



UNIVERSIDADE ESTADUAL PAULISTA
“JÚLIO DE MESQUITA FILHO”
INSTITUTO DE BIOCÊNCIAS DE BOTUCATU – IBB



Programa de Pós-Graduação em Ciências Biológicas (GENÉTICA)

DIVERSIDADE DO GENE *KIR3DL3* EM POPULAÇÕES MUNDIAIS E ASPECTOS EVOLUTIVOS

Heloísa de Souza Andrade

Botucatu - SP
2020



UNIVERSIDADE ESTADUAL PAULISTA
“JÚLIO DE MESQUITA FILHO”
INSTITUTO DE BIOCÊNCIAS DE BOTUCATU – IBB



Programa de Pós-Graduação em Ciências Biológicas (GENÉTICA)

DIVERSIDADE DO GENE *KIR3DL3* EM POPULAÇÕES MUNDIAIS E ASPECTOS EVOLUTIVOS

Heloísa de Souza Andrade

Orientador: Erick da Cruz Castelli

Dissertação apresentada ao Instituto de Biociências,
Universidade Estadual Paulista “Júlio de Mesquita
Filho”, Campus de Botucatu, como requisito parcial
para obtenção do título de Mestre no Programa de
Pós-Graduação em Ciências Biológicas (Genética).

Botucatu - SP
2020

FICHA CATALOGRÁFICA ELABORADA PELA SEÇÃO TÉC. AQUIS. TRATAMENTO DA INFORM.
DIVISÃO TÉCNICA DE BIBLIOTECA E DOCUMENTAÇÃO - CÂMPUS DE BOTUCATU - UNESP
BIBLIOTECÁRIA RESPONSÁVEL: ROSEMEIRE APARECIDA VICENTE-CRB 8/5651

Andrade, Heloísa de Souza.

Diversidade do gene KIR3DL3 em populações mundiais e aspectos evolutivos / Heloísa de Souza Andrade. - Botucatu, 2020

Dissertação (mestrado) - Universidade Estadual Paulista "Júlio de Mesquita Filho", Instituto de Biociências de Botucatu

Orientador: Erick da Cruz Castelli

Capes: 20200005

1. Polimorfismo (Genética). 2. Haplótipos. 3. Variação genética. 4. Ciências da vida.

Palavras-chave: Ancestralidade; Evolução; Haplótipo; KIR3DL3; Polimorfismo.

DEDICATÓRIA

Aos meus pais, Hugo e Edilane.
Por todo carinho, apoio, amor. Que
sorte a minha.

AGRADECIMENTOS

Agradeço aos meus pais, Hugo e Edilane, pelo amor, carinho, cuidado e apoio em todas as etapas da minha vida. Pelas acolhidas na casa azul da roça em MG, lugar que chamo de lar e que é o meu preferido do mundo. Sem vocês nada disso seria possível. Aos meus irmãos, Thaís e Daniel, pela parceria e carinho de sempre, mesmo com os 900km entre nós, e aos meus cunhados, Daniel e Juliana. Às minhas avós, Lourdes e Cidoca, por todo carinho e aconchego. Aos meus queridos sobrinhos Felipe, Clarice e Bento e ao meu afilhado João Pedro – o sorriso de uma criança traz uma leveza que cura. Amo vocês!

À toda minha família, pelo apoio e por sempre se preocuparem comigo. Em especial aos meus tios e primos de São Paulo (tia Dica, tio Carlinhos, Amanda, Lu e Vi) por terem me abrigado diversas vezes, sempre com muito amor. Obrigada madrinha Lê e prima/comadre Marina, por toda a força e torcida.

Muito obrigada prof. Erick, pela oportunidade. Sou muito grata por todos os ensinamentos, conversas, orientações e pela (muita) paciência em me ensinar bioinformática.

Amigos do GeMBio <3: as risadas, cafés e almoços no bom prato foram fundamentais para que o meu caminho em Botucatu fosse mais leve e a saudade de casa mais branda, obrigada mesmo! Vou levar cada um de vocês em meu coração com muito carinho, desejo o melhor para todos - Andréia, Emiliana, Marília, Thálitta, Nayane, Camila, Hector, Gabi, Arielle, Amanda.

Às minhas orientadoras de IC na UnB, Silviene e Fabíola, por todos os ensinamentos e pela oportunidade que me deram durante a graduação.

Aos meus grandes amigos – Naty, Carol, Pedro e Fernando – que continuaram ao meu lado apesar da distância, rotina doida e desencontrada e dos meus esquecimentos em responder WhatsApp. Obrigada pelas conversas e por estarem sempre perto mesmo longe.

Aos amigos Troteco, pelos ótimos momentos e risadas nas minhas breves passagens por Brasília.

Ao centro budista Caminho do Diamante. As meditações e conversas foram fundamentais para meu crescimento.

A Universidade Estadual Paulista “Júlio Mesquita Filho”, campus Botucatu, em especial a Unipex, onde realizei todos os experimentos deste projeto.

Aos professores que participaram dessa trajetória e ao Programa de Pós-Graduação em Ciências Biológicas (Genética) do Instituto de Biociências de Botucatu.

A Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) e a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) – com o Processo número 2018/05863-0, pelo apoio financeiro para a realização dessa pesquisa. Obrigada pelo investimento na minha formação científica por meio da bolsa de mestrado concedida.

Aos membros da banca avaliadora, por aceitarem o convite e dedicarem tempo para contribuir com a minha formação e trabalho.

Ao Yan, obrigada por tudo. Teria sido muito difícil sem você. Obrigada pela paciência durante as crises e por ter continuado ao meu lado.

Aos álbuns do Nação Zumbi, Grande Encontro, Academia da Berlinda, Novos Baianos, Seu Pereira e de toda trilha sonora que me acompanhou nesses dois anos, principalmente durante as fases de escrita e nos finais de semana.

Por fim, gostaria de agradecer minha filha de quatro patas, Lilica, infelizmente não chegamos juntas ao final dessa jornada, mas ela está todos os dias comigo em pensamento. Obrigada por ser a melhor companheira que eu poderia ter tido, é incrível como um serzinho me ensinou a ser mais forte e resiliente. Amo você Liliqueli e sei que vamos nos reencontrar.

EPÍGRAFE

“E um homem não me define
Minha casa não me define
Minha carne não me define
Eu sou meu próprio lar”

(Triste, louca ou má – Francisco, el hombre)

RESUMO

Os receptores *killer* semelhantes à imunoglobulina (KIRs) são moléculas receptoras transmembranares expressas em células *natural killer* e T citotóxicas. Esses receptores reconhecem moléculas do sistema imunológico, principalmente moléculas do complexo principal de histocompatibilidade (HLA) de classe I na superfície de células somáticas. Os genes *KIR* são altamente polimórficos tanto do ponto de vista de número de cópias de cada gene quanto do conteúdo de sequência. *KIR3DL3* é um gene *KIR* considerado *framework*, pois está presente em todos os indivíduos em duas cópias. Aqui, abordamos a diversidade genética e os aspectos evolutivos do gene *KIR3DL3* em populações humanas de todo o mundo, usando uma abordagem de bioinformática desenvolvida para este estudo e que minimiza viés de mapeamento e erros de genotipagem. O conjunto de amostras compreende 2504 indivíduos de 26 populações distintas, provenientes do consórcio *1000 Genomes Project*, além de 216 amostras brasileiras do estado de São Paulo que foram sequenciadas utilizando sequenciamento de segunda geração. Marcadores informativos de ancestralidade foram utilizados na população brasileira para determinar o perfil de ancestralidade genômica de cada indivíduo e para se obter uma média da população total. A variabilidade genética do *KIR3DL3* é muito maior do que conhecemos atualmente, especialmente quando consideramos íntrons e sequências regulatórias. A diversidade de proteínas de *KIR3DL3* é fortemente influenciada pelo perfil de ancestralidade do indivíduo, com proteínas completas e resíduos de aminoácidos apresentando diferentes frequências entre africanos, europeus e outras populações. Nesse caso, as populações do leste da Ásia apresentam um padrão diferente das demais considerando frequências de alelos, proteínas, sequências regulatórias e perfil de seleção natural. Por esse motivo, qualquer estudo de associação com doenças envolvendo *KIR3DL3* deve ser realizado considerando a estratificação populacional. Alguns segmentos de *KIR3DL3* são altamente conservados, como os éxons que codificam o peptídeo líder e o domínio extracelular D2, importantes para a interação com HLA, enquanto outros segmentos, como a região promotora e o éxon 6 (domínio transmembranar) são bastante variáveis. Encontramos evidências de seleção balanceadora na região promotora, provavelmente mantendo sequências com diferentes ações regulatórias. Essa evidência é mais forte entre europeus e no leste asiático. Também encontramos evidências de seleção balanceadora em íntrons, que devem ser considerados como possíveis portadores de elementos regulatórios.

Palavras-chave: *KIR3DL3*, polimorfismo, haplótipo, NGS, ancestralidade, seleção natural, evolução

ABSTRACT

Killer-cell Immunoglobulin-like Receptors (KIRs) are transmembrane receptors expressed in Natural Killer cells and cytotoxic T cells. These receptors recognize molecules from the immune system, mainly molecules from the major class I histocompatibility complex on somatic cells' surface. KIR genes are highly polymorphic regarding the number of copies of each locus and their sequence content. *KIR3DL3* is a framework KIR gene present in every individual. Here we addressed *KIR3DL3* genetic diversity and evolutionary aspects in worldwide human populations, using a bioinformatics pipeline created for this study, minimizing mapping bias and genotyping errors. The samples included 2504 individuals from 26 populations (1000 Genomes Project consortium), and 216 Brazilians samples sequenced using second-generation sequencing. Ancestry informative markers were used in the Brazilian population to access the contribution of the different ancestry profiles of each individual and to obtain an average of the whole population. *KIR3DL3* genetic variability is far more significant than we currently acknowledge, especially when considering intronic segments and regulatory sequences. *KIR3DL3* protein diversity is strongly influenced by the ancestry background, with different frequencies of protein versions and amino acid residues among Africans, Europeans, and other populations. In this matter, populations from East Asia present an unusual pattern considering protein frequencies, regulatory sequences diversity, and how natural selection has shaped *KIR3DL3* diversity and haplotype frequencies. Because of that, any association study involving *KIR3DL3* should be performed concerning for population stratification. Some *KIR3DL3* segments are highly conserved, such as the exons encoding the leader peptide and the extracellular domain D2, important for HLA interaction, while other segments, such as the promoter region and exon 6 (the transmembrane domain) are highly variable. We found evidence of balancing selection in the promoter region, probably maintaining sequences with different promotion capabilities, and it is evident among Europeans and East Asians. We also find evidence of balancing selection in some large intronic segments, and they must be considered as possibly carrying regulatory elements.

Keywords: *KIR3DL3*, polymorphism, haplotype, NGS, ancestry, natural selection, evolution

LISTA DE FIGURAS

Figura 1. Mecanismo de ação das células <i>natural killer</i> (NK) denominado <i>missing-self</i> , no qual a interação dos receptores das células NK com ligantes específicos vão induzir a uma resposta. Caso haja um predomínio de sinais inibidores (principalmente com a ligação a moléculas do complexo MHC), não há ativação das células NK. Se há um predomínio de sinais ativadores e inibição dos inibidores, a célula inicia um processo de defesa. Fonte: Hinton-Sheley, Phoebe (2019).	3
Figura 2. Composição do Complexo Receptor Leucocitário (do inglês, Leukocyte Receptor Complex, LRC), presente na região 19q13.4. Dentre eles, em destaque, os genes KIR. Em cinza os genes framework, vermelho os ativadores, verde inibidores e amarelo o pseudogene não framework. Fonte: Single et al., (2008).	4
Figura 3. Representação das moléculas KIR. Em azul os domínios semelhantes a imunoglobulina (dois ou três), em vermelho os motivos de inibição baseados em tirosina (ITIM) nas caudas citoplasmáticas longas e em verde os motivos de ativação baseados em tirosina (ITAM) nas caudas curtas. Imagem disponível em http://www.ebi.ac.uk/ipd/kir/images	5
Figura 4. Estrutura geral dos genes KIR. Os éxons 1 e 2 codificam o peptídeo líder, éxons 3, 4 e 5 os domínios semelhantes a Ig, éxon 6 a haste que liga o D2 ao domínio transmembranar (TM) que é codificado pelo éxon 7 e a cauda citoplasmática é codificada pelos éxons 8 e 9. Lembrando que há alguns KIR que possuem deleção em éxon.	6
Figura 5. Nomenclatura KIR. Baseada na figura disponível em https://www.ebi.ac.uk/ipd/kir/alleles.html	6
Figura 6. Haplótipos A e B. Em cinza pode-se observar os genes framework, e em azul os demais genes KIR de cada haplótipo, os quais podem ou não estar presentes e em diferentes quantidades de número de cópia. Adaptado de Parham et al., 2005.	7
Figure 1. Principal component analysis (PCA) considering variable sites in the KIR3DL3 locus. We pruned variants with a $r^2 < 0.2$. Each superpopulation is represented by a different color: red for Africans, blue for Americans, green for Brazilians, black for East Asians, brown for Europeans, and yellow for South Asians.	51
Figure 2. Multidimension scaling plot of the F_{ST} values among each superpopulation.	52
Figure 3. Tajima's D across the KIR3DL3 gene in different worldwide populations, considering sliding windows of 250bp and a step size of 1. The x-axis the central position of a window. The y-axis represents the Tajima's D value.	53
Figure S1. Linkage Disequilibrium (LD) between pair of single nucleotide polymorphisms (SNPs) of the KIR3DL3 locus, considering genomic positions from Promoter to 3'UTR. LD plot generated by Haploview 4.2, considering variable sites with a minor allele frequency (MAF) $\geq 5\%$ and the fraction of strong LD in informative comparison set to 0.9. Areas in red color indicate strong LD (LOD score ≥ 2 , D' score =1); areas in light red or pink indicate moderate LD (LOD ≥ 2 , $D' < 1$); areas in blue indicate weak LD (LOD < 2 , $D' = 1$); and areas in white indicate very weak LD (LOD < 2 , $D' < 1$).	62
Figure S2. KIR3DL3 promoter sequences and their frequency in worldwide populations.	63
Figure S3. KIR3DL3 3'UTR sequences and their frequency in worldwide populations.	64

LISTA DE TABELAS

Table 1. KIR3DL3 encoded proteins in worldwide population samples grouped by the ancestry background (superpopulations).	54
Table 2. KIR3DL3 encoded proteins and amino acid in each polymorphic residue in worldwide population samples.	55
Table 3. KIR3DL3 polymorphic residues (minor allele frequency > 1%) and their frequency in worldwide population samples.	57
Table 4: Nucleotide diversity in each KIR3DL3 segment in worldwide populations grouped according to their ancestry background.	58
Table 5: Tajima's D in each KIR3DL3 segment in worldwide populations grouped according to their ancestry background.	59
Table 06: dN/dS ratio test of the KIR3DL3 exons in worldwide population samples.	60
Table S1. List of variants and frequency of the reference allele in worldwide populations samples.	65
Table S2. KIR3DL3 encoded proteins in worldwide population samples.	82
Table S3. KIR3DL3 encoded proteins in worldwide population samples, and each amino acid exchange. Here are listed only the proteins that do not match any of the ones described in the IPD-IMGT/KIR database.	83
Table S4. KIR3DL3 coding sequences (only exons) in worldwide population samples grouped by the ancestry background.	85
Table S5. KIR3DL3 promoter haplotypes and their frequencies.	87
Table S6. The association between KIR3DL3 promoter lineages and genomic sequences encoding specific proteins.	88
Table S7. The association between KIR3DL3 3'UTR haplotypes and genomic sequences encoding specific proteins.	89

LISTA DE ABREVIATURAS E SIGLAS

AFR	Africano, do inglês African
AIM	Marcadores Informativos de Ancestralidade, do inglês Ancestry Informative Markers
bp	Pares de base
BWA	Burrows-Wheeler Aligner
CDS	Sequências codificadoras
CNV	Variação no Número de Cópias, do inglês Copy Number Variations
D0, D1 e D2	Domínio 0, 1 e 2 dos receptores KIR
DNA	Ácido Desoxirribonucleico
EAS	Leste da Ásia, do inglês East Asian
EBV	Vírus Epstein-Barr
EUR	Europeu, do inglês European
GATK	Genome Analysis Toolkit
hg 38	Homo sapiens (human) genome assembly GRCh38
HIV	Vírus da imunodeficiência humana
HLA	Antígeno Leucocitário Humano, do inglês Human Leukocyte Antigen
HWE	Hardy-Weinberg equilibrium
IFN- γ	Interferon Gamma
Ig	Imunoglobulina
IL-10	Interleucina 10
IPD	The Immuno Polymorphism Database
	Motivo de ativação baseado em tirosina, do inglês Immunotyrosine-based activatory
ITAM	motif
	Motivo de inibição baseado em tirosina, do inglês Immunotyrosine-based inhibitory
ITIM	motif
	Receptor de células killer semelhante à imunoglobulina, do inglês Killer-cell
KIR	Imunoglobulin-like Receptor
LD	Desequilíbrio de ligação, do inglês Linkage Disequilibrium
LRC	Complexo de receptores leucocitários, do inglês Leukocyte Receptor Complex
MAF	Frequência alélica mínima, do inglês Minor Allele Frequency
	Complexo Principal de Histocompatibilidade, do inglês Major Histocompatibility
MHC	Complex
miRNA	microRNA

mRNA	RNA mensageiro
NGS	Next Generation Sequencing
NK	Células Natural Killer
PCR	Reação em Cadeia da Polimerase
PCR-SSP	Reação em Cadeia da Polimerase utilizando primer de sequência específica
PF	Pênfigo Foliáceo
ProI	Promotora Intermediária
RBP	ReadBackedPhasing
SNP	Polimorfismo de base única, do inglês Single nucleotide polymorphism
STR	Microssatélite
TCR	Receptor de Célula T
TNF- α	Fator de necrose tumoral- α , do inglês Tumor Necrosis Factor- α
UTR	Região não traduzida, do inglês Untranslated Region
VCF	Variant Call Format

SUMÁRIO

Capítulo 1 – Revisão de Literatura	01
1. As células Natural Killer e os receptores KIR	02
2. Receptores KIR e associação com doenças	08
3. <i>KIR3DL3</i>	09
a. Estrutura e Polimorfismo	09
b. Regiões regulatórias	11
c. Aspectos evolutivos e seleção natural atuando sobre o gene <i>KIR3DL3</i>	12
4. Ancestralidade da população brasileira e Projeto 1000 Genomes	13
Justificativa	15
Objetivos	16
Capítulo 2 – Artigo	25
Abstract	26
Introduction	26
Methods	28
1. Population data and samples	28
2. 1000Genomes data retrieval	29
3. <i>KIR3DL3</i> amplification and sequencing in Brazilian samples.....	29
4. Informative Ancestry Markers (AIMs) for the Brazilian samples.....	30
5. Raw sequencing data processing	30
6. <i>KIR3DL3</i> promoter, coding, and 3'UTR diversity	32
Results	32
1. <i>KIR3DL3</i> genomic diversity, population differentiation, and linkage disequilibrium.....	33
2. <i>KIR3DL3</i> worldwide protein diversity	33
3. <i>KIR3DL3</i> regulatory sequences: promoter and 3'UTR.....	34
4. <i>KIR3DL3</i> extended haplotypes: the association among promoters, coding, and 3'UTR haplotypes	35
5. <i>KIR3DL3</i> evolutionary aspects in worldwide human populations	35
Discussion	36
Tables and figure	51
Supplementary Material	61
Conclusão	91

Capítulo I

REVISÃO DE LITERATURA

1. As células Natural Killer e os receptores KIR

As células *Natural Killer* (NK) atuam principalmente na primeira linha de defesa do sistema imune, a imunidade inata. Elas participam do controle contra infecções por patógenos (vírus, bactérias, parasitas), do reconhecimento de células tumorais e de processos gestacionais. As NK atuam por meio de diversos receptores localizados em sua membrana, os quais são capazes de reconhecer padrões nas células ligantes. Estes receptores não são antígeno-específicos e não sofrem rearranjos genéticos como os receptores de células T (TCR), atuantes na resposta adaptativa (Vivier *et al.*, 2011). Os mecanismos de regulação das células NK envolvem diversas vias celulares e bioquímicas (Jager e Kuchroo, 2010).

Um conjunto de sinais são produzidos por esses receptores, os quais podem levar as células NK a desencadear uma resposta imunológica ou inibi-las, para que células em determinadas condições (como células saudáveis) não sejam alvo de ataque (Hou *et al.*, 2012). As NK possuem atividade citotóxica - eliminam uma célula alvo pela secreção de grânulos líticos contendo perforinas e granzimas que induzem a apoptose (Michelet *et al.*, 2018). Além disso, elas também estão envolvidas com a resposta adaptativa, produzindo quimiocinas e citocinas que ativam outras vias responsáveis por esta resposta. Por exemplo, a produção de interferon- γ (IFN- γ) em condições fisiológicas e patológicas, e de TNF- α e IL-10 com funções pró inflamatórias e imunossupressoras (Vivier *et al.*, 2011). Já as quimiocinas estão envolvidas na colocação com outras células hematopoiéticas, como as células dendríticas em áreas de inflamação (Moretta *et al.*, 2005). Estas diferenças em sua função podem ser observadas ao comparar células CD56^{bright}, as quais possuem atividade citotóxica limitada e estão mais relacionadas com a produção de citocinas e imunomodulação, enquanto CD56^{dim} são altamente citotóxicas, contendo uma maior quantidade de perforinas, granzimas e grânulos citolíticos (Pesce *et al.*, 2018).

Os receptores ativadores em contato com os ligantes podem detectar moléculas do próprio organismo que estão sob condição de estresse. Por outro lado, os receptores inibidores quando interagem, principalmente, com moléculas HLA de classe I em condições normais inibem a ação da NK (Bottino *et al.*, 2005; Vivier *et al.*, 2011). Este mecanismo das células NK funciona como uma balança entre os receptores ativadores e inibidores. Um dos mecanismos de funcionamento das NK é denominado *missing self* (perda do próprio) (Figura 1) (Karre *et al.*, 1986). Neste mecanismo, a perda ou diminuição da expressão de moléculas HLA de classe I e outros ligantes específicos na superfície celular da célula alvo será interpretada como uma

situação anormal, pois não haverá a interação com os receptores inibidores. Com isso, haverá mais interações com receptores ativadores do que inibidores, o que direcionará a célula NK a uma resposta mediada por estes sinais de ativação (Orr e Lanier, 2010; Vivier *et al.*, 2011). Essa situação manifesta-se quando a célula está infectada ou sofre processo neoplásico. Caso a ligação entre o ligante e os receptores inibitórios ocorra normalmente e haja um predomínio desses sinais, não há a ativação das NK. Desta forma, a ligação do receptor com os ligantes, principalmente com as moléculas HLA de classe I, é vital para o reconhecimento de antígenos não próprios e tolerância imunológica (Purdy e Campbell, 2009).

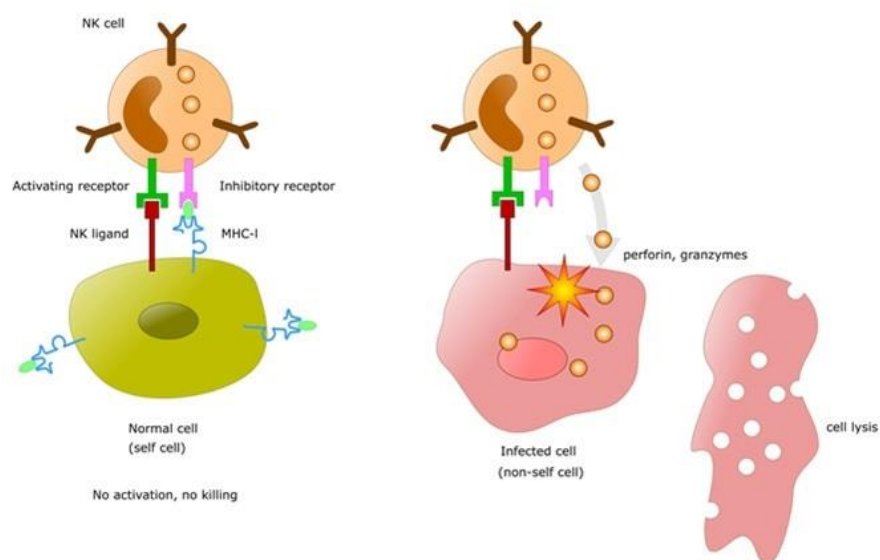


Figura 1. Mecanismo de ação das células *natural killer* (NK) denominado *missing-self*, no qual a interação dos receptores das células NK com ligantes específicos vão induzir a uma resposta. Caso haja um predomínio de sinais inibidores (principalmente com a ligação a moléculas do complexo MHC), não há ativação das células NK. Se há um predomínio de sinais ativadores e inibição dos inibidores, a célula inicia um processo de defesa. Fonte: Hinton-Sheley, Phoebe (2019).

Dentre os receptores presentes nas células NK, há os receptores KIR (receptores *killer* semelhantes à imunoglobulina, do inglês *Killer cell Immunoglobulin-like Receptor*, KIR), que também são expressos em células T citotóxicas (Dupont *et al.*, 1997; Single *et al.*, 2008). Os genes *KIR* constituem uma família gênica composta por 16 genes, sendo 2 pseudogenes (Figura 2), localizados no cromossomo 19 na região chamada Complexo Receptor Leucocitário (do inglês *Leukocyte Receptor Complex*, LRC). Dentre os genes *KIR*, quatro estão presentes em praticamente todos os indivíduos e são considerados genes *framework* – *KIR3DL3*, *KIR3DL2*, *KIR2DL4* e *KIR3DP1*.

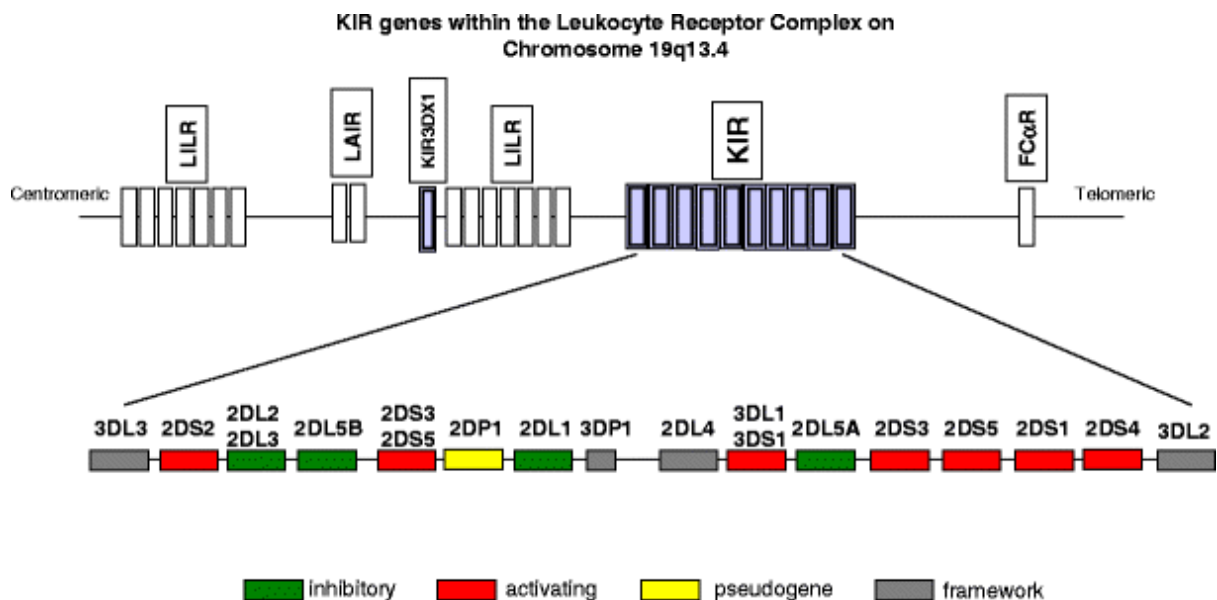


Figura 2. Composição do Complexo Receptor Leucocitário (do inglês, Leukocyte Receptor Complex, LRC), presente na região 19q13.4. Dentre eles, em destaque, os genes KIR. Em cinza os genes *framework*, vermelho os ativadores, verde inibidores e amarelo o pseudogene não *framework*. Fonte: Single et al., (2008).

Os principais ligantes dos receptores KIR são as moléculas do antígeno leucocitário humano de classe I, ou HLA de classe I, presentes na superfície da maioria das células somáticas (Campbell e Purdy, 2011). Os genes HLA localizam-se no Complexo Principal de Histocompatibilidade (MHC, do inglês *Major Histocompatibility Complex*), no braço curto do cromossomo seis. Os genes de classe I codificam proteínas que atuam na apresentação de antígenos intracelulares aos linfócitos T citotóxicos ou CD8⁺ na resposta imune adaptativa (Norman *et al.*, 2016). Além disso, as proteínas produzidas pelos genes não clássicos (*HLA-G*, *-E*, *-F*) estão envolvidos com a modulação do sistema imune (Geraghty *et al.*, 1992).

Os receptores KIR podem ter função ativadora ou inibidora, o que vai depender da formação da sua cauda citoplasmática. Os receptores de inibição possuem cauda citoplasmática longa com motivo de inibição baseado em tirosina (ITIM) e são muito variáveis em termos de polimorfismos (Boudreau e Hsu, 2018a). Os receptores de ativação possuem cauda curta com motivo de ativação baseado em tirosina (ITAM) e tendem a ser menos variáveis (Humphrey *et al.*, 2005). Até o momento, são conhecidos 1110 alelos (ou sequências distintas) de genes KIR (*The Immuno Polymorphism Database*, v. 2.9.0), porém boa parte dessas sequências estão incompletas, principalmente em relação a íntrons ou regiões regulatórias.

A classificação dos genes *KIR* é feita por meio dos seguintes critérios: número de domínios Ig extracelulares que o gene codifica, podendo ser dois (KIR2D) ou três (KIR3D); tamanho da cauda citoplasmática, sendo a inibidora longa (*L* - *long*) e a ativadora curta (*S* -

short); e sua similaridade com outras sequências e ordem de descrição, o que é indicado pelo número final da nomenclatura, após S ou L, como indicado na figura 3 (Bashirova *et al.*, 2006).

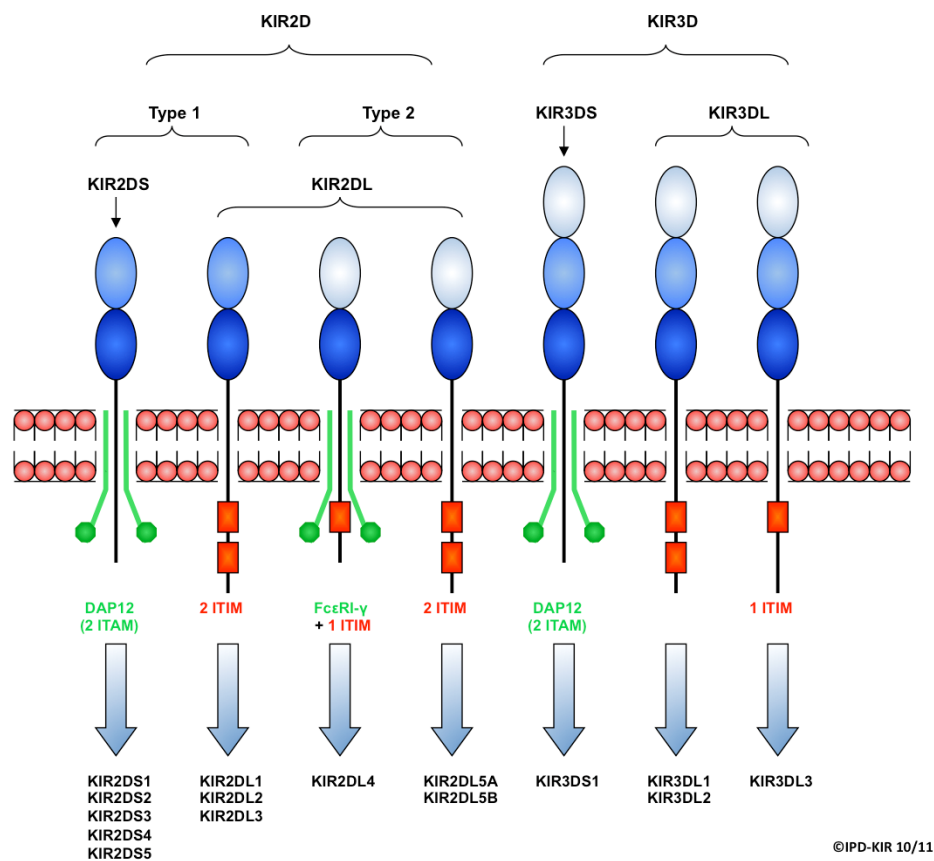


Figura 3. Representação das moléculas KIR. Em azul os domínios semelhantes a imunoglobulina (dois ou três), em vermelho os motivos de inibição baseados em tirosina (ITIM) nas caudas citoplasmáticas longas e em verde os motivos de ativação baseados em tirosina (ITAM) nas caudas curtas. Imagem disponível em <http://www.ebi.ac.uk/ipd/kir/images>.

A maioria dos genes *KIR* apresenta em sua estrutura nove éxons e oito íntrons, com tamanho variando entre 4 e 16 mil pares de base (Figura 4). Os éxons 1 e 2 codificam o peptídeo líder; éxons 3, 4 e 5 os domínios semelhantes à imunoglobulina, chamados de D0, D1 e D2, respectivamente; o éxon 6 codifica uma haste que liga o que está no meio extracelular ao intracelular; éxon 7 codifica o domínio transmembrana e os éxons 8 e 9 a cauda citoplasmática. Alguns genes *KIR* possuem deleção de éxons inteiros quando comparado à estrutura comum dos genes *KIR*, como o gene *KIR3DL3* que não possui o éxon 6 e o *KIR2DL4* que não possui o éxon 4 (Torkar *et al.*, 1998).

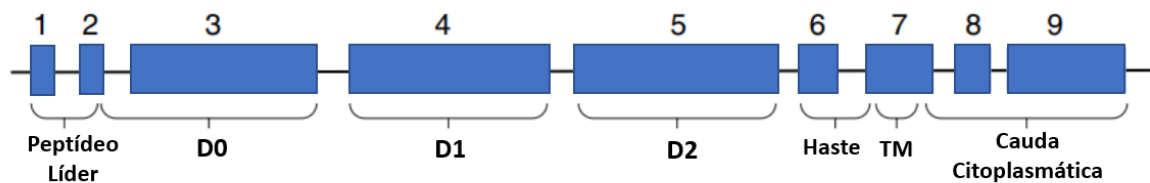


Figura 4. Estrutura geral dos genes *KIR*. Os éxons 1 e 2 codificam o peptídeo líder, éxons 3, 4 e 5 os domínios semelhantes a Ig, éxon 6 a haste que liga o D2 ao domínio transmembranar (TM) que é codificado pelo éxon 7 e a cauda citoplasmática é codificada pelos éxons 8 e 9. Lembrando que há alguns *KIR* que possuem deleção em éxon.

A nomenclatura dos genes *KIR* se dá de forma semelhante à dos HLA (Figura 5). O nome do gene é separado por um asterisco (*), os três primeiros dígitos indicam variações não sinônimas em éxons, os dois seguintes variações sinônimas em éxons e os dois últimos variações na região intrônica, promotora ou não codificante (Marsh *et al.*, 2003).

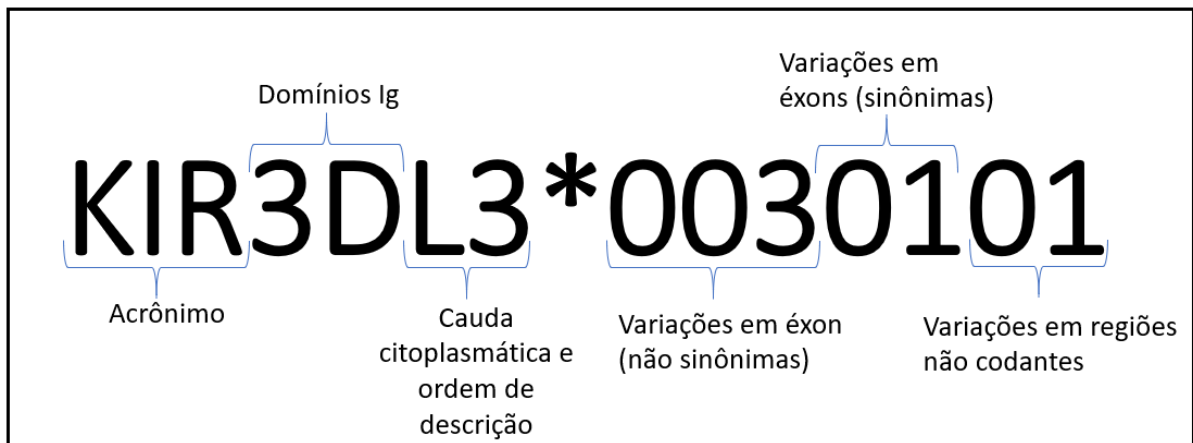


Figura 5. Nomenclatura *KIR*. Baseada na figura disponível em <https://www.ebi.ac.uk/ipd/kir/alleles.html>.

Assim como seus principais ligantes, os genes HLA, a variabilidade dos *KIR* varia conforme o gene estudado. Há genes com um menor número de polimorfismos e outros com muitas variantes (Guethlein *et al.*, 2015). Embora haja diferenças no número de haplótipos e de variantes polimórficas, além de diferenças na estrutura do gene, como já discutido (Suto *et al.*, 1996; Bashirova *et al.*, 2006), os genes *KIR* compartilham uma alta similaridade de sequência (85-99%), fazendo com que recombinações possam ocorrer com frequência, resultando em expansão e contração de haplótipos e rearranjos na ordem dos genes ao longo do haplótipo (Single *et al.*, 2008). Esta variabilidade (que inclui SNPs e variação no número de cópias) influencia na resposta e no funcionamento das células NK, pois a diversidade alélica está relacionada com os níveis de expressão dos receptores e podem influenciar a interação com os

ligantes (Parham, 2005). Além disso, a formação do haplótipo de um indivíduo está associada a especificidade da modulação da resposta imune, se há ou não a presença de determinado gene KIR o perfil de expressão será diferente (Vilches e Parham, 2002).

Há dois principais haplogrupos descritos para os genes *KIR* – A e B, como ilustrado na figura 6 (Single *et al.*, 2008; Hou *et al.*, 2012). O haplótipo A possui apenas um gene de ativação (*KIR2DS4*), cinco de inibição (*KIR2DL1*, *KIR2DL3*, *KIR3DL1*, *KIR3DL2*, *KIR3DL3*), um com dupla função (*KIR2DL4*) e os dois pseudogenes. O haplótipo B possui diferentes combinações, já foram descritos mais de 20 haplótipos diferentes, e há vários genes de ativação. Os *framework*, representados em cinza (*KIR3DL3*, *KIR3DL2*, *KIR2DL4* e *KIR3DP1*) estão presentes em todos os haplótipos, apesar de haver estudos em que há ausência de algum gene *framework*, mas em frequência muito baixa (Hsu, Liu, *et al.*, 2002). No geral, a frequência de genes KIR inibidores é maior e mais homogênea do que os ativadores (Augusto, Zehnder-Alves, *et al.*, 2012). A distribuição dos haplótipos de *KIR* varia entre populações de localizações geográficas distintas (Nemat-Gorgani *et al.*, 2019). Por exemplo, entre indivíduos caucasianos a frequência dos haplótipos A e B é balanceada (Hsu, Chida, *et al.*, 2002). No entanto, o haplótipo A é menos frequente em aborígenes australianos (13%) (Toneva *et al.*, 2001). Os genes KIR possuem potencial para servirem como marcadores de ancestralidade, mas, para isso, é preciso uma melhor descrição da sua diversidade alélica (Augusto, Zehnder-Alves, *et al.*, 2012).

