130 kb of DNA sequence reveals two new genes and a regional duplication distal to the human iduronate-2-sulfate sulfatase locus.

K M Timms, F Lu, Y Shen, et al.

Genome Res. 1995 5: 71-78
Access the most recent version at doi:10.1101/gr.5.1.71

RESEARCH

130 kb of DNA Sequence Reveals Two New Genes and a Regional Duplication Distal to the Human Iduronate-2-sulfate Sulfatase Locus

Kirsten M. Timms,[1] Fei Lu, Ying Shen, Craig A. Pierson, Donna M. Muzny, Yanghong Gu, David L. Nelson, and Richard A. Gibbs

Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, Texas 77030

Deficiency of IDS activity results in Hunter Syndrome (mucopolysaccharidosis type II), a fatal X-linked recessive disorder. We report characterization of 28 cosmids around the IDS locus in Xq28. Four overlapping cosmids have been sequenced in their entirety generating a 130-kb contig. These studies show the fine structure of the IDS gene and identify an IDS pseudogene-like structure located 20 kb distal to the active gene. Two novel genes have also been identified in this sequence, and one of these genes is also locally duplicated. Both homologs are expressed, and a number of alternative transcript products have been characterized. The presence of a highly conserved pseudogene-like structure within a larger duplicated region close to the IDS gene has significant implications for the study of mutations at this locus.

Iduronate-2-sulfate sulfatase (IDS, EC 3.1.6.13) deficiency results in Hunter Syndrome [mucopolysaccharidosis type II, (MPS-II)], a lysosomal storage disorder leading to abnormal storage of heparin sulfate and dermatin sulfate. The severe form of this X-linked recessive disorder causes progressive mental retardation, physical disability, and death before the age of 15 (OMIM, 309900). In a few patients the severe metabolic deficiency is complicated further by the occurrence of seizures (Wraith et al. 1991).

The IDS gene maps to the boundary of the Xq27.3 and q28 bands of the X chromosome, in a region that is relatively rich in anonymous markers. The coding region of the IDS gene has been sequenced (Wilson et al. 1993), and an exonic structure has been determined by PCR and direct DNA sequencing (Flomen et al. 1993). These data have been used for molecular analysis of IDS patients who exhibit a range of point mutations and deletions. In general, the severity of the disease correlates with the class of the mutation.

Further diagnostic studies have suggested that the molecular organization of the human IDS locus is more complicated than can be explained by the occurrence of a single gene. A number of male patients have been observed who are apparently heterozygous for a mutation at the genomic level, whereas only the altered sequence can be detected in the cDNA (Rathmann et al. 1995). Recently, homologous IDS genomic sequences have been found by PCR (Rathmann et al. 1995), and mapping (Bondeson et al. 1995) that suggest the presence of a linked IDS pseudogene-like structure.

To characterize the complete structure of the IDS locus, identify new IDS-related sequences, and search for additional genes in the region, we have generated 130 kb of contiguous DNA sequence from four overlapping cosmids (GenBank accession no. L43581). In addition to the IDS sequences two new genes and evidence for a regional duplication have been found.

RESULTS

Preliminary Studies

Fourteen cosmids that hybridized to an IDS cDNA probe were selected from a flow-sorted

[1]**Corresponding author.**
E-MAIL ktimms@bcm.tmc.edu; FAX (713) 798-5741.

X-chromosome library. Each was analyzed by fingerprinting, and the results were correlated with data from previous hybridization studies. It was not possible to determine an unambiguous map of overlapping cosmids from these data. Therefore, two different cosmids were selected for analysis on the basis of their unique restriction digestion fingerprinting patterns and because they each hybridized to anonymous probes believed to be a least 200 kb apart. The first cosmid, 112G10 (C2), contained homology to the probe TH4 (Willard et al. 1994), whereas the second, 169A5 (C6), had been identified by hybridization to DXS1113 (Willard et al. 1994).

Cosmid Sequencing

Cosmids C2 and C6 were sequenced in their entirety using the random shotgun method and automated fluorescent DNA sequencers. The sequence was edited to 99.99% accuracy and was entirely covered in both strands. Sequence strategies are discussed in Richards et al. (1994).

