ELECTRONIC LETTER

Partial *NSD1* deletions cause 5% of Sotos syndrome and are readily identifiable by multiplex ligation dependent probe amplification

This article is available free on JMG online via the **JMG Unlocked** open access trial, funded by the Joint Information Systems Committee. For further information, see http://jmg.bmjjournals.com/cgi/content/full/42/2/97

J Douglas, K Tatton-Brown, K Coleman, S Guerrero, J Berg, T R P Cole, D FitzPatrick, Y Gillerot, H E Hughes, D Pilz, F L Raymond, I K Temple, A Irrthum, J P Schouten, N Rahman

J Med Genet 2005;**42**:e56 (http://www.jmedgenet.com/cgi/content/full/42/9/e56). doi: 10.1136/jmg.2005.031930

Background: Most cases of Sotos syndrome are caused by intragenic *NSD1* mutations or 5q35 microdeletions. It is uncertain whether allelic or genetic heterogeneity underlies the residual cases and it has been proposed that other mechanisms, such as 11p15 defects, might be responsible for Sotos cases without *NSD1* mutations or 5q35 microdeletions.

Objective: To develop a multiplex ligation dependent probe amplification (MLPA) assay to screen *NSD1* for exonic deletions/duplications.

Methods: Analysis was undertaken of 18 classic Sotos syndrome cases in which *NSD1* mutations and 5q35 microdeletions were excluded. Long range polymerase chain reaction (PCR) was used to characterise the mechanism of generation of the partial *NSD1* deletions.

Results: Eight unique partial *NSD1* deletions were identified: exons 1–2 (n = 4), exons 3–5, exons 9–13, exons 19–21, and exon 22. Using long range PCR six of the deletions were confirmed and the precise breakpoints in five cases characterised. This showed that three had arisen through *Alu-Alu* recombination and two from non-homologous end joining.

Conclusions: MLPA is a robust, inexpensive, simple technique that reliably detects both 5q35 microdeletions and partial *NSD1* deletions that together account for ~15% of Sotos syndrome.

Sotos syndrome (MIM 117580) is a condition characterised by overgrowth, distinctive facial features, and learning disability.[1 2] Recently, we and others have shown that small intragenic *NSD1* mutations and 5q35 microdeletions that delete one whole copy of *NSD1* are responsible for the majority of Sotos syndrome cases.[3–6] There has been debate as to whether allelic or genetic heterogeneity is responsible for the remaining cases and it has been proposed that *GPC3* mutations or 11p15 abnormalities, or both, may be responsible for some Sotos cases without *NSD1* abnormalities.[7 8]

Current diagnostic *NSD1* screening strategies typically employ fluorescent in situ hybridisation with a single *NSD1* probe to detect 5q35 microdeletions, together with mutational screening of the 23 *NSD1* exons by heteroduplex analysis or direct sequencing.[3–6 9 10] These techniques detect whole gene deletions and small intra-exonic deletions/insertions or point mutations. However, they are rarely able to detect deletions or duplications that encompass one or more exons. Partial gene deletions/duplications have been identified in several conditions such as Fanconi anaemia group A, hereditary non-polyposis colorectal cancer (HNPCC), and hereditary breast-ovarian cancer syndrome (HBOS),[11–13] but have not been reported in Sotos syndrome.

Several methods for detecting partial gene deletions/duplications have been described, including Southern blotting, quantitative fluorescent polymerase chain reaction (QF-PCR), real time PCR, multiplex amplification and probe hybridisation (MAPH), and multiplex ligation dependent probe amplification (MLPA).[14] MLPA is a high throughput, sensitive technique for detecting copy number variations in genomic sequences.[15] Two adjacently hybridised probes are ligated and then PCR amplified using fluorescent labelled universal primers that correspond to sequence tags present in every probe. Each probe is designed to give a uniquely sized product resulting in a ladder of amplified products that can be quantified by fluorescent electrophoretic analysis.

We have analysed over 400 overgrowth cases, or cases with facial features similar to Sotos syndrome but no overgrowth, for intragenic *NSD1* mutations and 5q35 microdeletions. The majority of negative cases after these analyses were clinically considered to have conditions other than Sotos syndrome because they lacked the characteristic facial features of this condition. However, in 18 cases we considered the facial features to be classic of Sotos syndrome. These cases were clinically very similar to the cases with *NSD1* mutations and we suspected covert *NSD1* abnormalities were responsible. We therefore developed an MLPA probe set that interrogates all *NSD1* exons and analysed the 18 cases for exonic deletions/duplications.

METHODS

Cases

The research was approved by the London Multi-Research Ethics Committee and consent was obtained from all cases or their parents. DNA was extracted using standard methods. Intragenic *NSD1* mutations were identified by methods described in Douglas *et al*,[4] and 5q35 microdeletions were detected by microsatellite analyses, QF-PCR, or FISH, or combinations of these.[16] Cases negative after these analyses were reviewed if clinical photographs were available. Cases with a clinical phenotype of Sotos syndrome were screened by MLPA.

Abbreviations: HBOS, hereditary breast-ovarian cancer syndrome; HNPCC, hereditary non-polyposis colorectal cancer; MLPA, multiplex ligation dependent probe amplification; NAHR, non-allelic homologous recombination; NHEJ, non-homologous end joining

J Med Genet: first published as 10.1136/jmg.2005.031930 on 1 September 2005. Downloaded from http://jmg.bmj.com/ on December 5, 2020 by guest. Protected by copyright.

