SHORT REPORT

A syndrome of short stature, microcephaly and speech delay is associated with duplications reciprocal to the common Sotos syndrome deletion

Luis M Franco[1], Thomy de Ravel[2], Brett H Graham[1], Stephanie M Frenkel[3], Jozef Van Driessche[4], Pawel Stankiewicz[1,5], James R Lupski[1,3,6], Joris R Vermeesch[2] and Sau Wai Cheung*,[1]

Genomic rearrangements are an increasingly recognized mechanism of human phenotypic variation and susceptibility to disease. Sotos syndrome is characterized by overgrowth, macrocephaly, developmental delay and advanced osseous maturation. Haploinsufficiency of *NSD1*, caused by inactivating point mutations or deletion copy number variants, is the only known cause of Sotos syndrome. A recurrent 2 Mb deletion has been described with variable frequency in different populations. In this study, we report two individuals of different ethnic and geographical backgrounds, with duplications reciprocal to the common Sotos syndrome deletion. Our findings provide evidence for the existence of a novel syndrome of short stature, microcephaly, delayed bone development, speech delay and mild or absent facial dysmorphism. The phenotype is remarkably opposite to that of Sotos syndrome, suggesting a role for *NSD1* in the regulation of somatic growth in humans.

INTRODUCTION

Structural rearrangements in the human genome are an increasingly recognized mechanism of human phenotypic variation and susceptibility to disease.[1] Nonallelic homologous recombination (NAHR), nonhomologous end joining and fork stalling and template switching have been proposed as mechanisms by which recurrent and non-recurrent genomic rearrangements occur.[2,3] The primary mechanism for the generation of recurrent rearrangements causing genomic disorders is NAHR between region-specific low-copy repeats (LCRs). This mechanism predicts that, for each deletion in which the rearrangement breakpoints map within flanking LCRs, there is a reciprocal duplication. Such deletions are a known cause of human disease; identifying the reciprocal duplications and documenting their role in generating specific phenotypes is one of the challenges of modern human genetics. Duplications reciprocal to the recurrent deletions associated with velo-cardio-facial syndrome,[4] the Williams–Beuren syndrome[5] and the Smith–Magenis syndrome[6] have been documented. In addition, newly described phenotypes associated with deletions and duplications in regions of the genome previously not known to be causative of human disease[7–12] have arisen.

The Sotos syndrome (MIM no. 117550), described clinically in 1964,[13] involves prenatal-onset overgrowth, macrocephaly, scaphocephaly, expressive and receptive language delay, motor delay, advanced osseous maturation and enlargement of the cerebral ventricles and extra-cerebral fluid spaces. The only known cause of this syndrome is haploinsufficiency of the *NSD1* (nuclear receptor SET domain containing protein-1) gene in the distal long arm of chromosome 5

(5q35.2q35.3).[14] Although its precise functions remain largely unknown, NSD1 is believed to have a role in transcriptional regulation through methylation of histone lysine residues, the differential binding of its two nuclear receptor-interacting domains and chromatin–chromatin interactions.[15–17] The portion of the human genome surrounding *NSD1*, often referred to as the Sotos critical region, encompasses 1.1 Mb of DNA, contains approximately 21 genes (Figure 1a) and is flanked by two well-characterized low-copy repeats (Figure 1c): a proximal 390 kb repeat (Sos-PREP) and a distal 429 kb repeat (Sos-DREP).[18] A common microdeletion[19] mediated by directly oriented subunits within Sos-PREP and Sos-DREP[18] is the most common mutational mechanism in the Japanese population, whereas intragenic mutations are responsible for at least 80% of the reported cases in European and North American patients.

A large duplication including this and adjacent regions, detected by fluorescence *in situ* hybridization (FISH),[20] and a small duplication contained within the region, detected by multiplex ligation-dependent probe amplification (MLPA),[21] have been reported. However, the duplication reciprocal to the common deletion of the Sotos critical region predicted by nonallelic homologous recombination at the flanking LCRs has not been described to date, and its phenotypic consequences in humans are not known. In this study, we present two individuals of different ethnic and geographical backgrounds, with a similar clinical presentation and a duplication of this genomic region reciprocal to the common Sotos syndrome deletion.

