ORIGINAL ARTICLE

LCR-initiated rearrangements at the *IDS* locus, completed with *A/u*-mediated recombination or non-homologous end joining

Junko Oshima^{1,4}, Jennifer A Lee¹, Amy M Breman¹, Priscilla H Fernandes¹, Dusica Babovic-Vuksanovic², Patricia A Ward¹, Lynne A Wolfe^{3,5}, Christine M Eng¹ and Daniela del Gaudio^{1,6}

Mucopolysaccharidosis type II (MPS II) is caused by mutations in the *IDS* gene, which encodes the lysosomal enzyme iduronate-2-sulfatase. In ~20% of MPS II patients the disorder is caused by gross *IDS* structural rearrangements. We identified two male cases harboring complex rearrangements involving the *IDS* gene and the nearby pseudogene, *IDSP1*, which has been annotated as a low-copy repeat (LCR). In both cases the rearrangement included a partial deletion of *IDS* and an inverted insertion of the neighboring region. *In silico* analyses revealed the presence of repetitive elements as well as LCRs at the junctions of rearrangements. Our models illustrate two alternative consequences of rearrangements initiated by non-allelic homologous recombination of LCRs: resolution by a second recombination event (that is, *A/u*-mediated recombination), or resolution by non-homologous end joining repair. These complex rearrangements have the potential to be recurrent and may be present among those MPS II cases with previously uncharacterized aberrations involving *IDS*.

Journal of Human Genetics (2011) 56, 516–523; doi:10.1038/jhg.2011.51; published online 19 May 2011

Keywords: array CGH; chromosomal rearrangements; clinical molecular genetics; Hunter syndrome; lysosomal storage diseases; mucopolysaccharidoses; X-linked diseases

INTRODUCTION

Mucopolysaccharidosis type II (MPS II; Hunter syndrome; OMIM # 309900) is a rare X-linked recessive lysosomal storage disorder. It is caused by deficiency of the enzyme iduronate-2-sulfatase (IDS; EC 3.1.6.13), one of the enzymes in the pathway of heparan sulfate and dermatan sulfate degradation. The clinical severity of MPS II occurs along a spectrum with the severe form being characterized by progressive cognitive decline, developmental regression and involvement of multiple organ systems occurring during the first few years of life. Progressive respiratory and cardiac failure may lead to death in the first or second decade. In patients with the attenuated form, the central nervous system is not (or is minimally) compromised, although skeletal deformities and other organ involvement can be present to the same degree as in the severe form. Cardiac or respiratory failure may lead to death during the third to fourth decades of life. Treatment for MPS II is currently focused on enzyme replacement therapy and on the management of the clinical symptoms.¹

MPS II is caused by mutations in the *IDS* gene (RefSeq NM_000202.4), which consists of nine exons and spans ~24 kb on

chromosome Xq28.2.² Its 1650-bp open reading frame encodes the 550 amino-acid lysosomal enzyme iduronate-2-sulfatase.³ A 1.4-kb alternative transcript that lacks exons 8 and 9 but contains an alternative terminal exon 7 has been reported, although the significance of this alternative form is not well known.⁴ At 20 kb telomeric to the *IDS* locus, there is a predicted locus, *LOC727913* (RefSeq XM_001126136), that is highly homologous to part of exon 2, intron 2 and exon 3 of the *IDS* gene (Figure 1c).^{5–7} The *IDSP1* locus, which is the only pseudogene of *IDS* in the human genome, is located ~25 kb telomeric to *IDS* and spans 1.6 kb in an inverse orientation. *IDSP1* is highly homologous to a 1.6-kb region within *IDS* intron 7 (Figure 1c).^{8,9}

To date, more than 350 *IDS* disease-causing mutations have been described (Human Genome Mutation Database),¹⁰ with the majority represented by missense, nonsense, splicing and small deletion mutations. Genomic rearrangements account for ~20% of *IDS* cases.¹⁰ There have been a number of previous reports documenting various rearrangements involving *IDS* and *IDSP1*.^{8,9} Among them is a common inversion involving low-copy repeats (LCRs) in *IDS* intron 7 and *IDSP1* that has been observed in multiple independent

¹Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, USA; ²Department of Medical Genetics, Mayo Clinic, Rochester, MN, USA and

³Department of Genetics, Yale University School of Medicine, New Haven, CT, USA

⁴Present address: Department of Pathology, University of Washington, Seattle, WA, USA.

⁵Present address: National Human Genome Research Institute, Bethesda, MD, USA.

⁶Present address: Department of Human Genetics, University of Chicago, IL, USA.

Correspondence: Dr D del Gaudio, Department of Human Genetics, University of Chicago, 5841 S. Maryland Avenue, Room L155, MC 0077, Chicago, IL 60637, USA.