Cosmid C2 was shown to contain the entire IDS gene by comparison to the known cDNA sequence. A sequence comparison revealed two types of similarity between cosmids C2 and C6. First, C6 was found to contain copies of exons 2 and 3 and introns 2 and 7 of the IDS gene. This pseudogene-like structure showed an overall >88% homology to corresponding regions in the bona fide IDS gene. There was no apparent difference between the homology of exonic versus intronic regions of the two IDS genes. Second, the ends of the two cosmids were complementary, suggesting an unexpected 1-kb overlap to form a single 76-kb contig. The overlap was confirmed by PCR spanning the common region using two sets of primers from C2 and C6 (data not shown).

Comparison of C2 sequence with previously published intron–exon boundaries generated by PCR revealed one discrepancy. Flomen et al. (1993) reported a purine-rich tract adjacent to exon 4 that was not present in our sequence. We discounted the presence of exon 4 in the IDS pseudogene-like structure and, therefore, are unable to offer any explanation for this difference between our IDS sequence and the published version (Fig. 1).

Cosmid-walking Procedure

To identify minimally overlapping cosmids extending in both orientations from the C2/C6 contig we utilized a hybridization walking method (see Fig. 2 and Methods, below). This approach makes use of the clones generated during construction of shotgun sequencing libraries to identify neighboring minimally overlapping cosmids. Using this approach we were able to identify cosmids that overlap by as little as 1.8 kb.

Cosmids 145C10(D4) (proximal of C2) and 84H1(F3) (distal of C6) were isolated by this approach. Each was sequenced in its entirety. D4 overlaps C2 by 1.8 kb. F3 has an extensive overlap with C6 but extends the length of the contig by 15 kb to give a total of 130 kb of complete genomic sequence (Fig. 1).

Identification of Novel Genes

Computer analysis was carried out on all available sequence. Two previously unidentified genes were detected by similarities to entries in the expressed sequence tags data base (dBEST). The first (gene *X*), had similarities ranging from 85%–100% to a total of 11 ESTs, whereas the second (gene *Y*) identified 6 entries in dBEST. Exhaustive computer searching, including protein motif and stringent data base comparisons, failed to reveal any clues to the functional identity of gene *X* or gene *Y*. Gene *Y*, however, contains several repeat-like elements.

To analyze further the expressed sequences related to gene *X*, five separate EST clones were obtained from the laboratories where they were originally characterized. An additional cDNA was isolated from a placental cDNA library. Complete high fidelity sequence was obtained for all six clones, yielding a total of 6152 bases (GenBank accession nos.: L43575–L43580).

A comparison of gene *X* genomic sequence with all of the expressed sequence data is shown in Figure 3. The cDNAs are in two distinct groups: (1) four cDNAs identical to the genomic cosmid sequence, and (2) two sequences that had an apparent similar exonic distribution but with similarities of 85%–99%. These data are consistent with expression of gene X as well as one or more separate, related genes (labeled *X′*). Prediction of open reading frames (ORFs) from our sequence of the gene *X* and *X′* transcripts showed a complex pattern of potential translation products.

dBEST contains six entries with similarity to genomic sequences of gene *Y*. Expression of two exons have been confirmed by RT–PCR. The six ESTs showing homology to this gene also are not identical to the genomic sequence (98.5%–99.7%

homologous). These differences may represent errors in the single-pass EST data.

