Genes and Immunity (2008) 9, 470–480

www.nature.com/gene

ORIGINAL ARTICLE

KIR2DL5 alleles mark certain combination of activating KIR genes

Z Du[1], SK Sharma[2], S Spellman[3], EF Reed[1] and R Rajalingam[1]

[1]Department of Pathology and Laboratory Medicine, UCLA Immunogenetics Center, David Geffen School of Medicine at UCLA, University of California at Los Angeles, Los Angeles, CA, USA; [2]Department of Medicine, All India Institute of Medical Sciences, New Delhi, India and [3]National Marrow Donor Program (NMDP), Minneapolis, MN, USA

*Killer cell Ig-like receptors (KIR) control the immune response of NK cells and some T cells to infections and tumors. KIR genes evolve rapidly and are variable between individuals in their number, type and sequence. Here, we determined the nature of KIR2DL5 gene polymorphism in four ethnic groups using direct DNA sequencing method. Nine new sequences were discovered. Within the panel of 248 KIR2DL5-positive individuals, 14 KIR2DL5-sequences differing in coding regions were observed. They differed at only seven amino acid positions, and such limited polymorphism is consistent with its conserved nature throughout the hominoid lineage. Ethnic deviation was seen in the distribution of KIR2DL5A, KIR2DL5B and their alleles. African Americans had more KIR2DL5 alleles than other populations indicating that more polymorphisms are yet to be discovered in Africans. Linkage between KIR2DL5-alleles and certain activating-KIR genes were observed, but frequency of these linked clusters differed substantially between populations. Consequently, KIR2DL5 alleles can be used as markers to predict the activating-KIR gene content. Typing system distinguishing A*001 and B*002 alleles can serve as a powerful screening test to assess the content of most variable activating-KIR genes that have been implicated in human disease and in the outcome of hematopoietic stem cell transplantation.*
Genes and Immunity (2008) **9,** 470–480; doi:10.1038/gene.2008.39; published online 29 May 2008

Keywords: *killer cell immunoglobulin-like receptor; KIR2DL5; NK cell receptors; polymorphism; population diversity*

Introduction

By interacting with specific HLA class I molecules, the killer cell immunoglobulin-like receptors (KIR) control the effector function of natural killer (NK) cells and subsets of T cells against infection and tumor transformation.[1,2] The KIR receptors can be grouped into three distinct lineages based on the configuration of the extracellular immunoglobulin (Ig)-like domains.[3] The first lineage comprises two Ig domain-containing KIRs with D0-D2 configuration (KIR2DL4 and 2DL5), the second lineage comprises three Ig domain containing KIRs with D0-D1-D2 configuration (KIR3DL1-3 and 3DS1) and the third lineage comprises two Ig domain containing KIRs with D1-D2 configuration (KIR2DL1-3 and 2DS1-5). Lineage-1 has been conserved among rhesus monkeys, gorillas, chimpanzees and humans, whereas the other two lineages revealed extensive diversity within and between species.[4] Despite clustering together into a discrete lineage with high structural homologies, the KIR2DL5 and 2DL4 differ substantially in their amino-acid sequences, genomic locations, population distributions, level of transcriptions, cell-surface expressions, ligand specificities and signaling functions.[5–9]

The *KIR2DL4* gene is present on most *KIR* haplotypes and occurs at a frequency of 100% in most populations, whereas *KIR2DL5* is variable among *KIR* haplotypes and thus differs considerably between populations in its frequency.[10–12] Two copies of the *KIR2DL5* genes have been characterized, *KIR2DL5A* and *KIR2DL5B*, that show 99.5–99.7% identity in their coding sequences.[5] *KIR2DL5A* is located in the telomeric half the *KIR* gene complex whereas the *KIR2DL5B* is located in the centromeric half.[10,13] Haplotypes carrying both *2DL5A* and *B* have been described, and thus individuals homozygous for these haplotypes may carry four copies of *KIR2DL5* sequences.[14] KIR2DL5 displays a variegated distribution on the surface of $CD56^{dim}$ NK cells, whereas the surface expression of KIR2DL4 appears to be restricted to the minority subset of KIR^{neg} $CD56^{bright}$ NK cells.[8,15] Although KIR2DL4 has been implicated in both inhibitory and activating functions,[6,16–18] KIR2DL5 appears to be solely an inhibitory receptor.[7]

Multiple sequences for both *KIR2DL5A* and *KIR2DL5B* loci have been characterized from individuals of different ethnic origins.[13] However, the nature of *KIR2DL5* sequence polymorphism within distinct ethnic populations is not known. Here, we have developed a direct DNA sequencing method and characterized the *KIR2DL5*

Correspondence: Dr R Rajalingam, Department of Pathology and Laboratory Medicine, UCLA Immunogenetics Center, David Geffen School of Medicine at UCLA, University of California at Los Angeles, 1000 Veteran Avenue, Room no. 1–536, Box: 951652, Los Angeles, CA 90095-1652, USA.
E-mail: rrajalingam@mednet.ucla.edu
Received 19 March 2008; revised 23 April 2008; accepted 25 April 2008; published online 29 May 2008

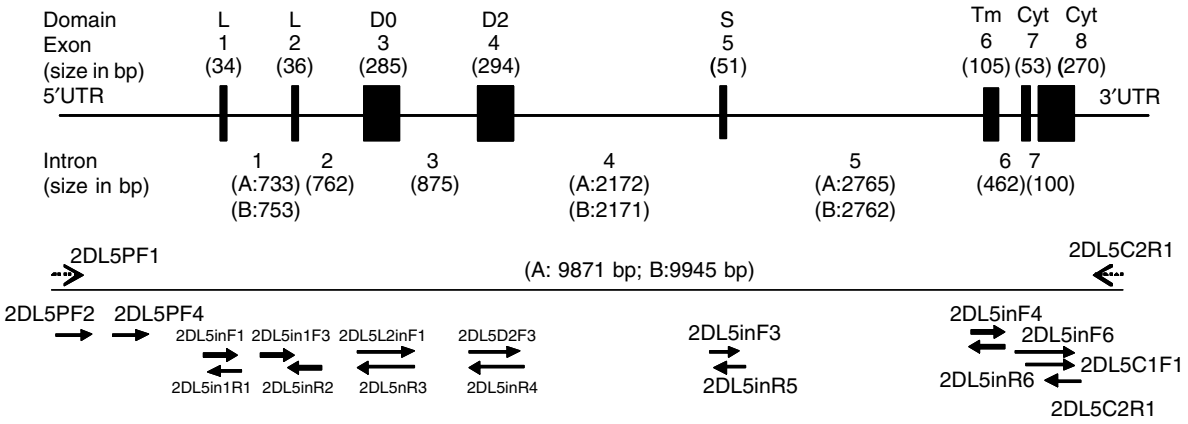

Figure 1 Schematic representation of the strategy used for sequencing the *KIR2DL5* gene. The exon (dark boxes) and intron (lines connecting the dark boxes) organization of the *KIR2DL5* gene is depicted at the top. The size of each exon and intron is provided in parenthesis; only three introns (1, 4 and 5) differ between *KIR2DL5A* and *B*. The complete *KIR2DL5* gene was PCR amplified using a forward primer 2DL5PF1 (anneals at 5′UTR, ~670 bp upstream of the start codon) and a reverse primer 2DL5C2R1 (anneals at 3′UTR, ~262 bp downstream of the stop codon). The amplification yielded products of 9872 bp (*2DL5A*) or 9950 bp (*2DL5B*). The annealing location and direction of the sequencing primers is depicted by solid arrows at the bottom, and their sequences are listed in Table 1.

gene polymorphism in 248 unrelated individuals carrying this gene from four distinct ethnic populations. Further, we determined the link between *KIR2DL5* alleles and presence and absence of other *KIR* genes.

Results

Discovery of nine novel KIR2DL5 sequences

To determine the nature of *KIR2DL5* gene polymorphism in four distinct ethnic populations, we developed a direct DNA-sequencing method (Figure 1, Table 1). We detected all known *KIR2DL5* sequences in this study of 248 *KIR2DL5*-positive unrelated individuals with the exception of *KIR2DL5B*003, B*007* and *B*009*. Furthermore, we discovered nine novel *KIR2DL5* sequences in this study (Figure 2a and b). Four of the novel sequences (*2DL5A*00102, *00103, *00104* and **00105*) differed from one another by a single synonymous substitution located either in exon-3 (D0 domain) or exon-4 (D2 Domain), and their predicted amino-acid sequences were identical. Except for *2DL5A*00103* that differed by a single nucleotide at position -177, the three other new *2DL5A*001* variants were identical to *2DL5A*0010101* in the promoter and intron-1 regions (Figure 2b).