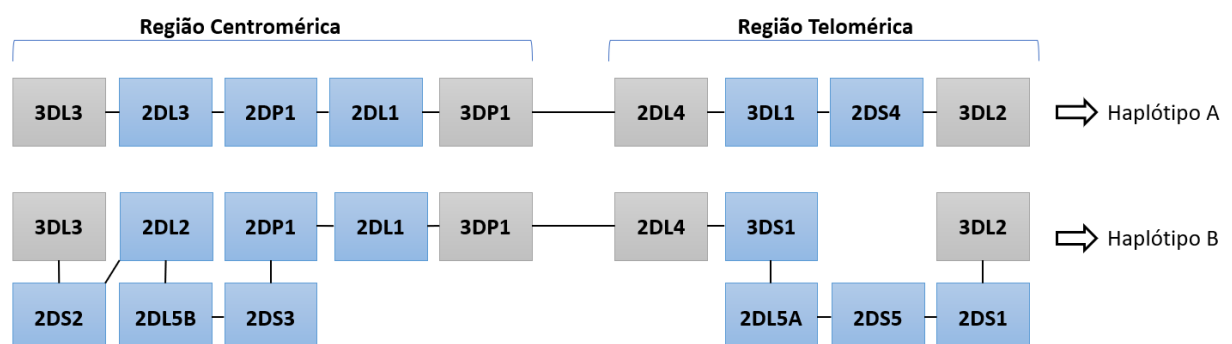


Figura 6. Haplótipos A e B. Em cinza pode-se observar os genes *framework*, e em azul os demais genes *KIR* de cada haplótipo, os quais podem ou não estar presentes e em diferentes quantidades de número de cópia. Baseado na figura de Parham *et al.*, 2005.

2. Receptores KIR e associação com doenças

Há diversos estudos relacionando a variabilidade de alelos, haplótipos e a interação HLA/KIR com diferentes doenças e a resistência a infecções virais (Augusto, Zehnder-Alves, *et al.*, 2012). A grande maioria dos estudos de associação utiliza metodologias de PCR-SSP (Ordonez *et al.*, 2012; Bossi *et al.*, 2015; Wang *et al.*, 2018), que permitem uma análise da composição haplotípica dos genes *KIR* dos indivíduos - quais KIR estão presentes ou ausentes e sua frequência em um grupo caso e um controle. Em geral, haplótipos com maior presença de genes *KIR* ativadores são importantes para combater infecções virais, como HIV e citomegalovírus, pois com o reconhecimento das células infectadas, os genótipos ativadores codificam moléculas que iniciam uma resposta nas células levando à morte celular (Qi *et al.*, 2006).

Em relação à predisposição a respostas autoimunes, os receptores ativadores vêm sendo considerados como um fator de susceptibilidade para algumas doenças, como diabetes tipo I e esclerose múltipla (Kulkarni *et al.*, 2008). Outro exemplo é a artrite reativa, uma doença autoimune com mecanismos pouco conhecidos. Foi observado que altos níveis de expressão de receptores ativadores e baixos níveis de inibidores podem afetar o funcionamento das células NK, contribuindo para a patogênese desta doença, devido superprodução de citocinas locais (Sun *et al.*, 2019). Em outros estudos observou-se que a interação KIR ativador/HLA e sua taxa de expressão podem ter papel de proteção em algumas doenças, se assemelhando ao mecanismo observados em doenças infecciosas, como é o caso do Pênfigo Foliáceo (PF) (Kulkarni *et al.*, 2008; Augusto, Lobo-Alves, *et al.*, 2012). Também em relação ao PF, foi encontrada uma relação de predisposição em um haplótipo composto por dois SNPs localizados a jusante do pseudogene *LILRP2* e um no éxon 3 do gene *KIR3DL3* (Farias *et al.*, 2019).

Em tumores, a baixa ou ausência de expressão de moléculas HLA de classe I leva a eliminação de células transformadas por meio da ação das NK. Foi observado que genótipos compostos por muitos KIR inibidores podem facilitar a evasão das células tumorais, mas não é uma regra, já que há grande variação de acordo com a neoplasia estudada (Verheyden *et al.*, 2004; Besson *et al.*, 2007; Kulkarni *et al.*, 2008). Por exemplo, um estudo sobre câncer colorretal mostrou uma maior frequência de *KIR2DS3* no grupo caso (Diaz-Pena *et al.*, 2020), já em um estudo familiar sobre Linfoma de Hodgkin, indivíduos com o vírus Epstein-Barr (EBV) detectável no sangue apresentaram um fator de proteção mais alto com a presença de genes KIR ativadores, comparado aos que não possuem EBV (Besson *et al.*, 2007). Portanto cada tipo de câncer deve ser analisado em relação aos receptores/ligantes e a sua patogênese,

para que seja possível perceber se há ou não influência desses genes no desenvolvimento ou proteção a neoplasias.

Os genes *KIR* também estão relacionados a gestação. HLA de classe I (*HLA-G*, *-C* e *-E*) estão presentes no trofoblasto extraviloso e há uma alta concentração de células NK ($CD56^{\text{bright}}$) na decídua que expressam receptores KIR, fazendo com que eles sejam críticos para uma placentação apropriada (Trowsdale e Moffett, 2008). A pré-eclâmpsia ocorre quando há uma baixa perfuração da placenta na decídua por parte do trofoblasto extraviloso. Foi estudado que uma maior frequência de homozigose para o haplótipo A materno em combinação com HLA-C2 no feto está associada com o aumento de risco desta condição, pois há uma alta inibição de células NK da decídua, por meio da interação *KIR2DL1* e *HLA-C2*, que prejudica a remodelação dos vasos sanguíneos (Kulkarni *et al.*, 2008; Trowsdale e Moffett, 2008; Huhn *et al.*, 2018). Também foi observado em um estudo recente que o gene *KIR3DL2* possui um fator de proteção e os genes ativadores *KIR2DS2* e *KIR2DS3* são fatores de risco para o aborto espontâneo recorrente (Akbari *et al.*, 2020).

Os receptores KIR possuem um importante papel na regulação da resposta imune e podem desempenhar papel de proteção e susceptibilidade em diversas doenças, de acordo com suas patogêneses e especificidades. Ter uma melhor compreensão a respeito da estrutura e diversidade alélica de cada gene poderia contribuir para que estudos de associação sejam mais realizados nesta escala. Além disso, ter estudos em populações com diferentes ancestralidades e background genético, pois pode haver variações de acordo com a população estudada.

3. *KIR3DL3*

❖ Estrutura e Polimorfismo genético

O gene *KIR3DL3* possui apenas oito éxons e 11.561 pares de base. O éxon 6, que codifica a haste que liga os domínios semelhantes à imunoglobulina a região transmembranar, está ausente nesse gene, provavelmente por processos de recombinação no íntron 5 (Torkar *et al.*, 1998; Jones *et al.*, 2006). Assim como nos demais genes da família *KIR*, os éxons 1 e 2 codificam os peptídeos líderes, os éxons 3, 4 e 5 os domínios semelhantes a imunoglobulinas (D0, D1 e D2), éxon 7 o domínio transmembrana e os éxons 8 e 9 a cauda citoplasmática.

A molécula de *KIR3DL3* possui um único domínio de ITIM, devido truncamento que ocorre antes da posição do segundo ITIM (Trundley *et al.*, 2006). Estas características estruturais podem indicar diferenças na ligação e especificidade em relação aos ligantes, sugerindo uma função diferenciada em relação aos outros KIR. Os receptores KIR possuem

expressão variada nas células NK, alguns podem ser expressos em algumas células e outros não e há variação na intensidade de expressão (Kulkarni *et al.*, 2008).

O *KIR3DL3* possui expressão baixa ou indetectável em células NK no sangue periférico e os transcritos foram detectados em maiores níveis em células NK CD56^{bright}, encontradas no sangue e no sítio de implantação do embrião (Trundley *et al.*, 2006). Este gene provavelmente está envolvido com um importante processo de modulação do sistema imune materno na relação mãe-feto, pois há uma maior expressão na gravidez, no sangue periférico e na decídua materna. Acredita-se que ele pode induzir uma sinalização inibitória específica para esta relação em células NK uterinas, para evitar um conflito com células do embrião e permitir uma infiltração efetiva das células do trofoblasto na parede uterina (Wilson *et al.*, 2000; Jones *et al.*, 2006; Leaton *et al.*, 2019). Também foi encontrado RNAm do gene *KIR3DL3*, mas ausência de expressão celular em células NK92 (Nutalai *et al.*, 2019).

Mesmo com esta baixa expressão, o gene é funcional e codifica proteínas expressas na membrana celular na ausência de proteínas acessórias (Trundley *et al.*, 2006). No geral, o receptor *KIR3DL3* está mais presente nas células NK circulantes de mulheres do que de homens (Wilson *et al.*, 2000). Ainda não se sabe qual é o seu ligante principal, mas há evidências que sugerem que seja uma molécula com a estrutura similar a HLA de classe I clássico, mas com polimorfismo limitado (Leaton *et al.*, 2019).

O alto polimorfismo encontrado em outros genes desta família também é visto em *KIR3DL3*, que atualmente conta com 165 alelos e 92 variantes proteicas catalogadas. O *KIR3DL3* é o gene *KIR* com maior número de alelos descritos até o momento (IPD-KIR, acesso em 12/05/2020). Esta alta diversidade alélica pode estar relacionada a eventos de recombinação que ocorrem principalmente no íntron 5, os quais podem influenciar na ligação com o ligante e nos níveis de expressão (Jones *et al.*, 2006). Ele apresenta elevada heterozigose alélica, com exceção de populações da América do Sul e leste da Ásia (Leaton *et al.*, 2019). A região do gene com maior variabilidade é a que codifica os domínios extracelulares (éxons 3 a 5), enquanto que o segmento intracelular (éxon 7 e 8) são mais conservados (Hou *et al.*, 2012).

Apesar de existirem quatro genes considerados *framework*, apenas o *KIR3DL3* é realmente encontrado em todos os indivíduos, taxa de ocorrência dos demais é de cerca de 99% nas populações mundiais. Ele está na posição 5' terminal do complexo *KIR*, provavelmente sendo protegido de remoções em eventos de recombinação desbalanceada. Há um *hotspot* de recombinação localizado no íntron 5, o qual possui mais de cinco mil pares de base (Leaton *et al.*, 2019). Devido à falta de informações sobre o gene em questão, não se sabe ao certo como

é a relação entre o *KIR3DL3* e a molécula ligante, qual é a diversidade real deste gene em diferentes populações humanas e quais os aspectos funcionais dessa molécula.

❖ Regiões regulatórias do gene *KIR3DL3*

As células NK possuem padrões de expressão variado, em número e identidade. Estes receptores expressos geram subconjuntos de células NK que reconhecem expressão alterada das moléculas HLA de classe I (Li *et al.*, 2016b). No caso dos receptores KIR, a regulação gênica ocorre de forma independente, apesar da alta similaridade de sequência, inclusive na região promotora, entre os genes desta família (Chan *et al.*, 2003). A região promotora de *KIR3DL3* e *KIR2DL4* são as que apresentam menor homologia em relação aos demais genes e se diferem em questão funcional (fatores de transcrição, expressão) (Van Bergen *et al.*, 2005). A variação na expressão de genes *KIR* pode ser explicada pela hipótese do estado de metilação de alelos ativos e inativos. KIR não expressos encontram-se metilados e quando tratados com agentes desmetilantes apresentam uma expressão *de novo* (Stulberg *et al.*, 2007).

A baixa expressão de *KIR3DL3* está relacionada a silenciamento epigenético por meio de forte metilação do promotor proximal (Trompeter *et al.*, 2005), de forma que as poucas moléculas de RNAm de *KIR3DL3* encontradas em algumas amostras devem ser resultado da transcrição de um pequeno número de promotores desmetilados, portanto a maioria fica em um estado inativo de transcrição. Apesar da baixa expressão, é o KIR com atividade promotora mais forte, em especial na região -225, diferentes dos demais em que a região -65 que é a que apresenta uma maior atividade (posição de acordo com o IPD-KIR). A presença de um fator de transcrição (AML) na região proximal também influencia na atividade da região promotora, podendo atuar na desmetilação do promotor (Trompeter *et al.*, 2005; Wright *et al.*, 2014). O silenciamento do gene *KIR3DL3* é, então, associado à hipótese de que sua expressão é condicionada a estímulos relacionados ao desenvolvimento do organismo ou a contextos patológicos, por meio de uma maior regulação da transcrição ou por mecanismos pós-transcricionais (Trundley *et al.*, 2006).

A região promotora intermediária (ProI) influencia na expressão de proteínas por células NK maduras. Foi observado em *KIR2DL1* que, uma deleção nesta região diminuiu a atividade da promotora e os níveis de expressão proteica. *KIR3DL3* é o gene menos conservado nesta região dentre os genes KIR (Wright *et al.*, 2014). Estudos sobre a promotora distal e intermediária mostram que a região intergênica do complexo KIR possui um papel importante na regulação dos transcritos dessa família gênica (Li *et al.*, 2016a). Existe a hipótese de que elementos da promotora intermediária controlam a expressão de KIR, enquanto a promotora

proximal leva a expressões variadas dos KIRs por meio do silenciamento epigenético e coopera com o ProI, possibilitando ou não a expressão – se ele estiver metilado, não há produção de transcritos (Li *et al.*, 2016b).

Os miRNAs são microRNA dentro da família de *small non-coding* RNA (pequenos RNA não codificadores de proteínas) que se liga a região 3'UTR de mRNA, eles podem alterar a estabilidade do RNAm e limitar processos da tradução de proteínas, sendo um importante mecanismo de regulação celular pós-transcricional (Bartel, 2009; Pesce *et al.*, 2018). Um estudo recente mostrou evidências que miRNAs - miR-26a-5p, miR-26b-5p e miR-185-5p - podem interagir com a região 3'UTR de *KIR3DL3*. Foi observado que, além da regulação por meio da metilação da região promotora, a expressão de proteínas na superfície celular é induzida quando há o silenciamento dos miRNAs (Nutalai *et al.*, 2019).

Outro ponto atuante na regulação e expressão de *KIR3DL3* é o íntron 1, o qual possui uma região de VNTR (variação no número de repetições em tandem, do inglês “variable number of tandem repeats”) em uma região rica em GC. Esta combinação pode levar a consequências na expressão gênica de *KIR3DL3*, *imprinting*, eventos de recombinação e *splicing* de RNAm (Trompeter *et al.*, 2005; Jones *et al.*, 2006).

❖ Aspectos evolutivos e seleção natural atuando sobre o gene *KIR3DL3*

O gene *KIR3DL3* está presente em todos os haplótipos humanos e é uma característica conservada dos Hominoidea (Khakoo *et al.*, 2000; Pena *et al.*, 2011). Ele é o único gene que prevaleceu em todos os primatas Catarrhini, mesmo após a rápida evolução desses genes. É observado um considerável nível de conservação de sequências entre os grandes primatas e evidências evolutivas sugerem que é um membro antigo desta família gênica (Jones *et al.*, 2006). É possível que ele tenha sido conservado para agir como uma reserva de sequências para recombinação ou como um elemento estrutural para auxiliar o alinhamento do cromossomo durante a meiose (Trundley *et al.*, 2006). Os genes *KIR* ativadores são mais recentes na escala evolutiva e provavelmente surgiram a partir dos genes inibidores homólogos (Abi-Rached e Parham, 2005).

Os genes *KIR* são divididos em cinco linhagens. A linhagem I se liga a moléculas MHC não polimórficas, como a *HLA-G*. As linhagens II e III possuem ligantes bem estabelecidos e se ligam a motivos polimórficos (*HLA-A, B, C*). A linhagem IV é como a linhagem dois, mas é específica de *Macaca*. No entanto, esta designação não é mais utilizada visto que foi observado que eles se enquadram na linhagem II. A linhagem V não possui ligante definido em nenhuma espécie e apresentam domínios de sinalização pertencentes à linhagem III, este é o caso do gene

KIR3DL3. A linhagem V é a mais conservada ao longo do processo evolutivo (Guethlein *et al.*, 2015).

Em estudo realizado por Leaton e colaboradores (2019), foi realizada uma análise filogenética em cada um dos domínios de *KIR3DL3* em primatas Catarrhini (humanos, gorilas, chimpanzés, orangotango, gibão, macaca) e em seis espécies de macacos do velho mundo (OWM). Foi constatado que, para os domínios semelhantes a Ig, há a presença da linhagem V para a maioria das espécies (exceto D2 de orangotango). Nenhum homínídeo (chimpanzé, gorila, humano) apresentou o domínio citoplasmático de linhagem V, apenas na linhagem III.

Em uma análise populacional é possível avaliar se há algum tipo de seleção natural atuando em uma região gênica. Esta seleção pode favorecer o aumento da frequência de alelos que aumentam o fitness (seleção positiva ou balanceadora) ou diminuir a frequência e até mesmo eliminar variantes que reduzam o fitness do indivíduo (seleção negativa ou purificadora) (Fu e Akey, 2013). Regiões sob seleção balanceadora tendem a acumular um maior número de polimorfismos, em frequências intermediárias, aumentando o grau de variabilidade e a heterozigose se comparado aos outros tipos de seleção (Key *et al.*, 2014).

Esta família gênica possui grande importância na resposta imune e na reprodução, por esse motivo a variabilidade desses genes estão sob pressão seletiva (Augusto *et al.*, 2015). Em uma análise dos domínios semelhantes a imunoglobulina de *KIR3DL3*, foi sugerida a presença de seleção purificadora, com valores baixos de dN/dS (mutações não sinônimas/ mutações sinônimas), e seleção positiva atuando no éxon 6 (região transmembrana) (Jones *et al.*, 2006). Em estudo mais recente, foi observado uma alta taxa de mutações não sinônimas no domínio D2 e a indicação de seleção positiva no éxon 6 (Leaton *et al.*, 2019).

4. Ancestralidade da população brasileira e Projeto 1000 genomas

A composição genética da população brasileira é reflexo de mais de cinco séculos de mistura inter-étnica entre europeus, africanos e ameríndios, tornando-a uma das populações mais heterogêneas do mundo. Portanto, a população brasileira é um modelo tri-híbrido, formada pela miscigenação entre três principais grupos étnicos – europeus, africanos e ameríndios. Por conta disso, ela possui elevada taxa de miscigenação. Tal heterogeneidade étnica se reflete na alta diversidade fenotípica observada entre os brasileiros (Pena *et al.*, 2011). Dessa forma, pode-se considerar os brasileiros como uma população que permite a identificação de novos alelos e haplótipos em qualquer gene, principalmente os considerados muito variáveis.

O estado de São Paulo, alvo do presente estudo, é o mais populoso do Brasil e contém uma considerável miscigenação entre os grupos citados acima. Para essa região do país a maior

contribuição é Europeia (79%), seguida da Africana (14%) e Ameríndia (7%) (Ferreira et al., 2006). Embora dezenas de alelos diferentes de KIR3DL3 tenham sido descritos, não se sabe a respeito da variabilidade inerente desse gene em populações humanas miscigenadas como a brasileira.

Marcadores informativos de ancestralidade (AIMs, do inglês Ancestry Informative Markers) são muito utilizados para análises de ancestralidade em indivíduos. Estes marcadores utilizam tipos de variabilidade genética (SNPs, STRs, inserções/ deleções, entre outros) que apresentam alelos com frequências significativamente diferente entre populações provenientes de localidades geográficas distintas (Martinez-Fierro et al., 2009; Pereira et al., 2012). Um painel com 34 AIMs não relacionados foi desenvolvido com SNPs que diferenciam perfis europeus, africanos e do leste asiático (Phillips et al., 2007; Fondevila et al., 2013). A busca por marcadores indicativos de ancestralidade ameríndia é desafiadora, devido ao pouco espaço de tempo da separação deste grupo com o pertencente ao leste asiático, além da mistura com europeus (Phillips et al., 2007). Por esse motivo utilizamos esses marcadores neste grupo amostral.

O projeto 1000 genomas utilizou diversas metodologias de sequenciamento para avaliar o genoma de 2.504 indivíduos de 26 populações ao redor do mundo. As 26 populações estão divididas cinco grupos: leste asiático, europeu, africano, ameríndio e sul asiático. O objetivo principal do projeto foi genotipar esses indivíduos e fornecer informações a respeito da formação haplotípica e dos polimorfismos presentes nas diferentes populações. Foram classificadas variantes que possuem uma frequência de pelo menos 1% em diferentes populações, variantes raras (0,1-0,5%) no genoma funcional e variantes estruturais, como variações no número de cópias, inversões, inserções e deleções (NCBI). Todos os indivíduos possuem sequenciamento de exoma e de baixa cobertura, além disso 24 indivíduos foram sequenciados com alta cobertura para validação. O mapeamento foi realizado por meio da ferramenta BWA. Não há dados sobre a população brasileira e de nenhuma população miscigenada como a nossa neste consórcio.

JUSTIFICATIVA

Com esse projeto, pretendemos identificar a variabilidade genética, em nível alélico e haplotípico, de um importante gene imunomodulador, em amostras de diferentes populações mundiais, incluindo a brasileira. Os receptores KIR são importantes na modulação de respostas imunes, principalmente nas que há a participação das células NK, fundamentais para a manutenção da gestação, combate a patógenos e neoplasias. Assim, variantes alélicas dos genes *KIR* e alterações nos seus níveis de expressão vem sendo constantemente associadas a susceptibilidade ou proteção a doenças (Parham, 2005). Neste trabalho, exploramos o receptor inibidor KIR3DL3. Além disso, os dados gerados constituirão um banco de dados bastante completo da variabilidade inerente desse gene, considerando que a maioria das sequências conhecidas de genes *KIR*, disponíveis no IPD-KIR (banco de dados específicos para sequências KIR e HLA), não apresentam regiões regulatórias, íntrons e alguns éxons caracterizados.

Não há uma metodologia estabelecida para a análise de genes *KIR* em estudos de sequenciamento de genoma completo. Ainda, devido ao tamanho desses genes, e as dificuldades acarretadas pela similaridade de sequência entre os genes da família *KIR*, sua identificação até o momento vinha sendo feita principalmente com técnicas de baixa resolução, que avaliavam fragmentos pequenos e isolados, ou apenas a presença/ausência de determinado gene. Desta forma, a constituição de uma metodologia adequada e posteriormente de um banco de dados de variantes encontradas no Brasil contribuirá para o entendimento das interações entre esses receptores e HLA, sobre as forças evolutivas que atuam nesses genes, e como esses genes podem influenciar a resposta imune.

OBJETIVOS

O objetivo deste trabalho é analisar a diversidade genética do gene KIR3DL3 em amostras de populações mundiais, incluindo populações europeias, africanas e asiáticas, além de populações miscigenadas como a brasileira. Avaliaremos todos os segmentos do gene, incluindo as regiões regulatórias (promotora e 3' não traduzida), por meio de sequenciamento de segunda geração. Os objetivos específicos são:

1. Elaborar uma metodologia molecular e computacional para analisar a variabilidade do gene KIR3DL3 utilizando dados gerados por sequenciamento de segunda geração;
2. Determinar as frequências alélicas e genotípicas das variantes encontradas nesse gene;
3. Definir os haplótipos encontrados em cada indivíduo;
4. Caracterizar as regiões regulatórias, exônicas e intrônicas deste gene;
5. Classificar funcionalmente cada variante encontrada em éxons;
6. Avaliar a intensidade do desequilíbrio de ligação entre as variantes encontradas;
7. Correlacionar a variabilidade encontrada com o perfil de ancestralidade genômica;
8. Avaliar o perfil de seleção atuando ao longo de todos os segmentos do gene KIR3DL3;
9. Alimentar um banco de dados de variabilidade desse gene para uso em estudos futuros de associação com doenças e estudos evolutivos;

REFERÊNCIAS BIBLIOGRÁFICAS

- ABI-RACHED, L.; PARHAM, P. Natural selection drives recurrent formation of activating killer cell immunoglobulin-like receptor and Ly49 from inhibitory homologues. **J Exp Med**, v. 201, n. 8, p. 1319-32, Apr 18 2005. ISSN 0022-1007.
- AKBARI, S. et al. Recurrent Spontaneous Abortion (RSA) and Maternal KIR Genes: A Comprehensive Meta-Analysis. **JBRA Assist Reprod**, Feb 12 2020. ISSN 1517-5693.
- AUGUSTO, D. G. The Impact of KIR Polymorphism on the Risk of Developing Cancer: Not as Strong as Imagined? **Front Genet**, v. 7, p. 121, 2016. ISSN 1664-8021.
- AUGUSTO, D. G.; HOLLENBACH, J. A.; PETZL-ERLER, M. L. A deep look at KIR-HLA in Amerindians: comprehensive meta-analysis reveals limited diversity of KIR haplotypes. **Hum Immunol**, v. 76, n. 4, p. 272-80, Apr 2015. ISSN 0198-8859.
- AUGUSTO, D. G. et al. Activating KIR and HLA Bw4 ligands are associated to decreased susceptibility to pemphigus foliaceus, an autoimmune blistering skin disease. **PLoS One**, v. 7, n. 7, p. e39991, 2012. ISSN 1932-6203.
- AUGUSTO, D. G.; PETZL-ERLER, M. L. KIR and HLA under pressure: evidences of coevolution across worldwide populations. **Human Genetics**, v. 134, n. 9, p. 929-940, Sep 2015. ISSN 0340-6717.
- AUGUSTO, D. G. et al. Diversity of the KIR gene cluster in an urban Brazilian population. **Immunogenetics**, v. 64, n. 2, p. 143-52, Feb 2012. ISSN 0093-7711.
- BARRETT, J. C. et al. Haploview: analysis and visualization of LD and haplotype maps. **Bioinformatics**, v. 21, n. 2, p. 263-265, Jan 2005. ISSN 1367-4803.
- BARTEL, D. P. MicroRNAs: target recognition and regulatory functions. **Cell**, v. 136, n. 2, p. 215-33, Jan 23 2009. ISSN 0092-8674.
- BASHIROVA, A. A. et al. The killer immunoglobulin-like receptor gene cluster: tuning the genome for defense. **Annu Rev Genomics Hum Genet**, v. 7, p. 277-300, 2006. ISSN 1527-8204.
- BESSON, C. et al. Association of killer cell immunoglobulin-like receptor genes with Hodgkin's lymphoma in a familial study. **PLoS One**, v. 2, n. 5, p. e406, May 2 2007. ISSN 1932-6203.
- BEZMAN, N. A. et al. Distinct requirements of microRNAs in NK cell activation, survival, and function. **J Immunol**, v. 185, n. 7, p. 3835-46, Oct 1 2010. ISSN 0022-1767.
- BOSSI, G. et al. Genetic epistasis between killer immunoglobulin-like receptors and human leukocyte antigens in Kawasaki disease susceptibility. **Genes Immun**, v. 16, n. 7, p. 481-7, Oct 2015. ISSN 1466-4879.
- BOTTINO, C. et al. Cellular ligands of activating NK receptors. **Trends Immunol**, v. 26, n. 4, p. 221-6, Apr 2005. ISSN 1471-4906

BOUDREAU, J. E.; HSU, K. C. Natural Killer Cell Education and the Response to Infection and Cancer Therapy: Stay Tuned. **Trends in Immunology**, v. 39, n. 3, p. 222-239, Mar 2018a. ISSN 1471-4906.

_____. Natural killer cell education in human health and disease. **Current Opinion in Immunology**, v. 50, p. 102-111, Feb 2018b. ISSN 0952-7915.

BROWNING, S. R.; BROWNING, B. L. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. **Am J Hum Genet**, v. 81, n. 5, p. 1084-97, Nov 2007. ISSN 0002-9297

BUTTURA, R. V. et al. HLA-F displays highly divergent and frequent haplotype lineages associated with different mRNA expression levels. **Hum Immunol**, v. 80, n. 2, p. 112-119, Feb 2019. ISSN 0198-8859.

CAMPBELL, K. S.; PURDY, A. K. Structure/function of human killer cell immunoglobulin-like receptors: lessons from polymorphisms, evolution, crystal structures and mutations. **Immunology**, v. 132, n. 3, p. 315-325, Mar 2011. ISSN 0019-2805.

CASTELLI, E. C. et al. A Comprehensive Study of Polymorphic Sites along the HLA-G Gene: Implication for Gene Regulation and Evolution. **Molecular Biology and Evolution**, v. 28, n. 11, p. 3069-3086, Nov 2011. ISSN 0737-4038.

CASTELLI, E. C. et al. Hla-mapper: An application to optimize the mapping of HLA sequences produced by massively parallel sequencing procedures. **Hum Immunol**, v. 79, n. 9, p. 678-684, Sep 2018. ISSN 0198-8859.

CHAN, H. W. et al. DNA methylation maintains allele-specific KIR gene expression in human natural killer cells. **J Exp Med**, v. 197, n. 2, p. 245-55, Jan 20 2003. ISSN 0022-1007

DELANEAU, O.; MARCHINI, J.; ZAGURY, J. F. A linear complexity phasing method for thousands of genomes. **Nat Methods**, v. 9, n. 2, p. 179-81, Dec 4 2011. ISSN 1548-7091.

DIAZ-PENA, R. et al. Analysis of Killer Immunoglobulin-Like Receptor Genes in Colorectal Cancer. **Cells**, v. 9, n. 2, Feb 24 2020. ISSN 2073-4409.

EXCOFFIER, L.; LISCHER, H. E. L. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. **Molecular Ecology Resources**, v. 10, n. 3, p. 564-567, May 2010. ISSN 1755-098X.

FARIAS, T. D. J. et al. Screening the full leucocyte receptor complex genomic region revealed associations with pemphigus that might be explained by gene regulation. **Immunology**, v. 156, n. 1, p. 86-93, Jan 2019. ISSN 0019-2805.

FERREIRA, L. B. et al. Genomic ancestry of a sample population from the state of Sao Paulo, Brazil. **American Journal of Human Biology**, v. 18, n. 5, p. 702-705, Sep-Oct 2006. ISSN 1042-0533.

FONDEVILA, M. et al. Revision of the SNPforID 34-plex forensic ancestry test: Assay enhancements, standard reference sample genotypes and extended population studies. **Forensic Science International-Genetics**, v. 7, n. 1, p. 63-74, Jan 2013. ISSN 1872-4973.

FU, W.; AKEY, J. M. Selection and adaptation in the human genome. **Annu Rev Genomics Hum Genet**, v. 14, p. 467-89, 2013. ISSN 1527-8204.

GERAGHTY, D. E. et al. The HLA class I gene family includes at least six genes and twelve pseudogenes and gene fragments. **J Immunol**, v. 149, n. 6, p. 1934-46, Sep 15 1992. ISSN 0022-1767

GONZALEZ-GALARZA, F. F. et al. Allele frequency net 2015 update: new features for HLA epitopes, KIR and disease and HLA adverse drug reaction associations. **Nucleic Acids Res**, v. 43, n. Database issue, p. D784-8, Jan 2015. ISSN 0305-1048.

GUETHLEIN, L. A. et al. Co-evolution of MHC class I and variable NK cell receptors in placental mammals. **Immunol Rev**, v. 267, n. 1, p. 259-82, Sep 2015. ISSN 0105-2896.

HINTON-SHELEY, P. Missing Self Hypothesis for Natural Killer Cells. News-Medical. (2019, February 26). Retrieved on June 16, 2020 from <https://www.news-medical.net/life-sciences/Missing-Self-Hypothesis-for-Natural-Killer-Cells.aspx>.

HOU, L. et al. Conserved KIR allele-level haplotypes are altered by microvariation in individuals with European ancestry. **Genes and Immunity**, v. 13, n. 1, p. 47-58, Jan 2012. ISSN 1466-4879.

HOU, L. et al. Seventeen novel alleles add to the already extensive KIR3DL3 diversity. **Tissue Antigens**, v. 70, n. 6, p. 449-454, Dec 2007. ISSN 0001-2815.

HSU, K. C. et al. The killer cell immunoglobulin-like receptor (KIR) genomic region: gene-order, haplotypes and allelic polymorphism. **Immunological Reviews**, v. 190, n. 1, p. 40-52, Dec 2002. ISSN 0105-2896.

HSU, K. C. et al. 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. **Journal of Immunology**, v. 169, n. 9, p. 5118-5129, Nov 2002. ISSN 0022-1767.

HUHN, O. et al. High-Resolution Genetic and Phenotypic Analysis of KIR2DL1 Alleles and Their Association with Pre-Eclampsia. **J Immunol**, v. 201, n. 9, p. 2593-2601, Nov 1 2018. ISSN 0022-1767.

HUMPHREY, M. B.; LANIER, L. L.; NAKAMURA, M. C. Role of ITAM-containing adapter proteins and their receptors in the immune system and bone. **Immunological Reviews**, v. 208, p. 50-65, Dec 2005. ISSN 0105-2896.

JAGER, A.; KUCHROO, V. K. Effector and Regulatory T-cell Subsets in Autoimmunity and Tissue Inflammation. **Scandinavian Journal of Immunology**, v. 72, n. 3, p. 173-184, Sep 2010. ISSN 0300-9475.

JONES, D. C. et al. Nature of allelic sequence polymorphism at the KIR3DL3 locus. **Immunogenetics**, v. 58, n. 8, p. 614-627, Aug 2006. ISSN 0093-7711.

KARRE, K. et al. Selective rejection of H-2-deficient lymphoma variants suggests alternative immune defence strategy. **Nature**, v. 319, n. 6055, p. 675-8, Feb 20-26 1986. ISSN 0028-0836

KEY, F. M. et al. Advantageous diversity maintained by balancing selection in humans. **Curr Opin Genet Dev**, v. 29, p. 45-51, Dec 2014. ISSN 0959-437x.

KHAKOO, S. I. et al. Rapid evolution of NK cell receptor systems demonstrated by comparison of chimpanzees and humans. **Immunity**, v. 12, n. 6, p. 687-98, Jun 2000. ISSN 1074-7613.

KULKARNI, S.; MARTIN, M. P.; CARRINGTON, M. The Yin and Yang of HLA and KIR in human disease. **Seminars in Immunology**, v. 20, n. 6, p. 343-352, Dec 2008. ISSN 1044-5323.

KUMAR, S.; STECHER, G.; TAMURA, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. **Mol Biol Evol**, v. 33, n. 7, p. 1870-4, Jul 2016. ISSN 0737-4038.

LEATON, L. A. et al. Conservation, Extensive Heterozygosity, and Convergence of Signaling Potential All Indicate a Critical Role for KIR3DL3 in Higher Primates. **Front Immunol**, v. 10, p. 24, 2019. ISSN 1664-3224.

LI, H. et al. Characterization of KIR intermediate promoters reveals four promoter types associated with distinct expression patterns of KIR subtypes. **Genes and Immunity**, v. 17, n. 1, p. 66-74, Jan-Feb 2016a. ISSN 1466-4879.

_____. Characterization of KIR intermediate promoters reveals four promoter types associated with distinct expression patterns of KIR subtypes. **Genes Immun**, v. 17, n. 1, p. 66-74, Jan-Feb 2016b. ISSN 1466-4879.

LIMA, T. H. A. et al. HLA-A promoter, coding, and 3'UTR sequences in a Brazilian cohort, and their evolutionary aspects. **Hla**, v. 93, n. 2-3, p. 65-79, Feb 2019. ISSN 2059-2302.

MARSH, S. G. E. et al. Killer-cell immunoglobulin-like receptor (KIR) nomenclature report, 2002. **Journal of Immunological Methods**, v. 281, n. 1-2, p. 1-8, Oct 2003. ISSN 0022-1759.

MARTINEZ-FIERRO, M. L. et al. Ancestry informative markers and admixture proportions in northeastern Mexico. **J Hum Genet**, v. 54, n. 9, p. 504-9, Sep 2009. ISSN 1434-5161.

MCKENNA, A. et al. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. **Genome Research**, v. 20, n. 9, p. 1297-1303, Sep 2010. ISSN 1088-9051.

MICHELET, X. et al. Metabolic reprogramming of natural killer cells in obesity limits antitumor responses. **Nat Immunol**, v. 19, n. 12, p. 1330-1340, Dec 2018. ISSN 1529-2908.

MORETTA, A. et al. Early liaisons between cells of the innate immune system in inflamed peripheral tissues. **Trends Immunol**, v. 26, n. 12, p. 668-75, Dec 2005. ISSN 1471-4906

MURRELL, B. et al. FUBAR: a fast, unconstrained bayesian approximation for inferring selection. **Mol Biol Evol**, v. 30, n. 5, p. 1196-205, May 2013. ISSN 0737-4038 (

NEMAT-GORGANI, N. et al. Diversity of KIR, HLA Class I, and Their Interactions in Seven Populations of Sub-Saharan Africans. **J Immunol**, v. 202, n. 9, p. 2636-2647, May 1 2019. ISSN 0022-1767.

NORMAN, P. J. et al. Co-evolution of Human Leukocyte Antigen (HLA) Class I Ligands with Killer-Cell Immunoglobulin-Like Receptors (KIR) in a Genetically Diverse Population of Sub-Saharan Africans. **Plos Genetics**, v. 9, n. 10, Oct 2013. ISSN 1553-7404.

NORMAN, P. J. et al. Defining KIR and HLA Class I Genotypes at Highest Resolution via High-Throughput Sequencing. **American Journal of Human Genetics**, v. 99, n. 2, p. 375-391, Aug 2016. ISSN 0002-9297.

NUTALAI, R. et al. Regulation of KIR3DL3 Expression via Mirna. **Genes (Basel)**, v. 10, n. 8, Aug 9 2019. ISSN 2073-4425.

O'CONNOR, G. M. et al. Mutational and structural analysis of KIR3DL1 reveals a lineage-defining allotypic dimorphism that impacts both HLA and peptide sensitivity. **J Immunol**, v. 192, n. 6, p. 2875-84, Mar 15 2014. ISSN 0022-1767.

ORDONEZ, D. et al. KIR typing by non-sequencing methods: polymerase-chain reaction with sequence-specific primers. **Methods Mol Biol**, v. 882, p. 415-30, 2012. ISSN 1064-3745.

ORR, M. T.; LANIER, L. L. Natural Killer Cell Education and Tolerance. **Cell**, v. 142, n. 6, p. 847-856, Sep 2010. ISSN 0092-8674.

PARHAM, P. MHC class I molecules and KIRs in human history, health and survival. **Nature Reviews Immunology**, v. 5, n. 3, p. 201-214, Mar 2005. ISSN 1474-1733.

PENA, S. D. et al. The genomic ancestry of individuals from different geographical regions of Brazil is more uniform than expected. **PLoS One**, v. 6, n. 2, p. e17063, Feb 16 2011. ISSN 1932-6203.

PEREIRA, R. et al. Straightforward inference of ancestry and admixture proportions through ancestry-informative insertion deletion multiplexing. **PLoS One**, v. 7, n. 1, p. e29684, 2012. ISSN 1932-6203.

PESCE, S. et al. New miRNA Signature Heralds Human NK Cell Subsets at Different Maturation Steps: Involvement of miR-146a-5p in the Regulation of KIR Expression. **Front Immunol**, v. 9, p. 2360, 2018. ISSN 1664-3224.

PHILLIPS, C. Forensic genetic analysis of bio-geographical ancestry. **Forensic Science International-Genetics**, v. 18, p. 49-65, Sep 2015. ISSN 1872-4973.

PHILLIPS, C. et al. Inferring ancestral origin using a single multiplex assay of ancestry-informative marker SNPs. **Forensic Sci Int Genet**, v. 1, n. 3-4, p. 273-80, Dec 2007. ISSN 1872-4973.

POSTH, C. et al. Reconstructing the Deep Population History of Central and South America. **Cell**, v. 175, n. 5, p. 1185-1197.e22, Nov 15 2018. ISSN 0092-8674.

PRITCHARD, J. K.; STEPHENS, M.; DONNELLY, P. Inference of population structure using multilocus genotype data. **Genetics**, v. 155, n. 2, p. 945-959, Jun 2000. ISSN 0016-6731.