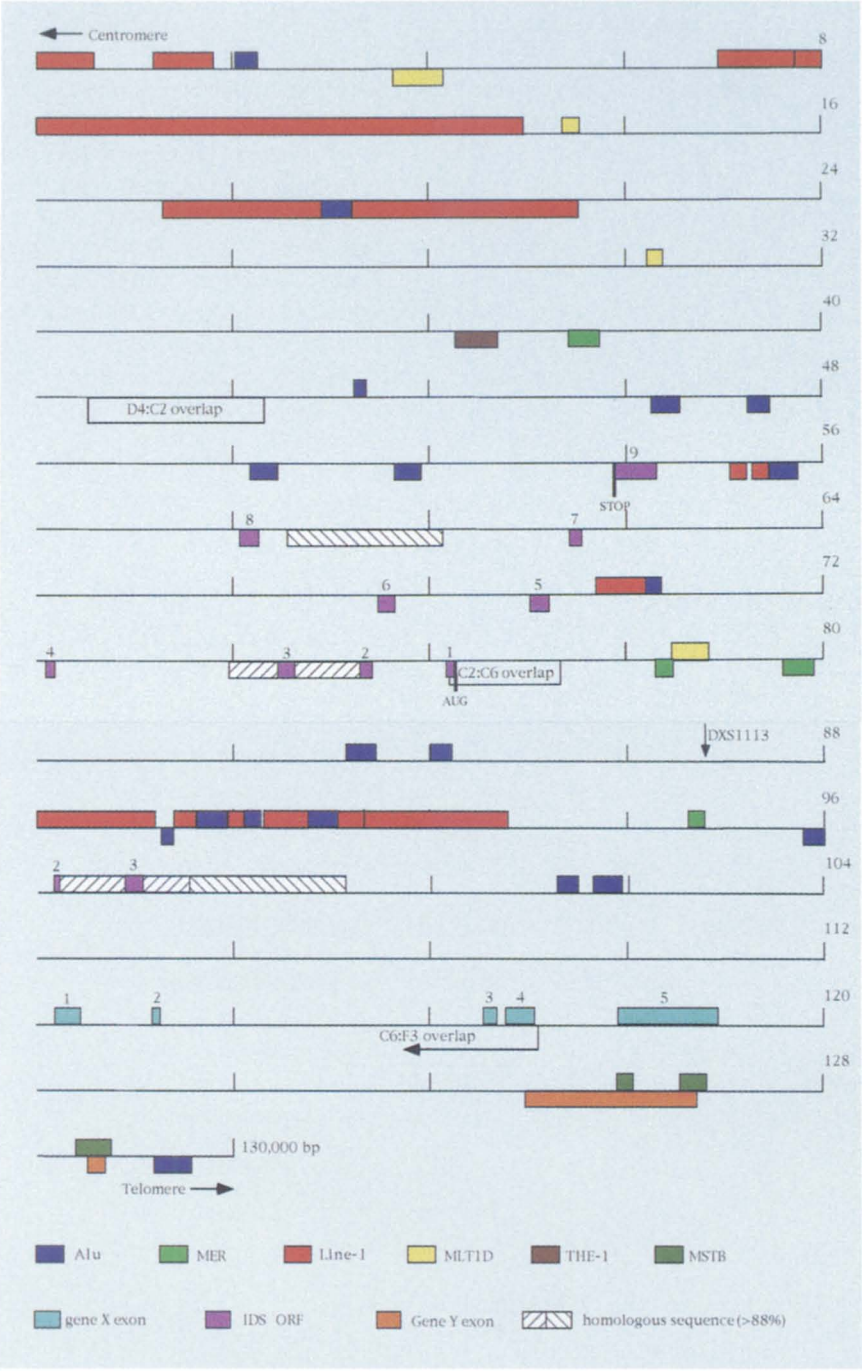

Figure 1 Schematic representation of a 130-kb sequence contig from the IDS region. The IDS gene, IDS pseudogene-like structure, and two additional genes of unknown function are represented by pink, pale blue, and light brown boxes. The other colored boxes indicate the positions of repetitive elements. Boxes above the line represent features oriented toward the telomere; below the line, toward the centromere. Hatched boxes represent homologous sequence (homology >88%) in the IDS gene and pseudogene-like structure.

Other Sequence Features

Computer analysis of this contig identified a number of repetitive elements interspersed within the region. Overall, there are 19 Alus and ~18 kb of sequence with homology to LINE-1 (long interspersed elements), or partial LINE-1, elements. Twelve other regions of similarity with known repeat families were identified throughout the region (Fig. 1).

Mapping the Distance from Gene *X* to Gene *X′*

To determine the location of the duplicated gene, probes, and PCR primers specific to either gene *X*, gene *X′* or both (*X* + *X′*) were generated using the exonic sequences of greatest divergence. These were then used to map gene *X′* to a position within 400 kb of gene *X*.

The X-chromosomal assignment was based on probing a somatic hybrid panel with probe *X* + *X′*, where the hybridization pattern revealed two strong bands from DNA digested with the restriction endonuclease *Bam*HI (data not shown). This pattern is consistent with the model of gene *X* duplication, as there is no *Bam*HI site within the probe that was used.