[1]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, USA; [2]Centre for Human Genetics, Universitaire Ziekenhuizen Leuven, Leuven, Belgium; [3]Texas Children's Hospital, Houston, TX, USA; [4]VZW Martine Van Camp, Diest, Belgium; [5]Department of Medical Genetics, Institute of Mother and Child, Warsaw, Poland; [6]Department of Pediatrics, Baylor College of Medicine, Houston, TX, USA
*Correspondence: Professor SW Cheung, Department of Molecular and Human Genetics, One Baylor Plaza, Mail Stop NAB 2015, Houston, TX 77030, USA.
Tel: +1 713 798 4991; Fax: +1 713 798 3157; E-mail: scheung@bcm.edu
Received 28 April 2009; revised 9 July 2009; accepted 10 August 2009; published online 21 October 2009
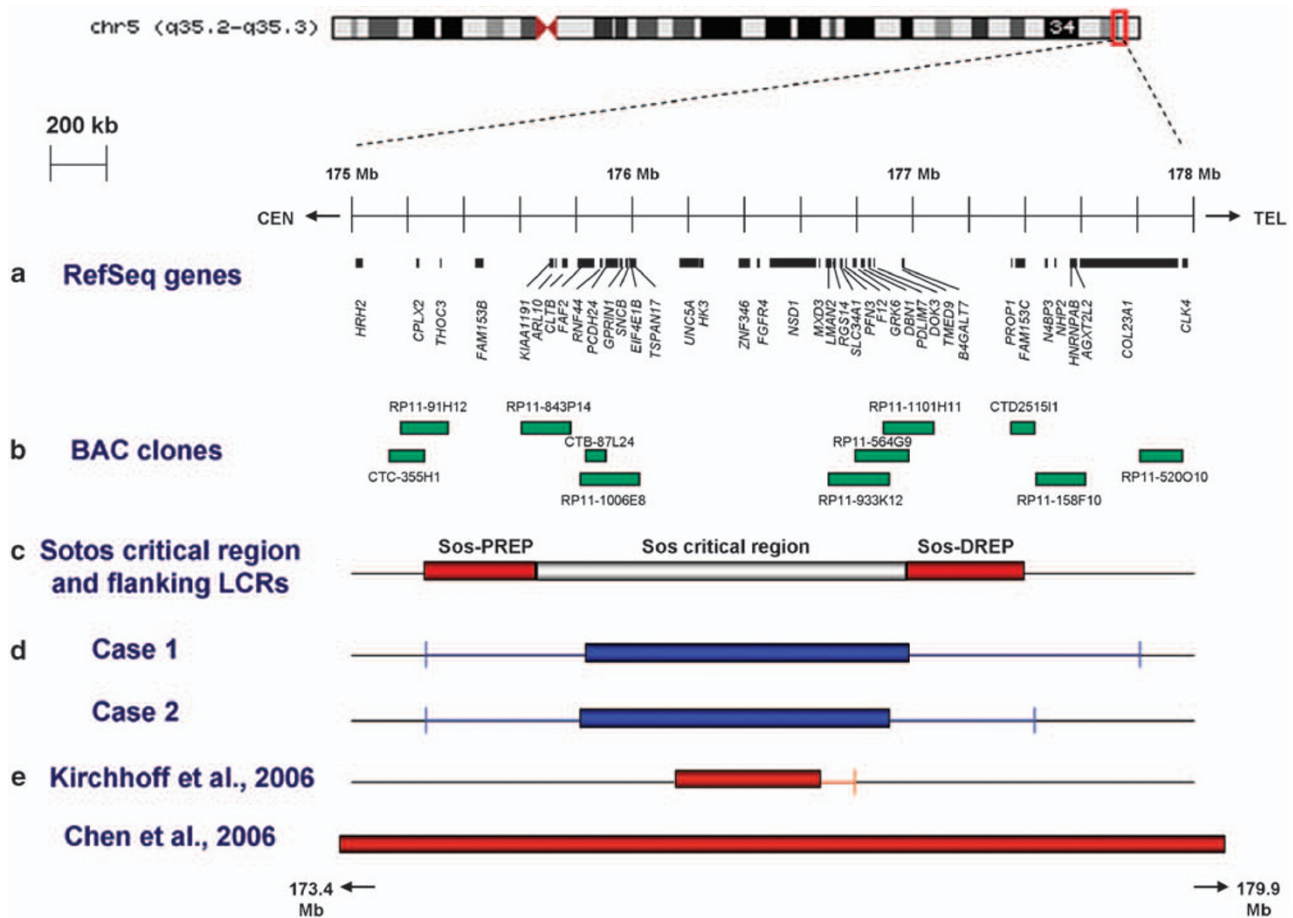
Figure 1 Distal long arm of chromosome 5. Scale representation of the duplicated region in the distal long arm of chromosome 5. (**a**) Location of RefSeq genes in the region. (**b**) Location of BAC clones cited in the text. (**c**) Location of the low-copy repeats flanking the Sotos syndrome critical region. (**d**) Duplicated fragments in cases 1 and 2. (**e**) Previously described duplications in the region. In **d** and **e**, boxes indicate the minimum boundaries of the duplicated segments and whiskers indicate the maximum boundaries of the duplicated fragments. RefSeq: National Center for Biotechnology Information Reference Sequence. BAC: bacterial artificial chromosome. LCR: low-copy repeat.
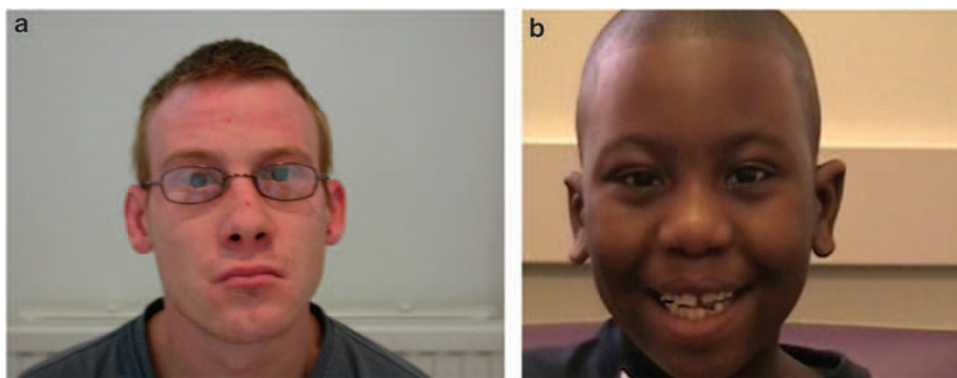