E-mail: ddelgaudio@bsd.uchicago.edu

Received 13 January 2011; revised 11 February 2011; accepted 16 March 2011; published online 19 May 2011

Figure 1 RT-PCR analysis for case 1. (a) Ethidium bromide staining of the RT-PCR products spanning *IDS* exons 6 through 9 (left), nested PCR reaction spanning exons 7 through 9 (middle) and the repeated nested PCR reaction (right). M: 100 bp marker, P: patient, C: control. B: water blank. (b) Nucleotide sequences and the alignments of the junction of the *IDS* exons and cryptic inserts from RT-PCR products showed in a. The 5' (left) and 3' (right) junctions of the longer insert are shown. Nucleotide positions follow the numbering in RefSeq NC_000023.9. (c) Diagram of the region surrounding the *IDS* locus. Loci in this chromosome Xq (ChXq) region are shown; centromere, cen; telomere, tel. The relevant repeat and repetitive sequences are shown: LCR, low-copy repeat; L1-Alu-L1, a combination of LINE and Alu repeats: L1MD2-AluSg-L1MD2 (left) and L1MB3-AluSc-L1MB3 (right). (d) Diagram of the six cryptic exons inserted between *IDS* exons 7 and 9. Genomic coordinates of the exons and the locations of the loci annotated in the Entrez Gene Database are shown.

pedigrees.¹¹ Also reported are large deletions of the *IDS* locus ranging from single or multiple exon deletions to deletions of the entire gene, with or without the involvement of neighboring genes. Complex rearrangements reported previously include recombination between the *IDS* gene and the *IDSP1* pseudogene,^{8,9} a deletion with an inverted duplication,¹² and an *Alu*-mediated deletion with an insertion.¹³ In particular, two highly homologous regions within *IDS* and *IDSP1* appear to facilitate intrachromosomal recombination events leading to complex rearrangements.^{5,8,14} These two regions span ~1.6 kb and are designated as LCRs (UCSC Genome Browser, Human March 2006 Assembly (hg18); <http://genome.ucsc.edu/>). LCRs are thought to mediate non-allelic homologous recombination (NAHR) events resulting in recurrent rearrangements of a various number of loci.^{15,16} Non-homologous end joining (NHEJ) is another well-known rearrangement repair mechanism that is thought to be facilitated by the presence of repetitive DNA sequences, such as *Alu* or LINE repeats, although the precise causal mechanisms remain elusive.^{15–17}

In this report we present two novel rearrangements involving the *IDS* gene and its pseudogene, *IDSP1*, in two unrelated patients with MPS II. One consists of a deletion including *IDS* exon 8 and an inverted duplication of an ~49-kb region distal to *IDS* that is inserted into the deleted region. The other consists of an ~36-kb deletion encompassing *IDS* exons 1 through 7 and an ~8-kb deletion, interrupted by an inverted 2-kb region, and a point mutation that destroys the usual translation termination signal. *In silico* analysis of the breakpoints, aimed at identifying consensus sequences of motifs involved in complex rearrangements, suggests that the close proximity of the two LCRs, combined with the nearby *Alu* and other repetitive sequences, confers significant instability to this region, potentially contributing to these novel genomic rearrangements.

MATERIALS AND METHODS

Patient samples

Peripheral blood samples from the patients and one mother were submitted for *IDS* clinical testing to the DNA Diagnostic Laboratory at Baylor College of

Medicine. The families subsequently gave their consent for further analysis of the pertinent genomic region on a research basis. This study was approved by the Baylor College of Medicine Institutional Review Board.

Mutation analysis

Genomic DNA was isolated from blood leukocytes on the AutoPure LS robotic workstation (Gentra Systems, Minneapolis, MN, USA) using a Pure-gene DNA isolation kit (Gentra Systems). Total RNA was isolated from whole blood using Trizol LS (Invitrogen, Carlsbad, CA, USA) as previously described.¹⁸ PCR primers were designed for amplification of all coding exons and flanking intronic sequences of *IDS* (RefSeq NM_000202.2) and are available upon request. PCR products were sequenced using the Big Dye terminator version 3.1 chemistry on an ABI 3730 XL Genetic Analyzer (Applied Biosystems, Foster City, CA, USA).

Reverse transcriptase (RT)-PCR analysis

Three hundred nanograms of total RNA were reverse-transcribed in a 20 µl reaction volume with Superscript III Reverse Transcriptase (Invitrogen) according to the manufacturer's instructions. Primer sets and, when necessary, nested primer sets were designed based on the GenBank reference sequence NM_000202.2 at the exon–exon junctions. PCR was performed with FastTaq (Applied Biosystems). RT-PCR products were sequenced as described above.

