ORIGINAL ARTICLE

Multiple mechanisms are implicated in the generation of 5q35 microdeletions in Sotos syndrome

K Tatton-Brown, J Douglas, K Coleman, G Baujat, K Chandler, A Clarke, A Collins, S Davies, F Faravelli, H Firth, C Garrett, H Hughes, B Kerr, J Liebelt, W Reardon, G B Schaefer, M Splitt, I K Temple, D Waggoner, D D Weaver, L Wilson, T Cole, V Cormier-Daire, A Irrthum, N Rahman on behalf of the Childhood Overgrowth Collaboration

Background: Sotos syndrome (MIM 117550) is characterised by learning difficulties, overgrowth, and a typical facial appearance. Microdeletions at 5q35.3, encompassing *NSD1*, are responsible for ~10% of non-Japanese cases of Sotos. In contrast, a recurrent ~2 Mb microdeletion has been reported as responsible for ~50% of Japanese cases of Sotos.

Methods: We screened 471 cases for *NSD1* mutations and deletions and identified 23 with 5q35 microdeletions. We investigated the deletion size, parent of origin, and mechanism of generation in these and a further 10 cases identified from published reports. We used ''in silico'' analyses to investigate whether repetitive elements that could generate microdeletions flank *NSD1*.

Results: Three repetitive elements flanking *NSD1*, designated REPcen, REPmid, and REPtel, were identified. Up to 18 cases may have the same sized deletion, but at least eight unique deletion sizes were identified, ranging from 0.4 to 5 Mb. In most instances, the microdeletion arose through interchromosomal rearrangements of the paternally inherited chromosome.

Conclusions: Frequency, size, and mechanism of generation of 5q35 microdeletions differ between Japanese and non-Japanese cases of Sotos. Our microdeletions were identified from a large case series with a broad range of phenotypes, suggesting that sample selection variability is unlikely as a sole explanation for these differences and that variation in genomic architecture might be a contributory factor. Non-allelic homologous recombination between REPcen and REPtel may have generated up to 18 microdeletion cases in our series. However, at least 15 cannot be mediated by these repeats, including at least seven deletions of different sizes, implicating multiple mechanisms in the generation of 5q35 microdeletions.

See end of article for authors' affiliations

Correspondence to:
Dr N Rahman, Section of Cancer Genetics, Brookes Lawley Building, Institute of Cancer Research, 15 Cotswold Road, Sutton, Surrey SM2 5NG, UK; nazneen.rahman@icr.ac.uk

Sotos syndrome (OMIM 117550) is characterised by a typical facial appearance, learning difficulties, and overgrowth in childhood. Additional features include neonatal jaundice and hypotonia, cardiac and renal anomalies, scoliosis, seizures, and tumours.[1–3] Sotos syndrome is caused by mutations and deletions of *NSD1*, a histone methyltransferase implicated in transcriptional regulation, which is located at chromosome 5q35.[4–7]

We previously reported that *NSD1* intragenic mutations cause ~76% of UK cases of Sotos syndrome, whereas microdeletions encompassing *NSD1* cause ~10% of cases.[5] Similar results were reported for cases of Sotos from France, Germany, Italy, and the USA.[6–9] However, intragenic mutations were reported in only 12% of Japanese cases of Sotos, whereas a recurrent ~2 Mb microdeletion encompassing *NSD1* was identified in ~50%.[10] It has been suggested that these contrasting results could be attributed to differences in sample selection, as the Japanese microdeletion cases did not all meet the stringent criteria for classic Sotos syndrome.

In this study we screened 471 cases with varying degrees of overgrowth and/or phenotypic overlap with Sotos syndrome for *NSD1* mutations and deletions, to further evaluate the 5q35 microdeletion frequency in the UK and the possible reasons for differences between Japanese and non-Japanese populations. We performed microsatellite analyses in 33 cases with an *NSD1* microdeletion to establish the size, parental origin, and mechanism of generation of the deletions, and undertook ''in silico'' analyses to identify sequence elements that may mediate the deletions.

SUBJECTS AND METHODS

Cases

The research was approved by the London Multicentre Research Ethics Committee, and consent was obtained from participating cases and/or parents. Through analyses of 471 cases ascertained through the Childhood Overgrowth Collaboration, 23 cases with microdeletions were identified. These 471 cases included ~200 with a clinical diagnosis of Sotos syndrome or a Sotos-like syndrome. However, the majority of cases either had overgrowth and/or macrocephaly but did not have the facial gestalt of Sotos syndrome, or had facial features similar to Sotos syndrome, but no overgrowth. All 471 cases were screened for mutations and whole gene deletions of *NSD1* as previously described.[5] Fifteen microdeletion cases identified in these analyses were from the UK, and three (COG025, COG044, COG070) have been previously published.[5] The remaining cases were from the Republic of Ireland, Australia, the USA, and France. Patients COG508a
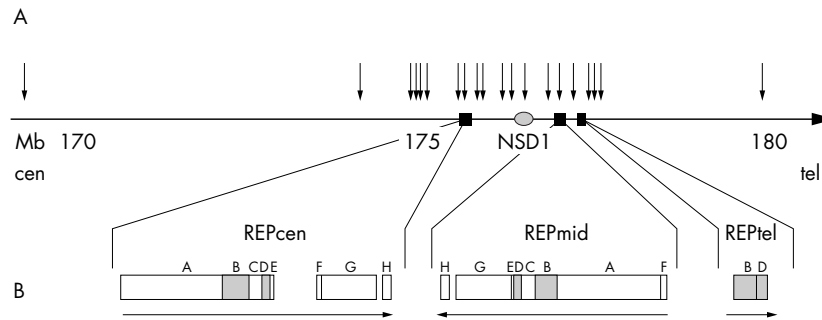
Figure 1 Schematic representation of 10 Mb region surrounding NSD1. (A) Positions of low copy repeat elements REPcen, REPmid, and REPtel, and NSD1. Black arrows represent the microsatellite markers analysed. The names and positions of theses microsatellite markers are shown in fig 2. SOT26 is shown twice as it is duplicated in REPcen[H] and REPmid[H]. (B) Enlarged representation demonstrating block composition and orientation.

and COG508b were identical twins. *NSD1* mutations were identified in 146 of the 471 cases.

