Killer Ig-Like Receptor Haplotype Analysis by Gene Content: Evidence for Genomic Diversity with a Minimum of Six Basic Framework Haplotypes, Each with Multiple Subsets

Katharine C. Hsu, Xiao-Rong Liu, Annamalai Selvakumar, Eric Mickelson, Richard J. O'Reilly and Bo Dupont

J Immunol 2002; 169:5118-5129; ;
doi: 10.4049/jimmunol.169.9.5118
http://www.jimmunol.org/content/169/9/5118

References
This article cites 48 articles, 17 of which you can access for free at:
http://www.jimmunol.org/content/169/9/5118.full#ref-list-1

Subscription
Information about subscribing to The Journal of Immunology is online at:
http://jimmunol.org/subscription

Permissions
Submit copyright permission requests at:
http://www.aai.org/About/Publications/JI/copyright.html

Email Alerts
Receive free email-alerts when new articles cite this article. Sign up at:
http://jimmunol.org/alerts
Killer Ig-Like Receptor Haplotype Analysis by Gene Content: Evidence for Genomic Diversity with a Minimum of Six Basic Framework Haplotypes, Each with Multiple Subsets

Katharine C. Hsu,*† Xiao-Rong Liu,* Annamalai Selvakumar,* Eric Mickelson,‡ Richard J. O’Reilly,*† and Bo Dupont2*†

Killer Ig-like receptor (KIR) genes constitute a multigene family whose genomic diversity is achieved through differences in gene content and allelic polymorphism. KIR haplotypes containing a single activating KIR gene (A-haplotypes), and KIR haplotypes with multiple activating receptor genes (B-haplotypes) have been described. We report the evaluation of KIR gene content in extended families, sibling pairs, and an unrelated Caucasian panel through identification of the presence or absence of 14 KIR genes and 2 pseudogenes. Haplotype definition included subtyping for the expressed and nonexpressed KIR2DL5 variants, for two alleles of pseudogene 3DP1, and for two alleles of 2DS4, including a novel 2DS4 allele, KIR1D. KIR1D appears functionally homologous to the rhesus monkey KIR1D and likely arose as a consequence of a 22 nucleotide deletion in the coding sequence of 2DS4, leading to disruption of Ig-domain 2D and a premature termination codon following the first amino acid in the putative transmembrane domain. Our investigations identified 11 haplotypes within 12 families. From 49 sibling pairs and 17 consanguineous DNA samples, an additional 12 haplotypes were predicted. Our studies support a model for KIR haplotype diversity based on six basic gene compositions. We suggest that the centromeric half of the KIR genomic region is comprised of three major combinations, while the telomeric half can assume a short form with either 2DS4 or KIR1D or a long form with multiple combinations of several stimulatory KIR genes. Additional rare haplotypes can be identified, and may have arisen by gene duplication, intergenic recombination, or deletions. The Journal of Immunology, 2002, 169: 5118–5129.

Bridging innate and adaptive immunity, the NK cell is an important effector lymphocyte that participates in the early immune response to pathogens through the production of cytokines and chemokines (1). Furthermore, the NK cell has also been found to mediate cytolytic activity against virally infected cells and malignant cells (1, 2). Following the principle of the “missing self” hypothesis, NK recognition of “self” MHC Ags on putative target cells leads to inhibition of effector functions. Accordingly, target cell loss of self-MHC class I expression releases NK cell effector functions by removing the MHC-mediated inhibition (3). Regulation of NK cell function is accomplished through a diverse complement of receptors mediating, activating, and inhibiting signals in response to ligand interactions.

In humans, receptors that signal activation include the NK cytotoxicity receptors (4), whose ligands remain unclear, and NKG2D, which has been shown to recognize MHC class I chain-related proteins A and B and UL16-binding proteins (5). Inhibitory receptors include the heterodimer molecules CD94:NKG2A which recognize complexes of HLA-E and peptides encoded from the receptors include the heterodimer molecules CD94:NKG2A which

\*Immunology Program, Sloan-Kettering Institute for Cancer Research and \*Allogeneic Bone Marrow Transplantation Service, Department of Medicine, Memorial Sloan-Kettering Cancer Center, New York, NY 10021; and \*Human Immunogenetics Program, Clinical Research Division, Fred Hutchinson Cancer Research Center, Seattle, WA 98109

Received for publication May 30, 2002. Accepted for publication August 30, 2002.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

1 This work was supported by Grants CA 08748, CA23766, and U24 AI49213 from the National Institutes of Health.

2 Address correspondence and reprint requests to Dr. Bo Dupont, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, NY 10021. E-mail address: bdupont@ski.mskcc.org

Abbreviations used in this paper: KIR, killer Ig-like receptor; ILT, Ig-like transcript; BLCL, B-lymphoblastoid cell line; SSP, sequence-specific primer; UTR, untranslated region; LD, linkage disequilibrium.
The ligand specificity for HLA-A, -B, -C, and -G has been demonstrated for certain KIR, while specificity for other KIR remains unknown (28). The KIR region exhibits an extensive degree of diversity, which it achieves through a combination of variable gene content and polymorphism (29–38). Population studies have generally used KIR gene typing methods on genomic DNA to determine the presence or absence of each KIR gene and have demonstrated that between individuals, KIR gene content can vary widely. Recently, these studies were extended to include evaluation of allelic differences at polymorphic sites to document an additional dimension of KIR diversity achieved through polymorphism (36). Indeed, estimates of the extent of KIR genotype diversity within the population suggest that far <0.24% of unrelated individuals can expect to have identical genotypes (36).

However, underlying the diversity of the KIR genomic region are patterns that appear conserved within the population. The chromosomal arrangement of KIR genes, for instance, maintains a certain regularity as exemplified by the regular spacing of KIR genes ~2.4 kb from each other and the presence of the framework genes 3DL3 and 3DL2 at either terminus of the region and 2DL4 in the middle (24). There is suggestion of conservation of haplotypes as well, with the early description of two main haplotype groups A and B found within the population (29–30). Estimated to be as frequent as 47–59% within the Caucasian population, haplotype A comprises a common complement of KIR genes 3DL3, -2DL3, -2DL1, -2DL4, -3DL1, -2DS4, and -3DL2. In contrast, haplotype B has been defined as a more varied haplotype group that encompasses genotypes containing more activating receptor genes, including KIR2DS1, -2DS2, -2DS3, and -2DS5. The definition of these haplotypes, particularly haplotype B, has been limited due to an inability to accurately resolve haplotypes, short of sequencing the genomic region, which has only been achieved for three haplotypes (Refs. 24 and 39 and accession no. AC011501.7). Family studies are a useful means by which to clarify haplotypes, if the families are large and informative for haplotype segregation.