PURDY, A. K.; CAMPBELL, K. S. Natural killer cells and cancer Regulation by the killer cell Ig-like receptors (KIR). **Cancer Biology & Therapy**, v. 8, n. 23, p. 2211-2220, Dec 2009. ISSN 1538-4047.

PYO, C. W. et al. Different Patterns of Evolution in the Centromeric and Telomeric Regions of Group A and B Haplotypes of the Human Killer Cell Ig-Like Receptor Locus. **Plos One**, v. 5, n. 12, Dec 2010. ISSN 1932-6203.

PYO, C. W. et al. Recombinant structures expand and contract inter and intragenic diversification at the KIR locus. **BMC Genomics**, v. 14, p. 89, Feb 8 2013. ISSN 1471-2164.

QI, Y. et al. KIR/HLA pleiotropism: protection against both HIV and opportunistic infections. **PLoS Pathog**, v. 2, n. 8, p. e79, Aug 2006. ISSN 1553-7366.

ROBINSON, J. et al. The IPD and IMGT/HLA database: allele variant databases. **Nucleic Acids Res**, v. 43, n. Database issue, p. D423-31, Jan 2015. ISSN 0305-1048.

ROE, D. et al. Revealing complete complex KIR haplotypes phased by long-read sequencing technology. **Genes Immun**, v. 18, n. 3, p. 127-134, Sep 2017. ISSN 1466-4879.

SINGLE, R. M. et al. Methods for assessing gene content diversity of KIR with examples from a global set of populations. **Immunogenetics**, v. 60, n. 12, p. 711-725, Dec 2008. ISSN 0093-7711.

STULBERG, M. J. et al. Identification of distal KIR promoters and transcripts. **Genes Immun**, v. 8, n. 2, p. 124-30, Mar 2007. ISSN 1466-4879.

SUN, H. S. et al. Disease-association of different killer cell immunoglobulin-like receptors (KIR) and HLA-C gene combinations in reactive arthritis. **Mod Rheumatol**, v. 29, n. 3, p. 531-537, May 2019. ISSN 1439-7595.

SUTO, Y. et al. Chromosomal localization of the human natural killer cell class I receptor family genes to 19q13.4 by fluorescence in situ hybridization. **Genomics**, v. 35, n. 1, p. 270-272, Jul 1996. ISSN 0888-7543.

TAKESHITA, L. Y. et al. A database for curating the associations between killer cell immunoglobulin-like receptors and diseases in worldwide populations. **Database (Oxford)**, v. 2013, p. bat021, 2013. ISSN 1758-0463.

TONEVA, M. et al. Genomic diversity of natural killer cell receptor genes in three populations. **Tissue Antigens**, v. 57, n. 4, p. 358-62, Apr 2001. ISSN 0001-2815

TORKAR, M. et al. Isotypic variation of novel immunoglobulin-like transcript/killer cell inhibitory receptor loci in the leukocyte receptor complex. **Eur J Immunol**, v. 28, n. 12, p. 3959-67, Dec 1998. ISSN 0014-2980

TROMPETER, H. I. et al. Three structurally and functionally divergent kinds of promoters regulate expression of clonally distributed killer cell Ig-like receptors (KIR), of KIR2DL4, and of KIR3DL3. **Journal of Immunology**, v. 174, n. 7, p. 4135-4143, Apr 2005. ISSN 0022-1767.

TROWSDALE, J.; MOFFETT, A. NK receptor interactions with MHC class I molecules in pregnancy. **Semin Immunol**, v. 20, n. 6, p. 317-20, Dec 2008. ISSN 1044-5323

TRUNDLEY, A. E. et al. Molecular characterization of KIR3DL3. **Immunogenetics**, v. 57, n. 12, p. 904-916, Jan 2006. ISSN 0093-7711.

UHRBERG, M. et al. Human diversity in killer cell inhibitory receptor genes. **Immunity**, v. 7, n. 6, p. 753-763, Dec 1997. ISSN 1074-7613.

VAN BERGEN, J. et al. Structural and functional differences between the promoters of independently expressed killer cell Ig-like receptors. **Eur J Immunol**, v. 35, n. 7, p. 2191-9, Jul 2005. ISSN 0014-2980

VAN DER AUWERA, G. A. et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. **Curr Protoc Bioinformatics**, v. 43, p. 11.10.1-33, 2013. ISSN 1934-3396.

VANDENBUSSCHE, C. J. et al. Dramatically reduced surface expression of NK cell receptor KIR2DS3 is attributed to multiple residues throughout the molecule. **Genes Immun**, v. 10, n. 2, p. 162-73, Mar 2009. ISSN 1466-4879.

VENDELBOSCH, S. et al. Extensive Variation in Gene Copy Number at the Killer Immunoglobulin-Like Receptor Locus in Humans. **Plos One**, v. 8, n. 6, Jun 2013. ISSN 1932-6203.

VERHEYDEN, S.; BERNIER, M.; DEMANET, C. Identification of natural killer cell receptor phenotypes associated with leukemia. **Leukemia**, v. 18, n. 12, p. 2002-7, Dec 2004. ISSN 0887-6924.

VILCHES, C.; PARHAM, P. KIR: Diverse, rapidly evolving receptors of innate and adaptive immunity. **Annual Review of Immunology**, v. 20, p. 217-251, 2002. ISSN 0732-0582.

VIVIER, E. et al. Innate or Adaptive Immunity? The Example of Natural Killer Cells. **Science**, v. 331, n. 6013, Jan 2011. ISSN 0036-8075.

WANG, L. et al. Killer-cell immunoglobulin-like receptors associate with HIV-1 infection in a narrow-source Han Chinese cohort. **PLoS One**, v. 13, n. 4, p. e0195452, 2018. ISSN 1932-6203.

WILSON, M. J. et al. Plasticity in the organization and sequences of human KIR/ILT gene families. **Proc Natl Acad Sci U S A**, v. 97, n. 9, p. 4778-83, Apr 25 2000. ISSN 0027-8424

WRIGHT, P. W. et al. Characterization of a weakly expressed KIR2DL1 variant reveals a novel upstream promoter that controls KIR expression. **Genes Immun**, v. 15, n. 7, p. 440-8, Oct 2014. ISSN 1466-4879.

Capítulo II

INSIGHTS ON THE *KIR3DL3* GENETIC DIVERSITY AND EVOLUTIONARY HISTORY IN WORLDWIDE POPULATIONS

Heloísa S. Andrade, Erick C. Castelli

Abstract

Killer-cell Immunoglobulin-like Receptors (KIRs) are transmembrane receptors expressed in Natural Killer cells and cytotoxic T cells that recognize molecules from the immune system, mainly molecules from the major class I histocompatibility complex on the surface of somatic cells. *KIR* genes are highly polymorphic considering the number of copies of each locus and their sequence content. *KIR3DL3* is a framework *KIR* gene present in every individual. Here we addressed *KIR3DL3* genetic diversity and evolutionary aspects in worldwide human populations, using a bioinformatics approach that minimizes mapping bias and genotyping errors. *KIR3DL3* genetic variability is far greater than we currently acknowledge, especially when we consider intronic segments and regulatory sequences. *KIR3DL3* protein diversity is strongly influenced by the ancestry background, with different frequencies of protein versions and amino acid residues among Africans, Europeans, and other populations. In this matter, populations from East Asia present a different pattern considering protein frequencies, regulatory sequences diversity, and how natural selection has shaped *KIR3DL3* diversity and haplotype frequencies. Because of that, any association study involving *KIR3DL3* should be performed concerning for population stratification. Some *KIR3DL3* segments are highly conserved, such as the exons encoding the leader peptide and the extracellular domain D2, important for HLA interaction, while other segments, such as the promoter region and exon 6 (the transmembrane domain) are highly variable. We found evidence of balancing selection in the promoter region, probably maintaining sequences with different promotion capabilities. This evidence is stronger among Europeans and East Asians. We also find evidence of balancing selection in some large intronic segments, and they must be considered as possibly carrying regulatory elements.

Keywords: *KIR3DL3*, polymorphism, haplotype, NGS, ancestry, natural selection, evolution

Introduction

Killer-cell Immunoglobulin-like Receptors (KIRs) are transmembrane receptors that recognize molecules from the immune system, mainly molecules from the major class I histocompatibility complex on the surface of somatic cells (Campbell e Purdy, 2011). They are mostly expressed in Natural Killer cells and cytotoxic T cells (Single *et al.*, 2008). The KIR complex is located on chromosome 19 in a region called the Leukocyte Receptor Complex

(LRC), with 16 different genes and some pseudogenes. KIR receptors present either activator or inhibitory function, depending on their cytoplasmic tail structure and the ligand binding to them (Single *et al.*, 2008). Inhibitory receptors have a long cytoplasmic tail with tyrosine-based inhibitory motif (ITIM) and are very polymorphic (Boudreau e Hsu, 2018a; b). Activating receptors have a tyrosine-based activation motif (ITAM), a short cytoplasmic tail and are usually less variable (Humphrey *et al.*, 2005).

The HLA class I molecules are the major ligands for KIR receptors and both gene families (HLA and KIR) are highly polymorphic (Guethlein *et al.*, 2015). KIR genes are polymorphic in their genomic sequence and the number of copies of each locus. According to the IPD-KIR database, version 2.9.0, there are 1110 different KIR sequences considering all KIR genes, but most of these sequences are incomplete with no information regarding introns or regulatory regions. Despite their polymorphisms and some gene structure differences among them (Suto *et al.*, 1996; Bashirova *et al.*, 2006), KIR genes share a high sequence similarity rate (85-99%), leading to frequent recombination and resulting in the expansion/contraction of haplotypes and rearrangements in the order of genes in the KIR complex (Single *et al.*, 2008). This variability (single nucleotide variants and copy number variation) influences which KIR genes will be expressed and their expression rate, and also the binding strength between ligand and receptor, and thus it influences NK cell response and function (Vilches e Parham, 2002). KIR allelic variation is poorly known and they are rarely studied in diseases (Augusto, 2016) or different populations. However, a recent study has evaluated KIR3DL3 in samples from the 1000Genomes and higher primates (Leaton *et al.*, 2019), indicating that *KIR3DL3* is highly conserved throughout primate evolution.

KIR3DL3 presents eight exons and seven introns in a segment of about 12,000 nucleotides (Torkar *et al.*, 1998). Exons 1 and 2 encode the leader peptide, exons 3, 4 and 5 the immunoglobulin-like domains (D0, D1, and D2), exon 7 the transmembrane domain, and exons 8 the cytoplasmic tail. *KIR3DL3* molecules have a single ITIM domain. These structural characteristics may indicate differences in binding and specificity related to its ligands, suggesting a differentiated function compared to other KIRs. *KIR3DL3* has low or undetectable protein expression in peripheral blood but, even though, the gene is functional and encodes proteins expressed in the absence of accessory proteins (Trundley *et al.*, 2006). Its major ligand is not yet known, but evidence suggests the ligand has a similar structure as classical HLA class I, but with limited polymorphism (Leaton *et al.*, 2019).

The low *KIR3DL3* expression is also related to epigenetic silencing by promoter methylation (Trompeter *et al.*, 2005). Few *KIR3DL3* mRNA molecules were found in a study sample and they may be a result of the transcription of a small number of demethylated promoters, as most are in an inactive state of transcription. Such silencing is associated with the hypothesis that its expression is conditioned to stimulation during the development or in pathological contexts (Trundley *et al.*, 2006).

Previous studies have indicated that only three KIR genes are usually present in all individuals in two copies. These framework genes are *KIR3DL2*, *KIR3DL3*, and *KIR2DL4* (Vendelbosch *et al.*, 2013). However, *KIR3DL3* seems to be the only true framework gene since some individuals carrying one copy of the others have been described (Pyo *et al.*, 2010; Pyo *et al.*, 2013; Roe *et al.*, 2017; Leaton *et al.*, 2019). *KIR3DL3* is in the 5' terminal position of the KIR complex, probably being protected from removals in unbalanced recombination events (Leaton *et al.*, 2019). The high polymorphism found in other genes in this family is also seen in *KIR3DL3* since it is the KIR gene with the largest number of alleles described so far.

Here we will address *KIR3DL3* genetic diversity and evolutionary aspects in worldwide populations, considering a segment of about 12,000 nucleotides that includes *KIR3DL3* promoter and the coding sequence. We also introduce a new molecular and computational method to evaluate *KIR3DL3* complete genetic diversity.

Methods

Population data and samples

For this study we retrieved data from the newly released high-coverage sequencing data for all the 2,504 samples of the 1000Genomes dataset, comprehending 26 different population samples across the world. Different from previous studies, here we retrieved all that from approximately 2Kb upstream the first translated ATG to the last exon, including introns.

Additionally, we included 216 unrelated Brazilian samples from the São Paulo state, Brazil. Brazilians are one of the most admixed populations in the world (Ferreira *et al.*, 2006) and it is not included in the 1000Genomes dataset. All participants from Brazil signed informed consent before blood draw and the local Human Research Ethics Committee (Protocol 24157413.7.0000.5411) approved the study. Ancestry Informative Markers (AIMs) were used to infer the ancestry of these samples. In this Brazilian sample ancestry composition was 75.5% European, 16.1% African, and 8.4% Amerindian/Asian, but we also assessed the ancestry composition of each individual. This is the expected proportion for populations from the region

that samples were collected (Ferreira *et al.*, 2006). DNA was obtained using salting out procedure, quantified with Qubit Broad Range Assays (Thermo Fisher Scientific Inc., Waltham, MA) and normalized in a concentration of 50ng/μL.

1000Genomes data retrieval

We downloaded the aligned file using the high-speed protocol ASPERA. After, we retrieved all reads mapped to the *KIR* complex, and all unmapped reads, using *samtools*. Unmapped reads are important because there may be many *KIR*-related reads among them, especially when the individual's variability creates sequences that are too divergent from the reference genome. A local Perl script converted the aligned data into raw fastq data, that were further processed to avoid mapping bias (discussed later). Throughout only the results for *KIR3DL3* are presented here, we retrieved all reads mapped to the *KIR* complex, since there may be cross-mappings among paralogous genes.

***KIR3DL3* amplification and sequencing in Brazilian samples**

KIR3DL3 is a large gene. Because of that, we amplified *KIR3DL3* in two overlapping fragments, encompassing 14.190 bp that includes its promoter region, the coding segment with all exons and introns, and the complete 3'UTR (untranslated) region. Polymerase Chain Reaction (PCR) was carried out with primers: (1) 5'-GGTGCATAAGGTTGGGTGTTG-3' and 5'-GGCTTCCAGTCCTAGATCATTC-3', which amplify a segment of 5.460 bp from 19:54722518 to 19:54727978 (human genome assembly hg38) comprehending the full promoter up to begin of intron 4; and (2) 5'-GATAGACACCATGGAGGGGA-3' e 5'-CAGATGGGGTTATGTGGACG-3', which amplify 9.193 bp, from intron 3 until the end of the 3'UTR (19:54727515 – 54736708, hg38). Both fragments overlap in about 460 nucleotides.

The final PCR volume was 25μL for both reactions. For the first fragment, the protocol was: 1.25U of TaKaRa GXL Taq® DNA polymerase and 1X buffer solution, 0.20mM of dNTP and 0.20μM of each primer. The cycling conditions were: 30 cycles 98°C by 10 seconds, 60°C for 15 seconds, 68°C for 6 minutes. For the second fragment, we used the same protocol, but with 0.25 μM of each primer. The cycling conditions were: 30 cycles of 98°C by 10 seconds and 68°C for 10 minutes. We used 100ng of genomic DNA for each reaction. Amplicons were evaluated on 1% agarose gel stained with GelRed® (Biotium, Inc. Hayward, CA).

We prepared sequencing libraries following the recommendations for Agilent SureSelect[^]QXT Reagent Kits. Although here we present *KIR3DL3* data only, other *KIR* genes were amplified and included in this sequencing procedure. We sequenced 96 samples per run. Libraries quantification and fragmentation pattern were evaluated using Qubit dsDNA High Sensitivity assays (ThermoFisher) and Agilent High Sensitivity DNA Kits (Agilent Technologies, CA, USA), respectively. Libraries were normalized to the recommended concentration for MiSeq Reagent Kit version V2 (500 cycles, 2 x 250-bp).

Informative Ancestry Markers (AIMs) for the Brazilian samples

A panel with 34 AIMs was used to estimate the contribution of European, African, and Amerindian/Asian ancestry for this sample. These markers consist of the SNPforID 34-plex panel (Phillips, 2015). Each variant was amplified in a series of multiplex amplification and pooled together with the *KIR* amplicons before the library preparation step. We used STRUCTURE v2.3.4 (Pritchard *et al.*, 2000) to infer the population structure and individual ancestry rates, with K=3. For parental samples, we used data from the 1000Genomes Project data for European, African and East Asian populations (404 Europeans from populations TSI, FIN, GBR, and IBS, 504 Africans from YRI, LWK, GWD, MSL, and ESN, and 504 East Asians from CHB, JPT, CHS, CDX, and KHV). Native American markers are a challenge to be detected because of the brief time of separation between this group and the East Asian, besides the admixture with the European group (Phillips *et al.*, 2007). For this reason, the East Asian group was used as a representative for Amerindians (Posth *et al.*, 2018).

Raw sequencing data processing

KIR genes are prone to mapping errors and mapping bias because of their complex structure and high similarity within genes of the *KIR* cluster. Some tools have been developed to deal with *KIR* data from massively parallel sequencing, such as PING (Norman *et al.*, 2016), but these tools do not aim to map reads against the reference genome. Moreover, when the original *KIR3DL3* data from the 1000Genomes is retrieved from the phased VCF for Phase 3, most of the known *KIR3DL3* alleles described in the IPD-KIR database are not there.

To circumvent this issue, we used a new version of the hla-mapper software (Castelli *et al.*, 2018) that supports a *KIR* database. This software applies a series of filters to define where each read pair should be mapped [method described in (Castelli *et al.*, 2018)]. The *KIR* database

includes sequences for all KIR genes retrieved from the IPD- KIR database, and also from public databases such as GENBANK. This new version is available at www.castelli-lab.net/apps/hla-mapper.

We obtained genotypes using the Genome Analysis Toolkit HaplotypeCaller version 4.1.2 in the GVCF mode (Mckenna *et al.*, 2010; Van Der Auwera *et al.*, 2013) and GATK GenotypeGVCFs. Each variant was annotated using the dbSNP dataset. For variant refinement and selection, we applied three filters in this order: vcfx checkpl, which introduced missing alleles in genotypes with low likelihoods; vcfx checkad, which introduced missing alleles in unbalanced genotypes and homozygous genotypes inferred on segments with low sequence depth, and vcfx evidence, which annotates each variant for balanced heterozygosis, homozygosis for alternative alleles, the proportion of missing alleles, and other quality parameters. vcfx is available at www.castelli-lab.net/apps/vcfx. Each variant that was not annotated as “PASS” by vcfx evidence was manually inspected observing the alignments (BAM file) and, if necessary, they were removed or included in the final VCF. This step assures that only high-quality genotypes have been passed forward for the imputation and haplotyping procedure discussed later.

Although hla-mapper optimizes mappings along the entire *KIR3DL3* sequence, we detected an issue that prevented the correct genotype of variants in intron 1. *KIR3DL3* intron 1 is a repetitive segment, with the same sequence repeated many times (Jones *et al.*, 2006). The read size of the 1000Genome data (about 150bp or less) led to misalignments in this segment and genotypes with low likelihoods. Because of that, we decided to remove all variants detected in intron 1 from further analysis.

We inferred haplotypes combining two computational strategies. First, GATK ReadBackedPhasing (RBP) inferred the relationship between close heterozygous variants, with the minimal Phase Quality Threshold set to 500 (25x the default value). This procedure resulted in phasing sets, i.e., blocks of known phases, but unphased among them. Multi-allelic variants, indels, and missing alleles are not considered by RBP. Then, the second step consisted of inferring the final haplotypes considering these phase sets. For that, we used an in-house software named *phasesx*. This local software use Shapeit4 (Delaneau *et al.*, 2011) to phase all bi-allelic variants considering the phase sets, in 50 independent runs, comparing the results afterward. We preserve the phase of all samples in which the same pair of haplotypes was observed in at least 48 runs ($P > 0.95$), passing this forward to the next round of runs. This procedure is repeated until the number of samples with $P > 0.95$ no longer increases. Then, the

samples with defined haplotypes are passed forward to the next step, while the others are passed as completely unphased. The next step consisted of using Beagle 4.1 (Browning e Browning, 2007) to infer the final set of haplotypes, now including the multi-allelic variants. Here we also used 50 independent runs, fixing the threshold for detecting a haplotype as $P > 0.95$ as described using Shapeit. The final set of haplotypes are the ones with $P > 0.7$ after the final run. Shapeit4 and Beagle 4.1 also imputed the bi-allelic and multi-allelic missing alleles, respectively. We have removed all singletons before the haplotyping procedure. This step is necessary because singletons are ambiguous by definition. When possible, singletons were manually inserted in the final VCF file comparing the RBP results and the BAM files. Some samples (2.1%) did not achieve the threshold after the final run and they were removed.

***KIR3DL3* promoter, coding, and 3'UTR diversity**

We used vcfx fasta and vcfx transcript to create complete genomic and CDS sequences for each individual (two per sample), but intron 1 was replaced by a series of N nucleotides. We used a local BLAST server containing all known genomic and nucleotide sequences from IPD-IMGT/KIR Database version 2.9.0 to define the closest known allele for each different sequence. Moreover, sequences of the promoter and 3'UTR regions were isolated, counted and named manually according to sequence similarities and frequency of each different haplotype. The CDS sequences were translated into full-length proteins using EMBOSS *transeq* and named according to protein sequences from IPD-IMGT/KIR using the same BLAST server.

Allele, genotype, and haplotype frequencies were calculated by direct counting. The VCF file was converted into an ARP file for ARLEQUIN 3.5 (Excoffier e Lischer, 2010), to evaluate F_{ST} . We used variscan to evaluate *Tajima's D* and nucleotide diversity, for each segment and in 250bp windows with a step size of 1. Linkage disequilibrium (LD) was assessed using Haploview 4.2 for variants with minor allele frequency (MAF) higher than 5% (Barrett *et al.*, 2005). The dN/dS ratio was calculated using MegaX and megacc and one copy of each sequence, for each segment (Kumar *et al.*, 2016). We used SPCC to perform a multidimensional scaling ($K=2$) of the F_{ST} data, plotted using ggplot in R. We used plink (Purcell *et al.*, 2007) to perform a Principal Component analysis (PCA) of the phased data using R (library tidyverse).

Results

Here we have addressed *KIR3DL3* diversity and evolutionary aspects in worldwide human populations, combining data from the new release of the 1000Genomes high-coverage

sequencing data, and our own high-coverage data from Brazil, considering about 13,702 nucleotides, from 19:54722834 to 19:54736536 (hg 38), encompassing the full promoter, all exons, and all introns except intron 1. The full list of variants detected in this sample and their frequency among each superpopulation (Europeans, Africans, East Asians, South Asians, Americans, and Brazilians) is given as supplemental material, Table S1. The phased VCF data is available upon request.

***KIR3DL3* genomic diversity, population differentiation, and linkage disequilibrium**

The 13Kb segment comprehending *KIR3DL3* presented 793 variable sites as listed in Table S1. Many of these variants are frequent in all superpopulations, but some are particularly frequent among Europeans, Africans, or East Asians. For instance, rs62132664 (exon 4) range from 1.1% among East Asians to 26.54% among Africans (Table S1). We detected more than a thousand full haplotypes, mostly because of the presence of many variants in intron 5. Intron 5 corresponds to 40.5% of all variable sites described in Table S1, many of which occurring in few individuals. Since we have not evaluated intron 1, and since there are many new variants described in this dataset when compared to the IPD-KIR database, mostly in intron 5, we will not track specific *KIR3DL3* genomic alleles.

In spite of some frequency differences among populations, most of the common variants are found worldwide. This is supported by a principal component analysis, with poor clusterization of each superpopulation except maybe the East Asia (EAS) samples (Figure 1). This different pattern observed for EAS is supported by a multidimension scale plot of the F_{ST} values among populations (Figure 2), illustrating a high differentiation of EAS samples (CHB, CHS, JPT, CDX, and KHV). Brazilians follow the same pattern observed for other admixed samples such as CLM (Colombia) and PUR (Puerto Rico).

We also assessed linkage disequilibrium (LD) across the *KIR3DL3* gene pooling all populations. There is a high LD comprehending all variants from the promoter to intron 5, and from intron 5 to the 3'UTR segment (Figure S1). This LD pattern is compatible with a recombination hotspot already reported in intron 5, and it may explain the high number of different full haplotypes observed in this study.

***KIR3DL3* worldwide protein diversity**

Altogether we detected 95 different *KIR3DL3* proteins, 51 of them already reported in the IPD-KIR database and with a summed frequency of 98.5%. The remaining 1.5% are new

proteins due to recombination or novel non-synonymous mutations. Encoded protein frequencies are given in Table 1. Some protein variants are highly frequent among Europeans and rare among Africans (e.g., KIR3DL3*001), others are frequent among East Asians and rare in others (e.g. KIR3DL3*010), and some are frequent worldwide (e.g., KIR3DL*004). The highest frequency difference is related to KIR3DL3*010, which is very common among EAS but uncommon in the rest of the world. Thus, KIR3DL3 protein diversity is strongly influenced by the ancestry background. In this matter, Brazilians follow the same pattern observed for other admixed populations, with a trend for the European pattern. The same pattern is observed for superpopulations (Table 1) and each population separately (Table S2).

We also characterized the amino acids in each polymorphic residue (Table 2 and S3) and their frequencies (Table 3) among superpopulations. We detected 16 residues that are truly polymorphic, with a minor allele frequency greater than 1% in at least one superpopulation (Table 3). KIR3DL3 amino acid frequencies are different among superpopulations, especially among EAS. Residue 77R, for instance, has a low frequency in any other population except the ones in East Asia. Residue 168I is twice as frequent among Asians than Africans. Residue 231D is an African variant.

Finally, the frequency of each CDS sequence (many encoding the same KIR3DL3 protein) is given in Table S4. The summed frequency of sequences with assigned names by the IPD-KIR database is 97.15%, and the remaining ones are related either to the new proteins described in Table S3 or they are associated with synonymous mutations.

***KIR3DL3* regulatory sequences: promoter and 3'UTR**

There were 61 different promoter haplotypes considering 2Kb upstream the first translated ATG (Figure S2). Since there is no nomenclature pattern established for KIR promoter sequences, and the IPD-KIR database reports only partial promoter sequence for each allele, we named haplotypes following sequence similarities and haplotype frequencies. We clustered promoter sequences into two groups (P01 and P02) since most of the variable sites are in absolute LD (Figure S1). There are essentially two different promoter lineages, P01, and P02, with summed frequencies of 66.94% and 33.06%, respectively. Nevertheless, only nine promoter haplotypes are frequent, seven from lineage P01 e two from lineage P02. The proximal and core promoter was virtually invariable, with only two frequent variants between positions -1 and -450 (Figure S2). There is a highly polymorphic segment between nucleotides -1414 and -1490. While promoters from lineage 1 are the most frequent among Europeans,

Africans, South Asians, and Admixed populations, both P01 and P02 are evenly distributed among East Asians (Table S5).

The 3'UTR segment is conserved among different populations since two haplotypes present a summed frequency of 92.31%. These haplotypes are frequent worldwide, but EAS presented the lowest frequency of 3UTR-02. Moreover, 83.27% of the EAS samples present the same 3'UTR segment. Since there is no nomenclature pattern established for *KIR* 3'UTR sequences, we named haplotypes following their frequencies (Figure S3). Following the promoter pattern, most of the 3'UTR variants are in absolute LD and associated with the same haplotype, here named 3UTR-03 (Figure S3), which is absent among East Asians.

***KIR3DL3* extended haplotypes: the association among promoters, coding, and 3'UTR haplotypes**

As presented earlier, because of the high number of uncommon variants in intron 5 (Table S1) and the presence of a recombination hotspot in intron 5 (Figure S1), we detected thousands of full haplotypes. The relationship between each promoter and the coding region can be obtained directly on the phased VCF data (available upon request). In general, each *KIR3DL3* allele encoding the same protein is associated with the same promoter sequence. For instance, more than 99.5% of all genomic sequences with *KIR3DL3**001, *003, *011, *012, *015 and others, are associated with promoters from lineage 1, while the opposite is observed for genomic sequences with *005, *010, and *019 (Table S6). Genomic sequences encoding *KIR3DL3**002 and 014 are associated with both promoter lineages. The relationship among 3'UTR haplotypes and encoded proteins are available in Table S7.

***KIR3DL3* evolutionary aspects in worldwide human populations**

First, we have evaluated nucleotide diversity in each *KIR3DL3* segment for all superpopulations (Table 4). Nucleotide diversity is low throughout *KIR3DL3*, especially in all extracellular domains (encoded by exons 3, 4, and 5). However, we detected a high nucleotide diversity in exon 6 (the transmembrane domain) for all superpopulations, and also in intron 6 and intron 7 for all samples except EAS. The leader peptide was highly conserved especially the segment encoded by exon 1. Moreover, none of the amino acid exchanges in the leader peptide is polymorphic, and all EAS samples presented the same leader peptide.

When we calculate Tajima's D for all *KIR3DL3* segments, we observed different patterns among superpopulations (Table 5). We observed high positive values compatible with

a signature of balancing selection in the promoter segment of Europeans and East Asians. However, this signal moves to intron 3 among Africans, and intron 6 among South Asians and admixed samples. Since Tajima's D may be influenced by the size of the segment, we calculated this statistic in 250bp sliding windows with a step size of 1. While EAS present the strong signals of balancing selection at the promoter round position -1470 and in intron 5, and for AFR the strongest signal is in intron 3, the rest of the superpopulations has a strong signal at intron 5 (Figure 3). The high positive values in the promoter for EAS is supported by the even frequency of both promoter lineages in this sample.

When we calculate the dN/dS ratio in each superpopulation (Table 6), we detected a trend to purifying selection in exon 5 (the D2 domain) among Europeans, Africans, and Admixed samples from the 1000Genomes project, and significant deviations towards purifying selection among Brazilians and in South Asia. Only one variable site is frequent in exon 5 (amino acid 290, Table 3). The pattern observed in EAS is different from all others.

Discussion

KIR genes are highly polymorphic considering both nucleotide diversity, copy number variation and haplotypes. They have suffered extension, contraction, and duplication over time, making them very similar to each other (Uhrberg *et al.*, 1997; Wilson *et al.*, 2000). Three genes and one pseudogene are considered frameworks (*KIR3DL3*, *KIR3DL2*, *KIR2DL4*, and *KIR3DP1*), with all individuals presenting two copies, while other *KIR* loci might be present or not. Thus, there are a variety of haplotypes (in terms of CNVs and polymorphic sites within), making unlikely two non-related individuals presenting the same sequence in the *KIR* complex (Augusto, 2016). The *KIR* allelic diversity is still underestimated as it became clear in this study, and the functional and evolutionary implications of this diversity are not yet completely understood (Augusto e Petzl-Erler, 2015).

KIR3DL3 is the most polymorphic *KIR* gene so far, with 165 alleles encoding 114 proteins according to the IPD-IMGT/*KIR* database version 2.9.0. *KIR3DL3* was present in all haplotypes already described (Pyo *et al.*, 2010; Pyo *et al.*, 2013; Roe *et al.*, 2017; Leaton *et al.*, 2019), and is very conserved throughout primate's evolution (Guethlein *et al.*, 2015), indicating a key role for *KIR3DL3* in immune function and modulation. However, its expression pattern is still unclear. Here we evaluated *KIR3DL3* genetic diversity in worldwide populations, including 26 population samples from the newly released high-coverage 1000Genomes data, and a sample from a Brazil, detecting 793 variable sites in 13,702 base pairs. Moreover, to minimize mapping bias

and genotyping errors, we have developed a bioinformatics pipeline lying on a new version of the hla-mapper software fed with a database of known *KIR* sequences and haplotypes (Castelli *et al.*, 2018), an SNP by SNP genotype call using GATK HaplotypeCaller, a variant refinement method created specifically for this approach, and a phasing method that combines either experimental phasing (data obtained directly in the sequencing data) and probabilistic models. This methodology allowed us to identify each *KIR3DL3* variant in regulatory segments and the coding region, except for intron 1 here a repetitive sequence jeopardize the genotyping procedure due to the low read size in the 1000Genome data.

The quality of the data obtained with our pipeline can be assigned by the following: the majority of variants and sequences here detected for exons had already been described by the IPD-KIR database; allele imputation was kept to a minimum; the translation of the full-haplotypes to predicted proteins reveals the presence of known *KIR3DL3* proteins with a summed frequency of 98.5%; there is good overlap between our data and the one reported by Leaton and colleagues using Exome sequencing from the 1000Genomes dataset (Leaton *et al.*, 2019) for exonic sequences.

However, we observed a *KIR3DL3* diversity much higher than the one reported in the IPD-IMGT/KIR database and even in the original data of the *1000Genomes* project. The presence of many singletons (Table S1) and many new variants in intronic segments show that the *KIR3DL3* genetic diversity is far underestimated, at least in regulatory segments and intronic regions. The use of a computational approach that optimizes NGS sequence mapping in the *KIR* cluster minimizes mapping errors and the detection of false positives, but also decrease false negatives. Because of that, we detected more variants in our Brazilian sample (2n=216) than in the entire 1000Genomes data from phase 3 in 2,504 samples (before the optimization performed here), and the majority of the Brazilian exonic sequences are compatible with the ones already reported, demonstrating that this gene is indeed more variable than we currently acknowledge. After the mapping optimization, this became clear since we detected thousands of full haplotypes mostly because of uncommon variants in intronic segments, and these haplotypes encode more than 70 different *KIR3DL3* proteins in real population samples. It is important to complete characterize and submit these sequences to the IPD-IMGT/KIR database. However, this high number of new sequences makes the cloning and sequencing process impracticable. Earlier studies from our research group also found inconsistency in the 1000Genomes data regarding other gene families, such as *HLA-C* and *KIR2DL4* (Weiss E. et al, personal communication).

KIR3DL3 present seven introns, and some of them are larger than the genes encoding HLA ligands (intron 5 is larger than an HLA class I gene, for instance). There were many intronic variants, most of them new considering the IPD-KIR database. The nucleotide diversity of introns 6 and 7 are among the highest ones in all super populations except EAS, which present a different pattern from all others. Nucleotide diversity at exon 6 is also high in all populations (Table 4). It is not clear whether the high LD in intron 6 and 7 are a hitchhiking effect because of the high LD among them and exon 6. Interestingly, intron 6 also presents the highest Tajima's D values among Brazilians, South Asians, and admixed populations from the 1000Genomes, and the second-highest among Africans and Europeans (Table 5). It is possible that this intron has some regulatory activity still to be unfolded, and balancing selection is maintaining divergent sequences in this segment. The Tajima's D analysis in sliding windows also indicates evidence of balancing selection in intronic segments. *KIR* genes might also present other regulatory segments than the promoter region. Previous studies have suggested that the first *KIR2DL4* intron, for instance, could overcome the function of the promoter region when silencing elements bond to the promoter (Trompeter *et al.*, 2005).

Apart from intronic variants, *KIR3DL3* is polymorphic mainly on its external domains, encoded by exons 3, 4, and 5 (Tables 2 and 3). In terms of the number of amino acid exchanges and frequency, domain D0 (exon 3) and the transmembrane segment (exon 6) are the most variable ones. Exon 6 presented the highest nucleotide diversity observed across the *KIR3DL3* CDS. It is not clear whether these amino acid exchanges modify the *KIR3DL3* function or its anchorage to the cell membrane. Interestingly, the three most frequent *KIR3DL3* proteins globally, but also among Brazilians, Europeans, and South Asians, present different transmembrane segments. Africans are also polymorphic in this segment, but with only two frequent protein versions. East Asians are the most conserved in this segment (Tables 2 and 3).

In terms of protein diversity, the most conserved segment is the leader peptide (exons 1 and 2) and the cytoplasmatic domain (exons 7 and 8). Variations in the peptide leader may influence signal recognition and decrease the number of mature receptors for transport to the cell surface (Vandenbussche *et al.*, 2009). For *KIR3DL3* these functions are maintained for all proteins encoded by this dataset, especially for EAS with no variation whatsoever.

To date, there is no available *KIR3DL3* crystallography to evaluate *KIR3DL3* tertiary structure and which amino acids are important for binding. Considering a *KIR3DL1*/HLA-Bw4 binding (O'connor *et al.*, 2014) crystallography (3VH8), and *KIR3DL1* also presents three extracellular domains as *KIR3DL3*, we observed three key bind regions, one in D1 and two in

D2. Here we detected a signature of purifying selection in exon 5 among Brazilians, South Asians, and a trend among Africans, Europeans and admixed populations. Exon 5 encodes the extracellular domain D2. Thus, D2 conservation may be explained maintaining key amino acid residues for specific-ligand binding and to the KIR3DL3 modulatory function. Since D1 and D2 are conserved, with very few frequent amino acid exchanges, we may expect that the receptors binding to KIR3DL3 are also conserved in terms of protein sequence, such as the non-classical HLA molecules HLA-E, HLA-G, and HLA-F (Robinson *et al.*, 2015). However, it is not clear which molecules bind to KIR3DL3. In contrast, a previous study detected the opposite, i.e., a signature of positive selection and a higher rate of non-synonymous exchanges in this segment, but here we are addressing a real population sample with the KIR3DL3 alleles that occur worldwide, while the previous study addressed KIR3DL3 variability mostly in databases of known sequences (Leaton *et al.*, 2019).

An important KIR3DL3 feature is that allele frequencies are strongly influenced by the ancestry background. This became evident when comparing CDS sequences (Table S4), protein sequences (Table 1), each protein residue (Table 3), and the F_{ST} among populations (Figure 2). Some variants are frequent in Africa, others are frequent in Europe, and others are frequent in Asia. For instance, the major KIR3DL3 protein in East Asia is KIR3DL3*010 and *008 (Table 1), and they are uncommon in other parts of the world. These proteins present two amino acid exchanges that are highly frequent in East Asia (52H and 77W, Table 3), both in the D0 domain. Likewise, residue 321N is common in all populations except Africans, and residue 348P is common worldwide except in East Asia. These frequency differences among populations are supported by previous studies. For instance, KIR3DL3*005 is frequent in a population from Ghana (Hou *et al.*, 2007; Norman *et al.*, 2013; Gonzalez-Galarza *et al.*, 2015) and KIR3DL3*010 is frequent among East Asian samples, including Chinese (Leaton *et al.*, 2019). We detected the influence of ancestry on allele frequencies even in Brazil. Brazilians are highly admixed, with contributions of the native Amerindians, European immigrants, and Africans (mostly brought centuries ago because of slavery) (Ferreira *et al.*, 2006). According to our AIM analysis, our samples are mostly Europeans, but with some individuals with a major African or Asian/Amerindian background. KIR3DL3*005, for instance, is more frequent among samples with a higher African background compared to Europeans ($P = 0.0488$), while KIR3DL3*10 is quite frequent in samples with a major Asian/Amerindian background compared to Europeans ($P=0.0004$). Moreover, some *KIR3DL3* alleles are linked to specific KIR haplotypes. For instance, KIR3DL3*009 (common worldwide) is usually associated with

haplotype A, while KIR3DL3*010 (common only in EAS) with haplotype B (Pyo *et al.*, 2010). Haplotypes A and B have distinct combinations of activating and inhibitory *KIR* genes and may present different immunomodulation properties. Nevertheless, it is not clear whether these differences influence KIR3DL3 function or the expression profile.

