Downloaded from genome.cshlp.org on August 4, 2022 - Published by Cold Spring Harbor Laboratory Press

Single haplotype analysis demonstrates rapid evolution of the killer immunoglobulin-like receptor (*KIR*) loci in primates

Jennifer G. Sambrook,[1] Arman Bashirova,[2] Sophie Palmer,[1] Sarah Sims,[1] John Trowsdale,[4] Laurent Abi-Rached,[5] Peter Parham,[5] Mary Carrington,[3] and Stephan Beck[1,6]

[1]*The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SA, United Kindgom;* [2]*Laboratory of Genomic Diversity, National Cancer Institute, Frederick, Maryland 21702, USA;* [3]*Basic Research Program, SAIC-Frederick, Inc., Frederick, Maryland 21702, USA;* [4]*Immunology Division, Pathology Department, University of Cambridge, Cambridge, CB2 1QP, United Kingdom;* [5]*Department of Structural Biology, Department of Microbiology, and Department of Immunology, Stanford University School of Medicine, Stanford, California 94305, USA*

The human killer immunoglobulin-like receptors (*KIR*) are encoded within the Leukocyte Receptor Complex (LRC) on chromosome 19q13.4. Here we report the comparative genomic analysis of single *KIR* haplotypes in two other primates. In the common chimpanzee (*Pan troglodytes*), seven *KIR* genes (*ptKIRnewI*, *ptKIRnewII*, *ptKIR2DL5*, *ptKIRnewIII*, *ptKIR3DP1*, *ptKIR2DL4*, *ptKIR3DL1/2*) have been identified, and five *KIR* genes (*mmKIRnewI*, *mmKIR1D*, *mmKIR2DL4*, *mmKIR3DL10*, *mmKIR3DL1*) are present in the haplotype sequenced for the rhesus macaque (*Macaca mulatta*). Additional cDNA analysis confirms the genes predicted from the genomic sequence and reveals the presence of a fifth novel *KIR* gene (*mmKIRnewII*) in the second haplotype of the rhesus macaque. While all known human haplotypes contain both activating and inhibitory *KIR* genes, only inhibitory *KIR* genes (characterized by long cytoplasmic tails) were found by in silico and cDNA analyses in the two primate haplotypes studied here. Comparison of the two human and the two non-human primate haplotypes demonstrates rapid diversification of the *KIR* gene family members, many of which have diverged in a species-specific manner. An analysis of the intronic regions of the two non-human primates reveals the presence of ancient repeat elements, which are indicative of the duplication events that have taken place since the last common ancestor.

[The sequence data from this study have been submitted to EMBL/GenBank/DDBJ under accession nos. BX842589–BX842591 and AY728181–AY728190. The following individuals kindly provided reagents, samples, or unpublished information as indicated in the paper: M.B. McChesney.]

Human Leukocyte Antigen (HLA) class I molecules, expressed on the surface of most nucleated cells, and Natural Killer (NK) cell receptors are both involved in immune recognition in humans (Trowsdale 2001). They are encoded by large, polymorphic, gene-dense clusters in the Major Histocompatibility Complex (MHC), the Leukocyte Receptor Complex (LRC), and the Natural Killer Complex (NKC) found on chromosomes 6p21.3, 19q13.4, and 12p13.1, respectively. The killer immunoglobulin-like receptors (*KIR*) are members of a superfamily of immunoglobulin genes found within the LRC (Wende et al. 2000) and interact with the gene products of *HLA-A*, *HLA-B*, and *HLA-C* found in the MHC classical class I region. There is evidence for epistatic interactions between *KIR* and *HLA* class I genes in susceptibility to disease (Carrington and Norman 2003). On the other hand, the nonclassical class I molecule HLA-E is recognized by a C-type lectin family member in the NKC. While the KIRs represent the major functional NK cell receptors in humans, in mice Ly49 molecules carry out the analogous function (Natarajan et al. 2002). However, the two are not phylogenetically or structurally related, and have diverged between primates and rodents since the last common ancestor about 87 million years ago (Mya) (Springer et al. 2003). The one identified human *LY49* gene appears to be nonfunctional (Westgaard et al. 1998), and only two homologous *KIR*-like genes, mapping to chromosome X, have been identified in mouse (Welch et al. 2003). Here we report the genomic characterization of the *KIR* regions of two other primates: the common chimpanzee (*Pan troglodytes*) and the rhesus macaque (*Macaca mulatta*). These two species diverged from humans ~5 and ~30 Mya, respectively (Horai 1995; Kumar and Hedges 1998). A family of *KIR* cDNA has been characterized in common chimpanzee (Khakoo et al. 2000) and rhesus macaque (Grendell et al. 2001; Hershberger et al. 2001) as well as bonobo (Rajalingam et al. 2001), orangutan (Guethlein et al. 2002), and gorilla (Rajalingam et al. 2004). The observation that only a minority of *KIR* genes are conserved between species, with the majority being specific to one or two species, makes it of considerable interest to compare the organization of the *KIR* gene family in different species.

Human *KIR* haplotypes differ in gene content and in the combination of *KIR* alleles, allowing for great diversity in the number and combination of *KIR* genes in different individuals (Trowsdale et al. 2001; Vilches and Parham 2002). To date, at least 14 expressed *KIR* genes (*KIR3DL1-3*, *KIR3DS1*, *KIR2DL1-5*, and *KIR2DS1-5*), two gene fragments (*KIR3DP1* and *KIR2DP1*), and >100 cDNAs (deposited in public databases) have been iden-

[6]**Corresponding author.**
E-mail beck@sanger.ac.uk; fax 44-(0)1223-494919.
Article and publication are at http://www.genome.org/cgi/doi/10.1101/gr.2381205.

Downloaded from genome.cshlp.org on August 4, 2022 - Published by Cold Spring Harbor Laboratory Press

tified. From the two major groups of haplotypes (named A and B), there are four framework genes that are conserved (*KIR3DL3*, *KIR3DP1*, *KIR2DL4*, and *KIR3DL2*). Genes with only two immunoglobulin (Ig) domains have either the D0 and D2 (D0+D2) or D1 and D2 (D1+D2) configurations, and seem to be derived from genes with three domains by either exon skipping, where exon 3 is a pseudoexon (*KIR2DL1-3*, *KIR2DS1-5*), or exon loss (*KIR2DL4*, *KIR2DL5*), where the D1 domain is absent (Vilches et al. 2000). The *KIR3DL3* gene is missing exon 6. The cytoplasmic tails can be long (L) or short (S), with the former carrying one or two immunoreceptor tyrosine-based inhibitory motifs (ITIM) that transduce inhibitory signals. In contrast, the short tail *KIR*s possess a charged residue in the transmembrane (TM) region that mediates an association with *DAP12*, which contains an immunoreceptor tyrosine-based activating motif (ITAM) and transduces activating signals. The prototypical *KIR* gene product, named *KIR3DL*, has a signal peptide (encoded by exons 1 and 2), the extracellular three Ig domains (exons 3 to 5) that attach via a stem (exon 6) to the transmembrane region (exon 7), and the cytoplasmic tail (exons 8 and 9). In this manner, the NK cells express two distinct sets of HLA class I-specific receptors that control NK cell function and lysis of foreign cells.

Both the LRC and MHC are regions of high plasticity within the human genome containing members of multicopy gene families. The evolutionary forces driving the genesis of NK receptors and their HLA ligands, as a concerted response to pathogens, may have taken place in parallel. Successive duplication events have resulted in the genomic clustering of both the *HLA* and *KIR* genes and pseudogenes (Shiina et al. 1999; Wilson et al. 2000). Furthermore, the presence of repetitive sequences, such as short and long interspersed repeats (*SINEs* and *LINEs*) and elements with long terminal repeats (*LTRs*), would contribute toward the diversity within these segments of DNA (Gaudieri et al. 1999; Martin et al. 2000). These may also be responsible, in part, for the expansion and contraction of the locus that appears to occur by unequal crossing over (Martin et al. 2003). Propagation of *Alu SINEs* in primate genomes has resulted in a series of subfamilies of different ages (Batzer and Deininger 2002). The human *KIR* genes, for example, contain *AluS* sequences but no *AluJ* repeats (Trowsdale et al. 2001), emphasizing that they have evolved relatively recently. We have carried out phylogenetic analyses (exonic sequences) and repeat analyses (intronic sequences) of *KIR* genes in chimpanzee, rhesus macaque, and human to help elucidate the complex evolution of these genes in primates.

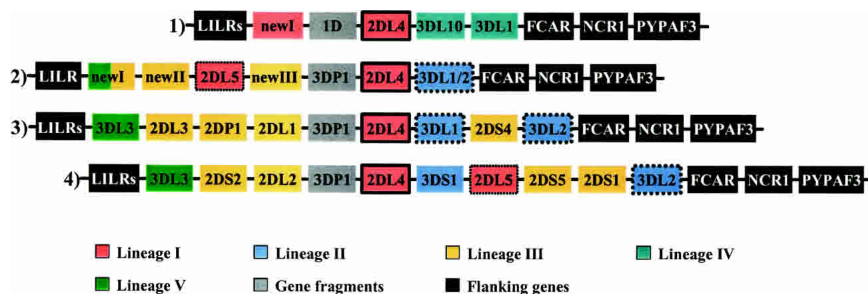

Figure 1. Comparison of *KIR* haplotypes in three primates (not to scale): rhesus macaque (1), common chimpanzee (2), human haplotype type A (3), and human haplotype type B (4). *KIR* lineages have been color-coded. Genes that are found in common between the three species have corresponding gene box borders. Only the extracellular immunoglobulin domains of *ptKIRnewI* are related to human *KIR3DL3*. *PtKIR3DL1/2* shares common features with both human *KIR3DL1* and *KIR3DL2*. Flanking framework genes surrounding the *KIR* loci are shown in dark gray.