Multiplex ligation-dependent probe amplification analysis

Dosage analysis of the *IDS* gene was performed by multiplex ligation-dependent probe amplification using commercially available reagents (SALSA P164 probe mix) and instructions from MRC-Holland (Amsterdam, The Netherlands). This probe mixture contains 11 probes specific to *IDS* and two probes specific to the *IDSPI* pseudogene, along with other control fragments. Ligation products were amplified and resolved on an ABI 3730 XL Genetic Analyzer and the data were analyzed using Genescan software (Applied Biosystems). Patient data were normalized to gender-matched controls for copy-number differences using Gene Marker Software (Softgenetics, State College, PA, USA).

Affymetrix genome-wide SNP array 6.0 comparative genomic hybridization analysis

Array-comparative genomic hybridization (array-CGH) using the Affymetrix Genome-Wide SNP Array 6.0 was performed according to the Genome-Wide Human SNP Nsp/Sty Assay Kit 5.0/6.0 protocol provided by the manufacturer (Affymetrix, Santa Clara, CA, USA).¹⁹ The arrays were scanned using

a GeneChip Scanner 3000 7G (Affymetrix) and results were analyzed using Genotyping Console version 2.1 software (Affymetrix).

Determination of rearrangement breakpoints

The rearrangement breakpoints were ascertained by PCR primer walking using FastTaq (Applied Biosystems) and the Expand Long Template PCR System (Roche Diagnostics, Indianapolis, IN, USA). PCR primers were designed based on the GenBank reference sequence NC_000023.9, and are listed in Table 1. PCR products were sequenced as described above. All chromosome positions were determined using the Human March 2006 Assembly (hg18).

RESULTS

Clinical presentations of MPS II patients

The patient (case 1) is a 3-year-old European Caucasian male with clinical features consistent with a diagnosis of MPS II. He was born after a normal term pregnancy by Cesarean section to a 25-year-old mother. His birth weight was 3450 g. His Apgar scores were high, but on the second day of life he aspirated and developed respiratory distress requiring ventilator support for 2 weeks. In addition, he had mild jaundice, which persisted for the first month of life. The patient's development was globally delayed and he displayed progressive facial coarsening, hepatomegaly, joint stiffness, fine papular skin rash and an umbilical hernia. An echocardiogram showed mild thickening of the mitral valve. He developed hearing loss and was given hearing aids. Magnetic resonance imaging of the head showed generalized dilatation of perivascular spaces, confluent bi-hemispheric white matter, T-signal hyperintensity and mild narrowing of foramen magnum due to deformity of the dens and thickening of the ligamentous structures along the posterior aspect of the dens, but without impingement on the cervicomedullary junction. The skeletal survey was consistent with mild dysostosis multiplex. Urinary mucopolysaccharides elevated with excretion of dermatan and heparan sulfate, along with an analysis of IDS activity in peripheral blood leukocytes, confirmed the diagnosis of MPS II. The patient was started on enzyme replacement therapy at age 2.5 years. At age 3.5 years he presented with sleep abnormalities and was diagnosed with confusional arousals, without evidence of obstructive sleep apnea or electroencephalogram abnormalities suggesting seizures.

The second patient (case 2) is a 6-year-old male of Pakistani origin with clinical features consistent with a diagnosis of MPS II.

Table 1 PCR primers for genomic PCR amplification for junction analysis

Primer name	Oligonucleotide sequence	Primer location/gene ^a
<i>Case 1, 5' junction</i>		
F1	5'-TCCTGGGACATCCTCCATCAG-3'	chrX:148 378 477–148 378 497/ <i>IDS</i>
R1	5'-CCTGAGAGCTGCTTTGATCTTGTG-3'	chrX:148 422 490–148 422 513/ <i>IDSPI</i>
<i>Case 1, 3' junction</i>		
F2	5'-TGTCTACAGGATGAAGTCATCGAG-3'	chrX:148 460 267–148 460 290/ <i>FLJ16423</i>
R2	5'-AACCTGCAGTCCTGCAAGTCC-3'	chrX:148 372 581–148 372 601/ <i>IDS</i>
<i>Case 2, 5' junction</i>		
F3	5'-GAGGAGGTGCTCGCATGTGTCCAG-3'	chrX:148 425 528–148 425 505/prox. to <i>CXorf40A</i>
R3	5'-CCACCGACATGGTCACATAGC-3'	chrX:148 415 327–148 415 307/ <i>IDSPI</i>
<i>Case 2, 3' junction</i>		
F4	5'-ACCTCTGCTTCTGCTCCCTAA-3'	chrX:148 415 144–148 415 164/ <i>IDSPI</i>
R4	5'-CCTTACTCTAAGCTCTGGAGC-3'	chrX:148 376 024–148 376 044/ <i>IDS</i>

^aGenBank Accession number NC_000023.9.

This diagnosis was confirmed through biochemical enzyme analysis. Detailed clinical information could not be obtained for this patient.

