UNIVERSIDADE FEDERAL DO PARANÁ LUCIANA DE BRITO VARGAS

CONDUZINDO O ESTUDO DE *KIR* PARA A NOVA GERAÇÃO: CARACTERIZAÇÃO EM ALTA RESOLUÇÃO DE TODOS OS GENES *KIR* E LIGANTES *HLA* EM POPULAÇÕES SUL-AMERICANAS



UNIVERSIDADE FEDERAL DO PARANÁ LUCIANA DE BRITO VARGAS

CONDUZINDO O ESTUDO DE *KIR* PARA A NOVA GERAÇÃO: CARACTERIZAÇÃO EM ALTA RESOLUÇÃO DE TODOS OS GENES *KIR* E LIGANTES *HLA* EM POPULAÇÕES SUL-AMERICANAS

Dissertação de Mestrado apresentado ao Programa de Pós-Graduação em Genética, Setor de Ciências Biológicas, Universidade Federal do Paraná, como requisito à obtenção do título de mestre em Genética.

Orientador: Prof. Dr. Danillo G. Augusto.

Coorientadora: Profa. Dra. Marcia Holsbach Beltrame

CURITIBA 2020

Universidade Federal do Paraná. Sistema de Bibliotecas. Biblioteca de Ciências Biológicas. (Rosilei Vilas Boas – CRB/9-939).

Vargas, Luciana de Brito.

Conduzindo o estudo de KIR para a nova geração: caracterização em alta resolução de todos os genes KIR e ligantes HLA em populações sulamericanas. / Luciana de Brito Vargas. – Curitiba, 2020. 119 f. : il.

Orientador: Danillo G. Augusto. Coorientadora: Marcia Holsbach Beltrame.

Dissertação (Mestrado) – Universidade Federal do Paraná, Setor de Ciências Biológicas. Programa de Pós-Graduação em Genética.

1. Células killer. 2. Seleção natural. 3. Coevolução. 4. Expressão gênica. 5. Genética da população humana. I. Título. II. Augusto, Danillo G. III. Beltrame, Marcia Holsbach. IV. Universidade Federal do Paraná. Setor de Ciências Biológicas. Programa de Pós-Graduação em Genética.

CDD (20. ed.) 575



MINISTÉRIO DA EDUCAÇÃO SETOR DE CIENCIAS BIOLOGICAS UNIVERSIDADE FEDERAL DO PARANÁ PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO PROGRAMA DE PÓS-GRADUAÇÃO GENÉTICA -40001016006P1

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em GENÉTICA da Universidade Federal do Paraná foram convocados para realizar a arguição da dissertação de Mestrado de LUCIANA DE BRITO VARGAS intitulada: Conduzindo o estudo de KIR para a nova geração: caracterização em alta resolução de todos os genes KIR e ligantes HLA em populações sul-americanas, sob orientação do Prof. Dr. DANILLO GARDENAL AUGUSTO, que após terem inquirido a aluna e realizada a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

A outorga do título de mestre está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

CURITIBA, 30 de Março de 2020.

Assinatura Eletrônica 08/04/2020 17:34:03.0 DANILLO GARDENAL AUGUSTO Presidente da Banca Examinadora

Assinatura Eletrônica 27/04/2020 13:25:13.0 DIOGO MEYER Avaliador Externo (UNIVERSIDADE DE SÃO PAULO) Assinatura Eletrônica 13/04/2020 10:18:08.0 LIANA ALVES DE OLIVEIRA Avaliador Externo (FACULDADES INTEGRADAS DO BRASIL)

Assinatura Eletrônica 04/04/2020 11:24:18.0 MARIA LUIZA PETZL ERLER Avaliador Interno (UNIVERSIDADE FEDERAL DO PARANÁ)

AGRADECIMENTOS

Agradeço à CAPES e ao Governo Brasileiro pelo financiamento e apoio para a realização desse projeto. Agradeço também à UFPR, meu lar pelos últimos oito anos, por propiciar os recursos financeiros, estruturais, e, principalmente, os recursos humanos para minha formação acadêmica. Ainda, agradeço a oportunidade de fazer parte do LGMH, de conviver e colaborar com os seus mentores, e de usufruir de seu ambiente intensamente colaborativo e inclusivo. Não fosse pelo trabalho de excelente qualidade dos seus membros ao longo dos anos, não seria possível desenvolver este projeto. Além disso, o grupo que formamos me proporcionou grandes amizades que levarei para a vida toda.