Finer scale mapping was achieved by the amplification of DNA from a collection of hybrid clones with human X-chromosome dele-

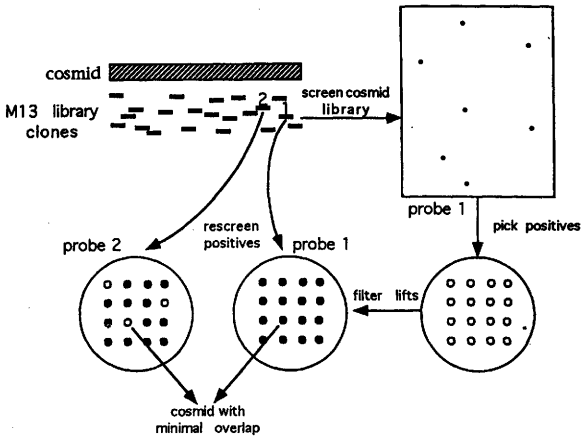

Figure 2 Walking to select cosmids with minimal overlap. Clones from a shotgun sequencing library are used to isolate neighboring minimally overlapping cosmids for large-scale DNA sequencing. A detailed description of this technique is given in Methods.

tions (Suthers 1991). Both gene *X* and *X'*-specific PCR primers amplified all hybrid DNAs except those defining an ~1 Mb region surrounding the IDS gene.

Finally, the distance from gene *X* to gene *X'* was narrowed further by indirect evidence from yeast artificial chromosome (YAC) hybridization experiments. First, 13 additional cosmids from the flow-sorted X-chromosome library were isolated using the *X* + *X'* hybridization probe. These were then divided into *X*- or *X'*-containing cosmids by PCR and hybridization. Reference to a local data base showed two of the *X'*-containing cosmids were found to have hybridized to YAC yWXD1904 (J. Parrish, pers. comm.). As this YAC had also hybridized to the distal three of the four cosmids in the 130-kb sequence contig, and the YAC is ~400 kb in length, then the maximum distance likely to separate gene *X* from gene *X'* is ~350 kb. These data are summarized in Fig. 4.

DISCUSSION

Genomic DNA sequencing has been used to characterize the region surrounding the human IDS gene. The sequence reveals the complete structure of the expressed IDS gene, shows the position of an IDS pseudogene-like element, and has revealed two novel genes within 40 kb distal to the IDS locus. Highly confident assembly of all the data was achieved by combining the genomic sequencing with a hybridization/walking strategy and extensive PCR verification of clones and genomic sequence. The structure of the entire region is consistent with the occurrence of one or more ancestral duplications leading to the present organization.