Table 1 Primers for long range amplification across exon 1 and 2 deletions

Forward primers		Reverse primers	
Distance centromeric of exon 1 (kb)	Primer sequence	Distance telomeric of exon 2 (kb)	Primer sequence
22.5	CCTCGTAGGCTGTGGGTTTA	2	TTTGACGAGAAACAGGTTTACAA
20	TCCCAGTCCTGTCTCCAGTC	4.5	ATCTCTCAAGGGCTGCCAGT
18	TAGGCCATCAGTCCTGGTTC	7	CCCCTTCCCAAACTGTACCT
16.5	GCTGTGGGACAAATCTAGGC	8.5	GGAATTGCTTTAAAAGGCTACA
14.5	CCACAGTCCCTTCACTTGGT	14.5	CAATTGCTAGAAGAGCATTATGG
10.5	CCTACCTTTGCCCCAGCTC	17.5	TCCAAAAAGATACATCAGCAAAGA
8	CAGGCCCTGTGATAGAGAGG	18.5	GAAAAAGGAAGACCAAGTGTGC
6	GAGCTGCTTTTCAGCCTTTG	20	TGAAAAGAGCGAGGGAAGAG
3.5	AGGGGAGGAACCAGTGTCTT		
1.5	CAAAGAGCTGGCCCTTGTAG		

MLPA analyses

MLPA reactions using 150 ng genomic DNA were carried out as previously described.[15] We included known 5q35 microdeletion cases and control samples in all experiments, which were repeated in triplicate. MLPA PCR products were run on an ABI 3100 and analysed using Genescan and Genotyper software (Applied Biosystems, Foster City, California, USA). Chromatograms were initially checked individually by eye and samples were excluded from further analysis if the MLPA reaction control peaks were high. Sample peak areas were exported to an Excel template. For each sample, the peak area for each *NSD1* and *FGFR4* probe was normalised by dividing by the sum of the peak areas for all control probes (excluding the Y chromosome probe). The normalised ratio for each probe was then compared between samples. Deletion of an exon was indicated by a ratio that was approximately half that in controls and approximately equal to that obtained in 5q35 microdeletion cases.

Long range PCR analyses to define deletion breakpoints

We carried out long range PCR using BIO-X-ACT^{TM} long DNA polymerase with 2.5 mM $MgCl_2$ and a touch down 60–50° protocol. For exon 1–2 deletions we used ''Hi-Spec Additive'' according to suppliers instructions for GC-rich templates. We designed primers at approximately 2 kb intervals in the 20 kb flanking regions 5′ to exon 1 and 3′ to exon 2 (table 1). We used various combinations of these primers in cases with exon 1–2 deletions identified on MLPA and ran the PCR products on 1% agarose gels to detect shortened products. We purified the PCR products using QIAquick spin columns (Qiagen, Valencia, California, USA) and these were sequenced using the BigDye 3.1 cycle sequencing kit (Applied Biosystems). We designed additional internal primers to sequence across the breakpoints as necessary (primers and conditions available on request). For deletions of all other exons we used combinations of the forward and reverse primers for *NSD1* mutation screening reported in Douglas *et al*.[4]

RESULTS

NSD1 MLPA probe set

We designed target probes for each *NSD1* exon that avoided polymorphisms identified through our *NSD1* mutation analyses (table 2). To optimise the probe set we analysed normal controls and samples from 5q35 microdeletion cases, in which every *NSD1* exon is deleted. We replaced probes that gave poor or inconsistent results such that the probes in the final mix robustly detected exonic deletions and produced normalised ratios of ~0.5 for each exon in microdeletion cases compared with controls. Only the probe for exon 1 and occasionally exon 2 gave variable results, particularly if poorer quality DNA was used. To offset this variability we added another probe that is 1 kb 5′ to exon 1 and we recommend that exon 1 and 2 deletions are confirmed, where possible, by long range PCR using the primers in table 1. The final probe set includes probes for each of the 23 *NSD1* exons, the probe that is 1 kb 5′ to *NSD1*, two probes for *FGFR4*, which is 35 kb centromeric to *NSD1*, probes for *FLT4* and *TRIM52*, which are telomeric to *NSD1*, and 14 control probes (table 2).

Analyses of Sotos syndrome cases

We used the *NSD1* MLPA probe mix to analyse 18 classic Sotos cases negative for *NSD1* mutations and 5q35 microdeletions, of which 17 were from the UK and one was from Belgium. We identified eight partial *NSD1* deletions involving one or more exons (figs 1 and 2). All eight cases had the classic facial and clinical features of Sotos syndrome (table 3). In six cases we were able to amplify a long range PCR product across the deletion, thus confirming the MLPA results. In the remaining two cases—COG567 and COG151—the substantial size of flanking sequence in which the breakpoints were potentially located precluded long range PCR analyses (fig 2). We did not identify any abnormalities on MLPA in the remaining 10 cases. We suspect these are caused by covert *NSD1* abnormalities because the phenotype of these cases was very similar to that of *NSD1* positive cases.