Dentre essas amizades, gostaria de agradecer principalmente ao meu orientador e amigo Danillo! Agradeço por aceitar nos orientar a distância nos últimos anos, e por continuar disposto a abrir portas e lutar por seus orientados. Principalmente, sou muito grata a você e ao Billy por terem me recebido na sua casa, por todos os cafés da minha, pizzas e cervejas, para que eu pudesse ter a experiência de levar esse trabalho desde à bancada até a bioinformática. Sou muito grata a você pela incrível experiência que tive durante o mestrado.

Agradeço a meus amigos de laboratório pelo apoio científico e psicológico durante esse período. À Verónica, que passou de minha primeira mentora à minha grande amiga. Ao Leonardo, meu amigo KIRólogo por responder todas as minhas mensagens pedindo ajuda e opiniões no meio da noite. À Hellen, à Luana e ao Willian por oferecerem apoio incondicional ao longo desse período, pela jarra sempre pronta de café, e pelas longas conversas sobre propósitos e sobre a vida. Novamente, agradeço à toda equipe do LGMH e ao trabalho de anos que permitiu a realização desse projeto.

Agradeço esse trabalho à minha coorientadora Márcia, por sua orientação cientifica, pelos interesses compartilhados que deram origem ao *Journal Club* de GenePop, pelos diversos livros emprestados e pelas conversas sobre carreira, ciência e humanidade. Levarei muitos dos seus ensinamentos para a vida!

Finalmente, agradeço aos meus pais por me incentivarem na carreira acadêmica e por me proporcionarem a oportunidade de dedicar diversos anos de estudo para fazer o que eu gosto. No país e no momento político em que estamos, a educação é a minha maior herança e privilégio. Agradeço especialmente ao meu pai, meu grande mentor e maior ouvinte.

"Só um sentido de invenção e uma necessidade intensa de criar levam o homem a revoltar-se, a descobrir e a descobrir-se com lucidez". (Pablo Picasso)

RESUMO

Os genes dos receptores semelhantes a imunoglobulina das células natural killer (KIR) codificam proteínas transmembrana que medeiam a função efetora das células NK (do inglês, natural killer) contra células próprias infectadas e neoplásicas, através da interação com antígenos leucocitários humanos (HLA, do inglês human leukocyte antigen) de classe I. Combinações de genes KIR e HLA já foram associadas a susceptibilidade diferencial a doenças autoimunes, infecciosas, câncer e também com o sucesso reprodutivo. O polimorfismo dos genes KIR é complexo e incomum, uma vez que determinados genes podem ou não estar presentes no cromossomo. A família KIR contém até 13 genes e 2 pseudogenes localizados no cromossomo 19g13.4, que, por sua vez, possuem grande diversidade alélica. A complexidade do estudo de KIR se amplifica pela homologia e alta similaridade de seguência entre os diversos genes. Como conseguência, poucos são os estudos que caracterizaram a diversidade alélica de KIR, especialmente em populações. Portanto, o objetivo desse trabalho foi utilizar um método pioneiro e inovador através de seguenciamento de nova geração (NGS) para caracterizar, em alta resolução, a diversidade alélica de todos os loci KIR e dos loci de seus ligantes HLA em oito populações sul-americanas (uma eurodescente, uma descendente de japoneses, e seis ameríndias dos grupos Guarani, Kaingang e Aché; n=707). Além disso, buscamos evidências de coevolução envolvendo o receptor KIR2DL1, que reconhece com alta afinidade moléculas HLA-C do grupo C2. O NGS foi realizado no aparelho HiSeq 4000 após o enriquecimento das bibliotecas a partir de um conjunto de mais de 20 mil sondas biotiniladas específicas para aos genes KIR e HLA. A filtragem, alinhamento e identificação de genótipos de seguências KIR foi feita através do algoritmo PING; e o mesmo foi feito para os genes HLA, que foram analisados utilizando o programa HLA Explore[™]. Para a análise de coevolução, variantes em KIR2DL1 foram relacionadas aos níveis de expressão in silico de HLA-C a partir de dados públicos do 1000 genomes e relacionada à expressão de KIR2DL1 em uma amostra de Curitiba, Paraná (n=48). Para a análise de expressão, foi coletado sangue periférico de indivíduos eurodescendentes que não relataram nenhuma doença aguda ou crônica, seguido da separação de linfócitos mononucleados (PBMC) e extração de DNA. A genotipagem alélica de KIR2DL1 foi feita através de sequenciamento Sanger e os níveis de expressão foram avaliados por citometria de fluxo. Foram encontradas nas populações desse estudo 209 variantes nos 13 genes KIR, organizadas em 37 haplótipos centroméricos e 44 haplótipos teloméricos. Os ameríndios apresentaram menor número de alelos, cerca de 58 por população, e menor diversidade de haplótipos, com até 9 haplótipos centroméricos e 12 haplótipos teloméricos. A diversidade de KIR mantida em ameríndios pode ser considerada a mínima já observada em populações humanas, e esta é principalmente concentrada nas interações com HLA-C (até 97,2%). Ainda, nós encontramos evidência de coevolução entre KIR e HLA ao observar uma variante em KIR2DL1, rs2304224T, que marca os níveis de expressão de KIR2DL1 (p = 0,0002) e HLA-C2 (p < 0,05). Esta variante está em deseguilíbrio de ligação com outras duas variantes, rs4806553G e rs687000G, que apresentam elevada homozigose compatível com seleção positiva. Os resultados desse estudo têm o potencial de contribuir para a compreensão da variação normal dos genes KIR e na compreensão das relações evolutivas entre as famílias gênicas KIR e HLA, bem como suas implicações na susceptibilidade diferencial a doenças.