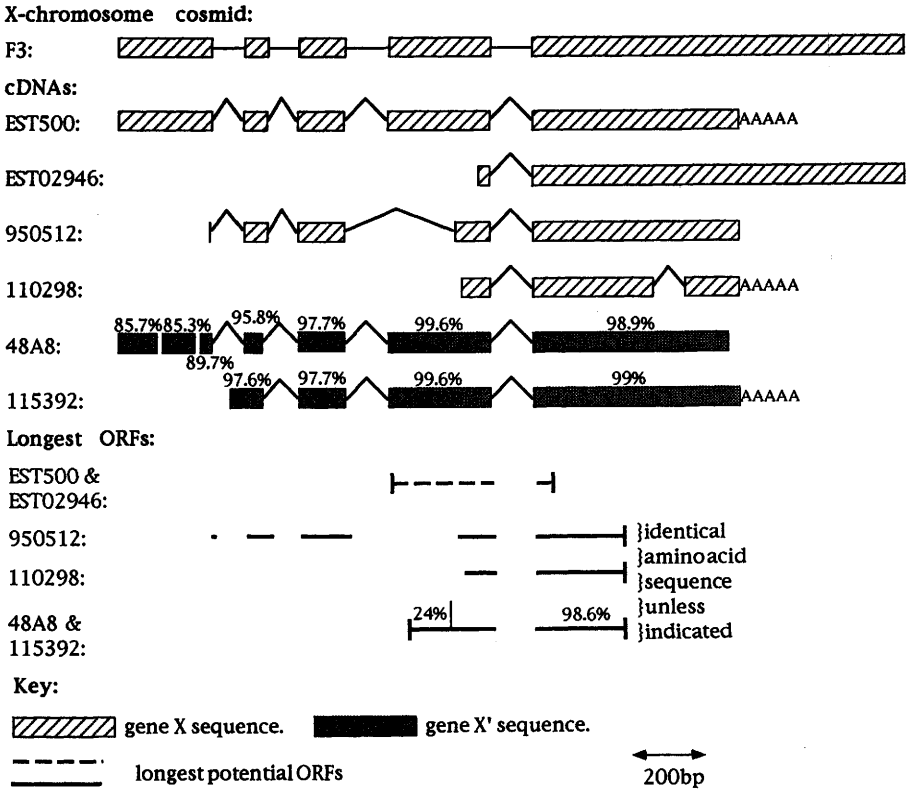

Figure 3 Alignment of F3 cosmid genomic sequence with gene X/X' cDNAs. Exons are drawn to scale; introns are not. Hatching patterns indicate where sequence homology is 100%. Homology between gene X' sequences and F3 cosmid sequence is indicated above each putative X' exon. The longest predicted ORF from each transcript is shown as either solid or broken lines. Vertical bars represent the position of either start or stop codons. The ORFs represented by the solid lines have identical amino acid sequence except where indicated.

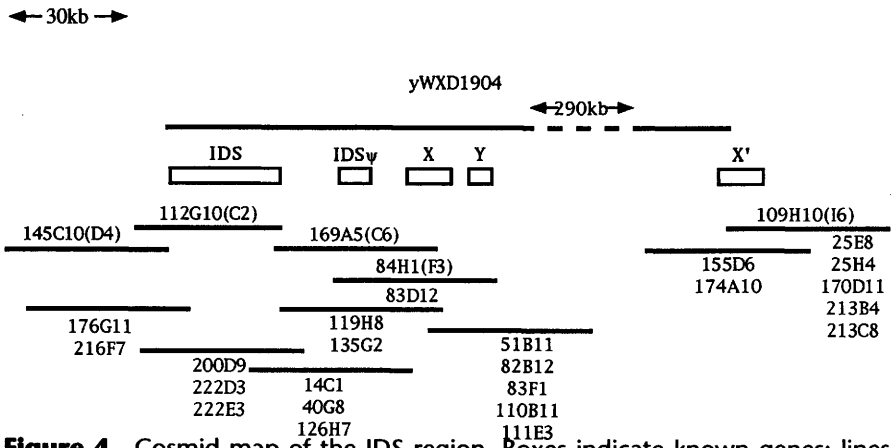

Figure 4 Cosmid map of the IDS region. Boxes indicate known genes; lines represent either cosmids or YAC yWXD1904. A total of 28 cosmids have been mapped to this region.

The presence of an IDS pseudogene-like structure only 20 kb distal to the active gene has important implications for the study of Hunter Syndrome patients. A number of patients have mutational lesions that could be best explained by homologous recombinations and/or deletions resulting from the proximity of these conserved sequences (Bondeson et al. 1995; Rathmann et al. 1995). Detailed comparison of junctional sequence in these patients should now be possible.

The identification of novel expressed genes within 40 kb of the IDS gene is also important for Hunter Syndrome patients. Severely affected patients showing unusual symptoms might show loss of these loci. Involvement of these loci in variants of Hunter Syndrome would also suggest that alterations in gene *X* or gene *Y* might cause recognizable disease even in the absence of IDS mutations. These issues can be readily resolved using the information reported here.

The events that lead to the current molecular organization are difficult to interpret, but the sequence of the pseudogene-like structure gives some indication of the possible mechanism of IDS evolution. Both exonic and intronic sequences are highly conserved between the two genes. Presumably if the pseudogene-like structure arose from the expressed gene, then this high level of conservation would suggest that it was a recent event. The presence of unusually long tracts of LINE-1 sequences near and between the two IDS gene is also intriguing and may represent the boundary of an early recombination event. The possible presence of additional undiscovered IDS-like genomic sequences nearby the current region cannot be overruled.

As gene *X* is also duplicated at a site within 350 kb, it is possible that the greater region evolved through a series of duplications. To understand this process it will be necessary to obtain additional genomic sequence, including that surrounding the gene *X'* exons.