The remaining five novel sequences belong to *2DL5B* types. *KIR2DL5B*00801* differs from **00802* by a single synonymous substitution at nucleotide position 582, and thus amino-acid sequences of these alleles are identical, and differ from their closest sequence *2DL5A*001* by a single amino-acid valine residue at position-16 in the leader sequence, which is conserved in all *2DL5B* sequences (Figure 2a). *KIR2DL5B*00801* also differs from **00802* at position 882 in the intron-1 region. Both of these alleles differ substantially from the corresponding regions of *2DL5A*001* (Figure 2b). The new allele *2DL5B*00602* differs from the known *2DL5B*00601* sequence by two synonymous substitutions located in exon-5 (stem) and exon-7 (cytoplasmic tail). However, within the promoter and intron-1 regions, the *2DL5B*00602* differs substantially from *2DL5B*00601* and displays the highest homology with the expressed allele *2DL5B*003* (Figure 2b). As seen in all *2DL5A* sequences and *2DL5B*003*, an acute myeloid leukemia gene 1 (AML1) site in the promoter region (also known as CBP, CBFα, PEBP2 or RUNX),[19,20] the factor that has been implicated for *KIR2DL5* expression is intact in *2DL5B*00602* alleles, and therefore this allele is likely to be expressed on the cell surface. As the *2DL5B*00602* is a rare allele (identified in just one of the 248 *KIR2DL5*-positive individuals) and lymphocytes from this individual were not available, we could not confirm the cell surface expression of 2DL5B*00602.

The *2DL5B*011* differs from its closest sequence *2DL5B*00601* by a novel nonsynonymous substitution at amino-acid position 284 located in the cytoplasmic tail, and by a single nucleotide in the intron-1 region. The *2DL5B*010* differs from its closest sequence *2DL5B*002* by a nonsynonymous substitution at amino-acid position-1 located in the leader peptide, and by nine nucleotides in the promoter and intron-1 regions. Three samples carrying both the *2DL5A* and *B* were found to carry more than two KIR2DL5 sequences, and potentially carrying some novel substitutions indicating the existence of additional new *2DL5* sequences (*N5*, *N10* and *N12*). However, we were unable to separate *2DL5* sequences from these samples.

It is interesting to note that the amino-acid sequences of 2DL5A*001, 2DL5B*006, 2DL5B*008 and 2DL5B*011 are identical in the extracellular segment that comprise the D0-domain, D2-domain and stem region (Figure 2a). Similarly, 2DL5A*005, 2DL5B*002, 2DL5B*009 and 2DL5*010 are identical in their extracellular segments. KIR2DL5B*003 and 2DL5B*007 encode identical extracellular segments, and 2DL5B*004 encodes a unique extracellular segment. Therefore, all known KIR2DL5 sequences produce only four types of D0-D2-stem domains (termed ectotypes) that differ from one another by one or two amino-acid substitutions (Figure 2a).

Three distinct lineages of KIR2DL5 sequences

To assess the structural relationship between different *KIR2DL5* sequences, we constructed a phylogenetic tree from a nucleotide sequence alignment of all *KIR2DL5*

a

Domain	Leader		D0					D2			Stem	Tm	Cyto	
Codon	-16	-1	4	16	79	86	95	152	173	174	210	215	284	345
Nuclotide	16	61	75	109	300	321	346	517	582	583	693	707	913	1098
*A*0010101*	ATC(I)	A--(-)	GGT(G)	GCT(A)	CCA(P)	TCA(S)	GTG(V)	AAT(N)	TTC(F)	GGC(G)	TCC(S)	CGC(P)	GTT(V)	GCT(A)
*A*0010102*	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)
A*00102	---(-)	---(-)	---(-)	---(-)	--G(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)
A*00103	---(-)	---(-)	---(-)	---(-)	---(-)	--G(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)
A*00104	---(-)	---(-)	--A(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)
A*00105	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	--T(-)	---(-)	---(-)	---(-)	---(-)	---(-)
*A*0050101*	---(-)	---(-)	--A(-)	---(-)	--G(-)	--G(-)	---(-)	G--(D)	---(-)	A--(S)	---(-)	---(-)	---(-)	---(-)
*A*0050102*	---(-)	---(-)	--A(-)	---(-)	--G(-)	--G(-)	---(-)	G--(D)	---(-)	A--(S)	---(-)	---(-)	---(-)	---(-)
*B*0020101*	G--(V)	---(-)	--A(-)	---(-)	--G(-)	--G(-)	---(-)	G--(D)	---(-)	A--(S)	---(-)	---(-)	---(-)	---(-)
*B*0020102*	G--(V)	---(-)	--A(-)	---(-)	--G(-)	--G(-)	---(-)	G--(D)	---(-)	A--(S)	---(-)	---(-)	---(-)	---(-)
*B*0020103*	G--(V)	---(-)	--A(-)	---(-)	--G(-)	--G(-)	---(-)	G--(D)	---(-)	A--(S)	---(-)	---(-)	---(-)	---(-)
*B*003*	G--(V)	C--(P)	---(-)	---(-)	---(-)	---(-)	A--(M)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)
*B*004*	G--(V)	---(-)	---(-)	--T(T)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	--C(-)
*B*00601*	G--(V)	C--(P)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)
B*00602	G--(V)	C--(P)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	--G(-)	---(-)	---(-)	--C(-)
*B*007*	G--(V)	---(-)	---(-)	---(-)	---(-)	---(-)	A--(M)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)
B*00801	G--(V)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)
B*00802	G--(V)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	--T(-)	---(-)	---(-)	---(-)	---(-)	---(-)
*B*009*	G--(V)	---(-)	--A(-)	---(-)	--G(-)	--G(-)	---(-)	G--(D)	---(-)	A--(S)	---(-)	-T-(L)	---(-)	---(-)
B*010	G--(V)	C--(P)	--A(-)	---(-)	--G(-)	--G(-)	---(-)	G--(D)	---(-)	A--(S)	---(-)	---(-)	---(-)	---(-)
B*011	G--(V)	C--(P)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	---(-)	A--(I)	---(-)

b

Nucleotide	Promoter																	
	-602	-521	-421	-376-319	-263	-239	-215	-208	-177	-159	-154	-116	-104	-97	-84	-27	-23	-10
*A*0010101*	G	G	T	*	G	G	C	C	C	A	C	C	G	G	A	C	C	C
*A*0010102*	.	-	-	*	-	-	-	-	-	-	-	-	-	-	-	-	-	-
A*00102	-	-	-	*	-	-	-	-	-	-	-	-	-	-	-	-	-	-
A*00103	-	-	-	*	-	-	-	-	T	-	-	-	-	-	-	-	-	-
A*00104	-	-	-	*	-	-	-	-	-	-	-	-	-	-	-	-	-	-
A*00105	-	-	-	*	-	-	-	-	-	-	-	-	-	-	-	-	-	-
*A*0050101*	-	-	-	*	-	-	-	-	-	-	-	-	-	-	-	-	-	-
*A*0050102*	-	-	-	*	-	-	-	-	-	-	-	-	A	-	-	-	-	-
*B*0020101*	.	-	-	#	-	-	T	-	-	-	T	-	A	A	G	-	T	-
*B*0020102*	.	.	.	.	.	-	T	-	-	-	T	A	A	A	G	-	T	-
*B*0020103*	.	-	-	#	-	-	T	-	-	-	T	-	A	A	G	-	T	-
*B*003*	T	A	C	@	A	T	-	T	-	G	-	-	A	-	-	T	-	T
*B*004*	-	-	-	#	-	-	T	-	-	-	T	-	A	A	G	-	T	-
*B*00601*	.	-	-	#	-	-	T	-	-	-	T	-	A	A	G	-	-	-
B*00602	T	A	C	@	A	T	-	T	-	G	-	-	A	-	-	T	-	T
*B*007*	.	-	-	#	-	-	T	-	-	-	T	-	A	A	G	-	T	-
B*00801	-	-	-	#	-	-	T	-	-	-	T	-	A	A	G	-	T	-
B*00802	-	-	-	#	-	-	T	-	-	-	T	-	A	A	G	-	T	-
*B*009*	.	.	.	.	.	-	T	-	-	-	T	A	A	A	G	-	T	-
B*010	-	-	-	#	-	-	T	-	-	-	T	-	A	A	G	-	-	-
B*011	-	-	-	#	-	-	T	-	-	-	T	-	A	A	G	-	-	-

Nucleotide	Intron 1																							
	66	68	141	277	283	286	296	309	361	391	456-475	488	499-615	630	702	712	729	764	773	776	795	855	857	882
*A*0010101*	T	C	T	G	C	A	G	A	G	T	*	T	*	T	G	G	G	T	A	A	C	A	T	T
*A*0010102*	-	-	-	-	-	-	-	-	-	-	*	-	*	-	-	-	-	-	-	-	-	-	-	-
A*00102	-	-	-	-	-	-	-	-	-	-	*	-	*	-	-	-	-	-	-	-	-	-	-	-
A*00103	-	-	-	-	-	-	-	-	-	-	*	-	*	-	-	-	-	-	-	-	-	-	-	-
A*00104	-	-	-	-	-	-	-	-	-	-	*	-	*	-	-	-	-	-	-	-	-	-	-	-
A*00105	-	-	-	-	-	-	-	-	-	-	*	-	*	-	-	-	-	-	-	-	-	-	-	-
*A*0050101*	-	-	-	-	-	-	-	-	-	-	*	-	*	-	-	-	-	-	-	-	-	-	-	-
*A*0050102*	-	-	-	-	-	-	-	-	-	-	*	-	*	-	-	-	-	-	-	-	-	-	-	-
*B*0020101*	C	-	-	-	T	-	T	-	-	C	x	-	*	-	-	-	C	-	G	T	-	-	-	-
*B*0020102*	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
*B*0020103*	C	-	-	-	T	-	T	-	-	C	x	-	*	-	-	-	C	-	G	T	-	-	-	-
*B*003*	-	T	A	A	T	-	-	G	A	-	x	-	*	C	T	C	-	-	-	-	-	T	C	-
*B*004*	C	-	-	-	T	-	T	-	-	C	x	-	*	-	-	-	C	-	G	T	-	-	-	-
*B*00601*	C	-	-	-	T	-	T	-	-	C	x	C	$	-	.	.	.	.	.	.	.	.	.	.
B*00602	-	T	A	A	T	-	-	G	A	-	x	-	*	C	T	C	-	-	-	-	A	T	C	-
*B*007*	C	-	-	-	T	-	T	-	-	C	x	-	*	-	-	-	C	-	G	T	-	-	-	-
B*00801	C	-	-	-	T	-	T	-	-	C	x	-	*	-	-	-	C	-	G	T	-	-	-	-
B*00802	C	-	-	-	T	-	T	-	-	C	x	-	*	-	-	-	C	-	G	T	-	-	-	C
*B*009*	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
B*010	C	-	-	-	T	-	T	-	-	C	x	-	*	-	T	C	-	C	-	-	-	T	C	-
B*011	C	-	-	-	T	-	T	-	-	C	x	-	*	-	T	-	C	-	G	T	-	-	-	-