Palavras-chave: Células NK. Seleção natural. Desequilíbrio de ligação. Coevolução. Expressão gênica. Genética de populações.

ABSTRACT

Killer-cell immunoglobulin-like receptor (KIR) genes encode surface transmembrane molecules that mediate NK (natural killer) cell function against neoplastic and infected cells. KIR interact with human leukocyte antigen (HLA) class I molecules on the surface of target cells. Combinations of KIR and HLA alleles have been associated with susceptibility to autoimmune diseases, viral infections, cancer and reproductive success. The KIR family comprises up to 13 genes and 2 pseudogenes located on chromosome 19q13.4 that exhibit an unusual and complex presence and absence polymorphism. In addition, each single locus exhibits extensive allelic diversity, which is difficult to analyze due to several technical difficulties imposed by high sequence similarity among KIR genes. Consequently, studies that characterize KIR alleles are scarce, especially among populations. Thus, the aim of this project is to use a pioneer and cutting-edge next generation sequencing (NGS) method to characterize, in high resolution, the allelic diversity of all KIR genes and HLA ligands in eight South American populations (European and Japanese descendants from Curitiba, and six Amerindian populations belonging to the Guarani, Kaingang and Aché groups, n=707). Besides describing the allelic and haplotype frequencies, we aim to search for evidences of coevolution involving KIR2DL1, which is the stronger receptor for C2 allotypes of HLA-C. NGS was performed in HiSeq 4000 after library enrichment with over 20,000 biotinylated probes to capture the KIR and HLA regions. Sequence filtering, alignment and genotype calling of KIR genes were performed using PING pipeline; for HLA genes sequence reads were analyzed using the software HLA Explore[™]. For coevolution analysis, we searched for *KIR2DL1* variants which could influence HLA-C expression levels in silico using public data from 1000 genomes, and verified if those also marked KIR2DL1 expression levels in a group of Euro descendants from Curitiba (n=48). In order to assess KIR2DL1 expression levels, we collected peripheral blood from euro-descendant individuals from Curitiba with no report of acute or chronic disease, followed by mononuclear cells (PBMC) isolation and DNA extraction. KIR2DL1 was sequenced using Sanger method and expression levels were measured through flow cytometry. In the 13 KIR loci analyzed, we found 209 variants organized in 13 centromeric and 44 telomeric haplotypes. Amerindians present reduced KIR diversity, averaging 58 KIR alleles per population, and a reduced diversity of haplotypes, with a maximum of 9 centromeric and 12 telomeric haplotypes. This reduced diversity in KIR maintains mainly HLA-C interactions (up to 97.2%). We also present evidence of coevolution between HLA-C and KIR2DL1 by showing that a KIR2DL1 variant, rs2304224T, marks KIR2DL1 expression (p = 0.002) and HLA-C expression (p < 0.05). This variant is in linkage disequilibrium with other two KIR2DL1 variants, rs4806553G and rs687000G, which present elevated homozygosity compatible with positive selection. The results of this study may provide the basis for the comprehension of KIR diversity and susceptibility to diseases, as well as insights on KIR-HLA coevolution.

