Differential Divergence of Three Human Pseudoautosomal Genes and Their Mouse Homologs: Implications for Sex Chromosome Evolution

Fernando Gianfrancesco,1,6,7 Remo Sanges,1 Teresa Esposito,1,6 Sergio Tempesta,2 Ercole Rao,3 Gudrun Rappold,3 Nicoletta Archidiacono,2 Jennifer A.M. Graves,4 Antonino Forabosco,5 and Michele D’Urso1

1International Institute of Genetics and Biophysics, CNR, 80125 Naples, Italy; 2 Institute of Genetics, University of Bari, Via Amendola, 165/A, 70126 Bari, Italy; 3 Institute of Human Genetics, Heidelberg University, 69120 Heidelberg, Germany; 4 School of Genetics and Human Variation, La Trobe University, Melbourne, Victoria 3083, Australia; 5 Department of Morphology and Legal Medicine Sciences, Medical Genetics, University of Modena, 41100 Modena, Italy

The human pseudoautosomal region 1 (PAR1) is essential for meiotic pairing and recombination, and its deletion causes male sterility. Comparative studies of human and mouse pseudoautosomal genes are valuable in charting the evolution of this interesting region, but have been limited by the paucity of genes conserved between the two species. We have cloned a novel human PAR1 gene, DHRSXY, encoding an oxidoreductase of the short-chain dehydrogenase/reductase family, and isolated a mouse ortholog Dhrsxy. We also searched for mouse homologs of recently reported PGPL and TRAMP genes that flank it within PAR1. We recovered a highly conserved mouse ortholog of PGPL by cross-hybridization, but found no mouse homolog of TRAMP. Like Csf2ra and Il3ra, both mouse homologs are autosomal; Pgpl on chromosome 5, and Dhrsxy subtelomeric on chromosome 4. TRAMP, like the human genes within or near PAR1, is probably very divergent or absent in the mouse genome. We interpret the rapid divergence and loss of pseudoautosomal genes in terms of a model of selection for the concentration of repetitive recombinogenic sequences that predispose to high recombination and translocation.

[The sequence data described in this paper have been submitted to the EMBL data library under accession nos. AJ293620, AJ296079, and AJ293619.]

The human sex chromosomes, heteromorphic in both size and gene content, are fossils of an ancestral, homologous chromosome pair (Ohno 1967). Over the 200 million years of mammalian evolution, the X and Y chromosomes lost homology as the Y chromosome was progressively degraded as the result of drift in nonrecombining regions and/or selection at linked loci on the Y (Charlesworth 1990). However, pseudoautosomal regions 1 and 2 (PAR1 and PAR2) at the termini of the short and long arms of X and Y chromosomes still recombine during male meiosis, ensuring X-Y nucleotide sequence identity that is necessary for normal male meiosis (Rappold 1993). In this paper, we will consider only the short-arm PAR1.

Between-species comparisons of the PAR and neighboring X-specific regions have aided our understanding of the genesis of this region, but also have presented mysteries. Cloning and mapping the marsupial homologs of genes within and near the human PAR1 revealed that this region was added to the X and Y since the divergence of marsupials and eutherian mammals 130 million years ago (Toder and Graves 1998). Comparisons with carnivores and artiodactyls showed that the human PAR1 has been recently reduced from a larger homologous region that included the steroid sulfatase (STS) gene, which lies just outside the PAR on the human X chromosome, and detects an inactive copy on the long arm of the human Y (Toder et al. 1997).

Comparison with mouse genes is harder to interpret. Human and mouse pseudoautosomal regions appear to have distinct evolutionary origin, as there are no genes that are pseudoautosomal in both species. Of the 10 cloned genes assigned to the human PAR1, eight appear to have no mouse homologs. The two that do recognize mouse homologs, Csf2mgm (the homolog of CSF2RA) and Il3ra (the homolog of IL3RA), both have diverged considerably from their human homologs (Hara and Miyajima 1992; Park et al. 1992) and both have been assigned to mouse autosomes (chromosomes 19 and 14, respectively) (Distech et al. 1992; Ellison et al. 1996). The only gene assigned to the PAR of the mouse is steroid sulfatase (Sts), which is pseudoautosomal in other eutherians, but not human. Mouse Sts is very divergent from human STS (63% identity at nucleotide and 59% at amino-acid levels [Salido et al. 1996]).