Results and Discussion

Analysis of the common chimpanzee *KIR* region

The 201-kb single chimpanzee sequence haplotype was analyzed and found to contain a cluster of *KIR* genes and four surrounding genes (Fig. 1). The flanking genes include a member of the leukocyte immunoglobulin-like receptor (*LILR*) family, the receptor for the Fc fragment of IgA (*FCAR*), natural cytotoxicity triggering receptor 1 (*NCR1*), and the terminal end of the PYRIN-containing Apaf1-like protein 3 (*PYPAF3*) gene. All seven *ptKIR* genes are tightly clustered within 106 kb and are arranged in a head-to-tail arrangement. The *ptKIR3DP1* fragment is missing the last four exons at the genomic level, and a 4-bp insertion in the exon encoding the Ig D1 domain results in a frameshift and early termination site. From the remaining six putatively expressed genes, four have exons for all three Ig domains, two have exons for only two Ig domains, and they all possess long cytoplasmic tails. As all the encoded *KIR*s contain one or two ITIM motifs, their putative function would be to convey inhibitory signals to NK or T-cells expressing these *KIR*s. In contrast to activating *KIR* receptors, they also lack a positively charged lysine residue in the transmembrane domain that associates with the *DAP12* signaling molecule. No *KIR* genes with short cytoplasmic tails were identified in this genomic region. The majority of *ptKIR* cDNAs identified previously also have long tails (Khakoo et al. 2000).

Only two chimp orthologs of the human *KIR* genes were clearly identifiable: *ptKIR2DL4* and *ptKIR2DL5*. They are both 94% identical at the protein level with their human orthologs. The presence of *ptKIR2DL5*, in particular, in clone RPCI-43-61P22 confers a resemblance to human haplotypes of type B. Thus the two *ptKIR* genes that have human orthologs are those that lack an exon encoding the D1 domain.

The gene at the 5′-end of the chimpanzee *KIR* haplotype *ptKIRnewI* is similar to *KIR3DL3*, the human gene at the 5′-end of the locus, and like *KIR3DL3* it lacks an exon 6. In the exons encoding the extracellular Ig domains, *ptKIRnewI* and *KIR3DL3* appear orthologous, but in the exons encoding the transmembrane region and cytoplasmic tail, they diverge. This pattern of similarity and divergence is likely the result of a recombination (Table 1). A cDNA that may correspond to an allele of *ptKIRnewI* has been identified (Rajalingam et al. 2001). The gene at the 3′-end of the chimpanzee *KIR* haplotype is identical to the *ptKIR3DL1/2* cDNA (Khakoo et al. 2000), although the translated cDNA sequence (AF258798) is missing the first 14 amino acids. It appears to be orthologous to human *KIR3DL1* in the extracellular Ig domains and to human *KIR3DL2* in the stem, transmembrane, and cytoplasmic tail. Again the human and chimpanzee *KIR* genes are seen to have diverged through recombination. The result is that the telomeric part of the *KIR* locus comprises a single gene, *ptKIR3DL1/2*, in the chimpanzee haplotype, whereas it consists of three or more genes in human *KIR* haplotypes. Note that the two ends of the *KIR* locus, corresponding to the 5′ part of the human *KIR3DL3* gene and the 3′ part of the *KIR3DL2* gene, are conserved in the chimpanzee *KIR* haplotype as well as the central region containing *KIR3DP1* and *KIR2DL4*.

Downloaded from genome.cshlp.org on August 4, 2022 - Published by Cold Spring Harbor Laboratory Press

Table 1. **Domain and full-length nucleotide comparisons between the chimpanzee and macaque novel genes with previously identified cDNAs and *KIR2DP1***

	Closest sequences (number of differences)				
	Full-length	**D0**	**D1**	**D2**	**S/TM/CYT**
ptKIRnewI	*ptKIRCI* (23)	*ptKIRCI* (1)	*ptKIRCI* (1)	*ptKIRCI* (0)	*KIR2DPI* (12) *ptKIRCI* (13)
ptKIRnewII	*ptKIR2DL6* (23)	*ptKIR2DL6* (1)	*ptKIR3DL6* (1) *ptKIR2DL6* (2)	*ptKIR3DL6* (0) *ptKIR2DL6* (0)	*ptKIR2DL3* (13) *ptKIR2DL6* (20)
ptKIRnewIII	*ptKIR2DL6* (39)	*ptKIR2DL6* (4)	*ptKIR3DL6* (10) *ptKIR3DL7* (10) *ptKIR2DL6* (11)	*ptKIR3DL7* (9) *ptKIR2DL6* (9)	*ptKIR3DL4* (10) *ptKIR2DL6* (15)
mmKIRnewI	*mmKIR2DL5.1* (3)	*mmKIR2DL5.1* (0)	*ggKIR3DL3* (25)	*mmKIR2DL5.1* (3)	*mmKIR2DL5.1* (0)

Domains have been divided into Ig D0, Ig D1, Ig D2 and the combined sequence of the stem, transmembrane, and cytoplasmic tail (S/TM/CYT).

The remaining two *ptKIR* genes (*ptKIRnewII* and *ptKIRnewIII*) are found in the centromeric part of the chimpanzee *KIR* haplotype, where they flank *ptKIR2DL5*. These two genes are similar to each other and also to the cDNA *ptKIR2DL6* (Table 1; Fig. 2). The gene encoding *ptKIR2DL6* has been shown to contain a pseudoexon 3, which is homologous to exon 3 of *KIR3DL* genes but is not incorporated into mature mRNA. The sequence similarity between pseudoexon 3 of *ptKIR2DL6* and the homologous sequences of *ptKIRnewII* and *ptKIRnewIII* (1-nt and 4-nt differences, respectively) reveals the possibility that these genes could encode *KIR2D*, although *KIR3D* is not ruled out. *PtKIRnewII*, *ptKIRnewIII*, and *ptKIR2DL6* form a closely related trio of *KIR* that are very similar to each other in the extracellular domains and are more divergent in the stem, transmembrane, and cytoplasmic domains. Uncertain is whether *ptKIR2DL6* is paralogous to both *ptKIRnewII* and *ptKIRnewIII* or whether it is allelic to one of them.

```
SIGNAL PEPTIDE
ptnewII   MSLMVVSMACVGLFLLQGAWPHE
pt2DL6    --------------M........
ptnewIII  ....A.......F..........
pt3DL6    --------------........M
pt3DL4    -----------------.RR.LM
pt3DL5    -----------------.RR.LM

DO DOMAIN
ptnewII   GGQDKPLLSAWPSLVVPLG-HVILWCHSYLGFKNFSLYKEGGVPVPELYNRIFWKSLFMGPVTPAHTGTYRCRGSHPHSPSGWPAPSNPLVIVVT
pt2DL6    ...................-............N..............................................................
ptnewIII  ...N...............-....R.......N.....................N..............................S.........
pt3DL6    ......F......A..SE.EY.T.Q.R.H...NE...S..D.M........V.RN...........A...........FLT..S.A......M..
pt3DL4    ......F......A..SE.EY.T.Q.R.R...NE...S..D.M........V.QN...........A...........FLT..S........M..
pt3DL5    ......F......A..SE.EY.T.Q.R.R...NE...S..D.M........V.RN...........A...........FLT..S........M..

D1 DOMAIN
ptnewII   GVHRKLSLLAYPGPLVKSEEAVILQCWSDVMFEHFLLHREGKFNDTLRLTGELHDGVSKANFSIGHMTQDLAGTYRCYGSLTHSPYLLSAPSDPLDIVIT
pt2DL6    .................................................................R..............V...................
ptnewIII  .....P....H.........T............................................P..E.................Q.............
pt3DL6    .................................................................R..................................
pt3DL4    .....P....H.........T....................M......HI..H.........P..R....................Q.............
pt3DL5    .....P.F..H..R......T..................K.M.....S....H............P..E...........VP....Q.............

D2 DOMAIN
ptnewII   GLYEKPSLSAQPGPTVQAGESVTLSCSSQSSYDMYHLSREEGGHERRLPAGPKVNGTFQADFSLGPATHGGTYRCFGSFRDSPYEWSDPSDPLLVSVT
pt2DL6    ..................................................................................................
ptnewIII  ................W..........................R...........R..........................................
pt3DL6    ..................................................................................................
pt3DL4    ...............IL...N.......RR.....R....GES...................P..............Y.........KS.........
pt3DL5    ................L...N.......R.........G.GEA.......VT........E.P.....Q...........V......NS.........

STEM
ptnewII   GNPSNSWPSPTEPSSKTGNPRHLH
pt2DL6    ....................S...
ptnewIII  E...................S...
pt3DL6    ........................
pt3DL4    ...............T...IR...
pt3DL5    ...............T....S...

TRANSMEMBRANE
ptnewII   LLIGTSVAIILFIPLLLFLL
pt2DL6    V.....AV...L.L..F...
ptnewIII  V.....AV...L.L..F...
pt3DL6    V......VK.P.TI..F...
pt3DL4    V.....AV...L.L......
pt3DL5    V......V...L.L..F...

CYTOPLASMIC TAIL
                                                   ******                        ******
ptnewII   HRWCSNKKNAAVMDQEPAGNRTVSREDSDEQDPQEVTYAQLNHCVFTQRKITRPSQRPKTPPTDIIVYTELPNAEPRSKVVSCP
pt2DL6    .....K.........Q.....I.K...............H.D.........NP..R........T...................
ptnewIII  .....K.................N...........................SP..E........TS..................
pt3DL6    ..C......................................D.R................T...TS..I.P...........--
pt3DL4    .....K...............I.K...........................SP..E........TS..................
pt3DL5    .....K...............I.K...............H.D.........NP..R........T...................
```

