Differentially regulated and evolved genes in the fully sequenced Xq/Yq pseudoautosomal region

Alfredo Ciccodicola, Maurizio D’Esposito, Teresa Esposito, Fernando Gianfrancesco, Carmela Migliaccio, Maria Giuseppina Miano, Maria Rosaria Matarazzo, Marcella Vacca, Annamaria Franzè, Monica Cuccurese, Massimo Cocchia, Anna Curci, Antonio Terracciano, Anna Torino, Silvia Cocchia, Grazia Mercadante, Emilio Pannone1, Nicoletta Archidiacono1, Mariano Rocchi1, David Schlessinger2 and Michele D’Urso+

International Institute of Genetics and Biophysics, CNR, Via Marconi 10, 80125 Naples, Italy, 1Institute of Genetics, University of Bari, Via Amendola 165/A, 70126 Bari, Italy and 2GRC/NIA/NIH, 5600 Nathan Shock Drive, Baltimore, MD 21224, USA

Received 13 October 1999; Revised and Accepted 7 December 1999

Human sex chromosomes, which are morphologically and genetically different, share few regions of homology. Among them, only pseudoautosomal regions (PARs) pair and recombine during meiosis. To better address the complex biology of these regions, we sequenced the telomeric 400 kb of the long arm of the human X chromosome, including 330 kb of the human Xq/YqPAR and the telomere. Sequencing reveals subregions with distinctive regulatory and evolutionary features. The proximal 295 kb contains two genes inactivated on both the inactive X and Y chromosomes [SYBL1 and a novel homologue (HSPRY3) of Drosophila sprouty]. The GC-rich distal 35 kb, added in stages and much later in evolution, contains the X/Y expressed gene IL9R and a novel gene, CXYorf1, only 5 kb from the Xq telomere. These properties make Xq/YqPAR a model for studies of region-specific gene inactivation, telomere evolution, and involvement in sex-limited conditions.

INTRODUCTION

During the progressive attrition of genes from the Y chromosome during evolution, the requisite pairing of Y and X at meiosis is maintained by the retention of some DNA that is the same on both chromosomes. Recombination between this ‘pseudoautosomal’ DNA at the end of the short arm of the chromosomes (Xp and Yp) assures the homogenization of sequence in the region (1,2). Like most autosomal genes, those in the 2.6 Mb Xp/YpPAR are expressed from both X and Y alleles (3).

A formally comparable second pseudoautosomal region, Xq/YqPAR, was discovered during the mapping of the X chromosome (4). It contains sequences that had been earlier recovered from both X and Y, and shows recombination over its entire extent (5). However, it also shows properties distinct from those of the Xp/YpPAR. The level of recombination is 6-fold higher than in the bulk of the X chromosome, though much less than the requisite recombination at every meiosis in the Xp/YpPAR (5,6). Perhaps more striking, when two genes were mapped to the Xq/YqPAR and studied further, IL9R showed the biallelic expression characteristic of Xp/YpPAR genes (7), but SYBL1 was found to be inactivated on both the inactive X and Y chromosomes, achieving dosage compensation in a unique manner (8).

To analyze the nature and regulation of Xq/YqPAR genes, we have sequenced the 330 270 bp region. Analysis shows that the two genes found earlier are characteristic of two distinct subregions. The two segments differ sharply in GC and repetitive sequence content and show different modes of gene expression. The proximal portion contains a second gene, HSPRY3, with the same unusual inactivation pattern reported for SYBL1: the distal portion contains not only IL9R but also a gene that is embedded in the GC-rich region adjoining the telomeric TTAGGG repeats. Consistent with recommendations of the HUGO nomenclature committee, we have named it CXYorf1. Like IL9R, it can be expressed even from inactive X chromosome.

Using probes for these genes, analyses of evolutionary dynamics reveal a quite complex stepwise addition of autosomal portions to the ancestral region, which subsequently translocated to the Y chromosome, thus forming the modern Xq/YqPAR. The analysis of allele-specific expression reveals a uniquely human situation for the Xq/YqPAR, in which, for two of the four constituent genes (SYBL1 and HSPRY3), the Y chromosome provides a reserve of inactive alleles that can be activated by recombination. We speculate on a possible relation of these inactivation properties to potential sex-limitation of some conditions mapped to distal Xq.

