ARTICLE

The IL-9 receptor gene, located in the Xq/Yq pseudoautosomal region, has an autosomal origin, escapes X inactivation and is expressed from the Y chromosome

Joris Robert Vermeech1, Paul Petit1, Abdennaïm Kermouni2, Jean-Christophe Renauld2, Herman Van Den Berghe1 and Peter Marynen1,*

1Center for Human Genetics, Flanders Interuniversity Institute for Biotechnology, University of Leuven, Leuven, Belgium and 2The Ludwig Institute for Cancer Research, Brussels Branch, Brussels, Belgium

Received August 12, 1996; Revised and Accepted October 28, 1996

All human X-linked genes known so far, except for the Xp/Yp pseudoautosomal genes, are conserved as a single linkage group on the murine X chromosome. We show that the interleukin-9 (IL-9) receptor gene (IL9R), which is located within the human Xq/Yq homology region, maps to the murine chromosome 11. The Xq/Yq pseudoautosomal region (Xq PAR) thus represents a second region on the human X chromosome which is not X linked in mice. Furthermore, we show that IL9R is absent on the Y of great apes. IL9R is thus exceptional among X/Y genes in that it is X linked in some mammals, but autosomal or pseudoautosomal in others. Genes located on the X and the Y generally escape X inactivation. An exception to this rule is SYBL1, a gene located in Xq PAR. SYBL1 is X inactivated and is inactive on the Y chromosome. In contrast, we show that IL9R expression does occur from the Y, the active and the inactive X chromosomes. This finding raises the question of how the transcriptional regulation of genes within Xq PAR occurs and how the X inactivation status of IL9R has evolved following the autosome to X and the X to X/Y translocation. The evolutionary analysis of the IL9R gene, which is located at 10 kb from the telomere, and its pseudogenes at several telomeres, also provides insight into the evolution of these loci and of subtelomeric regions in general.

INTRODUCTION

The human sex chromosomes are structurally and genetically different from one another, except for two regions of 2.5 Mbp and 320 kbp located respectively near the short and long arm telomeres. These regions can pair and recombine during meiosis and are thus termed pseudoautosomal regions (PAR) (1). Linkage analysis indicates that a single obligatory cross-over occurs within the short arm PAR which is probably involved in proper sex chromosome segregation (2,3). The crossing-over frequency of the long arm PAR is much lower (4,5).

Two genes located in the short arm PAR in man, CSF2RA and IL3RA, are autosomal in mice, suggesting that this region is of autosomal origin (6,7). The origin of the long arm PAR is unknown. Bickmore et al. (8) demonstrated by Southern analysis that a sequence, which was later shown to originate from the long arm PAR (4,9,10), is present on the X but absent from the Y chromosome of great apes. Sequence analysis of the Xq/Yq pseudoautosomal boundary further supported the notion that the formation of this region occurred recently during human evolution (10).

Genes present on the mammalian X chromosome are generally subject to X inactivation, which results in an equal dosage of expressed sequences in males and females. A small minority of genes, most of which have copies on the X and the Y chromosome, escape X inactivation (reviewed by Disteche, 11). Surprisingly, SYBL1, a gene located within Xq PAR does not escape X inactivation and, in addition, is also inactive on the Y chromosome, thereby maintaining dosage compensation in an unprecedented way (12).

The interleukin-9 (IL-9) receptor gene (IL9R) is also located within Xq PAR at 10–25 kbp from the telomere. IL9R pseudogenes, lacking at least the promoter and exon 1a, are present at 9qter, 10pter, 16pter and 18pter (13). Using comparative mapping and analysis of sequence variation, we were able to investigate the evolutionary origin of the IL9R gene and pseudogenes. A polymorphism within the coding region of IL9R enables us to show expression from the X, inactive X and Y chromosome.