It is interesting to note that there are now several other examples of regional duplications that have occurred in human evolution. In addition to well-characterized gene families such as the globin loci and the color blindness loci (Nathans et al. 1986) we note the recent report of genomic duplication in the polycystic kidney disease gene locus (Hughes et al. 1995). The IDS region now represents an additional example of this phenomenon.

Overall, this study represents an example of how a DNA sequence-based approach can efficiently resolve questions that are based on understanding the fine structure of a region. In this case, the study was hampered by apparent anomalies in the physical map and by a paucity of well-characterized cosmids from the region. Instead of delaying the study to further articulate the physical map we chose to press ahead with the genomic sequencing and continue the mapping as we proceeded. This led us to develop a hybridization-based strategy that uses prior sequence information to move rapidly from a sequenced cosmid to a minimally overlapping neighbor for further analysis. Combining the phases of mapping and sequencing allowed a sure path to the completion of the region and the fastest progress along the way.

METHODS

Cosmids

All cosmids were derived from the human X chromosome flow-sorted library prepared at Lawrence Livermore National Laboratories (LLNL). The cosmids characterized in this study are identified using the LLNL nomenclature (e.g., 112G10); this name can be used to isolate the origi-

nal clones from the LLNL X-chromosome cosmid library. Cosmid DNA was isolated and purified using Qiagen Plasmid Maxi Kit columns and then purified further by cesium chloride banding.

Sources of cDNAs

EST500 was from a colorectal cancer cDNA library; EST02946 was from a fetal brain library; 950512 was a clone from a Burkitt's lymphoma library; 48A8 was isolated from a placental library; both 110298 and 115392 were from a fetal liver/spleen cDNA library.

Walking to Select Cosmids with Minimum Overlap

In this scheme individual M13 clones are used to probe a cosmid library to identify minimally overlapping cosmids. The position of the M13 probes in the first cosmid are known from the sequence data. Probe 1 is chosen to be near to the end of the cosmid insert, and probe 2 is internal to the first probe. Cosmids from the primary screen that are found to hybridize to probe 1 are rescreened with both probes. Cosmids that hybridize to probe 1, but not to probe 2, overlap minimally and are selected for M13 library construction and sequencing (Fig. 2).

Shotgun Library Construction and Template Preparation

M13 shotgun libraries were prepared from nebulized or sonicated cosmid DNA using an adapter-based strategy (Andersson et al. 1994). In summary, after shearing, cosmid DNA was repaired to form blunt ends, ligated to adaptors generating an 11 bp overhang, and size fractionated to isolate fragments between 1.0 and 2.5 kb. ΔM13 vector with a complementary 11-bp overhang was prepared as described (Andersson et al. 1994). Inserts were annealed to vector DNA and transformed into XL-1 Blue-competent cells (Stratagene). White plaques were picked, grown for 6 hr, and stored at 4°C. Sequence template was prepared from these cultures as described (Kristensen et al. 1987). Reverse sequencing template was prepared using a modified asymmetric PCR protocol (Muzny et al. 1994).

Sequencing

Random and directed sequencing of clones was performed using fluorescence-labeled primers. Dye primer sequencing utilizing a Biomek 1000 workstation (Beckman Instruments Inc.) and Perkin-Elmer Cetus (PEC) 9600 thermocyclers was performed as described in Civitello et al. (1993). Reagent kits were supplied by Applied Biosystems (ABI). Dye terminator sequencing using sequence-specific primers and M13 templates was performed for gap closure. Reagent kits were provided by ABI and reactions were performed on PEC 9600 thermocyclers. Sequence reactions were electrophoresed on ABI 373A DNA sequencers.

Sequence Assembly

Sequence reads were edited using the software SEQPREP developed by the Molecular Biology Computational Resource Center at Baylor College of Medicine. After editing, sequence was assembled using the Staden XDAP software (Dear and Staden 1991). Gap closure was performed as described previously (Muzny et al. 1994; Richards et al. 1994).