We also ascertained microdeletion cases from published reports. Six cases were from France (COG342, COG343, COG344, COG345, COG346, COG347[6]) three from the USA (COG183, COG184, COG399[8]) and one from Italy (COG548[9]). DNA was obtained from both parents for 19 cases, from one parent for three cases, and was unobtainable for the remaining 11 cases. DNA from grandparents and/or siblings was obtained for 11 cases.

All cases with a microdeletion had been clinically diagnosed with Sotos syndrome prior to the molecular analyses. Clinical details were obtained in all cases except COG179, and photographs were available for 21 cases. These were independently assessed by five of us (TC, HEH, IKT, NR, and KT-B) and all were considered typical of Sotos syndrome.

Identification and delineation of 5q35 microdeletions

Microdeletions were identified and confirmed using at least two separate methods. Cases ascertained from the literature were identified by fluorescent in situ hybridisation using at least one intragenic *NSD1* probe.[6 8 9] The microdeletions were confirmed by multiplex ligation dependent probe amplification (MLPA) using the SALSA P026 *NSD1* test kit and the methods described in Schouten *et al*.[10] This kit contains probes for *NSD1* exons 1, 2, 3, 5, 6, 11, 14, 17, and 23, *FGFR4* exon 2, and 17 control probes. Microdeletion cases ascertained through the Childhood Overgrowth Collaboration were initially identified by analysis of the microsatellite marker, SOT3.[5] This intragenic marker is highly polymorphic and within intron 2 of *NSD1*. Cases homozygous at SOT3 were further investigated with quantitative fluorescent PCR as previously described[5] and identified microdeletions were confirmed using MLPA.

To determine the size of the microdeletions, polymorphic microsatellite markers within and surrounding *NSD1* in cases and parental DNA were analysed. We had previously published four of these markers, SOT1, SOT12, SOT3, and SOT19,[5] and we developed new markers using the UCSC Human Genome Project Working Draft sequence. The 2 Mb region encompassing *NSD1* was searched for dinucleotide, trinucleotide, and tetranucleotide repeat elements and amplifying primers were designed using Primer 3 software. Twelve markers were informative and worked reliably in the analyses (Appendix 1). We also used three known microsatellite markers, D5S400, D5S2008, and D5S2073, and one *FGFR* probe from the SALSA P026 *NSD1* kit. The position of the 20 markers relative to *NSD1* is shown in figs 1A and 2. The forward primer for each marker was labelled with $\gamma[^{32}P]$-ATP, amplified using PCR and the product was electrophoresed on a denaturing polyacrylamide gel. The gels were exposed to *x* ray film and the positions of the product bands were scored relative to each other.

To determine the parental origin and mechanism of generation of microdeletions we analysed 12 microsatellite markers from chromosome 5q35 in DNA from cases, parents, grandparents, and siblings. The order and distance between these markers was cen-D5S436-17 Mb-D5S422-6 Mb-D5S400-2.9 Mb-D5S429-3.7 Mb-SOT30-1 Mb-SOT27-0.3 Mb - SOT1 - 0.2 Mb-SOT3-1 Mb - D5S2008-19 kb - SOT23 - 1.6 Mb-D5S2073-1.2 Mb-D5S2006-tel. Haplotypes of marker alleles were determined for each family.

In silico analyses to identify and characterise repetitive elements flanking *NSD1*

Bioinformatic analyses were based on the May 2004 human genome sequence assembly, accessed at UCSC. To identify low copy repeat elements within the 3 Mb region surrounding *NSD1*, we divided the sequence into 2 kb fragments and

Table 1 Size and position of low copy repeat elements flanking NSD1

Block	Size (kb)	REP-cen	REP-mid		REP-tel	
		Position* (kb)	% homology to REPcen	Position* (kb)	% homology to REPcen	Position* (kb)
A	156	175262	99.2	177140		
B	36	175418	99.2	177104	98.9	177364
C	14	175454	99.5	177092		
D	14	175468	99.2	177080	98.6	177400
E	10	175482	98.6	177070		
F	10	175552	98.4	177296		
G	72	175562	98.6	176998		
H	16	175646	96.9	176980		

*The precision of the coordinates is ±2 kb.