Four cases showed deletions of exons 1–2 on MLPA and each had a unique deletion with different breakpoints (figs 1 and 2). In COG003 a de novo 23.8 kb deletion encompassing exons 1 and 2 had occurred between highly homologous, directly orientated *Alu* repeats. Perfect sequence alignment of 11 bp was present at the breakpoint junction (fig 3). This suggests that non-allelic homologous recombination (NAHR) between *Alu* elements was the likely mechanism of generation.[17 18] A similar mechanism is implicated in COG068 in whom an 11.4 kb deletion between different directly orientated *Alu* elements had occurred. Sequence overlap of 17 bp was present at the breakpoint junction. Parental samples were not available, but neither parent is affected and the deletion is therefore likely to have arisen de novo. In COG001 we were not able to generate fully readable sequence at the breakpoint junction, but the de novo deletion was ~15 kb and the breakpoints were in directly orientated *Alu* repeats. In the final case with an exon 1–2 deletion, COG567, the two FGFR4 probes in the MLPA mix were deleted, as was the maternal allele at the microsatellite marker SOT3 in intron 2.[4] This indicates that the deletion was maternally derived, at least 67 kb in size, and differs from the other cases with deletions of exons 1–2.

J Med Genet: first published as 10.1136/jmg.2005.031930 on 1 September 2005. Downloaded from http://jmg.bmj.com/ on December 5, 2020 by guest. Protected by copyright.

Table 2 Chromosomal position and sequence information for 42 probes in the NSD1 MLPA kit

Probe	Chr	kb to next probe	Length (nt)	LIG sequence	Clone sequence
FGFR4 probe	5q35	8	373	CTTGGGTCCCTGAGAGCTGTGAGAAGG	AGATGCGGCTGCTGCTGGCCCTGTTGGGGGT
FGFR4 probe	5q35	35	355	CCTGCTCCTCCAGCGATTCTGTCT	TCAGCCACGACCCCCTGCCATTGGGATCCAGCTCC TTCCC
NSD1 1 Kb 5′ of ex1	5q35	1.4	337	GCAGGAAACAAAATGGAGGCAGCA	GCTCGGGTCTGCCTCCCTGAGCCAAGCCTGT
NSD1 exon 1	5q35	0.9	166	GTCGGCAGCAGCCATGTTTTTCGA	GCTGTAGCAGCTGCTGCTACCCTGACTGGGCTTCGCT
NSD1 exon 2	5q35	56.6	319	CTGCGGAGACTACAGGATTTGG CCTCCATG	ATCAATGTAGAGTATTTAAATGGGTCTGCTGATGGAT CAGAAT
NSD1 exon 3	5q35	12.2	301	GTCTACGCCACTGAAGTATGAAGT TGGAGA	TCTCATCTGGGCAAAATTCAAGAGACGCCCATGGTGG
NSD1 exon 4	5q35	6.9	142	GTGGAGGCTTTTGGAGATCCTTCTGA	GAGAGCCTGGGTGGCTGGAAAAGCAATCGTCATG TTTG
NSD1 exon 5	5q35	24.8	283	GTCAGTGACTCTGGAACATCAAAGCCA	TCAAAACCATTACTTTTCTCTTCTGCTTCTAGTCAG AATCACA
NSD1 exon 6	5q35	2.4	265	GGTCAGAGAAGAAACGCCTTAGGAA GCCAA	GCAAGTGGCTTTTGGAATATACAGAAGAATATGATCA GATATT
NSD1 exon 7	5q35	1.5	184	CTTAGGTAAGTTCCCGCTGTGAAGAG GAAA	GCCTTCTAGCCCGAGGTCGATCTAGTGCTCAGAACAA GCAG
NSD1 exon 8	5q35	4.4	238	CGTCAAAGAAAACCAACTAAGAAACTT CTTGA	ATCCAATGATTTAGACCCTGGATTTATGCCCAAGAA GGGGG
NSD1 exon 9	5q35	2.5	292	GGTGTTGTTTTCACTTATTTATAGTGCTATG AAGCT	GGTCACCTGGAGAATGGCATAACTGAATCTTGTGCCA