Case 1—a case of a deletion and an inverted duplication

Standard mutation analysis performed on a clinical basis failed to amplify *IDS* exon 8 with two different sets of PCR primers, which suggested the presence of a hemizygous deletion spanning exon 8. Messenger RNA studies by RT-PCR analysis of *IDS* exons 6 through 9 (c.852 through c.1303) yielded the expected 452-bp product in two control samples, but no significant products were seen in the patient sample (Figure 1a, left). However, nested RT-PCR with internal primers to amplify exons 7 through 9 (c.947 through c.1254) revealed a larger, 1.8-kb product in the patient sample and the expected 308-bp product in two controls (Figure 1a, middle). In nested RT-PCR experiments, two RT-PCR products, one of which was 1.8 kb as previously observed, and the other 1.1 kb, were observed in separate reactions for the patient sample (Figure 1a, right). Sequence analysis of the longer RT-PCR product revealed that it contained an insertion of 1639 bp between *IDS* exons 7 and 9 (Figure 1b). This 1639-bp insertion aligned with six regions interspersed within a 10-kb genomic

region that maps ~50 kb telomeric to the *IDS* gene (Figures 1c and d). These six regions are each flanked by AG nucleotides on the 5' ends and GT nucleotides on the 3' ends, suggesting that they could represent cryptic exons. The first four cryptic exons aligned with exons 2, 3, 4 and a portion of exon 5 of the putative gene, *CXorf40A* (Figure 1d). The fifth cryptic exon aligned with a part of *LOC100133153* in a reverse orientation. The sixth cryptic exon aligned with a portion of another putative gene, *FLJ16423*. The shorter RT-PCR product contained an 898-bp insertion between *IDS* exons 7 and 9. This 898-bp insertion aligned with the same first five cryptic exons derived from *CXorf40A* and *LOC100133153* that were observed within the longer RT-PCR product; however, it was missing the sixth cryptic exon from *FLJ16423* (Figure 1d). Thus, the 1639-bp and 898-bp insert are likely to represent alternative splice isoforms. These data indicate that a genomic fragment of at least 10 kb (chrX:148 431 470–148 441 869) was inserted between *IDS* exons 7 and 9 in a reverse orientation.

MLPA analysis confirmed the presence of the deletion of *IDS* exon 8 in the proband (Figure 2a) and his carrier mother (Figure 2b). In addition, one of the two *IDSP1*-specific probes showed a probe

Figure 2 Multiplex ligation-dependent probe amplification (MLPA) and array-CGH analyses for case 1. (a, b) MLPA data are shown for the patient (a) and his carrier mother (b). The deletion of *IDS* exon 8 was detected by two probes (red square) in both the patient (peak ratio of 0) and mother (peak ratio of 0.5). The partial duplication of *IDSP1* detected (blue square) in the proband (peak ratio of 2) and his mother (peak ratio of 1.5) is also shown. (c) A scatter plot (top panel) and copy-number plot (middle panel) of the Affymetrix Genome-Wide SNP array 6.0 results. The horizontal line represents the physical position of the oligonucleotide probes. Blue square, probes within a duplicated segment, detected as increase in the $\log_2$ (test/control) ratio; red square, deleted region with negative $\log_2$ (test/control) ratio; green squares, probes within these regions have normal copy number. The thick blue bar indicates the duplicated region (bottom panel). (d) The diagram shows the annotated loci and sizes of segments with respect to the array-CGH data. The minimum duplication size is ~45 kb.

height ratio of ~ 2 in the patient and of ~ 1.5 in his mother relative to the corresponding peak in the control samples. This suggests that the insertion between *IDS* exons 7 and 9 may include a partially duplicated portion of the *IDSP1* pseudogene, and that the lower probe height ratio for the mother is due to her being heterozygous for this mutation. In the attempt to delineate this rearrangement further, we conducted array-CGH analysis using the Affymetrix Genome-Wide SNP Array 6.0 (Figure 2c). This analysis revealed a loss of two probes (red square in Figure 2c), indicating the presence of a deletion involving *IDS* exon 8 and surrounding intronic regions, and a gain of an ~ 45 -kb (minimum) to an ~ 75 -kb (maximum) region including the *IDSP1*, *CXorf40A*, *LOC100133153* and *FLJ16423* loci (blue square in Figure 2c).