REP-cen^BD   REP-tel^BD

NSD1

	DS5400, 168375448*	SOT21, 174051635	SOT32, 174859934	SOT31, 174912519	SOT30, 174951061	SOT29, 175073217	SOT28, 175532388	SOT26, 175647901	SOT12, 175875110	SOT27, 175973600	SOT1, 176312978	FGFR4, 176449173‡	SOT3, 176516994	SOT19, 176940350	SOT26, 176993492	SOT25, 177327307	DS52008, 177498020	SOT24, 177457647	SOT23, 177517458	D5S2073, 179112542
C0G025†										D§	U	D	D	D	D	D	D	D	U	D
C0G231		U	D	D	D	D	D	D	D	D	D	D	D	D	D	D	U	D	D	
C0G342		U	D	D	D	D	D	D	D	D	D	D	D	D	D	D	U	D	U	
C0G183			U	U	U	D	D	D	D	D	D	D	D	D	D	D	U	U	U	
C0G070			U	U	U	U	D	D	D	D	D	D	D	D	D	D				
C0G064				N		U	U	U	N	U	U	D	D	D	D	N	U	N	N	
C0G547					N	N	U	N	U	U	D	D	N	U	N	U	N	N	N	
C0G218					U	N	N	U	U	U	D	D	D	N	U	U				
C0G347					U	D	D	D	D	D	D	D	D	D						
C0G344					U	D	D	D	D	D	D	D	D	D						
C0G172					U	D	U	U	U	D	D	D	U	U	U					
C0G548					U	U	U	N	D	U	D	D	N	U	N					
C0G177					U	U	U	U	U	D	D	D	U	U	U					
C0G179					U	U	U	U	U	D	D	D	U	U	U					
C0G211					U	U	U	U	U	D	D	D	U	U	U					
C0G146					U	U	U	U	U	D	D	D	U	U	U					
C0G160						U	U	U	U	D	D	D	U	U	U					
C0G346						D	D	D	D	D	D	D	D	D						
C0G343						D	D	D	D	D	D	D	D	D						
C0G184						U	D	D	D	D	D	D	D							
C0G330						U	U	U	U	U	D	D	U							
C0G532										N	U	U	D	D	N	U	U	U	U	
C0G395										N	U	U	D	D	N	U	U	N	U	U
C0G450										N	U	U	D	D	N	N	U			
C0G512										N	U	U	D	D	N	U	U			
C0G399									U	U	U	U	D	D	U	U				
C0G340					U	U	D	D	D	D	D	D	D							
C0G044									U	U	U	D	D	D	D					
C0G508¶										D	D	D	D	D	D					
C0G345											U	D	D	D						
C0G367												D	D	D						
C0G111												D								

Figure 2 Results of microsatellite analyses in cases with a 5q35 microdeletion. *Name and position of first base of primer. †The cases are approximately arranged in order of the largest to smallest deletion size. Data were uninformative (U) or unobtainable (N) at some markers in some cases and thus precise deletion size could not always be determined. ‡FGFR4 is an intragenic probe in the SALSA PO26 NSD1 MLPA kit. §D, deleted; shaded box, retained; U, uninformative data; N, no data. ¶COG508 represents deletion in identical twins COG508a and COG508b.

compared them to one another using the BLAT program[12] and a Perl script. Segment pairs with >90% sequence similarity over more than 200 bp were considered for further analysis, and plotted onto a 1500×1500 sequence similarity matrix, where duplicated regions clearly appeared as diagonals. To confirm that these duplicated regions were low copy repeats, the analysis was repeated with genomic sequence pre-masked for high copy repeats with RepeatMasker. The web based BLAT server at UCSC was used to validate these results and to obtain the percentage of sequence identity over the duplicated regions. The high copy repeat content of the NSD1 region was calculated with RepeatMasker using default settings.

RESULTS
NSD1 is flanked by three homologous low copy repeat elements
We identified three duplicated areas flanking NSD1 (fig 1). The homology between these regions is very high, generally >98% (table 1). We designated the flanking low copy repeat elements REPcen, REPmid, and REPtel. REPcen is centromeric to NSD1 and consists of eight blocks (REPcen^A to REPcen^H) ranging in size from 10 to 156 kb (table 1). The

total duplicated region is 328 kb, but this is distributed over 400 kb, as there is a 60 kb region between REPcen^E and REPcen^F and a 12 kb region between REPcen^G and REPcen^H that is not duplicated in REPmid or REPtel. REPmid is telomeric to NSD1 and is inversely orientated to REPcen. It contains all eight blocks but their order is cen-H-2 kb-GEDCBAF-tel. REPtel is 60 kb telomeric to REPmid and in the same orientation as REPcen. It consists only of blocks B and D (fig 1B).

Sotos microdeletions are variable in size
We analysed 20 microsatellite markers in 33 cases with a microdeletion. Parental samples, if available, were also analysed. These analyses confirmed the presence of a microdeletion in all cases. SOT26 was duplicated in block REPcen^H and REPmid^H and therefore provided information about centromeric and telomeric breakpoints. The size of the deletion ranged from 482 kb (COG111) in whom NSD1 was the only known gene deleted, to 5 Mb (COG025) in whom 54 known genes were deleted. There were at least eight unique deletion sizes (fig 2). Deletion mapping in four cases, COG343, COG344, COG346, and COG347, suggested the breakpoints may have occurred in REPcen and REPtel giving a deletion size of ~1.9 Mb. A further 14 cases may also have breakpoints within these repeat elements, but the remaining 15 cases are not consistent with breakpoints in both REPcen and REPtel. The clinical features of the microdeletion cases are shown in table 2.

Sotos microdeletions are primarily generated by interchromosomal rearrangements and are usually paternally derived
We investigated the mechanism generating microdeletions in 11 cases where DNA from grandparents and/or siblings was available. Eight deletions, of varying sizes, arose through interchromosomal rearrangements (COG064, COG111, COG183, COG184, COG344, COG346, COG347, COG512). Two deletions of different sizes arose through intrachromosomal rearrangements (COG231, COG044). One cases was the result of a terminal deletion (COG025) (fig 3).

We determined the parent of origin of the deletion in 21 cases for whom parental DNA was available. The paternal allele was deleted in 18/21 cases and the maternal allele was deleted in the remaining three cases. The paternally derived deletions included several that were of different sizes and the three maternally derived deletions included the smallest and largest deletion. The bias towards deletion of the paternal allele was statistically significant using the two tailed exact binomial test (p = 0.0015).