Figure 2 Facial characteristics. Facial characteristics of case 1 (**a**) and case 2 (**b**). Microcephaly was present. Note the absence of other dysmorphic features.

CLINICAL REPORTS

Case 1

The patient is a 33-year-old man, referred to the genetics clinic for developmental delay. He was the product of an uneventful pregnancy. A fetus papyraceus twin was detected at the time of delivery. At 4 years of age, the patient's parents noted that he was slow in the acquisition of psychomotor milestones. He required special assistance during primary education and attended a career-oriented high school. His IQ was assessed as 69 on WAIS testing, but he functions at a lower level. He has mood swings, which take him from being extremely active and motivated to dysphoric and withdrawn. He was admitted into a psychiatric institution at the age of 21 years for treatment of aggressive behavior and alcohol abuse. He has a dependent personality, now lives in a supervised environment and volunteers at a local

institution for the elderly. His speech is fluent. Physical examination at 33 years of age revealed microcephaly (head circumference: 51.3 cm; −3.5 SD), short stature (166.5 cm; percentile 6 for age) and mild myopia. No other dysmorphic features were present (Figure 2a) and the remainder of the physical examination was unremarkable. Karyotype was 46,XY.

Case 2

The patient is an 8-year-old boy, referred to the genetics clinic for developmental delay. He was the product of an uncomplicated pregnancy and was born after 40 weeks of gestation. There were no perinatal complications. He met gross and fine motor developmental milestones at appropriate ages, did not exhibit repetitive behaviors and was social and interactive. His medical and surgical histories were unremarkable and he did not require the use of any medications. Developmental delay was initially suspected at 5 years of age, when he entered school. His teachers noted that the patient had difficulty understanding them, learning to read and following two-step commands. The acquisition of language skills was slower than that of his peers and he was unable to learn how to read. At the age of 7 years, the patient's reported Full-Scale IQ on the Stanford-Binet test was 62 (mean=100, SD=15). At 8 years of age, Clinical Evaluation of Language Fundamentals (CELF-4) revealed below-average core language skills (score=44; mean=100, SD=15), with a discrepancy between his receptive language skills (score=65) and his expressive language skills (score=49). Physical examination revealed microcephaly (head circumference 48.5 cm; −3.5 SD) and short stature (116 cm; −2.4 SD). No other dysmorphic features were present (Figure 2b). His weight was in the 13th percentile for age, his height-to-weight ratio was in the 75th–85th percentile for age and the remainder of the physical examination was unremarkable. Radiological bone age, estimated by the standard of Greulich and Pyle at 7 years and 8 months of age, was delayed by 20 months (SD=10.8 months). Magnetic resonance imaging of the brain at 8 years of age showed a normal anatomy. Karyotype was 46, XY.