Figure 2 *KIR2DL5* gene polymorphism. (**a**) Polymorphic nucleotides and codons of *KIR2DL5* sequences. The predicted amino acid changes are marked in parentheses. (**b**) Polymorphic nucleotides of promoter and intron-1 regions of *KIR2DL5* alleles. A dash denotes positional identity with *2DL5A*0010101* sequence. The new sequences characterized in this study are shown in bold type. Adenine substitution at position-97 mutates the acute myeloid leukemia gene 1 (AML1) site 'tgtggt' into 'tgtagt', which is implicated as the factor that stops the transcription of certain *2DL5B* alleles.[13] # Indicates a 58 bp (AAAAAATACATGAAAAGTCTTTCATGTTAGCACAGATTTAGGCCATC TCGTGTTCGGA) insertion; @ indicates a 58 bp (TAAAAATACATGAAAAGTCTTTCATGTTAGCACAGATTTACGCATCTCGTGTTCGGA) insertion; x indicates a 20 bp (CTGGGTGTGGAGATATGGGC) insertion. $ indicates a 117 bp (AGGTGGAGATATGGGCCTGGAGGTGGAGA TATGGGCCTAGAGGTGGATATCTGGGCCTGGAGTGGACATATGGGCCTAGGATGGAGATATGGGCTTGGGGTGGAGATATGGGCCTGG) insertion. Asterisks (*) indicate the absence of these insertions. Periods (.) represent the sequence is not available. *KIR2DL2A*0010101* is identical to *2DL5A*0010102* in promoter, intron-1 and exons 1–8 regions, but differed by two nucleotide substitutions in intron-6 (not shown).

alleles covering the regions of promoter, exons 1 to 8, and intron-1. The phylogenetic tree clustered *KIR2DL5* sequences into three distinct lineages (Figure 3). Substantial sequence variations observed in the promoter and intron-1 regions potentially contributed to the split of these lineages. Trees constructed on these three distinct regions (promoter, intron-1, coding regions) are provided in Supplementary Figure 1a–c. One of the lineages comprises two *2DL5B* sequences (*2DL5B*003* and **00602*), another lineage comprises the remainder of the 12 KIR*2DL5B* sequences, and the third lineage comprises all eight *2DL5A* sequences. These three lineages also differ in frequency and in their capacity to be transcribed. The *KIR2DL5B*003/00602* lineage occurred only in African Americans at low frequency. This lineage and the lineage comprising all *KIR2DL5A* sequences carry an AML1 motif in the promoter region, suggesting that these two lineages are expressed on the

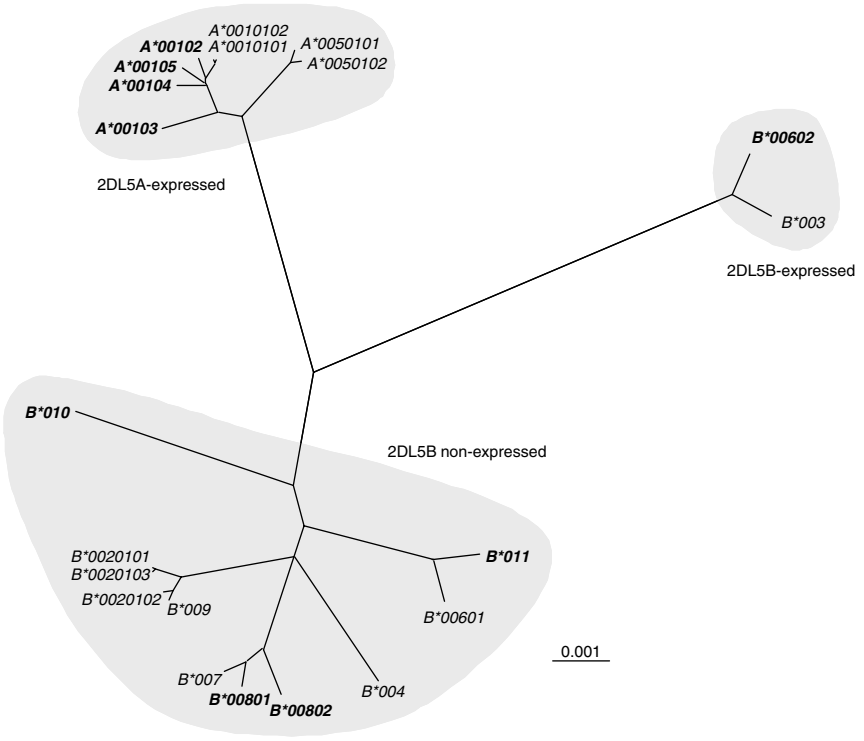

Figure 3 Unrooted neighbor-joining phylogenic tree of *KIR2DL5* sequences. The tree was constructed by the DNA Maximum Likelihood method with molecular clock (DNAMLK), version 3.5c.[21] The alignment used to construct this tree included sequences covering the region ~670 bp upstream of the start codon, all exons and intron-1. The unique insertion of 117 bp in intron-1 region of *KIR2DL5B*00601* was removed for this tree analysis. The new sequences characterized from this study are shown in bold type. The trees for individual regions (promoter, intron-1 and coding regions) are shown in the Supplementary Figure 2.

Table 1 Oligonucleotide primers used to analyze *KIR2DL5* polymorphism

Name	*Sequence (5′–3′)*	*Annealing site*	*Purpose*
2DL5PF1	GTTGGGACTCTGGAATCTTGG	Promoter region	LT-PCR
2DL5PF2	GGAATCTTGGGTCATGAGAC	Promoter region	Sequencing of promoter region
2DL5PF4	CACGTTAGCACAGATTTTA	Promoter region	Sequencing of promoter & exon 1 regions
2DL5inF1	GAGCCTGCGTACGTCACC	Promoter region	Sequencing of exon 1 & intron 1
2DL5in1R1	GAGTCTTGGTAGCCAGGC	Intron 1	Sequencing of intron 1
2DL5In1F3	GCCTGGCTACCAAGACTC	Intron 1	Sequencing of exon 2
2DL5inR2	GAACTGTGGGCTGAGCAC	Intron 2	Sequencing of exon 2
2DL5L2inF1	CCAAGTGTGGTAGGAGCC	Intron 2	Sequencing of exon 3
2DL5inR3	TAGCCATGCTGCCGACAG	Intron 3	Sequencing of exon 3
2DL5D2F3	CTATTTGGGAAACCTTCA	Exon 4	Sequencing of exon 4
2DL5inR4	CCTCCCTCACACCATGCT	Intron 4	Sequencing of exon 4
2DL5inF3	GTGTGAGGCCTCCAGCAC	Inton 4	Sequencing of exon 5
2DL5inR5	GTACTCGTCCGTCTAGCTTG	Intron 5	Sequencing of exon 5
2DL5inF4	GCCTCCCACACTGCGAGA	Inton 5	Sequencing of exon 6
2DL5inR6	CGTGAGGATACAGTTCAG	Intron 6	Sequencing of exon 6
2DL5inF6	GCACCTACGGCCTCCCGC	Inton 6	Sequencing of exons 7–8
2DL5C1F1	AATGGACCAAGAGCCTGCC	Exon 7	Sequencing of exons 7–8
2DL5C2R1	CCATGTTAAGAGGGAGCCTC	3′-UTR	LT-PCR & sequencing of exons 7–8

cell surface, whereas the other *2DL5B*-lineage is non-expressed.