DISCUSSION
Identified low copy repeats flanking NSD1 may mediate some, but not all, microdeletions
Through in silico analyses, we identified three low copy repeats (LCRs); one centromeric (REPcen) and two telomeric (REPmid, REPtel) to NSD1. REPcen^BD and REPtel^BD are ~50 kb in size, in the same orientation and show >98% homology, and are therefore potential substrates for non-allelic homologous recombination (NAHR), a mechanism implicated in the generation of several microdeletion syndromes.[13] Up to 18 of 33 microdeletions may be attributable to NAHR between REPcen and REPtel, resulting in a recurrent ~1.9 Mb deletion. However, in 14 of these cases, the microsatellite analyses were uninformative at multiple markers, and therefore these deletions may not be of uniform size and some may be attributable to other mechanism(s). Fifteen microdeletions were not consistent with NAHR between REPcen and REPtel, including at least seven distinct deletions. We did not identify other LCRs

Table 2 The clinical features of cases with 5q35 microdeletions

Case	Age[a]	Facial gestalt	LD	OFC	Height	Weight	Bone age	Cardiac	Genitourinary	Scoliosis	Seizures
COG025	14.3	+	+++	≥98	≥2	≥9	Adv	PDA	–	+	+
COG231	8.7	+	+++	≥99.6	≥98	≥99.6	Adv	–	–	–	+
COG342	2.5	+	+++	≥99.6	≥99.6	≥91	Norm	–	–	+	+
COG183	7.7	N	+++	≥99.6	≥75	≥75	Norm	–	–	–	–
COG070	3.1	+	+++	≥99.6	≥98	N	Adv	–	–	–	+
COG064	13.3	+	+++	≥98	≥50	≥75	Norm	PDA	Phimosis	+	–
COG547	0.4	N	N	≥99.6	≥99.6	≥75	N	ASD/PDA	VUR	N	–
COG218	4.1	+	++	≥98	≥91	≥91	Norm	–	Nephrocalcinosis	–	–
COG347	8.5	+	+++	≥91	≥99.6	≥75	Norm	–	–	–	–
COG344	5.2	+	++	≥99.6	≥99.6	≥91	Del	–	Cystic kidney	–	–
COG172	0.9	+	+++	≥99.6	≥50	≥91	N	VSD	–	+	N
COG548	9.0	N	+++	≥50	≥91	≥91	Adv	–	–	–	–
COG177	9.0	N	+++	≥98	≥9	≥2	Adv	–	–	+	+
COG211	1.1	+	+++	≥99.6	N	N	Norm	–	–	+	+
COG146	1.8	+	+++	≥91	≥91	≥75	Adv	ASD	Vur	–	+
COG160	3.9	N	+++	≥99.6	≥99.6	≥99.6	Norm	–	–	–	–
COG346	2.3	+	+++	≥91	≥75	≥50	Del	ASD/VSD	–	–	–
COG343	3.5	+	++	≥99.6	≥75	≥98	Adv	PDA	–	–	+
COG184	11.0	+	+++	≥91	≥50	≥50	Adv	PDA	Kidney failure	+	–
COG330	1.8	N	+++	≥98	≥75	≥75	Adv	N	N	N	N
COG532	1.5	+	+++	≥91	≥91	≥75	N	ASD/PDA	–	–	–
COG450	2.9	+	+++	≥91	≥50	N	N	Fallots	–	–	–
COG395	9.2	N	+++	≥99.6	≥99.6	≥99.6	Adv	–	Nephrocalcinosis	–	+
COG512	24.0	+	+++	N	≥75	N	Adv	–	–	–	–
COG399	9.1	+	+++	N	≥99.6	≥98	N	–	–	–	+
COG340	3.8	N	+++	≥98	≥0.4	≥25	Del	–	–	+	–
COG044	20.0	+	+++	≥99.6	≥75	≥75	N	–	–	+	+
COG508a	2.0	N	+	≥99.6	≥98	≥99.6	N	–	–	–	+
COG508b	2.0	N	+++	≥99.6	≥99.6	≥98	N	–	–	–	+
COG345	1.3	N	++	≥99.6	≥99.6	≥75	N	Ebstein	–	+	+
COG367	2.4	N	++	≥99.6	≥99.6	≥99.6	Adv	cardiomegaly	–	–	–
COG111	8.0	+	++	≥98	≥75	≥91	Adv	–	PUJ obstruction	+	–

The cases are ordered from largest to smallest deletion size as in fig 2. [a]Age at which individual was last reviewed and growth parameters were measured. ASD, atrial septal defect; PDA, patent ductus arteriosus; VSD, ventricular septal defect; VUR, vesicoureteral reflux; PUJ, pelviureteric junction; OFC, occipitofrontal circumference; LD, learning difficulties; +++, severe LD; ++, moderate LD; +, mild LD; Adv, advanced; Norm, normal; Del, delayed; N, no data. No clinical information is available for case COG179.

flanking *NSD1*, which could mediate these non-recurrent deletions, although it is possible these are present but were not detected by our analyses, as recently demonstrated in non-recurrent deletions in Smith-Magenis syndrome.[14] It is also noteworthy that the region encompassing *NSD1* has a high density of *Alu* repeats; 18.8% compared with an average of 10.6% for the human genome.[15] *Alu* repeats may act as substrates for homologous recombination, and an increased density of *Alu* elements is often observed in regions associated with genomic rearrangements.[13] Detailed mapping of deletion breakpoints will be required to clarify the mechanisms generating recurrent and non-recurrent Sotos microdeletions.

The paternal allele is preferentially deleted in the majority of cases with *NSD1* microdeletions

There was a significant bias towards deletion of the paternally derived allele in our cases with Sotos microdeletions, consistent with previous data from Japanese cases.[16] The combined data demonstrate that the paternally derived chromosome was deleted in 36/41 reported cases (p<0.001). It is likely that this bias is, at least in part, attributable to the greatly increased recombination rate in men compared with women at the 5q telomere. In general, the rate of recombination in men is increased at telomeres compared with centromeres, whereas women show a more even recombination rate along the length of a chromosome.[17] Other telomeric microdeletions, for example of 4p16.3 in Wolf-Hirschhorn syndrome,[18] and of 22q13,[19] also show a paternal bias and markedly increased recombination rates in men compared with women, supporting a role for sex dependent recombination rates in the generation of parental bias in some microdeletion syndromes.