*To whom correspondence should be addressed
Figure 1. FISH was performed with cosmid 4A6. (A) Analysis of the male mouse cell line L1210 (ATCC). (B) Analysis of metaphase spreads from strain C57BL/6J-Ei-Rb2H (The Jackson Laboratory, Bar Harbor, ME), a mouse containing the Robertsonian translocation Rb(11,16). Arrows indicate the signal at chromosome 11. Cosmid 4A6 also produces a strong hybridization signal to the centromeric region of the Y chromosome of the C57BL/6J-Ei-Rb2H mouse, but not to the Y chromosomes of other mouse strains (A and data not shown).

Table 1. Percentage of individual great ape chromosomes at which a FISH signal could be detected with the IL9R-containing cosmid

<table>
<thead>
<tr>
<th>Human homolog chromosome</th>
<th>Gorilla gorilla</th>
<th>Pan troglodytes</th>
<th>Pan paniscus</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>100%(X)</td>
<td>90%(X)</td>
<td>60%(X)</td>
</tr>
<tr>
<td>9</td>
<td>40%(13)</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>12</td>
<td>40%(10)</td>
<td>55%(10)</td>
<td>40%(10)</td>
</tr>
<tr>
<td>16</td>
<td>60%(17)</td>
<td>40%(18)</td>
<td>60%(18)</td>
</tr>
<tr>
<td>Y</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

The assignment of the signal is based upon the human chromosome numbering, while the homolog great ape chromosome numbering is indicated between parentheses (28).

RESULTS

Mapping of the murine Il9r

A cosmid probe, 4A6, containing the mouse IL-9 receptor gene was used to localize the Il9r gene by fluorescent in situ hybridization (FISH) on metaphase chromosomes from the male mouse cell line L1210 (ATCC). A clear fluorescent signal was visible on chromosome band 11A4 (Fig. 1A). This was confirmed by FISH on metaphase spreads of the C57BL/6J-Ei-Rb2H mouse, which contains a Rb(11,16) translocation and the WMP/PasDn mouse, which contains a Rb(1,11) translocation. Signals were observed on the long arm of the Rb(11,16) chromosome of the C57BL/6J-Ei-Rb2H mouse (Fig. 1B) and the short arm of the Rb(1,11) chromosome of the WMP/PasDn mouse. Cosmid 4A6 also produced a strong hybridization signal on the centromeric region of the Y chromosome of the C57BL/6J-Ei-Rb2H mouse, but not on the Y chromosomes of other mice strains (Fig. 1A and data not shown). Since Il9r cDNA does not produce a hybridization signal on this Y chromosome, it appears that the Y chromosome of this specific mouse strain contains amplified non-coding sequences which are also present in the 4A6 cosmid.

Comparative mapping of IL9R in great apes

To investigate the evolution of the IL-9 receptor locus, cosmid AK604-37, containing human IL9R, was hybridized to metaphase chromosomes of the gorilla, the common chimpanzee and the pygmy chimpanzee (Fig. 2A–C). The chromosomal assignment was based on the banding patterns generated by the DAPI counterstain and was confirmed by performing two-color FISH using a human centromere X probe (data not shown). For the gorilla, 15 metaphases were scored. For each chimpanzee species, 30 metaphases were scored. Non-random signals observed at different chromosomes are indicated in Table 1. The location of the signals produced by the AK604-37 cosmid on the gorilla are homologous to human chromosome regions 9q34, 12q13, 16p13.3 and Xq28 and on the chimpanzee they


Figure 2. FISH was performed with cosmid AK608-37. The chromosomes showing a positive signal are indicated. No signals were observed at chromosome Yq. (A) Gorilla. Signals were observed at the telomeres of all chromosome X, on 60% of chromosome arm 17p and on 40% of chromosome arm 13q and above the centromere on 40% of chromosome arm 10p. (B) Common chimpanzee. Signals were observed on 90% of chromosome Xq, 40% of chromosome 18p and at 55% of chromosome arm 10p. (C) Pygmy chimpanzee. FISH was performed by using the Tyramide system (Dupont). The Tyramide system results in stronger signals but higher signal intensity variation. Signals were observed at 60% of chromosome arms Xq and 18p and 40% of 10p1.

correspond to human chromosome band 12q13, 16p13.3 and Xq28. In humans, the same locations contain IL9R pseudogenes and light up with AK604-37. However, in contrast to humans, no signal was detected on any Y chromosome in the gorilla, the common chimpanzee or the pygmy chimpanzee, nor on the chromosome 10 or 18 homologs. Therefore, the translocation of IL9R to Yq postdates the human–chimpanzee divergence.