Differential distribution of KIR2DL5 alleles in populations
Only *KIR2DL5A*001*, *A*005*, *B*002* and *B*008* were detected in all four ethnic populations studied (Table 2). *KIR2DL5B*003*, *2DL5B*007* and *2DL5B*009* were not detected in this study. Eight sequences (*A*00102, A*00104, A*00105, B*00802, B*004, B*00602, B*010* and *B*011*) were encountered just once, and the latter four were found only in African Americans. The highest number of *2DL5* alleles were detected in African Americans (two alleles of *2DL5A* and seven alleles of *2DL5B*), whereas the lowest number was detected in Asians (three of *2DL5A* and two of *2DL5B*). Caucasians and Asian Indians displayed seven alleles each with

Table 2 Frequency of *KIR2DL5* alleles in ethnic populations

	Caucasians (n = *250*) % (N+)	*Asian Indians* (n = *96*) % (N+)	*Africans* (n = *32*) % (N+)	*Asians* (n = *35*) % (N+)
2DL5-positive	48.4 (121)	84.4 (81)	56.3 (18)	57.1 (20)
2DL5-negative	51.6 (129)	15.6 (15)	43.7 (14)	42.9 (15)
2DL5A	34.4 (86)	64.6 (62)	12.5 (4)	42.9 (15)
*A*0010101-2*	25.2 (63)	59.4 (57)	6.3 (2)	28.6 (10)
A*00102	0.4 (1)			
A*00103	0.4 (1)			2.9 (1)
A*00104	0.4 (1)			
A*00105		1.0 (1)		
*A*0050101-2*	8.8 (22)	5.2 (5)	6.3 (2)	8.6 (3)
2DL5B	28.4 (71)	56.3 (54)	50.0 (16)	20.0 (7)
*B*0020101-3*	25.2 (63)	46.9 (45)	15.6 (5)	14.3 (5)
*B*003*				
*B*004*			3.1 (1)	
*B*00601*		2.1 (2)	18.8 (6)	
B*00602			3.1 (1)	
*B*007*				
B*00801	2.8 (7)	4.2 (4)	6.3 (2)	2.9 (1)
B*00802		1.0 (1)		
*B*009*				
B*010			3.1 (1)	
B*011			3.1 (1)	
2DL5*N5	0.4 (1)			
2DL5*N10				2.9 (1)
2DL5*N12		1.0 (1)		

Frequencies (%F) of *KIR2DL5* variants are expressed as a percentage and defined as the number of individuals having the variant (N+) divided by the number of individuals studied (n) in the study group. New alleles characterized from the present investigation are shown in bold type. *2DL5*N5, N10* and *N12* are heterozygous sequences and need to be separated by cloning methods.

different allelic components (five of *2DL5A* and two of *2DL5B* in Caucasians; three of *2DL5A* and four of *2DL5B* in Asian Indians). The most common allele in Asian Indians, Asians and Caucasians was *KIR2DL5A*00101* that occurs at frequencies of 59.4, 28.6 and 25.2% respectively, whereas it occurs only in 6.3% of African Americans. The most common allele in African Americans was *2DL5B*00601* occurring at a frequency of 18.8%. This allele was totally absent in Caucasians and Asians, but detected in two (2.1%) Asian Indians.

Comparative analysis of *KIR* profiles published in the literature revealed that the frequency of *KIR2DL5* differs substantially among ethnic populations (Figure 4a). This ethnic deviation is also seen in the distribution of *KIR2DL5A* and *KIR2DL5B* subtypes (Figure 4b). Although the African Americans and Asians carry similar frequencies of the *KIR2DL5* gene (~57%), the *2DL5B* gene was seen most frequently in African Americans whereas *2DL5A* was most common in Asians. In contrast, both the *2DL5A* and *2DL5B* loci were equally distributed in Caucasians and Asian Indians.

KIR2DL5 and its subtypes identify certain combinations of KIR genes

To assess the link between *KIR2DL5* and other *KIR* genes, we compared the frequency of distinct *KIR* genes between the groups that differ in the presence of the *KIR2DL5* gene (Figure 5). Overall, the *KIR* genes associated with group-*B* haplotypes are more frequently encountered in the *2DL5*-positive group compared to the *2DL5*-negative group. Particularly, the *KIR3DS1, 2DS3* and *2DS5* genes are present only in *2DL5*-positive individuals. *KIR2DS1, 2DL2* and *2DS2* occur more frequently in the *2DL5*-positive group than in *2DL5*-negative group. In contrast, the genes associated with the group-*A* haplotype (*KIR3DL1, 2DL3* and *2DS4*) are significantly decreased in the *2DL5*-positive group as compared with the *2DL5*-negative group.

To refine the association with *KIR2DL5*-subtypes, we compared the *KIR* gene frequencies between the *KIR2DL5A+* and *KIR2DL5B+* groups (Figure 5). Although, none of the *KIR* genes were restricted to a particular group, the frequency of individual *KIR* genes differed among these two groups. *KIR2DL3, 3DS1, 2DS5* and *2DS1* genes more frequently occurred in the *2DL5A+* group than in the *2DL5B+* group. Conversely, *KIR2DS2, 2DL2, 2DS3, 3DL1* and *2DS4* occurred more frequently in the *2DL5B+* group than in *2DL5A+* group. Individuals positive for both *KIR2DL5A* and *B* genes carry most *KIR* genes, and their *KIR* gene profile markedly differs from that of *KIR2DL5*-negative subjects.

Association between particular alleles of *KIR2DL5A/B* and the presence and absence of certain *KIR* genes was also observed (Figure 6a). At least three linked clusters were identified occurring at variable frequencies in all ethnic groups studied (Figure 6b). Except one Asian Indian sample (genotype 5, Figure 6a), all individuals carrying the *KIR2DL5A*001* allele were also found to carry *KIR3DS1, 2DS5* and *2DS1* genes (genotypes 1–4, 6–15, 21–24 and 31–41). The *KIR2DS3* gene was always found in individuals carrying either *KIR2DL5A*005* (genotypes 16–24 and 42–47) or *KIR2DL5B*002* (genotypes 25–36, 42–45 and 48–49). Majority of the individuals carrying *2LD5A*005* allele were found to carry *KIR3DS1, 2DS3* and *2DS1* genes.

Discussion

The direct DNA sequencing method developed here elucidates the sequence polymorphism of the entire *KIR2DL5* gene from its promoter region to 3′UTR region, and identifies all known as well as new alleles. However, this method will not distinguish an individual carrying three alleles (A*001, A*005 and B*002) from those carrying two alleles (A*001 and B*002) due to shared polymorphisms among these alleles. Additional cloning of the long-template PCR products is necessary to separate three alleles. This is a simple and robust method that can be easily adopted in any laboratory. Within our study panel of 248 *KIR2DL5*-positive unrelated individuals representing four ethnic populations, we found only 14 *KIR2DL5* sequences that differ from each other within the coding region, suggesting that *KIR2DL5* has limited polymorphism. Although the African-American group was small, comprising only 18 *KIR2DL5*-positive individuals, more *KIR2DL5* alleles were detected in this group than in other population groups. This suggests that more *KIR2DL5* polymorphisms are likely to be discovered in Africans. Consistent with this finding, extensive polymorphism was also observed at the *KIR3DL1/S1* locus in African populations[22] indicating

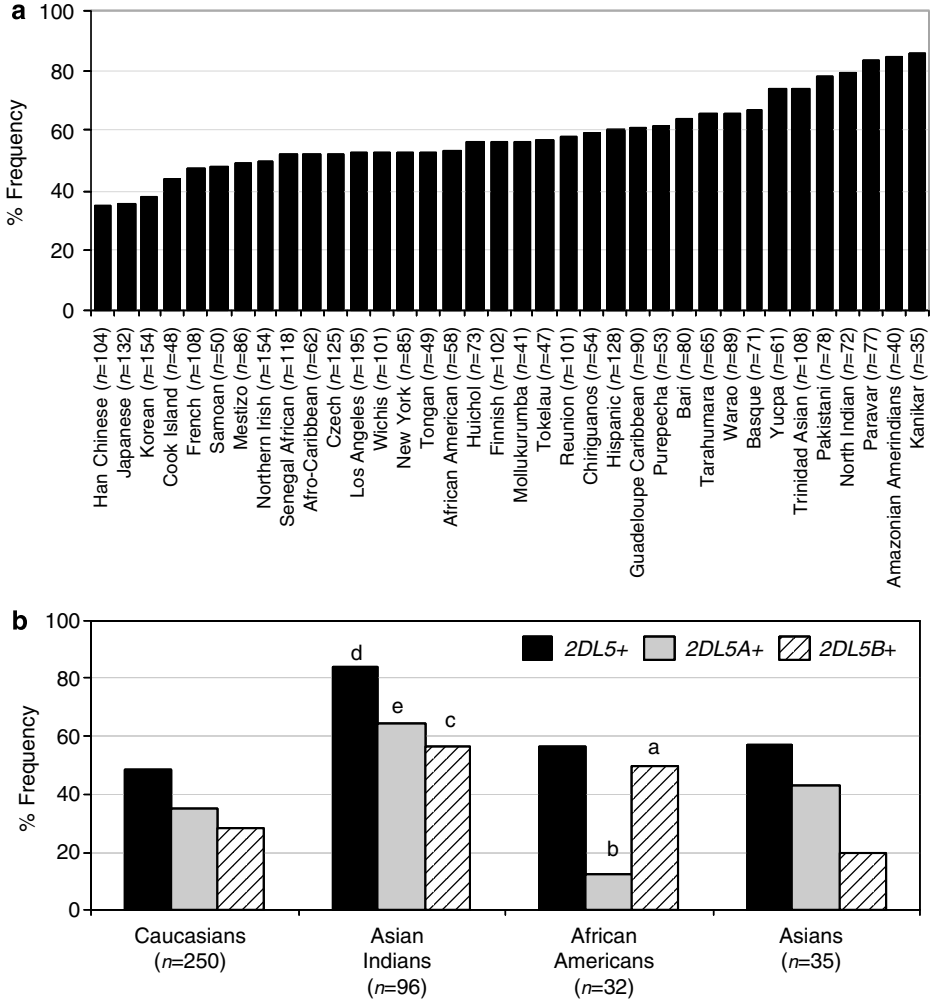

Figure 4 Ethnic diversity in the distribution of *KIR2DL5* and its variants. (**a**) Frequency of individuals carrying *KIR2DL5* in world populations published in the literature. (**b**) Frequency of *KIR2DL5A* and *KIR2DL5B* in four distinct populations determined in the present study. Only a single African American carried the expressed *2DL5B*00602* sequence and further, had an additional expressed *2DL5A*001* sequence and thus the frequency *2DL5A* is identical to the frequency of individuals carrying the expressed gene (carrying at least one of the expressed alleles *2DL5A*001*, *A*005* and *B*00602*). Similarly, the frequency of *2DL5B* corresponds to the frequency of individuals carrying the unexpressed *KIR2DL5* gene. Only the significant *p*-values compared to Caucasians are provided: $a = 0.041$; $b = 0.015$; $c = 0.0000042$; $d = 1.4 \times 10^{-8}$; $e = 6.2 \times 10^{-6}$.

