleic Acids Research* 2012 **40** D881–D886. (doi:10.1093/nar/gkr974)
 - 25 Seelow D, Schwarz JM & Schuelke M. GeneDistiller – distilling candidate genes from linkage intervals. *PLoS ONE* 2008 **3** e3874. (doi:10.1371/journal.pone.0003874)
 - 26 Shprintzen RJ. Velo-cardio-facial syndrome: 30 years of study. *Developmental Disabilities Research Reviews* 2008 **14** 3–10. (doi:10.1002/ddrr.2)
 - 27 Yatsenko SA, Kruer MC, Bader PI, Corzo D, Schuette J, Keegan CE, Nowakowska B, Peacock S, Cai WW, Peiffer DA *et al.* Identification of critical regions for clinical features of distal 10q deletion syndrome. *Clinical Genetics* 2009 **76** 54–62. (doi:10.1111/j.1399-0004.2008.01115.x)
 - 28 Strehle EM, Yu L, Rosenfeld JA, Donkervoort S, Zhou Y, Chen TJ, Martinez JE, Fan YS, Barbouth D, Zhu H *et al.* Genotype–phenotype analysis of 4q deletion syndrome: proposal of a critical region. *American Journal of Medical Genetics. Part A* 2012 **158A** 2139–2151. (doi:10.1002/ajmg.a.35502)
 - 29 Girirajan S, Campbell CD & Eichler EE. Human copy number variation and complex genetic disease. *Annual Review of Genetics* 2011 **45** 203–226. (doi:10.1146/annurev-genet-102209-163544)
 - 30 Dauber A, Yu Y, Turchin MC, Chiang CW, Meng YA, Demerath EW, Patel SR, Rich SS, Rotter JJ, Schreiner PJ *et al.* Genome-wide association of copy-number variation reveals an association between short stature and the presence of low-frequency genomic deletions. *American Journal of Human Genetics* 2011 **89** 751–759. (doi:10.1016/j.ajhg.2011.10.014)
 - 31 Guzman ML, Delgado I, Lay-Son G, Willans E, Puga A & Repetto GM. Growth in Chilean infants with chromosome 22q11 microdeletion syndrome. *American Journal of Medical Genetics. Part A* 2012 **158A** 2682–2686. (doi:10.1002/ajmg.a.35525)
 - 32 Witte S, Villalba M, Bi K, Liu Y, Isakov N & Altman A. Inhibition of the c-Jun N-terminal kinase/AP-1 and NF- κ B pathways by PICOT, a novel protein kinase C-interacting protein with a thioredoxin homology domain. *Journal of Biological Chemistry* 2000 **275** 1902–1909. (doi:10.1074/jbc.275.3.1902)
 - 33 Cheng NH, Zhang W, Chen WQ, Jin J, Cui X, Butte NF, Chan L & Hirschi KD. A mammalian monothiol glutaredoxin, Grx3, is critical for cell cycle progression during embryogenesis. *FEBS Journal* 2011 **278** 2525–2539. (doi:10.1111/j.1742-4658.2011.08178.x)
 - 34 Itoh M, Yoshida Y, Nishida K, Narimatsu M, Hibi M & Hirano T. Role of Gab1 in heart, placenta, and skin development and growth factor- and cytokine-induced extracellular signal-regulated kinase mitogen-activated protein kinase activation. *Molecular and Cellular Biology* 2000 **20** 3695–3704. (doi:10.1128/MCB.20.10.3695-3704.2000)
 - 35 Cookson W, Liang L, Abecasis G, Moffatt M & Lathrop M. Mapping complex disease traits with global gene expression. *Nature Reviews. Genetics* 2009 **10** 184–194. (doi:10.1038/nrg2537)
 - 36 Christiansen J, Kolte AM, Hansen T & Nielsen FC. IGF2 mRNA-binding protein 2: biological function and putative role in type 2 diabetes. *Journal of Molecular Endocrinology* 2009 **43** 187–195. (doi:10.1677/JME-09-0016)
 - 37 Lango Allen H, Estrada K, Lettre G, Berndt SI, Weedon MN, Rivadeneira F, Willer CJ, Jackson AU, Vedantam S, Raychaudhuri S *et al.* Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* 2010 **467** 832–838. (doi:10.1038/nature09410)
 - 38 Lolli G, Pinna LA & Battistutta R. Structural determinants of protein kinase CK2 regulation by autoinhibitory polymerization. *ACS Chemical Biology* 2012 **7** 1158–1163. (doi:10.1021/cb300054n)
 - 39 Lou DY, Dominguez I, Toselli P, Landesman-Bollag E, O'Brien C & Seldin DC. The α catalytic subunit of protein kinase CK2 is required for mouse embryonic development. *Molecular and Cellular Biology* 2008 **28** 131–139. (doi:10.1128/MCB.01119-07)
 - 40 Hanemaaijer NM, Sikkema-Raddatz B, van der Vries G, Dijkhuizen T, Hordijk R, van Essen AJ, Veenstra-Knol HE, Kerstjens-Frederikse WS, Herkert JC, Gerkes EH *et al.* Practical guidelines for interpreting copy number gains detected by high-resolution array in routine diagnostics. *European Journal of Human Genetics* 2012 **20** 161–165. (doi:10.1038/ejhg.2011.174)
 - 41 Nishiyama M, Skoultchi AI & Nakayama KI. Histone H1 recruitment by CHD8 is essential for suppression of the Wnt– β -catenin signaling pathway. *Molecular and Cellular Biology* 2012 **32** 501–512. (doi:10.1128/MCB.06409-11)
 - 42 Hao Y, Liu X, Lu X, Yang X, Wang L, Chen S, Li H, Li J, Cao J, Chen J *et al.* Genome-wide association study in Han Chinese identifies three novel loci for human height. *Human Genetics* 2013 **132** 681–689. (doi:10.1007/s00439-013-1280-9)
 - 43 Hendriksz CJ, Harmatz P, Beck M, Jones S, Wood T, Lachman R, Gravance CG, Orii T & Tomatsu S. Review of clinical presentation and diagnosis of mucopolysaccharidosis IVA. *Molecular Genetics and Metabolism* 2013 **110** 54–64. (doi:10.1016/j.ymgme.2013.04.002)
 - 44 Lugtenberg D, de Brouwer AP, Oudakker AR, Pfundt R, Hamel BC, van Bokhoven H & Bongers EM. Xq13.2q21.1 duplication encompassing the ATRX gene in a man with mental retardation, minor facial and genital anomalies, short stature and broad thorax. *American Journal of Medical Genetics. Part A* 2009 **149A** 760–766. (doi:10.1002/ajmg.a.32742)
 - 45 Hao J, Zhang S, Zhou Y, Liu C, Hu X & Shao C. MicroRNA 421 suppresses DPC4/Smad4 in pancreatic cancer. *Biochemical and*