Figure 2. Multiple sequence alignment of the novel *ptKIRnewII* and *ptKIRnewIII* genes (shown in bold) in the common chimpanzee. Periods (.) indicate identity with *ptKIRnewII* and dashes (-) indicate the absence of amino acids. Conserved cysteine residues in the immunoglobulin domains are highlighted. ITIM motifs are shown by asterisks. Although closely related to *ptKIR2DL6*, which is shown to include the D0 domain encoded by pseudoexon 3, it is not clear whether *ptKIRnewII* and *ptKIRnewIII* belong to the *KIR2D* or *KIR3D* families. The cDNA sequences for *ptKIR2DL6* and *ptKIR3DL4-6* have been included, and are available under accession numbers AF258806 (*ptKIR2DL6*), AY122876 (pseudoexon 3 of *ptKIR2DL6*), AF258800 (*ptKIR3DL4*), AF258801 (*ptKIR3DL5*), and AF258802 (*ptKIR3DL6*).

Downloaded from genome.cshlp.org on August 4, 2022 - Published by Cold Spring Harbor Laboratory Press

Analysis of the rhesus macaque *KIR* region

The 314-kb macaque sequence was assembled from two clones with a 51-kb 100% identical overlap demonstrating that it represents a single haplotype. Five *KIR* genes, five *LILR* genes, and the *FCAR*, *NCR1*, and *PYPAF3* genes were identified in the sequence (Fig. 1). The *LILR* gene cluster appears to be more divergent than previously assumed (Wilson et al. 2000) and will be described elsewhere. Further cDNA analysis, in order to confirm all *KIR* genes and gene structures predicted by genomic sequencing, was carried out using blood from the same rhesus macaque animal (25311) sequenced in this study. Within the *KIR* cluster of genes that spans 86 kb of DNA, one was identified as *mmKIR1D*, three contained three Ig domains and long cytoplasmic tails (*mmKIR3DL*), and one contained two Ig domains (*mmKIR2DL4*). Activating receptors, characterized by short cytoplasmic tails, were absent in both the in silico analysis and in 30 cDNA clones analyzed in the study. Although >100 cDNAs were sequenced, only genes represented by two or more clones were included in the analysis.

The *mmKIR1D* gene encodes a molecule with only one complete Ig D1 domain. Depending on alternative splicing (Hershberger et al. 2001), variants may contain a novel D2 domain resulting from a frameshift, or have deletions encompassing the entire D2 domain but retain the stem, transmembrane, and cytoplasmic domains. Two cDNA clones sequenced for this haplotype correspond to the variant with a novel stretch of 55 amino acids in the Ig D2 domain resulting in early termination of the protein. Alternatively spliced variants, with exon 4 being completely or partially deleted, were also observed, but these were only represented by one cDNA clone.

The *mmKIR2DL4* gene has the highest similarity with a previously identified cDNA sequence named *mmKIR2DL4.2* (AF334645), which was known to have an unidentified 3′-end. The last exon, as identified by genomic and cDNA sequencing in this study, is divergent from the human and chimpanzee orthologs in both sequence composition and length, in particular the terminal 67 amino acids in the cytoplasmic tail. Although the primate *KIR2DL4* genes have a long cytoplasmic tail and ITIM motifs (only one in human and chimpanzee), the presence of a charged arginine residue in the transmembrane domain indicates an activating function (Vilches and Parham 2002; Kikuchi-Maki et al. 2003). Functional testing of NK cytotoxicity using effector cells from the same animal whose *KIR* haplotype was sequenced herein (animal 25311) against a rhesus macaque-derived target cell line with down-regulated cell surface MHC-I expression showed levels of activity within the range observed in 15 unrelated rhesus macaques (H. Andersen, pers. comm.). Thus, *mmKIR2DL4* or some other activating NK cell receptor encoded by a gene outside of the *KIR* gene cluster, such as mmCD94/NKG2C, is capable of sending an activating signal and triggering NK cell cytotoxicity (LaBonte et al. 2001).

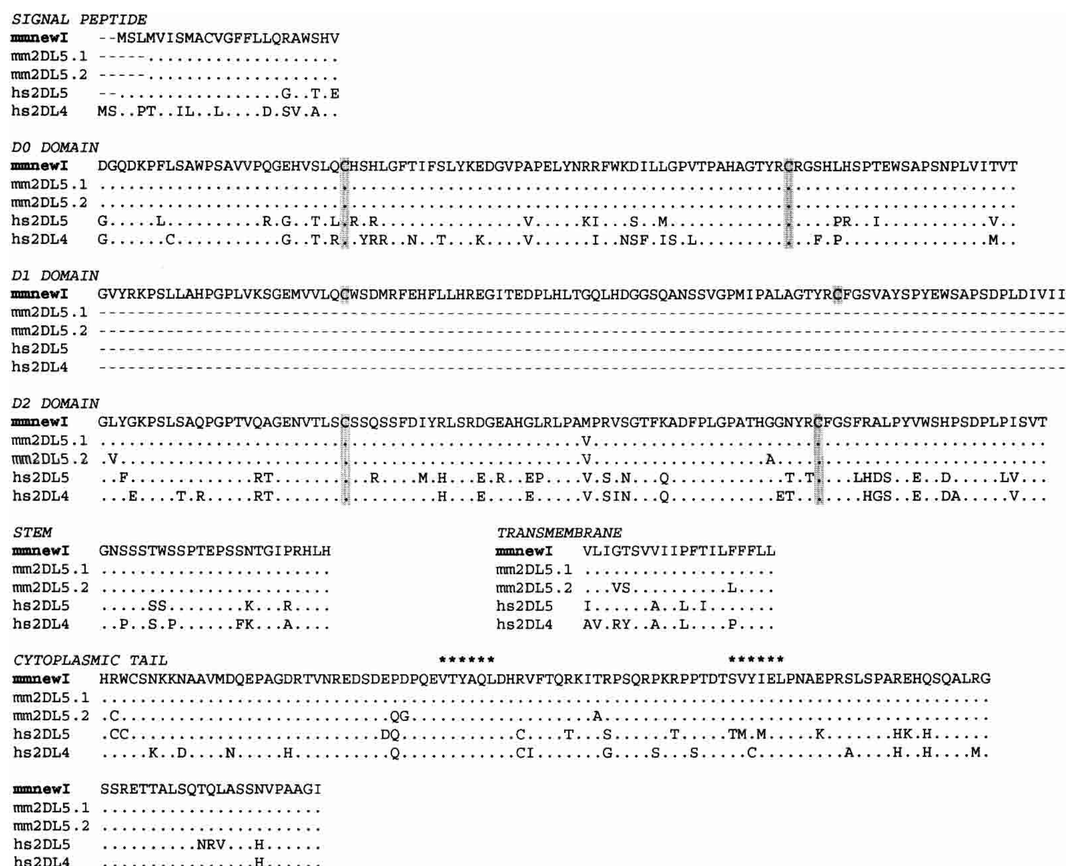

Figure 3. Multiple sequence alignment of the novel *mmKIRnewI* gene in the rhesus macaque (shown in bold). Periods (.) indicate identity with *mmKIRnewI* and dashes (-) indicate the absence of amino acids. Human *KIR2DL4* and *KIR2DL5* genes have the D0+D2 structure. Conserved cysteine residues in the immunoglobulin domains are highlighted. ITIM motifs are shown by asterisks. The cDNA sequences are available under accession numbers AAK26807 (*mmKIR2DL5.1*) and AAK26808 (*mmKIR2DL5.2*).

Downloaded from genome.cshlp.org on August 4, 2022 - Published by Cold Spring Harbor Laboratory Press

Two *KIR* genes containing three Ig domains correspond to previously identified cDNA sequences: *mmKIR3DL1* and *mmKIR3DL10*, and apart from the first six amino acids missing in both cDNAs (Q8MK40 and Q8MK31), the sequences are identical at the protein level. The third novel gene, provisionally called *mmKIRnewI*, is highly related to the cDNA sequence named *mmKIR2DL5.1*, differing by only three nucleotide substitutions in the coding region (Table 1; Fig. 3). While the exon for Ig D1 in *mmKIRnewI* is incorporated into mRNA, it may be that *mmKIR2DL5.1* is the transcript from an allele in which the exon is not present. Alternatively, it could represent a different but closely related gene. There were no *mmKIR2DL5* cDNAs found for the haplotype analyzed here. What is most intriguing is that *mmKIRnewI* appears to represent an evolutionary intermediate between *KIR3D* genes and the genes for human *KIR2DL4* and *KIR2DL5*, which have lost exon 4 corresponding to the D1 domain.