RESULTS

The sequence reported here (398 661 bp) extends from the first reported X-specific marker (DXS1108) (5) to the telomeric TTAGGG repeat, and is deposited in the EMBL database under

DDBJ/EMBL/GenBank accession no. Y19189

*To whom correspondence should be addressed. Tel/Fax: +39 081 7257247; Email: durso@iigbna.iigb.na.cnr.it
accession no. Y19189. The border of X and X/Y homologous sequence is in an L1 element, as previously described by Kvaloy et al. (9), where we have assigned the effective start of the PAR at nucleotide 68 391.

GC composition and repetitive elements define two zones

The entire region has an average G+C content of 39.43%, placing it in the GC-poorest class of human DNA ['isochore 1', up to 43% GC, as described by Bernardi (10)]. However, closer examination reveals two subregions. In the proximal 295 kb, the modal G+C value is 34.5% between the JXYQ (6) marker and the X22 marker defined here, and rises sharply to >51% in the distal 35 kb, between the IL9R and TelBam3.4 (11) markers (Fig. 1). This difference defines the boundary between zone 1 and zone 2 of the Xq/YqPAR (see below).

The two zones also differ in their content of highly repetitive Alu and LINE sequence elements, which constitute 67% of the nucleotide sequence in zone 1 but only 29% in zone 2. Furthermore, the LINE:Alu ratio falls from 4 in zone 1 to 0.4 in zone 2, with an abrupt transition corresponding to the shift in GC content.

Less repetitive sequence elements, including MERs and runs of more than five consecutive di-, tri-, tetra- and pentanucleotide repeats, are found throughout the region. A complete list is given in the EMBL submission (accession no. Y19189), including five variable number of tandem repeats elements for which we have thus far developed PCR tests that show polymorphism in human populations. These markers should help to characterize recombination and localize disease states in distal Xq. [Initial analyses with the X22 marker in 10 CEPH families, including the family 1377 used in earlier studies of recombination in the region (5,6), reveal differential recombination levels across the region (Cirigliano et al., manuscript in preparation)].

Four PAR genes include telomere-associated CXYorf1

Far more effective than computer-assisted predictions of gene content (see Materials and Methods) was the use of the expressed sequence tag (EST) database, which identified four genes in the Xq/YqPAR, positioned in Figure 1. The comparison of various ESTs and recovery of additional segments of cDNAs defined isoforms as well as exons and introns for four PAR genes.

Zone 1 of the PAR contained the known SYBL1 gene and the previously undetected HSPRY3. SYBL1 (synaptobrevin-like; EMBL accession no. X92396) (8,12) starts at 180 345 bp in the PAR, and is inactivated on both X and Y with an evolutionary history similar to HSPRY3 (see below). Its monoallelic expression has recently been confirmed by RFLP-based RT–PCR analyses of heterozygous individuals (M.R. Matarazzo et al., manuscript in preparation).

HSPRY3 starts 67 161 bp from the start of the PAR. Its two exons are separated by a 5524 bp intron. The 3' end was found as an EST match in GenBank (accession no. X97511); the first exon was recovered by a 5' RACE protocol (see Materials and Methods), and extends into a putative promoter region with no CpG island. The entire open reading frame (ORF) of 288 amino acids is in exon 2, and is widely expressed, in particular in fetal tissues, where transcripts of 4.4, 7.5 and >9 kb are seen (unpublished data). The existence of a long mRNA transcript with variable lengths of 3' untranslated sequences was confirmed by multiple ESTs extending matches over ~8 kb up to the stop codon of the ORF.

Blast search against the non-redundant sequence database revealed very high homology (68% identity and 86% similarity) with the Drosophila melanogaster gene sprouty. This gene has been implicated as an inhibitor of FGF signaling (13) and also as a regulator of Ras signaling (14). Three human and four mouse homologues have been reported (13,15), with the PAR gene identified previously only by a partial sequence that had not been localized in the genome.

Its ubiquitous expression facilitates the assessment of sex chromosome-specific expression of HSPYR3 in somatic cell hybrids retaining a Y or an active or inactive X chromosome as the sole human DNA (8). RT–PCR assays with primers derived from exons 1 and 2 demonstrated that the gene, like the nearby SYBL1 gene (8), is subject to X inactivation (Fig. 2). Apparently, X inactivation is maintained in all of zone 1 (see Discussion).