- Biophysical Research Communications* 2011 **406** 552–557. (doi:10.1016/j.bbrc.2011.02.086)
- 46 Le Goff C, Mahaut C, Abhyankar A, Le Goff W, Serre V, Afejar A, Destree A, di Rocco M, Heron D, Jacquemont S *et al.* Mutations at a single codon in Mad homology 2 domain of SMAD4 cause Myhre syndrome. *Nature Genetics* 2011 **44** 85–88. (doi:10.1038/ng.1016)
- 47 Klopocki E, Schulze H, Strauss G, Ott CE, Hall J, Trotier F, Fleischhauer S, Greenhalgh L, Newbury-Ecob RA, Neumann LM *et al.* Complex inheritance pattern resembling autosomal recessive inheritance involving a microdeletion in thrombocytopenia-absent radius syndrome. *American Journal of Human Genetics* 2007 **80** 232–240. (doi:10.1086/510919)
- 48 Kumar RA, KaraMohamed S, Sudi J, Conrad DF, Brune C, Badner JA, Gilliam TC, Nowak NJ, Cook EH Jr, Dobyns WB *et al.* Recurrent 16p11.2 microdeletions in autism. *Human Molecular Genetics* 2008 **17** 628–638. (doi:10.1093/hmg/ddm376)
- 49 Sandbacka M, Laivuori H, Freitas E, Halttunen M, Jokimaa V, Morin-Papunen L, Rosenberg C & Aittomaki K. TBX6, LHX1 and copy number variations in the complex genetics of Mullerian aplasia. *Orphanet Journal of Rare Diseases* 2013 **8** 125. (doi:10.1186/1750-1172-8-125)

Received 20 March 2014

Revised version received 28 May 2014

Accepted 30 May 2014