**Genotyping IL9R**

Both the peculiar location and the exceptional chromosomal evolution of IL9R prompted us to examine the chromosomal origin of IL9R transcription. As IL9R transcripts are absent in Epstein–Barr virus (EBV) transformed and hybrid cell lines (data not shown), we searched for a polymorphism within the coding
sequence of *IL9R* which would allow us to assess the origin of *IL9R* transcripts. Upon comparison of the *IL9R* cDNA and its genomic sequence, a stretch of eight AGC trinucleotides (encoding for eight serine amino acids) was found in the cDNA sequence (bp 1469–1496, GenBank accession no. M84747), while in the corresponding genomic sequence of exon 9 a series of nine AGC trinucleotides was found. To determine whether this difference represents a polymorphism, a PCR using primer pair JVIL5/JVIL9 amplifying the putative polymorphism within exon 9 was developed and performed on X- and Y-containing hybrid cell lines. Separation of the amplification products on a sequencing gel showed only one amplification product on a sequencing gel showed only one 

### Table 1: Amplification products on a sequencing gel showed only one

<table>
<thead>
<tr>
<th>primers</th>
<th>JVIL5</th>
<th>JVIL8</th>
</tr>
</thead>
<tbody>
<tr>
<td>cDNA</td>
<td>(N) 100</td>
<td>(N) 100</td>
</tr>
<tr>
<td>X</td>
<td>(N) 100</td>
<td>(N) 100</td>
</tr>
<tr>
<td>16</td>
<td>(N) 100</td>
<td>(N) 100</td>
</tr>
<tr>
<td>Y</td>
<td>(N) 100</td>
<td>(N) 100</td>
</tr>
<tr>
<td>10</td>
<td>(N) 100</td>
<td>(N) 100</td>
</tr>
</tbody>
</table>

*Figure 3.* Sequence comparison of *IL9R* and its chromosome 16 allele A and 10 pseudogenes around the S8/S9 polymorphism. Primers used in the genotyping PCR reactions are indicated above the sequences. *IL9R* gene and pseudogene sequences previously published (Genbank accession no. L39062-4 and M84747) are in regular type while the newly derived sequence from the Y chromosome is shown in bold type. The intron sequences are in lower case and the exon sequences in upper case. The number of nucleotides are indicated with subscripts and behind the dash the number of mutations compared with the sequence derived from the X chromosome.

**DISCUSSION**

**Chromosomal evolution of the *IL9R* locus**

The genes present on the mammalian X chromosome have been strongly conserved as a single linkage group in all eutherian mammals. It was postulated by Ohno that this strong linkage is due to the constraints imposed by X chromosome inactivation (16). A translocation of an X-linked gene subject to X inactivation to an autosome or an autosome to X translocation would alter the gene dosage and such a change would be selected against. So far, only three genes are known that defy X linkage in all eutherian mammals. *CLCN4* was shown recently to be present on the X chromosome of *Mus musculus* but on autosome 7 in *M. spreus* (17). In human, two genes present within the short arm PAR, *CSF2RA* and *IL3RA*, map autosomal in mice (6.7). Our results demonstrate that the IL-9 receptor is located on mouse chromosome 11. Thus, the *IL9R* locus represents a second region of the human X chromosome, besides the short arm PAR, which
chimpanzee divergence (8,10). As all human receptors gene was duplicated recently onto the human Y chromosome, the IL-9 homologous site on the great apes Y indicates that the IL-9 pseudogenes are located on chromosome 11A, while in human the HBA gene cluster is located near the chromosome 16 telomere, ~10 kb centromeric from the chromosome 16 IL9R pseudogene (19). We postulate that IL9R is physically linked with the Hba locus and that the human 16p IL9R pseudogene is a remnant of the original ancestral IL-9 receptor.