NSD1 exon 10	5q35	1.5	202	CAAGATATTTGACAAGCCAAGGAAGCG	AAAACGACAGAGGCATGCTGCAGCCAAGATGCA GTGT
NSD1 exon 11	5q35	3.6	247	CGGCCACAAGCCCCAAGGAG	ACTGTTGAGGAAGGTGTAGAACACGATCCCGGGA TGCCTGC
NSD1 exon 12	5q35	5.3	445	GTGAAAAATTGGGTGAGCTGCTGTTATGTG	AGGCTCAGTGCTGTGGGGCTTTCCACCTGGAGTGCCT
NSD1 exon 13	5q35	3.0	148	GTGTGTCCAGAAGTACCCACCCACTG	TTATGCAGAACAAGGGCTTCCGGTGCTCCCTCCA CATCTGT
NSD1 exon 14	5q35	7.6	229	GTCCTGTGGCATACCACGCCAAT	GACTTTTGCCTGGCTGCTGGGTCAAAGATCCTTGC ATCTAA
NSD1 exon 15	5q35	2.1	346	GGAAACTGGTATTGCAATGACTGTAAA GCA	GGCAAAAAGCCACACTACAGGGAGATTGTCTGGGT AAAAG
NSD1 exon 16	5q35	4.0	160	GAGGGTGACGTGAGCAGCAAGGA	TAAGATGGGCAAAGGAGTGGATGGGACATATAAAA AAGGTAACTTTATCC
NSD1 exon 17	5q35	7.0	211	GGCCCAAAAAGAGCTAAGACAGCTGCA	GGAAGACCGAAAGAATGACAAGAAGCCACCACC TTATAAA
NSD1 exon 18	5q35	1.7	400	CGCTGTCAAAACCAGTGCTTTTCCAA	GCGCCAATATCCAGAGGTTGAAATTTTCCGCACATTA CAGC
NSD1 exon 19	5q35	1.3	427	GTCTAGGGTGAGCATATAGAAATTAGTGA TATCATG	TTCTTGAGCATAGCGAATTCGAGCTCTGCATTC TTCTTCATCTATAAGC
NSD1 exon 20	5q35	5.0	418	GACCGAATCATTGATGCTGGTCCCAA	AGGAAACTATGCTCGGTTCATGAATCATTGCTGCCAG CCCA
NSD1 exon 21	5q35	3.2	454	CTAGAATGTCTTGGGAATGGAAAGACTGTT	TGCAAATGTGGAGCCCCGAACTGCAGTGGCTTCTTGG
NSD1 exon 22	5q35	2.7	364	GGAGCGAGAAGATGAGTGTTTTAGTTGTG	GGGATGCTGGCCAGCTCGTCTCCTGCAAGAAACCA
NSD1 exon 23	5q35	3500	178	CCAGAAGGAGCGGGCAGCTTCACCTCA	TCAGGTCACACCACAGGCTGATGAGAAGATGCCA GTGTTG
FLT4 probe	5q35	629	391	CCTTGAACATCACGGAGGAGTCACACG	TCATCGACACCGGTGACAGCCTGTCCATCTCCTG
TRIM52 probe	5q35		193	ACCGATCCCAGAGAGAGGCTGTGAGTCT	GGATCTGGGCGTTCGCACCCTGAGAGAGGGGGTCA AAGCTC
Control probe	Y		118	GTTTATTCTAACCTAGGCAAACGGCATGCT	ATCACAAGAAAGGTTTAAAGCTTTGATAAAATGGG GGAGATTT
Control probe	5q31		130	ACATTGTCACTGCAAATCGACACCTAT	TAATGGGTCTCACCTCCCAACTGCTTCCCCCT
Control probe	4q11		136	GCGTAAGAGCAAAAAGCGAAGGCGCAA	TCTGGACACTGGGAGATTCGGAGCGCAGGGAG TTTGAGAG
Control probe	6p21		154	CTGAAACTCAGCACCGAGGCTCT	GCTCCCCACTCTGAGAGTGATCTACCAGAGCAGG AAGA
Control probe	3p22		172	GGCATCCAGATGGTGTGTGAGACGTT	GACTGAGTGCTGGGACCACGACCCAGAGGCCCGTC
Control probe	10q22		220	GCAGGCCACAGAGATGCTGGT	TCACAATGCCCAGAACCTCATGCAGTCTGTGAAG GAGACT
Control probe	16q24		256	CAGCATATTCAGGAGGCCTTACTACGT	GTCCCACTTCCTCCCCGCCCTGCTCACACCTCGAGTG
Control probe	16q22		274	GCCTCCGTTTCTGGAATCCAAGCA	GAATTGCTCACATTTCCCAACTCCTCTCCTG GCCTCAGAAGAC
Control probe	3p25		310	GAGCAAAGCCACTGAGGTATCTCTACAAA	ACCCACCAGAGTCTGGCACTGATGGTTGCATTTTG TTAATT
Control probe	3q12		328	GAAGACGCTGTCCATTTTCACAGAACTCT	GAAGGAGGCTTGTGACCAGCATGGTCCAGATC TCTACC
Control probe	5q33		382	TTGGTATCCGGATGCCCCTGG	AGAAATGGTGGTCCTCACCTGTGACACCCCTGAA
Control probe	2p14		409	GTTGTTGACCTCCTGTACTGGAGAGACAT	TAAGAAGACTGGAGTGGTGTTTGGTGCCAGCCTATT CCTGC
Control probe	1p36		436	CCTGCACAGATGGCGGCTATC	AGGCCCTTCCGGTTTTCCTGCTCCTTCAGTATCTCCT
Control probe	13q14		463	GGCCCTAGAGTGGGAGTCCTGATAA	CCCAGGCCTGTCTGACTACTTTGCCTTCTTTTGTAGCA TATAGG