In this study, we report the gene content evaluation of 12 families, 6 of which are multigenerational with large sibships, and the elucidation of specific haplotypes within these families. In addition, the investigation presented in this study includes typing for the two major subtypes of a pseudogene, for the expressed and nonexpressed forms of KIR2DL5, and for a new KIR gene that previously was included in the typing for KIR2DS4. Typing for these loci provides additional insight into distinct haplotype gene combinations based on gene content. These findings were then extended to a set of 49 sibling pairs and a panel of unrelated donors. The studies indicate the existence of a limited number of gene combinations centromeric to KIR2DL4 and greater variability telomeric to this anchor gene. We present a model for the genomic organization of the human KIR region where the gene content can be defined by six major haplotype gene combinations, each with multiple permutations.

Materials and Methods

**Populations**

The normal Caucasian population consisted of 85 randomly selected unrelated individuals. Sixty-one of these individuals were relatives of patients from the New York tristate area referred to Memorial Sloan-Kettering Cancer Center (New York, NY) for allogeneic hematopoietic stem cell transplantation, and the remaining 24 individuals were parents from the Center d’Etude du Polymorphism Humain (CEPH) cell bank. An additional 49 unrelated Caucasian individuals, each being a sibling of one of the panel donors from the New York tristate area, were included for comparison of KIR genotypes between siblings. These 49 individuals were not included in the analysis of the unrelated panel. Family studies were performed on B-lymphoblastoid cell lines (BLCL) derived from the family members of 12 CEPH families (consisting of at least two parents and two children). Six CEPH families, consisting of multiple generations and sibships containing at least eight children, were selected for extended typing studies for a total of 98 individuals. Seventeen BLCLs derived from offspring of consanguinous parents and obtained from the International Histocompatibility Workshop repositories were included as a potential source of KIR homoygous cells (40). BLCL were maintained in RPMI 1640 supplemented with 10% FBS at 5% CO₂, 37°C.

**Sample preparation**

Genomic DNA was extracted from 5 × 10⁶ PBMCs, bone marrow mononuclear cells, or BLCL tissue culture cells using the Puregene DNA Isolation kit (Genta Systems, Minneapolis, MN) according to the manufacturer’s instructions.

**Polymerase chain reaction**

PCR-sequence-specific primers (SSP) for the detection of KIR gene loci in genomic DNA are detailed in Table I. To detect the gene encoding all known alleles at a given locus and to achieve consistent results, alternative primer sets to those previously published (29, 32) were designed for the detection of KIR2DS1, 2DS5, 2DL1, 2DL4, 2DL5, 3DS1, 3DL1, 3DL2, 3DL3, and pseudogenes 3DP1 and 2DP1. Pseudogene KIR1DP1 has previously been designated KIRX (24), KIR2DS6 (39), KIR48 (41), and CD158c (8). Pseudogene KIR2DP1 has previously been designated KIRZ (24), KIR15 (41), and KIRY (20). KIR genes are reported in this study according to the guidelines by the Human Gene Nomenclature Committee (HGNC) (42). Results obtained with these primer sets were compared with results using primers designed by other groups (29, 32). Primers were also designed in the 5’-untranslated region (UTR) and 3’-UTR as forward or reverse primers, respectively, for use with a KIR gene-specific primer partner (Table I). For many KIR genes, two different primer sets were used to establish the presence or absence of the gene. The primer pair for pseudogene 3DP1 amplifies two different-sized amplicons (344 and 1817 bp) corresponding to KIR3DP1 and KIR3DP1+. In addition, primer sets were designed to accurately identify 2DS4 and distinguish it from a novel gene (designated KIR1D), which contains a 22-bp deletion in 2D2, leading to a frame shift and premature stop codon, with a predicted protein product truncated just within the transmembrane domain. Primers were also designed for KIR2DL5 subtyping to distinguish between expressed and non-expressed variants of 2DL5 (44). Amplifications using primers labeled 2DL5.1 identify the expressed variants 2DL5.1 and 2DL5.3, and amplifications using primers labeled 2DL5.2 identify the nonexpressed variants 2DL5.2 and 2DL5.4 (nomenclature 2DL5.1-2DL5.4 as previously described; Ref. 44). All primers were confirmed to be KIR gene-specific by basic local alignment sequence tool searches and verification with the National Center for Biotechnology Information KIR database (http://www.ncbi.nlm.nih.gov/IEB/Research/GVWG/MHC/MHC.cgi) (3) and were subsequently validated by amplicon sequencing. Nucleotide sequencing of amplicons was performed using dye terminators and automated sequencing (ABI 377 instrument; PE Applied Biosystems, Foster City, CA) in the Sloan-Kettering Institute Sequencing Core Facility (New York, NY).

Amplification conditions were optimized for each primer pair set as follows: µl volume reactions were prepared to include 0.1 µg of test DNA, PCR buffer, 0.2 mM of dNTPs, 2.5 U Taq DNA polymerase (Roche Molecular Biochemicals, Mannheim, Germany), 0.4 µM (with the exception of KIR2DS1 primers which were used at 0.8-µM final concentration). Internal controls for KIR2DL4 were used to confirm PCR amplifications. All amplifications were performed in PerkinElmer GeneAmp 9700 or 2400 thermocyclers (Wellesley, MA) programmed with a 2-min denaturing step, followed by 30 cycles of 92°C for 10s, 65°C for 30s, 68°C for 1 min 30 s, followed by 72°C for 10 min. Annealing temperatures were modified for primers amplifying KIR2DL2 (63°C), 2DS3 (63°C), KIR1D (63°C), 2DS4 (61°C), and 2DS5 (63°C). Extension time was modified for long-range amplification of gene KIR2DL1 (10 min).