The human IL9R probe detects homologous sites on the gorilla 9, 12, 16 and X and on the chimpanzee 12, 16 and X homologous chromosomes but not on their Y chromosome. The absence of an IL9R homologous site on the great apes Y indicates that the IL-9 receptor gene was duplicated recently onto the human Y chromosome, a finding in agreement with previous observations suggesting the duplication of Xq PAR following the human–chimpanzee divergence (8,10). As all human IL9R pseudogenes branched before the great apes divergence (13) and as a recent translocation created the Xq/Yq pseudoautosomal region, IL9R exists most likely as an X linked gene in the great apes. The finding that the strongest AK4604-37 cross-hybridizing locus is located on the great apes X chromosome (Table 1) further validates this hypothesis. IL9R is thus exceptional among X/Y genes in that it is X linked in some mammals (great apes), but autosomal or pseudoautosomal in others (mice and human).

Xp PAR genes are thought to have evolved by succeeding additions onto one sex chromosomal PAR region which subsequently recombined onto the other (20). The transition from an autosomal to the PAR does not integrate the gene within the X inactivation system and, consequently, preserves the gene dosage. In contrast, the addition of IL9R onto the long arm of X was not onto a pre-existing PAR, thus altering the IL9R gene dosage. The intriguing question which arises is how the X inactivation status of IL9R has evolved following the autosomal to X and later following the X to XY translocation. SYBL1, another gene within Xq PAR, is not only X inactivated but is also silenced on Y (12). Silencing of genes, possibly by position effect variegation (21–23), could have provided a way to keep the gene dosage constant following the duplication of Xq PAR genes. However, we show here that IL9R escapes X inactivation and is transcribed from the Y chromosome. The different transcriptional activities between SYBL1 and IL9R on the Y chromosome could be due to the fact that the cis-acting elements do not exert their silencing effect into the IL9R locus, which is located distally to SYBL1. Alternatively, an IL9R promoter may specifically escape silencing. As IL9R is not X inactivated and is expressed from Y in humans, the question remains whether or not this gene is X inactivated in the great apes. Further analysis of the IL9R X inactivation status in the great apes may provide more insights into the evolution of the X inactivation and X inactivation escape mechanisms.

Taking all information into consideration, we propose the following evolutionary history for the IL9R locus (Fig. 6). An IL9R ancestral gene resided near the telomere on what became the present-day human chromosome 16. During the mouse chromosomal evolution, a telomeric translocation, inversion or fusion internalized IL9R (24). The chromosome 16 telomere-associated region translocated from chromosome 16 onto a region that is now defined as the long arm PAR. A deletion caused the loss of exon 1a on the chromosome 16 IL9R.
after and well before the divergence of the great apes, other terminal translocations created the pseudogenes on chromosomes 9 and 10. Since no hybridization signals were detected on the chromosome 10 homologs of gorilla and chimpanzee nor on the chromosome 9 homolog of chimpanzee, the chromosome 10 and 9 pseudogenes have probably been lost during gorilla and chimpanzee chromosome evolution. Later, during human evolution, a second translocation may have added Telbam3.4 (25) near the IL9R gene and shortened the original subtelomeric region (10pter allele B-like) on chromosome X. After the human–chimpanzee divergence, the CA repeat present within intron 8 of the IL9R expanded on the X chromosome followed by a translocation of the Xq terminal region to the Yq chromosome, which created the present day long arm PAR (10, this paper). Finally, ectopic gene conversion or recombination between the IL-9 receptor gene present on chromosome X and chromosome 16 (allele B) generated allele A on chromosome 16 and is responsible for part of the allelic variation described at chromosome 16p13.3 (13,14,26).