Differences in Sotos microdeletion frequency are not due to case ascertainment bias and may reflect differences in genomic architecture

We identified 15 microdeletions and 123 intragenic mutations in 366 cases from the UK who were all fully screened for mutations and deletions of *NSD1*. These cases were ascertained from across the UK and consisted of a broad range of phenotypes, including individuals with overgrowth but no other features of Sotos syndrome and cases with facial similarity to Sotos syndrome but no overgrowth. There was no obvious bias towards ascertainment of cases with mutations in these analyses. Indeed, as the sensitivity of mutation detection is likely to be less than that of deletions, any bias is likely to be towards identification of cases with microdeletions. Furthermore, all the cases with deletions had been clinically diagnosed with Sotos syndrome prior to the molecular analyses. Our results therefore indicate that ~10% of Sotos cases in the UK are caused by 5q35 microdeletions. Kurotaki *et al* reported 49 microdeletions in 95 cases from Japan.[11] The identification of such a large number of deletions in Japan, despite the smaller number of cases analysed, suggests there is a genuine difference in microdeletion frequency between the two populations. Our data do not support the hypothesis that this is due to under-ascertainment of cases with deletions compared with mutations.[11] An alternative hypothesis is that differences in genomic architecture in Japanese and non-Japanese populations influences the microdeletion frequency. It is interesting in this regard that only 2/11 cases in our series were generated through intrachromosomal rearrangements, compared with 6/8 Japanese cases.[16] Moreover, we identified at least eight distinct deletions, and at least 15/33 deletion cases were not consistent with a recurrent microdeletion, whereas
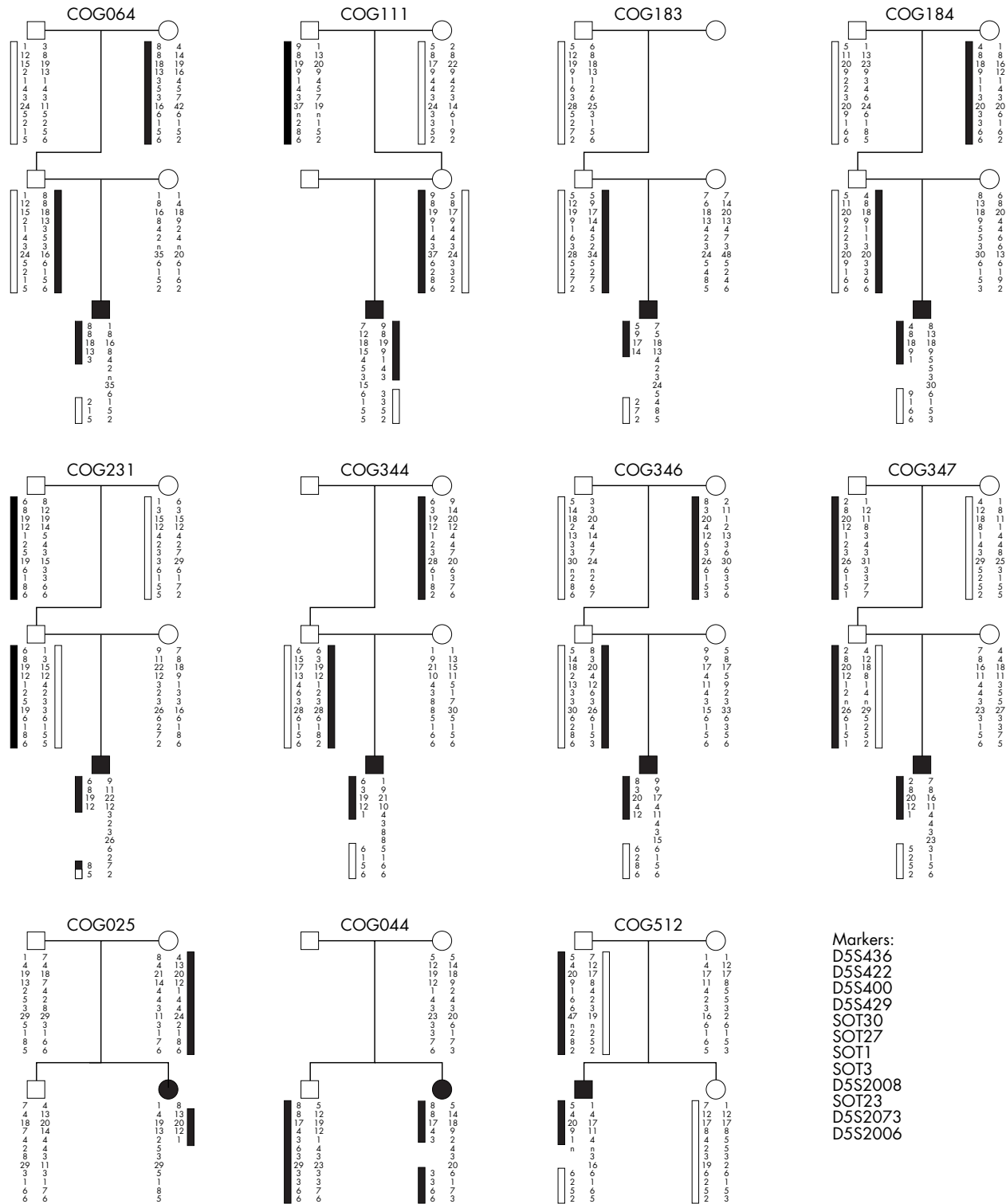
Figure 3 Microsatellite marker analyses in 11 families showing mechanism of generation of 5q35 microdeletions. Haplotypes of marker alleles at 12 microsatellite markers are demonstrated. Filled and open bars next to marker alleles represent the most likely segregation of parental haplotypes above and below deletions. In some cases, other interpretations of the haplotypes are possible but would require additional recombination events. n, no data.