We should not rule out co-evolution between KIR3DL3 (and other KIR genes) and HLA genes, which encodes their main ligands, influencing the frequency shifts observed among populations (Augusto e Petzl-Erler, 2015). When we reevaluate the *HLA-C* data from 1000Genomes using the same pipeline described here and calculating F_{ST} among each population (data not shown), we detected the same pattern as observed for KIR3DL3, i.e., a great differentiation between EAS and the rest of the world, and a complete separation of Europeans and African populations. Thus, KIR3DL3 protein diversity might be influenced by HLA protein diversity, and vice-versa, keeping combinations with a greater fit in terms of ligand/receptor interaction.

Our study also evaluated the KIR3DL3 regulatory diversity in worldwide populations. KIR3DL3 promoter region has little homology with other KIR promoters, which in part explains the distinct expression pattern of this gene (Li *et al.*, 2016a). KIR3DL3 promoter region presented two major promoter lineages, with are frequent worldwide but presenting even frequencies among East Asian samples. We found evidence of balancing selection acting on the promoter segment, particularly among EAS and EUR (Table 5). This signal is stronger in a highly polymorphic promoter segment around position -1470 (Figure 3), which presents many variants in absolute LD (Figure S1). This may indicate the maintenance of two different transcriptional and regulation patterns, in which variable sites influence the methylation profile, the binding of transcriptional factors, or even chromatin assembly, leading to different expression levels. *HLA-A*, *HLA-G*, and *HLA-F* present this same pattern of balancing selection in their promoters (Castelli *et al.*, 2011; Buttura *et al.*, 2019; Lima *et al.*, 2019). Nevertheless, when we consider the transcription factors that co-express along with KIR3DL3, and the binding sites that are created or destroyed (data not shown), we found no evidence of a differential function between lineage 1 and 2. Moreover, although KIR3DL3 present two divergent promoter lineages, there were no frequent variations in its proximal and core promoters, with the exception, maybe, of the variant rs4806534 (19:54724332, position -165). The proximal promoter is usually methylated and this is related to the low mRNA and protein levels usually detected for this gene (Chan *et al.*, 2003; Trompeter *et al.*, 2005). KIR3DL3 protein expression is weak or undetected in peripheral blood because of epigenetic silencing

that was densely seen in NK3.3 and CD34⁺ (Trompeter *et al.*, 2005). Variants in the intermediate promoter (~300pb upstream) may affect gene expression, as observed for KIR2DL1 in which a deletion in a specific allele decreases the promoter activity reducing the expression levels (Wright *et al.*, 2014). The intermediate promoter is conserved in all *KIR* genes except in KIR3DL3 (Wright *et al.*, 2014), and some nucleotide exchanges in this region are frequent. As discussed elsewhere (Trundley *et al.*, 2006; Li *et al.*, 2016a), there may be a collaboration between the proximal and intermediate promoters influencing *KIR3DL3* expression levels, with activation and demethylation of the proximal promoter following the activation of the intermediate one. Besides that, exposure to a methyltransferase inhibitor induce KIR3DL3 expression, showing the importance of this mechanism on gene expression (Trompeter *et al.*, 2005). Variant *rs4806534* (position -165) may influence the proximal promoter methylation profile since it creates or destroys a CpG site, and both alleles are frequent worldwide, with the CpG site conserved among P02 promoters and destroyed among P01. This variant is in absolute LD with the ones around position -1470, where the strongest signal for balancing selection is located.

Gene expression may also be influenced by the 3'UTR segment, in which variable sites can influence the binding of miRNAs. MicroRNAs have an important role in regulating processes in Natural Killer cells (Bezman *et al.*, 2010). *KIR3DL3* presents a 3'UTR sequence very similar to the ones of other *KIR* genes (Nutalai *et al.*, 2019). Here we detected few variants in this segment, and most of them are related to a single haplotype 3UTR-03 (Figure S3). A recent study showed that *KIR3DL3* is also regulated by miRNAs, and this regulation may be related to the undetectable gene expression on the cell surface (Nutalai *et al.*, 2019). The binding sites for miRNAs miR-26, miR-185-5p, and miR-203a-3p, known to regulate KIR3DL3 expression, are conserved worldwide. However, the binding sites for other miRNAs are disrupted. For instance, when combining data from three difference algorithms to minimize false positives, we find 8 miRNAs that would bind differently to the three most frequent 3'UTR sequences. Of those, five binds preferentially to 3UTR-01 (miR-376a-5p, miR-1289, and miR-5007-5p, miR-5683, and miR-1273h-3p), and three binds preferentially to 3UTR-03 (miR-4478, miR-3161, and miR-6815-5p). At least miR-376a-5p is highly expressed in placenta.

KIR3DL3 is of great clinical interest because it always encodes a full-length protein (Leaton *et al.*, 2019) (only a rare protein variant has an in-frame indel and lacks one amino acid), and all individuals have two copies of the gene. Moreover, it is conserved in primates, thus it may present an important function for immune responses (Leaton *et al.*, 2019). Most of

the studies that looked for *KIR3DL3* associations with pathologies and infections used molecular techniques that indicate only the presence or absence of the gene (Ordonez *et al.*, 2012; Bossi *et al.*, 2015; Wang *et al.*, 2018), disregarding its inner genetic variability (Augusto, 2016). Because of that, there is few description of *KIR3DL3* allele associated with diseases (Takeshita *et al.*, 2013). This gene is probably expressed in some pathological situations or during development (Nutalai *et al.*, 2019). Understanding its inner genetic diversity may indicate variants and regions of interest for clinical and disease-association studies. Moreover, since protein frequencies (and specific amino acid exchanges) shift among samples with different ancestry backgrounds, such studies must control for population-stratification when evaluating *KIR3DL3*.

Here we addressed *KIR3DL3* genetic diversity and evolutionary aspects in worldwide human populations, using a bioinformatics approach that minimizes mapping bias and genotyping errors. *KIR3DL3* genetic variability is far greater than we currently acknowledge, especially when we consider intronic segments and regulatory sequences. *KIR3DL3* protein diversity is strongly influenced by the ancestry background, with different frequencies of protein versions and amino acid residues among Africans, Europeans, and other populations. In this matter, populations from East Asia present a different pattern considering protein frequencies, regulatory sequences diversity, and how natural selection has shaped *KIR3DL3* diversity and haplotype frequencies. Because of that, any association study involving *KIR3DL3* should be performed concerning for population stratification. Some *KIR3DL3* segments are highly conserved, such as the exons encoding the leader peptide and the extracellular domain D2, important for HLA interaction, while other segments, such as the promoter region and exon 6 (the transmembrane domain) are highly variable. We found evidence of balancing selection in the promoter region, probably maintaining sequences with different promotion capabilities. This evidence is stronger among Europeans and East Asians. We also find evidence of balancing selection in some large intronic segments, and they must be considered as possibly carrying regulatory elements. Finally, here we present a comprehensive dataset of the *KIR3DL3* genetic diversity in worldwide populations, and a method suitable to evaluate this gene in whole-genome, whole-exome, and targeted sequencing.

Conflicts of interest

The authors declare that they have no conflict of interest.

Data availability statement

The data that support the findings of this study are available from the corresponding author upon request, in the format of a phased Variant Call Format (VCF) file.

Acknowledgements

This work was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo – FAPESP/Brazil (Grant 2017/19223-0). This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) – Finance Code 001 and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq/Brazil) (Grant# 302590/2016-1).

REFERENCES

ABI-RACHED, L.; PARHAM, P. Natural selection drives recurrent formation of activating killer cell immunoglobulin-like receptor and Ly49 from inhibitory homologues. **J Exp Med**, v. 201, n. 8, p. 1319-32, Apr 18 2005. ISSN 0022-1007.

AKBARI, S. et al. Recurrent Spontaneous Abortion (RSA) and Maternal KIR Genes: A Comprehensive Meta-Analysis. **JBRA Assist Reprod**, Feb 12 2020. ISSN 1517-5693.

AUGUSTO, D. G. The Impact of KIR Polymorphism on the Risk of Developing Cancer: Not as Strong as Imagined? **Front Genet**, v. 7, p. 121, 2016. ISSN 1664-8021.

AUGUSTO, D. G.; HOLLENBACH, J. A.; PETZL-ERLER, M. L. A deep look at KIR-HLA in Amerindians: comprehensive meta-analysis reveals limited diversity of KIR haplotypes. **Hum Immunol**, v. 76, n. 4, p. 272-80, Apr 2015. ISSN 0198-8859.

AUGUSTO, D. G. et al. Activating KIR and HLA Bw4 ligands are associated to decreased susceptibility to pemphigus foliaceus, an autoimmune blistering skin disease. **PLoS One**, v. 7, n. 7, p. e39991, 2012. ISSN 1932-6203.

AUGUSTO, D. G.; PETZL-ERLER, M. L. KIR and HLA under pressure: evidences of coevolution across worldwide populations. **Human Genetics**, v. 134, n. 9, p. 929-940, Sep 2015. ISSN 0340-6717.

AUGUSTO, D. G. et al. Diversity of the KIR gene cluster in an urban Brazilian population. **Immunogenetics**, v. 64, n. 2, p. 143-52, Feb 2012. ISSN 0093-7711.

BARRETT, J. C. et al. Haploview: analysis and visualization of LD and haplotype maps. **Bioinformatics**, v. 21, n. 2, p. 263-265, Jan 2005. ISSN 1367-4803.

BARTEL, D. P. MicroRNAs: target recognition and regulatory functions. **Cell**, v. 136, n. 2, p. 215-33, Jan 23 2009. ISSN 0092-8674.

BASHIROVA, A. A. et al. The killer immunoglobulin-like receptor gene cluster: tuning the genome for defense. **Annu Rev Genomics Hum Genet**, v. 7, p. 277-300, 2006. ISSN 1527-8204

BESSON, C. et al. Association of killer cell immunoglobulin-like receptor genes with Hodgkin's lymphoma in a familial study. **PLoS One**, v. 2, n. 5, p. e406, May 2 2007. ISSN 1932-6203.

BEZMAN, N. A. et al. Distinct requirements of microRNAs in NK cell activation, survival, and function. **J Immunol**, v. 185, n. 7, p. 3835-46, Oct 1 2010. ISSN 0022-1767.

BOSSI, G. et al. Genetic epistasis between killer immunoglobulin-like receptors and human leukocyte antigens in Kawasaki disease susceptibility. **Genes Immun**, v. 16, n. 7, p. 481-7, Oct 2015. ISSN 1466-4879.

BOTTINO, C. et al. Cellular ligands of activating NK receptors. **Trends Immunol**, v. 26, n. 4, p. 221-6, Apr 2005. ISSN 1471-4906

BOUDREAU, J. E.; HSU, K. C. Natural Killer Cell Education and the Response to Infection and Cancer Therapy: Stay Tuned. **Trends in Immunology**, v. 39, n. 3, p. 222-239, Mar 2018a. ISSN 1471-4906.

_____. Natural killer cell education in human health and disease. **Current Opinion in Immunology**, v. 50, p. 102-111, Feb 2018b. ISSN 0952-7915.

BROWNING, S. R.; BROWNING, B. L. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. **Am J Hum Genet**, v. 81, n. 5, p. 1084-97, Nov 2007. ISSN 0002-9297

BUTTURA, R. V. et al. HLA-F displays highly divergent and frequent haplotype lineages associated with different mRNA expression levels. **Hum Immunol**, v. 80, n. 2, p. 112-119, Feb 2019. ISSN 0198-8859.

CAMPBELL, K. S.; PURDY, A. K. Structure/function of human killer cell immunoglobulin-like receptors: lessons from polymorphisms, evolution, crystal structures and mutations. **Immunology**, v. 132, n. 3, p. 315-325, Mar 2011. ISSN 0019-2805.

CASTELLI, E. C. et al. A Comprehensive Study of Polymorphic Sites along the HLA-G Gene: Implication for Gene Regulation and Evolution. **Molecular Biology and Evolution**, v. 28, n. 11, p. 3069-3086, Nov 2011. ISSN 0737-4038.

CASTELLI, E. C. et al. Hla-mapper: An application to optimize the mapping of HLA sequences produced by massively parallel sequencing procedures. **Hum Immunol**, v. 79, n. 9, p. 678-684, Sep 2018. ISSN 0198-8859.

CHAN, H. W. et al. DNA methylation maintains allele-specific KIR gene expression in human natural killer cells. **J Exp Med**, v. 197, n. 2, p. 245-55, Jan 20 2003. ISSN 0022-1007.

DELANEAU, O.; MARCHINI, J.; ZAGURY, J. F. A linear complexity phasing method for thousands of genomes. **Nat Methods**, v. 9, n. 2, p. 179-81, Dec 4 2011. ISSN 1548-7091.

DIAZ-PENA, R. et al. Analysis of Killer Immunoglobulin-Like Receptor Genes in Colorectal Cancer. **Cells**, v. 9, n. 2, Feb 24 2020. ISSN 2073-4409.

DUPONT, B.; SELVAKUMAR, A.; STEFFENS, U. The killer cell inhibitory receptor genomic region on human chromosome 19q13.4. **Tissue Antigens**, v. 49, n. 6, p. 557-63, Jun 1997. ISSN 0001-2815.

EXCOFFIER, L.; LISCHER, H. E. L. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. **Molecular Ecology Resources**, v. 10, n. 3, p. 564-567, May 2010. ISSN 1755-098X.

FARIAS, T. D. J. et al. Screening the full leucocyte receptor complex genomic region revealed associations with pemphigus that might be explained by gene regulation. **Immunology**, v. 156, n. 1, p. 86-93, Jan 2019. ISSN 0019-2805.

FERREIRA, L. B. et al. Genomic ancestry of a sample population from the state of Sao Paulo, Brazil. **American Journal of Human Biology**, v. 18, n. 5, p. 702-705, Sep-Oct 2006. ISSN 1042-0533.

FONDEVILA, M. et al. Revision of the SNPforID 34-plex forensic ancestry test: Assay enhancements, standard reference sample genotypes and extended population studies. **Forensic Science International-Genetics**, v. 7, n. 1, p. 63-74, Jan 2013. ISSN 1872-4973.

FU, W.; AKEY, J. M. Selection and adaptation in the human genome. **Annu Rev Genomics Hum Genet**, v. 14, p. 467-89, 2013. ISSN 1527-8204.

GERAGHTY, D. E. et al. The HLA class I gene family includes at least six genes and twelve pseudogenes and gene fragments. **J Immunol**, v. 149, n. 6, p. 1934-46, Sep 15 1992. ISSN 0022-1767

GONZALEZ-GALARZA, F. F. et al. Allele frequency net 2015 update: new features for HLA epitopes, KIR and disease and HLA adverse drug reaction associations. **Nucleic Acids Res**, v. 43, n. Database issue, p. D784-8, Jan 2015. ISSN 0305-1048.

GUETHLEIN, L. A. et al. Co-evolution of MHC class I and variable NK cell receptors in placental mammals. **Immunol Rev**, v. 267, n. 1, p. 259-82, Sep 2015. ISSN 0105-2896.

HOU, L. et al. Conserved KIR allele-level haplotypes are altered by microvariation in individuals with European ancestry. **Genes and Immunity**, v. 13, n. 1, p. 47-58, Jan 2012. ISSN 1466-4879.

HOU, L. et al. Seventeen novel alleles add to the already extensive KIR3DL3 diversity. **Tissue Antigens**, v. 70, n. 6, p. 449-454, Dec 2007. ISSN 0001-2815.

HSU, K. C. et al. The killer cell immunoglobulin-like receptor (KIR) genomic region: gene-order, haplotypes and allelic polymorphism. **Immunological Reviews**, v. 190, n. 1, p. 40-52, Dec 2002. ISSN 0105-2896.

HSU, K. C. et al. 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. **Journal of Immunology**, v. 169, n. 9, p. 5118-5129, Nov 2002. ISSN 0022-1767.

HUHN, O. et al. High-Resolution Genetic and Phenotypic Analysis of KIR2DL1 Alleles and Their Association with Pre-Eclampsia. **J Immunol**, v. 201, n. 9, p. 2593-2601, Nov 1 2018. ISSN 0022-1767.

HUMPHREY, M. B.; LANIER, L. L.; NAKAMURA, M. C. Role of ITAM-containing adapter proteins and their receptors in the immune system and bone. **Immunological Reviews**, v. 208, p. 50-65, Dec 2005. ISSN 0105-2896.

JAGER, A.; KUCHROO, V. K. Effector and Regulatory T-cell Subsets in Autoimmunity and Tissue Inflammation. **Scandinavian Journal of Immunology**, v. 72, n. 3, p. 173-184, Sep 2010. ISSN 0300-9475.

JONES, D. C. et al. Nature of allelic sequence polymorphism at the KIR3DL3 locus. **Immunogenetics**, v. 58, n. 8, p. 614-627, Aug 2006. ISSN 0093-7711.

KARRE, K. et al. Selective rejection of H-2-deficient lymphoma variants suggests alternative immune defence strategy. **Nature**, v. 319, n. 6055, p. 675-8, Feb 20-26 1986. ISSN 0028-0836

KEY, F. M. et al. Advantageous diversity maintained by balancing selection in humans. **Curr Opin Genet Dev**, v. 29, p. 45-51, Dec 2014. ISSN 0959-437x.

KHAKOO, S. I. et al. Rapid evolution of NK cell receptor systems demonstrated by comparison of chimpanzees and humans. **Immunity**, v. 12, n. 6, p. 687-98, Jun 2000. ISSN 1074-7613

KULKARNI, S.; MARTIN, M. P.; CARRINGTON, M. The Yin and Yang of HLA and KIR in human disease. **Seminars in Immunology**, v. 20, n. 6, p. 343-352, Dec 2008. ISSN 1044-5323.

KUMAR, S.; STECHER, G.; TAMURA, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. **Mol Biol Evol**, v. 33, n. 7, p. 1870-4, Jul 2016. ISSN 0737-4038.

LEATON, L. A. et al. Conservation, Extensive Heterozygosity, and Convergence of Signaling Potential All Indicate a Critical Role for KIR3DL3 in Higher Primates. **Front Immunol**, v. 10, p. 24, 2019. ISSN 1664-3224.

LI, H. et al. Characterization of KIR intermediate promoters reveals four promoter types associated with distinct expression patterns of KIR subtypes. **Genes Immun**, v. 17, n. 1, p. 66-74, Jan-Feb 2016a. ISSN 1466-4879.

_____. Characterization of KIR intermediate promoters reveals four promoter types associated with distinct expression patterns of KIR subtypes. **Genes and Immunity**, v. 17, n. 1, p. 66-74, Jan-Feb 2016b. ISSN 1466-4879.

LIMA, T. H. A. et al. HLA-A promoter, coding, and 3'UTR sequences in a Brazilian cohort, and their evolutionary aspects. **Hla**, v. 93, n. 2-3, p. 65-79, Feb 2019. ISSN 2059-2302.

MARSH, S. G. E. et al. Killer-cell immunoglobulin-like receptor (KIR) nomenclature report, 2002. **Journal of Immunological Methods**, v. 281, n. 1-2, p. 1-8, Oct 2003. ISSN 0022-1759.

MARTINEZ-FIERRO, M. L. et al. Ancestry informative markers and admixture proportions in northeastern Mexico. **J Hum Genet**, v. 54, n. 9, p. 504-9, Sep 2009. ISSN 1434-5161.

MCKENNA, A. et al. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. **Genome Research**, v. 20, n. 9, p. 1297-1303, Sep 2010. ISSN 1088-9051.

MICHELET, X. et al. Metabolic reprogramming of natural killer cells in obesity limits antitumor responses. **Nat Immunol**, v. 19, n. 12, p. 1330-1340, Dec 2018. ISSN 1529-2908.

MORETTA, A. et al. Early liaisons between cells of the innate immune system in inflamed peripheral tissues. **Trends Immunol**, v. 26, n. 12, p. 668-75, Dec 2005. ISSN 1471-4906

NEMAT-GORGANI, N. et al. Diversity of KIR, HLA Class I, and Their Interactions in Seven Populations of Sub-Saharan Africans. **J Immunol**, v. 202, n. 9, p. 2636-2647, May 1 2019. ISSN 0022-1767.

NORMAN, P. J. et al. Co-evolution of Human Leukocyte Antigen (HLA) Class I Ligands with Killer-Cell Immunoglobulin-Like Receptors (KIR) in a Genetically Diverse Population of Sub-Saharan Africans. **Plos Genetics**, v. 9, n. 10, Oct 2013. ISSN 1553-7404.

NORMAN, P. J. et al. Defining KIR and HLA Class I Genotypes at Highest Resolution via High-Throughput Sequencing. **American Journal of Human Genetics**, v. 99, n. 2, p. 375-391, Aug 2016. ISSN 0002-9297.

NUTALAI, R. et al. Regulation of KIR3DL3 Expression via Mirna. **Genes (Basel)**, v. 10, n. 8, Aug 9 2019. ISSN 2073-4425.

O'CONNOR, G. M. et al. Mutational and structural analysis of KIR3DL1 reveals a lineage-defining allotypic dimorphism that impacts both HLA and peptide sensitivity. **J Immunol**, v. 192, n. 6, p. 2875-84, Mar 15 2014. ISSN 0022-1767.

ORDONEZ, D. et al. KIR typing by non-sequencing methods: polymerase-chain reaction with sequence-specific primers. **Methods Mol Biol**, v. 882, p. 415-30, 2012. ISSN 1064-3745.

ORR, M. T.; LANIER, L. L. Natural Killer Cell Education and Tolerance. **Cell**, v. 142, n. 6, p. 847-856, Sep 2010. ISSN 0092-8674.

PARHAM, P. MHC class I molecules and KIRs in human history, health and survival. **Nature Reviews Immunology**, v. 5, n. 3, p. 201-214, Mar 2005. ISSN 1474-1733.

PENA, S. D. et al. The genomic ancestry of individuals from different geographical regions of Brazil is more uniform than expected. **PLoS One**, v. 6, n. 2, p. e17063, Feb 16 2011. ISSN 1932-6203.

PEREIRA, R. et al. Straightforward inference of ancestry and admixture proportions through ancestry-informative insertion deletion multiplexing. **PLoS One**, v. 7, n. 1, p. e29684, 2012. ISSN 1932-6203.

PESCE, S. et al. New miRNA Signature Heralds Human NK Cell Subsets at Different Maturation Steps: Involvement of miR-146a-5p in the Regulation of KIR Expression. **Front Immunol**, v. 9, p. 2360, 2018. ISSN 1664-3224.

PHILLIPS, C. Forensic genetic analysis of bio-geographical ancestry. **Forensic Science International-Genetics**, v. 18, p. 49-65, Sep 2015. ISSN 1872-4973.

PHILLIPS, C. et al. Inferring ancestral origin using a single multiplex assay of ancestry-informative marker SNPs. **Forensic Sci Int Genet**, v. 1, n. 3-4, p. 273-80, Dec 2007. ISSN 1872-4973.

POSTH, C. et al. Reconstructing the Deep Population History of Central and South America. **Cell**, v. 175, n. 5, p. 1185-1197.e22, Nov 15 2018. ISSN 0092-8674.

PRITCHARD, J. K.; STEPHENS, M.; DONNELLY, P. Inference of population structure using multilocus genotype data. **Genetics**, v. 155, n. 2, p. 945-959, Jun 2000. ISSN 0016-6731.

PURCELL, S. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. **Am J Hum Genet**, v. 81, n. 3, p. 559-75, Sep 2007. ISSN 0002-9297

PURDY, A. K.; CAMPBELL, K. S. Natural killer cells and cancer Regulation by the killer cell Ig-like receptors (KIR). **Cancer Biology & Therapy**, v. 8, n. 23, p. 2211-2220, Dec 2009. ISSN 1538-4047.

PYO, C. W. et al. Different Patterns of Evolution in the Centromeric and Telomeric Regions of Group A and B Haplotypes of the Human Killer Cell Ig-Like Receptor Locus. **Plos One**, v. 5, n. 12, Dec 2010. ISSN 1932-6203.

PYO, C. W. et al. Recombinant structures expand and contract inter and intragenic diversification at the KIR locus. **BMC Genomics**, v. 14, p. 89, Feb 8 2013. ISSN 1471-2164.

QI, Y. et al. KIR/HLA pleiotropism: protection against both HIV and opportunistic infections. **PLoS Pathog**, v. 2, n. 8, p. e79, Aug 2006. ISSN 1553-7366.

ROBINSON, J. et al. The IPD and IMGT/HLA database: allele variant databases. **Nucleic Acids Res**, v. 43, n. Database issue, p. D423-31, Jan 2015. ISSN 0305-1048.

ROE, D. et al. Revealing complete complex KIR haplotypes phased by long-read sequencing technology. **Genes Immun**, v. 18, n. 3, p. 127-134, Sep 2017. ISSN 1466-4879.

SINGLE, R. M. et al. Methods for assessing gene content diversity of KIR with examples from a global set of populations. **Immunogenetics**, v. 60, n. 12, p. 711-725, Dec 2008. ISSN 0093-7711.

STULBERG, M. J. et al. Identification of distal KIR promoters and transcripts. **Genes Immun**, v. 8, n. 2, p. 124-30, Mar 2007. ISSN 1466-4879.

SUN, H. S. et al. Disease-association of different killer cell immunoglobulin-like receptors (KIR) and HLA-C gene combinations in reactive arthritis. **Mod Rheumatol**, v. 29, n. 3, p. 531-537, May 2019. ISSN 1439-7595.

SUTO, Y. et al. Chromosomal localization of the human natural killer cell class I receptor family genes to 19q13.4 by fluorescence in situ hybridization. **Genomics**, v. 35, n. 1, p. 270-272, Jul 1996. ISSN 0888-7543.

TAKESHITA, L. Y. et al. A database for curating the associations between killer cell immunoglobulin-like receptors and diseases in worldwide populations. **Database (Oxford)**, v. 2013, p. bat021, 2013. ISSN 1758-0463.

TONEVA, M. et al. Genomic diversity of natural killer cell receptor genes in three populations. **Tissue Antigens**, v. 57, n. 4, p. 358-62, Apr 2001. ISSN 0001-2815.

TORKAR, M. et al. Isotypic variation of novel immunoglobulin-like transcript/killer cell inhibitory receptor loci in the leukocyte receptor complex. **Eur J Immunol**, v. 28, n. 12, p. 3959-67, Dec 1998. ISSN 0014-2980.

TROMPETER, H. I. et al. Three structurally and functionally divergent kinds of promoters regulate expression of clonally distributed killer cell Ig-like receptors (KIR), of KIR2DL4, and of KIR3DL3. **Journal of Immunology**, v. 174, n. 7, p. 4135-4143, Apr 2005. ISSN 0022-1767.

TROWSDALE, J.; MOFFETT, A. NK receptor interactions with MHC class I molecules in pregnancy. **Semin Immunol**, v. 20, n. 6, p. 317-20, Dec 2008. ISSN 1044-5323.

TRUNDLEY, A. E. et al. Molecular characterization of KIR3DL3. **Immunogenetics**, v. 57, n. 12, p. 904-916, Jan 2006. ISSN 0093-7711.

UHRBERG, M. et al. Human diversity in killer cell inhibitory receptor genes. **Immunity**, v. 7, n. 6, p. 753-763, Dec 1997. ISSN 1074-7613.

VAN BERGEN, J. et al. Structural and functional differences between the promoters of independently expressed killer cell Ig-like receptors. **Eur J Immunol**, v. 35, n. 7, p. 2191-9, Jul 2005. ISSN 0014-2980.

VAN DER AUWERA, G. A. et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. **Curr Protoc Bioinformatics**, v. 43, p. 11.10.1-33, 2013. ISSN 1934-3396.

VANDENBUSSCHE, C. J. et al. Dramatically reduced surface expression of NK cell receptor KIR2DS3 is attributed to multiple residues throughout the molecule. **Genes Immun**, v. 10, n. 2, p. 162-73, Mar 2009. ISSN 1466-4879.

VENDELBOSCH, S. et al. Extensive Variation in Gene Copy Number at the Killer Immunoglobulin-Like Receptor Locus in Humans. **Plos One**, v. 8, n. 6, Jun 2013. ISSN 1932-6203.

VERHEYDEN, S.; BERNIER, M.; DEMANET, C. Identification of natural killer cell receptor phenotypes associated with leukemia. **Leukemia**, v. 18, n. 12, p. 2002-7, Dec 2004. ISSN 0887-6924

VILCHES, C.; PARHAM, P. KIR: Diverse, rapidly evolving receptors of innate and adaptive immunity. **Annual Review of Immunology**, v. 20, p. 217-251, 2002. ISSN 0732-0582.

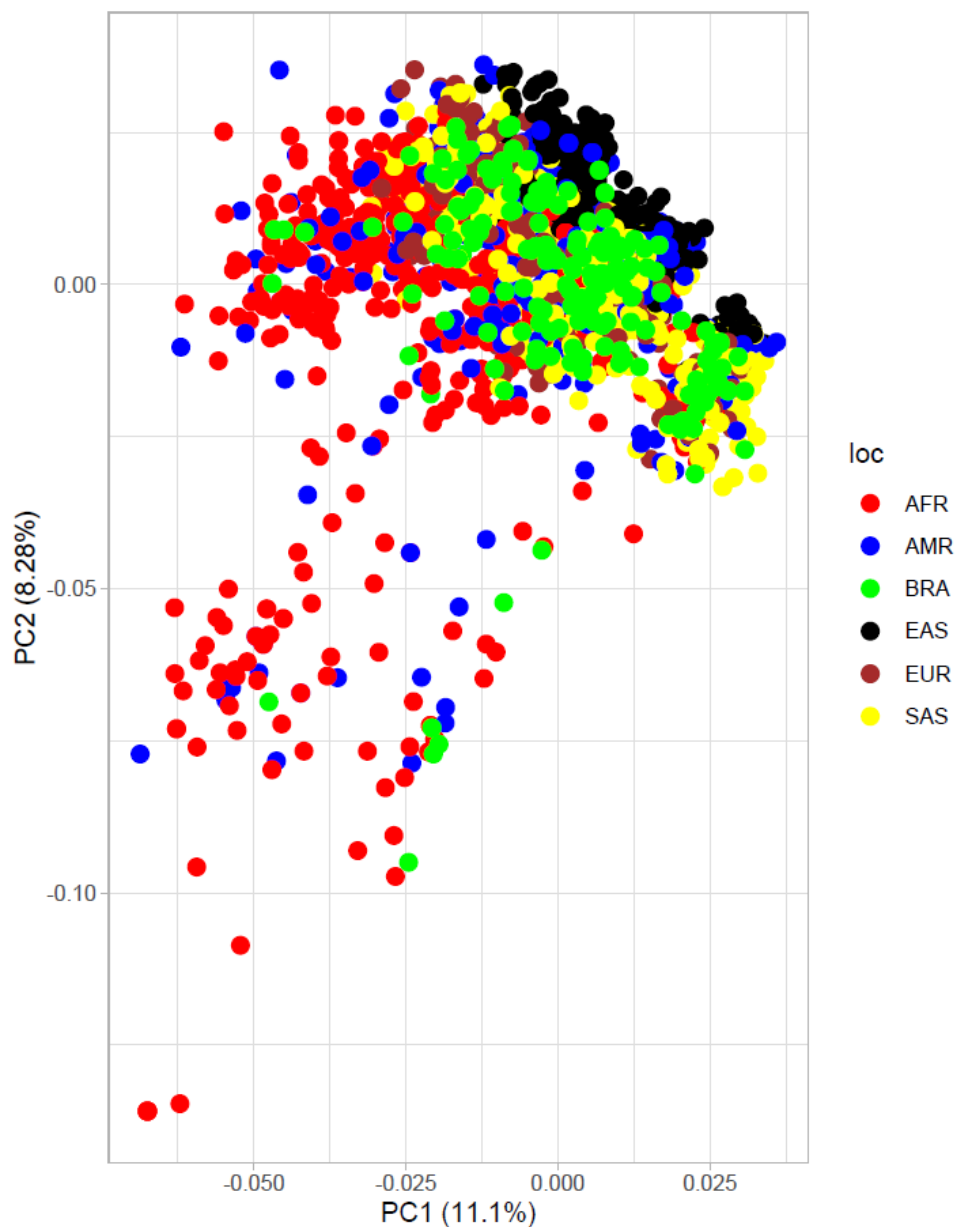
VIVIER, E. et al. Innate or Adaptive Immunity? The Example of Natural Killer Cells. **Science**, v. 331, n. 6013, Jan 2011. ISSN 0036-8075.

WANG, L. et al. Killer-cell immunoglobulin-like receptors associate with HIV-1 infection in a narrow-source Han Chinese cohort. **PLoS One**, v. 13, n. 4, p. e0195452, 2018. ISSN 1932-6203.

WILSON, M. J. et al. Plasticity in the organization and sequences of human KIR/ILT gene families. **Proc Natl Acad Sci U S A**, v. 97, n. 9, p. 4778-83, Apr 25 2000. ISSN 0027-8424 (Print) 0027-8424.

WRIGHT, P. W. et al. Characterization of a weakly expressed KIR2DL1 variant reveals a novel upstream promoter that controls KIR expression. **Genes Immun**, v. 15, n. 7, p. 440-8, Oct 2014. ISSN 1466-4879.

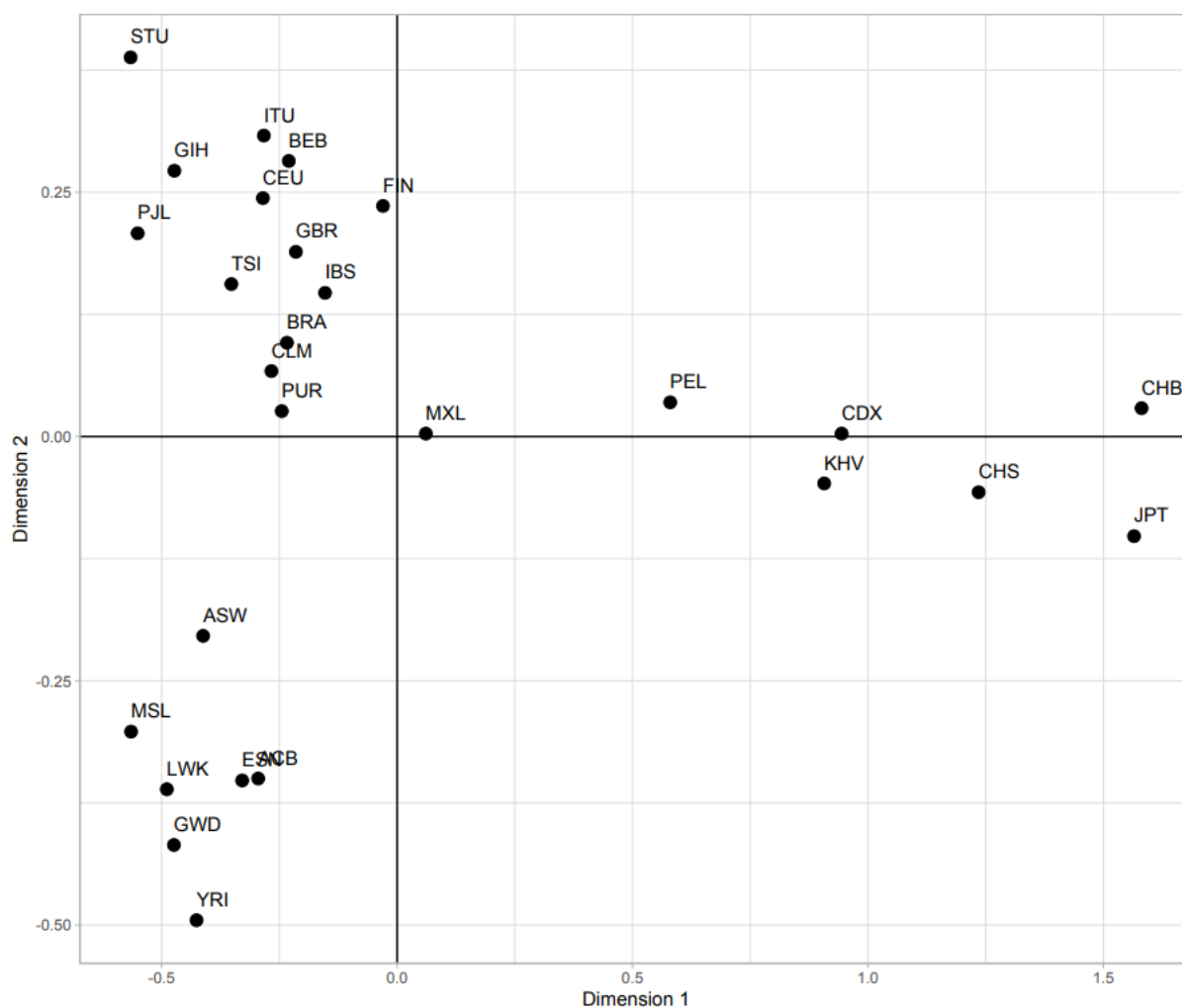
Figure 1. Principal component analysis (PCA) considering variable sites in the KIR3DL3 locus. We pruned variants with a $r^2 < 0.2$. Each superpopulation is represented by a different color: red for Africans, blue for Americans, green for Brazilians, black for East Asians, brown for Europeans, and yellow for South Asians.



European samples: CEU, TSI, FIN, GBR, IBS; African samples: YRI, LWK, GWD, MSL, ESN; East Asian samples: CHB, JPT, CHS, CDX, KHV; South Asian samples: GIH, PJL, BEB, STU, ITU; Admixed samples *: MXL, PUR, CLM, PEL, ASW, ACB. Since Brazilians (BRA) are not included in the 1000Genomes project and since we used a different sequencing method for Brazilians, this sample is presented isolated from the other admixed populations.

For a description of each of these population samples from the 1000Genomes dataset, please refer to <https://www.internationalgenome.org/category/population/>

Figure 2. Multidimension scaling plot of the F_{ST} values among each superpopulation.



European samples: CEU, TSI, FIN, GBR, IBS; African samples: YRI, LWK, GWD, MSL, ESN; East Asian samples: CHB, JPT, CHS, CDX, KHV; South Asian samples: GIH, PJI, BEB, STU, ITU; Admixed samples *: MXL, PUR, CLM, PEL, ASW, ACB. Since Brazilians (BRA) are not included in the 1000Genomes project and since we used a different sequencing method for Brazilians, this sample is presented isolated from the other admixed populations.

For a description of each of these population samples from the 1000Genomes dataset, please refer to

<https://www.internationalgenome.org/category/population/>

Figure 3. Tajima's D across the *KIR3DL3* gene in different worldwide populations, considering sliding windows of 250bp and a step size of 1. The x-axis the central position of a window. The y-axis represents the Tajima's D value.