**2DS4 variant identification and cloning**

Amplification of KIR2DS4 using previously published primer pairs (29) was performed on genomic DNA samples. PCR products were gel-purified using the GFX PCR DNA and Gel Band Purification kit (Amersham Pharmacia Biotech, Piscataway, NJ) and were subsequently cloned into pCR2.1-TOPO using the TOPO-TA Cloning kit (Invitrogen, Carlsbad, CA). Plasmids were isolated from individual bacterial clones using the QiAprep Miniprep kit (Qiagen, Chatsworth, CA). Plasmids were isolated from individual bacterial clones using the GFX PCR DNA and Gel Band Purification kit (Amersham Pharmacia Biotech, Piscataway, NJ) and were subsequently cloned into pCR2.1-TOPO using the TOPO-TA Cloning kit (Invitrogen, Carlsbad, CA). Vector-specific primers were used for sequence analysis of plasmid inserts using dye terminators and automated sequencing (ABI 377 instrument; PE Applied Biosystems) for identification of a 22-bp deletion. Primer pairs were then designed to specifically identify the intact KIR2DS4 gene and KIR1D from...
genomic DNA samples (Table I), NK cells isolated from individuals homozygous for KIR1D and individuals heterozygous for 2DS4 and KIR1D were used for mRNA isolation (MicroFastTrack 2.0 Kit; Invitrogen) and cDNA isolation (cDNA Cycle Kit; Invitrogen). KIR2DS4-specific primers (forward 5'-CCATGGTCCACATCGTGATC-3', reverse 5'-ATCCACGTTCTTGAGTCC-3') were then used to generate a 730-bp 2DS4 ampli-

KIR haplotypes were determined by segregation patterns in families. In general, genotypes were constructed using an algorithm that divided haplotypes into haplotypes shared with the family's founders (parents), haplotypes unique to the family's founder, and haplotypes specific to the family itself. This allowed for the identification of novel haplotypes not previously described. KIR1D cDNA clones were distinguished from 2DS4 clones by sequencing and by PCR analysis.

**Haplotypes analysis**

KIR haplotypes were determined by segregation patterns in families. In assigning genes to specific haplotypes, the following assumptions were made: 1) all haplotypes contained KIR3DL3, 2DL4, and 3DL2; 2) haplotypes contained either 2DL2 or 2DL3, but not both; and 3) haplotypes contained pseudogene 2DL5, but not both. All assumptions were supported by previous analyses of linkage disequilibrium (LD) (29, 31, 32, 34, 35) and by sequenced KIR haplotypes (Refs. 24 and 39 and accession no. AC011501).

Haplotype deduction for sibling pairs not included in the family studies was performed by first applying the haplotypes defined in the family studies and, in the case of siblings who were not part of the family studies, by analyzing the segregation patterns in the family. This allowed for the identification of novel haplotypes not previously described. KIR1D cDNA clones were distinguished from 2DS4 clones by sequencing and by PCR analysis.

**Statistical analysis**

The observed KIR Ag frequency was determined by the ratio of gene presence within the population to the total population number. Gene frequency was calculated by the formula: gene frequency = 1 - 2√(1 - f), where f is the observed Ag frequency in the population. Two-locus LD parameters (Δ) were calculated according to Mattiu et al. (45).

**Results**

**KIR genomic typing of unrelated individuals**

A PCR-SSP typing method was devised for the identification of 16 known KIR genes and pseudogenes, KIR2DS1-5, 2DL1-5, 3DS1, 3DL3, and pseudogenes 2DP1 and 2DL1 (Table I). The method includes alternative primer sets designed for inclusion of all known alleles for KIR2DS1, 2DS3, 2DS5, 2DL4, 2DL5, 3DS1, 3DL1, and 2DL2. Confirmation of new primer set specificity was achieved in several ways; positive amplification reactions were compared with results obtained using previously published primer sets (29, 32); amplification products were sequenced for specificity; and finally, typing of 85 unrelated Caucasian individuals yielded comparable estimated gene frequencies to those in previously published findings (29, 32, 34, 35). In this cohort, 36 different genotypes were identified, 23 of which were unique, with the remaining genotypes each observed from 2 to 14 times (Fig. 1). The most common genotype (observed 278 times) was KIR5DL3-2DL3-2DP1-2DL1-3DP1-2DL4-3DL1-1D-3DL2, corresponding to homozygosity for the major subtype of the previously reported A haplotype (see below). Observed nine times was the genotype KIR3DL3-2DL3-2DP1-2DL1-3DP1-2DL4-3DL1-2DS4-1D-3DL2, corresponding to heterozygosity for the two major subtypes of haplotype A. Genotypes represented more than once accounted for 73% of all genotyped samples. Included in this typing is the detection of a new gene KIR1D and typing for the two different variants of pseudogene 3DP1.
A novel 2DS4 variant resembling MmKIR1D

PCR-SSP typing for gene 2DS4 using a previously published primer set (29) revealed a nearly ubiquitous gene frequency within the population. However, sequencing of several genomic DNA amplicons using this primer set revealed the presence of a 2DS4 gene variant, characterized by a 22-bp deletion in the second extracellular Ig-like domain (D2). Although the remainder of the nucleotide sequence is identical to KIR2DS4, the amino acid sequence resulting from the deletion-generated frame shift is comprised of a novel stretch of 88 aa, loss of the D2 Ig domain, and termination at a final length of 239 aa, one amino acid into the putative transmembrane domain, with no discernible cytoplasmic domain (Fig. 2). Search of GenBank nucleotide sequences revealed the presence of this gene in the genomic DNA sequence of a human chromosome 19 haplotype (accession no. AC011501.7), placing this novel KIR gene in the KIR gene cluster on human chromosome 19q13.4. A GenBank amino acid search identified significant amino acid homology (72%) to a variant of the Mm-KIR1D receptor found in rhesus monkeys (accession no. AF334633) which has nucleotide homology to Mm-KIR2DL4, but contains only one complete Ig domain and no cytoplasmic tail as a consequence of a frame shift (17). Because of the amino acid homology between this novel human gene and Mm-KIR1D, we propose designation of the human gene as KIR1D.