46/50 Japanese microdeletions were reported to be the same size.[11] These data are consistent with possible mechanistic differences in microdeletion generation in Japanese and non-Japanese populations. One possible explanation could be an inversion polymorphism between the inversely orientated LCRs, REPcen and REPmid, being more frequent in Japan, which could predispose to deletions in offspring, as reported in Williams and Angelman syndromes.[20 21] However, further analyses of both Japanese and non-Japanese cases with Sotos and microdeletions and their parents will be required to elucidate the processes responsible for the difference in Sotos microdeletion frequency in these populations.

ELECTRONIC DATABASE INFORMATION

The URLs for data presented are as follows.

- BLAST, http://www.ncbi.nlm.nih.gov/blast
- BLAT, http://genome.ucsc.edu/cgi-bin/hgBlat
- Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/OMIM/ (for Sotos syndrome, Wolf-Hirschhorn syndrome, Williams-Beuren syndrome)
- Primer 3, http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi
- RepeatMasker, http://www.repeatmasker.org
- University of California-Santa Cruz Genome Bio-informatics, http://genome.ucsc.edu/

Authors' affiliations

K Tatton-Brown, J Douglas, K Coleman, A Irrthum, N Rahman, Section of Cancer Genetics, Institute of Cancer Research, Sutton, Surrey , UK
G Baujat, V Cormier-Daire, Department of Medical Genetics, Hopital Necker Enfants Malades, Paris, France
K Chandler, B Kerr, Regional Genetics Service, St Mary's Hospital, Manchester, UK
A Clarke, S Davies, H Hughes, Institute of Medical Genetics, University Hospital of Wales, Cardiff, UK
A Collins, I K Temple, Department of Human Genetics, Southampton University Hospital, Southampton, UK
F Faravelli, Laboratorio di Genetica Umana, Ospedali Galliera de Genova, Genova, Italy
H Firth, Medical Genetics, Addenbrooke's Hospital, Cambridge, UK
C Garrett, Kennedy Galton Centre, Northwick Park Hospital, Harrow, UK
J Liebelt, South Australian Clinical Genetics Service, North Adelaide, Australia
William Reardon, National Centre for Medical Genetics, Our Lady's Hospital for Sick Children, Crumlin, Dublin 12, Ireland
G Bradley Schaefer, D Waggoner, University of Nebraska Medical Center, Omaha, NE, USA
M Splitt, Department of Clinical Genetics, Guy's and St Thomas' Hospital NHS Trust, London, UK
D D Weaver, Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, USA
L Wilson, Great Ormond Street Hospital for Sick Children, London, UK
T Cole, Clinical Genetics Unit, Birmingham Women's Hospital, Birmingham, UK

Competing interests: none declared

The Childhood Overgrowth Collaboration includes the following contributors: M Addor, A Al Swaid, S Andries, H Archer, A Barnicoat, M Barrow, J Barwell, G Baujat, K Becker, J Berg, B Bernhard, M Bhat, M Bitner, E Blair, A Brady, L Brueton, K Chandler, C Christensen, A Clarke, J Clayton-Smith, T Cole, L Colleaux, A Colley, A Collins, V Cormier-Daire, S Danda, S Davies, R Day, De Roy Magali, N Dennis, A Dobbie, F Elmslie, F Faravelli, H Firth, D Fitzpatrick, N Foulds, J Franklin, A Fryer, S Garcia, C Gardiner, C Garrett, B Gener, R Gibbons, Y Gillerot, D Goudie, A Henderson, J Hirst, S Hodgson, S Holder, T Homfrey, H Hughes, B Kerr, A Kumar, D Kumar, W Lam, N Leonard, J Liebelt, P Lunt, S Lynch, A Magee, S Mansour, M McEntagart, C McKeown, S McKee, K Metcalfe, S Mohammad, A Murray, A Nemeth, S Park, M Patton, E Penny, D Pilz, B Plecko, C Pollitt, S Price, O Quarrell, A Raas-Rothschild, N Rahman, W Raith, J Rankin, L Raymond, W Reardon, E Reid, E Rosser, D Ruddy, H Santos, GB Schaeffer, A Schulze, A Shaw, S Smithson, M Splitt, F Stewart, H Stewart, M Suri, E Sweeney, K Tatton-Brown, I K Temple, E Thompson, M Tischowitz, J Tolmie, S Turkmen, P Turnpenny, Van Maldergem, P Vasudevan, I Vaz, D Waggoner, C Verellen, E Wakeling, D Weaver, K White, L Wilson, R Winter, P Zack, A Zankl.