Zone 2 of the PAR contains the previously identified IL9R gene (16), starting at nucleotide 295 352. In addition, a novel gene was found starting at nucleotide 319 633 and ending <5 kb from the

Figure 1. An overview of the Xq/Yq pseudoautosomal region. In the top frame, G+C% is plotted, using a window size of 20 kb and a shift of 10 kb (right scale). Genes and pseudogenes are shown as boxes (all are orientated towards the telomere). Below are shown sequence features discussed in the text. Boxes corresponding to polymorphic markers and degenerate telomere repeats are shown by a black outline. Detailed analysis showed the distribution along the region for the repetitive sequences (Alu, LINE), detailed along with other repeats in the EMBL submission. The scale is in base pairs.
telomere-associated repeat sequence. This 5.5 kb genomic region shows high homology with >180 human ESTs and is also homologous to (though not previously noted) in subtelomeric sequence from chromosome 16p (17). The ESTs can be grouped into different subclasses of highly homologous transcripts (>98% identity). One (GenBank accession no. A1042587, from an IMAGE clone) was distinguishable from 16p sequence and other subclasses but identical to PAR sequence. When it was sequenced, the identity extended throughout its 962 bp. Detailed comparison of this cDNA with the genomic sequence revealed the last 4 exons, part of a fifth and a putative polyadenylation signal (AATAAA) 16 bp from the observed poly(A) tract. This partial gene, named CXorf1, hybridized to at least four RNA species, including an abundant 2 kb transcript, in all adult human tissues examined (unpublished data). The sequence of CXorf1 gives no obvious clues to its function, though a biological role could conceivably be related to its localization near the telomere (see Discussion).

The assessment of chromosome- and sex-specific expression was attempted using primers from the PAR-specific sequence, in the hybrid cells employed to analyze HSPRY3 (Fig. 2). Expression was seen in two hybrids containing inactive X chromosomes (THX88 and HY70) (8), though not in a third, and weak signals were also seen for both active X or Y in other somatic cell hybrids (unpublished data).

Zones 1 and 2 also contained a number of fragmentary pseudogenes. They include, for example, one previously reported (18) (GenBank accession no. U02035) for AMD2, at nucleotide 127 489 in zone 1, and six other ESTs with up to 96% identity to sequences in the last 5 kb of zone 2 (for example, sequence similar to part of a CHL1-related helicase gene that maps to chromosome 12 and to other human telomeric-associated DNA) (19). These ESTs, however, deviate from PAR sequence and contain multiple stop codons, and do not amplify with Xq-specific primers. The PAR thus contains bits of DNA related to sequences at the ends of other human chromosomes, consistent with frequent genetic exchange, as suggested earlier by Kvaloy et al. (9).

**Evolution of zones 1 and 2 in the human Xq/YqPAR**

Fluorescence in situ hybridization (FISH) results with probes from the HSPRY3 and CXorf1 genes provide data that, along with base composition, recombination and transcriptional features, support a step-wise evolution (12) of the Xq/YqPAR in at least two distinct zones. To assess the localization of these genes in primates, cosmid clones were used in FISH assays of various Old and New World monkeys (Fig. 3a).

With HSPRY3 as a probe for zone 1, a single signal on the X chromosome was revealed in all species, as earlier shown for SYBL1 (12) (Fig. 3a). In contrast, a cosmid containing the zone 2 gene CXorf1 revealed the expected X and Y signals in normal males, but homologous sequences were also revealed on at least six other human chromosomes. Other weaker signals are visible in scattered metaphases (Fig. 3b, which leaves out signals on chromosomes 15 and 16, where partial homology to the X subtelomeric region has already been detected) (11). In other primates (Fig. 3c–e) only a subset of localizations was conserved, indicating that successive additions or losses occurred as evolution proceeded. From the FISH results (Fig. 3c–e), mobilization of the CXorf1 gene subregion during evolution may have started from chromosome 12.