We identified the breakpoints of this deletion/duplication by long-range PCR using various combinations of primers designed from the sequence surrounding this region, assuming that the duplicated fragment was inserted into the deleted region in a reverse orientation (Figure 3a). Long-range PCR using a primer mapping to *IDS* intron 7 (F1, Table 1, Figure 3a), and one between *IDSP1* (that is, LCR) and *CXorf40A* (R1, Table 1, Figure 3a), yielded a product only in the patient sample, but not in the control DNA (Figure 3b), indicating that the distal or 5' breakpoint junction (in *IDS* intron 7) is within

the LCRs. Based on the sequence similarity between the two LCRs, the recombination appears to have been mediated by identical 328-bp regions within the LCRs, chrX:148 377 890–148 378 207 in *IDS* and chrX:148 416 096–148 416 423 in *IDSP1*. Due to the absence of array-CGH probe coverage at the proximal or 3' end of the rearrangement (in *IDS* intron 8), a series of PCR primers within the 30-kb genomic region between the *FLJ16423* and *HSEF2* loci was tested in combination with primers mapping within *IDS* exon 9. One of the combinations, (F2 and R2, Table 1, Figure 3a), yielded a PCR product unique to the patient sample (Figure 3b). Junction sequence analysis revealed 9 bp of homology (5'-TCAAGACCA-3') at chrX:148 465 246–148 465 254 and chrX:148 373 971–148 373 979, both located within *Alu* elements (Figure 3c). An *Alu* repeat belonging to the S subfamily, *AluSc*, in the duplicated region (chrX:148 373 761–148 374 055) is flanked by L1MB3 LINE elements (chrX:148 373 068–148 373 760 and chrX:148 374 056–148 374 127) on both sides (Figures 1c and 3a). The length of the L1MB3-*AluSc*-L1MB3 is 1060 bp. Another *Alu* repeat of the S subfamily, *AluSg*, in *IDS* intron 9 (chrX:148 465 170–148 465 482) is also flanked by the LINE element L1MD2 (chrX:148 464 505–148 465 169 and chrX:148 465 510–148 465 761) (Figures 1d and 3a). The length of the L1MD2-*AluSg*-L1MD2 is 1257 bp. The overall similarity between the two L1-*Alu*-L1 regions, or the combination of an L1 element, an *Alu* repeat, and another L1 element, is 84%. The similarity between the two *Alu* repeats only is 98%. The total lengths of the duplicated and deleted segments are 49 159 and 6090 bp, respectively.

Case 2—a case of a complex rearrangement and a point mutation

Standard mutation analysis failed to amplify *IDS* exons 1 through 7 with two different sets of PCR primers, which suggested the presence of a hemizygous deletion spanning exons 1 through 7. In addition, sequence analysis of exons 8 and 9 identified a hemizygous c.1652G>C nucleotide change in *IDS* exon 9 (Figure 4a). The c.1652G>C substitution results in the replacement of the normal translation stop codon UGA, with a UCA serine codon (p.X551S). This single base-pair substitution is predicted to result in the addition of 17 amino acids to the *IDS* protein.

MLPA analysis confirmed the deletion of exons 1 through 7 (Figure 4c), and also revealed a deletion of part of the *IDSP1* region, suggesting the presence of a deletion partially involving the *IDSP1* pseudogene. In addition, the MLPA probe that hybridizes to both *IDS* exon 3 and a homologous region within the locus, *LOC727913* (Figure 4b), showed a probe height ratio of ~ 0.5 (pink box in Figure 4c). As *IDS* exon 3 appears to be deleted, this suggests that the deletion is interrupted by the region that includes a portion of *LOC727913*. Array-CGH analysis revealed an ~ 36 -kb deletion encompassing the *IDS* locus and an ~ 8 -kb deletion distal to the *IDS* locus (Figures 4d and e). Based on the genomic coordinates of the probes, this second deletion potentially could range from 7 to 14 kb in size. The ~ 36 -kb deletion encompasses *IDS* exons 1 through 7 and the ~ 8 -kb deletion includes a portion of *IDSP1*. These two deletions are interrupted by an ~ 2 -kb region that includes portions of *LOC727913* and *IDSP1* loci (Figure 4d).

As recombination between the LCRs within *IDS* intron 7 and *IDSP1* appears to be a recurrent and possibly common event, we first tested whether LCR-mediated recombination might be involved in this case. Long-range PCR with a forward primer mapping 5' or proximal to *IDSP1*, (F4, Table 1, Figures 5a and b), and a reverse primer mapping within *IDS* intron 8 (R4, Table 1, Figures 5a and b), yielded a distinct junction product in the patient but not in the control (Figure 5b). Sequence analysis of this proximal breakpoint junction revealed

Figure 3 Genomic analysis of the rearrangement for case 1. (a) Possible mechanism for the generation of the complex rearrangement mediated by homologous sequences in the *IDS* gene and the *IDSP1* pseudogene. The inserted ~ 49 -kb fragment between exons 7 and 9 is indicated by the thick line. The ~ 6 -kb deleted region is indicated by the dashed line. The putative recombination sites are shown by crosses. Regions amplified by long-range PCR primers F1 and R1, and F2 and R2, are shown. (b) Long-range PCR amplification across the breakpoint. (c) Sequence alignment for the 3' or proximal breakpoint junction. The red square indicates the 9 bp of homology within the *Alu* sequences. A full color version of this figure is available at the *Journal of Human Genetics* journal online.