MATERIALS AND METHODS

Cell growth and culture

Blood samples of the gorilla (Gorilla gorilla) and the pygmy chimpanzee (Pan paniscus) were provided by Dr De Meurichy from the Antwerp Zoo. A fibroblast cell line of the chimpanzee (Pan troglodytes) was obtained from the NIA Aging Cell Repository (Rep# AG06939A). Primary fibroblast cultures were initiated from skin biopsies of mice strains C57BL/6J-Ei-Rb2H and WMP/PasDn (The Jackson Laboratory, Bar Harbor, ME), and the murine L1210 cell line was obtained from the ATCC. Chromosome preparations were made by standard methods from peripheral blood or cell cultures. M07E was grown as described (27). White blood cells, obtained by fractionation of peripheral blood with Lymphoprep (GIBCO-BRL), were grown as described.

Fluorescent in situ hybridization (FISH)

The 4A6 cosmid containing the mouse IL-9 receptor gene was obtained by screening a genomic library from the DBA/2-derived P815 mastocytoma (cloned in the c2RB cosmid) with the murine Il9r cDNA as a probe (Renauld et al., manuscript in preparation). A total of 11 kb from cosmid 4A6 were sequenced covering exons 2–9 of the IL-9 receptor gene. The human cosmid AK608-37 (13) was used as a probe on human, gorilla and chimpanzee metaphase spreads. FISH was performed as described previously (13). Since the signals were weak on the pygmy chimpanzee metaphase spreads using the above described FISH method, we also used the TSA-Tyramid signal amplification system using fluorescein–Tyramide (Dupont, Boston, MA). The chromosomes were counterstained with 4, 6-diamidino-2-phenylindole (DAPI) and the slides were mounted in Vectashield mounting medium (Vector Laboratories, Inc., Burlingame, CA). The signal was visualized by digital imaging microscopy using a cooled charge-coupled device camera system (Photometrics Ltd., Tucson, AR). A G/Q banding pattern was generated using the DNA counterstain DAPI. Merging and pseudocoloring were performed using the Smartcapture software (Vysis, Stuttgart, Germany).

DNA and RNA isolation

Genomic DNA from peripheral blood cells was extracted using QIAamp (Qiagen, CA). Genomic DNA from established cell lines was isolated following standard procedures. DNA from human–rodent somatic cell hybrids containing chromosome Y (NA06317A), chromosome X (NA06318C) or chromosome 16 (NA10567) was obtained from NIGMS human genetic mutant cell repository. RNA was extracted from phytohemagglutinin-stimulated white blood cells and cell lines using Trisol (Life Technologies).
PCR and reverse transcription-based PCR

PCR of 0.1 µg of genomic DNA was performed within a 50 µl reaction mixture of 10 mM Tris–HCl pH 8.3, 50 mM KCl, 1.5 mM MgCl₂, 0.01% gelatine, 0.2 mM dNTPs, 2.5 U of Taq DNA polymerase and 25 pmol of each primer. Before the first cycle, the samples were incubated at 95°C for 5 min. Cycling conditions consisted of 20 s at 94°C, 20 s at 58°C and 20 s at 72°C for primer pair JVIL5 5′-FITC-CGCGTCCCCAGACTCAAG-3′ and JVIL9 5′-CTGAGGAGGGTCAAGGAGG-3′, of 30 s at 95°C, 30 s at 58°C, 30 s at 72°C for primer pair JVIL12 5′-TGCGCTTTCCAGAAGGAG-3′ and JVIL9.