To localize these genes in the mouse genome, yeast artificial chromosomes (YACs 217E2, 143G8 and 377C9) were screened from the Whitehead Institute/MIT YAC library with 3′-untranslated region (3′-UTR) primer pairs from mouse IL9R (7). YACs 392A3 and 288H2 were recovered with intronic primer pairs from the mouse SYBL1 locus (12). In agreement with their locations in primates, Southern analysis of these YACs showed that mouse SPRY3 and SYBL1 colocalize within single YACs, whereas IL9R and CXorf1 do not (Fig. 4).
portion of the Xq/YqPAR (SYBL1) and other genes (as in ref. 21). In contrast, the genes in the proximal hybrid cells was not concordant, but this is similar to findings with genes [expression from one of three inactive X chromosomes in mouse genomic DNA; lane 2, Y AC 392A3; lane 3, Y AC 288H2; lane 4, negative control; lane 5, YAC 271E2; lane 6, YAC 143G8; lane 7, YAC 377C9. YACs in lanes 2–3 were isolated with SYBL1-specific primers. YACs in lanes 5–7 were isolated with IL9R-specific primers (see Material and Methods). Sizes are listed as shown. (a) IL9R; (b) SYBL-1; (c) CXYorf1; and (d) HSPRY3.

**DISCUSSION**

**Definition of the PAR**

A number of regions on the X and Y chromosomes share sequence content. In general, X-Y homologous sequences in the body of the chromosome are relics remaining on the Y after the evolutionary attrition of its gene content. Because these regions do not recombine at any detectable level between X and Y, they are gradually diverging by mutation, so that sequences clearly differ [e.g. in direct sequencing of probes from the 4 Mb Xq21.3–Yq11 homologous region (20)]. In contrast, the Xp/YpPAR and Xq/YqPAR sequences recombine, and therefore remain homogenized; but they differ in idiosyncratic secondary features, including recombination rates and region-specific gene regulation. The difference in recombination is quantitative. Xp/YpPAR recombines at least once at every male meiosis; the Xq/YqPAR recombines at a rate of 2% (though this is still 6-fold greater than the average for X-specific DNA). In contrast, the difference in regulation of the two PAR regions is qualitative, and readily apparent at the gene level. Like the Xp/YpPAR, where all genes escape inactivation, biallelic expression is seen for the distal portion of the Xq/YqPAR for IL9R and very likely for CXYorf1 genes [expression from one of three inactive X chromosomes in hybrid cells was not concordant, but this is similar to findings with other genes (as in ref. 21)]. In contrast, the genes in the proximal portion of the Xq/YqPAR (SYBL1 and HSPRY3) show a distinctive pattern, with inactivation on both X and Y chromosomes.

**Zones of the Xq/YqPAR: structure and evolution**

Differences in regulation are correlated with evolution and structure of zones 1 and 2 of the Xq/YqPAR. Zone 1 was first translocated onto the X chromosome at a time after the development of marsupials but before the development of mice (22). Zone 2, with its different content of GC and repetitive elements, was probably added later, in a multistep process that first created the IL9R subregion. The presence of a third, separate region, which includes CXYorf1 and the telomeric repeat, is clearly demonstrated using CXYorf1 as an evolutionary marker. In fact, the anonymous marker TelBam3.4 has already showed that telomere-associated DNA is common to Xq/Yq and other chromosomes (11). It is interesting to note that the mobilization of these sequences may have started from chromosome 12 and later spread on a number of autosomes, thus forming neo-telomeres (Fig. 3c–e). This notion is further supported by the finding of a degenerate TTAGGG repeat between transition zones (see below). Thus, the Xq/YqPAR has two transcriptionally distinct subregions, zones 1 and 2, but in terms of evolutionary origin, zone 2 shows a further division into at least two distinct subregions.

Figure 5 shows putative steps in evolution from the simpler X structure in mouse to the more complex structure in primates. Presumably all of the steps shown are subsequent to the stages of Y attrition defined for the non-recombinating portion of the sex chromosomes by Lahn and Page (23). The oldest segment of the PAR, zone 1, is characteristically GC poor and richer in LINEs compared with Alu sequences and, in fact, the centromeric border of zone 1 is defined by a crossover in a LINE element that created the region. The addition of zone 1 to X-specific sequences may have been correlated with the local concentration of LINE elements to the exclusion of genes [the first gene centromeric of the Xq/YqPAR border is BBOX2, transcribed starting 116 kb centromeric and in the centromeric direction (Ciccodicola et al., manuscript in preparation); HSPRY3, the first gene in the PAR, is 67 kb distal].