Figure 4 Sequencing, multiplex ligation-dependent probe amplification (MLPA) and array-CGH analyses for case 2. **(a)** Sequence analysis of *IDS* exon 9. Nucleotide and amino acid sequences within the vicinity of the c.1652G>C (p.X551S) mutation and the novel stop codon created downstream are shown; ▼ indicates the site of the normal stop codon. **(b)** Diagram of the region surrounding the *IDS* locus. Locations of the *IDS* gene, two loci (*LOC727913* and *IDSP1*) highly homologous to *IDS*, and *CXorf40A* are shown. LCR and repetitive sequences, *AluSp* and a CT-rich region, are also shown. **(c)** MLPA result for *IDS* and *IDSP1* genes. Deletion of *IDS* exons 1 through 7 as well as *IDSP1* probes are marked in the red square as in Figure 2. The probe that hybridizes to both *IDS* exon 3 and the highly homologous *LOC727913* with a peak ratio of 0.5 is shown in the purple square. **(d)** Array-CGH results for copy-number determination within the *IDS* region. A scatter plot (top panel), copy-number plot (middle panel) and deleted regions (bottom panel, thick red bars) are shown as in Figure 2. Deleted regions and the regions of normal copy number are marked in red and green squares, respectively. Diagram below shows the relative locations of annotated loci with respect to the array-CGH data.

that the same 328-bp homologous region within the LCRs in *IDS* intron 7 and *IDSP1* may have mediated this rearrangement as well. No deletion or nucleotide insertion was found at the sequence around the breakpoint junction.

Based on array-CGH analysis and the evidence for recombination between two LCRs, we hypothesize a model in which the two deletions are interrupted by an inverted insertion of a region with identity to *IDSP1* (Figure 5a). Long-range PCR with primers

mapping proximal to *CXorf40A* (F3, Table 1, Figure 5a), and proximal to *IDSP1*, (R3, Table 1, Figure 5a), yielded a unique product in the patient, but no product in the control (Figure 5b). Sequence analysis revealed that the distal breakpoint joins a CT-rich low-complexity repeat (chrX:148 424 336–148 424 660) at nucleotide position chrX:148 424 658, to an *AluSp* element (chrX:148 414 067–148 414 377) at nucleotide position chrX:148 414 166 (Figure 5c). There is no sequence homology at this junction, suggesting repair

Figure 5 Genomic analysis of the rearrangement for case 2. (a) A model for generation of the complex rearrangement mediated by homologous sequences in the *IDS* gene and the *IDSP1* pseudogene. Deleted regions are shown as dotted lines. The inserted ~2-kb fragment including *IDSP1* is indicated by the thick line. Locations of *IDS* exons, LCRs and repetitive sequences are shown. Crosses indicate putative recombination sites. Regions amplified by long-range PCR primers F3 and R3, and F4 and R4, are shown. (b) Long-range PCR amplification across the breakpoint. (c) Sequence alignment for the 5' or distal breakpoint junction. Red squares indicate homology between the junction sequence and the reference sequences. A full color version of this figure is available at the *Journal of Human Genetics* journal online.

by NHEJ. The total sizes of the deletions are 36 275 and 8234 bp, respectively; the size of the insertion is 2248 bp.

DISCUSSION

In this study, we carried out detailed molecular analyses for two novel complex rearrangements involving the *IDS* locus in patients with MPS II. The first case involves an ~49-kb duplication of a region telomeric to *IDS* that was inserted into a deleted region of *IDS*. The second case involves an interrupted deletion including *IDS* exons 1 through 7 and part of the *IDSP1* region, and a point mutation that destroys the usual *IDS* termination codon. Junction analysis was facilitated by the use of array-CGH to assess the approximate locations of the breakpoints.

Collective junction analyses for the first case support our model that the insertion of a duplicated region may be mediated by a strand exchange involving the homologous LCR and L1-*Alu*-L1 regions (Figure 3a). Interestingly, the 3-kb-deletion including exon 8 reported by Ricci *et al.* involves an L1-*AluS* in *IDS* intron 7 and *AluSc* in intron 8, the latter of which we propose to have participated in the *Alu*-mediated recombination event in our case. Although we present a model in which the duplicated insertion was derived from the distal

region of its own strand through a loop formation, interchromatid or interchromosomal strand exchange is also a possibility. Because of the physical proximity, however, it would be more likely that the duplication was derived from its own strand. If this were the case, the strand exchange at the LCR would have to occur before the strand exchange at the L1-*Alu*-L1 to retain the region between LCRs in *IDS* and *IDSP1*. This raises the possibility that the *Alu*-mediated recombination was a subsequent event resulting from the genomic instability caused by recombination between LCRs. It is not clear why these two particular *Alu* sequences were chosen for the recombination event. There are 17 *Alu* repeats within this region and many of them are flanked by L1 LINE elements. Therefore, although this choice could have been stochastic, it may have involved an epigenetic mechanism. Nevertheless, considering the involvement of LCRs and L1-*Alu*-L1 repetitive elements, this particular complex rearrangement has the potential to be recurrent and may be present among those MPSII cases with previously uncharacterized aberrations involving *IDS*.