Spanning the boundary of the mouse PAR is the gene

6Present address: Institute of Molecular Genetics, CNR, 07041 Alghero, Italy.
7Corresponding author.
E-MAIL F.Gianfrancesco@igm.ss.cnr.it; FAX 39 079 946714.
Article and publication are at http://www.genome.org/cgi/doi/10.1101/gr.197001.
Fxy, which codes for a RING finger protein. The human homolog FXY is located on Xp22.3, near but not within PAR1 (Perry et al. 1998). The 5’ region (exons 1–3, including a significant portion of the coding region) has no Y homolog, while the 3’ region (exons 4–10) lies within PAR1 and therefore is present both on the X and Y chromosomes. Human FXY and mouse Fxy shows a very high sequence similarity at the DNA and protein levels, but the X-Y shared region is much less conserved (estimated at 170-fold higher for synonymous sites) than the 5’ portion (Perry and Ashworth 1999). It is hard to explain why genes in the pseudoautosomal region should diverge more rapidly than genes in the differentiated region of the X chromosome or the autosomes.

It is not clear, from the few human genes with mouse homologs, how general this observation is. It would be a great advantage to analyze other human PAR1 genes with mouse homologs. To further explore the relationship between human and mouse PARs, we have isolated a novel pseudoautosomal human gene called DHRSXY and undertaken comparative studies of this and two other genes that we have recently isolated: PGPL and TRAMP. We found a highly conserved mouse ortholog of PGPL and a less conserved mouse ortholog of DHRSXY, both of which are autosomal. There appears to be no mouse TRAMP ortholog. We discuss the implications of this finding in understanding the evolution of the pseudoautosomal region.

RESULTS

Cloning of a Novel Pseudoautosomal Gene

We previously isolated two human PAR1 genes, PGPL and TRAMP (Gianfrancesco et al. 1998; Esposito et al. 1999). To identify additional transcripts, we started from X-linked expressed sequence tags (ESTs) of the Human Transcript Map (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Gene). We used the IMAGE clone 447388 (AA702323) belonging to the most telomeric UniGene cluster (Hs.16129) of the map. A PCR product of 262 bp made from this EST was used to screen a cDNA library derived from a human undifferentiated male teratocarcinoma cell line, NT2/D1 (Skowronski et al. 1988). A cDNA clone of 2554 bp was isolated, sequenced, and submitted to the European Molecular Biology Laboratory (EMBL) with accession number AJ293620 (Fig. 1).

Sequence analysis of this cDNA clone revealed an open reading frame (ORF) of 993 bp encoding a putative protein of 330 amino acids. Using this protein as a query, the BLAST algorithm revealed homology with numerous proteins known as members of the short-chain dehydrogenase/reductase (SDR) family. Consistent with the recommendations of the Human Gene Organization nomenclature committee, we have named it DHRSXY (dehydrogenase/reductase of SDR family on X and Y chromosomes). The SDR family encompasses a wide variety of enzymes, most of which are known to be NAD- or NADP-dependent oxidoreductases of average-size 250 to 300 amino-acid residues. Although the similarity between the different SDR proteins can be as low as about 30%, two domains are conserved: the first binding the coenzyme, often NAD (GXXXGXG), and the second binding the substrate (YXXXK) (Jornvall et al. 1995). This latter domain determines the substrate specificity and contains amino acids involved in catalysis. Although DHRSXY is much larger than the average SDR enzymes, it exhibits the sequence features and distances between conserved motifs that are characteristic of SDR enzymes. In fact, DHRSXY amino-acid sequence analysis identified both the coenzyme, binding site GGTDGIG at positions 50–56 and the potential substrate binding site YAQSK at positions 208–212. At this stage, it is difficult to predict the role that DHRSXY may play in the metabolism of the cell, but its unique features make it an interesting target for further investigation.