Keywords: NK Cells. Natural Selection. Linkage disequilibrium. Coevolution. Gene expression. Population Genetics.

LISTA DE FIGURAS

FIGURA 1 - CONJUNTO DE RECEPTORES KIR, DESTACANDO SEUS DOMÍNIOS	S
EXTRACELULARES E MOTIVOS BASEADOS EM TIROSINA1	4
FIGURA 2 – ORGANIZAÇÃO DOS GENES <i>KIR</i> 1	7
FIGURA 3 - COMBINAÇÕES DE SEGMENTOS <i>KIR</i> CENTROMÉRICOS E <i>KIR</i>	
TELOMÉRICOS IDÊNTICOS1	9
FIGURA 4 - REPRESENTAÇÃO DE DIFERENTES HAPLÓTIPOS DE PRESENÇA	
DE GENES KIR, SEPARANDO REPRESENTANTES DOS	
HAPLOGRUPOS A E B2	20
FIGURA 5 - ORGANIZAÇÃO DO GENE <i>KIR2DL1</i> 2	22
FIGURA 6 – ESTRATÉGIA DE <i>GATES</i> PARA QUANTIFICAR A EXPRESSÃO DE	
KIR2DL1 POR CITOMETRIA3	\$5

LISTA DE TABELAS

TABELA 1 – INTERAÇÕES CONFIRMADAS ENTRE KIR E HLA (REVISADO F	POR
NEMAT-GORGANI et al., 2019)	15
TABELA 2 - CARACTERIZAÇÃO DAS OITO POPULAÇÕES INCLUÍDAS NO	
ESTUDO	30
TABELA 3 - OLIGONUCLEOTÍDEOS UTILIZADOS NESTE TRABALHO	32

LISTA DE SIGLAS

- CenA região centromérica A de KIR
- CenB região centromérica B de KIR
- dbSNP Database of Single Nucleotide Polymorphisms
- EHH Haplótipo de homozigose estendido
- Fst Fixation index do teste de diferenciação populacional
- HLA Antígeno Leucocitário Humano
- HUGO Organização do Genoma Humano (Human Genome Organization)
- Indel Polimorfismo de inserção e deleção
- KIR Receptores semelhantes a imunoglobulina das células NK
- LD Desequilíbrio de ligação
- LNA Ácido nucleico bloqueado (locked nucleid acid)
- LGMH Laboratório de Genética Molecular Humana
- p Valor de probabilidade
- PCA Análise de Componentes Principais
- PCR Reação em cadeia da Polimerase
- PING Pushing Immunogenetics to the Next Generation (software)
- SNP Polimorfismo de sítio único
- TelA região telomérica A de KIR
- TelB região telomérica B de KIR
- TFBS Sítio de ligação de fator de transcrição
- NK Células assassinas naturais (natural killer)
- ∆G Energia livre de Gibbs