KIR1D transcripts are transcribed in heterozygous individuals, efforts to isolate the full cDNA transcripts revealed the majority of transcripts to be 2DS4. In contrast, KIR1D transcripts were more readily isolated from individuals homozygous for KIR1D (Fig. 3). From a separate individual, a KIR1D variant was identified; while exhibiting the same 22-bp deletion, this variant lacked the stem region (corresponding to exon 5), indicating its likely identity as a splice variant (accession no. AY102633). A GenBank search also revealed a recently cloned KIR gene that is identical to human KIR1D except for a single nucleotide substitution at position 156 giving rise to a codon change from arginine to lysine in the exon 4-encoded gene product (accession no. AF417554).

LD patterns support KIR1D and KIR2DS4 as alleles

The percentage of Caucasian individuals exhibiting KIR1D is high (78.8%) with an estimated gene frequency of 0.54. Accurately revised typing for KIR2DS4 revealed a lower Ag frequency (35.3%) than previously determined with an estimated gene frequency of 0.20. LD analysis identified strong negativity between the two genes, supporting their possible identities as alleles. Strong positive LD with KIR1D was noted with KIR3DL1 and 2DL3, whereas there was only a very weak LD between KIR2DS4 and these genes, indicating that the previously noted positive LD between KIR2DS4 and KIR3DL1 and KIR2DS4 and 2DL3 was likely due to lack of distinction between KIR1D and 2DS4. Strong negative LD between KIR1D and 2DS1 is also seen. Other LD relationships are indicated in Table II.
A pseudogene 3DP1 variant defines a partial KIR haplotype

KIR typing for pseudogene 3DP1 revealed the presence of two predominant subtypes. These can be identified as two different amplicons, whose 1.5-kb size difference is due to the presence or absence of exon 2 and its flanking intron sequences. KIR3DP1, characterized by the absence of exon 2, was found in the genomic DNA sequence of a fully sequenced human chromosome 19 KIR haplotype (accession no. AC011501.7); KIR3DP1v, characterized by the presence of exon 2, was found in a separate chromosome 19 KIR haplotype genomic sequence (accession no. AL133414). LD for KIR3DP1 and KIR3DP1v reveal the two exhibiting strong negative LD, supporting a possible allelic relationship. Whereas KIR3DP1 has a high gene frequency within the population (0.72), KIR3DP1v exhibits a significantly lower gene frequency (0.17). KIR3DP1 displays a strong positive LD with pseudogene KIR2DP1 and with 2DL1, supporting their inclusion in a partial haplotype KIR2DP1-KIR2DL1-KIR3DP1. In contrast, KIR3DP1v displays an equally strong negative LD with KIR2DP1 and 2DL1; moreover, the identical frequency of these gene pairs in relation to each other further supports their definition of a KIR3DP1v partial haplotype distinguished by the absence of KIR2DP1 and 2DL1. In other words, when KIR3DP1v is present, KIR2DP1 and 2DL1 are absent. This partial haplotype lacking KIR2DP1-2DL1 en bloc can be observed in PCR-based genotyping only when the individual is homozygous for KIR3DP1v, a relatively rare combination noted in five of the genotypes from the Caucasian panel studied (see Fig. 1, genotypes Z, AA, AB, AC, and AD) and in extended family study haplotype elucidation. In addition, KIR3DP1v was found to be in strongly positive LD with KIR2DL2 and KIR2DS2 extending the partial haplotype to resemble KIR2DL2-(absent 2DP1-absent 2DL1)-3DP1v.

Family segregation studies define haplotypes

Family studies performed to date have not included typing for KIR1D or for the pseudogenes 3DP1, 3DP1v, and 2DP1, whose typing significantly helped to define distinct haplotypes. Twelve families consisting of both parents and two children were analyzed for gene content; of these, six families were selected for extended family analysis based on heterozygosity at the KIR3DL1-3DS1 locus and the likelihood of haplotype resolution. Through typing of multiple generations and sibships numbering no fewer than eight and in several cases more than 10, KIR haplotypes could be reliably resolved with nearly no ambiguities. Examples of genotype analysis for individuals from two families and the resolved parental and grandparental haplotypes are shown in Fig. 4. From 12 families, 11 different haplotypes were resolved by gene content analysis for individuals heterozygous for 3DP1 and KIR1D. KIR1D is a deletion variant of 2DS4 resulting in significant amino acid change due to a frame shift. A. Nucleotide sequence alignment of KIR2DS4 cDNA (allele L76672 is used for comparison), KIR1D genomic DNA (haplotype sequence AC011501.7), and KIR1D cDNA (accession no. AY102624) demonstrates a loss of 22 nucleotides from exon 4 in KIR1D. B. Amino acid alignment reveals no homology between KIR1D and 2DS4 after amino acid 151, but demonstrates 72% homology to a MnKIR1D variant found in the rhesus monkey (accession no. AF334633).
Table 2. LD for KIR loci