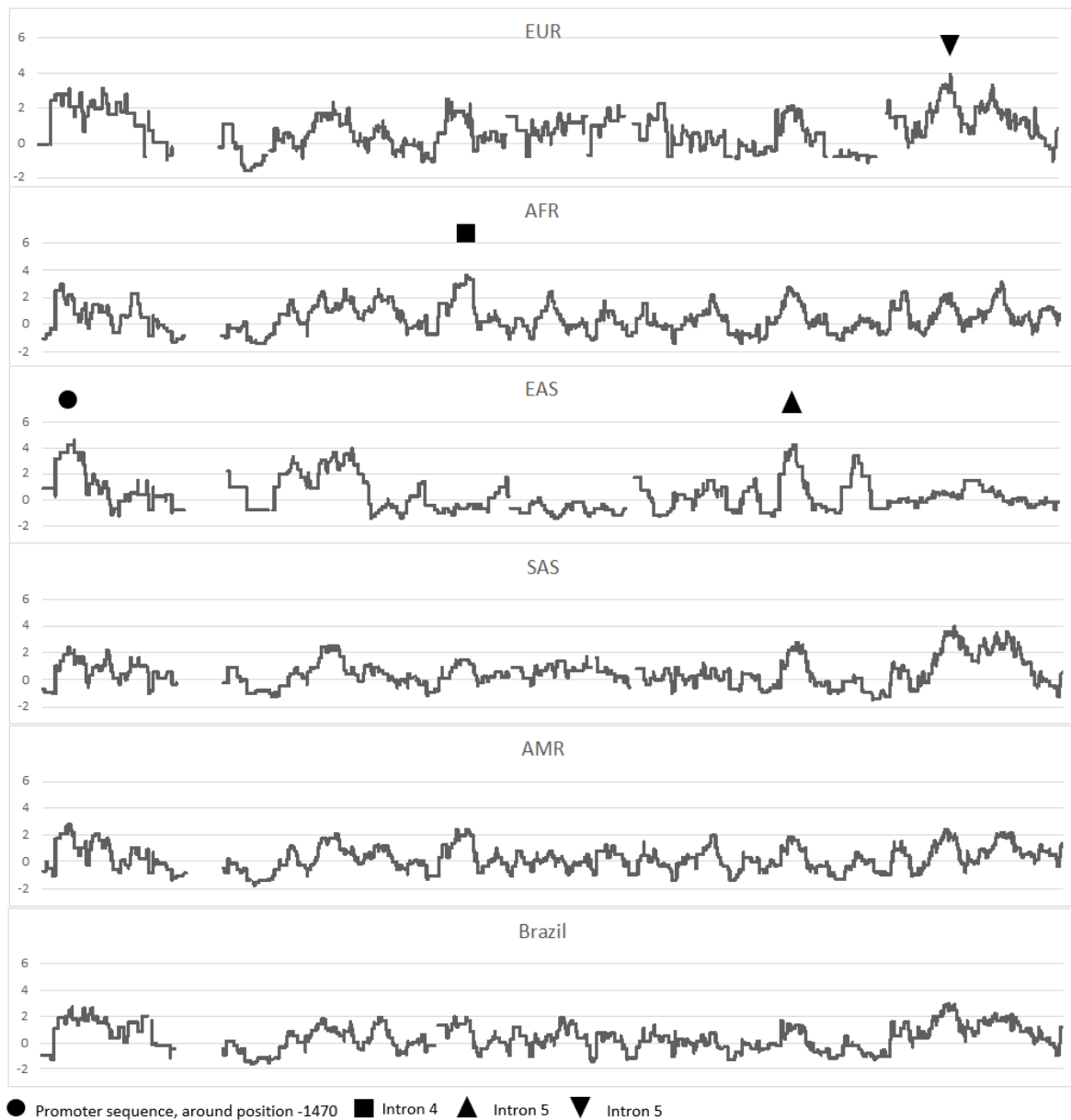


Table 1. *KIR3DL3* encoded proteins in worldwide population samples grouped by the ancestry background (superpopulations).

KIR3DL3 proteins	Global Frequency (2n=5330)	European samples (2n=986)	African samples (2n=968)	East Asian samples (2n=992)	South Asian samples (2n=958)	Admixed samples * (2n=996)	Brazilian samples (2n=430)
KIR3DL3*001	0.0859	0.1836	0.0041	0.0292	0.1451	0.0562	0.1140
KIR3DL3*002	0.0730	0.1166	0.0764	0.0060	0.0814	0.0843	0.0744
KIR3DL3*003	0.1210	0.1572	0.0506	0.0252	0.2422	0.1034	0.1884
KIR3DL3*004	0.0383	0.0172	0.0455	0.0544	0.0251	0.0552	0.0233
KIR3DL3*005	0.0272	0.0010	0.1002	-	-	0.0412	0.0140
KIR3DL3*006	0.0208	0.0243	-	0.0474	0.0177	0.0120	0.0256
KIR3DL3*007	0.0103	0.0203	-	-	0.0209	0.0090	0.0140
KIR3DL3*008	0.0426	0.0213	0.0072	0.1069	0.0167	0.0653	0.0279
KIR3DL3*009	0.1578	0.1166	0.1777	0.1835	0.1044	0.2129	0.1395
KIR3DL3*010	0.0994	0.0183	0.0227	0.3942	0.0344	0.0532	0.0302
KIR3DL3*011	0.0216	0.0071	0.0651	-	0.0198	0.0181	0.0186
KIR3DL3*012	0.0103	-	0.0434	-	-	0.0120	0.0023
KIR3DL3*013	0.0356	0.0842	0.0010	-	0.0428	0.0321	0.0767
KIR3DL3*014	0.0692	0.0740	0.1126	0.0030	0.0783	0.0633	0.1070
KIR3DL3*015	0.0251	0.0213	0.0227	0.0363	0.0261	0.0171	0.0302
KIR3DL3*016	0.0064	0.0030	0.0052	0.0060	0.0104	0.0080	0.0047
KIR3DL3*017	0.0114	0.0345	-	-	0.0125	0.0080	0.0163
KIR3DL3*018	0.0002	-	-	-	-	0.0010	-
KIR3DL3*019	0.0043	0.0132	-	-	0.0010	0.0050	0.0093
KIR3DL3*020	0.0073	0.0061	0.0134	-	0.0125	0.0050	0.0070
KIR3DL3*021	0.0019	0.0020	-	-	0.0021	0.0030	0.0070
KIR3DL3*022	0.0041	-	0.0134	-	-	0.0080	0.0023
KIR3DL3*023	0.0002	-	-	0.0010	-	-	-
KIR3DL3*025	0.0094	0.0162	0.0217	-	-	0.0090	0.0093
KIR3DL3*026	0.0028	0.0010	-	0.0060	0.0073	0.0010	-
KIR3DL3*027	0.0208	0.0132	0.0506	-	0.0136	0.0271	0.0209
KIR3DL3*028	0.0216	0.0081	0.0052	0.0383	0.0522	0.0110	0.0070
KIR3DL3*030	0.0023	0.0041	0.0021	-	0.0010	0.0040	0.0023
KIR3DL3*031	0.0009	0.0030	-	-	-	0.0010	0.0023
KIR3DL3*032	0.0062	-	0.0258	-	-	0.0070	0.0023
KIR3DL3*034	0.0045	-	0.0155	-	-	0.0080	0.0023
KIR3DL3*035	0.0101	-	0.0434	-	-	0.0100	0.0047
KIR3DL3*036	0.0002	0.0010	-	-	-	-	-
KIR3DL3*037	0.0002	-	-	-	0.0010	-	-
KIR3DL3*041	0.0083	0.0122	0.0103	-	0.0125	0.0080	0.0047
KIR3DL3*044	0.0009	0.0030	-	-	-	0.0010	0.0023
KIR3DL3*048	0.0060	0.0020	-	0.0202	0.0052	0.0040	0.0023
KIR3DL3*049	0.0023	0.0010	0.0062	-	-	0.0040	0.0023
KIR3DL3*050	0.0054	-	0.0227	-	-	0.0070	-
KIR3DL3*052	0.0013	-	0.0052	-	-	0.0020	-
KIR3DL3*053	0.0009	-	0.0021	-	-	0.0030	-
KIR3DL3*054	0.0004	-	0.0021	-	-	-	-
KIR3DL3*056	0.0015	-	0.0021	-	-	0.0050	0.0023
KIR3DL3*057	0.0004	0.0010	0.0010	-	-	-	-
KIR3DL3*062	0.0004	-	-	0.0020	-	-	-
KIR3DL3*063	0.0006	-	-	0.0030	-	-	-
KIR3DL3*064	0.0017	-	-	0.0050	0.0042	-	-
KIR3DL3*065	0.0011	-	-	0.0050	0.0010	-	-
KIR3DL3*066	0.0004	0.0010	-	-	0.0010	-	-
KIR3DL3*076	0.0002	0.0010	-	-	-	-	-
KIR3DL3*084	0.0002	0.0010	-	-	-	-	-
New proteins	0.0150	0.0091	0.0227	0.0272	0.0073	0.0141	0.0023

European samples: CEU, TSI, FIN, GBR, IBS; African samples: YRI, LWK, GWD, MSL, ESN; East Asian samples: CHB, JPT, CHS, CDX, KHV; South Asian samples: GIH, PJL, BEB, STU, ITU; Admixed samples *: MXL, PUR, CLM, PEL, ASW, ACB. Since Brazilians are not included in the 1000Genomes project and since we used a different sequencing method for Brazilians, this sample is presented isolated from the other admixed populations.

For a description of each of these population samples from the 1000Genomes dataset, please refer to

<https://www.internationalgenome.org/category/population/>

Proteins were named according to the IPD-IMGT/KIR database, version 2.9.0

Protein frequencies for each population separately is available at table S2.

Table 2. KIR3DL3 encoded proteins and amino acid in each polymorphic residue in worldwide population samples.

KIR3DL3 proteins	Amino acid positions starting from the first translated Methionine																																					Global Frequency							
	LP 6	LP 13	D0 52	D0 56	D0 77	D0 91	D0 100	D0 107	D0 109	D0 117	D1 132	D1 134	D1 139	D1 149	D1 152	D1 166	D1 168	D1 181	D2 219	D2 221	D2 231	D2 242	D2 246	D2 250	D2 259	D2 290	D2 291	D2 293	D2 294	D2 301	D2 306	D2 313	TM 321	TM 324	TM 336	TM 346	TM 348		CY 353	CY 373	CY 376	CY 382	CY 385	CY 388	CY 395
KIR3DL3*001	V	F	R	N	R	G	P	S	P	V	P	V	T	R	R	R	I	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	N	V	A	W	A	A	E	P	A	N	V	T	0.0859
KIR3DL3*002	V	F	R	N	R	G	P	S	P	V	P	V	T	R	R	R	V	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	H	V	A	W	A	A	E	P	A	N	V	T	0.0730
KIR3DL3*003	V	F	R	N	R	G	P	S	P	V	P	V	T	R	R	R	I	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	Y	A	A	W	A	A	E	P	A	N	V	T	0.1210
KIR3DL3*004	V	F	H	N	R	G	P	S	P	V	P	V	T	R	R	R	V	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	Y	A	A	W	A	A	E	P	A	N	V	T	0.0383
KIR3DL3*005	V	F	R	N	R	G	P	S	P	V	P	V	T	R	R	R	V	S	G	Y	D	L	S	F	A	N	Y	C	F	P	D	V	H	V	A	W	A	A	E	P	A	N	V	T	0.0272
KIR3DL3*006	V	F	R	K	R	G	P	S	P	V	P	V	T	S	R	R	V	S	G	Y	G	L	S	F	A	T	Y	C	F	P	D	V	N	V	A	W	A	A	E	P	A	N	V	T	0.0208
KIR3DL3*007	V	F	R	N	R	G	P	S	P	V	P	V	T	R	H	R	V	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	Y	A	A	W	A	A	E	P	A	N	V	T	0.0103
KIR3DL3*008	V	F	H	N	R	G	P	S	P	V	P	V	T	R	R	R	V	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	H	V	A	W	A	A	E	P	A	N	V	T	0.0426
KIR3DL3*009	V	F	R	N	R	G	P	S	P	V	P	V	T	R	R	R	I	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	H	V	A	W	A	A	E	P	A	N	V	T	0.1576
KIR3DL3*010	V	F	R	N	W	G	P	S	P	V	P	V	T	R	R	R	I	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	H	V	A	W	A	A	E	P	A	N	V	T	0.0994
KIR3DL3*011	V	F	R	N	R	G	P	S	P	V	P	V	T	R	R	R	V	S	G	Y	G	L	S	F	A	T	Y	C	F	P	D	V	H	V	A	W	A	A	E	P	A	N	V	T	0.0216
KIR3DL3*012	V	F	R	K	R	G	P	S	P	V	P	V	T	R	R	R	V	S	G	Y	G	L	S	F	A	T	Y	C	F	P	D	V	H	V	A	W	A	A	E	P	A	N	V	T	0.0103
KIR3DL3*013	V	F	R	N	R	G	P	S	P	V	P	V	T	R	R	R	V	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	N	V	A	W	A	A	E	P	A	N	V	T	0.0356
KIR3DL3*014	V	F	R	N	R	G	P	S	P	V	P	V	T	R	R	R	V	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	Y	A	A	W	A	A	E	P	A	N	V	T	0.0692
KIR3DL3*015	V	F	R	K	R	G	P	S	P	V	P	V	T	S	R	R	V	S	G	Y	G	L	S	F	A	T	Y	C	F	P	D	V	H	V	A	W	A	A	E	P	A	N	V	T	0.0251
KIR3DL3*016	V	F	R	K	R	G	P	S	P	V	P	V	T	S	R	R	V	S	G	Y	G	L	S	F	A	T	Y	C	F	P	D	V	Y	A	A	W	A	A	E	P	A	N	V	T	0.0064
KIR3DL3*017	V	F	R	N	R	G	P	S	P	V	P	V	T	R	R	R	V	S	G	Y	G	L	S	F	A	T	Y	C	F	P	D	V	N	V	A	W	A	A	E	P	A	N	V	T	0.0114
KIR3DL3*018	V	F	R	N	R	G	P	S	P	V	P	V	T	S	R	R	V	S	G	Y	G	L	S	F	A	T	Y	C	F	P	D	V	H	V	A	W	A	A	E	P	A	N	V	T	0.0002
KIR3DL3*019	V	F	R	N	R	G	P	S	P	V	P	V	T	R	H	R	V	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	N	V	A	W	A	A	E	P	A	N	V	T	0.0043
KIR3DL3*020	V	F	R	N	R	G	P	S	P	V	P	V	T	R	R	R	V	S	G	Y	G	L	S	F	A	T	Y	C	F	P	D	V	Y	A	A	W	A	A	E	P	A	N	V	T	0.0073
KIR3DL3*021	V	F	R	N	R	G	P	S	P	V	P	V	T	S	R	R	V	S	G	Y	G	L	S	F	A	T	Y	C	F	P	D	V	N	V	A	W	A	A	E	P	A	N	V	T	0.0019
KIR3DL3*022	V	F	R	K	R	G	P	S	P	V	P	V	T	R	R	R	V	S	G	Y	G	L	S	F	A	T	Y	C	F	P	D	V	Y	A	A	W	A	A	E	P	A	N	V	T	0.0041
KIR3DL3*023	V	F	R	K	R	G	P	S	P	V	P	V	T	R	R	R	I	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	H	V	A	W	A	A	E	P	A	N	V	T	0.0002
KIR3DL3*025	V	F	R	K	R	G	P	S	P	V	P	V	T	S	R	R	V	S	G	Y	G	L	S	F	A	T	Y	C	F	P	D	V	H	V	A	W	P	A	D	P	A	N	V	T	0.0094
KIR3DL3*026	V	F	H	N	R	G	P	S	P	V	P	V	T	R	R	R	V	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	N	V	A	W	A	A	E	P	A	N	V	T	0.0028
KIR3DL3*027	V	F	R	N	R	G	P	S	P	V	P	V	T	R	R	R	V	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	H	V	A	W	P	A	D	P	A	N	V	T	0.0208
KIR3DL3*028	V	F	R	N	W	G	P	S	P	V	P	V	T	R	R	R	I	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	Y	A	A	W	A	A	E	P	A	N	V	T	0.0216
KIR3DL3*030	V	F	R	N	R	G	P	S	P	V	P	V	T	R	H	R	V	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	H	V	A	W	A	A	E	P	A	N	V	T	0.0023
KIR3DL3*031	V	F	R	N	R	G	P	S	P	V	P	V	T	R	H	R	V	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	H	V	A	W	P	A	D	P	A	N	V	T	0.0009
KIR3DL3*032	V	F	R	N	R	G	P	S	P	V	L	V	T	R	R	R	I	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	Y	A	A	W	A	A	E	P	A	N	V	T	0.0062
KIR3DL3*034	V	F	R	N	R	G	P	S	P	V	P	V	T	R	R	R	V	S	G	Y	D	L	S	F	A	N	Y	C	F	P	D	V	Y	A	A	W	A	A	E	P	A	N	V	T	0.0045
KIR3DL3*035	V	F	R	N	R	G	P	S	P	V	P	V	T	R	R	R	V	S	G	Y	D	L	S	F	A	N	Y	C	F	P	D	V	H	V	A	W	P	A	D	P	A	N	V	T	0.0101
KIR3DL3*036	V	F	R	N	R	G	P	S	P	V	P	V	T	S	R	R	V	S	G	Y	G	L	S	F	A	T	Y	C	F	P	D	V	Y	A	A	W	A	A	E	P	A	N	V	T	0.0002
KIR3DL3*037	V	F	R	N	R	G	P	S	P	V	P	V	T	R	R	C	I	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	Y	A	A	W	A	A	E	P	A	N	V	T	0.0002
KIR3DL3*041	V	F	R	N	R	G	P	S	P	V	P	V	T	R	R	R	I	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	H	V	A	W	P	A	D	P	A	N	V	T	0.0083
KIR3DL3*044	V	F	R	N	R	G	P	S	P	V	P	V	T	R	R	R	I	S	D	Y	G	L	S	F	A	N	Y	C	F	P	D	V	H	V	A	W	A	A	E	P	A	N	V	T	0.0009
KIR3DL3*048	V	F	R	N	W	G	P	S	P	V	P	V	T	R	R	R	I	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	N	V	A	W	A	A	E	P	A	N	V	T	0.0060
KIR3DL3*049	V	F	H	N	R	G	P	S	P	V	P	V	T	R	R	R	V	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	H	V	A	W	P	A	D	P	A	N	V	T	0.0023
KIR3DL3*050	V	F	R	N	R	G	P	S	P	V	P	V	T	R	R	R	V	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	I	Y	A	A	W	A	A	E	P	A	N	V	T	0.0054
KIR3DL3*052	V	F	R	N	R	G	P	P	P	V	P	V	T	R	R	R	I	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	H	V	A	W	A	A	E	P	A	N	V	T	0.0013
KIR3DL3*053	V	F	R	N	R	G	P	S	P	V	P	V	T	R	R	R	V	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	I	H	V	A	W	A	A	E	P	A	N	V	T	0.0009
KIR3DL3*054	V	F	R	N																																									

KIR3DL3*056	V	F	R	N	R	G	P	S	P	V	P	V	T	R	R	R	V	S	G	Y	G	L	S	F	A	T	Y	C	F	P	D	V	H	V	A	W	P	A	D	P	A	N	V	T	0.0015
KIR3DL3*057	V	F	H	N	R	G	P	S	P	V	P	V	T	R	R	R	V	S	G	Y	G	L	S	F	A	T	Y	C	F	P	D	V	H	V	A	W	A	A	E	P	A	N	V	T	0.0004
KIR3DL3*062	V	F	H	N	R	G	P	S	P	V	P	V	T	R	R	R	V	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	N	V	A	W	A	A	E	P	A	N	V	I	0.0004
KIR3DL3*063	V	F	R	N	W	G	P	S	P	V	P	V	T	R	R	R	I	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	N	V	A	W	A	A	E	P	A	N	V	I	0.0006
KIR3DL3*064	V	F	R	K	R	G	P	S	P	V	P	V	T	S	R	R	V	S	G	Y	G	L	S	F	A	T	Y	C	F	P	D	V	N	V	A	W	A	A	E	P	A	N	V	I	0.0017
KIR3DL3*065	V	F	R	N	R	G	P	S	P	V	P	V	T	R	R	R	I	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	N	V	A	W	A	A	E	P	A	N	V	I	0.0011
KIR3DL3*066	V	F	H	N	R	G	P	S	P	V	P	V	T	R	R	R	I	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	N	V	A	W	A	A	E	P	A	N	V	T	0.0004
KIR3DL3*076	V	F	R	N	R	G	P	S	P	V	P	V	T	R	R	R	I	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	Y	V	A	W	A	A	E	P	A	N	V	T	0.0002
KIR3DL3*084	V	F	R	N	R	G	P	S	P	V	P	V	T	S	R	R	V	S	G	Y	G	L	S	F	A	T	Y	C	F	P	D	V	H	V	A	W	P	A	D	P	A	N	V	T	0.0002

Here we listed only known KIR3DL3 proteins, named according to the IPD-IMGT/HLA database, with a summed frequency of 98.48%. For a list of new proteins please refer to Table S3. Some variants listed in this table do not present an alternative amino acid because these positions are related to new proteins described in Table S3.

Table 3. KIR3DL3 polymorphic residues (minor allele frequency > 1%) and their frequency in worldwide population samples

Population Residue/Amino acid	Region	EUR 2n=986	AFR 2n=968	EAS 2n=992	SAS 2n=958	AMR 2n=996	BRA 2n=428
52/H	D0	0.0426	0.0620	0.1704	0.0532	0.1285	0.0537
52/R		0.9574	0.9380	0.8296	0.9468	0.8715	0.9463
56/K	D0	0.0649	0.1074	0.1028	0.0585	0.0663	0.0748
56/N		0.9351	0.8926	0.8972	0.9415	0.9337	0.9252
77/R	D0	0.9716	0.9711	0.5292	0.9050	0.9297	0.9603
77/W		0.0284	0.0289	0.4708	0.0950	0.0703	0.0397
107/P	D0	-	0.0124	-	-	0.0060	-
107/S		1.0000	0.9876	1.0000	1.0000	0.9940	1.0000
132/L	D1	0.0010	0.0289	-	-	0.0070	0.0023
132/P		0.9990	0.9711	1.0000	1.0000	0.9930	0.9977
149/R	D1	0.9310	0.9504	0.8962	0.9395	0.9498	0.9229
149/S		0.0690	0.0496	0.1038	0.0605	0.0502	0.0771
152/H	D1	0.0416	0.0031	-	0.0230	0.0191	0.0280
152/R		0.9584	0.9969	1.0000	0.9770	0.9809	0.9720
168/I	D1	0.5081	0.3151	0.7046	0.6054	0.4659	0.4930
168/V		0.4919	0.6849	0.2954	0.3946	0.5341	0.5070
231/D	D2	0.0010	0.1591	-	-	0.0612	0.0210
231/G		0.9990	0.8409	1.0000	1.0000	0.9388	0.9790
290/N	D2	0.8813	0.8048	0.8962	0.8946	0.8926	0.8738
290/T		0.1187	0.1952	0.1038	0.1054	0.1074	0.1262
313/I	D2	-	0.0248	-	-	0.0100	-
313/V		1.0000	0.9752	1.0000	1.0000	0.9900	1.0000
321/H	TM	0.3580	0.6684	0.7470	0.3111	0.5863	0.3902
321/N		0.3479	0.0052	0.1250	0.2401	0.1225	0.2523
321/Y		0.2941	0.3264	0.1280	0.4489	0.2912	0.3575
324/A	TM	0.2931	0.3264	0.1280	0.4489	0.2912	0.3575
324/V		0.7069	0.6736	0.8720	0.5511	0.7088	0.6425
348/A	TM	0.9533	0.8626	1.0000	0.9729	0.9337	0.9533
348/P		0.0467	0.1374	-	0.0271	0.0663	0.0467
373/D	CY	0.0467	0.1374	-	0.0271	0.0663	0.0467
373/E		0.9533	0.8626	1.0000	0.9729	0.9337	0.9533
395/I	CY	-	-	0.0151	0.0052	-	-
395/T		1.0000	1.0000	0.9849	0.9948	1.0000	1.0000

* Residues positions considers the first translated Methionine as residue 1.

European samples: CEU, TSI, FIN, GBR, IBS; African samples: YRI, LWK, GWD, MSL, ESN; East Asian samples: CHB, JPT, CHS, CDX, KHV; South Asian samples: GIH, PJL, BEB, STU, ITU; Admixed samples *: MXL, PUR, CLM, PEL, ASW, ACB. Since Brazilians are not included in the 1000Genomes project and since we used a different sequencing method for Brazilians, this sample is presented isolated from the other admixed populations.

For a description of each of these population samples from the 1000Genomes dataset, please refer to

<https://www.internationalgenome.org/category/population/>

Table 4: Nucleotide diversity in each KIR3DL3 segment in worldwide populations grouped according to their ancestry background.

Nucleotide diversity	AFR	AMR	EAS	EUR	BRA	SAS
Promoter	0.0032	0.0033	0.0034	0.0033	0.0033	0.0031
Exon 1	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000
Intron 1	ne	ne	ne	ne	ne	ne
Exon 2	0.0022	0.0016	0.0000	0.0029	0.0022	0.0025
Intron 2	0.0020	0.0015	0.0020	0.0012	0.0014	0.0013
Exon 3	0.0029	0.0033	0.0048	0.0025	0.0026	0.0031
Intron 3	0.0061	0.0054	0.0048	0.0053	0.0053	0.0049
Exon 4	0.0034	0.0034	0.0026	0.0039	0.0039	0.0035
Intron 4	0.0036	0.0032	0.0013	0.0034	0.0034	0.0030
Exon 5	0.0060	0.0046	0.0009	0.0047	0.0048	0.0035
Intron 5	0.0040	0.0035	0.0023	0.0032	0.0034	0.0034
Exon 6	0.0119	0.0147	0.0114	0.0128	0.0140	0.0134
Intron 6	0.0110	0.0104	0.0052	0.0113	0.0115	0.0116
Exon 7	0.0001	0.0012	0.0036	0.0031	0.0021	0.0020
Intron 7	0.0192	0.0156	0.0069	0.0143	0.0157	0.0158
Exon 8	0.0019	0.0010	0.0002	0.0007	0.0007	0.0006
3'UTR	0.0074	0.0052	0.0017	0.0046	0.0048	0.0044

European samples: CEU, TSI, FIN, GBR, IBS; African samples: YRI, LWK, GWD, MSL, ESN; East Asian samples: CHB, JPT, CHS, CDX, KHV; South Asian samples: GIH, PJL, BEB, STU, ITU; Admixed samples *: MXL, PUR, CLM, PEL, ASW, ACB. Since Brazilians (BRA) are not included in the 1000Genomes project and since we used a different sequencing method for Brazilians, this sample is presented isolated from the other admixed populations.

For a description of each of these population samples from the 1000Genomes dataset, please refer to <https://www.internationalgenome.org/category/population/>

Table 5: Tajima's D in each KIR3DL3 segment in worldwide populations grouped according to their ancestry background.

<i>Tajima's D</i>	AFR	AMR	EAS	EUR	BRA	SAS
Promoter	0.8805	0.9214	2.4064	2.7273	1.5887	1.2728
Exon 1	-	-0.8046	-	-	-	-
Intron 1	ne	ne	ne	ne	ne	ne
Exon 2	-0.7940	-0.4711	-	-0.2402	-0.4214	-0.2732
Intron 2	-0.2779	-0.9978	1.5012	-0.4528	-0.9390	-0.5245
Exon 3	0.0549	-0.2374	1.6750	-0.4546	-0.0264	0.1568
Intron 3	2.1687	1.0083	2.1354	1.0044	0.8575	1.3373
Exon 4	-0.1035	-0.0632	0.2988	-0.0660	-0.0346	-0.0564
Intron 4	1.1835	0.7680	-0.5738	1.4085	0.9126	1.0118
Exon 5	0.9079	-0.3155	-1.4076	0.2965	0.1050	0.5017
Intron 5	0.6055	0.3726	0.7655	1.4880	0.6216	1.0000
Exon 6	1.1032	0.8892	1.5024	1.3168	1.3343	1.4444
Intron 6	1.3452	1.5977	0.3115	2.1028	2.0053	3.1717
Exon 7	-1.0910	-0.4250	0.3405	0.1866	-0.2451	-0.1703
Intron 7	1.4133	1.1325	0.3843	2.3928	1.2920	2.7772
Exon 8	-0.5155	-0.5892	-0.6351	-0.7355	-0.3704	-1.0948
3'UTR	0.7521	0.8682	-0.0870	0.3071	0.3514	0.0098

European samples: CEU, TSI, FIN, GBR, IBS; African samples: YRI, LWK, GWD, MSL, ESN; East Asian samples: CHB, JPT, CHS, CDX, KHV; South Asian samples: GIH, PJL, BEB, STU, ITU; Admixed samples *: MXL, PUR, CLM, PEL, ASW, ACB. Since Brazilians (BRA) are not included in the 1000Genomes project and since we used a different sequencing method for Brazilians, this sample is presented isolated from the other admixed populations.

For a description of each of these population samples from the 1000Genomes dataset, please refer to

<https://www.internationalgenome.org/category/population/>

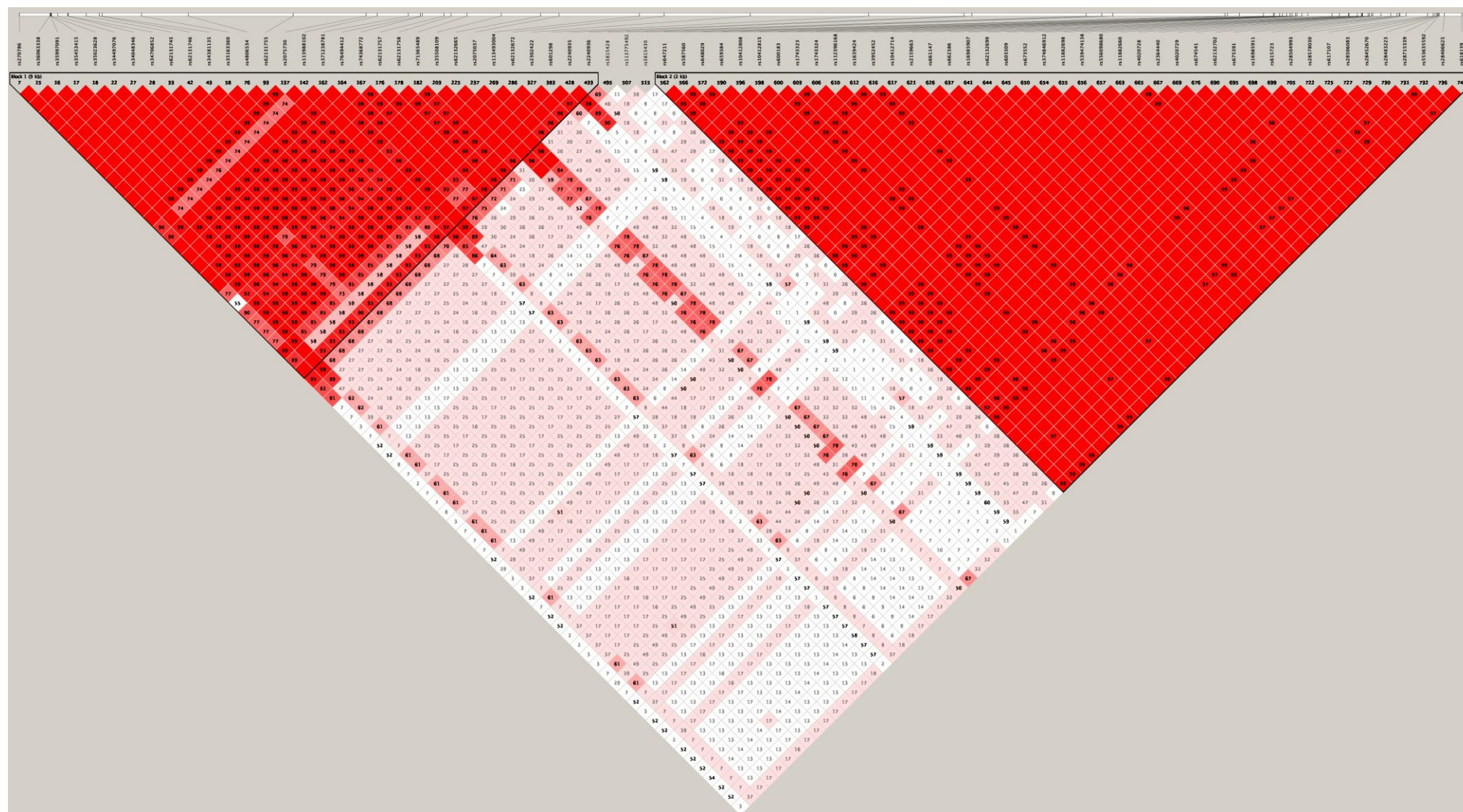
Table 6: dN/dS ratio test of the *KIR3DL3* exons in worldwide population samples.

Neutrality deviation (dn # ds)						
Exon	AFR	AMR	EAS	EUR	Brazilians	SAS
E1	0.9908, P=0.3238	0.9908, P=0.3238	-	-	-	-
E2	-1.6316, P=0.1054	-1.0983, P=0.2743	-	-1.1255, P=0.2626	-1.1223, P=0.2640	-1.1223, P=0.2640
E3	-0.6771, P=0.4996	-0.5998, P=0.5498	0.2225, P=0.8243	-1.1781, P=0.2411	-0.8585, P=0.3923	-0.7377, P=0.4621
E4	-0.0564, P=0.9551	-0.9661, P=0.3359	-0.4905, P=0.6247	-0.2966, P=0.7673	-0.9573, P=0.3403	-0.9777, P=0.3302
E5	-1.6524, P=0.1011	-1.4409, P=0.1522	-0.4294, P=0.6684	-1.4140, P=0.1600	-1.7124, P=0.0894	-2.0516, P=0.0424
E6	0.2473, P=0.8051	0.9206, P=0.3591	-0.0425, P=0.9662	0.3386, P=0.7355	0.2410, P=0.8100	0.2410, P=0.8100
E7	-0.6154, P=0.5394	-0.8279, P=0.4094	-0.8279, P=0.4094	-0.8279, P=0.4094	-0.8279, P=0.4094	-0.8998, P=0.3700
E8	1.7787, P=0.0778	-0.6477, P=0.5184	-1.0837, P=0.2807	-0.6230, P=0.5345	0.9417, P=0.3482	-0.2817, P=0.7786
Positive selection (dn > ds)						
Exon	AFR	AMR	EAS	EUR	Brazilians	SAS
E1	0.9908, P=0.1619	0.9908, P=0.1619	-	-	-	-
E2	-1.6316, P=1.0000	-1.0983, P=1.0000	-	-1.1255, P=1.0000	-1.1223, P=1.0000	-1.1223, P=1.0000
E3	-0.6771, P=1.0000	-0.5998, P=1.0000	0.2225, P=0.4122	-1.1781, P=1.0000	-0.8585, P=1.0000	-0.7377, P=1.0000
E4	-0.0564, P=1.0000	-0.9661, P=1.0000	-0.4905, P=1.0000	-0.2966, P=1.0000	-0.9850, P=1.0000	-0.9777, P=1.0000
E5	-1.6524, P=1.0000	-1.4409, P=1.0000	-0.4294, P=1.0000	-1.4140, P=1.0000	-1.7124, P=1.0000	-2.0516, P=1.0000
E6	0.2473, P=0.4025	0.9206, P=0.1796	-0.0425, P=1.0000	0.3386, P=0.3677	0.2410, P=0.4050	0.2410, P=0.4050
E7	-0.6154, P=1.0000	-0.8279, P=1.0000	-0.8279, P=1.0000	-0.8279, P=1.0000	-0.8279, P=1.0000	-0.8998, P=1.0000
E8	1.7787, P=0.0389	-0.6477, P=1.0000	-1.0837, P=1.0000	-0.6230, P=1.0000	0.9417, P=0.1741	-0.2817, P=1.0000
Purifying selection (ds > dn)						
Exon	AFR	AMR	EAS	EUR	Brazilians	SAS
E1	-0.9908, P=1.0000	-0.9908, P=1.0000	0.0000, P=1.0000	0.0000, P=1.0000	0.0000, P=1.0000	0.0000, P=1.0000
E2	1.6316, P=0.0527	1.0983, P=0.1371	0.0000, P=1.0000	1.1255, P=0.1313	1.1223, P=0.1320	1.1223, P=0.1320
E3	0.6771, P=0.2498	0.5998, P=0.2749	-0.2225, P=1.0000	1.1781, P=0.1205	0.8585, P=0.1962	0.7377, P=0.2311
E4	0.0564, P=0.4776	0.9661, P=0.1680	0.4905, P=0.3123	0.2966, P=0.3837	0.9850, P=0.1633	0.9777, P=0.1651
E5	1.6524, P=0.0505	1.4409, P=0.0761	0.4294, P=0.3342	1.4140, P=0.0800	1.7124, P=0.0447	2.0516, P=0.0212
E6	-0.2473, P=1.0000	-0.9206, P=1.0000	0.0425, P=0.4831	-0.3386, P=1.0000	-0.2410, P=1.0000	-0.2410, P=1.0000
E7	0.6154, P=0.2697	0.8279, P=0.2047	0.8279, P=0.2047	0.8279, P=0.2047	0.8998, P=0.1850	0.8998, P=0.1850
E8	-1.7787, P=1.0000	0.6477, P=0.2592	1.0837, P=0.1403	0.6230, P=0.2672	-0.9417, P=1.0000	0.2817, P=0.3893

dN/dS ratio test was performed using all sequences found for each segment of the *KIR3DL3* locus, using bootstrap method (5,000 replicates). Significant P-values are marked in boldface.

Supplementary material

Figure S1. Linkage Disequilibrium (LD) between pair of single nucleotide polymorphisms (SNPs) of the KIR3DL3 locus, considering genomic positions from Promoter to 3'UTR. LD plot generated by Haploview 4.2, considering variable sites with a minor allele a minor allele frequency (MAF) $\geq 5\%$ and the fraction of strong LD in informative comparison set to 0.9. Areas in red color indicate strong LD (LOD score ≥ 2 , D' score =1); areas in light red or pink indicate moderate LD (LOD ≥ 2 , $D' < 1$); areas in blue indicate weak LD (LOD < 2 , $D' = 1$); and areas in white indicate very weak LD (LOD < 2 , $D' < 1$).



0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99

[illegible]

Figure S3. *KIR3DL3* 3'UTR sequences and their frequency in worldwide populations.

Genomic position	SNPId	Reference allele	Alternative allele	Global	EUR	AFR	EAS	SAS	AMR	BRA
54736119	rs548786357	C	T	0.6306	0.6471	0.5258	0.8327	0.5230	0.6365	0.5972
54736120	rs565504003	G	A	0.2925	0.2931	0.3027	0.1280	0.4384	0.2801	0.3518
54736128	19:54736128	A	T	0.0546	0.0467	0.1374	-	0.0271	0.0663	0.0463
54736145	rs186885062	A	T	0.0049	-	0.0165	-	-	0.0090	-
54736147	rs676354	T	C	0.0039	-	-	0.0212	-	-	-
54736190	rs28578030	T	C	0.0019	0.0051	-	-	-	0.0040	0.0046
54736219	rs537445098	T	C	0.0017	-	0.0072	-	-	0.0010	-
54736224	19:54736224	T	C	0.0017	-	-	0.0081	0.0010	-	-
54736240	rs617107	C	A	0.0011	-	-	0.0060	-	-	-
54736254	rs574084778	G	A	0.0011	-	0.0031	-	-	0.0010	-
54736258	rs28506083	G	A	0.0011	-	-	-	0.0063	-	-
54736277	rs28452670	T	C	0.0009	0.0030	-	-	-	0.0010	-
54736285	rs28483223	A	G	0.0008	-	-	0.0040	-	-	-
54736290	rs28715339	G	A	0.0006	0.0030	-	-	-	-	-
54736296	19:54736296	A	G	0.0006	0.0010	0.0010	-	0.0010	-	-
54736300	rs553635592	A	T	0.0004	-	0.0020	-	-	-	-
54736326	19:54736326	T	C	0.0002	-	0.0010	-	-	-	-
54736337	rs545760514	C	G	0.0002	-	-	-	0.0010	-	-
54736349	rs28406625	G	A	0.0002	-	0.0010	-	-	-	-
54736360	rs369357381	C	T	0.0002	-	0.0010	-	-	-	-
54736381	rs909754396	C	T	0.0002	-	0.0010	-	-	-	-
54736437	rs575714276	CT	C	0.0002	-	0.0010	-	-	-	-
54736448	19:54736448	G	A	0.0002	-	-	-	0.0010	-	-
54736461	19:54736461	A	T	0.0002	-	-	-	0.0010	-	-
54736481	19:54736481	T	C	0.0002	-	0.0010	-	-	-	-
54736485	rs187183548	A	T	0.0002	-	-	-	0.0010	-	-
54736501	19:54736501	C	G	0.0002	-	-	-	0.0010	-	-
54736515	rs618139	G	A	0.0002	-	-	-	0.0010	-	-
54736520	rs372934281	A	G	0.0002	-	-	-	0.0010	-	-
54736536	rs189545361	G	A	0.0002	-	-	-	0.0010	-	-

Table S1. List of variants and frequency of the reference allele in worldwide populations samples.