SUMÁRIO

1 INTRODUÇÃO	8
2 REVISÃO DE LITERATURA	10
2.1 CÉLULAS NATURAL KILLER (NK)	10
2.2 RECEPTORES SEMELHANTES A IMUNOGLOBULINA DE CÉLULAS N	K (KIR).
	12
2.3 GENES KIR	15
2.4 GENE KIR2DL1	21
2.5 POPULAÇÕES HUMANAS E DIVERSIDADE GENÉTICA	22
2.5.1 POPULAÇÕES AMERÍNDIAS	23
2.5.2 JAPONESES NO BRASIL	25
2.6 JUSTIFICATIVA	26
3 OBJETIVOS	27
3.1 OBJETIVO GERAL	27
3.2 OBJETIVOS ESPECÍFICOS	27
4 MATERIAL E MÉTODOS	29
4.1 AMOSTRAS BIOLÓGICAS	29
4.2 GENOTIPAGEM DE PRESENÇA E AUSÊNCIA DO GENE KIR2DL1	30
4.3 IDENTIFICAÇÃO DOS ALELOS DE KIR2DL1	31
4.4 DETERMINAÇÃO DO NÚMERO DE CÓPIAS DE KIR2DL1	33
4.5 ENSAIO DE EXPRESSÃO GÊNICA DE KIR2DL1	33
4.6 GENOTIPAGEM DE GRUPOS ALÉLICOS DE HLA DE CLASSE I (-A, -B	E -C)36
4.7 SEQUENCIAMENTO DE NOVA GERAÇÃO DE TODOS OS GENES KIR	E DE
HLA DE CLASSE I	36
4.7.1 ANÁLISE DA QUALIDADE DO DNA	36
4.7.2 SEQUENCIAMENTO DE NOVA GERAÇÃO DOS GENES KIR E HLA	37
4.8 ANÁLISE DOS DADOS	
5 RESULTADOS	43
5.1 SINGLE NUCLEOTIDE POLYMORPHISM IN KIR2DL1 IS ASSOCIATED	WITH
HLA-C EXPRESSION IN GLOBAL POPULATIONS	43
5.2 KIR ALLELIC CHARACTERIZATION IN SOUTHEASTERN AMERINDIAN	IS AND
BRAZILIAN URBAN POPULATIONS	69

6 CONSIDERAÇÕES FINAIS	101
REFERÊNCIAS	

1 INTRODUÇÃO

As células *natural killer* (NK) são linfócitos grandes e granulares que participam da imunidade inata e adaptativa (KIESSLING et al., 1975; O'LEARY et al., 2006). Sua ação é mediada por receptores de membrana, que reconhecem principalmente os antígenos leucocitários humanos (HLA) de classe I na superfície de células alvo (PARHAM, 2004). Os genes da família *KIR* (*killer cell immunoglobulin-like receptors*) codificam a família de receptores semelhantes a imunoglobulina presentes na superfície das células NK e também NKT (*natural killer* T). A interação de KIR com as moléculas HLA da célula alvo podem gerar uma resposta de ativação ou inibição das células NK, modulando assim sua atividade citotóxica (PEGRAM et al., 2011).

Os genes *KIR* se localizam na região genômica 19q13.4 (WILSON et al., 2000) e apresentam uma complexa variação estrutural e de sequência. A homologia e alta similaridade de sequências resultam em muitos eventos de recombinação não recíproca entre diferentes *loci KIR*, causando a duplicação, deleção e até a formação de genes híbridos (MARTIN et al., 2003), caracterizando a rápida evolução dessa família gênica (VILCHES; PARHAM, 2002; TRAHERNE et al., 2010). Essas características impõem dificuldades técnicas para análise de*ss*a família, principalmente quando se leva em conta que cada gene *KIR* apresenta grande diversidade alélica (ROBINSON et al., 2015).

Tamanha a diversidade observada em KIR, é muito pouco provável encontrar dois indivíduos não relacionados geneticamente que possuam a mesma combinação de variantes *KIR*, quase tão improvável quanto encontrar indivíduos idênticos para os variantes *HLA*, que são os genes mais polimórficos do genoma humano (SHILLING et al., 2002). Além disso, os genes *KIR* apresentam uma expressão proteica variegada e aleatória (UHRBERG, 2005), o que significa que nem todas as células NK de um mesmo indivíduo apresentam as mesmas moléculas KIR na superfície celular. Para que o receptor expresso seja funcional, é ainda necessário que o ligante HLA estão localizados no cromossomo 6 (HORTON et al., 2004) e segregam independentemente dos genes *KIR*, além de poderem não apresentar certos genes *KIR*, os indivíduos podem possuir o gene mas não necessariamente o ligante HLA que torna o receptor funcional.