Fine Localization and Expression of DHRSXY

Because the EST sequence used to clone human DHRSXY came from the most distal portion of the X transcript map, DHRSXY must lie within PAR1. To map DHRSXY...
more precisely, YAC clones covering the PAR1 were screened by PCR using sequence from the putative gene to design primers (Ried et al. 1995). Two YAC clones (H1130 and D1116) were positive. We also analyzed the cosmids covering these two YACs (unpubl.), finding three (LLNOYCO3'M43C10, LLNOYCO3'M16C12 and LLNOYCO3'M21E1) that were positive for Dhrsxy. These cosmids lie distal to ASMT. One (LLNOYCO3'M21E1) overlaps with the ASMT-containing cosmid LLNOYCO3'M52D2, so Dhrsxy must be physically adjacent to ASMT (Fig. 2).

The PCR product used for cDNA library screening was used as a probe to examine the levels of expression in various tissue types. Northern analysis detected two RNA species of about 1800 and 5000 bp in all human adult tissues tested (Fig. 3). Expression studies by reverse transcriptase (RT) and PCR using various somatic cell hybrid lines (Esposito et al. 1997) showed that the RT-PCR products were present in lines containing an active X, inactive X, or Y chromosome (data not shown). Thus, this novel gene, like other pseudoautosomal genes, escapes X inactivation and has a functional copy on the Y chromosome.

**Isolation of Mouse Homologs of Human Dhrsxy and Pgpl**

To isolate the murine ortholog of the human Dhrsxy gene, we performed an EST database search to identify ESTs of the mouse-related gene. A variety of mouse ESTs were detected, but only a few showed significant homology at the amino acid levels using the TBLASTN program. The EST (AI466831) with the highest similarity with the putative Dhrsxy protein was analyzed. The sequence of this IMAGE clone 716487 of 852 bp (named Dhrsxy) was deposited in the EMBL (accession no. AJ296079). This clone does not contain the full-length sequence, as the termination codon was lacking. Comparison of human Dhrsxy to mouse Dhrsxy by "BLAST2" sequences revealed a 64% identity at nucleotide and a 59% identity (84% similarity) at protein level (data not shown). Using primer pairs based on the mouse cDNA sequence, a genomic DNA of about 3 kb, containing a portion of this gene, was amplified by PCR to obtain a probe for in situ localization. Fluorescence in situ hybridization (FISH) was performed and a specific hybridization signal on chromosome 4E was detected (Fig. 4A).

To isolate the mouse ortholog of the recently reported Pgpl gene (Gianfrancesco et al. 1998), its cDNA was used to screen a 10.5-d embryonic mouse cDNA library (Stratagene). A cDNA of almost 1300 bp was isolated, subcloned, and fully sequenced (EMBL accession no. AJ293619). Comparison of mouse Pgpl and human Pgpl revealed 71% identity at the nucleotide level (data not shown) and 68% identity and 87% similarity at the protein level (data not shown). Two mouse multiple-tissue Northern blots were hybridized with the full-length mouse cDNA. Low expression of a transcript of 1800 bp was detected in heart, brain, spleen, lung, liver, kidney, and testis, but not in skeletal muscle. The same transcript also was detected in fetal tissues (data not shown). A genomic clone obtained by screening a mouse λ library (129 strain) was used for FISH to chromosomes of a mouse cell line. Pgpl was localized to chromosome 5, portion 5E (Fig. 4B).

Recently, we reported the cloning of a novel human PAR1 gene called TRAMP, having some characteristics of an ancient transposable element (Esposito et al. 1999). To isolate the murine homolog, we used the human full-length cDNA to probe a mouse genomic library (129 strain). Even at reduced stringency, the human probes failed to detect any cross-hybridizing signal. Moreover, we found no sequences in a mouse EST database that showed significant homology to human TRAMP either at nucleotide or protein level. FISH, using the human TRAMP clone at low stringency, failed to detect any hybridization signal on mouse chromosomes. The results suggest that the TRAMP rodent gene, if it exists, is highly divergent from its human counterpart.

**DISCUSSION**

The pseudoautosomal region, as the last vestige of the autosomal pair that differentiated into the sex chromosomes, might be expected to maintain the characteristics of autosomal regions. Whereas the differential region of the Y chromosome is expected to have a high mutation and attrition rate, there was no reason to suppose that the pseudoautosomal region would be subject to any special evolutionary forces and should therefore be as conserved in gene content and sequence as any autosomal region (Charlesworth 1991).