that a higher level of *KIR* sequence diversity is a characteristic feature of African populations and is consistent to that seen in other genetic loci.[23,24]

Differences between *KIR2DL5* sequences were only seen at 14 nucleotide positions within the 1128 bp long coding region (Figure 2a), of which only seven were nonsynonymous substitutions causing amino acid change. The substitutions are scattered over the coding regions; none affected the predicted ligand-binding loops, stem/transmembrane regions, or ITIM motifs (Supplementary Figure 2). Moreover, all *KIR2DL5* sequences encoded only four distinct 'ectotypes' that differ from each other by just 1–2 amino-acid substitutions in the extracellular Ig-like domains and stem region. This is very different from the scenario seen in *KIR3DL1*, which has been shown to be highly polymorphic with enriched amino-acid substitutions in the ligand-binding loops.[22,25] Confined variations in the ligand-binding sites of KIR3DL1 have been considered to be driven by coevolving polymorphic HLA-B and HLA-A ligands.[22,26] The conserved nature of *KIR2DL5* suggests the possibility that it recognizes an invariant determinant. It is interesting to note that KIR2DL4 that displays limited polymorphism,[27–29] and shares high structural homologies to KIR2DL5, binds to the conserved non-classical HLA-G molecule.[6,9]

The known regulatory elements in the promoter region are conserved in all *KIR2DL5* sequences.[13] Of the 21 distinct *KIR2DL5* sequences, eight belong to *2DL5A* and the remaining 13 belong to *2DL5B*. Eight *KIR2DL5A* alleles encode only two distinct proteins (2DL5A*001 and 2DL5A*005). Although 13 *KIR2DL5B* alleles are predicted to encode nine distinct proteins, the mutation in the AML1 motif causes most alleles to be untranscribed. Thus only 2DL5B*003 and 2DL5B*00602 are likely to be expressed on the cell surface. *KIR3DL3*, another unexpressed gene, is also highly polymorphic.[30,31] *KIR3DL3* and *KIR2DL5B* share several common features including their location in the centromeric half of the *KIR* gene complex, conservation in hominoids and they are

predicted to encode inhibitory receptors of unknown ligand specificity.[4,10] However, it is not clear what drives the sequence polymorphism of these unexpressed loci. One possibility is that these null-loci may be expressed in certain physiological or pathological conditions, and thus are subjected to natural selection.[32]

Individuals carrying the *KIR2DL5* gene vary substantially among populations ranging in frequency from 35 to 85% (Figure 4a). Owing to the allelic variations, only a subset of these individuals carries a functional *KIR2DL5* gene (Figure 4b). Notably, 83% of the African Americans, who carry a *2DL5* gene do not express this inhibitory receptor as they carry only the unexpressed *2DL5B* variants. Analogous to this scenario, *KIR3DL1*004*[26,33] and *KIR2DL2*004*,[34] alleles of other inhibitory *KIR* genes lacking cell surface expression, are reported at high frequencies in African populations indicating that some inhibitory *KIR* genes may be negatively selected in Africans. In contrast, the majority of the *KIR2DL5*-positive Asian individuals carry an expressed variant.

A stretch of 14 kb enriched with L1 repeats at the upstream of *KIR2DL4* divides the *KIR* haplotype into two halves.[10] *KIR3DL3* at the 5′-end and *KIR3DP1* at the 3′-end mark the centromeric half, whereas *KIR2DL4* at the 5′-end and *KIR3DL2* at the 3′-end mark the telomeric half. These four end-marking framework genes are present on all *KIR* haplotypes and thus occur 100% in all populations.[11] *KIR2DL5A* and *2DL5B* genes are located in the telomeric and centromeric halves respectively, and their alleles display strong linkage with certain activating *KIR* genes positioned in close proximity. At least three ''linked-clusters'' were observed in all

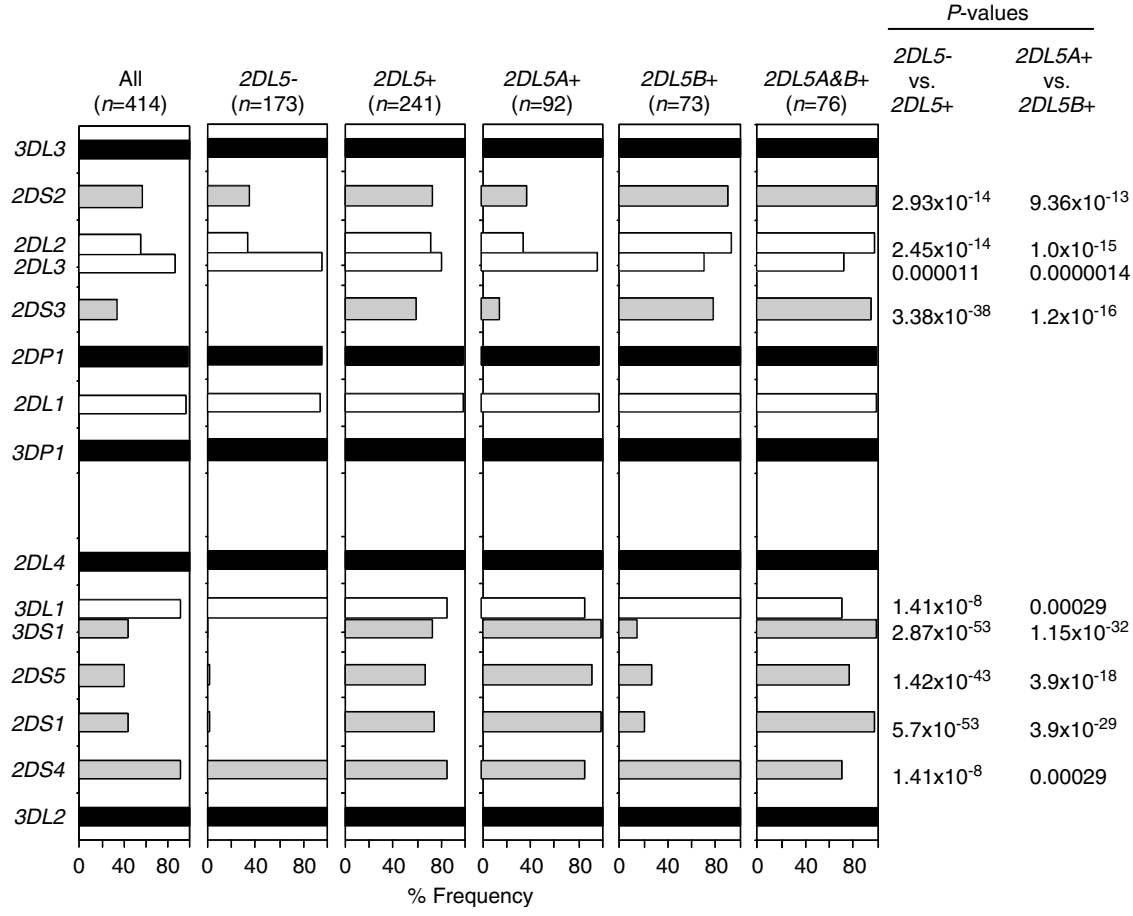

Figure 5 Link between *KIR2DL5* and other *KIR* gene content. Frequency (%F) of each genotype is expressed as a percentage and defined as the number of individuals having the genotype (N+) divided by the number of individuals studied (n) in the group. The framework genes and pseudogenes are depicted in black bars; the inhibitory *KIR*s are shown in white bars, whereas the activating *KIR*s are shown in gray bars.