<table>
<thead>
<tr>
<th></th>
<th>2DL2</th>
<th>2DL3</th>
<th>2DP1</th>
<th>2DL1</th>
<th>3DP1</th>
<th>3DP1v</th>
<th>3DL1</th>
<th>3DS1</th>
<th>2DL5</th>
<th>2DS3</th>
<th>2DS5</th>
<th>2DS1</th>
<th>2DS4</th>
<th>1D</th>
</tr>
</thead>
<tbody>
<tr>
<td>2DS2</td>
<td>0.22</td>
<td>-0.254</td>
<td>-0.18</td>
<td>-0.194</td>
<td>-0.194</td>
<td>0.109</td>
<td>-0.147</td>
<td>0.031</td>
<td>0.07</td>
<td>0.077</td>
<td>0.031</td>
<td>0.031</td>
<td>-0.005</td>
<td>-0.067</td>
</tr>
<tr>
<td>p</td>
<td>&lt;0.0005</td>
<td>&lt;0.001</td>
<td>&lt;0.025</td>
<td>&lt;0.025</td>
<td>&lt;0.025</td>
<td>&lt;0.0005</td>
<td>NS</td>
<td>NS</td>
<td>&lt;0.025</td>
<td>&lt;0.0005</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>2DL2</td>
<td>-0.254</td>
<td>-0.18</td>
<td>-0.194</td>
<td>-0.194</td>
<td>0.11</td>
<td>-0.147</td>
<td>0.031</td>
<td>0.07</td>
<td>0.077</td>
<td>0.012</td>
<td>0.031</td>
<td>-0.005</td>
<td>-0.067</td>
<td></td>
</tr>
<tr>
<td>p</td>
<td>&lt;0.001</td>
<td>&lt;0.025</td>
<td>&lt;0.025</td>
<td>&lt;0.025</td>
<td>&lt;0.0005</td>
<td>NS</td>
<td>NS</td>
<td>&lt;0.025</td>
<td>&lt;0.0005</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>2DL3</td>
<td>0.166</td>
<td>0.18</td>
<td>0.18</td>
<td>-0.099</td>
<td>0.072</td>
<td>0.026</td>
<td>-0.072</td>
<td>0.003</td>
<td>-0.088</td>
<td>-0.094</td>
<td>0.024</td>
<td>0.069</td>
<td></td>
<td></td>
</tr>
<tr>
<td>p</td>
<td>&lt;0.0005</td>
<td>&lt;0.0005</td>
<td>&lt;0.0005</td>
<td>&lt;0.0005</td>
<td>&lt;0.0005</td>
<td>NS</td>
<td>NS</td>
<td>&lt;0.025</td>
<td>&lt;0.025</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>2DP1</td>
<td>0.19</td>
<td>0.19</td>
<td>-0.224</td>
<td>0.051</td>
<td>0.018</td>
<td>-0.03</td>
<td>0.038</td>
<td>-0.062</td>
<td>-0.046</td>
<td>0.03</td>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p</td>
<td>&lt;0.0005</td>
<td>&lt;0.0005</td>
<td>&lt;0.0005</td>
<td>&lt;0.0005</td>
<td>&lt;0.0005</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2DL1</td>
<td>0.205</td>
<td>-0.242</td>
<td>0.046</td>
<td>0.028</td>
<td>-0.01</td>
<td>0.041</td>
<td>-0.045</td>
<td>-0.061</td>
<td>0.036</td>
<td>0.032</td>
<td>0.021</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p</td>
<td>&lt;0.0005</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3DP1</td>
<td>-0.242</td>
<td>0.046</td>
<td>0.028</td>
<td>-0.01</td>
<td>0.041</td>
<td>-0.045</td>
<td>-0.061</td>
<td>0.036</td>
<td>0.032</td>
<td>NS</td>
<td>NS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p</td>
<td>&lt;0.0005</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3DP1vA</td>
<td>-0.029</td>
<td>-0.019</td>
<td>-0.033</td>
<td>-0.021</td>
<td>-0.006</td>
<td>-0.001</td>
<td>0.018</td>
<td>-0.043</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3DL1</td>
<td>-0.163</td>
<td>-0.151</td>
<td>-0.078</td>
<td>-0.067</td>
<td>-0.163</td>
<td>0.043</td>
<td>0.117</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p</td>
<td>&lt;0.025</td>
<td>NS</td>
<td>&lt;0.025</td>
<td>NS</td>
<td>&lt;0.0005</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3DS1</td>
<td>0.165</td>
<td>0.081</td>
<td>0.102</td>
<td>0.139</td>
<td>-0.043</td>
<td>-0.058</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p</td>
<td>&lt;0.0005</td>
<td>&lt;0.0005</td>
<td>&lt;0.0005</td>
<td>&lt;0.0005</td>
<td>&lt;0.0005</td>
<td>&lt;0.005</td>
<td>NS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2DL5</td>
<td>0.089</td>
<td>0.132</td>
<td>0.165</td>
<td>-0.053</td>
<td>-0.075</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p</td>
<td>&lt;0.0005</td>
<td>&lt;0.0005</td>
<td>&lt;0.0005</td>
<td>&lt;0.0005</td>
<td>&lt;0.0005</td>
<td>&lt;0.005</td>
<td>NS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2DS3</td>
<td>0.03</td>
<td>-0.034</td>
<td>0.011</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2DS5</td>
<td>0.143</td>
<td>-0.043</td>
<td>-0.102</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p</td>
<td>&lt;0.0005</td>
<td>NS</td>
<td>NS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2DS1</td>
<td>-0.043</td>
<td>-0.102</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p</td>
<td>&lt;0.0005</td>
<td>NS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2DS4</td>
<td>-0.153</td>
<td>-0.018</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p</td>
<td>&lt;0.0005</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*LD coefficients for two-locus associations were calculated from unrelated Caucasian individuals according to Mattiaz et al. (45). Pairs with \( p < 0.05 \) are indicated.

(Fig. 5A), with 2 representing haplotype A (haplotypes 1 and 2), as previously defined by Shilling et al. (36) as the presence of KIR2DL3 and the lack of all activating genes with the exception of 2DS4, presented in this study as comprising both 2DS4 and KIR1D.

Haplotype analysis-centromeric half

Segregation analysis of haplotypes supported the following findings: as predicted by their negative LD relationship, KIR2DL2 and 2DL3 segregated exclusively onto separate chromosomes. In none of the families could both KIR genes be assigned to the same haplotype, supporting their relationship as possible alleles. Pseudogene 2DP1, KIR2DL1, and pseudogene 3DP1 segregated en bloc and could be associated with either KIR2DL3 or 2DL2, although more often with the former (Fig. 5A, haplotypes 1, 2, 5, 6, 7 vs haplotypes 8 and 9). In contrast, the presence of the pseudogene variant 3DP1v was associated with the absence of pseudogene 2DP1 and KIR2DL1 and was observed exclusively with KIR2DL2 in the family studies. This “deletion” partial haplotype was observed in three of the resolved haplotypes and occurred in 9 of the 24 parents (37.5%) (Fig. 5A, haplotypes 3, 4, and 10). Haplotype 11 is an unusual KIR haplotype, seen only in one family.