**Figure 2** Localization of the Dhrsxy gene in the human pseudoautosomal region.

**Figure 3** Gene expression of Dhrsxy in human tissues. The bars on the left indicate the position of migration of RNA markers. Reprobing of the same filter with a β-actin probe is shown in the bottom panel.

**Figure 4** Localization of the Dhrsxy gene in the human pseudoautosomal region.
Our results with previous data (Blaschke and Rappold 1997) demonstrate that this clearly is not the case. Only one human PAR gene (PGPL) is conserved in the mouse genome. The other three human PAR genes with mouse homologs (DHRXXY, CSF2RA, and IL3RA) and one that maps to the original larger eutherian PAR (STS) are poorly conserved in mouse and could be isolated only by functional strategies rather than by cross-hybridization experiments. FXY, whose mouse counterpart straddles the pseudoautosomal boundary, is much more divergent on the PAR than the differentiated side. Moreover, several human PAR genes (TRAMP, SHOX, XE7, ANT3, ASMT, MIC2, PBDX/Y), as well as several genes that were in the ancient eutherian PAR but differentiated recently in primate (KAL, PRKY, ARSD, ARSE), cannot be detected at all in mouse, implying that they have diverged beyond recognition, or have been lost from the genome (Ellison et al. 1996; Smith and Goodfellow 1994). The PAR also seems to have been rearranged or lost independently in primates and rodents. Comparative mapping in marsupials shows that this PAR was a part of a larger autosomal region translocated to the X and Y in an ancestral eutherian, as it includes CSF2RA, as well as STS and several genes from human Xp (Toder and Graves 1998). The difference in gene content between the human and mouse PAR therefore must represent losses from the mouse PAR, rather than gains to the primate PAR.

Why is the pseudoautosomal region undergoing this rapid evolution? An intriguing hypothesis involves the recombination rate. Recombination within the human PAR is 10 times more frequent during male meiosis than female meiosis (Rappold 1993). This is because there is an obligate crossing over event within this 2.6-Mb region in males that is required for correct segregation of the X and Y chromosomes. This high frequency of recombination may render the PAR more prone to mistakes and repair, creating sequence divergence between the two species. A higher recombination rate also may predispose the region to interchromosomal exchanges, making translocation to autosomes more frequent.

Why and how did the PAR evolve this high frequency of recombination? Heteromorphic sex chromosomes evolved originally from an autosomal pair, driven initially by differences at a sex-determining locus. Ohno (1967) suggested that subsequent differentiation of this pair to the sex pair was accomplished exclusively at the expense of one member of the pair (the Y chromosome in mammals) that was selected to accumulate sex-specific functions. Acquisition of a sex-determining function by locus on the proto-Y was the starting point for Y attrition (Fig. 5). As well as attrition, comparative mapping in marsupial and monotreme mammals demonstrates the addition of a large autosomal region to an ancient PAR early in eutherian divergence (Graves 1995). The process of attrition continued along this added region, ultimately leaving the present human PAR as a relic of this addition (Graves et al. 1998). The independent movement of the pseudoautosomal boundary contributed to the human/mouse differentiation. Thus the PAR is a product of two competing processes: addition of autosomal regions to the X and Y, and progressive attrition of the Y.

Attrition of the Y appears to be inexorable. Its end-point is likely to be complete loss of homology between the X and Y. This appears to have occurred in akodont rodents and in all marsupials, which do not undergo homologous pairing and recombination (Ashley et al. 1989; Sharp 1982; Toder et al. 2000), and must presumably have evolved alternate means of controlling chromosome segregation and avoiding nondisjunction.