Comparison of the genomic and cDNA data confirmed that, as in humans, different *KIR* haplotypes exist in non-human primates. Among the cDNAs obtained here were full-length sequences for *mmKIR3DL8* and *mmKIR3DH1*, and a novel sequence named *mmKIRnewII*, all pointing to the two *KIR* haplotypes in the animal analyzed being different. The *mmKIRnewII* gene encodes a protein with three Ig domains and is >2% divergent from other *mmKIR3DLs*. Its closest relative, with 90% protein identity, is *mmKIR3DL7*. Two new variants of *mmKIRnewI* and *mmKIR3DL1*, which are 98.5% and 98.2% identical at the protein level with the ones belonging to the sequenced haplotype, respectively, were also identified. The *KIR3DH* molecules resemble *KIR2DL4* in their TM domains with the presence of an arginine, but they lack ITIM motifs in the cytoplasmic tail. It is possible their hybrid nature allows them to act as activating receptors (Hershberger et al. 2001). The identification of the *KIR* sequences reported here, in addition to previously identified cDNAs (Hershberger et al. 2001) from five rhesus macaques, further highlights the diversity of *KIR* genes in non-human primates (Table 2).

Phylogenetic analysis of the primate *KIR* genes

Phylogenetic trees (Fig. 4) were constructed using four structural *KIR* domains, including the three Ig domains and the combined sequence comprising the stem, TM, and cytoplasmic tail. This approach allows for the formation of new *KIR* genes, for example, by recombination, to be examined (Rajalingam et al. 2004). The branching patterns reflect structural differences but lineage similarities between *KIR* members.

Lineage I contains the *KIR2DL4* and *KIR2DL5* genes, which have the Ig D0+D2 configuration in common, and *mmKIRnewI*, which may represent an evolutionary bridge with the D0+D1+D2 structure. *KIR2DL4* is the most conserved gene between human, common chimpanzee, and rhesus macaque. Whether it binds *HLA-G*, a nonclassical MHC class I region gene that is not polymorphic, is not clear (Boyson et al. 2002). Owing to variable phylogenetic clustering, *KIR3DL3* is not a representative member of this lineage (Vilches and Parham 2002) and has been assigned to lineage V, although *KIR3DL3* and *KIR2DL4* are predicted to be descendants from a common ancestor (Martin et al. 2000). The D1 domain of *mmKIRnewI* is the only rhesus macaque sequence to show affinity with *KIR3DL3*. All the remaining macaque genes that contain three Ig domains cluster to form lineage IV.

Lineage II encompasses human *KIR3DL2*, which is known to bind *HLA-A*, and *KIR3DL1*, which binds *HLA-B*. Both *ptKIR3DL1/2* and *KIR3DS1* are in this group. Lineage III consists of a mixture of *KIR* genes that have two (D1+D2) or three Ig (D0+D1+D2) domains. The two novel chimp genes, *ptKIRnewII* and *ptKIRnewIII*, are also found in the cluster. Species-specific evolutionary processes are likely to have driven the diversity within this branch.

Although the similarity between the human and chimpanzee MHC class I regions is high (Anzai et al. 2003) with the conservation of *HLA-A*, *HLA-B*, and *HLA-C* genes, the MHC organization in the rhesus macaque is more varied (Kulski et al. 2002). Of particular interest is the possibility that the recent duplication event leading to the *HLA-C* locus, estimated at 21–28 Mya, only occurred in the ape lineage, and to date has not been mapped in monkeys, such as the rhesus macaque (Piontkivska and Nei 2003). By extrapolation, the receptor to this ligand is predicted to be absent in these primates.

Repeat analysis and evolutionary implications

The DNA sequences encompassing the syntenic regions to the LRC of humans in two closely related primates have made it possible to examine the genomic structure of the *KIR* gene family members. In all three species, they are tightly clustered with relatively short intergenic sequences, share the same direction of transcription, and have similar exon–intron structures, making it likely that they are derived from the same ancestral gene. The insertion of ancient retroelements and *Alu* repeats within the introns of the *KIR* genes accounts for the differences in their lengths. In the common chimpanzee and rhesus macaque, for example, a single *KIR* gene span (including the pseudogenes) varies from 5.4 to 16.1 kb and 8.7 to 14.6 kb, respectively.

Table 2. Diversity of *KIR* genes in rhesus macaque

	1 Ig	2 Ig		3 Ig					
Haplotype	*mmKIR1D*	*mmKIR2DL4*	*mmKIR2DL5*	*mmKIRnewI*	*mmKIRnewII*	*mmKIR3DL1*	*mmKIR3DL8*	*mmKIR3DL10*	*mmKIR3DH1*
25311g	dark gray	dark gray		dark gray *		dark gray *		dark gray	
25311p	light gray	light gray		dark gray *	dark gray	dark gray *	dark gray	light gray	dark gray
173		dark gray				dark gray	dark gray *		
227		dark gray	dark gray			dark gray	dark gray		
223	dark gray	dark gray	dark gray				dark gray		dark gray
577		dark gray							dark gray
576		dark gray				dark gray *		dark gray	

The five *KIR* genes identified by genomic sequencing define the first haplotype (25311g). By deduction, the gene content of the second putative haplotype (25311p) consists of genes or alleles not identified genomically (shown in dark gray) but may also enclose genes in common with the first haplotype (shown in light gray). Asterisks (*) indicate the presence of sequences (alleles) that differ by <2% at the protein level within one individual. Comparisons are made to five rhesus monkeys (173, 227, 223, 577, 576) published by Hershberger and coworkers (2001). Genes are grouped according to the number of immunoglobulin domains they have.

Downloaded from genome.cshlp.org on August 4, 2022 - Published by Cold Spring Harbor Laboratory Press

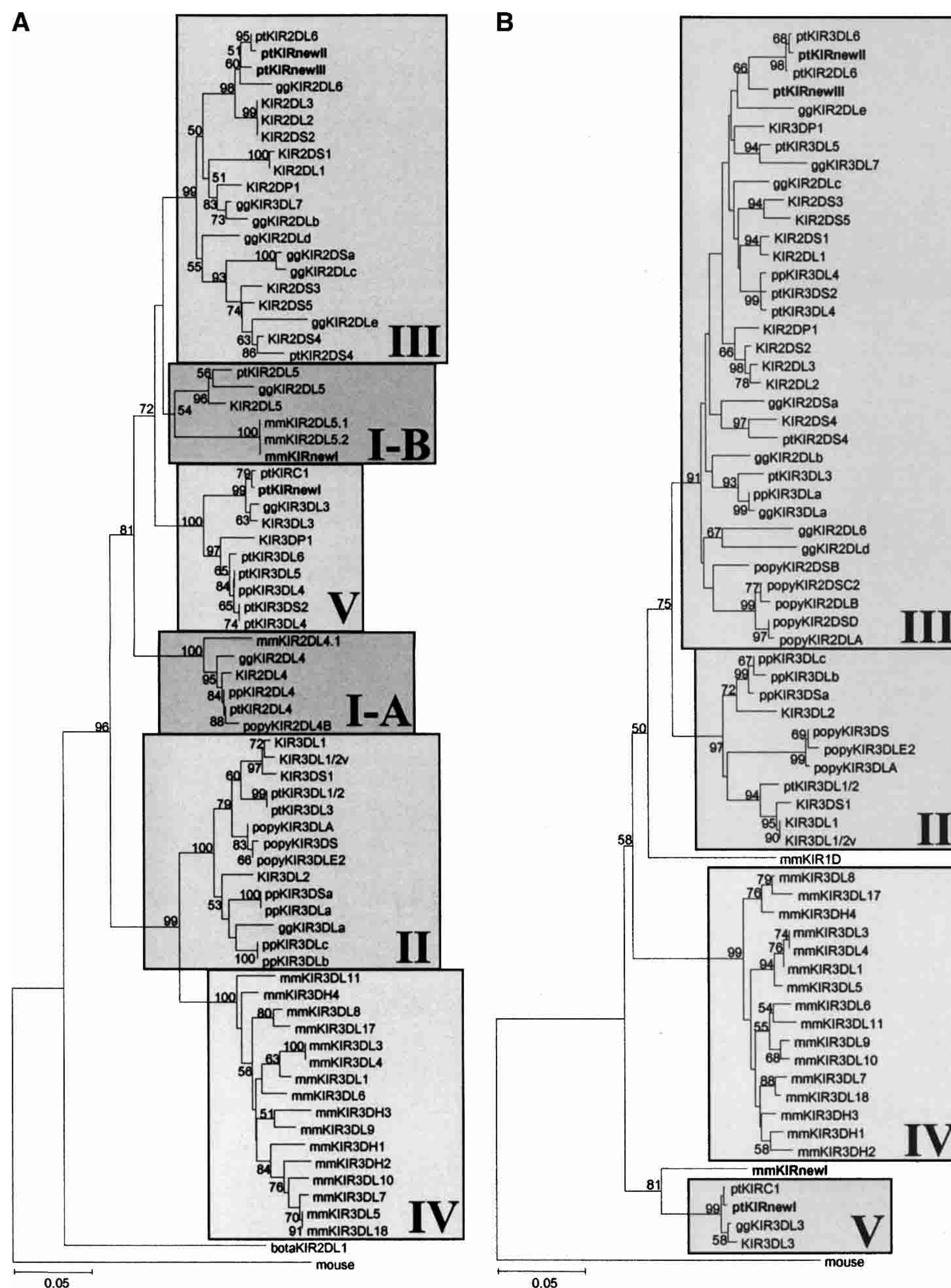

Figure 4. (Continued on next page)