Essa complexa combinação de genótipos de *KIR* e de *HLA* já foi associada a doenças autoimunes, susceptibilidade a infecções virais e câncer (RAJAGOPALAN; LONG, 2005; KHAKOO; CARRINGTON, 2006; AUGUSTO, 2016). Combinações *KIR-HLA* estão também envolvidas no processo de placentação (HIBY et al., 2004; TROWSDALE; MOFFETT, 2008; XIONG et al., 2013) e mostraram importância no aumento da sobrevida após transplantes terapêuticos (COOLEY et al., 2010). Essas implicações evidenciam um papel chave dos receptores KIR e das células NK na imunidade e sobrevivência humana.

Devido a dificuldade do estudo de *KIR*, a maioria dos trabalhos até o momento se limitaram à análise de presença e ausência de genes (GUINAN et al., 2010). Quando estudado, o polimorfismo alélico é normalmente restrito a um ou poucos genes. Ainda, os poucos estudos da diversidade de *KIR* em alta resolução são concentrados em populações europeias (NORMAN et al., 2016; SOLLOCH et al., 2020) e algumas poucas africanas (NEMAT-GORGANI et al., 2019) e asiáticas (RAJALINGAM et al., 2002). Até fevereiro de 2020, foram descritos 1110 variantes dentre os 12 genes e pseudogenes *KIR* no banco de dados IPD-KIR (Release 2.9.0, ROBINSON et al., 2015), que concentra dados submetidos de todas as populações mundiais já descritas. Desses, grande parte (551) é restrita aos três genes mais bem caracterizados: *KIR3DL1, KIR3DL2 e KIR3DL3*.

O presente trabalho compreende o estudo de todos os 13 genes *KIR* em oito populações sul-americanas, incluindo seis populações isoladas e duas urbanas, levando em consideração a presença ou ausência dos ligantes HLA. Ainda, estudamos o impacto da diversidade alélica nos níveis de expressão de KIR2DL1, um dos receptores KIR mais potentes.

2 REVISÃO DE LITERATURA

2.1 CÉLULAS NATURAL KILLER (NK)

As células *natural killer* (NK) são linfócitos grandes e granulares do sistema imune, e correspondem a 2 a 18% dos linfócitos circulantes no sangue periférico em humanos (GRÉGOIRE et al., 2007). A função efetora das NK está relacionada a exocitose de seus grânulos, que contém moléculas líticas, como perforinas e granzimas (SMYTH et al., 2005). Essas moléculas formam poros na membrana da célula alvo, que alteram sua permeabilidade e contribuem para lise osmótica, bem como ao desencadeamento de processos pró-apoptóticos. Ainda, as células NK liberam citocinas e quimiocinas, que atraem outras células do sistema imune (LANIER, 2008).

Devido a sua capacidade natural de atacar seu alvo somente pela falta de inibição, estas células foram nomeadas de células assassinas naturais, ou ainda exterminadoras naturais (do inglês, *natural killer cells*) (KIESSLING et al., 1975; HERBERMAN; ORTALDO, 1981). São classificadas tradicionalmente como componentes do sistema imune inato, mas apresentam características associadas a imunidade adaptativa. Dentre elas, estão a expansão clonal após a ativação (DOKUN et al., 2001; SUN et al., 2009), o fato dessas células continuarem circulantes mesmo após a cessão da resposta citotóxica contra a célula alvo, e a capacidade de reativação dessas células de memória frente a um novo estímulo (O'LEARY et al., 2006; SUN et al., 2009; PAUST et al., 2010).

A caracterização de células NK periféricas pode ser feita pela presença de proteínas em sua superfície, que constituem os marcadores de diferenciação. Na superfície das células NK, há a presença dos marcadores CD56 e CD16 e ausência de expressão de CD3. O CD16 (do inglês, *cluster of differentiation 16*) é um receptor do fragmento cristalizável (Fc) de imunoglobulinas, e, portanto, responsável pela resposta mediada por anticorpo das células NK (BULMER; LASH, 2015). Interessantemente, o CD16 está presente em menos de 20% das células NK uterinas (uNK) (KING; LOKE, 1991), já que a ação mediada por anticorpos não é a função primária desse subtipo celular, e isso faz com que o CD16 não seja um bom marcador de uNKs. O CD56 (do inglês, *cluster of differentiation 56*), é uma proteína de adesão celular já descrita em células NK, mas também em neurônios e subtipos de células T

CD4+ e CD8+. O CD3 (do inglês, *cluster of differentiation 3*) é um correceptor necessário para ativação de células T e, portanto, marcador desse subtipo celular (LANIER et al., 1989). A ausência do CD3 é importante na diferenciação entre células NK e células T – incluindo células T *natural killer* (NKT) (GODFREY et al., 2004).