MATERIALS AND METHODS

DNA samples

Case 1 was referred to the Centre for Human Genetics, Universitaire Ziekenhuizen, Leuven, Belgium, for clinical array comparative genomic hybridization (aCGH) analysis.

Case 2 was referred to the Medical Genetics Laboratories, Baylor College of Medicine, Houston, United States of America, for clinical aCGH analysis. DNA was extracted from whole blood using the Puregene DNA extraction kit (Gentra, Minneapolis, MN, USA) according to the manufacturer's instructions.

Array comparative genomic hybridization (aCGH)

DNA from Case 1 was hybridized to a custom BAC array, with clones spaced at approximately 1 Mb intervals throughout the genome, according to the protocol of Menten *et al*.[22] Regions were scored as copy number variants if one clone passed the threshold of $4\times SD$, and if two or more flanking clones passed the threshold of $\log_2(3/2)-2\times SD$.[23]

DNA from Case 2 was hybridized to a custom BAC-emulated oligonucleotide array (Agilent Technologies, Santa Clara, CA, USA). The array comprises a total of 42 640 oligonucleotide probes localized to the genomic positions of 1400 FISH-verified bacterial artificial chromosome (BAC) clones covering more than 140 regions implicated in genetic diseases, as well as clinically relevant subtelomeric and pericentromeric regions. This targeted oligonucleotide array was manufactured in a $4\times44K$ format with an average of 28–30 oligos per region previously covered by a single BAC clone and a minimum of 10–15 oligos whenever possible. The microarray was designed to include oligonucleotide probes within the Sotos syndrome critical region, including probes at the proximal end of Sos-PREP (corresponding to BAC RP11-91H12) and at the distal end of Sos-DREP (corresponding to BAC CTD-2515I1), in

addition to probes immediately proximal to Sos-PREP (corresponding to BAC CTC-355H1) and immediately distal to Sos-DREP (corresponding to BAC RP11-158F10) (Figure 1b). DNA digestion, labeling, hybridization and data analysis were performed as previously described.[24]

Real-time quantitative PCR

The duplication on Case 1 was confirmed using real-time quantitative PCR (RT-qPCR), as previously described,[19] with a probe within CTB-87L24 (Figure 1b) in the duplicated region. The parents' DNA was also investigated using RT-qPCR.

FISH analysis

The duplication on Case 2 was confirmed by FISH. BAC DNA probes specific for the Sotos region were isolated using the QIAprep Spin Mini kit (Qiagen Inc., Valencia, CA, USA). Probes were labeled by nick translation using Spectrum Red and Spectrum Green Vysis dUTPs (Vysis Inc., Des Plaines, IL, USA). Hybridization was performed according to the manufacturer's protocol (Abbott Molecular, Abbott Park, IL, USA). The mother's DNA was also investigated using FISH.

Genome sequence analysis of the 5q35.2q35.3 region

A scale representation of the human 5q35.2q35.3 region in the March 2006 reference sequence (NCBI Build 36.1) was generated (Figure 1) using the UCSC Genome Browser (http://genome.ucsc.edu/).[25] The specific locations in the current reference sequence of the BACs of interest were determined using the BLAST-like alignment tool (BLAT)[26] to map the most centromeric and telomeric sequences of each BAC.

RESULTS

Case 1

Array CGH revealed a gain in copy number in the long arm of chromosome 5 [arr cgh 5q35. 2q35.3(CTB-87L24→RP11-564G9)×3] (Figure 1d), corresponding to the Sotos syndrome critical region, with flanking clones CTC-355H1 and RP11-520O10 showing normal copy number (Figure 1b). The duplication was confirmed by RT-qPCR. The parents were shown not to have the duplication.

Case 2

Array CGH showed a gain in copy number in the long arm of chromosome 5 [arr cgh 5q35.2q35.3(RP11-1006E8→RP11-933K12)×3] (Figure 1d), corresponding to the Sotos syndrome critical region, with flanking clones CTC-355H1 and RP11-158F10 (Figure 1b) showing normal copy number. The duplication was confirmed by FISH [nuc ish. 5q35.2(RP11-1006E8×3)]. The patient's mother's DNA was shown not to have the duplication, using FISH analysis. A paternal sample could not be obtained.