Four major retroelements were found to be present in the chimp and macaque *KIR* intronic regions: *MLT1D*, *MSTB1*, *MER70B*, and *LIMA4* (Fig. 5). These have also been identified within the human *KIR* genes (Martin et al. 2000) and are present in similar locations. Interspersed between and within these retroelements are several *Alu* repeats, which are useful in dating genes according to the presence of differently aged subfamilies. According to Martin et al. (2000) and based on the analysis of the human *KIR* cluster, the ancestral *KIR* gene originated ~60 to 100 Mya when the four retroelements were inserted. Subsequent duplication events, ranging from 30 to 45 Mya, resulted in the diversification of the *KIR* genes, which can be classified according to the type and number of *Alu*s located within the ancestral *LIMA4* element. For example, *KIR2DL4* and *KIR3DL3* have an *AluSq*, indicating a common ancestry, as opposed to *KIR3DL1* and *KIR3DL2*, which have an *AluSp* and two *AluSx* elements and have evolved from a different progenitor.

The chimpanzee *KIR* genes appear to follow a similar pattern of evolution as proposed by Martin et al. (2000). The divergence of *ptKIR2DL4* from *ptKIR3DL1/2*, *ptKIRnewII*, and *ptKIRnewIII* is clearly defined by the distinct *Alu* elements found within the *LIMA4* retroelement. In concordance with the phylogenetic analysis (Fig. 4), members of lineage II, such as *KIR3DL1*, *KIR3DL2*, and *ptKIR3DL1/2*, share an *AluSx* upstream of *LIMA4*,

Downloaded from genome.cshlp.org on August 4, 2022 - Published by Cold Spring Harbor Laboratory Press

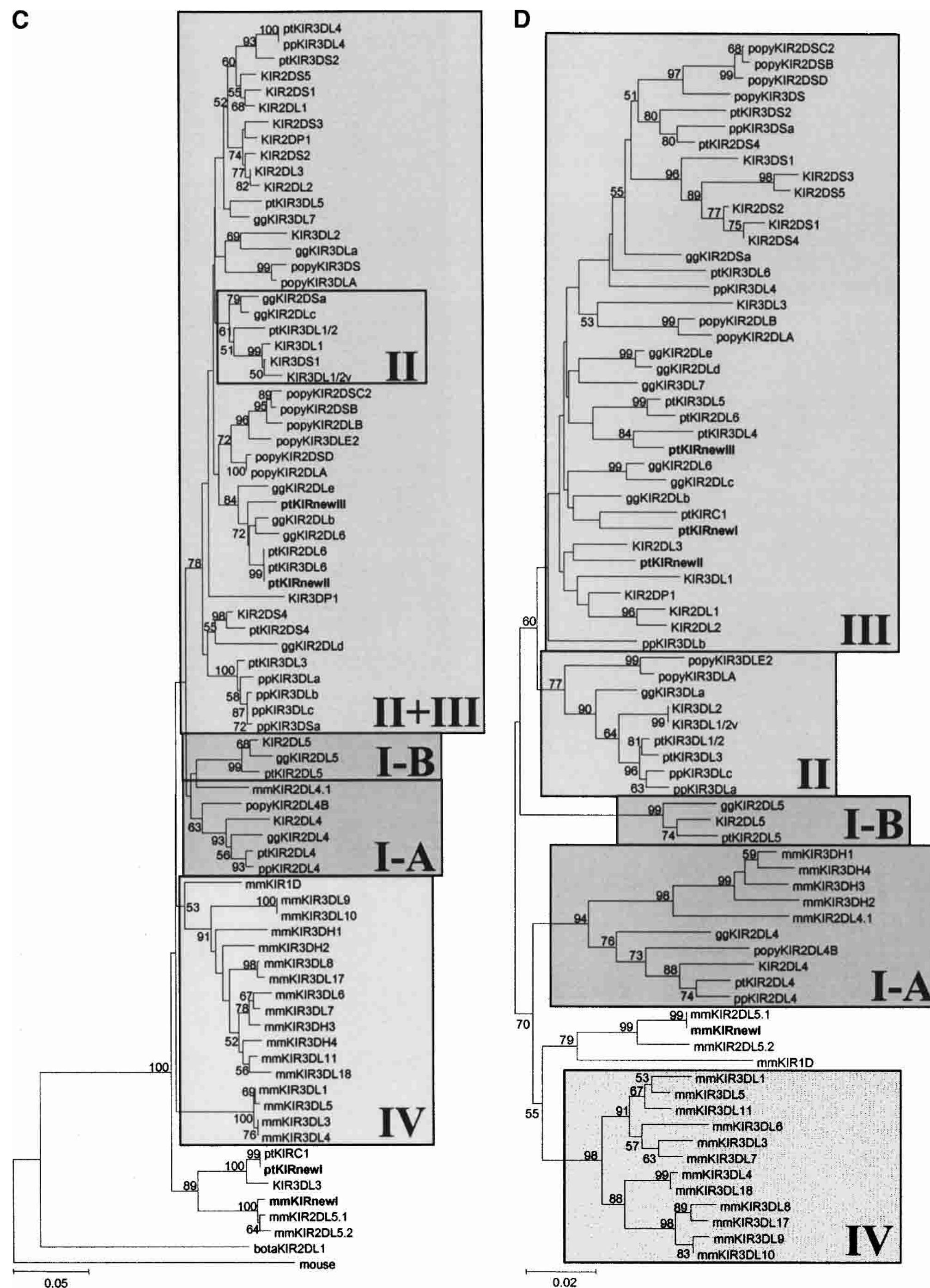

Figure 4. Phylogenetic analysis of the *KIR* genes carried out on a domain-by-domain basis, including Ig D0 (*A*), Ig D1 (*B*), Ig D2 (*C*), and the combined stem, transmembrane, and cytoplasmic tail (*D*). The tree has been constructed using the Neighbor Joining (NJ) method. Novel genes identified in the common chimpanzee and rhesus macaque haplotypes analyzed are shown in bold. Sequences have been grouped into five lineages (I, II, III, IV, V) and are enclosed within boxes. The previously identified *ptKIR3DL3* cDNA does not correspond to the human *KIR3DL3* gene but rather is allelic to *ptKIR3DL1/2*.

while *ptKIRnewII* and *ptKIRnewIII* have an *AluSg* and are part of lineage III. The *ptKIR3DL1/2* gene shares the same repeat elements of intron 3 with human *KIR3DL1* and intron 6 with human *KIR3DL2*, further demonstrating the commonality between these three genes. As with their human counterparts, the *ptKIRnewI* gene, which also contains the *AluSq* element within *LIMA4*, also appears to be related to *ptKIR2DL4*, providing further evidence that primate lineage I and lineage V genes share a common ancestor. Notable differences between intronic regions of *KIR2DL4* and *KIR2DL5* are the presence or absence of the *AluSq* repeat in the *LIMA4* element in intron 4, as well as two types of *Alu* repeats (*AluSx* and *AluSq/x*) in the latter gene.

Downloaded from genome.cshlp.org on August 4, 2022 - Published by Cold Spring Harbor Laboratory Press