Chr, chromosome.

J Med Genet: first published as 10.1136/jmg.2005.031930 on 1 September 2005. Downloaded from http://jmg.bmj.com/ on December 5, 2020 by guest. Protected by copyright.

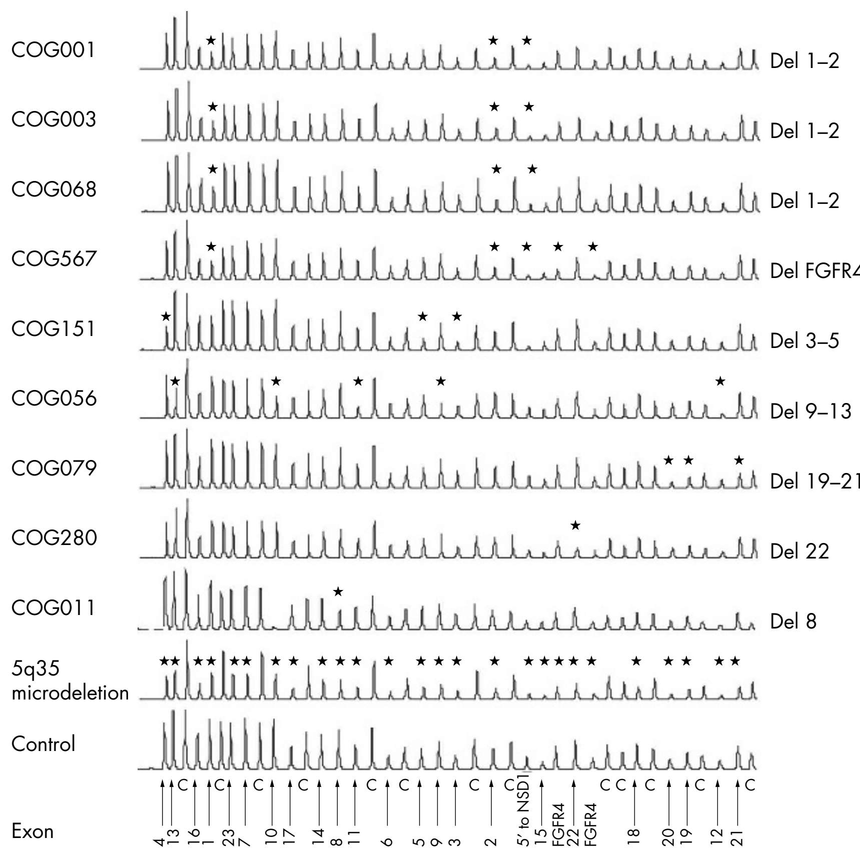

Figure 1 Multiplex ligation dependent probe amplification (MLPA) analysis in Sotos syndrome cases with partial *NSD1* deletions, 5q35 microdeletion, and a control. The asterisks indicate deleted exons. Numbers at the bottom of the figure refer to the *NSD1* exon; C indicates a control probe. The apparent deletion of exon 8 in COG011 is actually caused by an insertion of illegitimate sequence within the ligation sequence of the exon 8 probe.

The remaining partial gene deletions each involved different exons. In COG056 recombination between an *Alu*-Sp repeat in intron 8 and an *Alu*-Sq repeat in intron 13 resulted in deletion of 14.5 kb including exons 9–13. Perfect sequence overlap of 23 bp was present at the junction. In COG079 there was deletion of 7.9 kb including exons 19–21 between intron 18 and intron 22. The proximal breakpoint was not in a recognised repetitive element and the distal breakpoint was in an *Alu*-Y element. There was overlap of only a single thymidine residue at the breakpoint and there was little homology between sequences flanking the breakpoint (fig 3). This indicates that non-homologous end joining (NHEJ) is the likely mechanism of generation.[19] NHEJ is also implicated in the de novo 2.5 kb deletion of exon 22 in COG280. Neither breakpoint was in a known repeat or a recombination associated motif[20] and the sequence overlap at the junction was restricted to three base pairs, ''ACT''. In the final case, COG151, a de novo deletion of exons 3–5 was identified. Owing to the extensive size of intron 2 we did not attempt long range PCR analyses to define the precise breakpoints.

In one case, COG011, the MLPA analysis indicated a de novo deletion of exon 8 (fig 1). However, long range PCR showed that in fact an insertion of ~190 bp of *Alu*-Y sequence had occurred within the ligation sequence of the exon 8 MLPA probe. On reviewing the heteroduplex analysis from the *NSD1* mutation screen in this individual, an aberrant band was visible but had not been noted because the size of the insertion had resulted in the heteroduplex band being very substantially displaced from the wild type allele.

DISCUSSION

Our data show that partial deletions of one or more *NSD1* exons are a novel cause of Sotos syndrome and are readily identifiable by MLPA. Furthermore, MLPA is a robust method for detecting 5q35 microdeletions which occur in

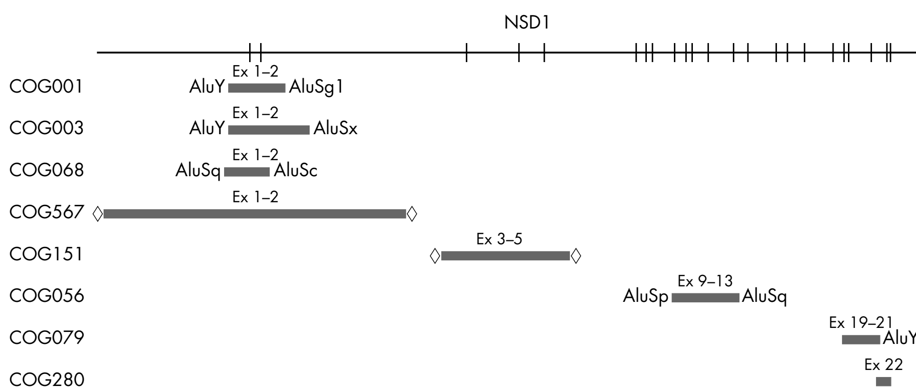

Figure 2 Schematic representation of NSD1, drawn to scale, showing size and position of partial *NSD1* deletions. ''Alu'' indicates that an Alu repeat element was present at the breakpoint with the subfamily indicated; diamonds indicate that the breakpoint was not defined.

J Med Genet: first published as 10.1136/jmg.2005.031930 on 1 September 2005. Downloaded from http://jmg.bmj.com/ on December 5, 2020 by guest. Protected by copyright.