Segment	Position (hg38)	SNPId	Reference Allele	Alternative Allele	Freq. Reference Allele	Freq. Reference Allele - AFR	Freq. Reference Allele - AMR	Freq. Reference Allele - EUR	Freq. Reference Allele - EAS	Freq. Reference Allele - SAS
Promoter	54722591	rs368034636	G	T	0.999812	1	1	0.998986	1	1
Promoter	54722631	rs796548904	G	A	0.999625	0.998967	0.998996	1	1	1
Promoter	54722634	rs890066138	C	T	0.998687	0.994835	0.997992	1	1	1
Promoter	54722649	.	G	A	0.999812	1	0.998996	1	1	1
Promoter	54722677	.	C	T	0.999812	1	0.998996	1	1	1
Promoter	54722687	.	A	AT	0.999812	0.998967	1	1	1	1
Promoter	54722701	rs1035539449	T	A	0.999625	0.997934	1	1	1	1
Promoter	54722714	rs270786	A	G	0.902439	0.933884	0.865462	0.938134	0.828629	0.942589
Promoter	54722745	.	C	T	0.999812	1	1	1	1	0.998956
Promoter	54722766	rs374781826	A	G	0.990994	0.964876	0.98996	1	1	1
Promoter	54722835	rs948642026	A	T	0.999625	1	1	1	1	0.998956
Promoter	54722859	rs909565298	G	A	0.999062	1	1	1	1	0.994781
Promoter	54722905	.	A	G	0.999625	1	0.997992	1	1	1
Promoter	54722977	rs1011177880	A	G	0.99925	1	1	1	1	1
Promoter	54723007	rs36063338	A	T	0.922139	0.899793	0.934739	0.935091	0.898185	0.940501
Promoter	54723012	rs35997091	T	C	0.330582	0.307851	0.302209	0.306288	0.479839	0.267223
Promoter	54723018	rs35453415	A	G	0.330582	0.307851	0.302209	0.306288	0.479839	0.267223
Promoter	54723022	rs35023628	A	C	0.330582	0.307851	0.302209	0.306288	0.479839	0.267223
Promoter	54723037	.	G	A	0.999812	0.998967	1	1	1	1
Promoter	54723039	.	C	G	0.999812	1	1	1	1	0.998956
Promoter	54723064	.	G	A	0.999812	0.998967	1	1	1	1
Promoter	54723083	rs34497076	G	A	0.330582	0.307851	0.302209	0.306288	0.479839	0.267223
Promoter	54723126	.	C	T	0.999812	1	1	1	1	0.998956
Promoter	54723127	rs972842909	G	A	0.999625	0.997934	1	1	1	1
Promoter	54723130	rs562865124	G	A	0.999812	0.998967	1	1	1	1
Promoter	54723164	rs933735370	G	A	0.999062	0.998967	0.997992	0.998986	1	1
Promoter	54723176	rs34048346	G	A	0.339587	0.342975	0.312249	0.306288	0.479839	0.267223
Promoter	54723179	rs34706852	T	C	0.330582	0.307851	0.302209	0.306288	0.479839	0.267223
Promoter	54723190	.	G	T	0.999812	0.998967	1	1	1	1
Promoter	54723204	rs946892485	A	G	0.999812	1	1	1	1	0.998956
Promoter	54723220	.	A	G	0.999812	1	0.998996	1	1	1
Promoter	54723247	rs999493554	A	G	0.999062	0.996901	0.997992	1	1	1
Promoter	54723355	rs62131745	C	T	0.330582	0.307851	0.302209	0.306288	0.479839	0.267223
Promoter	54723364	.	A	G	0.999812	0.998967	1	1	1	1
Promoter	54723387	rs374046553	T	C	0.999437	1	1	1	0.997984	0.998956
Promoter	54723413	rs1043927035	C	T	0.999625	1	1	1	0.997984	1
Promoter	54723428	rs935601288	G	A	0.997186	1	1	1	1	0.984342
Promoter	54723430	.	G	C	0.999812	1	0.998996	1	1	1
Promoter	54723449	rs770815275	G	A	0.998874	1	1	0.998986	1	0.994781
Promoter	54723454	rs896464798	C	T	0.999625	0.997934	1	1	1	1
Promoter	54723480	rs62131746	G	A	0.330582	0.307851	0.302209	0.306288	0.479839	0.267223
Promoter	54723509	rs34381135	T	C	0.330582	0.307851	0.302209	0.306288	0.479839	0.267223
Promoter	54723526	rs1016724977	C	A	0.999812	1	1	1	1	0.998956
Promoter	54723532	rs897053010	A	G	0.999812	1	1	1	1	1
Promoter	54723548	rs35409159	A	G	0.15272	0.100207	0.165663	0.278905	0.00907258	0.174322

Promoter	54723575	.	T	A	0.998124	1	1	1	0.989919	1
Promoter	54723649	rs1018004157	T	C	0.999437	0.998967	0.997992	1	1	1
Promoter	54723658	.	G	A	0.999812	1	1	0.998986	1	1
Promoter	54723662	.	A	AGG	0.998311	1	1	1	0.990927	1
Promoter	54723706	rs35596870	C	T	0.15272	0.100207	0.165663	0.278905	0.00907258	0.174322
Promoter	54723736	rs948040823	C	T	0.999625	1	1	1	1	0.997912
Promoter	54723751	rs887822823	G	A	0.981426	0.921488	0.983936	1	1	1
Promoter	54723779	.	G	C	0.999812	1	1	0.998986	1	1
Promoter	54723785	rs374023031	G	A	0.999625	0.998967	0.998996	1	1	1
Promoter	54723853	rs891053129	T	C	0.999812	1	1	0.998986	1	1
Promoter	54723865	rs1017972913	A	C	0.99925	1	0.997992	1	1	0.997912
Promoter	54723879	rs1030779287	C	T	0.999625	1	1	1	0.997984	1
Promoter	54723880	rs957090533	G	T,A	0.999625	1	1	1	0.997984	1
Promoter	54723883	rs35163360	A	G	0.330582	0.307851	0.302209	0.306288	0.479839	0.267223
Promoter	54723954	rs891013524	G	T	0.999812	1	1	1	1	1
Promoter	54723955	rs1006826642	G	C	0.999812	1	0.998996	1	1	1
Promoter	54724000	rs569602538	C	T	0.999625	1	0.998996	1	0.998992	1
Promoter	54724013	.	C	T	0.999812	1	0.998996	1	1	1
Promoter	54724014	rs270787	G	A	0.615947	0.738636	0.608434	0.513185	0.753024	0.492693
Promoter	54724018	.	CA	C,TA	0.961351	1	1	1	0.997984	1
Promoter	54724030	rs549045019	A	G	0.999812	1	1	1	0.998992	1
Promoter	54724031	rs565767195	A	G	0.998124	1	1	0.997972	1	0.991649
Promoter	54724080	.	C	T	0.999812	0.998967	1	1	1	1
Promoter	54724103	.	G	A	0.999812	1	0.998996	1	1	1
Promoter	54724115	.	C	T	0.999812	1	1	1	1	0.998956
Promoter	54724151	.	A	G	0.999812	1	1	1	1	1
Promoter	54724155	rs972049390	A	G	0.999625	0.997934	1	1	1	1
Promoter	54724185	.	A	C	0.999812	1	1	1	1	0.998956
Promoter	54724194	.	T	C	0.999812	1	1	1	0.998992	1
Promoter	54724258	rs1039638753	G	A	0.999437	1	1	1	0.998992	0.997912
Promoter	54724287	.	C	G	0.999812	1	1	1	1	0.998956
Promoter	54724332	rs4806534	C	T	0.330582	0.307851	0.302209	0.306288	0.479839	0.267223
Promoter	54724333	.	G	A	0.999625	1	1	1	0.997984	1
Promoter	54724334	rs767041959	C	T	0.998499	1	0.996988	0.996957	1	1
Promoter	54724339	rs1016439702	G	C	0.99212	0.969008	0.988956	1	1	1
Promoter	54724386	rs879192678	C	T	0.989681	1	1	0.998986	1	0.943633
Promoter	54724387	rs111269305	G	C	0.97167	0.97314	0.978916	0.953347	1	0.955115
Promoter	54724401	.	A	T	0.999812	0.998967	1	1	1	1
Promoter	54724440	.	T	A	0.999812	0.998967	1	1	1	1
Promoter	54724488	rs372292788	G	A	0.999625	0.998967	0.998996	1	1	1
Promoter	54724496	rs936588474	C	A	0.999437	1	1	1	0.996976	1
Exon 1	54724501	rs892116481	C	T	0.999812	1	1	1	1	0.998956
Exon 1	54724511	rs974148272	C	T	0.999812	0.998967	1	1	1	1
Exon 1	54724512	rs370062918	G	A	0.999625	0.998967	0.998996	1	1	1
Exon 1	54724523	rs375595220	G	A	0.999812	0.998967	1	1	1	1
Exon 2	54725248	rs570478218	GTTC	TTTC,G	0.999437	0.998967	1	0.997972	1	1
Exon 2	54725281	rs35488539	G	A	0.966417	0.960744	0.970884	0.950304	1	0.953027
Intron 2	54725291	rs565431320	T	G	0.997936	0.991736	0.997992	1	1	1
Intron 2	54725301	rs531250098	C	A	0.999812	1	1	0.998986	1	1

Intron 2	54725334	rs62131755	C	T	0.330582	0.307851	0.302209	0.306288	0.479839	0.267223
Intron 2	54725343	rs926581824	C	T	0.999812	1	0.998996	1	1	1
Intron 2	54725347	rs563169809	A	G	0.999812	1	0.998996	1	1	1
Intron 2	54725353	.	G	A	0.999812	1	1	1	0.998992	1
Intron 2	54725355	rs376760752	G	A	0.998499	0.991736	1	1	1	1
Intron 2	54725383	rs548853364	C	T	0.999812	0.998967	1	1	1	1
Intron 2	54725391	.	G	A	0.999812	1	1	1	1	0.998956
Intron 2	54725441	rs944913763	C	T	0.999812	1	1	1	1	1
Intron 2	54725443	rs1039299562	C	A	0.999812	1	1	1	1	1
Intron 2	54725447	.	A	T	0.999812	1	0.998996	1	1	1
Intron 2	54725451	rs900787469	C	G	0.999812	1	1	0.998986	1	1
Intron 2	54725460	rs369858884	C	A	0.996623	0.987603	0.993976	1	1	1
Intron 2	54725468	rs775576257	G	A	0.998311	1	0.995984	0.997972	1	0.998956
Intron 2	54725505	rs113609287	G	C	0.998499	1	0.998996	0.996957	1	0.998956
Intron 2	54725559	rs959283905	G	A	0.999812	0.998967	1	1	1	1
Intron 2	54725597	rs374283119	G	C	0.999437	0.997934	1	0.998986	1	1
Intron 2	54725606	.	G	A	0.999812	1	1	1	0.998992	1
Intron 2	54725634	rs768877837	G	A	0.999625	1	0.998996	0.998986	1	1
Intron 2	54725644	.	C	T	0.999437	1	1	1	0.996976	1
Intron 2	54725648	.	A	G	0.999437	0.996901	1	1	1	1
Intron 2	54725658	rs958867133	A	G	0.997936	0.991736	0.997992	1	1	1
Intron 2	54725667	rs776975112	C	T	0.997749	1	0.995984	0.992901	1	1
Intron 2	54725688	.	G	A	0.999812	0.998967	1	1	1	1
Intron 2	54725695	.	C	A	0.999812	0.998967	1	1	1	1
Intron 2	54725698	.	C	T	0.999625	1	0.997992	1	1	1
Intron 2	54725720	.	G	T	0.999437	0.996901	1	1	1	1
Intron 2	54725743	.	G	A	0.999812	1	1	1	0.998992	1
Intron 2	54725754	.	T	A	0.999812	1	1	0.998986	1	1
Intron 2	54725759	.	G	A	0.999625	1	1	1	1	0.997912
Intron 2	54725841	rs759576531	GA	G	0.994934	0.987603	0.997992	1	0.987903	0.998956
Intron 2	54725869	rs112755405	T	C	0.978049	0.909091	0.977912	1	1	1
Intron 2	54725897	rs947043468	C	T	0.999625	1	1	1	1	0.997912
Intron 2	54725898	.	G	T,C	0.999625	0.998967	0.998996	1	1	1
Intron 2	54725908	.	T	C	0.999812	1	1	1	0.998992	1
Intron 2	54725911	.	G	A	0.999625	1	1	1	1	0.997912
Intron 2	54725916	.	G	A	0.999812	1	1	0.998986	1	1
Intron 2	54725925	rs111721146	C	T	0.989493	0.957645	0.988956	1	1	1
Intron 2	54725931	.	C	T	0.999812	0.998967	1	1	1	1
Intron 2	54725940	.	C	T	0.999625	1	1	1	0.997984	1
Intron 2	54725967	rs2075729	G	A	0.823827	0.797521	0.86747	0.971602	0.529234	0.907098
Intron 2	54725976	.	G	A	0.999437	1	1	1	1	0.996868
Intron 2	54725993	.	C	T	0.999812	1	1	0.998986	1	1
Intron 2	54726008	.	G	A	0.998499	1	1	1	1	0.991649
Intron 2	54726018	.	C	T	0.999812	1	1	0.998986	1	1
Intron 2	54726033	rs2075730	G	A	0.324953	0.309917	0.292169	0.289047	0.479839	0.265136
Intron 2	54726036	.	G	T	0.999812	1	0.998996	1	1	1
Intron 2	54726050	.	C	G	0.999812	1	1	1	1	0.998956
Exon 3	54726066	rs765652002	G	A	0.999812	1	1	0.998986	1	1
Exon 3	54726131	.	G	A	0.999812	1	1	1	1	0.998956

Exon 3	54726137	rs113988102	G	A	0.911257	0.938017	0.871486	0.957404	0.829637	0.946764
Exon 3	54726144	.	G	T	0.999812	1	1	0.998986	1	1
Exon 3	54726150	rs2075731	T	C,A	0.328893	0.311983	0.293173	0.289047	0.479839	0.283925
Exon 3	54726175	.	G	A	0.999812	1	1	1	1	0.998956
Exon 3	54726179	rs750098175	T	A	0.999812	0.998967	1	1	1	1
Exon 3	54726211	rs202170060	C	T	0.86848	0.971074	0.929719	0.971602	0.529234	0.90501
Exon 3	54726212	rs370423217	G	A	0.999812	1	1	1	0.998992	1
Exon 3	54726231	rs113216501	C	T	0.966792	0.958678	0.968876	0.967546	1	0.935282
Exon 3	54726254	.	G	C	0.999625	1	0.997992	1	1	1
Exon 3	54726280	.	C	A	0.999812	1	0.998996	1	1	1
Exon 3	54726284	.	A	G	0.999812	1	1	1	1	0.998956
Exon 3	54726288	rs753457635	C	T	0.999812	1	1	0.998986	1	1
Exon 3	54726293	rs536966126	C	T	0.999812	0.998967	1	1	1	1
Exon 3	54726301	rs374399247	T	C	0.996623	0.987603	0.993976	1	1	1
Exon 3	54726308	rs756969033	C	T	0.999812	1	1	0.998986	1	1
Exon 3	54726312	.	C	G	0.999812	1	1	1	1	0.998956
Exon 3	54726331	rs557207322	G	T	0.999625	1	1	1	0.998992	0.998956
Intron 3	54726357	rs567583769	G	C	0.999812	1	1	0.998986	1	1
Intron 3	54726404	rs270789	C	T,A	0.466792	0.432851	0.460843	0.416836	0.664315	0.370564
Intron 3	54726409	.	A	G	0.99925	1	1	0.997972	1	1
Intron 3	54726422	rs536639660	C	A	0.999812	1	0.998996	1	1	1
Intron 3	54726433	rs34502004	C	A,T	0.317824	0.268595	0.279116	0.291075	0.480847	0.279749
Intron 3	54726435	rs371238781	A	G	0.921201	0.892562	0.932731	0.934077	0.898185	0.946764
Intron 3	54726436	rs545486946	C	T	0.999812	1	1	1	0.998992	1
Intron 3	54726550	rs76494412	T	C	0.309944	0.256198	0.272088	0.290061	0.474798	0.264092
Intron 3	54726551	rs201448412	A	T	0.870356	0.971074	0.932731	0.972617	0.534274	0.906054
Intron 3	54726560	rs776921836	G	A	0.999812	1	1	0.998986	1	1
Intron 3	54726570	rs74368772	A	G	0.312008	0.261364	0.276104	0.290061	0.476815	0.264092
Intron 3	54726587	.	G	A	0.999812	1	1	1	1	0.998956
Intron 3	54726597	rs62131756	A	G	0.604878	0.740702	0.599398	0.479716	0.752016	0.487474
Intron 3	54726626	.	G	C	0.999812	0.998967	1	1	1	1
Intron 3	54726662	.	C	A	0.999625	1	1	1	0.997984	1
Intron 3	54726663	.	T	C	0.997936	0.991736	0.997992	1	1	1
Intron 3	54726674	.	G	A	0.999625	1	1	1	1	1
Intron 3	54726735	rs79514652	G	A	0.824203	0.797521	0.86747	0.971602	0.532258	0.906054
Intron 3	54726745	.	GA	G	0.999812	1	1	1	1	0.998956
Intron 3	54726773	.	C	T	0.999812	1	1	1	1	1
Intron 3	54726776	rs62131757	G	T	0.322702	0.31095	0.293173	0.290061	0.465726	0.264092
Intron 3	54726828	.	CA	C	0.999812	1	1	1	1	0.998956
Intron 3	54726832	rs887067429	G	A	0.999812	1	0.998996	1	1	1
Intron 3	54726847	rs62131758	C	T	0.322889	0.31095	0.293173	0.290061	0.466734	0.264092
Intron 3	54726906	rs551782627	A	G	0.999625	1	1	1	1	0.997912
Intron 3	54726909	rs371200159	A	G	0.00581614	-	0.00702811	0.00709939	-	0.0125261
Intron 3	54726927	rs590524	C	T,A	0.953471	0.994835	0.962851	0.859026	1	0.965553
Intron 3	54726997	rs71365489	G	A	0.321388	0.304752	0.292169	0.290061	0.466734	0.264092
Intron 3	54727011	rs369657207	A	G	0.987242	0.94938	0.983936	1	1	1
Intron 3	54727022	rs75158558	G	T	0.829831	0.811983	0.871486	0.971602	0.543347	0.906054
Intron 3	54727036	.	T	C	0.999062	1	1	1	0.99496	1
Intron 3	54727039	rs112652930	T	C	0.966792	0.958678	0.968876	0.967546	1	0.935282

Intron 3	54727050	rs113582651	G	T	0.966792	0.958678	0.968876	0.967546	1	0.935282
Intron 3	54727152	rs111551907	G	A	0.957786	0.840909	0.938755	0.998986	1	1
Intron 3	54727160	.	A	G	0.999812	1	1	1	0.998992	1
Intron 3	54727162	rs929156449	G	C	0.999812	0.998967	1	1	1	1
Intron 3	54727164	rs34830246	A	C	0.193433	0.27686	0.2249	0.262677	0.00907258	0.170146
Intron 3	54727172	.	G	A	0.999812	1	1	0.998986	1	1
Intron 3	54727188	.	A	T	0.999812	1	1	1	0.998992	1
Intron 3	54727198	.	A	G	0.999812	1	0.998996	1	1	1
Intron 3	54727226	rs370722088	C	T	0.999812	1	1	1	0.998992	1
Intron 3	54727227	rs894513709	G	T	0.999812	1	1	0.998986	1	1
Intron 3	54727239	rs34614411	G	A	0.193433	0.27686	0.2249	0.262677	0.00907258	0.170146
Intron 3	54727248	.	A	G	0.999812	1	1	1	0.998992	1
Intron 3	54727256	.	C	T	0.999812	1	1	0.998986	1	1
Intron 3	54727258	.	A	G	0.999812	1	1	0.998986	1	1
Intron 3	54727269	rs527821909	T	C	0.999062	1	1	0.998986	1	0.995825
Intron 3	54727304	rs35910454	A	G	0.193809	0.278926	0.2249	0.262677	0.00907258	0.170146
Intron 3	54727323	rs35508109	C	T	0.874109	0.804752	0.89257	0.881339	0.896169	0.894572
Intron 3	54727329	.	AAC	A	0.999625	1	0.997992	1	1	1
Intron 3	54727332	rs774298329	AC	A	0.949906	1	0.96988	0.934077	0.908266	0.942589
Intron 3	54727363	rs1017217027	A	T	0.99925	1	0.998996	1	1	0.996868
Intron 3	54727369	rs34876880	C	A	0.319325	0.472107	0.332329	0.381339	0.112903	0.275574
Intron 3	54727371	rs373763045	A	C	0.980113	1	0.980924	0.925963	1	0.994781
Intron 3	54727374	.	A	C	0.998874	0.993802	1	1	1	1
Intron 3	54727408	.	G	A	0.999812	0.998967	1	1	1	1
Intron 3	54727429	rs774846039	G	A	0.999812	1	1	0.998986	1	1
Intron 3	54727446	.	C	T	0.999812	1	1	1	1	0.998956
Intron 3	54727484	rs111271439	A	G	0.942964	0.783058	0.917671	0.998986	1	1
Intron 3	54727491	rs759945936	G	A	0.99606	1	0.992972	0.995943	1	0.991649
Intron 3	54727502	rs35093205	C	G	0.208255	0.334711	0.245984	0.262677	0.00907258	0.170146
Intron 3	54727513	rs575080585	G	C	0.998311	1	0.996988	0.996957	1	0.997912
Intron 3	54727519	rs761336421	G	C	0.998874	1	0.997992	0.998986	1	1
Intron 3	54727547	rs540543236	T	C	0.999625	1	1	1	0.997984	1
Intron 3	54727581	rs547585792	G	A	0.971295	0.97314	0.978916	0.951318	1	0.955115
Intron 3	54727587	.	T	G	0.999812	1	1	1	1	0.998956
Exon 4	54727650	.	C	T	0.993058	0.971074	0.992972	0.998986	1	1
Exon 4	54727655	.	G	A	0.999625	0.997934	1	1	1	1
Exon 4	54727661	.	T	G	0.999812	1	1	1	1	0.998956
Exon 4	54727663	rs62132664	A	G,C	0.168293	0.265496	0.197791	0.181542	0.0110887	0.164927
Exon 4	54727671	rs544733307	C	T	0.999812	1	1	1	0.998992	1
Exon 4	54727672	.	G	T	0.996998	0.987603	0.997992	1	1	1
Exon 4	54727689	rs564636994	G	C	0.999812	1	1	0.998986	1	1
Exon 4	54727702	rs62132665	G	T	0.932458	0.950413	0.949799	0.931034	0.896169	0.939457
Exon 4	54727710	rs199609975	G	A	0.981801	0.996901	0.980924	0.958418	1	0.977035
Exon 4	54727742	.	GA	G	0.999812	1	1	1	1	0.998956
Exon 4	54727751	rs200260605	C	T	0.999062	0.996901	1	0.998986	1	0.998956
Exon 4	54727752	rs201591361	G	T	0.999812	0.998967	1	1	1	1
Exon 4	54727757	rs270790	G	A	0.481989	0.684917	0.534137	0.491886	0.295363	0.394572
Exon 4	54727758	.	T	C	0.999812	1	1	1	1	0.998956
Exon 4	54727761	.	G	C	0.999812	1	1	0.998986	1	1

Exon 4	54727771	rs546562878	C	T	0.994747	1	0.995984	0.985801	1	0.992693
Exon 4	54727772	.	G	A	0.999812	1	1	1	1	0.998956
Exon 4	54727797	.	C	T	0.999812	1	1	0.998986	1	1
Exon 4	54727820	rs113773573	C	T	0.999812	0.998967	1	1	1	1
Exon 4	54727855	.	T	A	0.999812	1	1	1	1	0.998956
Exon 4	54727876	rs2075037	G	A	0.949906	1	0.96988	0.934077	0.908266	0.942589
Intron 4	54727927	rs569383474	C	A	0.998874	0.996901	0.996988	1	1	1
Intron 4	54727972	rs777944452	G	A	0.999625	1	1	0.998986	1	0.998956
Intron 4	54727975	rs537906244	A	G	0.99925	1	0.995984	1	1	1
Intron 4	54727997	.	G	A	0.99925	1	1	1	0.995968	1
Intron 4	54728084	rs555166738	C	T	0.999625	1	1	0.997972	1	1
Intron 4	54728173	rs62132666	C	T	0.193433	0.27686	0.2249	0.262677	0.00907258	0.170146
Intron 4	54728183	.	A	C	0.999812	1	1	1	1	0.998956
Intron 4	54728192	.	T	A	0.999812	1	1	1	1	0.998956
Intron 4	54728225	.	A	G	0.999812	1	1	1	0.998992	1
Intron 4	54728227	.	A	G	0.999437	1	1	1	1	0.996868
Intron 4	54728289	rs367705134	A	G	0.998687	0.994835	0.998996	1	1	1
Intron 4	54728318	rs34086533	A	G	0.193433	0.27686	0.2249	0.262677	0.00907258	0.170146
Intron 4	54728348	.	T	C	0.999812	1	1	1	1	0.998956
Intron 4	54728350	rs34478220	T	C	0.193433	0.27686	0.2249	0.262677	0.00907258	0.170146
Intron 4	54728362	.	T	A	0.999812	1	1	0.998986	1	1
Intron 4	54728377	.	C	G	0.999812	1	1	1	0.998992	1
Intron 4	54728394	.	C	T	0.999812	1	1	1	1	0.998956
Intron 4	54728395	rs35550185	A	G	0.314447	0.448347	0.329317	0.381339	0.112903	0.275574
Intron 4	54728397	rs772521181	C	A	0.998311	1	1	0.996957	1	0.996868
Intron 4	54728398	.	G	A	0.999812	1	1	0.998986	1	1
Intron 4	54728401	rs34223723	C	G	0.319325	0.472107	0.332329	0.381339	0.112903	0.275574
Intron 4	54728417	.	C	A	0.99925	1	0.996988	1	1	1
Intron 4	54728451	rs35476492	C	T	0.193433	0.27686	0.2249	0.262677	0.00907258	0.170146
Intron 4	54728458	rs373135951	T	A	0.995685	1	1	1	1	0.975992
Intron 4	54728466	.	G	T	0.999812	1	1	1	0.998992	1
Intron 4	54728493	.	A	G	0.999812	1	1	1	0.998992	1
Intron 4	54728520	.	T	A	0.999812	1	1	1	1	0.998956
Intron 4	54728576	.	G	A	0.999812	1	1	1	1	0.998956
Intron 4	54728609	.	G	C	0.999812	1	1	1	1	1
Intron 4	54728640	rs369579518	T	A	0.990432	0.965909	0.985944	1	1	1
Intron 4	54728648	rs113493004	T	C	0.874109	0.804752	0.89257	0.881339	0.896169	0.894572
Intron 4	54728660	.	G	A	0.999812	1	1	1	0.998992	1
Intron 4	54728675	rs374848163	G	A	0.994371	0.980372	0.990964	0.998986	1	1
Intron 4	54728681	.	T	C	0.999437	1	1	1	1	0.996868
Intron 4	54728685	rs112921756	A	C	0.966792	0.958678	0.968876	0.967546	1	0.935282
Intron 4	54728690	.	G	A	0.998311	0.991736	0.998996	1	1	1
Intron 4	54728703	.	C	T	0.999812	1	1	1	0.998992	1
Intron 4	54728751	rs995518976	G	C	0.999812	1	1	1	1	0.998956
Intron 4	54728757	rs35097716	A	G	0.193433	0.27686	0.2249	0.262677	0.00907258	0.170146
Intron 4	54728793	.	A	G	0.999812	1	1	1	1	0.998956
Intron 4	54728824	.	G	C	0.999812	1	1	0.998986	1	1
Intron 4	54728826	.	G	C	0.998687	1	1	1	1	0.992693
Intron 4	54728847	rs35915667	C	T	0.193433	0.27686	0.2249	0.262677	0.00907258	0.170146

Intron 4	54728849	.	G	T	0.999812	1	1	1	1	0.998956
Intron 4	54728857	rs201849144	T	C	0.870169	0.96281	0.930723	0.973631	0.543347	0.906054
Intron 4	54728861	.	T	C	0.999812	0.998967	1	1	1	1
Intron 4	54728867	rs113804824	A	G	0.966792	0.958678	0.968876	0.967546	1	0.935282
Intron 4	54728881	rs62132672	G	T	0.874109	0.804752	0.89257	0.881339	0.896169	0.894572
Intron 4	54728908	.	TAGAC	T	0.999812	1	1	0.998986	1	1
Intron 4	54728938	.	T	C	0.999812	1	1	1	1	1
Intron 4	54728945	rs373754902	G	C	0.999437	0.997934	0.998996	1	1	1
Intron 4	54728962	rs529828095	A	G	0.999625	0.998967	0.998996	1	1	1
Intron 4	54728978	rs191937283	T	C	0.999812	1	0.998996	1	1	1
Intron 4	54728988	.	TAGAG	T	0.999812	1	1	1	0.998992	1
Intron 4	54729000	rs148747226	AAGAC	A	0.616698	0.740702	0.610442	0.518256	0.752016	0.488518
Intron 4	54729002	.	GACA	AACA,*	0.605066	0.696281	0.594378	0.517241	0.752016	0.488518
Intron 4	54729014	rs530377466	CATG	C	0.99925	1	1	1	0.997984	0.997912
Intron 4	54729043	.	A	G	0.999812	1	1	1	1	0.998956
Intron 4	54729056	rs1053108037	G	C	0.999812	1	1	0.998986	1	1
Intron 4	54729153	rs34823346	A	G	0.15122	0.117769	0.163655	0.261663	0.00907258	0.170146
Intron 4	54729188	.	A	T	0.999812	1	1	1	0.998992	1
Intron 4	54729209	.	TAAAG	T	0.999812	1	1	1	1	0.998956
Intron 4	54729235	.	C	G	0.999812	1	1	1	0.998992	1
Intron 4	54729280	.	A	T	0.999625	1	1	1	0.997984	1
Intron 4	54729287	rs111505776	C	A	0.957411	0.923554	0.953815	0.967546	1	0.935282
Intron 4	54729312	.	G	A	0.996998	1	1	1	1	0.983299
Intron 4	54729323	.	A	G	0.998687	0.994835	0.997992	1	1	1
Intron 4	54729362	.	TG	T	0.999812	0.998967	1	1	1	1
Intron 4	54729372	rs968294997	G	A	0.998874	0.994835	1	1	1	1
Intron 4	54729380	.	G	C	0.999812	0.998967	1	1	1	1
Intron 4	54729385	rs1031353359	A	G	0.999812	0.998967	1	1	1	1
Intron 4	54729404	rs199843005	A	G	0.942589	0.994835	0.899598	0.995943	0.894153	0.916493
Intron 4	54729414	.	A	T	0.999812	0.998967	1	1	1	1
Intron 4	54729417	rs374768578	G	A	0.996998	0.984504	0.998996	1	1	1
Intron 4	54729435	.	T	A	0.999812	1	1	1	0.998992	1
Intron 4	54729453	rs28406252	G	C	0.319325	0.472107	0.332329	0.381339	0.112903	0.275574
Intron 4	54729454	.	G	A	0.998687	1	1	1	0.992944	1
Intron 4	54729485	rs28637725	T	C	0.277111	0.313017	0.271084	0.380325	0.112903	0.275574
Exon 5	54729493	rs764066889	G	A	0.998874	1	0.997992	0.996957	1	1
Exon 5	54729498	.	T	C	0.999812	1	0.998996	1	1	1
Exon 5	54729500	.	T	C	0.999625	1	1	1	1	0.997912
Exon 5	54729527	rs550078338	G	A	0.999812	1	1	1	1	0.998956
Exon 5	54729529	rs112412050	G	A	0.957786	0.840909	0.938755	0.998986	1	1
Exon 5	54729536	rs112156439	G	A	0.981801	0.923554	0.983936	1	1	1
Exon 5	54729561	.	T	G	0.99925	1	1	1	0.996976	1
Exon 5	54729574	.	C	T	0.999812	1	1	1	0.998992	1
Exon 5	54729575	rs113504910	C	T	0.957786	0.840909	0.938755	0.998986	1	1
Exon 5	54729587	.	T	G	0.999812	1	1	0.998986	1	1
Exon 5	54729612	rs903951354	G	A	0.999437	1	1	1	0.996976	1
Exon 5	54729614	rs11575928	G	A	0.15122	0.117769	0.163655	0.261663	0.00907258	0.170146
Exon 5	54729640	.	C	A	0.999812	1	1	1	0.998992	1
Exon 5	54729641	rs62132678	A	G	0.193433	0.27686	0.2249	0.262677	0.00907258	0.170146

Exon 5	54729706	rs2302422	A	C	0.874109	0.804752	0.89257	0.881339	0.896169	0.894572
Exon 5	54729708	.	T	G	0.999812	1	1	0.998986	1	1
Exon 5	54729715	.	G	A	0.999812	0.998967	1	1	1	1
Exon 5	54729718	.	T	A	0.999812	1	0.998996	1	1	1
Exon 5	54729739	.	C	G	0.999812	1	0.998996	1	1	1
Exon 5	54729743	rs34847288	T	C	0.139587	0.0661157	0.149598	0.266734	0.00907258	0.170146
Exon 5	54729753	.	G	C	0.99925	1	1	1	0.995968	1
Exon 5	54729774	rs111516669	G	A	0.993621	0.975207	0.98996	1	1	1
Intron 5	54729809	.	C	T	0.999812	0.998967	1	1	1	1
Intron 5	54729837	.	T	G	0.999812	1	0.998996	1	1	1
Intron 5	54729855	.	T	C	0.999812	0.998967	1	1	1	1
Intron 5	54729882	.	G	A	0.999625	1	0.998996	1	1	0.998956
Intron 5	54729912	rs570352855	C	T	0.971107	0.97314	0.977912	0.951318	1	0.955115
Intron 5	54729913	.	G	A	0.999812	1	0.998996	1	1	1
Intron 5	54729918	.	CA	C	0.996248	0.983471	0.996988	1	1	1
Intron 5	54729923	rs62132684	C	G	0.319325	0.472107	0.332329	0.381339	0.112903	0.275574
Intron 5	54729924	.	G	A	0.999625	1	0.998996	0.998986	1	1
Intron 5	54729946	.	T	C	0.999812	0.998967	1	1	1	1
Intron 5	54729952	.	C	A	0.999812	1	1	1	1	0.998956
Intron 5	54729975	rs35590954	T	G	0.15122	0.117769	0.163655	0.261663	0.00907258	0.170146
Intron 5	54730017	rs373439815	G	C	0.999437	0.996901	1	1	1	1
Intron 5	54730034	rs377458147	A	G	0.987054	0.948347	0.983936	1	1	1
Intron 5	54730039	.	A	C	0.999812	1	1	1	0.998992	1
Intron 5	54730045	.	A	G	0.999812	1	1	1	1	1
Intron 5	54730109	.	G	A	0.999812	0.998967	1	1	1	1
Intron 5	54730154	rs371852753	C	T	0.999437	0.998967	1	1	0.997984	1
Intron 5	54730182	.	G	T	0.999812	0.998967	1	1	1	1
Intron 5	54730207	.	C	A	0.999437	1	1	1	0.996976	1
Intron 5	54730241	rs636129	G	T	0.996998	1	0.996988	0.98783	1	1
Intron 5	54730281	.	C	T	0.999812	0.998967	1	1	1	1
Intron 5	54730292	rs35705724	G	T	0.319325	0.472107	0.332329	0.381339	0.112903	0.275574
Intron 5	54730308	.	C	A	0.999812	1	1	1	1	0.998956
Intron 5	54730309	.	T	G	0.999812	0.998967	1	1	1	1
Intron 5	54730348	.	G	C	0.999812	1	1	1	1	1
Intron 5	54730361	.	C	T	0.999812	0.998967	1	1	1	1
Intron 5	54730383	.	A	G	0.99925	1	1	0.998986	1	1
Intron 5	54730388	.	T	A	0.999812	1	1	1	0.998992	1
Intron 5	54730399	.	C	A	0.999437	0.996901	1	1	1	1
Intron 5	54730411	rs34483320	T	A	0.193433	0.27686	0.2249	0.262677	0.00907258	0.170146
Intron 5	54730428	rs886680001	G	T	0.981989	0.923554	0.98494	1	1	1
Intron 5	54730430	.	C	T,A	0.999625	1	1	1	1	0.997912
Intron 5	54730431	.	G	A	0.999437	0.997934	0.998996	1	1	1
Intron 5	54730462	.	T	C	0.999812	0.998967	1	1	1	1
Intron 5	54730469	rs35469929	T	G	0.192683	0.272727	0.2249	0.262677	0.00907258	0.170146
Intron 5	54730478	.	G	A	0.999437	1	1	1	1	1
Intron 5	54730506	.	G	A	0.997186	1	1	1	1	0.984342
Intron 5	54730521	.	C	T	0.999625	0.997934	1	1	1	1
Intron 5	54730534	rs370045784	A	C	0.999812	1	1	1	0.998992	1
Intron 5	54730575	.	A	AG	0.999812	1	1	1	1	0.998956

Intron 5	54730599	.	ATTC	A	0.999812	1	1	1	0.998992	1
Intron 5	54730645	rs375000294	G	A	0.987054	0.948347	0.983936	1	1	1
Intron 5	54730660	.	T	C	0.999812	0.998967	1	1	1	1
Intron 5	54730671	.	G	A	0.999812	1	1	1	0.998992	1
Intron 5	54730681	.	T	G	0.99531	0.983471	0.992972	1	1	1
Intron 5	54730685	.	A	G	0.999812	1	1	1	0.998992	1
Intron 5	54730690	.	C	T	0.999812	1	1	1	1	1
Intron 5	54730700	.	G	A	0.999812	1	0.998996	1	1	1
Intron 5	54730723	.	T	G	0.999812	1	0.998996	1	1	1
Intron 5	54730785	.	T	C	0.999812	1	1	1	1	1
Intron 5	54730841	rs601298	A	G	0.769981	0.739669	0.751004	0.803245	0.709677	0.840292
Intron 5	54730916	rs538352369	C	T	0.997186	1	1	1	0.984879	1
Intron 5	54730942	rs770439636	C	T	0.997749	1	0.995984	0.996957	1	0.997912
Intron 5	54730960	.	C	T	0.999625	1	1	1	1	0.997912
Intron 5	54730963	.	G	A	0.999812	1	0.998996	1	1	1
Intron 5	54730987	.	A	G	0.999437	0.996901	1	1	1	1
Intron 5	54730992	rs903662648	G	T	0.995685	0.983471	0.992972	1	1	1
Intron 5	54730999	.	GGTTTT	G	0.999812	1	1	0.998986	1	1
Intron 5	54731013	.	T	C	0.999812	1	1	1	1	0.998956
Intron 5	54731031	rs34943893	T	A	0.137336	0.0599174	0.148594	0.261663	0.00907258	0.170146
Intron 5	54731048	.	C	T	0.999812	1	1	1	0.998992	1
Intron 5	54731056	.	G	C	0.999812	1	1	1	1	1
Intron 5	54731067	rs34802676	C	T	0.235272	0.349174	0.271084	0.295132	0.00907258	0.234864
Intron 5	54731068	.	G	A	0.998499	1	1	1	1	0.991649
Intron 5	54731073	.	C	T	0.999812	1	1	1	0.998992	1
Intron 5	54731091	.	C	T	0.981801	0.923554	0.983936	1	1	1
Intron 5	54731113	rs1018891652	T	G	0.999625	0.998967	0.998996	1	1	1
Intron 5	54731127	rs113673738	C	T	0.999812	0.998967	1	1	1	1
Intron 5	54731142	.	G	A	0.999812	0.998967	1	1	1	1
Intron 5	54731152	rs200351441	C	A	0.995872	1	0.990964	1	0.991935	1
Intron 5	54731203	rs775483394	G	A	0.998874	0.996901	0.996988	1	1	1
Intron 5	54731234	.	G	A	0.999812	1	1	1	0.998992	1
Intron 5	54731244	.	G	A	0.999437	0.996901	1	1	1	1
Intron 5	54731267	rs760631993	A	G	0.99925	1	0.998996	0.996957	1	1
Intron 5	54731287	.	T	C	0.999812	1	1	1	1	1
Intron 5	54731302	.	G	A	0.999625	1	1	1	1	0.997912
Intron 5	54731304	.	C	T	0.999812	1	1	1	1	0.998956
Intron 5	54731351	.	G	T	0.999062	0.994835	1	1	1	1
Intron 5	54731365	rs34344117	A	G	0.603565	0.735537	0.599398	0.48073	0.752016	0.484342
Intron 5	54731367	.	A	G	0.999625	1	1	0.998986	1	1
Intron 5	54731390	.	T	C	0.999812	1	1	1	0.998992	1
Intron 5	54731392	rs757339918	C	A	0.997749	1	0.995984	0.996957	1	0.997912
Intron 5	54731397	.	C	T	0.999812	1	1	1	0.998992	1
Intron 5	54731455	rs370535167	A	T	0.999437	1	1	1	1	0.996868
Intron 5	54731478	rs199505898	T	C	0.984991	1	0.997992	1	0.924395	0.997912
Intron 5	54731502	rs35403782	A	T	0.137336	0.0599174	0.148594	0.261663	0.00907258	0.170146
Intron 5	54731529	rs112711994	A	G	0.957223	0.921488	0.954819	0.967546	1	0.935282
Intron 5	54731572	rs112849678	A	G	0.94409	0.785124	0.924699	0.998986	1	1
Intron 5	54731576	.	T	C	0.999625	0.997934	1	1	1	1