Computer Analysis Programs

Computer analysis of sequence was via a suite of programs

Table 1. PCR primers used in this study

Name	Sequence	Application
IDS #1	5′-TCTGGTTCTGAGCTCCGTCTGC-3′	primers used in RT–PCR reactions to generate an IDS probe
IDS #2	5′-CTTGCCTTCTCTGCACAGCTCA-3′	
Gene Y RT–PCR #1	5′-CACATGGCGAGTGAGAGAGCAAG-3′	primers used in RT–PCR reactions to confirm the expression of gene *Y* exons identified by homology to dBEST entries
Gene Y RT–PCR #2	5′-CCCTCCAGTACCGTAAACATCAC-3′	
Gene Y RT–PCR #3	5′-GTGATGTTTACGGTACTGGAGGG-3′	
Gene Y RT–PCR #4	5′-GGCCAATTCACTCTGACTTCC-3′	
Gene X specific #1	5′-GGAGGAAACACTCTTCCACTTAGC-3′	PCR primers specific for gene *X* DNA
Gene X specific #2	5′-CAGCTTGGACACTAGCCAGGC-3′	
Gene X′ specific #1	5′-GGAGGAAACACTCTTCCACTTGGT-3′	PCR primers specific for gene *X′* DNA
Gene X′ specific #2	5′-CAGCCCTGACACGAGCCGGGT-3′	
Gene X/X′ #1	5′-GCTAGGATAGGAAGTAGCTG-3′	PCR primers which will amplify from either gene *X* to gene *X′* DNA
Gene X/X′ #2	5′-CTTTGCAATGCCCCGAAGAC-3′	
C6 #1	5′-CAGCTCTCAAGCCAGCTAGGG-3′	PCR primers used to confirm the overlap between C2 and C6 cosmids in reactions using genomic DNA as a template
C2 #1	5′-GGTAGGAAGGAGTGAAAAATGG-3′	
C2 #2	5′-GATGGTGGAAAGGAAAGATG-3′	

assembled by the Baylor College of Medicine genome informatics core. They include BLAST (Altschul et al. 1990) via the National Library of Medicine server, GRAIL (Uberbacher et al. 1991) and CENSOR (Jurka et al. 1995).

Hybridizations

Standard procedures for Southern hybridization were utilized (Sambrook et al. 1989), except that transfers were carried out using alkaline conditions. In addition, hybridization buffer contained 1 M $NaHPO_4$ (pH 7.0) and 7% SDS, and wash buffer contained 0.08–0.2 M $NaHPO_4$ (pH 7.0) and 1% SDS. Probes were random prime labeled using Rediprime kits provided by Amersham, and purified using NICK columns from Pharmacia.

Fingerprinting

Purified cosmid DNA was digested with restriction enzymes and electrophoresed in agarose gels using standard procedures (Sambrook et al. 1989). Visual comparison of band sizes between neighboring lanes was carried out to estimate overlap between cosmids.

PCR

PCR was performed using AmpliTaq DNA polymerase in a PEC 480 thermocycler. Each 50-µl reaction contained 1.5 mM $MgCl_2$, 50 mM KCl, 10 mM Tris-HCl (pH 8.3), 25 mM each dNTP, 20 pmoles each primer, and 2.5 units of AmpliTaq. DMSO [10% (vol/vol)] was included in reactions where template DNA had high GC content. For template, ~10 ng of cosmid DNA, 100 ng of genomic DNA, or 2 µl of reverse transcriptase reaction was used. PCR conditions were 94°C/5 min, followed by 30 reaction cycles of 94°C/30 sec, 55°C/30 sec, 68°C/2 min. After the completion of the cycle, samples were held at 68°C/7 min and then cooled to 4°C. Primers used in PCR reactions are shown in Table 1.

ACKNOWLEDGMENTS

This work was supported by grants ROI HG00823-OIAI and P30 HG00210 from the National Center for Human Genome Research. We thank Dr. Cheng Chi Lee for providing the human placental library and Dr. Julie Parrish for YAC hybridization data. The X-chromosome-specific cosmid library was kindly provided by Lawrence Livermore National Laboratory. We thank all of those individuals and laboratories who kindly provided the EST clones used in this study.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

REFERENCES

Altschul, S.F., W. Gish, W. Miller, E.W. Myers, and D.J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* **215:** 403–410.

Andersson, B., C.M. Povinelli, M.A. Wentland, Y. Shen, D.M. Muzny, and R.A. Gibbs. 1994. Adaptor-based uracil DNA glycosylase cloning simplifies shotgun library construction for large-scale sequencing. *Anal. Biochem.* **218:** 300–308.