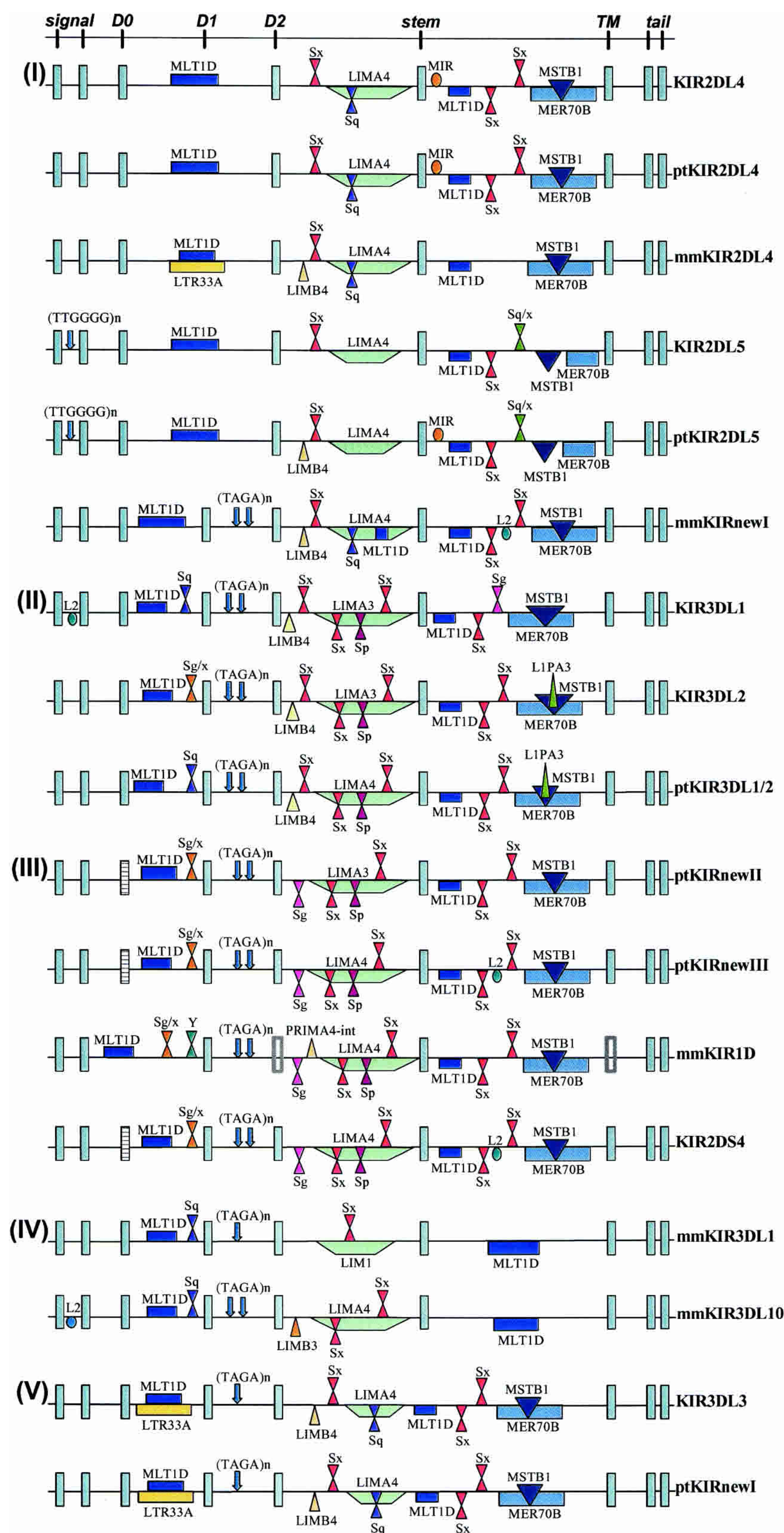

Figure 5. Repeat analysis of the *KIR* intronic regions showing the presence of ancient retroelements and *Alu* repeats (not to scale). Genes have been grouped according to their lineages (I, II, III, IV, V). Distinct domains are labeled as follows: signal peptide, D0, D1, and D2 for the Ig domains, stem, transmembrane domain, and cytoplasmic tail. Lineage I genes have a D0+D2 structure. Although the three Ig domains are shown for *ptKIRnewII* and *ptKIRnewIII*, it is likely that they contain a pseudoexon 3, indicated by the striped boxes, and express the 2D structure. Splice variants of *mmKIR1D* may be missing the D2 or TM domains, as indicated by the open boxes.

Repeat elements found in the intronic regions of the rhesus macaque genes are less conserved compared to the human and chimpanzee homologs, although similarities within respective lineages are apparent. Apart from *mmKIR2DL4*, which has the characteristic *AluSq* insertion in *LIMA4* found in all primate *KIR2DL4* genes, this was also observed in *mmKIRnewI*, which may indicate a unique shared common ancestor, and further supports its clustering with other lineage I genes. Both *mmKIR3DL1* and *mmKIR3DL10* do not have the *MSTB1* or *MER70B* retroelements in intron 6 that are present in all the other *KIR* genes analyzed here. Correspondingly, their protein sequences clustered phylogenetically within lineage IV, which comprises only rhesus macaque *KIR* genes (Fig. 4). The repeat elements in *mmKIR1D* are highly similar to those present in lineage II and III genes, including human *KIR2DS4*. Recently, a novel allele of *KIR2DS4* has been identified that resembles one of the splice variants of *mmKIR1D*, characterized by a 22-bp deletion that disrupts the Ig D2 domain and results in the absence of the cytoplasmic tail due to premature termination of the protein (Hsu et al. 2002). Although *mmKIR1D* found in this haplotype does not encode this particular variant, analysis of the intronic regions does reveal a shared ancestry with the human *KIR2DS4* allele. The youngest type of an *Alu* repeat, *AluY*, has only been observed in *mmKIR1D*.

Minisatellite sequences, originally identified in all human *KIRs* (Trowsdale et al. 2001), were also found in the first intron of all the *KIR* genes analyzed here, apart from *KIR2DL4*, in both non-human primates (Table 3). It is known that G+C-rich minisatellites are associated with recombination rates and/or variation of genes (Jeffreys et al. 2000; Boan et al. 2002).

Evolution of the *KIR* region in primates

Sequence comparisons and phylogenetic analyses have shown that the *KIR* gene fam-

Downloaded from genome.cshlp.org on August 4, 2022 - Published by Cold Spring Harbor Laboratory Press

Table 3. Identification of minisatellites in the first intron of the non-human *KIR* genes

Gene	Minisatellite repeat motif	Number
ptKIRnewI	CTGGAGTGGAGATATGGGC	30 (33)
ptKIRnewII	CCTGGAGTGGAGATATGCA	43
ptKIR2DL5	AGATATGGGCCTGGAGTGG	25 (23)
ptKIRnewIII	TGGAGATATGGGCCTGGAG	49
ptKIR3DP1	TGGAGATATGGGCCTGGAG	24
ptKIR2DL4	None	0 (0)
ptKIR3DL1/2	AGATCTGGGCCTGGAGTGG	16 (42/28)
mmKIRnewI	CTGGAGTGGAGATATGGGC	30
mmKIR1D	CTGGAGTGGAGATATGGGC	40
mmKIR2DL4	None	0 (0)
mmKIR3DL10	CTGGAGTGGAGATATGGGC	72
mmKIR3DL1	CTGGAGTGGAGATATGGGC	84 (42)

The numbers of repeats in the equivalent human *KIR* genes are shown in parentheses. The minisatellites consist of multiple copies of an imperfect 19-bp repeat.

ily members have diversified rapidly among primates, and only certain *KIR* genes remain orthologous among humans, apes, and monkeys. The presence of ancient retroelements indicates that these were derived from a common ancestor that was in existence ~60 to 100 Mya. The additional insertion of *Alu* elements are indicative of recent duplication events leading to the diversity in the *KIR* family.

The expansion of the *KIR* system in primates, as opposed to the diverse family of *Ly49* members in rodents, might have evolved as a functional means of replacing the single *LY49like* (*LY49L*) gene present in non-rodent mammals (McQueen et al. 2002). Point mutations have inactivated the *LY49L* gene in humans, chimpanzee, and gorilla (higher primates), although it appears to be functional in other species, such as cow, baboon, and orangutan. The *Ly49* gene also exists as a single copy in domestic cat, dog, and pig, as opposed to mouse, rat, and horse (Takahashi et al. 2004), which have multiple copies. Phylogenetic analysis suggests the *Ly49* genes are evolving more rapidly within rodents as compared to their non-rodent mammalian counterparts (Gagnier et al. 2003), mainly as a result of a series of large duplication events involving units of one or more genes (Wilhelm et al. 2002). To date, no mammals have been identified that have multiple *Ly49* and multiple *KIR* genes. Both systems ultimately provide inhibitory and activating receptors, expressed in a diverse NK cell repertoire, for MHC class I molecules.

Functional analysis of NK cell activity mediated through KIR in non-human primates is limited to a recent study showing that chimpanzee and human NK cells exhibit identical receptor specificities for HLA-C through nonorthologous KIR molecules (Khakoo et al. 2000). In the rhesus macaque, virtually all of the functional studies of NK activity performed to date have used the human target cell line K562, despite the phylogenetic divergence between the species and the fact that the rhesus monkey does not appear to have an *HLA-C* locus. However, recently developed rhesus NK target cell lines with down-regulated cell surface MHC-I expression (H. Andersen, pers. comm.) should help clarify the underlying molecular interactions between macaque KIR and their cognate ligands that regulate NK activity in this species.

The chimp and macaque genomic regions analyzed here represent one haplotype in each species, and do not encompass the full complement of *KIR* genes present in the genome or population. Interestingly, no activating *KIR* genes (characterized by short cytoplasmic tails) were identified, suggesting that they either reside in separate loci or their presence might vary within the population. As in humans, different haplotypes vary in gene content in non-human primates. Four novel *KIR* genes, preliminarily named *ptKIRnewI*, *ptKIRnewII*, *ptKIRnewIII*, and *mmKIRnewI*, have also been identified genomically in this study, and it is possible that more variants exist as shown by the presence of novel cDNAs, such as *mmKIRnewII*, in the rhesus macaque. The shortest *KIR* haplotype in the pygmy chimpanzee, a close relative of the common chimpanzee, only encompasses three genes, *KIR3DL3*, *KIR2DL4*, and *KIR3DL* (Rajalingam et al. 2001). These genes represent either orthologs or paralogs to one of the four framework genes present in all human haplotypes, and are also present in the analyzed common chimpanzee genomic region. Only *KIR2DL4* is held in common with rhesus macaque.