Intron 5	54731593	.	C	T	0.999437	0.996901	1	1	1	1
Intron 5	54731636	.	A	G	0.999812	1	1	0.998986	1	1
Intron 5	54731648	rs2240934	C	G	0.641088	0.453512	0.700803	0.821501	0.439516	0.735908
Intron 5	54731700	rs370087238	G	A	0.999625	1	1	1	0.997984	1
Intron 5	54731736	rs111458223	C	T	0.915197	0.76343	0.893574	0.966531	1	0.935282
Intron 5	54731749	.	C	A	0.999625	1	1	1	1	0.997912
Intron 5	54731752	rs2240935	C	T	0.911069	0.944215	0.871486	0.959432	0.81754	0.949896
Intron 5	54731768	.	A	G	0.999812	1	1	1	0.998992	1
Intron 5	54731789	.	G	A	0.999812	0.998967	1	1	1	1
Intron 5	54731811	rs113877412	C	A	0.913884	0.757231	0.89257	0.966531	1	0.935282
Intron 5	54731824	.	C	T	0.999812	1	1	1	1	0.998956
Intron 5	54731828	rs2240936	C	T	0.36848	0.386364	0.342369	0.321501	0.481855	0.329854
Intron 5	54731904	rs377062359	A	G	0.987054	0.948347	0.983936	1	1	1
Intron 5	54731908	.	T	C	0.999625	1	0.998996	0.998986	1	1
Intron 5	54731938	rs796586057	G	A	0.996435	0.988636	0.993976	1	0.998992	1
Intron 5	54731945	.	C	T	0.998874	1	1	1	1	0.993737
Intron 5	54731952	.	A	T	0.998874	0.996901	0.996988	1	1	1
Intron 5	54731963	.	T	C	0.981801	0.923554	0.983936	1	1	1
Intron 5	54731975	.	A	C	0.999812	1	1	1	1	0.998956
Intron 5	54732054	rs369006795	TG	T	0.986116	0.943182	0.983936	1	1	1
Intron 5	54732058	.	G	T	0.999812	0.998967	1	1	1	1
Intron 5	54732097	.	A	C	0.999812	1	1	0.998986	1	1
Intron 5	54732108	.	G	T	0.999812	1	1	1	0.998992	1
Intron 5	54732110	.	C	A	0.999812	1	1	1	0.998992	1
Intron 5	54732125	.	C	T	0.999812	1	1	0.998986	1	1
Intron 5	54732151	rs201660684	G	A	0.999062	1	1	1	0.99496	1
Intron 5	54732176	.	G	T	0.99925	0.995868	1	1	1	1
Intron 5	54732213	.	G	C	0.999812	1	0.998996	1	1	1
Intron 5	54732215	.	A	G	0.999625	1	1	1	1	0.997912
Intron 5	54732227	rs368274046	G	T,GGT	0.997749	0.991736	0.998996	1	0.996976	1
Intron 5	54732228	.	T	C	0.999812	1	1	1	0.998992	1
Intron 5	54732232	rs538239095	T	G,TTG,TG	0.988931	0.997934	1	0.998986	1	0.943633
Intron 5	54732234	rs200035731	G	T	0.868856	0.971074	0.931727	0.972617	0.529234	0.903967
Intron 5	54732243	.	TG	T	0.999437	1	1	1	0.996976	1
Intron 5	54732244	.	G	T,*	0.997561	0.990702	0.998996	1	0.996976	1
Intron 5	54732245	.	TG	T	0.999437	1	1	1	0.996976	1
Intron 5	54732246	rs2240938	G	T,*	0.539775	0.47314	0.539157	0.719067	0.265121	0.662839
Intron 5	54732248	rs370250465	T	G	0.963039	0.965909	0.975904	0.951318	0.992944	0.926931
Intron 5	54732250	.	TTTG	T	0.977486	1	0.975904	0.920892	1	0.994781
Intron 5	54732253	.	G	GTTTTTTTTT,*	0.976548	0.995868	0.9749	0.920892	1	0.994781
Intron 5	54732285	rs201899476	C	T	0.868668	0.970041	0.931727	0.973631	0.527218	0.90501
Intron 5	54732315	.	A	G	0.999812	1	0.998996	1	1	1
Intron 5	54732339	.	C	T	0.999625	0.997934	1	1	1	1
Intron 5	54732375	.	C	T	0.997749	1	0.996988	0.995943	1	0.997912
Intron 5	54732400	rs111865373	G	A	0.999812	0.998967	1	1	1	1
Intron 5	54732424	.	T	G	0.999812	1	0.998996	1	1	1
Intron 5	54732430	rs35813733	AT	A	0.145966	0.0599174	0.151606	0.262677	0.00907258	0.212944
Intron 5	54732432	.	T	C	0.998874	1	1	1	1	0.993737
Intron 5	54732469	rs200304792	G	C	0.954784	1	0.961847	0.93002	0.993952	0.908142

Intron 5	54732476	.	C	T	0.999625	1	1	1	0.998992	0.998956
Intron 5	54732486	.	A	G	0.999625	0.997934	1	1	1	1
Intron 5	54732548	.	C	T	0.999812	1	1	1	1	0.998956
Intron 5	54732586	.	G	A	0.999812	0.998967	1	1	1	1
Intron 5	54732593	rs4392183	T	C	0.90394	0.716942	0.88253	0.966531	1	0.935282
Intron 5	54732610	.	G	A	0.999812	1	0.998996	1	1	1
Intron 5	54732630	.	A	G	0.999812	0.998967	1	1	1	1
Intron 5	54732665	.	C	G	0.999812	0.998967	1	1	1	1
Intron 5	54732675	.	A	G	0.999812	1	1	1	0.998992	1
Intron 5	54732680	.	C	T	0.999625	1	0.998996	1	0.998992	1
Intron 5	54732687	.	A	G	0.999812	1	0.998996	1	1	1
Intron 5	54732710	rs186578807	C	T	0.999812	1	1	0.998986	1	1
Intron 5	54732734	.	T	C	0.999812	1	1	1	1	0.998956
Intron 5	54732738	.	T	C	0.999812	0.998967	1	1	1	1
Intron 5	54732772	.	G	A	0.999812	1	0.998996	1	1	1
Intron 5	54732778	rs150479380	G	A	0.693433	0.885331	0.759036	0.659229	0.518145	0.653445
Intron 5	54732779	rs137937971	A	G	0.693433	0.885331	0.759036	0.659229	0.518145	0.653445
Intron 5	54732798	rs112596073	A	G	0.923827	0.789256	0.909639	0.967546	1	0.935282
Intron 5	54732804	rs112806977	A	G	0.923827	0.789256	0.909639	0.967546	1	0.935282
Intron 5	54732811	rs113779134	T	C	0.923827	0.789256	0.909639	0.967546	1	0.935282
Intron 5	54732818	rs111977642	C	T	0.961914	0.855372	0.944779	0.998986	1	1
Intron 5	54732821	rs761723758	A	G	0.999812	1	0.998996	1	1	1
Intron 5	54732846	rs765347238	G	C	0.999812	1	1	0.998986	1	1
Intron 5	54732856	.	C	T	0.976173	0.89876	0.977912	1	1	1
Intron 5	54732857	rs3815425	A	G	0.541088	0.362603	0.571285	0.639959	0.518145	0.58977
Intron 5	54732861	rs3815426	T	G	0.537523	0.362603	0.570281	0.623732	0.518145	0.58977
Intron 5	54732867	rs112791358	T	C	0.852908	0.503099	0.829317	0.967546	1	0.935282
Intron 5	54732878	rs3815427	C	A	0.694559	0.880165	0.752008	0.673428	0.520161	0.654489
Intron 5	54732913	.	C	G	0.999812	1	1	0.998986	1	1
Intron 5	54732914	rs112894172	C	A	0.978424	0.914256	0.972892	1	1	1
Intron 5	54732916	.	G	C	0.999437	1	1	1	0.996976	1
Intron 5	54732919	rs111864792	T	A	0.981801	0.923554	0.983936	1	1	1
Intron 5	54732922	rs3815428	C	T	0.666604	0.759298	0.722892	0.673428	0.520161	0.660752
Intron 5	54732937	rs781447828	A	C	0.971482	0.974174	0.978916	0.953347	1	0.956159
Intron 5	54732958	.	G	A	0.999625	0.997934	1	1	1	1
Intron 5	54732978	rs3815429	C	G,T	0.671107	0.770661	0.728916	0.674442	0.52621	0.659708
Intron 5	54732996	.	C	T	0.999625	1	1	0.997972	1	1
Intron 5	54732997	.	G	A	0.999812	1	1	1	1	0.998956
Intron 5	54733011	rs191029781	A	G	0.99531	0.987603	0.997992	1	0.989919	0.998956
Intron 5	54733012	.	C	T	0.999625	1	0.997992	1	1	1
Intron 5	54733013	rs370823005	G	A	0.999812	1	1	1	0.998992	1
Intron 5	54733021	.	G	C	0.99925	0.995868	1	1	1	1
Intron 5	54733041	rs113622785	G	A	0.9803	0.923554	0.973896	1	1	1
Intron 5	54733049	.	G	T	0.999812	1	1	0.998986	1	1
Intron 5	54733066	.	T	G	0.999625	1	1	1	1	0.997912
Intron 5	54733070	rs113775492	G	A	0.901126	0.826446	0.908635	0.900609	1	0.868476
Intron 5	54733087	.	G	A	0.999812	1	1	1	1	1
Intron 5	54733112	rs369380140	C	T	0.999812	1	0.998996	1	1	1
Intron 5	54733116	.	T	C	0.999812	0.998967	1	1	1	1

Intron 5	54733126	.	C	G	0.999812	1	1	1	1	0.998956
Intron 5	54733135	.	T	A	0.996811	0.985537	0.997992	1	1	1
Intron 5	54733155	.	G	A	0.999812	1	1	1	0.998992	1
Intron 5	54733163	.	T	C	0.999812	1	1	1	1	1
Intron 5	54733177	rs3815430	A	C	0.811257	0.840909	0.811245	0.731237	0.900202	0.77453
Intron 5	54733194	rs199505043	C	T	0.989118	1	0.956827	0.998986	0.986895	0.998956
Intron 5	54733195	.	G	A	0.999625	1	0.997992	1	1	1
Intron 5	54733233	.	G	GTA	0.99925	0.995868	1	1	1	1
Intron 5	54733294	.	G	A	0.999812	1	1	1	1	1
Intron 5	54733311	.	C	G,T	0.999437	1	1	1	1	0.996868
Intron 5	54733316	rs917550268	A	G	0.999812	0.998967	1	1	1	1
Intron 5	54733320	.	G	A	0.999812	1	1	1	1	0.998956
Intron 5	54733361	.	G	A	0.998874	1	1	1	0.993952	1
Intron 5	54733395	rs112668569	A	G	0.999062	0.996901	0.998996	1	1	1
Intron 5	54733400	rs768495491	C	T	0.999812	1	1	0.998986	1	1
Intron 5	54733481	.	C	T	0.999812	1	1	1	1	0.998956
Intron 5	54733482	rs1052217470	G	A	0.999812	1	0.998996	1	1	1
Intron 5	54733494	.	T	A	0.999812	1	1	0.998986	1	1
Intron 5	54733515	rs892332135	T	A	0.999625	1	0.997992	1	1	1
Intron 5	54733522	rs1008063788	C	CA	0.997373	1	1	0.998986	0.986895	1
Intron 5	54733541	rs1019490743	A	C	0.998311	0.991736	1	0.998986	1	1
Intron 5	54733546	.	T	A	0.999437	1	1	1	0.997984	0.998956
Intron 5	54733549	rs199755001	T	TTATAA,TTAA,TATAA	0.696435	0.715909	0.736948	0.792089	0.517137	0.686848
Intron 5	54733573	.	G	A	0.999812	0.998967	1	1	1	1
Intron 5	54733611	rs76445578	G	A	0.790994	0.81405	0.825301	0.913793	0.514113	0.860125
Intron 5	54733617	.	T	C	0.999812	1	0.998996	1	1	1
Intron 5	54733639	.	C	T	0.999812	1	1	0.998986	1	1
Intron 5	54733691	rs568669706	G	A	0.996811	0.987603	0.995984	1	1	1
Intron 5	54733752	rs585270	G	A	0.832458	0.97314	0.90261	0.880325	0.541331	0.847599
Intron 5	54733754	rs112068566	G	A	0.949906	0.796488	0.944779	0.998986	1	0.998956
Intron 5	54733758	.	T	C	0.999812	1	1	0.998986	1	1
Intron 5	54733766	rs554086126	G	A	0.999625	0.997934	1	1	1	1
Intron 5	54733780	rs74533971	C	A	0.844841	0.82438	0.756024	0.975659	0.683468	0.95929
Intron 5	54733819	rs112166802	A	G	0.997186	0.989669	0.995984	1	1	1
Intron 5	54733858	rs558142447	T	C	0.999812	1	1	1	1	0.998956
Intron 5	54733859	rs17173055	A	G	0.918762	0.682851	0.891566	0.997972	1	0.997912
Intron 5	54733865	.	C	A	0.999812	1	1	1	1	0.998956
Intron 5	54733894	rs375293550	G	C	0.992683	0.980372	0.993976	1	0.988911	0.998956
Intron 5	54733926	rs560817663	T	C	0.999812	1	0.998996	1	1	1
Intron 5	54733968	rs1015035167	A	G	0.999812	0.998967	1	1	1	1
Intron 5	54733974	.	C	T	0.999812	1	1	1	0.998992	1
Intron 5	54733978	.	G	T	0.999812	1	1	0.998986	1	1
Intron 5	54733984	.	G	C	0.999812	1	1	1	1	0.998956
Intron 5	54733996	rs970041708	T	G	0.999812	0.998967	1	1	1	1
Intron 5	54734001	.	C	T	0.999625	1	1	1	1	0.997912
Intron 5	54734017	.	G	A	0.999812	1	1	0.998986	1	1
Intron 5	54734024	rs375100424	A	T	0.965478	0.864669	0.956827	1	1	1
Intron 5	54734049	.	T	G	0.999812	0.998967	1	1	1	1
Intron 5	54734053	rs540436475	A	G	0.999812	1	0.998996	1	1	1

Intron 5	54734088	.	G	C	0.999812	1	0.998996	1	1	1
Intron 5	54734091	.	T	A	0.999812	1	1	1	1	0.998956
Intron 5	54734121	.	T	A	0.999812	0.998967	1	1	1	1
Intron 5	54734162	rs67542354	A	AT,ATT	0.532833	0.556818	0.481928	0.64503	0.257056	0.685804
Intron 5	54734163	.	T	TA	0.999625	1	1	1	1	0.997912
Intron 5	54734173	.	T	G	0.998687	1	1	1	0.992944	1
Intron 5	54734174	rs148984435	C	G	0.993809	0.980372	0.988956	1	1	0.998956
Intron 5	54734250	rs647211	T	C	0.698687	0.677686	0.708835	0.70284	0.871976	0.551148
Intron 5	54734264	rs532470357	T	C	0.999812	1	1	1	0.998992	1
Intron 5	54734265	rs552465725	G	T	0.999812	0.998967	1	1	1	1
Intron 5	54734268	rs562938462	A	C	0.99925	1	1	1	0.995968	1
Intron 5	54734271	rs587560	C	T	0.699437	0.677686	0.710843	0.70284	0.871976	0.551148
Intron 5	54734304	.	G	A	0.996998	0.990702	0.995984	0.998986	1	0.998956
Intron 5	54734317	.	G	A	0.999625	1	1	1	1	1
Intron 5	54734346	rs531587774	G	T	0.999812	1	0.998996	1	1	1
Intron 5	54734363	.	C	T	0.999812	1	1	1	1	0.998956
Intron 5	54734374	rs548712626	C	T	0.999625	1	1	0.998986	1	0.998956
Intron 5	54734389	rs648029	A	G	0.644278	0.540289	0.64257	0.657201	0.871976	0.524008
Intron 5	54734390	.	C	A	0.999812	1	0.998996	1	1	1
Intron 5	54734405	rs369801958	G	C,A	0.997936	1	0.996988	1	0.991935	1
Intron 5	54734420	rs547603938	G	A	0.999625	1	1	1	1	0.997912
Intron 5	54734434	.	G	C	0.999812	1	0.998996	1	1	1
Intron 5	54734438	.	G	T	0.999812	1	1	1	1	0.998956
Intron 5	54734439	.	A	G	0.999812	1	0.998996	1	1	1
Intron 5	54734445	rs570674611	G	A	0.999812	1	0.998996	1	1	1
Intron 5	54734446	.	T	G	0.999812	0.998967	1	1	1	1
Intron 5	54734470	.	T	G	0.999812	1	1	1	0.998992	1
Intron 5	54734474	.	G	C	0.999812	0.998967	1	1	1	1
Intron 5	54734502	rs539774961	G	T	0.997749	0.988636	0.998996	1	1	1
Intron 5	54734515	rs556418443	G	A	0.997749	0.988636	0.998996	1	1	1
Intron 5	54734519	.	C	T	0.999812	1	1	1	1	0.998956
Intron 5	54734533	rs568881304	A	G	0.999812	1	0.998996	1	1	1
Intron 5	54734552	rs35835276	G	A,T	0.932083	0.809917	0.916667	0.952333	1	0.97286
Intron 5	54734573	rs574453246	G	A	0.999812	1	1	1	1	0.998956
Intron 5	54734590	.	T	C	0.999812	0.998967	1	1	1	1
Intron 5	54734598	rs200439305	T	C	0.999812	0.998967	1	1	1	1
Intron 5	54734605	.	C	G	0.999812	1	1	1	0.998992	1
Intron 5	54734619	rs659384	G	T	0.698124	0.673554	0.708835	0.703854	0.871976	0.551148
Intron 5	54734664	rs540303385	C	G	0.999625	0.997934	1	1	1	1
Intron 5	54734683	rs11671355	C	A	0.175422	0.00516529	0.12249	0.346856	0.128024	0.241127
Intron 5	54734698	.	C	T	0.999625	1	1	1	1	0.997912
Intron 5	54734703	rs577097617	C	G	0.999625	1	1	1	0.997984	1
Intron 5	54734704	rs546149628	A	C	0.998687	0.998967	0.997992	0.996957	1	0.998956
Intron 5	54734710	rs10412808	G	T	0.945403	0.862603	0.933735	0.953347	1	0.97286
Intron 5	54734713	rs1025555723	G	C	0.999812	0.998967	1	1	1	1
Intron 5	54734722	rs10412815	G	T	0.945403	0.862603	0.933735	0.953347	1	0.97286
Intron 5	54734728	.	A	G	0.999812	1	1	1	1	0.998956
Intron 5	54734772	rs600183	A	G	0.700375	0.680785	0.712851	0.704868	0.871976	0.551148
Intron 5	54734790	.	G	A	0.999812	1	1	0.998986	1	1

Intron 5	54734793	rs113747579	G	A	0.995122	0.983471	0.990964	1	1	1
Intron 5	54734798	rs1743323	G	A	0.698311	0.673554	0.708835	0.704868	0.871976	0.551148
Intron 5	54734804	.	G	T	0.999812	0.998967	1	1	1	1
Intron 5	54734805	.	C	T	0.999812	0.998967	1	1	1	1
Intron 5	54734839	rs1743324	C	G	0.64334	0.536157	0.64257	0.657201	0.871976	0.522965
Intron 5	54734843	.	G	A	0.999812	0.998967	1	1	1	1
Intron 5	54734846	rs142430937	G	A	0.984615	0.988636	0.981928	0.984787	1	0.967641
Intron 5	54734856	.	C	T	0.998499	1	0.998996	0.998986	0.998992	0.994781
Intron 5	54734872	rs112396168	C	T	0.698499	0.673554	0.708835	0.705882	0.871976	0.551148
Intron 5	54734874	rs375579673	C	CT	0.643902	0.536157	0.64257	0.659229	0.871976	0.524008
Intron 5	54734879	rs185715745	A	G	0.999812	1	0.998996	1	1	1
Intron 5	54734888	rs1639424	T	C	0.698499	0.673554	0.708835	0.705882	0.871976	0.551148
Intron 5	54734896	.	G	A	0.999812	0.998967	1	1	1	1
Intron 5	54734903	rs547466701	T	C	0.999812	1	1	1	0.998992	1
Intron 5	54734922	rs570736498	C	A	0.999812	1	1	1	1	0.998956
Intron 5	54734926	rs3952452	G	A	0.699062	0.672521	0.709839	0.708925	0.871976	0.551148
Intron 5	54734936	rs10412714	A	G	0.643902	0.536157	0.64257	0.659229	0.871976	0.524008
Intron 5	54734938	.	C	G	0.999812	1	1	1	0.998992	1
Intron 5	54734942	.	A	T	0.998874	1	1	1	1	0.993737
Intron 5	54734951	rs138816214	T	C	0.998687	0.994835	0.998996	1	1	1
Intron 5	54734983	rs2159663	A	C	0.698499	0.673554	0.708835	0.705882	0.871976	0.551148
Intron 5	54735007	.	C	T	0.999812	1	1	1	0.998992	1
Intron 5	54735008	rs62132698	T	C	0.174859	0.00516529	0.12249	0.34787	0.125	0.240084
Intron 5	54735026	rs554291646	A	C	0.999437	1	0.996988	1	1	1
Intron 5	54735033	rs111813709	G	T	0.982552	0.92562	0.98494	1	1	1
Intron 5	54735041	rs661147	G	T	0.698499	0.673554	0.708835	0.705882	0.871976	0.551148
Intron 5	54735080	rs568026554	C	T	0.999625	0.997934	1	1	1	1
Intron 5	54735081	rs533909615	G	A	0.999812	0.998967	1	1	1	1
Intron 5	54735128	rs553883286	C	T	0.999812	1	1	0.998986	1	1
Intron 5	54735144	rs372486889	G	T	0.991182	0.961777	0.990964	1	1	1
Intron 5	54735148	rs546013580	T	G	0.999625	1	0.997992	1	1	1
Intron 5	54735157	rs943784578	G	A	0.999812	1	1	1	1	1
Intron 5	54735196	rs556286170	A	G	0.999812	1	0.998996	1	1	1
Intron 5	54735203	rs368837758	T	A	0.999437	0.997934	0.998996	1	1	1
Intron 5	54735243	.	CT	C	0.999812	0.998967	1	1	1	1
Exon 6	54735264	rs602444	A	T,C	0.174859	0.00516529	0.12249	0.34787	0.125	0.240084
Exon 6	54735272	rs78691806	C	T	0.884053	0.995868	0.814257	0.973631	0.664315	0.956159
Exon 6	54735274	rs662386	T	C	0.698687	0.673554	0.708835	0.706897	0.871976	0.551148
Exon 6	54735304	rs760532224	C	T	0.999812	1	0.998996	1	1	1
Exon 6	54735307	rs759019961	T	TCA	0.999812	1	0.998996	1	1	1
Exon 6	54735309	rs766236460	G	C	0.999812	1	0.998996	1	1	1
Exon 6	54735310	rs777968232	CTA	C	0.999812	1	0.998996	1	1	1
Exon 6	54735341	rs189445349	G	C	0.999812	1	0.998996	1	1	1
Exon 6	54735345	rs16985907	G	C	0.945403	0.862603	0.933735	0.953347	1	0.97286
Intron 6	54735362	rs113225082	G	C	0.986679	0.947314	0.982932	0.998986	1	1
Intron 6	54735418	rs185549280	C	A	0.998687	1	0.996988	0.995943	1	1
Intron 6	54735430	rs62132699	T	A	0.942589	1	0.966867	0.909736	0.894153	0.943633
Intron 6	54735437	rs603309	C	T	0.643902	0.536157	0.64257	0.659229	0.871976	0.524008
Intron 6	54735441	.	G	GGA	0.999812	1	1	1	0.998992	1

Intron 6	54735442	.	ACT	GCT,A	0.174859	0.00516529	0.12249	0.34787	0.125	0.240084
Intron 6	54735444	rs62132701	T	A,*	0.174859	0.00516529	0.12249	0.34787	0.125	0.240084
Intron 6	54735457	rs603334	G	A,C	0.698499	0.673554	0.708835	0.705882	0.871976	0.551148
Intron 6	54735480	rs568091034	C	G	0.999437	0.997934	1	1	0.998992	1
Intron 6	54735497	.	A	G	0.999812	0.998967	1	1	1	1
Intron 6	54735506	.	C	T	0.999812	0.998967	1	1	1	1
Intron 6	54735513	.	C	T	0.999812	0.998967	1	1	1	1
Intron 6	54735515	rs673552	T	C	0.643902	0.536157	0.64257	0.659229	0.871976	0.524008
Intron 6	54735526	.	C	G	0.999812	1	1	1	1	0.998956
Intron 6	54735530	.	C	T	0.999812	1	1	1	1	0.998956
Intron 6	54735531	rs547559265	A	G	0.999812	1	0.998996	1	1	1
Intron 6	54735553	rs374846912	C	T	0.945216	0.862603	0.933735	0.952333	1	0.97286
Intron 6	54735554	rs11882698	G	A	0.698499	0.673554	0.708835	0.705882	0.871976	0.551148
Intron 6	54735563	rs539474138	C	T	0.945403	0.862603	0.933735	0.953347	1	0.97286
Intron 6	54735564	rs556098680	A	G	0.945403	0.862603	0.933735	0.953347	1	0.97286
Intron 6	54735587	.	A	G	0.999812	1	1	1	1	1
Intron 6	54735607	rs576168009	G	A	0.999812	1	1	1	0.998992	1
Intron 6	54735617	.	G	C	0.999812	1	0.998996	1	1	1
Intron 6	54735619	rs531529555	A	G	0.998499	1	1	0.992901	1	1
Intron 6	54735620	rs11882660	C	G	0.643715	0.535124	0.64257	0.659229	0.871976	0.524008
Intron 6	54735622	.	C	T	0.999812	1	0.998996	1	1	1
Intron 6	54735632	rs4020728	A	G	0.643527	0.535124	0.64257	0.659229	0.870968	0.524008
Intron 6	54735641	.	A	C	0.999812	1	1	1	1	1
Intron 6	54735653	rs2364440	A	T	0.643527	0.535124	0.64257	0.659229	0.871976	0.524008
Intron 6	54735658	rs188741619	G	C	0.999625	1	1	1	0.997984	1
Intron 6	54735664	rs4020729	A	G	0.643527	0.535124	0.64257	0.659229	0.871976	0.524008
Intron 6	54735672	.	G	A	0.999625	1	1	1	0.998992	1
Intron 6	54735680	.	A	G	0.999812	1	1	1	1	0.998956
Intron 6	54735724	rs543865012	G	C	0.999625	1	1	1	1	0.997912
Intron 6	54735726	rs563765694	A	T	0.999812	0.998967	1	1	1	1
Intron 6	54735728	.	C	T	0.999625	1	1	0.997972	1	1
Intron 6	54735734	.	C	T	0.999812	0.998967	1	1	1	1
Intron 6	54735744	rs674541	T	C	0.698499	0.673554	0.708835	0.705882	0.871976	0.551148
Intron 6	54735745	rs374765040	A	T	0.999812	1	1	0.998986	1	1
Intron 6	54735760	rs549461360	T	C	0.999812	1	1	1	1	0.998956
Intron 6	54735766	rs560082371	A	G	0.999812	1	1	1	0.998992	1
Intron 6	54735771	.	G	A	0.999437	1	1	1	0.996976	1
Intron 6	54735785	rs528757058	C	T	0.999812	1	0.998996	1	1	1
Intron 6	54735786	rs147206425	T	A	0.995497	0.978306	0.996988	1	1	1
Intron 6	54735788	rs756500958	C	T	0.999812	0.998967	1	1	1	1
Intron 6	54735800	rs747852122	A	T	0.999812	0.998967	1	1	1	1
Intron 6	54735801	rs757942332	C	T	0.999625	0.998967	1	1	0.998992	1
Intron 6	54735802	.	G	A	0.999812	1	1	1	1	0.998956
Intron 6	54735808	rs199782955	C	T	0.999625	1	1	0.997972	1	1
Intron 6	54735809	rs371297643	G	A	0.999812	1	1	1	1	0.998956
Exon 7	54735823	.	C	T	0.999812	0.998967	1	1	1	1
Exon 7	54735839	rs62132702	A	G	0.942589	1	0.966867	0.909736	0.894153	0.943633
Exon 7	54735869	rs549823982	G	A	0.999812	0.998967	1	1	1	1
Exon 7	54735872	rs141961795	G	A	0.999625	0.997934	1	1	1	1

Intron 7	54735874	rs535628491	T	C	0.999812	0.998967	1	1	1	1
Intron 7	54735882	rs28690883	C	G	0.981238	0.927686	0.972892	1	1	1
Intron 7	54735886	rs615369	C	A,G	0.643902	0.536157	0.64257	0.659229	0.871976	0.524008
Intron 7	54735897	rs675391	G	A	0.698499	0.673554	0.708835	0.705882	0.871976	0.551148
Intron 7	54735900	rs534575516	G	C	0.999625	1	1	1	0.997984	1
Intron 7	54735906	rs557747095	C	G	0.999812	1	1	1	0.998992	1
Intron 7	54735918	rs16985911	G	C	0.945403	0.862603	0.933735	0.953347	1	0.97286
Intron 7	54735926	rs615723	G	A	0.643902	0.536157	0.64257	0.659229	0.871976	0.524008
Intron 7	54735931	.	T	C	0.999812	0.998967	1	1	1	1
Intron 7	54735934	rs142677388	G	A	0.991182	0.961777	0.990964	1	1	1
Intron 7	54735958	rs574195799	T	C	0.999812	0.998967	1	1	1	1
Exon 8	54735982	rs28504993	A	T	0.945403	0.862603	0.933735	0.953347	1	0.97286
Exon 8	54735990	.	C	A	0.999812	1	1	0.998986	1	1
Exon 8	54736001	rs559942380	A	C	0.999812	1	1	1	0.998992	1
Exon 8	54736006	rs181518279	C	T	0.999812	1	1	1	0.998992	1
Exon 8	54736007	rs748343720	G	A	0.99925	0.998967	1	1	1	0.996868
Exon 8	54736017	.	A	G	0.999812	1	0.998996	1	1	1
Exon 8	54736022	rs545699217	T	G	0.999812	1	1	1	1	0.998956
Exon 8	54736025	rs564097015	G	A	0.999437	0.996901	1	1	1	1
Exon 8	54736047	rs199955684	C	T	0.996248	1	1	1	0.984879	0.994781
Exon 8	54736050	rs549649199	G	C	0.999812	1	1	1	1	0.998956
Exon 8	54736074	rs569741140	C	T	0.999812	1	1	1	1	0.998956
Exon 8	54736091	rs528935415	G	A	0.999812	1	1	1	0.998992	1
3'UTR	54736119	rs548786357	C	T	0.99606	1	1	1	0.978831	1
3'UTR	54736120	rs565504003	G	C	0.99606	1	1	1	0.978831	1
3'UTR	54736128	.	A	T	0.995122	0.983471	0.990964	1	1	1
3'UTR	54736145	rs186885062	A	G	0.999625	0.997934	1	1	1	1
3'UTR	54736147	rs676354	T	C,G	0.643715	0.535124	0.64257	0.659229	0.871976	0.524008
3'UTR	54736149	rs376880713	G	A	0.999812	1	1	1	0.998992	1
3'UTR	54736190	rs28578030	T	C	0.945216	0.862603	0.932731	0.953347	1	0.97286
3'UTR	54736219	rs537445098	T	C	0.999812	0.998967	1	1	1	1
3'UTR	54736224	.	T	C	0.998874	1	1	1	1	0.993737
3'UTR	54736240	rs617107	C	A	0.642026	0.536157	0.638554	0.654158	0.871976	0.524008
3'UTR	54736254	rs574084778	G	C	0.999812	1	1	0.998986	1	1
3'UTR	54736258	rs28506083	G	T	0.945403	0.862603	0.933735	0.953347	1	0.97286
3'UTR	54736266	.	G	A	0.999812	0.998967	1	1	1	1
3'UTR	54736277	rs28452670	T	C	0.945403	0.862603	0.933735	0.953347	1	0.97286
3'UTR	54736285	rs28483223	A	G	0.945403	0.862603	0.933735	0.953347	1	0.97286
3'UTR	54736290	rs28715339	G	A	0.944465	0.862603	0.932731	0.950304	1	0.97286
3'UTR	54736296	.	GAAC	G	0.945403	0.862603	0.933735	0.953347	1	0.97286
3'UTR	54736300	rs553635592	A	G	0.945216	0.862603	0.933735	0.953347	1	0.971816
3'UTR	54736315	rs573452122	T	G	0.999812	1	1	0.998986	1	1
3'UTR	54736326	.	TA	T	0.945403	0.862603	0.933735	0.953347	1	0.97286
3'UTR	54736337	rs545760514	C	T	0.99925	1	1	1	0.995968	1
3'UTR	54736347	.	G	C	0.999812	1	1	1	1	0.998956
3'UTR	54736349	rs28406625	G	A	0.945403	0.862603	0.933735	0.953347	1	0.97286
3'UTR	54736360	rs369357381	C	CTCCT	0.642777	0.536157	0.64257	0.659229	0.865927	0.524008
3'UTR	54736381	rs909754396	C	T	0.999625	0.997934	1	1	1	1
3'UTR	54736437	rs575714276	CT	C	0.998311	0.992769	0.998996	1	1	1

3'UTR	54736438	rs939910552	T	*	0.998311	0.992769	0.998996	1	1	1
3'UTR	54736447	.	C	T	0.999812	0.998967	1	1	1	1
3'UTR	54736448	.	G	A	0.99925	0.998967	1	0.998986	1	0.997912
3'UTR	54736461	.	A	G	0.998874	0.996901	0.998996	1	1	1
3'UTR	54736481	.	T	C	0.999812	1	1	1	1	0.998956
3'UTR	54736485	rs187183548	A	G	0.999437	1	1	0.996957	1	1
3'UTR	54736488	.	C	T	0.999812	1	1	1	1	0.998956
3'UTR	54736501	.	C	T	0.999812	0.998967	1	1	1	1
3'UTR	54736502	rs531069710	G	A	0.999812	1	1	1	0.998992	1
3'UTR	54736512	.	T	C	0.999812	1	1	1	1	0.998956
3'UTR	54736515	rs618139	G	A	0.643902	0.536157	0.64257	0.659229	0.871976	0.524008
3'UTR	54736520	rs372934281	A	ATTCCACCTTTCCTCATGTTG	0.643902	0.536157	0.64257	0.659229	0.871976	0.524008
3'UTR	54736536	rs189545361	G	C	0.998311	1	1	1	0.991935	0.998956

Table S2. *KIR3DL3* encoded proteins in population from the 1000 genomes project samples and our samples from Brazil.