Table 3 Deletion characteristics and clinical features of Sotos cases with partial *NSD1* deletions

ID	Deleted exons	Deletion size (kb)	Mechanism	Facial gestalt	LD	Height centile	OFC centile	Other features
COG001	1–2	~15	*NAHR*	+	+	>99.6	>99.6	Hydrocephalus, precocious puberty
COG003	1–2	23.8	NAHR	+	+	>99.6	99.6	Cryptorchidism, scoliosis
COG068	1–2	11.4	NAHR	+	+	99.6	>99.6	Seizures
COG567	1–2	U	U	+	+	>99.6	>99.6	Craniostenosis
COG151	3–5	U	U	+	+	>99.6	>99.6	Optic nerve hypoplasia, phobias
COG056	9–13	14.5	*NAHR*	+	+	98	>99.6	Scoliosis, seizures, VUR
COG079	19–21	7.9	NHEJ	+	+	>99.6	>99.6	
COG280	22	2.5	NHEJ	+	+	>99.6	>99.6	Scoliosis, seizures

LD, learning disability; NAHR, non-allelic homologous recombination; NHEJ, non-homologous end joining; OFC, occipito-frontal circumference; U, unknown; VUR, vesico-ureteric reflux; +, clinical feature present.

around 10% of non-Japanese Sotos cases and are the commonest cause of Sotos syndrome in Japan.[4 21]

Each of the eight partial gene deletions we identified was unique. *Alu* mediated recombination was the likely cause in at least four cases, similar to other conditions with partial deletions.[17] *NSD1* is enriched for *Alu* elements with a density of 40.2% compared with 10.6% for the genome generally.[22] In particular, intron 2—the largest intron in *NSD1*—contains 115 *Alu* repeats and the sequence 5′ to *NSD1* is also highly enriched with *Alu* elements. This probably explains why the most frequent partial deletion in our series encompassed exons 1–2. Despite the high density of *Alu* repeats in *NSD1*, two of six cases in which we defined the exact deletion breakpoints were not mediated by NAHR between *Alu* elements. In these, NHEJ was the likely mechanism as the breakpoints were not in repetitive elements, there was little homology between the breakpoint flanking sequences, and the region of sequence overlap at the junction was very short.[19]

As none of the parents was affected it is very likely that all the partial deletions occurred as a result of de novo events, as Sotos syndrome is essentially a fully penetrant disorder.[2] We confirmed this in six cases. Together with the observation that all the partial deletions were of unique size, these data suggest that founder or recurrent partial deletions are unlikely to be prevalent in Sotos syndrome and that strategies to identify such deletions will need to analyse the full gene. This differs from other conditions such as HNPCC and HBOS in which founder deletions are well recognised and specifically targeted in screening analyses.[23–25] This also contrasts with Sotos syndrome caused by 5q35 microdeletions in which a recurrent ~2 Mb deletion causes the majority of Japanese Sotos syndrome cases and a lesser proportion of non-Japanese cases.[16 26]

We have not seen false positive results for whole gene deletions or deletions involving more than one exon, and reproducibility of the deletion in a second separate MLPA analysis should be sufficient confirmation for such abnormalities.[27] For reproducible single exon deletions, sequencing of the relevant exon should be undertaken. This is because intra-exonic mutations within the probe sequences can prevent hybridisation or ligation, resulting in the profile of

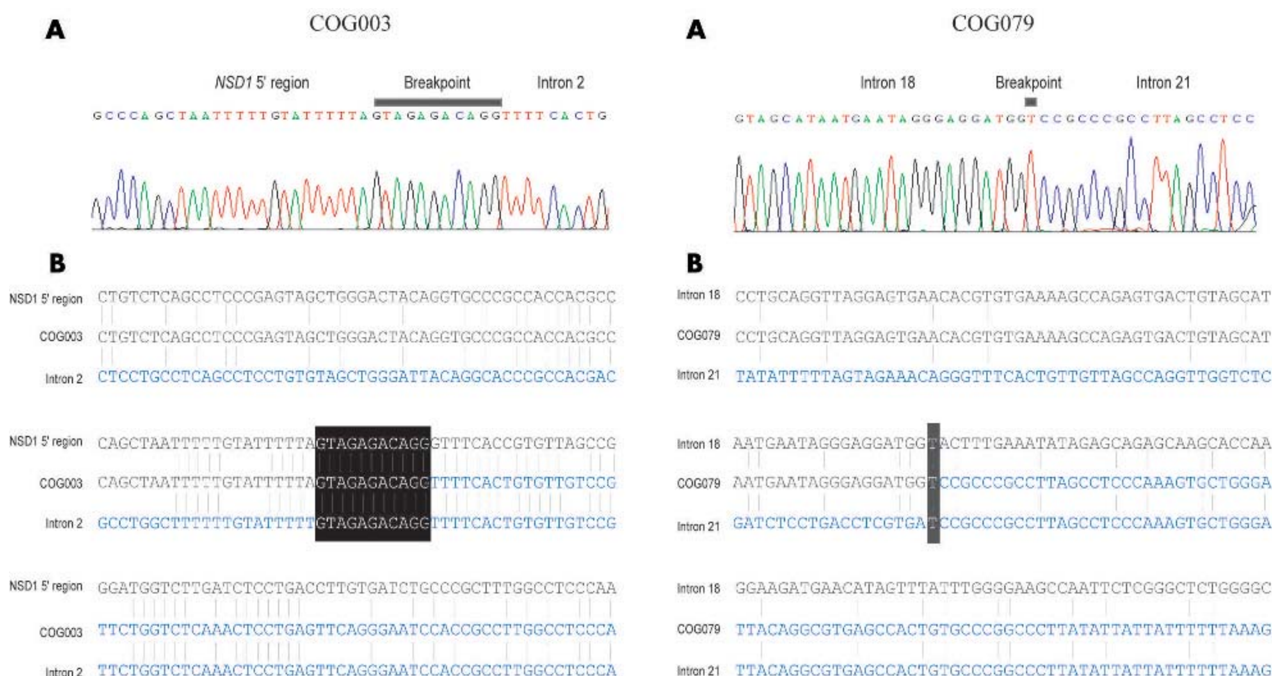

Figure 3 Characterisation of partial gene deletions in COG003 and COG079. (A) Chromatograms at the deletion breakpoints, with regions of sequence overlap indicated by black bars. (B) Sequence flanking the deletion breakpoints, with region of sequence overlap indicated by black boxes around the nucleotides. Identical residues are indicated by vertical bars.