For the second case, we have evidence for NAHR occurring at the same LCRs as in the first case, and that the breakpoints are located within an *AluSp* element and a CT-rich region is consistent with NHEJ. Although CT-rich regions are not commonly recognized as frequent NHEJ sites, they may not be as stable as the other regions. Our data support a mechanism of coupled homologous and non-homologous recombination initially proposed in an experimental system.²⁰ In this model, when double-strand breaks are introduced, the repair is initiated by homologous recombination but is completed by NHEJ. This model has been proposed to explain the mechanism of heterogeneous duplications of the *PLP1* gene.²¹ In our case, NAHR at the LCRs and NHEJ at the *AluSp* element and CT-rich region would have to occur as one event instead of two sequential events. Otherwise, if NAHR occurs without NHEJ, the region containing the *AluSp* would be deleted, and if NHEJ occurs without NAHR, *IDSP1* would be deleted. Replication-based mechanisms such as FoSTeS, or fork stalling and template switching,²² remain a possibility for case 1, but are unlikely for case 2 because there is no homology at this NHEJ junction. Thus, our findings support coupled homologous and non-homologous recombination as a molecular mechanism explaining the complex rearrangement of our second case.

An important finding for this second case is the presence of a point mutation, c.1652G>C, which alters the normal stop codon, TGA, to a serine codon, TCA (p.X551S). This results in the predicted addition of 17 amino acids at the C-terminus of the protein. This change has thus far not been reported in the literature as a disease-causing mutation, and it is not clear whether or to what extent this alteration might affect the biological activity of iduronate-2-sulfatase. We had reported previously a patient with Duchenne muscular dystrophy who carried a splicing mutation with an adjacent deletion of the *DMD* gene.²³ We proposed that the presence of the disease-causing mutation might have elicited genomic instability by driving a deletion event in the attempt to correct the pre-existing mutation. Although it is highly speculative, it is possible that the presence of the stop codon mutation might have had a role in the recombination event in our present case.

Among *IDS* rearrangements, LCR-mediated NAHR that does not involve other regions appears to be relatively common.^{9,11} Herein we present two cases of *IDS* complex rearrangements, that we propose to involve LCR-mediated NAHR, but to be completed by different mechanisms. It is not known what triggers such complex rearrangements or how complex rearrangements might proceed once initiated. Additional case studies, not only involving *IDS* but also other loci, are required to elucidate further the processes that might promote such

complex rearrangements. These studies are likely to be facilitated as array-CGH analysis becomes more routine in clinical diagnostic testing for various genetic disorders.

ACKNOWLEDGEMENTS

We thank Irene Miloslavskaya for her technical assistance. This work was partly funded by grants from the NIH (R24CA78088) and the Ellison Medical Foundation.