DISCUSSION

The phenotypic and genomic similarities between our two patients suggest a novel syndrome of microcephaly, short stature, developmental delay and delayed bone maturation, in the absence of facial or osseous dysmorphism, caused by duplications of the Sotos syndrome critical region. Remarkably, the observed phenotype is reciprocal to that of patients with Sotos syndrome, who commonly present with macrocephaly, overgrowth and advanced bone age.

The breakpoints of the duplication in both patients map to the proximal and distal LCRs that flank the Sotos critical region (Figure 1d), strongly suggesting NAHR as the underlying mechanism of their genomic imbalance and indicating that the predicted duplication reciprocal to the common deletion of this region can indeed be found in human populations.

The previously reported patient with a duplication involving this and adjacent regions of the genome (Figure 1e, bottom) had a more

severe phenotype than did our patients, including brachydactyly, abnormally shaped fifth digits, strabismus, bilateral inguinal hernias and facial dysmorphism, in addition to microcephaly, short stature and developmental delay.[20] A patient with a smaller (520–650 kb) duplication within the Sotos syndrome critical region, also encompassing the *NSD1* gene but not reciprocal to the common deletion (Figure 1e, top), had microcephaly and short stature in addition to mild facial dysmorphism.[21] This suggests that the gene dosage effect that seems to be responsible for this phenotype is limited to a region smaller than the 1.1 Mb minimum-duplicated fragment found in our patients, and that it is likely to involve the *NSD1* gene.

A reciprocal phenotype between deletion and duplication cases has been observed in some,[7] but not all, genomic disorders reported to date. The striking reciprocity of the phenotype between deletion and duplication cases in the Sotos syndrome critical region suggests a strong dosage effect of the *NSD1* gene. Although the exact physiological functions of *NSD1* remain largely unknown, our data support the hypothesis of a role for this gene as a regulator of somatic growth in humans. Functional studies and animal models may help to further clarify this.

When systematically studied, a majority of point mutations in the *NSD1* gene in patients with Sotos syndrome are nonsense or frameshift mutations.[27] Our data raise the possibility of gain-of-function mutations at this locus being a cause of microcephaly, short stature and developmental delay.

The paucity of dysmorphic features in our patients may help explain, at least in part, why deletions in this region are a well-known cause of human disease, whereas the phenotype associated with the reciprocal duplication is only now being uncovered. Given the population distribution of the common Sotos syndrome microdeletion, for example, one would expect the frequency of its reciprocal duplication to be relatively high in people of Japanese ancestry. We advocate the clinical use of techniques that allow the detection of copy number changes not visible by standard cytogenetic techniques, such as aCGH and MLPA, in patients with microcephaly and/or developmental delay, in the absence of other dysmorphic features. This should allow other cases to be uncovered, providing further insight into the clinical manifestations of this novel syndrome.

CONFLICT OF INTEREST
The authors declare no conflict of interest.

1 Lupski JR: Structural variation in the human genome. *N Engl J Med* 2007; **356**: 1169–1171.

2 Lupski JR, Stankiewicz P: Genomic disorders: molecular mechanisms for rearrangements and conveyed phenotypes. *PLoS Genet* 2005; **1**: e49.

3 Lee J, Carvalho CM, Lupski JR: A DNA replication mechanism for generating nonrecurrent rearrangements associated with genomic disorders. *Cell* 2007; **131**: 1235–1247.

4 Ensenauer RE, Adeyinka A, Flynn HC et al: Microduplication 22q11.2, an emerging syndrome: clinical, cytogenetic, and molecular analysis of thirteen patients. *Am J Hum Genet* 2003; **73**: 1027–1040.

5 Somerville MJ, Mervis CB, Young EJ et al: Severe expressive-language delay related to duplication of the Williams-Beuren locus. *N Engl J Med* 2005; **353**: 1694–1701.

6 Potocki L, Chen KS, Park SS et al: Molecular mechanism for duplication 17p11.2 – the homologous recombination reciprocal of the Smith-Magenis microdeletion. *Nat Genet* 2000; **24**: 84–87.

7 Brunetti-Pierri N, Berg JS, Scaglia F et al: Recurrent reciprocal 1q21.1 deletions and duplications associated with microcephaly or macrocephaly and developmental and behavioral abnormalities. *Nat Genet* 2008; **40**: 1466–1471.