The IL9R genotype is established by the combination of two PCR reactions. A first PCR using the primer pair JVIL5/JVIL9 amplifies fragments from the chromosome 16 allele A IL9R pseudogene and from IL9R on X and Y. Based upon the ratio of S8 and S9 products and the knowledge of whether alleles B or C on chromosome 16 are present, the genotype of IL9R can be deduced. The presence of chromosome 16 allele B or C can be determined by measuring the length of the (CA)n repeat polymerism in intron 8 by PCR (14), since a complete disequilibrium between the subtelomeric length polymorphism at polymorphism in intron 8 by PCR (14), since a complete disequilibrium between the subtelomeric length polymorphism at polymorphism in intron 8 by PCR (14), since a complete disequilibrium between the subtelomeric length polymorphism at polymorphism in intron 8 by PCR (14), since a complete disequilibrium between the subtelomeric length polymorphism at polymorphism in intron 8 by PCR (14), since a complete disequilibrium between the subtelomeric length polymorphism at polymorphism in intron 8 by PCR (14), since a complete disequilibrium between the subtelomeric length polymorphism at polymorphism in intron 8 by PCR (14), since a complete disequilibrium between the subtelomeric length polymorphism at polymorphism in intron 8 by PCR (14), since a complete disequilibrium between the subtelomeric length polymorphism at polymorphism in intron 8 by PCR (14), since a complete disequilibrium between the subtelomeric length polymorphism at polymorphism in intron 8 by PCR (14), since a complete disequilibrium between the subtelomeric length polymorphism at polymorphism in intron 8 by PCR (14).

For RT-PCR, first strand cDNA was synthesized from 2 µg of total RNA using 200 U of M-MLV reverse transcriptase and 50 ng of random hexamers in 20 µl of the buffer recommended by the manufacturer containing 0.5 mM of each dNTP (dATP, dCTP, dGTP and dTTP). The reaction proceeded for 1 h at 37°C and was stopped by heating to 95°C for 5 min. PCR was performed under the same conditions as for genomic DNA using the reverse transcription product of 0.2 µg of total RNA as template. For the first PCR reaction, cycling conditions consisted of 5 min denaturation at 95°C and 20 cycles of 40 s at 95°C, 40 s at 60°C, 40 s at 72°C using primer JVIL1 5′-GGATGCTGACGGTCTGTTGATGGG-3′ located within exon 1 and JVIL8 5′-AGCCAGCAACCCAGGAGGCAACTC-3′ located within exon 9. For the nested PCR, 1 µl of the first PCR product was used in a second amplification using the primer set JVIL5 and JVIL9 under the conditions described above.

Cloning of PCR fragment

To compare the sequences surrounding the allele S8/S9 polymorphism, the PCR amplification products derived from the DNA from cell hybrids containing only chromosome 16, X and Y using the primer pair JVIL12 and JVIL9 were digested by ScaI and PstI, ligated within pGEM3Z and transformed in Escherichia coli DH5α according to standard procedures. The plasmid DNA was purified with Qiaquick plasmid prep (Qiagen), dissolved in H₂O and sequenced with T7 DNA polymerase following standard procedures (Autoread, Pharmacia).

Analysis of X inactivation

Analysis of the X chromosome inactivation was performed as previously described (15). Briefly, genomic DNA was first digested with BstXI and PstI and subsequently the DNA was divided into two equal aliquots, one of which was digested with HpolI. DNA fragments were separated on a 1.5% agarose gel and transferred to nylon filters (Hybond-N+; Amersham,UK). Filters were hybridized with random primed 32P-labeled phospho-glycerate kinase (PGK) probe and washed under high stringency conditions.

ACKNOWLEDGEMENTS

We would like to thank Dr De Meurichy for providing blood samples of the gorilla and the pygmy chimpanzee, Dr G. Matthijs for help with the analysis of the methylation status of PGK and the helpful discussions, and R. Thoelen and L. Poleunis for their help with the karyotyping. These investigations have been supported by a ‘krediet aan navorser’, 1994’ from the ‘Nationaal Fonds voor Wetenschappelijk Onderzoek’ of Belgium, the Belgian Federal Service for Scientific, Technical and Cultural Affairs, the National Fund for Scientific Research (FNRS/NFWO) and the ‘Opération Télèvie’. Peter Marynen is an ‘Onderzoeksdirecteur’, Joris Robert Vermeesch an ‘Aangesteld Onderzoek’, Belgium and J.C.Renaudl is a research associate with the FNRS, Belgium.

REFERENCES