Bondeson, M.-L., N. Dahl, H. Malmgren, W.J. Kleijer, T. Tönnesen, B.-M. Carlberg, and U. Pettersson. 1995. Inversion of the IDS gene resulting from recombination with IDS-related sequences is a common cause of the Hunter syndrome. *Hum. Mol. Genet.* **4:** 615–621.

Civitello, A.B., S. Richards, and R.A. Gibbs. 1993. A simple protocol for the automation of DNA cycle sequencing reactions and polymerase chain reactions. *J. DNA Seq. Map* **3:** 17–23.

Dear, S. and R. Staden. 1991. A sequence and editing program for efficient management of large projects. *Nucleic Acids Res.* **19:** 3907–3911.

Flomen, R.H., E.P. Green, P.M. Green, D.R. Bentley, and F. Giannelli. 1993. Determination of the organization of coding sequences within the iduronate sulphate sulphatase (IDS) gene. *Hum. Mol. Genet.* **2:** 5–10.

Hughes, J., C.J. Ward, B. Peral, R. Aspinwall, K. Clark, J.L. San Millán, V. Gamble, and P.C. Harris. 1995. The polycystic kidney disease 1 (PKD1) gene encodes a novel protein with multiple cell recognition domains. *Nature Genet.* **10:** 151–159.

Jurka, J., P. Klonowski, V. Dagman, and P. Pelton. 1995. CENSOR—a program for identification and elimination of repetitive elements form DNA sequences. *Comput. & Chem.* (special issue, in press).

Kristensen, T., H. Voss, and W. Ansorge. 1987. A simple and rapid preparation of M13 sequencing templates for manual and automated dideoxy sequencing. *Nucleic Acids Res.* **15:** 5507–5516.

Muzny, D.M., S. Richards, Y. Shen, and R.A. Gibbs. 1994. PCR based strategies for gap closure in large-scale sequencing projects. In *Automated DNA sequencing and analysis* (ed. M.D. Adams, C. Fields, and J.C. Ventner), pp. 182–190. Academic Press, San Diego, CA.

Nathans, J., D. Thomas, and D.S. Hogness. 1986. Molecular genetics of human color vision: The genes encoding blue, green, and red pigments. *Science* **232:** 193–202.

Rathmann, M., S. Bunge, C. Steglich, E. Schwinger, and A. Gal. 1995. Evidence for an iduronate-sulfatase pseudogene near the functional Hunter syndrome gene in Xq27.3-28. *Hum. Genet.* **95:** 34–38.

Richards, S., D.M. Muzny, A.B. Civitello, F. Lu, and R.A. Gibbs. 1994. Sequence map gaps and directed reverse sequencing for the completion of large sequencing

projects. In *Automated DNA sequencing and analysis* (ed. M.D. Adams, C. Fields, and J.D. Ventner), pp. 191–198. Academic Press, San Diego, CA.

Sambrook, J., E.F. Fritsch, and T. Maniatis. 1989. *Molecular cloning: A laboratory manual*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Suthers, G. 1991. Estimation of an approximate confidence interval for FRAXA location using linkage data from many pedigrees. *Am. J. Hum. Genet.* **49:** 462–464.

Uberbacher, E.C. and R.J. Mural. 1991. Locating protein-coding regions in human DNA sequences by a multiple sensor-neural network approach. *Proc. Natl. Acad. Sci.* **88:** 11261–11265.

Willard, H.F., F. Cremers, J.L. Mandel, A.P. Monaco, D.L. Nelson, and D. Schlessinger. 1994. Report and abstracts of the Fifth International Workshop on Human X Chromosome Mapping. Heidelberg, Germany, April 24–27, 1994. *Cytogenet. Cell Genet.* **67:** 295–358.

Wilson, P.J., C.A. Meaney, J.J. Hopwood, and C.P. Morris. 1993. Sequence of the human iduronate 2-sulfatase (IDS) gene. *Genomics* **17:** 773–775.

Wraith, J.E., A. Cooper, M. Thornley, P.J. Wilson, P.V. Nelson, C.P. Morris, and J.J. Hopwood. 1991. The clinical phenotype of two patients with a complete deletion of the iduronate-2-sulphatase gene (mucopolysaccharidosis II–Hunter syndrome). *Hum. Genet.* **87:** 205–206.

Received June 15, 1995; accepted in revised form July 11, 1995.