The sites of translocation/recombination events that added parts of zone 2 to the end of the X chromosome may be marked by some degenerate sequence remnants of telomere (TTAGGG) repeats between the SYBL1 and IL9R genes (at positions 250 446–254 511 and 279 809–280 386) and between the IL9R and CXYorf1 genes (at position 318 737–319 056) (Fig. 1). This may be a general feature of transitions between proximal and distal portions of subtelomeric regions (cf. ref. 24). CXYorf1 might then have been added in another step(s), a possibility strengthened by the finding that, in all mammalian species analyzed but human, IL9R and CXYorf1 are physically separated.

The final step in the actual creation of the Xq/YqPAR occurred when, late in primate evolution, all of zones 1 and 2 were duplicated as a unit at the end of the Y chromosome. Presumably this event involved another, L1-mediated translocation (9).

**X- and X/Y-inactivated genes: regionalization and possible mechanism**

It may not be surprising that zone 1 genes show ‘normal’ X inactivation, as they were in residence on the X chromosome for a long evolutionary period, and could well have been co-opted into a standard X inactivation module. The general notion is that action
from the \textit{XIST} locus spreads \textit{in cis} across the entire chromosome, and it has been suggested that there are ‘boosters’ (25) that act as relay stations to reinforce the process. But in addition to a current lack of molecular detail about what is moving along the chromosome and what a booster might be, models have also found it awkward to accommodate the existence of some genes that escape inactivation. Some of these genes are embedded among others that are fully inactivated. An attractive model has been that local decisions for inactivation are made on a regional basis (rather than gene by gene). This would be consistent with the fact that genes that escape inactivation seem to cluster, for example, in Xp/YpPAR and in Xp11.3 (26). Also consistent with a regionalization model is the differential inactivation of genes in the zones of Xq/YqPAR. White \textit{et al}. (27) have proposed that sequence differences are paramount, and Lyon (28) has suggested that \textit{LINE} sequences may somehow promote inactivation, acting as a point in which \textit{XIST} makes contact with the inactive chromosome. It is certainly striking that zone 1 contains a relative concentration of \textit{LINE} sequences.

What remains unanticipated and unaccounted for, however, is the phenomenon of \textit{Y} inactivation. \textit{Y} inactivation has already been reported for \textit{SYBL1}, but could have required a specialized, gene-specific mechanism. Instead, it is now clear that \textit{Y} inactivation occurs across a region, including both \textit{SYBL1} and \textit{HSPRY3}, extending >100 kb. \textit{HSPRY3} thus represents a second exception, in addition to \textit{SYBL1}, to a recently proposed model for the evolution of mammalian \textit{X} inactivation (29). In the absence of an inactivation center on \textit{Y}, no simple sequence (like \textit{LINE}s, for example) can determine inactivation, because the same genes function fully on the active \textit{X}. But as suggested earlier (8), neighbouring heterochromatin might block the access of transcription machinery to zone 1, or might otherwise promote inactivation by mechanisms including methylation of the promoter regions of \textit{SYBL1} (30). This model should be relatively easy to test with transfection studies of gene expression from artificial chromosome constructs containing zone 1 combined with varying extents of upstream \textit{Y} or \textit{X} sequence.

\textbf{Consequences of differential inactivation of \textit{Y} and \textit{X} genes?}

Its proximity to the telomere of many chromosomes (Fig. 3b, and the subgroups of \textit{CXorfI-like ESTs}) suggests that \textit{CXorfI} might act, for example, in modulating the local action of telomerase. But the genes in the rest of Xq/YqPAR are apparently a mixed bag, with \textit{SYBL1} involved in transport activities in the brain, \textit{HSPRY3} in growth factor signaling and \textit{IL9R} in cytokine action. It seems likely that they were randomly selected for transfer to \textit{X} and then to \textit{Y} (Fig. 5). The unique mode of dosage compensation for zone 1 might then have occurred fortuitously, as a consequence of \textit{Y} sequence in the vicinity of the translocation breakpoint, or might have somehow been subsequently selected.