Figure 6 Link between *KIR2DL5*-alleles and other *KIR* gene content. (**a**) The *KIR* genotypes and their population frequencies. A shaded box indicates the presence of a gene and an unshaded box indicates the absence of a gene. The allele types are provided within the shaded boxes. Frequency (%F) of each genotype is expressed as a percentage and defined as number of individuals having the genotype (N+) divided by the number of individuals studied (n) in the population group. The data from eight *KIR2DL5*-positive Hispanic individuals is not listed separately, but has been included in the 'all' group. The framework genes are marked in dark boxes. Owing to shared polymorphism among A*001, A*005 and B*002 alleles, the present typing method does not distinguish an individual carrying three alleles (A*001, A*005 and B*002) from those carrying two alleles (A*001 and B*002), and thus the genotypes positive for A*001 and B*002 may further carry a third allele, A*005. (**b**) Three prominent *KIR2DL5*-linked gene clusters and their population frequencies are provided. Only the significant *P*-values compared to Caucasians are provided: $a = 0.014$; $b = 0.00018$; $c = 5.6 \times 10^{-8}$. #The significance of linkage disequilibrium (LD) estimated by two-tailed Fisher Exact probabilities for each cluster is given.

a

Genotype	Centromeric half: 3DL3	2DS2	2DL2/L3	2DL5B	2DS3	2DP1	2DL1	3DP1	2DL4	Telomeric half: 3DL1/S1	2DL5A	2DS5	2DS1	2DS4	3DL2	All (*n*=421) %	All (N+)	Caucasians (*n*=250) %	Caucasians (N+)	Asian Indians (*n*=96) %	Asian Indians (N+)	Africans (*n*=32) %	Africans (N+)	Asians (*n*=35) %	Asians (N+)
1			3							L/S	00101					9.7	(41)	7.2	(18)	15.6	(15)	3.1	(1)	20.0	(7)
2			3							L/S	00102					0.2	(1)	0.4	(1)						
3			2/3							L/S	00101					4.3	18	5.6	(14)	2.1	(2)			5.7	(2)
4			2/3							L/S	00103					0.2	(1)							2.9	(1)
5			2/3							L	00101					0.2	(1)			1.0	(1)				
6			3							L/S	00101					1.0	(4)	0.4	(1)	3.1	(3)				
7			2/3							L/S	00101					0.2	(1)	0.4	(1)						
8			2/3							S	00101					0.2	(1)	0.4	(1)						
9			2/3							S	00101, 00103					0.2	(1)								
10			3							S	00101					1.9	(8)	1.2	(3)	4.2	(4)			2.9	(1)
11			2/3							L/S	00101					0.2	(1)								
12			2							L/S	00101					0.2	(1)	0.4	(1)						
13			2/3							S	00101					0.2	(1)			1.0	(1)				
14			2							L/S	00101					0.5	(2)	0.8	(2)						
15			2/3							L/S	00101					0.2	(1)								
16			3							L/S	00501					0.2	(1)	0.4	(1)						
17			2/3							S	00501					0.2	(1)								
18			2/3							L/S	00501					0.7	(3)	1.2	(3)						
19			3							L/S	00501					0.7	(3)	0.4	(1)			3.1	(1)	2.9	(1)
20			3							S	00501					0.2	(1)							2.9	(1)
21			3							L/S	00101, 00501					0.2	(1)	0.4	(1)						
22			2/3							S	00101, 00501					0.5	(2)	0.8	(2)						
23			3							S	00101, 00501					0.2	(1)			1.0	(1)				
24			3							L/S	00105, 00501					0.2	(1)			1.0	(1)				
25			2/3	00201						L/S						1.9	(8)	2.0	(5)					8.6	(3)
26			2/3	00201						L						6.2	(26)	5.2	(13)	7.3	(7)	9.4	(3)	2.9	(1)
27			2	00201						L						3.8	(16)	3.2	(8)	7.3	(7)	3.1	(1)		
28			2/3	00201						L						0.5	(2)			2.1	(2)				
29			2/3	00201						L						0.2	(1)	0.4	(1)						
30			3	00201						L/S						0.2	(1)	0.4	(1)						
31			2/3	00201						L/S	00101					5.5	(23)	3.6	(9)	13.5	(13)				
32			2/3	00201						L/S	00104					0.2	(1)	0.4	(1)						
33			2	00201						S	00101					1.7	(7)	1.6	(4)	3.1	(3)				
34			2/3	00201						S	00101					2.1	(9)	1.2	(3)	6.3	(6)				
35			2	00201						L/S	00101					1.7	(7)	0.8	(2)	5.2	(5)				
36			2	00201						L/S	00103					0.2	(1)	0.4	(1)						
37			2/3	00601						L/S	00101					0.2	(1)					3.1	(1)		
38			3	00601						S	00101					0.2	(1)			1.0	(1)				
39			2	00801						S	00101					0.2	(1)			1.0	(1)				
40			2/3	00801						L/S	00101					0.2	(1)	0.4	(1)						
41			2/3	00802						S	00101					0.2	(1)			1.0	(1)				
42			2/3	00201						L/S	00501					0.5	(2)			2.1	(2)				
43			2	00201						L/S	00501					0.5	(2)	0.8	(2)						
44			3	00201						L/S	00501					2.9	(12)	4.4	(11)					2.9	(1)
45			2	00201						S	00501					0.5	(2)	0.8	(2)						
46			2/3	00601						L/S	00501					0.2	(1)			1.0	(1)				
47			2/3	00602						L/S	00501					0.2	(1)					3.1	(1)		
48			2	00201,00601						L						0.2	(1)					3.1	(1)		
49			2	00201,00801						L						0.2	(1)	0.4	(1)						
50			2/3	004						L						0.2	(1)					3.1	(1)		
51			3	00601						L						0.2	(1)					3.1	(1)		
52			3	00801						L						0.2	(1)					3.1	(1)		
53			2	00601						L						0.5	(2)					6.3	(2)		
54			2/3	00601						L						1.9	(8)	1.6	(4)	2.1	(2)	3.1	(1)		
55			2	00601						L/S						0.2	(1)					3.1	(1)		
56			3	00801						L						0.2	(1)	0.4	(1)						
57			2	00801						L						0.2	(1)			1.0	(1)				
58			2/3	00801						L/S						0.2	(1)							2.9	(1)
59			2/3	011						L						0.2	(1)					3.1	(1)		
60			2/3	010						L						0.2	(1)					3.1	(1)		
61			2/3	N5						S	N5					0.2	(1)	0.4	(1)						
62			2/3	N12						L/S	N12					0.2	(1)			1.0	(1)				
63			2	N10						L	N10					0.2	(1)							2.9	(1)
64			3							L						25.7	(108)	34.0	(85)	4.2	(4)	25.0	(8)	31.4	(11)
65			2/3							L						11.4	(48)	14.0	(35)	9.4	(9)	6.3	(2)	5.7	(2)
66			2							L						1.4	(6)	2.4	(6)						
67			3							L						0.5	(2)					6.3	(2)		
68			3							L						0.5	(2)			2.1	(2)				
69			2/3							L						0.5	(2)	0.8	(2)						
70			3							L						0.2	(1)					3.1	(1)		
71			2/3							L						0.2	(1)					3.1	(1)		
72			3							L/S						0.2	(1)							2.9	(1)
73			2							L						0.2	(1)	0.4	(1)						
74			2							L/S						0.2	(1)							2.9	(1)

b

Linkage clusters

Centromeric half	Telomeric half	#LD	Caucasians (*n*=250) % (N+)	Asian Indians (*n*=96) % (N+)	African Americans (*n*=32) % (N+)	Asians (*n*=35) % (N+)
	3DS1 – 2DL5A*001 – 2DS5 – 2DS1	p=1.2x10^{-112}	26.4 (66)	59.4 (57)[c]	6.3 (2)[a]	31.4 (11)
	3DS1 – 2DL5A*005 – 2DS3 – 2DS1	p=4.7x10^{-11}	9.2 (23)	3.1 (3)	6.3 (2)	8.6 (3)
2DL5B*002 – 2DS3		p=3.7x10^{-80}	25.6 (64)	46.9 (45)[b]	15.6 (5)	14.3 (5)

ethnic groups but at different frequencies, indicating that natural selection is probably affecting the entire group of *KIR* genes rather than the individual loci involved in these clusters. Two of the linked clusters carry a *KIR2DS3* gene in combination with either *2DL5A*005* in the telomeric half or *2DL5B*002* in the centromeric half, and thus two copies of the *KIR2DS3* gene were possibly present on the same haplotype, a situation previously observed by family segregation analysis and sequencing the inter-loci PCR amplicons.[35,36]

By simple subtyping of *KIR2DL5A* and *2DL5B* loci, one can predict the content of activating *KIR* genes. For instance, individuals carrying *2DL5A* (in the absence of *2DL5B*) will most likely carry *KIR3DS1, 2DS1* and *2DS5* genes. Individuals carrying *2DL5B* (in the absence of *2DL5A*) will most likely carry *KIR2DS3* and *2DS2*. Individuals carrying both *KIR2DL5A* and *2DL5B* will carry all five of these activating *KIR* genes. Conversely, individuals negative for *2DL5* will lack these five activating *KIR* loci. Consequently, a simple typing system distinguishing *2DL5A* (particularly the common allele *A*001*) and *2DL5B* (particularly the common allele *B*002*) can be a powerful screening test to assess the content of most variable activating *KIR* genes that have been implicated in human disease[37–39] and in the outcome of hematopoietic stem cell transplantation.[40–42]

Materials and methods

Samples

Genomic DNA samples from 250 Caucasian blood donors from the National Marrow Donor Program (NMDP), 96 Asian Indians from the All India Institute of Medical Sciences, New Delhi, 32 African Americans and 35 Asians (Korean, Vietnamese, Japanese and Filipino) from the UCLA International DNA exchange program were included in this study. The study was reviewed and approved by appropriate Institutional Review Boards. The DNA samples were isolated from peripheral blood samples using the QIAamp blood kit (Qiagen, Valencia, CA). Quality and quantity of DNA was determined by UV spectrophotometry and the concentration was adjusted to 100 ng/μl.