REFERENCES

1 **Cole TR**, Hughes HE. Sotos syndrome: a study of the diagnostic criteria and natural history. *J Med Genet* 1994;**31**:20–32.
2 **Visser R**, Matsumoto N. Genetics of Sotos syndrome. *Curr Opin Pediatr* 2003;**15**:598–606.
3 **Tatton-Brown K**. Rahman N Clinical Features of *NSD1* positive Sotos Syndrome. *Clin Dysmorphol* 2004;**13**:199–204.
4 **Kurotaki N**, Imaizumi K, Harada N, Masuno M, Kondoh T, Nagai T, Ohashi H, Naritomi K, Tsukahara M, Makita Y, Sugimoto T, Sonoda T, Hasegawa T, Chinen Y, Tomita HA, Kinoshita A, Mizuguchi T, Yoshiura KK, Ohta T, Kishino T, Fukushima Y, Niikawa N, Matsumoto N. Haploinsufficiency of NSD1 causes Sotos syndrome. *Nat Genet* 2002;**30**:365–6.
5 **Douglas J**, Hanks S, Temple IK, Davies S, Murray A, Upadhyaya M, Tomkins S, Hughes HE, Cole TR, Rahman N. NSD1 mutations are the major cause of Sotos syndrome and occur in some cases of Weaver syndrome but are rare in other overgrowth phenotypes. *Am J Hum Genet* 2003;**72**:132–43.
6 **Rio M**, Clech L, Amiel J, Faivre L, Lyonnet S, Le Merrer M, Odent S, Lacombe D, Edery P, Brauner R, Raoul O, Gosset P, Prieur M, Vekemans M, Munnich A, Colleaux L, Cormier-Daire V. Spectrum of NSD1 mutations in Sotos and Weaver syndromes. *J Med Genet* 2003;**40**:436–40.
7 **Turkmen S**, Gillessen-Kaesbach G, Meinecke P, Albrecht B, Neumann LM, Hesse V, Palanduz S, Balg S, Majewski F, Fuchs Z, Zschieschang P, Greiwe M, Mennicke K, Kreuz FR, Dehmel HJ, Rodeck B, Kunze J, Tinschert S, Mundlos S, Horn D. Mutations in NSD1 are responsible for Sotos syndrome, but are not a frequent finding in other overgrowth phenotypes. *Eur J Hum Genet* 2003;**11**:858–65.
8 **Raca G**, Waggoner DJ, Kamimura J, Matsumoto N, Schaefer GB. Mutation analysis of the NSD1 gene - genetic testing for Sotos syndrome. *Am J Hum Genet* 2003;**73**(suppl):2427.
9 **Faravelli F**, Cecconi M, Forzano F, Malacarne M, Cavani S, Baldo C, Milani D, Selicorni A, Silengo M, Ferrero GB, Scarano G, DellaMonica M, Memo L, Pantaleoni C, Pallotta R, Renieri A, Concolino D, Dagna Bricarelli F. Mutation analysis of Sotos syndrome. *Am J Hum Genet* 2003;**73**(suppl):179.
10 **Schouten JP**, McElgunn CJ, Waaijer R, Zwijnenburg, Diepvens F, Pals G. Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. *Nucleic Acids Res* 2002;**30**:e57.
11 **Kurotaki N**, Harada N, Shimokawa O, Miyake N, Kawame H, Uetake K, Makita Y, Kondoh T, Ogata T, Hasegawa T, Nagai T, Ozaki T, Touyama M, Shenhav R, Ohashi H, Medne L, Shiihara T, Ohtsu S, Kato Z, Okamoto N, Nishimoto J, Lev D, Miyoshi Y, Ishikiriyama S, Sonoda T, Sakazume S, Fukushima Y, Kurosawa K, Cheng JF, Yoshiura K, Ohta T, Kishino T, Niikawa N, Matsumoto N. Fifty microdeletions among 112 cases of Sotos syndrome: Low copy repeats possibly mediate the common deletion. *Hum Mutat* 2003;**22**:378–87.
12 **Kent WJ**, BLAT- The BLAST-like alignment tool. *Genome Research* 12;654–6..
13 **Shaw CJ**, Lupski JR. Implications of human genome architecture for rearrangement-based disorders: the genomic basis of disease. *Hum Mol Genet* 2004;**13**:R57–64.
14 **Shaw CJ**, Withers MA, Lupski JR. Uncommon deletions of the Smith-Magenis syndrome region can be recurrent when alternate low-copy repeats act as homologous recombination substrates. *Am J Hum Genet* 2004;**75**:75–81.
15 **Lander ES**, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann N, Stojanovic N, Subramanian A, Wyman D, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D, Gregory S, Hubbard T, Humphray S, Hunt A, Jones M, Lloyd C, McMurray A, Matthews L, Mercer S, Milne S, Mullikin JC, Mungall A, Plumb R, Ross M, Shownkeen R, Sims S, Waterston RH, Wilson RK, Hillier LW, McPherson JD, Marra MA, Mardis ER, Fulton LA, Chinwalla AT, Pepin KH, Gish WR, Chissoe SL, Wendl MC, Delehaunty KD, Miner TL, Delehaunty A, Kramer JB, Cook LL, Fulton RS, Johnson DL, Minx PJ, Clifton SW, Hawkins T, Branscomb E, Predki P, Richardson P, Wenning S, Slezak T, Doggett N, Cheng JF, Olsen A, Lucas S, Elkin C, Uberbacher E, Frazier M, Gibbs RA, Muzny DM, Scherer SE, Bouck JB, Sodergren EJ, Worley KC, Rives CM, Gorrell JH, Metzker ML, Naylor SL, Kucherlapati RS, Nelson DL, Weinstock GM, Sakaki Y, Fujiyama A, Hattori M, Yada T, Toyoda A, Itoh T, Kawagoe C, Watanabe H, Totoki Y, Taylor T, Weissenbach J, Heilig R, Saurin W, Artiguenave F, Brottier P, Bruls T, Pelletier E, Robert C, Wincker P, Smith DR, Doucette-Stamm L, Rubenfield M, Weinstock K, Lee HM, Dubois J, Rosenthal A, Platzer M, Nyakatura G, Taudien S, Rump A, Yang H, Yu J, Wang J, Huang G, Gu J, Hood L, Rowen L, Madan A, Qin S, Davis RW, Federspiel NA, Abola AP, Proctor MJ, Myers RM, Schmutz J, Dickson M, Grimwood J, Cox DR, Olson MV, Kaul R, Raymond C, Shimizu N, Kawasaki K, Minoshima S, Evans GA, Athanasiou M, Schultz R, Roe BA, Chen F, Pan H, Ramser J, Lehrach H, Reinhardt R, McCombie WR, de la Bastide M, Dedhia N, Blocker H, Hornischer K, Nordsiek G, Agarwala R, Aravind L, Bailey JA, Bateman A, Batzoglou S, Birney E, Bork P, Brown DG, Burge CB, Cerutti L, Chen HC, Church D, Clamp M, Copley RR, Doerks T, Eddy SR, Eichler EE, Furey TS, Galagan J, Gilbert JG, Harmon C, Hayashizaki Y, Haussler D, Hermjakob H, Hokamp K, Jang W, Johnson LS, Jones TA, Kasif S, Kaspryzk A, Kennedy S, Kent WJ, Kitts P, Koonin EV, Korf I, Kulp D, Lancet D, Lowe TM, McLysaght A, Mikkelsen T, Moran JV, Mulder N, Pollara VJ, Ponting CP, Schuler G, Schultz J, Slater G, Smit AF, Stupka E, Szustakowski J, Thierry-Mieg D, Thierry-Mieg J, Wagner L, Wallis J, Wheeler R, Williams A, Wolf YI, Wolfe KH, Yang SP, Yeh RF, Collins F, Guyer MS, Peterson J, Felsenfeld A, Wetterstrand KA, Patrinos A,