However, once \textit{Y} inactivation of these genes is in place, they represent a curious and unprecedented instance in which zone 1 on the \textit{Y} is a reserve of silent alleles that are nevertheless potentially transcribed and translated. They would be mobilized for expression each time that recombination in the Xq/YqPAR exchanged \textit{Y} alleles onto the \textit{X} chromosome.

Because recombination occurs in only 2% of cases, this could have several effects: (i) a null allele usually on the \textit{Y} could be exchanged onto the \textit{X}, leading to inactivation of the only gene on the single \textit{X} in males, which could be disease producing and even lethal, and could have detrimental action in females as well; (ii) an
active Y allele might be transferred to an X that has a mutated
gene, increasing the activity from that X chromosome; and (iii) a
polymorphic variant of a gene on Y might be transferred to X and
become functional, possibly changing the probability of certain
gene-dependent effects. The availability of the sequence permits
direct tests to see whether any such scenarios involve alleles of
Xq/YqPAR genes in sex-limited conditions.

MATERIALS AND METHODS

Sequencing strategy
DNA sequence of the telomeric region of Xq28 was derived from
12 overlapping cosmids from a variety of sources. Cosmids B9.3,
A2.3, H2.2, A9.1, C12.3, D9.3 and C8.2 (kindly provided by
W.R.A. Brown, University of Oxford, UK) (9) were isolated from
a library constructed by A. Monaco (University of Oxford), from
YAC CRFy900F0523. The cosmids U130F6, U46E4, U59F4,
U14F6 and U160G10 were isolated by screening the Lawrence
Livermore National Laboratory chromosome X specific library.
Shotgun sequencing for each cosmid was carried out as described
by Chen et al. (31). To sequence sizeable gaps (where no cosmids
were recovered from different libraries) between U130F6 and
U46E4 (gU130U46 of 9346 bp) and between U14F6 and
U160G10 (gU14U160 of 21 186 bp), and to cover the sub-
telomeric region between U160G10 and TelBam3.4 (gU160/Tel
of 13 451 bp), long-range PCR reactions were performed using
Expand Long Template PCR System kit (Boehringer Mannheim,
Mannheim, Germany). The amplifications were carried out with
flanking primer pairs, and each product was subcloned into
Expand Cloning vector (Boehringer Mannheim) and sequenced
flanking primer pairs, and each product was subcloned into
M13 (451 bp), long-range PCR reactions were performed using
Expand Long Template PCR System kit (Boehringer Mannheim,
Mannheim, Germany). The amplifications were carried out with
flanking primer pairs, and each product was subcloned into
Expand Cloning vector (Boehringer Mannheim) and sequenced
after subcloning into M13 (gU14U160 and gU160/Tel) or by
sequential sequencing steps with primer sequences inferred from
a sequence of a sequenced tract (gU130U46).

Computer-assisted sequence analysis
Programs based on the GCG package (32) and simple derivatives
were used to infer the content of repetitive sequences. More
complex repetitive elements, including Alu, LINE and moderately
repetitive (MER) sequences were identified using the CENSOR
program (33). These are tabulated in the database entry. The total
sequence was analysed by using six different algorithms (GRAIL
v1.3, GENEFINDER, GENEID, GENESCAN, GENIE and
FGENE programs) (34–37) to predict the presence of putative
exons. The unique portions, indexed to their positions in the total
sequence, were then checked for any clusters of a selected group of
promoter elements using PROSCAN v1.7 and NNPP programs
(38,39). In addition, unique sequences were screened against
current dbEST, EMBL protein and DNA databases with BLAST
programs (40,41).

5′ RACE strategy
Human brain cDNA (0.5 ng; Marathon-Ready cDNA; Clontech,
Palo Alto, CA) was amplified with 0.5 µM Adaptor Primer (AP)
1 and 0.5 µM SPRY3R1 primers (5′-GTC TCT GCC CCA ACA
TTC CT-3′), in 2× buffer (15 mM MgCl₂, 15 mM Tris–HCl pH
8.0, 15 mM HEPES, 100 mM KCl, 4 mM DTT, 0.02% gelatin),
3 mM dNTPs, 0.5 µ U of Taq, in total volume of 12 µL. After 5 min
at 96°C, we performed 24 cycles of PCR (15 s at 95°C, 1 min at
60°C and 2 min at 72°C) with a final cycle of 5 min at 72°C. One-