Haplotype analysis-telomeric half-KIR1D redefines haplotype A

Telomeric to KIR2DL4, the following findings were observed: KIR1D and 2DS4 segregated exclusively onto separate chromosomes (e.g. Fig. 4, C and D). Typing for KIR1D revealed that what was previously defined as haplotype A can be divided into two groups, one containing KIR1D, designated as haplotype A-1D (Fig. 5A, haplotype 1), and the other containing KIR2DS4, designated as haplotype A-2DS4 (Fig. 5A, haplotype 2). In the family analyses, haplotype A-1D was the more frequently observed, representing 20 of a total of 48 parental haplotypes (41.7% gene frequency). In contrast, haplotype A-2DS4 was present in only 6 of a total of 48 haplotypes (12.5%). These frequencies are comparable to the frequencies in the panel of 85 unrelated Caucasians (38.8% and 11.8%, respectively). Homozygosity for haplotype A-1D was seen in 14 of 85 unrelated Caucasian individuals (16.7%), whereas homozygosity for haplotype A-2DS4 was seen in only 1 of 85 unrelated individuals (1.1%). From the same panel, haplotype A-1D had the highest haplotype frequency of 38.8%, while haplotype A-2DS4 had a haplotype frequency of 11.8%.

In one extended family, the exclusive segregation of KIR2DS1 from haplotypes containing KIR1D or 2DS4 was unequivocally observed (Fig. 4, C and D). This mutual exclusion is supported by the highly significant negative LD between the genes, suggesting their possible relationship as alleles. Likewise, KIR2DS3 and 2DS5 were never identified on the same haplotype in our studies. Accompanying KIR2DS3 and 2DS5 is 2DL5, which was not found to segregate separately from either of the two activating receptor genes, except in one haplotype (Fig. 5B, haplotype 12), which was later found in a sibling pair that displayed full KIR identity. Subtyping for the expressed (labeled 2DL5.1) and nonexpressed forms of KIR2DL5 (labeled 2DL5.2) identified a haplotype containing both forms (Fig. 4, A and B, and Fig. 5A, haplotype 9). Other haplotypes contained either the nonexpressed or the expressed KIR2DL5 subtype, paired with either 2DS3 or 2DS5. Finally, KIR2DL2 and 2DS2 were found to segregate together for all resolved haplotypes. To address the issue of nonspecific amplification or coamplification, >20 ampiclons for both KIR2DL2 and KIR2DS2 were sequenced, repeatedly confirming gene specificity of the primers.

Sibling studies

KIR genotyping of 49 sibling pairs allowed for the informed deduction of 10 additional haplotypes (Fig. 5B). Of the 49 pairs, 17 pairs (34.7%) yielded identical genotypes, which could imply identical KIR haplotype pair combinations. Although this ratio was
FIGURE 4. Continues.
higher than the 25% expected from normal Mendelian segregation, it could be explained by high frequency of haplotype A-1D in these sibling pairs (12 of 17). Given the high frequency of this haplotype in the population, it might be expected that in some cases where both siblings exhibit this haplotype, they may not be identical haplotypes by descent. Eleven pairs (22.5%) exhibited no shared haplotypes, while 21 pairs (42.9%) exhibited one shared haplotype. Of the 10 new KIR haplotypes deduced among the sibling pairs, one haplotype contained both expressed and nonexpressed variants of KIR2DL5 (Fig. 5B, haplotype 22). Combinations from the deduced sibling haplotypes in addition to the 11 identified through family studies could account for all 36 genotypes seen in the unrelated Caucasian panel (Fig. 1). Fifteen unrelated individuals among the sibling group were found to be homozygous for 1 of the 24 identified or deduced haplotypes. This included eight instances of haplotype 1, two of haplotype 5, and one each of haplotypes 2, 4, 8, 14, and 17.

**Haplotype frequency in the Caucasian panel**

The unrelated Caucasian panel described in Fig. 1 was analyzed for deduced KIR haplotypes based on the haplotypes characterized from family studies and sibling pair analysis (Fig. 5, A and B). All 36 genotypes could be resolved into corresponding pairs of haplotypes as shown in Fig. 1. Of the 170 haplotypes exhibited in 85 individuals, the most commonly observed haplotype was haplotype 1, occurring 66 times (38.8%). In contrast, haplotype 2 was found 20 times (11.8%). Comprising the “classical haplotype A” frequency, the sum of frequencies for haplotype 1 (haplotype A-1D) and haplotype 2 (haplotype A-2DS4) yielded a total haplotype A frequency of 50.6%, consistent with previous estimates (29, 36). Less frequent, but very common, were the haplotypes characterized by the deletion partial haplotype (KIR2DS2-2DL2-3DP1v): haplotype 3 had a frequency of 7.7% while haplotype 4 had a frequency of 6.5%. In total, “short” haplotypes with 3DL1-2DS4 or 3DL1-1D accounted for 64.7% of all KIR haplotypes in this Caucasian population.

**Consanguineous individuals**

Consanguineous cell lines have proved useful for HLA allele identification and HLA haplotype analysis. A panel of 17 cell lines derived from individuals from consanguineous families was analyzed for KIR gene content (Fig. 6). Among these individuals, two...
more unique genotypes were found, yielding two additional unique haplotypes (Fig. 5B, haplotypes 18 and 23). Based on the conservative estimates for the haplotypes described in this study, five of the consanguineous individuals appear to be homozygous for haplotype A-1D, while one is homozygous for haplotype A-2DS4. These cell lines could very well be KIR homozygous by descent. For cell lines with larger gene content, it was not possible through this analysis to document homozygosity.

**Haplotype A** is comprised of multiple stereotypical KIR combinations

Haplotype A has been defined as containing more variable KIR gene combinations and characterized by KIR2DS genes other than KIR2DS4. According to this definition, 21 different B haplotypes were resolved. These haplotypes included most of the partial haplotypes described by other groups (31, 36); however, in contrast to one previous report (31), there was no representation of the 2DS3-2DS5-2DS1 complement in our studies. Evaluation of the haplotypes revealed specific patterns that are supported by KIR locus LD analysis. Although KIR3DS1 is most commonly associated with the presence of multiple 2DS genes, KIR3DL1 could occasionally, but not as commonly, also be found to be associated with these genes. KIR2DS3 and 2DS5 were each almost always paired with 2DL5, whose subtyping revealed the nonexpressed 2DL5.2 variant exhibiting a stronger positive LD with 2DS3 (p < 0.0005) in comparison to 2DS5 (p < 0.05). In addition, the nonexpressed 2DL5.2 variant exhibited a strong positive LD (p < 0.0005) with 2DS3 and 2DS5.2 variant showed strong positive LD with 3DS1 (p < 0.0005), 2DS3 (p < 0.0005), and 2DS2 (p < 0.0005) and no significant LD with 2DL2. These analyses support earlier reports of the placement of KIR2DL5.2 adjacent to 2DL2 on the chromosome and 2DL5.1 adjacent to 3DS1 (44). Although pair KIR2DL5-2DS3 could be seen with KIR1D, 2DS4, or 2DS1, pair KIR2DL5-2DS5 was seen only with 2DS1 in this analysis, consistent with the strongly positive LD between 2DS5 and 2DS1. KIR2DS2 displayed a strong positive LD with 2DS3, but its main association appeared to be with 2DL2. Although not exclusive, these patterns were supported by the overwhelming majority of identified and deduced haplotypes. Only rarely would haplotypes be identified that did not adhere to these patterns (see Fig. 5, haplotypes 11 and 12).