We propose that to avoid this scenario, a
fierce selection is working to accumulate and concentrate repetitive recombinogenic sequences on the sex chromosomes, for more robust pairing and/or increased recombination, which are essential (at least in mouse and man) for male fertility (Fig. 5). Only some of these sites may be active during female meiosis, in which the two X chromosomes may pair along their full length. The mouse PAR is rich in repetitive sequences and the sole surviving PAR gene Sts has a highly diverged sequence to the extent that it is not recognizable by hybridization with human STS (Salido et al. 1996). This may represent the penultimate stage, in which selection is for conservation of a pairing function even at the expense of gene activity. Studies of a proviral insertion in the mouse PAR (Harbers et al. 1986) suggested that repetitive flanking regions are rearranged and duplicated frequently, perhaps by unequal recombination. The accumulation of these repetitive recombinogenic sequences also may act to promote translocations of pseudoautosomal genes to autosomes.

METHODS

**cDNAs and Sequence Analysis**

The cDNA library for the isolation of DHRSX gene was from a human uninduced male teratocarcinoma cell line, NT2/D1 (Skowronska et al. 1988). Mouse PGPL-related cDNA clone was isolated from a mouse 11 days embryo cDNA library (Clontech). The cDNAs were subcloned into pGEM-4Z vector (Promega-Biotech) and analyzed by Dye-Terminator cycle sequencing on an Applied Biosystem 377 automated sequencer. The Pgp1 genomic clone, used for the FISH assay, was isolated from a mouse genomic library (Stratagene) using the Pgp1 cDNA as a probe.

**Northern Blot Analysis**

Two human and two mouse multiple-tissue Northern blots (Clontech) were hybridized with the PCR product of DHRSX gene and with a Pgp1 cDNA, respectively. The Northern blots were prehybridized, hybridized, and washed according to the manufacturer’s directions (Clontech).

**Yeast Artificial Chromosome and Cosmid Screening**

For the localization of the DHRSX gene, YACs from the PAR1 were screened (Ried et al. 1995). Cosmids identified by PCR with the DHRSX primers were derived from the ICRF X chromosome-specific cosmids library and the Lawrence Livermore X chromosome-specific cosmids library.

**Cell Lines**

The panel of somatic-cell hybrid lines used in RT-PCR assays comprises two hybrids retaining the active human X chromosome; three hybrids retaining an inactive human X chromosome; a hybrid retaining two inactive human X chromosomes; and two hybrids retaining the human Y chromosome (Esposito et al. 1997).

**Reverse Transcriptase-PCR**

The RT-PCR for X inactivation experiments of DHRSX were carried out using 100 ng of RNA and 5 ng of cDNA as the template in a 10 µL PCR reaction containing 1 x TNAK, 100 buffer (Blanchard et al. 1993), 0.2 mM dNTPs, 0.35 units of AmpliTaq polymerase (Boehringer Mannheim), and 0.5 µM each of the primer sequences derived from cDNA. Using a DNA Thermal Cycler MJR (M.J. Research Inc.), we carried out 35 cycles of amplification using: 1 min at 94°C; 2 min at 56°C; and 2 min at 72°C. The DHRSX primers used in RT-PCR assay are 5’ ATA GTC AGT GGG ATT TTG TC 3’ and 5’ AAA ACC TCA CAC AGC CAA AG 3’.

**FISH Analysis**

Mouse spreads were prepared from an EM mouse cell line, which is mosaic for a Robertsonian fusion between chromosomes 2 and 15. Chromosome identification was facilitated by co-hybridizing Dhrsx and Pgp1 probes with PACs 372I17 and 426A18 (de Jong RP-21 library) mapping to chromosomes 4, band b, and chromosome 5, band G, respectively (http://www.biologia.uniba.it/mrc/9-MOUSE/-MOUSE.html). Digital images were obtained using a Leica DMRXA epifluorescence microscope equipped with a cooled CCD camera (Princeton Instruments). Cy3 (red), FluorX (green), and DAPI (blue) fluorescence signals were detected using specific filters and recorded separately as gray-scale images. Pseudocoloring and merging of images were performed using Adobe Photoshop software.

**ACKNOWLEDGMENTS**

We gratefully acknowledge M. D’Esposito and V. Ursini for critical reading of the manuscript. This work was supported by Telethon Italy to M.D., Ricerca Avanzata prot.18899.2000/University of Modena to A.F., the financial support cofin98-MURST to N.A., and Deutsche Forschungsgemeinschaft to G.R.

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**


Received May 17, 2001; accepted in revised form September 10, 2001.