Methods

Contig construction and sequencing

The RPCI-43 chimpanzee and CHORI-250 rhesus macaque BAC libraries (Children's Hospital Oakland Research Institute, Oakland, CA) were hybridized with a human *KIR2DL4* probe (nucleotides 57132–57418, GenBank accession number AC011501) to identify both inhibitory and activating *KIR*. The libraries were screened, and positive clones were rescreened by hybridization with a *KIR2DL4* probe using Southern and dot blotting. Then 12 chimpanzee and eight macaque BAC clones were selected for further hybridization with human probes for *LILRA2* (nucleotides 1340–1686, GenBank accession number U82275) and *NCR1* (nucleotides 771–1138, GenBank accession number AJ001383) in order to identify BAC clones that contained the centromeric and/or telomeric ends of the *KIR* complex, respectively. The BAC clones were mapped into contigs using restriction-digest fingerprinting (Gregory et al. 1997). One chimp clone (RPCI-43-61P22) and two macaque clones (CHORI-250-178N19 and CHORI-250-242L13), covering the respective *KIR* complexes, were selected for sequencing. The virtual restriction maps from the clones sequenced were compared to the restriction-enzyme fingerprints in order to confirm correct assembly. These clones have flanking framework *non-KIR* genes, indicating that the complete *KIR* haplotypes have been sequenced for both species.

The BAC DNA was randomly subcloned (Bankier et al. 1987) into pUC plasmids and sequenced from both ends using the dideoxy chain termination method (Sanger et al. 1977) with different versions of big dye terminator chemistry (Rosenblum et al. 1997). The resulting sequencing reactions were analyzed on various models of ABI sequencing machines, and the generated data were processed by a suite of in-house programs (http://www.sanger.ac.uk/Software/) prior to assembly with the PHRED (Ewing et al. 1998) and PHRAP (http://www.phrap.org) algorithms. The GAP4 program (Bonfield et al. 1995) was used to view and edit the resulting sequence contigs during the finishing process. The finished sequences were submitted to the EMBL/GenBank/DDBJ databases under BX842589 for the common chimpanzee, and BX842590 and BX842591 for rhesus macaque.

Gene identification and genomic characterization

The DNA sequences were analyzed using NIX, a gene identification program (http://www.hgmp.mrc.ac.uk/NIX). The homologous *KIR* genes in the common chimpanzee and rhesus macaque were classified into distinct groups based on their genomic organization, that is, whether they encoded proteins with two or

Downloaded from genome.cshlp.org on August 4, 2022 - Published by Cold Spring Harbor Laboratory Press

three Ig domains, and either long or short cytoplasmic tails. Furthermore, following the nomenclature used to describe *KIR* homologs, the prefixes "pt" for *Pan troglodytes* and "mm" for *Macaca mulatta* have been used in the annotation. The naming of *KIR* alleles has recently been standardized (Marsh et al. 2003) and is followed here. Pairwise alignments were carried out in EMBOSS (Rice et al. 2000) with a gap opening penalty of 10 and a gap extension penalty of 0.5. Intronic sequences were analyzed for retroelements and tandem repeats using the RepeatMasker program (http://ftp.genome.washington.edu/RM/RepeatMasker.html; A.F. Smit and P. Green, unpubl.) and a tandem repeat finder program (Benson 1999).

cDNA analysis

Peripheral blood mononuclear cells (PBMC) were isolated from whole blood of rhesus monkey Mmu-25311 provided by M.B. McChesney (Virology and Immunology Unit, California National Primate Research Center, UC Davis, CA). Total RNA was isolated from the PBMC using Trizol (Invitrogen) and transcribed to cDNA using Superscript II (Invitrogen). Macaque *KIR* cDNAs were amplified in PCR reactions using eight different pairs of primers. Primers were designed based on alignment of GenBank sequences as well as predicted cDNA structures from the macaque *KIR* haplotype described in this paper. PCR primer pair sequences were as follows: (1) CATGTYGCTCAYGGTCGTC and CTGGGCTGGAGACAACGA; (2) ATGGTCGTCAGCGTGGYG and TTGTGTCCCTRGARGACCCC; (3) AGCACCATGTCGCTCATGGTCA and GTCGCGCCTTCAGATTCCTG; (4) ATGTCGCTCATGGTCGTTAGCG and CCCTAAGATGCAGACTCACAG; (5) CATGTCGCCCACGGTCGTCAT and CTAAGCAAAGGAGTGCGTTTTC; (6) AGCACCATGTCGCTCATGGTCC and TTGTCTCCCTAGAAGACCCCT; (7) AGCACCATGTCGCTCATGGTCG and ACGGTGGTGCTCATGGATAGA; (8) AGCACCATGTCGCTCATGGTCA and AGGCCTGACTCTGGTGCTCAC. Owing to the high level of sequence similarity, multiple genes were amplified in all reactions, except that for *KIR3DH1* and *KIR2DL4* (primer pairs 1 and 5, respectively). PCR products were cloned into pCR2.1-TOPO vector (Invitrogen), and 10–20 clones were isolated and sequenced for each PCR reaction. Although >100 cDNA clones were sequenced, only sequences represented by two or more clones were submitted to GenBank and described in this paper.

Phylogenetic analysis

KIR full-length nucleic acid sequences were aligned using CLUSTALX (Higgins et al. 1992), and then corrected manually. Neighbor Joining (Saitou and Nei 1987) phylogenetic trees were reconstructed for each domain using MEGA version 2.1 (Kumar et al. 2001). The *p*-distance was used with pairwise deletion and 500 bootstrap replicates. All the trees were rooted at the midpoint and bootstrap proportion values <50 were removed.

Acknowledgments

We thank Jie Cheng, Audrey Fraser, Karen Halls, and Matt Jones for technical assistance and the Wellcome Trust for financial support. Part of this work was supported with federal funds from the National Cancer Institute, National Institutes of Health under contract no. NO1-CO-12400. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organization imply endorsement by the U.S. Government.

References

Anzai, T., Shiina, T., Kimura, N., Yanagiya, K., Kohara, S., Shigenari, A., Yamagata, T., Kulski, J.K., Naruse, T.K., Fujimori, Y., et al. 2003. Comparative sequencing of human and chimpanzee MHC class I regions unveils insertions/deletions as the major path to genomic divergence. *Proc. Natl. Acad. Sci.* **100:** 7708–7713.

Bankier, A.T., Weston, K.M., and Barrell, B.G. 1987. Random cloning and sequencing by the M13/dideoxynucleotide chain termination method. *Methods Enzymol.* **155:** 51–93.

Batzer, M.A. and Deininger, P.L. 2002. *Alu* repeats and human genomic diversity. *Nat. Rev. Genet.* **3:** 370–379.

Benson, G. 1999. Tandem repeats finder: A program to analyze DNA sequences. *Nucleic Acids Res.* **27:** 573–580.

Boan, F., Rodriguez, J.M., Mourino, S., Blanco, M.G., Vinas, A., Sanchez, L., and Gomez-Marquez, J. 2002. Recombination analysis of the human minisatellite MsH42 suggests the existence of two distinct pathways for initiation and resolution of recombination at MsH42 in rat testes nuclear extracts. *Biochemistry* **41:** 2166–2176.

Bonfield, J.K., Smith, K., and Staden, R. 1995. A new DNA sequence assembly program. *Nucleic Acids Res.* **23:** 4992–4999.

Boyson, J.E., Erskine, R., Whitman, M.C., Chiu, M., Lau, J.M., Koopman, L.A., Valter, M.M., Angelisova, P., Horejsi, V., and Strominger, J.L. 2002. Disulfide bond-mediated dimerization of HLA-G on the cell surface. *Proc. Natl. Acad. Sci.* **99:** 16180–16185.

Carrington, M. and Norman, P. 2003. *The KIR gene cluster.* Available from http://ncbi.nlm.nih.gov/entrez/query.fcgi?db=Books, Bethesda, MD.

Ewing, B., Hillier, L., Wendl, M.C., and Green, P. 1998. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res.* **8:** 175–185.

Gagnier, L., Wilhelm, B.T., and Mager, D.L. 2003. Ly49 genes in non-rodent mammals. *Immunogenetics* **55:** 109–115.

Gaudieri, S., Kulski, J.K., Dawkins, R.L., and Gojobori, T. 1999. Different evolutionary histories in two subgenomic regions of the major histocompatibility complex. *Genome Res.* **9:** 541–549.

Gregory, S.G., Howell, G.R., and Bentley, D.R. 1997. Genome mapping by fluorescent fingerprinting. *Genome Res.* **7:** 1162–1168.

Grendell, R.L., Hughes, A.L., and Golos, T.G. 2001. Cloning of rhesus monkey killer-cell Ig-like receptors (KIRs) from early pregnancy decidua. *Tissue Antigens* **58:** 329–334.

Guethlein, L.A., Flodin, L.R., Adams, E.J., and Parham, P. 2002. NK cell receptors of the orangutan (*Pongo pygmaeus*): A pivotal species for tracking the coevolution of killer cell Ig-like receptors with MHC-C. *J. Immunol.* **169:** 220–229.

Hershberger, K.L., Shyam, R., Miura, A., and Letvin, N.L. 2001. Diversity of the killer cell Ig-like receptors of rhesus monkeys. *J. Immunol.* **166:** 4380–4390.