**Discussion**

The human immune system has adopted a strategy of immense diversity, exemplified by the MHC, whose variable gene content and allelic polymorphism combine to individualize the HLA genotype. Like the HLA, the KIR region is rapidly emerging as an equally diverse region that uses the similar evolutionary strategies of variable gene content and polymorphism to achieve similar depths of diversity. Addressing the need to identify unique KIR genotypes is increasingly refined methods of KIR gene detection. Initial studies first revealed variations in KIR gene content from individual to individual, whose KIR gene combinations roughly corresponded to two KIR haplotype groups, A and B (29, 30). These gene content studies have been extended to population analyses, which while revealing a great diversity in KIR genotypes, still supported the model of two general haplotype groups (31–35). Providing a still higher resolution of the KIR haplotypes has been achieved by analyses of specific KIR gene alleles (36–38) and sequence analysis of three haplotypes (Refs. 24 and 39 and accession no. AC011501.7). Although these two approaches have provided detailed analyses of KIR haplotypes, one based on sequencing, the other based on allele detection, they have also supported the presence of two main haplotype groups.

Haplotype A has been defined as containing KIR3DL3, -2DL2, -2DL1, -2DL4, -3DL1, -2DS4, and -3DL2, and haplotype B has been defined as having more KIR genes than the A haplotype; generally, these additional genes are the activating KIR genes characterized by short cytoplasmic tails. The ability to unambiguously define haplotypes has previously been approached by family studies; however, these studies have typically included fairly small family cohorts, consisting of three to six members. In this study, we report the elucidation of haplotypes through gene content analysis of extended family pedigrees whose large sibships and multiple generations could reliably identify specific haplotypes. Eleven different haplotypes were resolved through these studies, many of which have not previously been described. Contributing significantly to the resolution of these haplotypes was the typing for two KIR loci in particular, the 2DS4/KIR1D locus and the pseudogene 3DP1 locus. In this study, we report for the first time the cDNA isolation of human KIR1D, whose typing has not previously been distilled from that of 2DS4 and which appears to be the more prevalent of the two putative alleles of the 2DS4 locus. In addition, typing for the two known variants of pseudogene 3DP1 has revealed the presence of a “deletion” partial haplotype that is distinguished by the presence of the pseudogene 3DP1v with KIR2DS2 and 2DL2, and the absence of 2DL1 and pseudogene 2DP1.

KIR1D has not before been identified in humans and was recently identified in the rhesus monkey as Mm-KIR1D (17). Although the nucleotide sequences do not show significant homology, the amino

---

**FIGURE 6.** KIR locus genotypes in consanguineous individuals. Genotype AG, corresponding to homozygosity for haplotype A-1D, is represented most frequently. These individuals may be homozygous for the KIR region by descent.
acid sequences between the two gene products display a high degree of homology, suggesting they may be functional orthologs. In addition, while the Mm-KIR1D gene appears to be a deletion variant of Mm-KIR3DL7, the human KIR1D is likely a deletion variant of 2DS4. Interestingly, the respective deletions in the rhesus and human genes result in frame shifts that yield homologous predicted protein products. As in the rhesus monkey, it also appears that several putative splice variants of human KIR1D are present, with at least one isolated clone having a deletion of the exon encoding the stem region. KIR1D encodes a predicted protein product that has a complete D1 Ig domain and a portion of the D2 domain, before the frame shift deletion abrogates any transmembrane or cytoplasmic region. Splice variants of other KIR/ILT genes have been identified (46, 47), some of which are lacking the sequences encoding the transmembrane and cytoplasmic regions, leading one to speculate about the possibility of a secreted truncated protein product. However, the presence of an encoded secreted protein such as KIR1D has not heretofore been described. The ligands for the KIR1D proteins, both in the human and the rhesus monkey, remain unknown; however, the likelihood that that it is an MHC-like molecule is remote. Of interest is the difficulties in isolating KIR1D relative to 2DS4 transcripts in a heterozygous individual, despite the high prevalence of the gene within the population. The significance of this apparent transcriptional discord remains unclear. In contrast, transcripts were easily obtainable from an individual found to be homozygous for KIR1D.

A new model for KIR haplotypes

In this study, we describe extended family studies with very large sibships to define specific haplotypes, following them through three generations within the families. Analysis of the resolved haplotypes invites a reassessment of the initial two haplotype groups, first proposed by Uhrberg et al. (29) and further supported by the genomic sequencing of the complete KIR region for a representative of the A and B haplotypes (24). From our studies, it is evident that what was traditionally viewed as haplotype A is actually a combination of two different genotype groups, one containing KIR2DS4 and the other containing KIR1D in association with pseudogene 3DP1. The latter subtype is the more common, comprising 73% of haplotype A cases, and observed in 39% of unrelated Caucasian individuals. In contrast, the 2DS4-positive haplotype A has a haplotype frequency of 12%. The combined frequency of both haplotype A subtypes of 51% is consistent with previously published estimates (29, 36). It is interesting to note that because the most common genotype in the Caucasian population reflects homozygosity for the haplotype A containing KIR1D (14%), these individuals are lacking all activating KIR receptors, with the exception of KIR2DLA, a ubiquitously found receptor characterized by a long cytoplasmic tail, but exhibiting activating function (48, 49).