8 Ballif BC, Theisen A, Coppinger J et al: Expanding the clinical phenotype of the 3q29 microdeletion syndrome and characterization of the reciprocal microduplication. *Mol Cytogenet* 2008; **1**: 8.

9 Weiss LA, Shen Y, Korn J et al: Association between Microdeletion and Microduplication at 16p11.2 and Autism. *N Engl J Med* 2008; **358**: 667–675.

10 Sharp AJ, Hansen S, Selzer RR et al: Discovery of previously unidentified genomic disorders from the duplication architecture of the human genome. *Nat Genet* 2006; **38**: 1034–1042.

11 Koolen DA, Vissers LE, Pfundt R et al: A new chromosome 17q21.31 microdeletion syndrome associated with a common inversion polymorphism. *Nat Genet* 2006; **38**: 999–1001.

12 Shaw-Smith C, Pittman AM, Willatt L et al: Microdeletion encompassing MAPT at chromosome 17q21.3 is associated with developmental delay and learning disability. *Nat Genet* 2006; **38**: 974–976.

13 Sotos JF, Dodge PR, Muirhead D, Crawford JD, Talbot NB: Cerebral gigantism in childhood. A syndrome of excessively rapid growth and acromegalic features and a nonprogressive neurologic disorder. *N Engl J Med* 1964; **271**: 109–116.

14 Kurotaki N, Imaizumi K, Harada N et al: Haploinsufficiency of *NSD1* causes Sotos syndrome. *Nat Genet* 2002; **30**: 365–366.

15 Rayasam GV, Wendling O, Angrand PO et al: *NSD1* is essential for early post-implantation development and has a catalytically active SET domain. *EMBO J* 2003; **22**: 3153–3163.

16 Huang N, vom Baur E, Garnier JM et al: Two distinct nuclear receptor interaction domains in NSD1, a novel SET protein that exhibits characteristics of both corepressors and coactivators. *EMBO J* 1998; **17**: 3398–3412.

17 Stec I, Nagl SB, van Ommen GJ, den Dunnen JT: The PWWP domain: a potential protein-protein interaction domain in nuclear proteins influencing differentiation? *FEBS Lett* 2000; **473**: 1–5.

18 Kurotaki N, Stankiewicz P, Wakui K, Niikawa N, Lupski JR: Sotos syndrome common deletion is mediated by directly oriented subunits within inverted Sos-REP low-copy repeats. *Hum Mol Genet* 2005; **14**: 535–542.

19 Kurotaki N, Harada N, Shimokawa O et al: Fifty microdeletions among 112 cases of Sotos syndrome: low copy repeats possibly mediate the common deletion. *Hum Mutat* 2003; **22**: 378–387.

20 Chen CP, Lin SP, Lin CC et al: Molecular cytogenetic analysis of de novo dup(5)(q35.2q35.3) and review of the literature of pure partial trisomy 5q. *Am J Med Genet A* 2006; **140**: 1594–1600.

21 Kirchhoff M, Bisgaard AM, Bryndorf T, Gerdes T: MLPA analysis for a panel of syndromes with mental retardation reveals imbalances in 5.8% of patients with mental retardation and dysmorphic features, including duplications of the Sotos syndrome and Williams-Beuren syndrome regions. *Eur J Med Genet* 2007; **50**: 33–42.

22 Menten B, Maas N, Thienpont B et al: Emerging patterns of cryptic chromosomal imbalance in patients with idiopathic mental retardation and multiple congenital anomalies: a new series of 140 patients and review of published reports. *J Med Genet* 2006; **43**: 625–633.

23 Vermeesch J, Melotte C, Froyen G et al: Molecular karyotyping: array CGH quality criteria for constitutional genetic diagnosis. *J Histochem Cytochem* 2005; **53**: 413–422.

24 Ou Z, Kang SH, Shaw CA et al: Bacterial artificial chromosome-emulation oligonucleotide arrays for targeted clinical array-comparative genomic hybridization analyses. *Genet Med* 2008; **10**: 278–289.

25 Kent WJ, Sugnet CW, Furey TS et al: The human genome browser at UCSC. *Genome Res* 2002; **12**: 996–1006.

26 Kent WJ: BLAT – the BLAST-like alignment tool. *Genome Res* 2002; **12**: 656–664.

27 Saugier-Veber P, Bonnet C, Afenjar A et al: Heterogeneity of *NSD1* alterations in 116 patients with Sotos syndrome. *Hum Mutat* 2007; **28**: 1098–1107.