Higgins, D.G., Bleasby, A.J., and Fuchs, R. 1992. CLUSTAL V: Improved software for multiple sequence alignment. *Comput. Appl. Biosci.* **8:** 189–191.

Horai, S. 1995. Evolution and the origins of man: Clues from complete sequences of hominoid mitochondrial DNA. *Southeast Asian J. Trop. Med. Public Health* **26:** 146–154.

Hsu, K.C., Chida, S., Geraghty, D.E., and Dupont, B. 2002. The killer cell immunoglobulin-like receptor (KIR) genomic region: Gene-order, haplotypes and allelic polymorphism. *Immunol. Rev.* **190:** 40–52.

Jeffreys, A.J., Ritchie, A., and Neumann, R. 2000. High resolution analysis of haplotype diversity and meiotic crossover in the human TAP2 recombination hotspot. *Hum. Mol. Genet.* **9:** 725–733.

Khakoo, S.I., Rajalingam, R., Shum, B.P., Weidenbach, K., Flodin, L., Muir, D.G., Canavez, F., Cooper, S.L., Valiante, N.M., Lanier, L.L., et al. 2000. Rapid evolution of NK cell receptor systems demonstrated by comparison of chimpanzees and humans. *Immunity* **12:** 687–698.

Kikuchi-Maki, A., Yusa, S., Catina, T.L., and Campbell, K.S. 2003. KIR2DL4 is an IL-2-regulated NK cell receptor that exhibits limited expression in humans but triggers strong IFN-γ production. *J. Immunol.* **171:** 3415–3425.

Kulski, J.K., Shiina, T., Anzai, T., Kohara, S., and Inoko, H. 2002. Comparative genomic analysis of the MHC: The evolution of class I duplication blocks, diversity and complexity from shark to man. *Immunol. Rev.* **190:** 95–122.

Kumar, S. and Hedges, S.B. 1998. A molecular timescale for vertebrate evolution. *Nature* **392:** 917–920.

Kumar, S., Tamura, K., Jakobsen, I.B., and Nei, M. 2001. MEGA2: Molecular evolutionary genetics analysis software. *Bioinformatics* **17:** 1244–1245.

LaBonte, M.L., Hershberger, K.L., Korber, B., and Letvin, N.L. 2001. The KIR and CD94/NKG2 families of molecules in the rhesus monkey. *Immunol. Rev.* **183:** 25–40.

Marsh, S.G., Parham, P., Dupont, B., Geraghty, D.E., Trowsdale, J.,

Downloaded from genome.cshlp.org on August 4, 2022 - Published by Cold Spring Harbor Laboratory Press

Middleton, D., Vilches, C., Carrington, M., Witt, C., Guethlein, L.A., et al. 2003. Killer-cell immunoglobulin-like receptor (KIR) nomenclature report, 2002. *Tissue Antigens* **62:** 79–86.
Martin, A.M., Freitas, E.M., Witt, C.S., and Christiansen, F.T. 2000. The genomic organization and evolution of the natural killer immunoglobulin-like receptor (KIR) gene cluster. *Immunogenetics* **51:** 268–280.
Martin, M.P., Bashirova, A., Traherne, J., Trowsdale, J., and Carrington, M. 2003. Cutting edge: Expansion of the KIR locus by unequal crossing over. *J. Immunol.* **171:** 2192–2195.
McQueen, K.L., Wilhelm, B.T., Harden, K.D., and Mager, D.L. 2002. Evolution of NK receptors: A single Ly49 and multiple KIR genes in the cow. *Eur. J. Immunol.* **32:** 810–817.
Natarajan, K., Dimasi, N., Wang, J., Mariuzza, R.A., and Margulies, D.H. 2002. Structure and function of natural killer cell receptors: Multiple molecular solutions to self, nonself discrimination. *Annu. Rev. Immunol.* **20:** 853–885.
Piontkivska, H. and Nei, M. 2003. Birth-and-death evolution in primate MHC Class I genes: Divergence time estimates. *Mol. Biol. Evol.* **20:** 601–609.
Rajalingam, R., Hong, M., Adams, E.J., Shum, B.P., Guethlein, L.A., and Parham, P. 2001. Short KIR haplotypes in pygmy chimpanzee (*Bonobo*) resemble the conserved framework of diverse human KIR haplotypes. *J. Exp. Med.* **193:** 135–146.
Rajalingam, R., Parham, P., and Abi-Rached, L. 2004. Domain shuffling has been the main mechanism forming new hominoid killer cell Ig-like receptors. *J. Immunol.* **172:** 356–369.
Rice, P., Longden, I., and Bleasby, A. 2000. EMBOSS: The European Molecular Biology Open Software suite. *Trends Genet.* **16:** 276–277.
Rosenblum, B.B., Lee, L.G., Spurgeon, S.L., Khan, S.H., Menchen, S.M., Heiner, C.R., and Chen, S.M. 1997. New dye-labeled terminators for improved DNA sequencing patterns. *Nucleic Acids Res.* **25:** 4500–4504.
Saitou, N. and Nei, M. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4:** 406–425.
Sanger, F., Nicklen, S., and Coulson, A.R. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci.* **74:** 5463–5467.
Shiina, T., Tamiya, G., Oka, A., Takishima, N., Yamagata, T., Kikkawa, E., Iwata, K., Tomizawa, M., Okuaki, N., Kuwano, Y., et al. 1999. Molecular dynamics of MHC genesis unraveled by sequence analysis of the 1,796,938-bp HLA class I region. *Proc. Natl. Acad. Sci.* **96:** 13282–13287.
Springer, M.S., Murphy, W.J., Eizirik, E., and O'Brien, S.J. 2003. Placental mammal diversification and the Cretaceous–Tertiary boundary. *Proc. Natl. Acad. Sci.* **100:** 1056–1061.
Takahashi, T., Yawata, M., Raudsepp, T., Lear, T.L., Chowdhary, B.P., Antczak, D.F., and Kasahara, M. 2004. Natural killer cell receptors in the horse: Evidence for the existence of multiple transcribed LY49 genes. *Eur. J. Immunol.* **34:** 773–784.
Trowsdale, J. 2001. Genetic and functional relationships between MHC and NK receptor genes. *Immunity* **15:** 363–374.
Trowsdale, J., Barten, R., Haude, A., Stewart, C.A., Beck, S., and Wilson, M.J. 2001. The genomic context of natural killer receptor extended gene families. *Immunol. Rev.* **181:** 20–38.
Vilches, C. and Parham, P. 2002. KIR: Diverse, rapidly evolving receptors of innate and adaptive immunity. *Annu. Rev. Immunol.* **20:** 217–251.
Vilches, C., Pando, M.J., and Parham, P. 2000. Genes encoding human killer-cell Ig-like receptors with D1 and D2 extracellular domains all contain untranslated pseudoexons encoding a third Ig-like domain. *Immunogenetics* **51:** 639–646.
Welch, A.Y., Kasahara, M., and Spain, L.M. 2003. Identification of the mouse killer immunoglobulin-like receptor-like (Kirl) gene family mapping to chromosome X. *Immunogenetics* **54:** 782–790.
Wende, H., Volz, A., and Ziegler, A. 2000. Extensive gene duplications and a large inversion characterize the human leukocyte receptor cluster. *Immunogenetics* **51:** 703–713.
Westgaard, I.H., Berg, S.F., Orstavik, S., Fossum, S., and Dissen, E. 1998. Identification of a human member of the Ly-49 multigene family. *Eur. J. Immunol.* **28:** 1839–1846.
Wilhelm, B.T., Gagnier, L., and Mager, D.L. 2002. Sequence analysis of the ly49 cluster in C57BL/6 mice: A rapidly evolving multigene family in the immune system. *Genomics* **80:** 646–661.
Wilson, M.J., Torkar, M., Haude, A., Milne, S., Jones, T., Sheer, D., Beck, S., and Trowsdale, J. 2000. Plasticity in the organization and sequences of human KIR/ILT gene families. *Proc. Natl. Acad. Sci.* **97:** 4778–4783.

Web site references

http://ftp.genome.washington.edu/RM/RepeatMasker.html; RepeatMasker.
http://www.hgmp.mrc.ac.uk/NIX; NIX server at the MRC Rosalind Franklin Centre for Genomics Research (RFCGR).
http://www.phrap.org; the Phred/Phrap/Consed System.
http://www.sanger.ac.uk/Software/; the Sanger Institute Production Sequencing Software.

Received January 23, 2004; accepted in revised form September 14, 2004.

Downloaded from genome.cshlp.org on August 4, 2022 - Published by Cold Spring Harbor Laboratory Press

Single haplotype analysis demonstrates rapid evolution of the killer immunoglobulin-like receptor (*KIR*) loci in primates

Jennifer G. Sambrook, Arman Bashirova, Sophie Palmer, et al.

Genome Res. 2005 15: 25-35
Access the most recent version at doi:10.1101/gr.2381205

References This article cites 46 articles, 15 of which can be accessed free at:
http://genome.cshlp.org/content/15/1/25.full.html#ref-list-1

License

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here.**

Affordable, Accurate Sequencing.

To subscribe to *Genome Research* go to:
https://genome.cshlp.org/subscriptions

Cold Spring Harbor Laboratory Press