Although our studies are consistent with the broad classification of KIR regions into A and B haplotypes, the inclusion of typing in families for the KIR pseudogenes and the discovery of the human KIR1D as a likely allele of the 2DS4 locus has provided a more detailed view of the KIR genomic region, as also recently described by Shilling et al. (36). Our data can best be accommodated by an alternative model for KIR haplotypes (Fig. 7). The KIR haplotype can be considered as two halves, the centromeric half bordered upstream by KIR3DL3 and comprised of those KIR genes upstream of anchor gene KIR2DL4 and the telomeric half bordered downstream by 3DL2 and comprised of those KIR genes downstream of 2DL4. The centromeric half may be characterized by the presence of KIR2DL3 or 2DL2, but not both, and rarely, neither. In the rare case where neither KIR2DL3 nor 2DL2 was identified, the

![Figure 7](http://www.jimmunol.org/)
posibility exists for the presence of a new KIR gene (S. Chida and D. Geraghty, personal communication). In our studies, KIR2DL2 was found always to segregate with 2DS2, the latter occupying the more upstream position of the two. Alteration of our model to reflect the gene position of KIR2DS2 is consistent with a reinterpretation of the RP5 chromosomal KIR genomic sequence (accession no. AL133414) and with gene order studies performed by S. Chida and D. Geraghty (personal communication). Following the locus occupied by either KIR2DL3 or 2DL2 is either pseudogene 3DP1v or the trio of pseudogene 2DP1-KIR2DL1-pseudogene 3DP1. In our studies, KIR2DL3 was exclusively associated with the trio 2DP1-2DL1-3DP1 to define a common partial haplotype (Fig. 7, haplotype models 1 and 4). Although this trio could also be seen to associate with KIR2DL2 and 2DS2 (Fig. 7, haplotype models 2 and 5), 2DL2 and 2DS2 are more commonly found to associate with the pseudogene 3DP1v, in a partial haplotype defined by the en bloc absence of 2DP1-2DL1-3DP1 (Fig. 7, haplotype models 3 and 6). This partial haplotype is consistent with an available genomic sequence of this region of the 2DL2-positive chromosome (24).

The telomeric half of the complex is alternately defined by the presence of KIR3DL1 or 3DS1, or rarely, neither. Most commonly, KIR3DL1 is associated with a “short” haplotype, one that is lacking most activating KIR genes. In these “short” haplotypes, the locus adjacent to KIR3DL2 may be occupied by 2DS4, or more commonly, KIR1D (Fig. 7, haplotype models 1–3). In the “long” haplotypes, a number of activating KIR genes may be seen, almost always accompanied by KIR2DL5. KIR2DL5 has four known variants, two of which (2DL5.1 and 2DL5.3) are expressed and, in this paper, collectively labeled 2DL5.1. The two nonexpressed variants (2DL5.2 and 2DL5.4) are collectively labeled 2DL5.2. Previous gene order studies (44) and LD analysis performed in this report support the likelihood that the nonexpressed 2DL5 variants are located in the centromeric half adjacent to 2DL2 and are most often accompanied by 2DS3, while the expressed 2DL5 variants remain in the telomeric half accompanied by 2DS3 or 2DS5 (Fig. 7, haplotype models 5 and 6). In our studies, haplotypes containing both expressed and nonexpressed KIR2DL5 variants have been identified. Finally, in the long haplotypes, the locus adjacent to KIR3DL2 can be occupied by KIR1D, 2DS4, or 2DS1 (Fig. 7, haplotype models 4–6). Accordingly, the vast majority of our data fits the model presented in Fig. 7, which classifies KIR haplotypes into one of six major groups. Haplotype 1 in Fig. 7 is synonymous with the canonical haplotype A, with the remainder comprising the classical B haplotype. The documentation of the three short haplotypes can readily be seen in the family studies. The high frequency of these haplotypes with their small gene content in the telomeric KIR region is evident also in the sibling analysis and in the unrelated panel studies.

Not all 30 permutations within the three major subtypes of the long B haplotypes (Fig. 7, haplotypes 4–6) were identified in the family studies, in part due to the limitations of the family materials, but also likely in part due to differences in prevalence for each permutation. For instance, haplotype 5 (Fig. 5) identified in the pedigree studies was the most common long KIR haplotype found in the population analysis, occurring in 21 of 85 individuals (24.7%) with a haplotype frequency of 12.4%. Because KIR2DL5 was observed to pair with either 2DS3 or 2DS5, with one rare exception (Fig. 5B, haplotype 12), it is therefore very possible that KIR2DS3 and 2DS5 are alleles of the same locus, but our studies cannot resolve this. Although we did not observe KIR2DL5, 2DS3, and 2DS5 segregating together, it is theoretically possible for such a haplotype to exist if both 2DL5 subtypes are present. Similarly, we did not observe KIR1D, 2DS1, and 2DS4 together either as pairs or as triplets, which in combination with their negative LD, suggests their relationship as alleles. Although the nucleotide sequencing data would support KIR1D and 2DS4 being alleles, we cannot unequivocally claim 2DS1 as an allele of the same locus. Identification of possible recombinant haplotypes or direct sequencing of appropriate KIR genomic regions will be needed to resolve this issue. The finding that the pairing of 2DL5/2DS5 most frequently segregated with 2DS1 was also very characteristic for well-defined haplotypes in the families. We have no explanation for the mechanisms underlying these phenomena, but some functional interactions between the gene products may have preserved this relationship during evolution.

It is to be expected that “new” unusual “recombiant” haplotypes or haplotypes with gene duplication of fragments of the KIR region will be identified which cannot be accommodated by our model. This would be expected for a genomic region with multiple highly homologous genes, as has been observed for the HLA chromosomal region (50). The emerging picture of the gene content within the KIR region is becoming very similar to the findings made for the HLA class II region. In this study, the HLA-DR region demonstrates a variable gene content with functional class II genes and pseudogenes. Recently, several other genomic regions within the human genome have displayed similar variable gene compositions (50).

Acknowledgments

We thank Drs. Yatin M. Vyas, Hina Maniar, and Linda Butros for helpful advice and suggestions and Karen Reed for preparing the manuscript. We also thank Dr. A. Sauvageau for performing some of the statistical computations. We also thank Drs. Shohei Chida and Dan Geraghty (Fred Hutchinson Cancer Research Center, Seattle, WA) for permission to cite their unpublished findings regarding KIR gene order.

References