Morgan MJ, de Jong P, Catanese JJ, Osoegawa K, Shizuya H, Choi S, Chen YJ, International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature* 2001;**409**:860–921.

16 **Miyake N**, Kurotaki N, Sugawara H, Shimokawa O, Harada N, Kondoh T, Tsukahara M, Ishikiriyama S, Sonoda T, Miyoshi Y, Sakazume S, Fukushima Y, Ohashi H, Nagai T, Kawame H, Kurosawa K, Touyama M, Shiihara T, Okamoto N, Nishimoto J, Yoshiura K, Ohta T, Kishino T, Niikawa N, Matsumoto N. Preferential paternal origin of microdeletions caused by prezygotic chromosome or chromatid rearrangements in sotos syndrome. *Am J Hum Genet* 2003;**72**:1331–7.

17 **Broman KW**, Murray JC, Sheffield VC, White RL, Weber JL. Comprehensive human genetic maps: individual and sex-specific variation in recombination. *Am J Hum Genet* 1998;**63**:861–9.

18 **Quarrell OW**, Snell RG, Curtis MA, Roberts SH, Harper PS, Shaw DJ. Paternal origin of the chromosomal deletion resulting in Wolf-Hirschhorn syndrome. *J Med Genet* 1991;**28**:256–9.

19 **Luciani JJ**, de Mas P, Depetris D, Mignon-Ravix C, Bottani A, Prieur M, Jonveaux P, Philippe A, Bourrouillou G, de Martinville B, Delobel B, Vallee L, Croquette MF, Mattei MG. Telomeric 22q13 deletions resulting from rings, simple deletions, and translocations: cytogenetic, molecular, and clinical analyses of 32 new observations. *J Med Genet* 2003;**40**:690–6.

20 **Bayes M**, Magano LF, Rivera N, Flores R, Perez Jurado LA. Mutational mechanisms of Williams-Beuren syndrome deletions. *Am J Hum Genet* 2003;**73**:131–51.

21 **Gimelli G**, Pujana MA, Patricelli MG, Russo S, Giardino D, Larizza L, Cheung J, Armengol L, Schinzel A, Estivill X, Zuffardi O. Genomic inversions of human chromosome 15q11–q13 in mothers of Angelman syndrome patients with class II (BP2/3) deletions. *Hum Mol Genet* 2003;**12**:849–58.

Appendix 1   Primer sequences for 12 new 5q35 microsatellite markers

Marker	Primer sequence (5'→3') Forward	Reverse
SOT21	CACTTCATGGAGGATTTCAGC	TCAGAAGGGGAATGATGTGG
SOT23	GCGTCAGTGCTTCACCATAA	CGATCTCAGCTCACTGCAAG
SOT24	CCAGGAGATCGAGGTTTCAG	AGGTCGAGGCAGGTAGATCA
SOT25	TTGGCCAGGCATAGTAGCTC	CATCCCAGGTTCAAGCAATT
SOT26	AAGGCCATCAAAGGTCTGAA	CTCACCCTCCCAAGTAGCTG
SOT27	CACTCGTACGCACCAGAAGA	TATCTTGCCCTCTTTCCCCT
SOT28	TAATGAAGCCCCATCCTCAG	CAGGTTCATGCCATTCTCCT
SOT29	CGTGTGCAAGTGCAGTTTTT	CCCTACTTCCAGTGCAGGAC
SOT30	TGGTGATGCATGCCTGTAAT	AGACATCCACAGGGTTCCTG
SOT31	TGTGCATCCCTTATGCCTTAT	CCACATTCCTTGACTCCACA
SOT32	AATACCATTTTCAGTAGCAACAAA	TCATCTCTTCCCAAATTGTTG

ECHO

Posterior polar cataract is the predominant consequence of a recurrent mutation in the PITX3 gene

P K F Addison, V Berry, A C W Ionides, P J Francis, S S Bhattacharya, A T Moore

Please visit the *Journal of Medical Genetics* website [www. jmedgenet. com] for a link to the full text of this article.

Background: The authors recently identified three large genetically unrelated families with an identical 17 base pair duplication mutation in exon 4 of the PITX3 gene. Here, they report the detailed clinical phenotype.

Methods: Affected and unaffected individuals in the three families with autosomal dominant posterior polar cataract underwent full clinical examination and donated blood samples for DNA extraction and molecular genetic studies.

Results: In all three families, an identical 17 base pair duplication mutation in PITX3 was identified which co-segregated with disease status in the family. All affected individuals had bilateral progressive posterior polar cataracts. In one family, posterior polar cataract was the only clinical abnormality but in the other two families, one of 10 affected individuals and four of 11 affected individuals also had anterior segment mesenchymal dysgenesis (ASMD).

Conclusion: Mutations in the PITX3 gene in humans result in posterior polar cataract and variable ASMD. The gene encodes a transcription factor which has a key role in lens and anterior segment development. The mechanism by which the mutant protein gives rise to such a regional pattern of lens opacity remains to be elucidated.

▲ *British Journal of Ophthalmology* 2005;**89**:138–141.