The presence and absence of known *KIR* genes was determined using our recently described gene-specific PCR typing system.[43] A total of 240 individuals (121 Caucasians, 81 Asian Indians, 18 African Americans and 20 Asians) identified as *KIR2DL5*-positive were further characterized for sequence polymorphism of the *KIR2DL5* gene. Furthermore, eight Hispanic individuals carrying the *KIR2DL5* gene were included in the *KIR2DL5*-sequencing analysis. The DNA samples from blood donors RR and WC in whom the *KIR2DL5B*002* and **003* were originally isolated,[5,13] were included as controls for the direct sequencing method.

Direct DNA sequencing method for KIR2DL5 gene

To determine the nature of *KIR2DL5* gene polymorphism, we developed a direct DNA-sequencing method. The strategy includes the amplification of the complete *KIR2DL5* gene using gene-specific primers and direct sequencing of the PCR amplicons (without cloning) following enzymatic purification. We have recently used a similar approach successfully to determine the polymorphism of *KIR3DS1* and *2DS3* genes.[44]

KIR2DL5 gene was amplified from the promoter region to 3′untranslated region (3′UTR) using a high fidelity Expand Long Template PCR system (Roche Applied Science, Indianapolis, IN). A *KIR2DL5*-specific forward primer 2DL5PF1 recognizing the promoter region (at ~670 bp upstream of the translation-start codon) and a reverse primer 2DL5C2R1 recognizing the 3′UTR (at ~262 bp downstream of the translation-stop codon) were used to amplify both *KIR2DL5A* and *B* genes. The PCR amplifications were performed in a reaction volume of 30 μl with a final concentration of 1 × Expand Long Template buffer-3 (2.75 mM $MgCl_2$ and detergents, Roche Applied Science), 500 μM of each deoxynucleotide triphosphate (dNTPs) (Applied Biosystems, Foster City, CA), 0.3 μM of each forward and reverse primers, 2.25 unit Expand Long Template enzyme mix (Roche Applied Science) and 200 ng genomic DNA. The following thermal cycling was performed in a ABI 9700 GeneAmp PCR system (Applied Biosystems): initial denaturation for 2 min at 92 °C, 10 cycles at 92 °C for 10 s, 60 °C for 30 s and 68 °C for 10 min, 22 cycles at 92 °C for 15 s, 58 °C for 30 s and 68 °C for 10 min, and a final extension at 68 °C for 20 min. Two μl of PCR amplified products were electrophoresed on a 1% agarose gel supplemented with 0.1 μg/ml ethidium bromide, and examined for the single band of ~10 kb size.

The *KIR2DL5* gene amplicons were purified from unincorporated primers and dNTPs by digesting with ExoSAP-IT Exonuclease-I according to the manufacturer's instructions (USB Corporation, Cleveland, OH). The purified PCR products were used as template in the sequencing reactions. Sequencing was performed using BigDye terminator V1.1 cycle sequencing kit (Applied Biosystems) and ABI PRISIMTM 3100 capillary sequencer (Applied Biosystems) as we described recently.[44] All coding exons were sequenced in both directions using primers that were generally annealed at the intron regions (Figure 1). The promoter region and intron-1 were sequenced to determine *KIR2DL5A* and *B* types based on two unique insertions in the *2DL5B* gene, a 58 bp insertion in the promoter region and a 20 bp insertion in the intron-1 region.

Cloning of KIR2DL5 gene

The samples indicative of carrying novel mutations on direct sequencing analysis of *KIR2DL5* gene were subjected to cloning to separate the novel alleles from the known sequences. The *KIR2DL5* gene amplicons were purified by High Pure PCR template preparation kit (Roche Applied Science). The purified PCR fragments were polished, phosphorylated and then ligated into the cosmid vector (Expand vector I) by the blunt-end cloning method according to the manufacturer's instructions (Roche Applied Science). The vectors were packaged into λ bacteriophages, and then infected into *E. Coli* DH5α. The positive clones selected by ampicillin resistance were used to isolate *2DL5*-carrying vectors using Qiaprep Spin Miniprep kit (Qiagen). Nucleotide sequences were determined using the BigDye terminator V1.1 cycle sequencing kit (Applied Biosystems) and ABI PRISIMTM 3100 capillary sequencer (Applied Biosystems). For each variant, at least five clones derived from two

different PCR amplifications were characterized, a strategy adopted to eliminate the PCR-artifacts and errors.

Sequence analysis and statistical methods
The Assign SBT v3.5.1 software (Conexio Genomics, Australia) was used to combine both forward and reverse sequence files to inspect and edit the electropherograms. The Assign program assigned the alleles by comparing the test sequences with a library of known sequences downloaded from the Immuno Polymorphism Database (IPD) (http://www.ebi.ac.uk/ipd/kir). Sequence alignment and comparison were performed using ClustalX v1.83[45] and BioEdit v7.0.5.3 (by Tom Hall: http://www.mbio.ncsu.edu/BioEdit/page2.html) software. New sequences were submitted to the GenBank and IPD databases,[46] and were assigned the following accession numbers and official names: *KIR2DL5A*00102* (**EF473257**), *KIR2DL5A*00103* (**EF473258**), *KIR2DL5A*00104* (**EF473261**), *KIR2DL5A*00105* (**EF473264**), *KIR2DL5B*00602* (**EF473266**), *KIR2DL5B*00801* (**EF473256**), *KIR2DL5B*00802* (**EF473262**), *KIR2DL5B*010* (**EF473259**) and *KIR2DL5B*011* (**EF473263**). The complete sequences of promoter and intron-1 regions for four known alleles were further submitted to GenBank and were assigned the following accession numbers: *2DL5B*003* (**EU450878**), *2DL5B*004* (**EU450879**), *2DL5B*0050101* (**EU489567**) and *2DL5B*0050102* (**EU489568**). Three heterozygous sequences comprising novel substitutions (indicative of additional new *2DL5* alleles) were deposited to GenBank under the accession numbers: **EF473260** (*2DL5*-New5), **EF473265** (*2DL5*-New10) and **EF473267** (*2DL5*-New12). KIR2DL5A*00103 appears to be identical to a recently characterized sequence.[47] The complete KIR genotyping, ethnic background, and source of the DNA from which the new alleles were isolated are available at the IPD databases.[46] The KIR genotyping of these DNA samples are also provided in Figure 6a.

The frequencies of *KIR2DL5* alleles in the study populations were determined by direct counting. Differences in the distribution of *KIR2DL5* variants and the significance of linkage disequilibrium were estimated by two-tailed Fisher's Exact probabilities (p), where $P<0.05$ is considered to be statistically significant. A phylogenetic tree of *KIR2DL5* sequences was constructed by the DNA Maximum Likelihood method with molecular clock (DNAMLK) version 3.5c[21] and TREEVIEW[48] programs.

Acknowledgements

This work was supported by start-up funds from the UCLA Department of Pathology and Laboratory Medicine to Dr Rajalingam, and by funding from the National Marrow Donor Program (NMDP) and the Department of the Navy, Office of Naval Research Cooperative Agreement no. N00014-99-2-0006 and Grant no. N00014-05-1-0859 to the NMDP. We thank Damian Goodridge, Conexio Genomics, Western Australia for his help in building and optimizing *KIR* sequence libraries for the use in Assign software, John Muramoto for providing DNA samples from the UCLA International DNA Exchange Program, and Cynthia Vierra-Green and Rebecca Cullen for providing DNA samples from the NMDP repository. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the Office of Naval Research or the NMDP.

Conflict of interest

The authors declare no financial or commercial conflict of interest.