- 1 Martin, R. A. Mucopolysaccharidosis type II. In: *GeneReviews at GeneTests* (database online). <http://www.genetests.org>. University of Washington: Seattle, WA, (2007).
- 2 Flomen, R. H., Green, E. P., Green, P. M., Bentley, D. R. & Giannelli, F. Determination of the organisation of coding sequences within the iduronate sulphate sulphatase (IDS) gene. *Hum. Mol. Genet.* **2**, 5–10 (1993).
- 3 Wilson, P. J., Morris, C. P., Anson, D. S., Occhiodoro, T., Bielicki, J., Clements, P. R. *et al.* Hunter syndrome: isolation of an iduronate-2-sulfatase cDNA clone and analysis of patient DNA. *Proc. Natl Acad. Sci. USA* **87**, 8531–8535 (1990).
- 4 Malmgren, H., Carlberg, B. M., Pettersson, U. & Bondeson, M. L. Identification of an alternative transcript from the human iduronate-2-sulfatase (IDS) gene. *Genomics* **29**, 291–293 (1995).
- 5 Bondeson, M. L., Malmgren, H., Dahl, N., Carlberg, B. M. & Pettersson, U. Presence of an IDS-related locus (IDS2) in Xq28 complicates the mutational analysis of Hunter syndrome. *Eur. J. Hum. Genet.* **3**, 219–227 (1995).
- 6 Rathmann, M., Bunge, S., Steglich, C., Schwinger, E. & Gal, A. Evidence for an iduronate-sulfatase pseudogene near the functional Hunter syndrome gene in Xq27.3-q28. *Hum. Genet.* **95**, 34–38 (1995).
- 7 Timms, K. M., Lu, F., Shen, Y., Pierson, C. A., Muzny, D. M., Gu, Y. *et al.* 130 kb of DNA sequence reveals two new genes and a regional duplication distal to the human iduronate-2-sulfate sulfatase locus. *Genome Res.* **5**, 71–78 (1995).
- 8 Bunge, S., Rathmann, M., Steglich, C., Bondeson, M. L., Tylki-Szymanska, A., Popowska, E. *et al.* Homologous nonallelic recombinations between the iduronate-sulfatase gene and pseudogene cause various intragenic deletions and inversions in patients with mucopolysaccharidosis type II. *Eur. J. Hum. Genet.* **6**, 492–500 (1998).
- 9 Lualdi, S., Regis, S., Di Rocco, M., Corsolini, F., Stroppiano, M., Antuzzi, D. *et al.* Characterization of iduronate-2-sulfatase gene-pseudogene recombinations in eight patients with Mucopolysaccharidosis type II revealed by a rapid PCR-based method. *Hum. Mutat.* **25**, 491–497 (2005).
- 10 Froissart, R., Da Silva, I. M. & Maire, I. Mucopolysaccharidosis type II: an update on mutation spectrum. *Acta. Paediatr. Suppl.* **96**, 71–77 (2007).
- 11 Lagerstedt, K., Karsten, S. L., Carlberg, B. M., Kleijer, W. J., Tonnesen, T., Pettersson, U. *et al.* Double-strand breaks may initiate the inversion mutation causing the Hunter syndrome. *Hum. Mol. Genet.* **6**, 627–633 (1997).
- 12 Karsten, S., Voskoboieva, E., Krasnopolskaja, X. & Bondeson, M. L. Novel type of genetic rearrangement in the iduronate-2-sulfatase (IDS) gene involving deletion, duplications, and inversions. *Hum. Mutat.* **14**, 471–476 (1999).
- 13 Ricci, V., Regis, S., Di Duca, M. & Filocamo, M. An Alu-mediated rearrangement as cause of exon skipping in Hunter disease. *Hum. Genet.* **112**, 419–425 (2003).
- 14 Hopwood, J. J., Bunge, S., Morris, C. P., Wilson, P. J., Steglich, C., Beck, M. *et al.* Molecular basis of mucopolysaccharidosis type II: mutations in the iduronate-2-sulfatase gene. *Hum. Mutat.* **2**, 435–442 (1993).
- 15 Gu, W., Zhang, F. & Lupski, J. R. Mechanisms for human genomic rearrangements. *Pathogenetics* **1**, 4 (2008).
- 16 Lupski, J. R. Genomic rearrangements and sporadic disease. *Nat. Genet.* **39**, S43–47 (2007).
- 17 Stankiewicz, P., Shaw, C. J., Dapper, J. D., Wakui, K., Shaffer, L. G., Withers, M. *et al.* Genome architecture catalyzes nonrecurrent chromosomal rearrangements. *Am. J. Hum. Genet.* **72**, 1101–1116 (2003).
- 18 Chen, L., Lee, L., Kudlow, B. A., Dos Santos, H. G., Sletvold, O., Shafeghati, Y. *et al.* LMNA mutations in atypical Werner's syndrome. *Lancet* **362**, 440–445 (2003).
- 19 Nishida, N., Koike, A., Tajima, A., Ogasawara, Y., Ishibashi, Y., Uehara, Y. *et al.* Evaluating the performance of Affymetrix SNP Array 6.0 platform with 400 Japanese individuals. *BMC Genomics* **9**, 431 (2008).
- 20 Richardson, C. & Jasin, M. Coupled homologous and nonhomologous repair of a double-strand break preserves genomic integrity in mammalian cells. *Mol. Cell. Biol.* **20**, 9068–9075 (2000).
- 21 Woodward, K. J., Cundall, M., Sperle, K., Sistermans, E. A., Ross, M., Howell, G. *et al.* Heterogeneous duplications in patients with Pelizaeus-Merzbacher disease suggest a mechanism of coupled homologous and nonhomologous recombination. *Am. J. Hum. Genet.* **77**, 966–987 (2005).
- 22 Lee, J. A., Carvalho, C. M. & Lupski, J. R. A DNA replication mechanism for generating nonrecurrent rearrangements associated with genomic disorders. *Cell* **131**, 1235–1247 (2007).
- 23 Oshima, J., Magner, D. B., Lee, J. A., Breman, A. M., Schmitt, E. S., White, L. D. *et al.* Regional genomic instability predisposes to complex dystrophin gene rearrangements. *Hum. Genet.* **126**, 411–423 (2009).