KIR3DL3 proteins	Global 2n=5330	African samples (2n=968)					Admixed samples * (2n=996)						East Asian samples (2n=992)					European samples (2n=986)					South Asian samples (2n=958)					BRA
		YRI	LWK	GWD	MSL	ESN	ASW	ACB	MXL	PUR	CLM	PEL	CHB	JPT	CHS	CDX	KHV	CEU	TSI	FIN	GBR	IBS	GIH	PJL	BEB	STU	ITU	BRA
		206	196	210	168	188	122	190	128	198	188	170	206	204	200	186	196	192	210	194	180	210	204	188	172	196	198	430
KIR3DL3*001	8.59	-	1.53	0.48	-	-	5.74	3.16	4.69	8.08	7.98	3.53	3.40	5.39	3.00	1.08	1.53	19.79	18.57	19.59	18.33	15.71	12.75	11.17	13.95	15.31	19.19	11.40
KIR3DL3*002	7.30	4.37	4.08	11.90	10.12	7.98	9.02	3.68	10.94	14.65	8.51	4.12	-	-	2.50	-	0.51	12.50	11.43	9.79	11.11	13.33	6.86	9.57	8.72	8.67	7.07	7.44
KIR3DL3*003	12.10	4.37	8.16	4.76	4.76	3.19	12.30	4.74	7.03	15.66	16.49	4.71	1.46	-	2.50	4.84	4.08	16.15	16.67	13.92	17.22	14.76	25.49	20.74	23.84	30.61	20.20	18.84
KIR3DL3*004	3.83	3.40	6.63	3.33	3.57	5.85	4.10	4.74	4.69	7.07	2.66	9.41	3.40	3.43	5.00	8.60	7.14	0.52	1.43	3.09	2.22	1.43	1.47	5.32	1.74	1.02	3.03	2.33
KIR3DL3*005	2.72	13.11	12.76	7.14	6.55	10.11	11.48	11.58	-	1.01	1.06	0.59	-	-	-	-	-	-	0.48	-	-	-	-	-	-	-	-	1.40
KIR3DL3*006	2.08	-	-	-	-	-	0.82	0.53	-	3.54	1.06	0.59	1.94	2.94	6.00	6.99	6.12	3.65	3.33	1.03	2.22	1.90	0.49	-	4.07	1.53	3.03	2.56
KIR3DL3*007	1.03	-	-	-	-	-	-	-	1.56	2.02	1.60	-	-	-	-	-	-	2.60	3.81	-	2.22	1.43	1.47	3.19	2.33	1.53	2.02	1.40
KIR3DL3*008	4.26	0.49	-	0.48	-	2.66	3.28	2.63	12.50	4.04	3.72	14.71	11.65	14.71	8.00	8.60	10.20	2.08	1.90	4.64	1.11	0.95	1.47	1.06	4.65	1.02	0.51	2.79
KIR3DL3*009	15.78	14.56	18.88	13.81	20.24	22.34	11.48	21.05	27.34	14.65	20.21	32.94	19.42	17.16	17.00	18.28	19.90	13.54	7.14	13.40	8.89	15.24	10.78	6.38	11.05	11.22	12.63	13.95
KIR3DL3*010	9.94	3.40	3.06	3.33	1.19	-	1.64	2.11	4.69	3.03	2.66	17.65	46.12	42.65	40.50	32.80	34.18	1.04	0.95	3.09	2.22	1.90	0.98	6.91	4.65	1.53	3.54	3.02
KIR3DL3*011	2.16	7.28	6.63	6.67	2.98	8.51	2.46	4.74	0.78	1.01	1.60	-	-	-	-	-	-	-	0.95	1.55	1.11	-	1.96	2.13	-	3.06	2.53	1.86
KIR3DL3*012	1.03	5.83	0.51	2.38	4.76	8.51	3.28	3.68	0.78	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.23
KIR3DL3*013	3.56	-	-	0.48	-	-	0.82	2.11	3.91	5.56	5.32	0.59	-	-	-	-	-	5.73	8.57	12.37	7.78	7.62	4.41	5.85	2.91	4.59	3.54	7.67
KIR3DL3*014	6.92	11.65	6.63	12.86	13.10	12.23	5.74	8.42	3.91	8.59	9.57	-	-	-	-	1.53	-	10.42	10.48	2.58	5.56	7.62	8.82	10.11	5.23	7.14	7.58	10.70
KIR3DL3*015	2.51	4.85	1.53	1.90	1.79	1.06	3.28	1.05	1.56	1.52	1.60	1.76	1.94	1.96	5.00	4.30	5.10	1.04	3.33	0.52	3.33	2.38	3.92	2.13	2.33	2.55	2.02	3.02
KIR3DL3*016	0.64	0.97	1.02	-	0.60	-	0.82	1.05	0.78	0.51	1.06	0.59	0.49	0.49	0.50	0.54	1.02	-	-	-	1.11	0.48	1.47	0.53	1.74	1.53	-	0.47
KIR3DL3*017	1.14	-	-	-	-	-	-	0.53	1.56	1.01	1.60	-	-	-	-	-	-	3.13	2.86	2.06	5.56	3.81	2.45	-	1.16	0.51	2.02	1.63
KIR3DL3*018	0.02	-	-	-	-	-	-	-	-	-	0.53	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
KIR3DL3*019	0.43	-	-	-	-	-	-	0.53	1.56	0.51	-	0.59	-	-	-	-	-	-	1.90	1.03	2.22	1.43	-	-	0.58	-	-	0.93
KIR3DL3*020	0.73	0.97	2.04	2.38	0.60	0.53	-	-	0.78	0.51	1.06	0.59	-	-	-	-	-	1.04	0.48	0.52	-	0.95	-	4.26	0.58	1.02	0.51	0.70
KIR3DL3*021	0.19	-	-	-	-	-	-	-	-	1.01	0.53	-	-	-	-	-	-	-	0.48	-	-	0.48	-	-	-	0.51	0.51	0.70
KIR3DL3*022	0.41	1.46	2.04	0.95	1.19	1.06	1.64	2.11	-	1.01	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.23
KIR3DL3*023	0.02	-	-	-	-	-	-	-	-	-	-	-	0.49	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
KIR3DL3*025	0.94	0.97	0.51	3.81	4.17	1.60	4.10	0.53	-	-	1.06	0.59	-	-	-	-	-	0.52	0.95	4.12	2.22	0.48	-	-	-	-	-	0.93
KIR3DL3*026	0.28	-	-	-	-	-	-	-	-	0.51	-	-	1.46	0.49	0.50	0.54	-	-	0.48	-	-	-	1.47	1.06	0.58	0.51	-	-
KIR3DL3*027	2.08	4.85	6.63	5.24	3.57	4.79	3.28	4.21	2.34	1.52	3.72	1.18	-	-	-	-	-	1.56	1.43	-	-	3.33	1.96	2.13	0.58	1.02	1.01	2.09
KIR3DL3*028	2.16	0.97	-	-	0.60	1.06	-	1.05	2.34	-	1.60	1.76	5.34	3.43	4.00	5.38	1.02	0.52	0.48	2.06	0.56	0.48	6.37	5.85	6.40	3.06	4.55	0.70
KIR3DL3*030	0.23	-	1.02	-	-	-	-	0.53	0.78	-	1.06	-	-	-	-	-	-	0.52	0.48	-	-	0.95	-	0.53	-	-	-	0.23
KIR3DL3*031	0.09	-	-	-	-	-	-	0.53	-	-	-	-	-	-	-	-	-	-	-	-	1.11	0.48	-	-	-	-	-	0.23
KIR3DL3*032	0.62	2.43	0.51	3.33	3.57	3.19	2.46	1.05	0.78	-	0.53	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.23
KIR3DL3*034	0.45	1.46	2.04	1.90	2.38	-	0.82	3.68	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.23
KIR3DL3*035	1.01	5.34	1.53	6.19	5.36	3.19	-	3.68	0.78	1.01	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.47
KIR3DL3*036	0.02	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.56	-	-	-	-	-	-	-
KIR3DL3*037	0.02	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.49	-	-	-	-	-
KIR3DL3*041	0.83	0.49	3.57	-	1.19	-	1.64	0.53	0.78	-	1.06	1.18	-	-	-	-	-	2.08	1.43	2.06	0.56	-	3.43	0.53	0.58	0.51	1.01	0.47
KIR3DL3*044	0.09	-	-	-	-	-	-	-	-	-	0.53	-	-	-	-	-	-	1.04	-	-	-	0.48	-	-	-	-	-	0.23
KIR3DL3*048	0.60	-	-	-	-	-	-	-	-	0.51	0.53	1.18	1.94	3.92	1.00	1.61	1.53	-	-	1.03	-	-	0.49	-	1.16	-	1.01	0.23
KIR3DL3*049	0.23	0.49	0.51	1.43	0.60	-	1.64	1.05	-	-	-	-	-	-	-	-	-	-	-	-	0.56	-	-	-	-	-	-	0.23
KIR3DL3*050	0.54	2.91	3.06	0.48	3.57	1.60	2.46	1.05	0.78	0.51	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
KIR3DL3*052	0.13	0.49	-	1.43	0.60	-	0.82	0.53	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
KIR3DL3*053	0.09	0.49	-	0.48	-	-	0.82	-	0.78	0.51	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
KIR3DL3*054	0.04	-	-	-	1.19	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
KIR3DL3*056	0.15	-	1.02	-	-	-	0.82	1.05	-	0.51	0.53	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.23
KIR3DL3*057	0.04	-	0.51	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.56	-	-	-	-	-	-	-
KIR3DL3*062	0.04	-	-	-	-	-	-	-	-	-	-	-	0.49	-	-	-	0.51	-	-	-	-	-	-	-	-	-	-	-
KIR3DL3*063	0.06	-	-	-	-	-	-	-	-	-	-	-	-	-	1.00	-	0.51	-	-	-	-	-	-	-	-	-	-	-
KIR3DL3*064	0.17	-	-	-	-	-	-	-	-																			

For a description of each of these population samples from the 1000Genomes dataset, please refer to <https://www.internationalgenome.org/category/population/>
Proteins were named according to the IPD-IMGT/KIR database, version 2.9.0

Table S3. *KIR3DL3* encoded proteins in worldwide population samples, and each amino acid exchange. Here are listed only the proteins that do not match any of the ones described in the IPD-IMGT/KIR database.

KIR3DL3 proteins	Amino acid positions starting from the first translated Methionine																																								Global frequency					
	LP	LP	D0	D0	D0	D0	D0	D0	D0	D0	D1	D1	D1	D1	D1	D1	D1	D1	D2	D2	D2	D2	D2	D2	D2	D2	D2	D2	D2	D2	D2	TM	TM	TM	TM	TM	CY	CY	CY	CY		CY	CY			
	6	13	52	56	77	91	100	107	109	117	132	134	139	149	152	166	168	181	219	221	231	242	246	250	259	290	291	293	294	301	306	313	321	324	336	346	348	353	373	376	382	385	388	395		
unknown_01	V	F	H	N	R	G	P	S	P	V	P	V	T	R	R	R	V	S	G	H	G	L	S	F	A	N	Y	C	F	P	D	V	H	V	A	W	A	A	E	P	A	N	V	T	0.0002	
unknown_02	V	F	R	N	W	G	P	S	P	V	P	V	T	R	R	R	I	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	H	V	A	W	P	A	D	P	A	N	V	T	0.0004	
unknown_03	V	F	R	N	R	G	P	S	P	V	L	V	T	R	R	R	I	S	G	Y	G	L	S	F	A	N	Y	Y	F	P	D	V	Y	A	A	W	A	A	E	P	A	N	V	T	0.0002	
unknown_04	V	F	R	N	R	G	P	S	P	V	P	V	T	R	H	R	V	S	G	Y	G	L	S	F	A	N	D	C	F	P	D	V	Y	A	A	W	A	A	E	P	A	N	V	T	0.0002	
unknown_05	V	F	R	N	R	G	P	S	P	V	P	V	T	R	R	R	V	S	G	Y	D	L	S	F	A	N	Y	C	F	P	D	V	H	V	A	W	P	A	D	P	A	S	V	T	0.0002	
unknown_06	V	F	R	N	R	G	P	S	P	V	P	V	T	R	R	R	V	S	G	Y	G	L	S	L	A	N	Y	C	F	P	D	V	Y	A	A	W	A	A	E	P	A	N	V	T	0.0002	
unknown_07	V	F	R	N	R	G	P	S	P	V	P	V	T	R	R	C	V	S	G	Y	G	L	S	F	A	T	Y	C	F	P	D	V	H	V	A	W	A	A	E	P	A	N	V	T	0.0002	
unknown_08	V	F	R	N	W	G	P	S	P	V	P	V	T	S	R	R	V	S	G	Y	G	L	S	F	A	T	Y	C	F	P	D	V	H	V	A	W	A	A	E	P	A	N	V	T	0.0002	
unknown_09	V	F	H	N	R	G	P	S	P	V	P	V	T	R	R	R	V	S	G	Y	G	L	S	F	A	T	Y	C	F	P	D	V	Y	A	A	W	A	A	E	P	A	N	V	T	0.0004	
unknown_10	V	F	R	N	R	G	P	S	L	V	P	V	T	R	R	R	V	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	N	V	A	W	A	A	E	P	A	N	V	T	0.0002	
unknown_11	V	F	R	N	R	G	P	S	P	V	P	V	T	R	R	C	V	S	G	Y	G	L	S	F	A	T	Y	C	F	P	D	V	H	V	A	W	A	A	E	P	A	N	I	T	0.0002	
unknown_12	V	F	R	N	R	G	P	S	P	V	P	V	T	R	R	R	V	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	Y	A	A	W	A	A	E	P	A	N	I	T	0.0004	
unknown_13	V	F	R	K	R	G	P	S	P	V	P	V	T	S	R	R	V	S	G	Y	G	L	S	F	T	T	Y	C	F	P	D	V	H	V	A	W	A	A	E	P	A	N	V	T	0.0006	
unknown_14	V	F	H	N	W	A	P	S	P	V	P	V	T	R	R	R	V	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	H	V	A	W	A	A	E	P	A	N	V	T	0.0004	
unknown_15	V	F	R	N	R	G	P	S	P	V	P	V	T	R	R	R	V	S	G	Y	D	L	S	F	A	N	Y	C	F	P	D	V	H	V	P	W	P	A	D	P	A	N	V	T	0.0002	
unknown_16	V	F	R	N	W	G	P	S	P	V	P	V	T	R	R	R	V	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	N	V	A	W	A	A	E	P	A	N	V	T	0.0002	
unknown_17	V	F	R	N	R	G	P	S	P	V	P	V	T	R	R	R	I	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	Y	A	A	W	A	A	E	P	T	N	V	T	0.0002	
unknown_18	V	F	R	N	W	G	P	S	P	V	P	V	T	S	R	R	V	S	G	Y	G	L	S	F	A	T	Y	C	F	P	D	V	N	V	A	A	W	A	A	E	P	A	N	V	T	0.0002
unknown_19	V	F	R	N	R	G	P	S	P	V	P	V	T	R	R	R	I	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	N	V	A	C	A	A	E	P	A	N	V	T	0.0002	
unknown_20	I	F	R	N	R	G	P	S	P	V	P	V	T	R	R	R	V	S	G	Y	G	L	S	F	A	T	Y	C	F	P	D	V	H	V	A	W	A	A	E	P	A	N	V	T	0.0002	
unknown_21	V	F	R	N	R	G	P	S	P	V	P	V	T	R	H	R	V	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	Y	A	A	W	A	A	E	P	T	N	V	T	0.0002	
unknown_22	V	F	R	N	R	G	P	S	P	V	P	V	T	R	R	R	I	S	G	Y	G	V	S	F	A	N	Y	C	F	P	D	V	H	V	A	W	A	A	E	P	A	N	V	T	0.0008	
unknown_23	V	F	R	N	R	G	P	S	P	V	L	V	T	R	R	R	I	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	H	V	A	W	P	A	D	P	A	N	V	T	0.0002	
unknown_24	V	F	R	N	W	G	P	S	P	V	P	V	T	R	R	R	I	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	Y	A	A	W	A	A	E	P	T	N	V	T	0.0004	
unknown_25	V	F	H	N	R	G	P	S	P	F	P	V	T	R	R	R	V	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	Y	A	A	W	A	A	E	P	A	N	V	T	0.0004	
unknown_26	V	F	R	N	R	G	P	S	P	V	P	V	T	R	R	R	I	S	G	Y	G	L	S	F	A	N	Y	C	F	R	D	V	Y	A	A	W	A	A	E	P	A	N	V	T	0.0002	
unknown_27	V	F	R	N	R	G	P	S	P	V	L	V	T	R	R	R	I	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	H	V	A	W	A	A	E	P	A	N	V	T	0.0004	
unknown_28	V	F	H	N	R	G	P	S	P	V	P	V	T	R	R	R	I	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	Y	A	A	W	A	A	E	P	A	N	V	T	0.0004	
unknown_29	V	F	R	N	W	G	P	S	P	V	P	V	T	R	R	R	V	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	H	V	A	W	A	A	E	P	A	N	V	T	0.0021	
unknown_30	V	F	R	N	R	G	P	S	P	V	P	V	T	R	R	R	I	S	D	Y	G	L	S	F	A	N	Y	C	F	P	D	V	Y	A	A	W	A	A	E	P	A	N	V	T	0.0002	
unknown_31	V	-	R	N	R	G	P	S	P	V	P	V	T	R	R	R	I	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	Y	A	A	W	A	A	E	P	A	N	V	T	0.0004	
unknown_32	V	F	R	N	R	G	P	S	P	V	P	V	T	R	R	R	I	S	G	Y	G	L	S	F	A	N	Y	C	Y	P	D	V	H	V	A	W	A	A	E	P	A	N	V	T	0.0002	
unknown_33	V	F	R	N	R	G	P	S	P	V	P	V	T	R	R	R	V	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	Y	A	A	W	A	V	E	P	A	N	V	T	0.0002	
unknown_34	V	F	R	N	R	G	T	S	P	V	P	V	T	R	R	R	V	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	H	V	A	W	A	A	E	P	A	N	V	T	0.0002	
unknown_35	I	F	R	N	R	G	P	S	P	V	P	V	T	R	R	R	V	S	G	Y	G	L	S	F	A	T	Y	C	F	P	D	V	Y	A	A	W	A	A	E	P	A	N	V	T	0.0002	
unknown_36	V	F	R	N	R	G	P	S	P	V	P	V	T	R	R	C	I	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	N	V	A	W	A	A	E	P	A	N	V	T	0.0002	
unknown_37	V	F	R	N	W	G	P	S	P	V	P	V	M	R	R	R	I	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	H	V	A	W	A	A	E	P	A	N	V	T	0.0002	

unknown_38	V	F	R	N	R	G	P	S	P	V	P	V	T	R	R	C	V	S	G	Y	G	L	S	F	A	T	Y	C	F	P	D	V	Y	A	A	W	A	A	E	P	A	N	V	T	0.0002
unknown_39	V	F	R	N	R	G	P	S	P	V	P	V	T	R	R	R	I	S	G	Y	G	L	F	F	A	N	Y	C	F	P	D	V	N	V	A	W	A	A	E	P	A	N	V	T	0.0002
unknown_40	V	F	R	K	R	G	P	S	P	V	P	V	T	S	R	R	V	S	G	Y	G	L	S	F	A	T	Y	C	F	P	H	V	N	V	A	W	A	A	E	P	A	N	V	T	0.0008
unknown_41	V	F	R	N	R	G	P	S	P	V	P	V	T	R	R	R	I	F	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	Y	A	A	W	A	A	E	P	A	N	V	T	0.0002
unknown_42	V	F	R	K	R	G	P	S	P	V	P	V	T	R	R	R	V	S	G	Y	G	L	S	F	A	T	Y	C	F	P	D	V	H	V	A	W	P	A	D	P	A	N	V	T	0.0002
unknown_43	V	F	R	N	R	G	P	P	P	V	P	V	T	R	R	R	I	S	G	Y	G	L	S	F	A	N	Y	C	F	P	D	V	Y	A	A	W	A	A	E	P	A	N	V	T	0.0021
unknown_44	V	F	R	N	R	G	P	S	P	V	P	V	T	R	R	R	V	S	G	Y	G	L	S	F	A	T	Y	C	F	P	D	V	Y	A	A	W	A	A	E	H	A	N	V	T	0.0002

The proteins listed here have a summed frequency of 1.58%. The remaining 98.48% corresponds to KIR3DL3 proteins with defined names as listed by the IPD-IMGT/KIR database, listed in Table 02. Some variants listed in this table do not present an alternative amino acid because these positions are related to the proteins described in Table 2.

Table S4. *KIR3DL3* coding sequences (only exons) in worldwide population samples grouped by the ancestry background.

KIR3DL3 CDS sequence	Global Frequency (2n=5330)	European samples (2n=986)	African samples (2n=968)	EastAsian samples (2n=992)	SouthAsian samples (2n=958)	Admixed samples* (2n=996)	Brazilian samples (2n=430)
KIR3DL3*00101	0.0662	0.1410	0.0041	0.0131	0.1180	0.0432	0.0953
KIR3DL3*00102	0.0141	0.0264	-	0.0161	0.0177	0.0100	0.0140
KIR3DL3*00103	0.0043	0.0112	-	-	0.0073	0.0030	0.0047
KIR3DL3*00104	0.0008	0.0041	-	-	-	-	-
KIR3DL3*00201	0.0038	0.0132	0.0010	-	0.0010	0.0040	0.0023
KIR3DL3*00202	0.0206	0.0223	0.0093	0.0060	0.0355	0.0261	0.0302
KIR3DL3*00205	0.0111	-	0.0434	-	-	0.0151	0.0047
KIR3DL3*00206	0.0184	0.0619	0.0021	-	0.0052	0.0211	0.0209
KIR3DL3*00207	0.0118	0.0152	0.0072	-	0.0219	0.0151	0.0116
KIR3DL3*00208	0.0030	-	0.0124	-	-	0.0020	0.0047
KIR3DL3*00209	0.0008	0.0020	-	-	0.0021	-	-
KIR3DL3*00210	0.0030	0.0010	-	-	0.0146	0.0010	-
KIR3DL3*00301	0.1195	0.1552	0.0486	0.0252	0.2390	0.1024	0.1884
KIR3DL3*00302	0.0002	-	-	-	-	0.0010	-
KIR3DL3*00402	0.0383	0.0172	0.0455	0.0544	0.0251	0.0552	0.0233
KIR3DL3*005	0.0289	0.0010	0.1002	-	-	0.0412	0.0140
KIR3DL3*00601	0.0184	0.0172	-	0.0474	0.0146	0.0100	0.0233
KIR3DL3*00602	0.0024	0.0071	-	-	0.0031	0.0020	0.0023
KIR3DL3*007	0.0103	0.0203	-	-	0.0209	0.0090	0.0140
KIR3DL3*00801	0.0178	0.0193	0.0062	0.0343	0.0136	0.0181	0.0116
KIR3DL3*00802	0.0246	0.0020	-	0.0726	0.0031	0.0472	0.0163
KIR3DL3*00901	0.1096	0.0963	0.1560	0.1099	0.0720	0.1165	0.1023
KIR3DL3*00902	0.0432	0.0193	0.0041	0.0736	0.0313	0.0924	0.0279
KIR3DL3*00903	0.0017	-	0.0072	-	-	0.0010	0.0023
KIR3DL3*00904	0.0004	-	-	-	-	0.0010	0.0023
KIR3DL3*00905	0.0024	-	0.0093	-	-	0.0020	0.0047
KIR3DL3*01001	0.0591	0.0162	0.0227	0.2298	0.0292	0.0141	0.0163
KIR3DL3*01002	0.0398	0.0020	-	0.1623	0.0052	0.0382	0.0140
KIR3DL3*01003	0.0002	-	-	0.0010	-	-	-
KIR3DL3*01101	0.0090	0.0010	0.0362	-	-	0.0100	0.0047
KIR3DL3*01102	0.0122	0.0061	0.0279	-	0.0198	0.0080	0.0116
KIR3DL3*01103	0.0002	-	0.0010	-	-	-	-
KIR3DL3*01105	0.0002	-	-	-	-	-	0.0023
KIR3DL3*012	0.0103	-	0.0434	-	-	0.0120	0.0023
KIR3DL3*01302	0.0240	0.0487	-	-	0.0376	0.0221	0.0512
KIR3DL3*01303	0.0028	0.0101	-	-	-	0.0010	0.0093
KIR3DL3*01305	0.0013	0.0041	-	-	-	0.0020	0.0023
KIR3DL3*01306	0.0017	0.0081	-	-	-	-	0.0023
KIR3DL3*01307	0.0030	0.0061	-	-	0.0031	0.0050	0.0047
KIR3DL3*01308	0.0006	0.0030	-	-	-	-	-
KIR3DL3*01309	0.0002	0.0030	0.0010	-	0.0021	0.0020	0.0070
KIR3DL3*01401	0.0009	-	0.0031	-	-	0.0010	0.0023
KIR3DL3*01402	0.0334	0.0456	0.0300	0.0030	0.0438	0.0301	0.0674
KIR3DL3*01403	0.0083	0.0122	0.0103	-	0.0031	0.0100	0.0209
KIR3DL3*01404	0.0030	0.0030	0.0072	-	-	0.0040	0.0047
KIR3DL3*01405	0.0062	0.0132	0.0010	-	0.0136	0.0060	-
KIR3DL3*01406	0.0143	-	0.0610	-	-	0.0120	0.0116
KIR3DL3*01501	0.0148	0.0203	-	0.0171	0.0240	0.0080	0.0256
KIR3DL3*01502	0.0071	-	0.0227	0.0091	0.0010	0.0050	0.0023
KIR3DL3*01503	0.0021	0.0010	-	0.0091	0.0010	-	-
KIR3DL3*01601	0.0047	0.0030	-	0.0040	0.0104	0.0060	0.0047
KIR3DL3*01602	0.0017	-	0.0052	0.0020	-	0.0020	-
KIR3DL3*01701	0.0096	0.0274	-	-	0.0125	0.0060	0.0140
KIR3DL3*01702	0.0017	0.0061	-	-	-	0.0020	0.0023
KIR3DL3*01801	0.0002	-	-	-	-	0.0010	-
KIR3DL3*019	0.0043	0.0132	-	-	0.0010	0.0050	0.0093
KIR3DL3*020	0.0060	0.0061	0.0072	-	0.0125	0.0050	0.0047
KIR3DL3*02101	0.0019	0.0020	-	-	0.0021	0.0030	0.0070
KIR3DL3*022	0.0041	-	0.0134	-	-	0.0080	0.0023
KIR3DL3*023	0.0002	-	-	0.0010	-	-	-

KIR3DL3*02501	0.0041	0.0162	-	-	-	0.0040	0.0047
KIR3DL3*02502	0.0053	-	0.0217	-	-	0.0050	0.0047
KIR3DL3*02601	0.0013	0.0010	-	0.0010	0.0042	0.0010	-
KIR3DL3*02602	0.0015	-	-	0.0050	0.0031	-	-
KIR3DL3*02701	0.0099	0.0122	0.0124	-	0.0073	0.0171	0.0116
KIR3DL3*02702	0.0077	-	0.0320	-	-	0.0060	0.0093
KIR3DL3*02703	0.0004	0.0010	-	-	-	0.0010	-
KIR3DL3*02705	0.0002	-	0.0010	-	-	-	-
KIR3DL3*028	0.0210	0.0071	0.0052	0.0383	0.0501	0.0110	0.0070
KIR3DL3*030	0.0023	0.0041	0.0021	-	0.0010	0.0040	0.0023
KIR3DL3*031	0.0009	0.0030	-	-	-	0.0010	0.0023
KIR3DL3*032	0.0062	-	0.0258	-	-	0.0070	0.0023
KIR3DL3*034	0.0045	-	0.0155	-	-	0.0080	0.0023
KIR3DL3*035	0.0101	-	0.0434	-	-	0.0100	0.0047
KIR3DL3*036	0.0002	0.0010	-	-	-	-	-
KIR3DL3*037	0.0002	-	-	-	0.0010	-	-
KIR3DL3*041	0.0083	0.0122	0.0103	-	0.0125	0.0080	0.0047
KIR3DL3*044	0.0009	0.0030	-	-	-	0.0010	0.0023
KIR3DL3*04801	0.0026	-	-	0.0040	0.0052	0.0040	0.0023
KIR3DL3*04802	0.0034	0.0020	-	0.0161	-	-	-
KIR3DL3*04901	0.0023	0.0010	0.0062	-	-	0.0040	0.0023
KIR3DL3*050	0.0054	-	0.0227	-	-	0.0070	-
KIR3DL3*052	0.0013	-	0.0052	-	-	0.0020	-
KIR3DL3*053	0.0009	-	0.0021	-	-	0.0030	-
KIR3DL3*054	0.0004	-	0.0021	-	-	-	-
KIR3DL3*056	0.0015	-	0.0021	-	-	0.0050	0.0023
KIR3DL3*057	0.0004	0.0010	0.0010	-	-	-	-
KIR3DL3*062	0.0004	-	-	0.0020	-	-	-
KIR3DL3*063	0.0006	-	-	0.0030	-	-	-
KIR3DL3*064	0.0017	-	-	0.0050	0.0042	-	-
KIR3DL3*065	0.0011	-	-	0.0050	0.0010	-	-
KIR3DL3*076	0.0002	0.0010	-	-	-	-	-
New sequences	0.0285	0.0193	0.0393	0.0292	0.0418	0.0221	0.0070

European samples: CEU, TSI, FIN, GBR, IBS; African samples: YRI, LWK, GWD, MSL, ESN, ASW, ACB; East Asian samples: CHB, JPT, CHS, CDX, KHV; South Asian samples: GIH, PJL, BEB, STU, ITU; Admixed samples *: MXL, PUR, CLM, PEL. For a description of each of these population samples from the 1000Genomes dataset, please refer to <https://www.internationalgenome.org/category/population/>
CDS sequences were named according to the IPD-IMGT/KIR database, version 2.9.0

Table S5. *KIR3DL3* promoter haplotypes and their frequencies.

Haplotype	EUR	AFR	EAS	SAS	AMR	BRA
P01:01	0.4848	0.2552	0.2298	0.4374	0.3855	0.4444
P01:02	0.0619	0.0661	0.1714	0.0574	0.1335	-
P01:03	0.0629	0.1002	0.1008	0.0365	0.0653	0.0764
P01:04	0.0335	0.0878	-	0.0647	0.0462	0.1111
P01:05	0.0467	0.0248	-	0.0449	0.0201	0.0370
P01:06	-	-	0.0010	-	-	-
P01:07	-	0.0785	-	-	0.0161	0.0162
P01:08	0.0010	-	-	0.0543	-	-
P01:09	-	0.0351	-	-	0.0100	0.0069
P01:10	-	0.0310	-	-	0.0110	0.0023
P01:11	-	-	-	-	-	-
P01:12	-	-	-	-	-	-
P01:13	0.0020	-	-	0.0084	-	-
P01:14	-	-	0.0101	-	-	-
P01:15	-	-	-	0.0084	-	-
P01:16	-	0.0052	-	-	0.0020	-
P01:17	0.0010	-	-	0.0052	-	-
P01:18	-	0.0031	-	-	0.0020	-
P01:19	-	-	-	-	-	-
P01:20	-	-	-	0.0052	-	-
P01:21	-	-	-	-	-	0.0093
P01:22	-	-	-	0.0021	0.0020	-
P01:23	-	-	0.0010	0.0021	-	-
P01:24	-	-	0.0020	0.0010	-	-
P01:25	-	-	0.0030	-	-	-
P01:26	-	-	-	0.0021	-	-
P01:27	-	-	-	0.0021	-	-
P01:28	-	0.0010	-	-	0.0010	-
P01:29	-	-	-	-	0.0010	-
P01:30	-	0.0010	-	-	-	-
P01:31	-	-	-	-	-	0.0023
P01:32	-	-	0.0010	-	-	-
P01:33	-	0.0010	-	-	-	-
P01:34	-	0.0010	-	-	-	-
P01:35	-	-	-	0.0010	-	-
P01:36	-	-	-	-	0.0010	-
P01:37	-	-	-	-	0.0010	-
P01:38	-	0.0010	-	-	-	-
P02:01	0.0274	0.2056	0.4567	0.0919	0.1345	0.0625
P02:02	0.2748	0.0950	0.0091	0.1722	0.1556	0.2245
P02:03	-	-	-	-	-	-
P02:04	-	-	0.0010	-	-	-
P02:05	0.0030	-	-	-	0.0030	0.0046
P02:06	-	-	0.0071	-	-	-
P02:07	-	0.0010	-	-	0.0020	-
P02:08	-	0.0010	-	-	0.0020	-
P02:09	-	-	0.0020	-	-	-
P02:10	-	0.0021	-	-	-	-
P02:11	-	-	-	-	0.0020	-
P02:12	-	-	0.0020	-	-	-
P02:13	-	0.0021	-	-	-	-
P02:14	-	-	0.0020	-	-	-
P02:15	0.0010	-	-	-	-	-
P02:16	-	-	-	-	0.0010	-
P02:17	-	0.0010	-	-	-	-
P02:18	-	-	-	0.0010	-	-
P02:19	-	-	-	-	-	0.0023
P02:20	-	-	-	0.0010	-	-
P02:21	-	-	-	-	0.0010	-
P02:22	-	-	-	0.0010	-	-
P02:23	-	-	-	-	0.0010	-

The structure of each of these haplotypes is available at Figure S2.

Table S6. The association between *KIR3DL3* promoter lineages and genomic sequences encoding specific proteins.

Promoter Lineage	KIR3DL3 protein	Global Frequency	
		Global Count	$2n = 5330$
P01	KIR3DL3*001	456	0.0856
P02	KIR3DL3*001	2	0.0004
P01	KIR3DL3*002	119	0.0223
P02	KIR3DL3*002	270	0.0507
P01	KIR3DL3*003	644	0.1208
P02	KIR3DL3*003	1	0.0002
P01	KIR3DL3*004	184	0.0345
P02	KIR3DL3*004	20	0.0038
P01	KIR3DL3*005	1	0.0002
P02	KIR3DL3*005	144	0.0270
P01	KIR3DL3*006	111	0.0208
P02	KIR3DL3*007	55	0.0103
P01	KIR3DL3*008	215	0.0403
P02	KIR3DL3*008	12	0.0023
P01	KIR3DL3*009	831	0.1559
P02	KIR3DL3*009	10	0.0019
P01	KIR3DL3*010	1	0.0002
P02	KIR3DL3*010	529	0.0992
P01	KIR3DL3*011	115	0.0216
P01	KIR3DL3*012	55	0.0103
P01	KIR3DL3*013	30	0.0056
P02	KIR3DL3*013	160	0.0300
P01	KIR3DL3*014	166	0.0311
P02	KIR3DL3*014	203	0.0381
P01	KIR3DL3*015	134	0.0251
P01	KIR3DL3*016	34	0.0064
P01	KIR3DL3*017	61	0.0114
P01	KIR3DL3*018	1	0.0002
P02	KIR3DL3*019	23	0.0043
P01	KIR3DL3*020	39	0.0073
P01	KIR3DL3*021	10	0.0019
P01	KIR3DL3*022	22	0.0041
P01	KIR3DL3*023	1	0.0002
P01	KIR3DL3*025	50	0.0094
P01	KIR3DL3*026	14	0.0026
P02	KIR3DL3*026	1	0.0002
P01	KIR3DL3*027	47	0.0088
P02	KIR3DL3*027	64	0.0120
P01	KIR3DL3*028	5	0.0009
P02	KIR3DL3*028	110	0.0206
P02	KIR3DL3*030	12	0.0023
P02	KIR3DL3*031	5	0.0009
P01	KIR3DL3*032	33	0.0062
P02	KIR3DL3*034	24	0.0045
P02	KIR3DL3*035	54	0.0101
P01	KIR3DL3*036	1	0.0002

P01	KIR3DL3*037	1	0.0002
P01	KIR3DL3*041	44	0.0083
P01	KIR3DL3*044	5	0.0009
P02	KIR3DL3*048	32	0.0060
P01	KIR3DL3*049	11	0.0021
P02	KIR3DL3*049	1	0.0002
P01	KIR3DL3*050	29	0.0054
P01	KIR3DL3*052	7	0.0013
P01	KIR3DL3*053	5	0.0009
P01	KIR3DL3*054	2	0.0004
P01	KIR3DL3*056	8	0.0015
P01	KIR3DL3*057	2	0.0004
P01	KIR3DL3*062	2	0.0004
P02	KIR3DL3*063	3	0.0006
P01	KIR3DL3*064	9	0.0017
P01	KIR3DL3*065	6	0.0011
P01	KIR3DL3*066	2	0.0004
P01	KIR3DL3*076	1	0.0002
P02	KIR3DL3*084	1	0.0002

Table S7. The association between *KIR3DL3* 3'UTR haplotypes and genomic sequences encoding specific proteins.

3'UTR haplotype	KIR3DL3 protein	Global count	Global frequency
3UTR-01	KIR3DL3*001	452	0.0848
3UTR-12	KIR3DL3*001	1	0.0002
3UTR-14	KIR3DL3*001	2	0.0004
3UTR-06	KIR3DL3*001	2	0.0004
3UTR-25	KIR3DL3*001	1	0.0002
3UTR-01	KIR3DL3*002	383	0.0719
3UTR-07	KIR3DL3*002	6	0.0011
3UTR-16	KIR3DL3*003	2	0.0004
3UTR-10	KIR3DL3*003	1	0.0002
3UTR-02	KIR3DL3*003	635	0.1191
3UTR-23	KIR3DL3*003	1	0.0002
3UTR-24	KIR3DL3*003	1	0.0002
3UTR-11	KIR3DL3*003	1	0.0002
3UTR-04	KIR3DL3*003	1	0.0002
3UTR-15	KIR3DL3*003	3	0.0006
3UTR-10	KIR3DL3*004	5	0.0009
3UTR-02	KIR3DL3*004	199	0.0373
3UTR-01	KIR3DL3*005	145	0.0272
3UTR-09	KIR3DL3*006	2	0.0004
3UTR-01	KIR3DL3*006	89	0.0167
3UTR-06	KIR3DL3*006	1	0.0002
3UTR-05	KIR3DL3*006	19	0.0036
3UTR-02	KIR3DL3*007	55	0.0103
3UTR-08	KIR3DL3*008	1	0.0002
3UTR-01	KIR3DL3*008	226	0.0424
3UTR-08	KIR3DL3*009	1	0.0002
3UTR-01	KIR3DL3*009	834	0.1565
3UTR-07	KIR3DL3*009	3	0.0006
3UTR-20	KIR3DL3*009	1	0.0002
3UTR-06	KIR3DL3*009	2	0.0004
3UTR-08	KIR3DL3*010	7	0.0013

3UTR-13	KIR3DL3*010	4	0.0008
3UTR-01	KIR3DL3*010	519	0.0974
3UTR-01	KIR3DL3*011	115	0.0216
3UTR-01	KIR3DL3*012	55	0.0103
3UTR-01	KIR3DL3*013	183	0.0343
3UTR-12	KIR3DL3*013	4	0.0008
3UTR-14	KIR3DL3*013	1	0.0002
3UTR-06	KIR3DL3*013	2	0.0004
3UTR-19	KIR3DL3*014	1	0.0002
3UTR-02	KIR3DL3*014	360	0.0675
3UTR-11	KIR3DL3*014	5	0.0009
3UTR-04	KIR3DL3*014	3	0.0006
3UTR-01	KIR3DL3*015	134	0.0251
3UTR-02	KIR3DL3*016	33	0.0062
3UTR-04	KIR3DL3*016	1	0.0002
3UTR-01	KIR3DL3*017	58	0.0109
3UTR-06	KIR3DL3*017	3	0.0006
3UTR-01	KIR3DL3*018	1	0.0002
3UTR-01	KIR3DL3*019	23	0.0043
3UTR-02	KIR3DL3*020	39	0.0073
3UTR-01	KIR3DL3*021	10	0.0019
3UTR-02	KIR3DL3*022	4	0.0008
3UTR-04	KIR3DL3*022	18	0.0034
3UTR-01	KIR3DL3*023	1	0.0002
3UTR-03	KIR3DL3*025	50	0.0094
3UTR-01	KIR3DL3*026	15	0.0028
3UTR-03	KIR3DL3*027	111	0.0208
3UTR-18	KIR3DL3*028	1	0.0002
3UTR-02	KIR3DL3*028	114	0.0214
3UTR-01	KIR3DL3*030	12	0.0023
3UTR-03	KIR3DL3*031	5	0.0009
3UTR-17	KIR3DL3*032	1	0.0002
3UTR-02	KIR3DL3*032	32	0.0060
3UTR-02	KIR3DL3*034	21	0.0039
3UTR-04	KIR3DL3*034	3	0.0006
3UTR-03	KIR3DL3*035	54	0.0101
3UTR-02	KIR3DL3*036	1	0.0002
3UTR-02	KIR3DL3*037	1	0.0002
3UTR-03	KIR3DL3*041	44	0.0083
3UTR-01	KIR3DL3*044	5	0.0009
3UTR-01	KIR3DL3*048	31	0.0058
3UTR-05	KIR3DL3*048	1	0.0002
3UTR-03	KIR3DL3*049	12	0.0023
3UTR-02	KIR3DL3*050	29	0.0054
3UTR-01	KIR3DL3*052	7	0.0013
3UTR-01	KIR3DL3*053	5	0.0009
3UTR-01	KIR3DL3*054	2	0.0004
3UTR-03	KIR3DL3*056	8	0.0015
3UTR-01	KIR3DL3*057	2	0.0004
3UTR-01	KIR3DL3*062	2	0.0004
3UTR-01	KIR3DL3*063	3	0.0006
3UTR-01	KIR3DL3*064	9	0.0017
3UTR-01	KIR3DL3*065	6	0.0011
3UTR-01	KIR3DL3*066	2	0.0004
3UTR-02	KIR3DL3*076	1	0.0002
3UTR-03	KIR3DL3*084	1	0.0002

Conclusão

O gene *KIR3DL3* foi avaliado em 27 populações mundiais, incluindo populações europeias, africanas, asiáticas e algumas populações miscigenadas (incluindo Brasil). Essas populações são provenientes do banco de dados 1000 *genomes*, exceto a população brasileira. Para isso, desenvolvemos uma metodologia de sequenciamento de segunda geração para sequenciar os dados da população brasileira e uma estratégia computacional específica para trabalhar com esta família gênica. A partir disso podemos destacar alguns pontos em relação ao gene *KIR3DL3*:

1. Foram encontradas duas linhagens de promotores divergentes, com muitas variantes em absoluto LD, baixa diversidade no promotor proximal e evidencia de seleção balanceadora nos promotores principalmente em populações europeias e do leste asiático;
2. Observamos seleção purificadora no domínio D2 (éxon 5), alta diversidade de proteínas preditas e seleção balanceadora no domínio transmembrana (éxon 6);
3. O peptídeo líder e a região citoplasmática são muito conservadas em todas as populações estudadas;
4. Confirmamos a presença de um hotspot de recombinação no íntron 5 do gene *KIR3DL3*;
5. A diversidade genética desse gene é muito superior à já relatada em bancos de dados especializados;
6. A ancestralidade e região de origem do indivíduo está diretamente relacionada com as frequências de alguns alelos e proteínas preditas.