References

1 Biron CA, Nguyen KB, Pien GC, Cousens LP, Salazar-Mather TP. Natural killer cells in antiviral defense: function and regulation by innate cytokines. *Annu Rev Immunol* 1999; **17**: 189–220.
2 Cerwenka A, Lanier LL. Natural killer cells, viruses and cancer. *Nat Rev Immunol* 2001; **1**: 41–49.
3 Khakoo SI, Rajalingam R, Shum BP, Weidenbach K, Flodin L, Muir DG *et al*. Rapid evolution of NK cell receptor systems demonstrated by comparison of chimpanzees and humans. *Immunity* 2000; **12**: 687–698.
4 Rajalingam R, Parham P, Abi-Rached L. Domain shuffling has been the main mechanism forming new hominoid killer cell Ig-like receptors. *J Immunol* 2004; **172**: 356–369.
5 Vilches C, Rajalingam R, Uhrberg M, Gardiner CM, Young NT, Parham P. KIR2DL5, a novel killer-cell receptor with a D0-D2 configuration of Ig-like domains. *J Immunol* 2000; **164**: 5797–5804.
6 Rajagopalan S, Bryceson YT, Kuppusamy SP, Geraghty DE, van der Meer A, Joosten I *et al*. Activation of NK cells by an endocytosed receptor for soluble HLA-G. *PLoS Biol* 2006; **4**: e9.
7 Yusa S, Catina TL, Campbell KS. KIR2DL5 can inhibit human NK cell activation via recruitment of Src homology region 2-containing protein tyrosine phosphatase-2 (SHP-2). *J Immunol* 2004; **172**: 7385–7392.
8 Estefania E, Flores R, Gomez-Lozano N, Aguilar H, Lopez-Botet M, Vilches C. Human KIR2DL5 is an inhibitory receptor expressed on the surface of NK and T lymphocyte subsets. *J Immunol* 2007; **178**: 4402–4410.
9 Selvakumar A, Steffens U, Dupont B. NK cell receptor gene of the KIR family with two IG domains but highest homology to KIR receptors with three IG domains. *Tissue Antigens* 1996; **48**: 285–294.
10 Wilson MJ, Torkar M, Haude A, Milne S, Jones T, Sheer D *et al*. Plasticity in the organization and sequences of human KIR/ILT gene families. *Proc Natl Acad Sci USA* 2000; **97**: 4778–4783.
11 Yawata M, Yawata N, Abi-Rached L, Parham P. Variation within the human killer cell immunoglobulin-like receptor (KIR) gene family. *Crit Rev Immunol* 2002; **22**: 463–482.
12 Kulkarni S, Single RM, Martin MP, Rajalingam R, Badwe R, Joshi N *et al*. Comparison of the rapidly evolving KIR locus in Parsis and natives of India. *Immunogenetics* 2008 (in press).
13 Vilches C, Gardiner CM, Parham P. Gene structure and promoter variation of expressed and nonexpressed variants of the KIR2DL5 gene. *J Immunol* 2000; **165**: 6416–6421.
14 Gomez-Lozano N, Gardiner CM, Parham P, Vilches C. Some human KIR haplotypes contain two KIR2DL5 genes: KIR2DL5A and KIR2DL5B. *Immunogenetics* 2002; **54**: 314–319.
15 Jacobs R, Hintzen G, Kemper A, Beul K, Kempf S, Behrens G *et al*. CD56bright cells differ in their KIR repertoire and cytotoxic features from CD56dim NK cells. *Eur J Immunol* 2001; **31**: 3121–3127.
16 Kikuchi-Maki A, Yusa S, Catina TL, Campbell KS. KIR2DL4 is an IL-2-regulated NK cell receptor that exhibits limited expression in humans but triggers strong IFN-gamma production. *J Immunol* 2003; **171**: 3415–3425.

17 Faure M, Long EO. KIR2DL4 (CD158d), an NK cell-activating receptor with inhibitory potential. *J Immunol* 2002; **168**: 6208–6214.
18 Kikuchi-Maki A, Catina TL, Campbell KS. Cutting edge: KIR2DL4 transduces signals into human NK cells through association with the Fc receptor gamma protein. *J Immunol* 2005; **174**: 3859–3863.
19 Meyers S, Downing JR, Hiebert SW. Identification of AML-1 and the (8;21) translocation protein (AML-1/ETO) as sequence-specific DNA-binding proteins: the runt homology domain is required for DNA binding and protein-protein interactions. *Mol Cell Biol* 1993; **13**: 6336–6345.
20 Levanon D, Groner Y. Structure and regulated expression of mammalian RUNX genes. *Oncogene* 2004; **23**: 4211–4219.
21 Felsenstein J. Evolutionary trees from DNA sequences: a maximum likelihood approach. *J Mol Evol* 1981; **17**: 368–376.
22 Norman PJ, Abi-Rached L, Gendzekhadze K, Korbel D, Gleimer M, Rowley D *et al*. Unusual selection on the KIR3DL1/S1 natural killer cell receptor in Africans. *Nat Genet* 2007; **39**: 1092–1099.
23 Tishkoff SA, Verrelli BC. Patterns of human genetic diversity: implications for human evolutionary history and disease. *Annu Rev Genomics Hum Genet* 2003; **4**: 293–340.
24 Marth GT, Czabarka E, Murvai J, Sherry ST. The allele frequency spectrum in genome-wide human variation data reveals signals of differential demographic history in three large world populations. *Genetics* 2004; **166**: 351–372.
25 Gardiner CM, Guethlein LA, Shilling HG, Pando M, Carr WH, Rajalingam R *et al*. Different NK cell surface phenotypes defined by the DX9 antibody are due to KIR3DL1 gene polymorphism. *J Immunol* 2001; **166**: 2992–3001.
26 Single RM, Martin MP, Gao X, Meyer D, Yeager M, Kidd JR *et al*. Global diversity and evidence for coevolution of KIR and HLA. *Nat Genet* 2007; **39**: 1114–1119.
27 Schellekens J, Tilanus MG, Rozemuller EH. The elucidation of KIR2DL4 gene polymorphism. *Mol Immunol* 2008; **45**: 1900–1906.
28 Zhu FM, Jiang K, Lv QF, He J, Yan LX. Investigation of killer cell immunoglobulin-like receptor KIR2DL4 diversity by sequence-based typing in Chinese population. *Tissue Antigens* 2006; **67**: 214–221.
29 Shulse C, Steiner NK, Hurley CK. Allelic diversity in KIR2DL4 in a bone marrow transplant population: description of three novel alleles. *Tissue Antigens* 2007; **70**: 157–159.
30 Hou L, Chen M, Steiner NK, Belle I, Turino C, Ng J *et al*. Seventeen novel alleles add to the already extensive KIR3DL3 diversity. *Tissue Antigens* 2007; **70**: 449–454.
31 Jones DC, Hiby SE, Moffett A, Trowsdale J, Young NT. Nature of allelic sequence polymorphism at the KIR3DL3 locus. *Immunogenetics* 2006; **58**: 614–627.
32 Trundley AE, Hiby SE, Chang C, Sharkey AM, Santourlidis S, Uhrberg M *et al*. Molecular characterization of KIR3DL3. *Immunogenetics* 2006; **57**: 904–916.
33 Pando MJ, Gardiner CM, Gleimer M, McQueen KL, Parham P. The protein made from a common allele of KIR3DL1 (3DL1*004) is poorly expressed at cell surfaces due to substitution at positions 86 in Ig domain 0 and 182 in Ig domain 1. *J Immunol* 2003; **171**: 6640–6649.
34 VandenBussche CJ, Dakshanamurthy S, Posch PE, Hurley CK. A single polymorphism disrupts the killer Ig-like receptor 2DL2/2DL3 D1 domain. *J Immunol* 2006; **177**: 5347–5357.
35 Hsu KC, Chida S, Geraghty DE, Dupont B. The killer cell immunoglobulin-like receptor (KIR) genomic region: gene-order, haplotypes and allelic polymorphism. *Immunol Rev* 2002; **190**: 40–52.
36 Hsu KC, Liu XR, Selvakumar A, Mickelson E, O'Reilly RJ, Dupont B. 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. *J Immunol* 2002; **169**: 5118–5129.
37 Khakoo SI, Carrington M. KIR and disease: a model system or system of models? *Immunol Rev* 2006; **214**: 186–201.
38 Arnheim L, Dillner J, Sanjeevi CB. A population-based cohort study of KIR genes and genotypes in relation to cervical intraepithelial neoplasia. *Tissue Antigens* 2005; **65**: 252–259.
39 Scquizzato E, Teramo A, Miorin M, Facco M, Piazza F, Noventa F *et al*. Genotypic evaluation of killer immunoglobulin-like receptors in NK-type lymphoproliferative disease of granular lymphocytes. *Leukemia* 2007; **21**: 1060–1069.
40 Chewning JH, Gudme CN, Hsu KC, Selvakumar A, Dupont B. KIR2DS1-positive NK cells mediate alloresponse against the C2 HLA-KIR ligand group *in vitro*. *J Immunol* 2007; **179**: 854–868.
41 McQueen KL, Dorighi KM, Guethlein LA, Wong R, Sanjanwala B, Parham P. Donor-recipient combinations of group A and B KIR haplotypes and HLA class I ligand affect the outcome of HLA-matched, sibling donor hematopoietic cell transplantation. *Hum Immunol* 2007; **68**: 309–323.
42 Schellekens J, Rozemuller EH, Petersen EJ, van den Tweel JG, Verdonck LF, Tilanus MG. Activating KIRs exert a crucial role on relapse and overall survival after HLA-identical sibling transplantation. *Mol Immunol* 2008; **45**: 2255–2261.
43 Du Z, Gjertson DW, Reed EF, Rajalingam R. Receptor-ligand analyses define minimal killer cell Ig-like receptor (KIR) in humans. *Immunogenetics* 2007; **59**: 1–15.
44 Luo L, Du Z, Sharma SK, Cullen R, Spellman S, Reed EF *et al*. Chain-terminating natural mutations affect the function of activating KIR receptors 3DS1 and 2DS3. *Immunogenetics* 2007; **10**: 779–792.
45 Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 1997; **25**: 4876–4882.
46 Robinson J, Waller MJ, Stoehr P, Marsh SG. IPD—the immuno polymorphism database. *Nucleic Acids Res* 2005; **33**: D523–D526.
47 Houtchens KA, Nichols RJ, Ladner MB, Boal HE, Sollars C, Geraghty DE *et al*. High-throughput killer cell immunoglobulin-like receptor genotyping by MALDI-TOF mass spectrometry with discovery of novel alleles. *Immunogenetics* 2007; **59**: 525–537.
48 Page RD. TreeView: an application to display phylogenetic trees on personal computers. *Comput Appl Biosci* 1996; **12**: 357–358.

Supplementary Information accompanies the paper on Genes and Immunity website (http://www.nature.com/gene)