thirtieth of the first PCR was amplified with 0.25 µM AP 2 and
0.25 µM SPRY3R1 primers (5′-AAT AAG GGA AGG CTG
GGA GA-3′), in 1× buffer (7.5 mM MgCl₂, 7.5 mM Tris–HCl pH
8.0, 7.5 mM HEPES, 50 mM KCl, 2 mM DTT, 0.01% gelatin),
5 mM dNTPs, 0.5 U of Taq, in a total volume of 12 µL. After 5 min
at 96°C, we performed 24 cycles of PCR (15 s at 95°C, 1 min at
60°C and 2 min at 72°C) with a final cycle of 5 min at 72°C. The
PCR products were cloned with TOPO cloning kit (Invitrogen,
San Diego, CA).

FISH analysis
Cell lines. Human metaphase spreads were obtained from PHA-
stimulated peripheral blood lymphocytes of a normal human
donor. Cell lines from primates have been previously reported
(42). Five primate species were used in this study: HSA (Homo
sapiens); three great apes: common chimpanzee Pan troglodytes
(PT), gorilla Gorilla gorilla (GGO) and orang utan Pongo
pygmaeus (PPY); and the Black Lemur Lemur macaco (EMA).

FISH. Cosmid U130F6 and plasmid gU160/Tel, containing part
of HSPRY3 and CX041/1 genes, respectively, were labelled with
biotin by nick-translation and hybridized in situ essentially as
described by Lichter et al. (43) with minor modifications.
Detection was performed using avidin-conjugated Cy3
(Amersham, Little Chalfont, UK). Chromosome identification
was obtained by simultaneous DAPI staining. Digital images were
obtained using a Leica DMRXA epifluorescence microscope
 equipped with a cooled CCD camera (Princeton Instruments,
Trenton, NJ). Cy3 and DAPI fluorescence signals, detected using
specific filters, were recorded separately as grey-scale images.
Pseudocoloring and merging of images were performed using the
Adobe Photoshop commercial software.

Isolation of mouse YAC clones
YAC clones were provided by UK HGMP Resource Centre
(Oxford, UK), using the Whitehead Institute/MIT Mouse YAC
Library. Primers used are: mIL9R1, 5′-GTC TCT GCC CCA
TTC CT-3′; mIL9R2, 5′-AAG CTC AAA GAT GGA CAA
CC-3′; syblex8r1, 5′-TCT ATT CAG TTT GGG TTT
3′; syblex8r2, 5′-CTT TAC TAT TAG AGA GC-3′. Syb1
primers are intron 7 specific, given the presence in the mouse
genome of a processed Syb1 pseudogene. We isolated three YAC
clones for IL9R (217E2, 143G8 and 377C9) and two for SYBL1
(392A3 and 288H2) with an average length of 820 kb.

Southern blot and hybridization
Mouse genomic DNA (15 µg) and 1 µg of each YAC were
EcoRI-digested, run in 1% agarose and blotted on Hybond N+
membrane. These blots were subsequently hybridized with the
following probes, using the Amersham Rediprime kit: mIL9R,
1.5 kb genomic probe, PCR amplified using the primers
mIL9R-f1, 5′-GGG TGG AGC ATT CAC CTG TC-3′;
mIL9R-r1, 5′-CCC CCT CTG TCT TGG CCT CA-3′; SYBL1,
669 bp cDNA probe, coding +154 bp of 3′-UTR; HSPRY3,
1140 bp cDNA probe, amplified by the primers SPRY3F1,
5′-GGG CTG AGC TTT GTT GGT TT-3′; SPRY3F1,
5′-ATC ACT CCT CCC CCT CTG CA-3′; CX041/1, 962 bp,
IMAGE clone as probe (GenBank accession no. AI042587).
ACKNOWLEDGEMENTS
The authors gratefully acknowledge L. Luzzatto, A. Ballabio, M.V. Ursini and B. Franco for helpful suggestions and the Sequencing Unit Core of Area di Ricerca di Napoli, CNR where the sequencing was performed. The financial support of Telethon, Italy, is gratefully acknowledged. The work was also funded by EEC grants BMH4-CT96-0338 and BMH4-CT96-1134.

REFERENCES

Human Molecular Genetics, 2000, Vol. 9, No. 3, 401