J Med Genet: first published as 10.1136/jmg.2005.031930 on 1 September 2005. Downloaded from http://jmg.bmj.com/ on December 5, 2020 by guest. Protected by copyright.

a single exon deletion, as exemplified by COG011 in this study. If MLPA were used as the primary *NSD1* screening method before mutational screening, small deletions, insertions, or point mutations may be detected in this way. If the exon sequence is wild type we recommend confirmation of a single exon deletion by long range PCR.

To date, we have undertaken all *NSD1* testing in clinically diagnosed cases of Sotos syndrome in the UK; seven of 124 cases were caused by partial *NSD1* deletions and 12 by 5q35 microdeletions. Our data therefore suggest that MLPA (or a similar assay) could usefully be undertaken as the primary screen in individuals with a clinical diagnosis of Sotos syndrome to detect partial and whole gene deletions. MLPA analyses are less expensive than fluorescent in situ hybridisation (FISH) and can be carried out on an aliquot of the DNA obtained for mutation screening, so they do not require a separate sample to be taken. We estimate that primary MLPA analyses would detect around 15% of *NSD1* abnormalities in non-Japanese individuals and a high proportion of *NSD1* abnormalities in individuals from Japan.

ACKNOWLEDGEMENTS
We are grateful to all the families that have participated in these analyses and the many clinicians that have recruited them to the Childhood Overgrowth Study. We thank Lisa Walker, Shan Owens, and Julian Sampson for providing clinical information. JD is supported by the Medical Research Council (UK), KT-B is supported by the Birth Defects Foundation, and AI is supported by Tenovus the Cancer Charity. This research was funded by the Institute of Cancer Research and the Medical Research Council (UK).

ELECTRONIC DATABASE INFORMATION
Online Mendelian Inheritance in Man (OMIM) http://www.ncbi.nlm.nih.gov/Omim/ (for Sotos syndrome)

Repeat Masker http://www.repeatmasker.org (for identification of Alu repeats)

Authors' affiliations
J Douglas, K Tatton-Brown, K Coleman, A Irrthum, N Rahman, Section of Cancer Genetics, Institute of Cancer Research, Sutton, Surrey, UK
S Guerrero, J P Schouten, MRC-Holland, Amsterdam, The Netherlands
J Berg, Clinical Genetics, Human Genetics Unit, Ninewells Hospital and Medical School, Dundee, UK
T R P Cole, Clinical Genetics Unit, Birmingham Women's Hospital, Birmingham, UK
D FitzPatrick, MRC Human Genetics Unit, Edinburgh, Scotland, UK
Y Gillerot, Institut de Pathologie et de Génétique, Loverval, Belgium
H E Hughes, D Pilz, Institute of Medical Genetics, University Hospital of Wales, Cardiff, UK
F L Raymond, Medical Genetics Department, Addenbrooke's Hospital, Cambridge, UK
I K Temple, Department of Human Genetics, Southampton University Hospital, Southampton, UK

Competing interests: none declared

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

Revised version received 18 March 2005
Accepted for publication 20 March 2005

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 **Tatton-Brown K**, Rahman N. Clinical features of NSD1-positive Sotos syndrome. *Clin Dysmorphol* 2004;**13**:199–204.
3 **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 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.
4 **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.
5 **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.
6 **Turkmen S**, Gillessen-Kaesbach G, Meinecke P, Albrecht B, Neumann LM, Hesse V, Palanduz S, Balg S, Majewski F, Fuchs S, 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.
7 **Li M**, Shuman C, Ling Fei Y, Cutiongco E, Bender HA, Stevens C, Wilkins-Haug L, Day-Salvatore D, Yong SL, Geraghty MT, Squire J, Weksberg R. GPC3 mutation analysis in a spectrum of patients with overgrowth expands the phenotype of Simpson-Golabi-Behmel syndrome. *Am J Med Genet* 2001;**102**:161–8.
8 **Baujat G**, Rio M, Rossignol S, Sanlaville D, Lyonnet S, Le Merrer M, Munnich A, Gicquel C, Cormier-Daire V, Colleaux L. Paradoxical NSD1 mutations in Beckwith-Wiedemann syndrome and 11p15 anomalies in Sotos syndrome. *Am J Hum Genet* 2004;**74**:715–20.
9 **Faravelli F**, Cecconi M, Forzano F, Malacarne M, Cavani S, Baldo C, Milani D, Selicorni A, Silengo M, Ferroro 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 **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.
11 **Morgan NV**, Tipping AJ, Joenje H, Mathew CG. High frequency of large intragenic deletions in the Fanconi Anemia Group A gene. *Am J Hum Genet* 1999;**65**:1330–41.
12 **Hogervorst FB**, Nederlof PM, Gille JJ, McElgunn CJ, Grippeling M, Pruntel R, Regnerus R, van Welsem T, van Spaendonk R, Menko FH, Kluijt I, Dommering C, Verhoef S, Schouten JP, van't Veer LJ, Pals G. Large genomic deletions and duplications in the BRCA1 gene identified by a novel quantitative method. *Cancer Res* 2003;**63**:1449–53.
13 **Tournier I**, Bressac-de Paillerets B, Sobol H, Stoppa-Lyonnet D, Lidereau R, Barrois M, Mazoyer S, Coulet F, Hardouin A, Chompret A, Lortholary A, Chappuis P, Bourdon V, Bonadona V, Maugard C, Gilbert B, Nogues C, Frebourg T, Tosi M. Significant contribution of germline BRCA2 rearrangements in male breast cancer families. *Cancer Res* 2004;**64**:8143–7.
14 **Sellner LN**, Taylor GR. MLPA and MAPH: New techniques for detection of gene deletions. *Hum Mutat* 2004;**23**:413–19.
15 **Schouten JP**, McElgunn CJ, Waaijer R, Zwijnenburg D, Diepvens F, Pals G. Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. *Nucleic Acids Res* 2002;**30**:e57.
16 **Tatton-Brown K**, Douglas J, Coleman K, Baujat G, Chandler K, Clarke A, Collins A, Davies S, Faravelli F, Firth H, Garrett C, Hughes HE, Kerr B, Liebelt J, Reardon W, Schaeffer GB, Splitt M, Temple IK, Waggoner D, Weaver DD, Wilson L, Cole T, Cormier-Daire V, Rahman N. Multiple mechanisms are implicated in the generation of 5q35 microdeletions in Sotos syndrome. *J Med Genet* 2005;**42**:307–13.
17 **Deininger PL**, Batzer MA. Alu repeats and human disease. *Mol Genet Metab* 1999;**67**:183–93.
18 **Batzer MA**, Deininger PL. Alu repeats and human genomic diversity. *Nat Rev Genet* 2002;**3**:370–80.
19 **Roth DB**, Wilson JH. Nonhomologous recombination in mammalian cells: Role for short sequence homologies in the joining reaction. *Mol Cell Biol* 1986;**6**:4295–304.
20 **Abeysinghe SS**, Chuzhanova N, Krawczak M, Ball EV, Cooper DN. Translocation and gross deletion breakpoints in human inherited disease and cancer. I: Nucleotide composition and recombination-associated motifs. *Hum Mutat* 2003;**22**:229–44.
21 **Kurotaki N**, Harada N, Shimokawa O, Miyake N, Kawame H, Uetake K, Makita Y, *et al*. Fifty microdeletions among 112 cases of Sotos syndrome: low copy repeats possibly mediate the common deletion. *Hum Mutat* 2003;**22**:378–87.
22 **Lander ES**, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon KK. Initial sequencing and analysis of the human genome. *Nature* 2001;**409**:860–921.
23 **Nystrom-Lahti M**, Kristo P, Nicolaides NC, Chang SY, Aaltonen LA, Moisio AL, Jarvinen HJ, Mecklin JP, Kinzler KW, Vogelstein B, de la Chapelle A, Peltomaki P. Founding mutations and Alu-mediated recombination in hereditary colon cancer. *Nat Med* 1995;**1**:1203–6.
24 **Petrij Bosch A**, Peelen T, van Vliet M, van Eijk R, Olmer R, Drusedau M, Hogervorst FB, Hageman S, Arts PJ, Ligtenberg MJ, Meijers Heijboer H, Klijn JG, Vasen HF, Cornelisse CJ, van't Veer LJ, Bakker E, van Ommen GJ, Devilee P. BRCA1 genomic deletions are major founder mutations in Dutch breast cancer patients. *Nat Genet* 1997;**17**:341–5.
25 **Wagner A**, Barrows A, Wijnen JT, van der Klift H, Franken PF, Verkuijlen P, Nakagawa H, Geugien M, Jaghmohan-Changur S, Breukel C, Meijers-Heijboer H, Morreau H, van Puijenbroek M, Burn J, Coronel S, Kinarski Y, Okimoto R, Watson P, Lynch JF, de la Chapelle A, Lynch HT, Fodde R. Molecular analyses of hereditary nonpolyposis colorectal cancer in the United States: high mutation detection rate among clinically selected families and characterisation of an American founder genomic deletion on the MSH2 gene. *Am J Hum Genet* 2003;**72**:1088–100.

J Med Genet: first published as 10.1136/jmg.2005.031930 on 1 September 2005. Downloaded from http://jmg.bmj.com/ on December 5, 2020 by guest. Protected by copyright.

26 **Taylor F**, Charlton RS, Burn J, Sheridan E, Taylor GR. Genomic deletions in MSH2 or MLH1 are a frequent cause of hereditary non-polyposis colorectal cancer: identification of novel and recurrent deletions by MLPA. *Hum Mutat* 2003;**22**:428–33.

27 **Visser R**, Shimokawa O, Harada N, Kinoshita A, Ohta T, Niikawa N, Matsumoto N. Identification of a 3.0-kb major recombination hotspot in patients with Sotos syndrome who carry a common 1.9-Mb microdeletion. *Am J Hum Genet* 2005;**76**:52–67.

J Med Genet: first published as 10.1136/jmg.2005.031930 on 1 September 2005. Downloaded from http://jmg.bmj.com/ on December 5, 2020 by guest. Protected by copyright.