As células NK reconhecem, principalmente, células infectadas e neoplásicas (ROBERTSON; RITZ, 1990; BIRON, 1997). Além disso, um conjunto de células NK conhecidas como uterinas (uNK) têm importante função reprodutiva na implantação e estabelecimento da placenta durante os estágios intermediários de secreção menstrual e na gravidez inicial (KÄMMERER et al., 1999; BULMER; LASH, 2015). Essas funções fazem das células NK um importante alvo evolutivo.

A interação das células NK com outras células, normais ou alteradas, se dá por meio de receptores de membrana, que transduzem estímulos principalmente inibidores, mas também ativadores (LANIER, 1998). Os receptores que medeiam a atividade de células NK (NKR) pertencem a duas classes principais, os semelhantes a lectina, como CD94 e NKG2A, e os semelhantes a imunoglobulina, os KIR (PARHAM; GUETHLEIN, 2018). Os receptores das NK interagem predominantemente com regiões oligomórficas de moléculas HLA próprias, mas também com outras moléculas, como CD58, CD66 e CD99, que são reguladoras do sistema imune (LANIER, 2008). Juntas, essas três principais regiões genômicas – o MHC (do inglês, *major histocompatibility complex*, região no cromossomo 6 na qual se encontram os genes *HLA*), o NKC (do inglês, *natural killer complex*, região no cromossomo 12 que codifica os receptores semelhantes a lectina) e o LRC (do inglês, *leukocyte receptor complex*, região no cromossomo 19 que codifica os receptores semelhantes a imunoglobulina) – controlam a educação e a resposta das células NK (PARHAM; GUETHLEIN, 2018).

As células NK interagem e desencadeiam respostas imunes de acordo com o reconhecimento do perfil de expressão de moléculas de HLA de classe I na superfície de outras células. Células próprias não alteradas normalmente exibem alta expressão de moléculas HLA de classe I em sua superfície (APPS et al., 2015). Essa expressão normal de HLA é reconhecida e transduzida pelos NKR predominantemente na forma de sinais inibidores. Tais sinais levam a inibição global das células NK, que as impede de atacar células saudáveis, e assim, evitam respostas autoimunes (PARHAM, 2004; KIM et al., 2005).

Já células não-próprias, infectadas, ou neoplásicas, normalmente apresentam padrões anormais de expressão de HLA na sua superfície. Mais especificamente, células infectadas por vírus (APPS et al., 2015) ou neoplásicas (GARRIDO et al., 1997; KANEKO et al., 2011) frequentemente apresentam baixa expressão de HLA. A hipótese do "missing self" ou "perda do próprio" sugere que, quando as células NK reconhecem células que possuam baixos níveis de HLA na superfície, os sinais de inibição das células NK são diminuídos, o que desencadeia a ação citotóxica contra células alteradas (STORKUS et al., 1987; LJUNGGREN; KÄRRE, 1990; HILTON; PARHAM, 2017). Esse mecanismo é especialmente importante na proteção contra a evasão de patógenos, pois alguns patógenos se especializaram em escapar da resposta de linfócitos T CD8+ através da supressão da expressão de moléculas HLA nas células hospedeiras (JONES et al., 1995; SCHWARTZ et al., 1996; BENNETT et al., 1999; ISHIDO et al., 2000). O reconhecimento de baixos níveis de HLA por células NK teria sido vantajoso para evitar a evasão de respostas por linfócitos T CD8+. No entanto, para que as células NK reconheçam apenas perfis alterados de expressão de HLA, é necessário que haja um processo de educação das NK nas fases iniciais de seu desenvolvimento. Esse fenômeno também é conhecido como licenciamento das células NK (KÄRRE et al., 1986; KIM et al., 2005).

Alternativamente, o aumento anormal da expressão de HLA também pode levar à ativação das células NK (VIVIER et al., 2012). As NK podem desencadear citotoxicidade direcionada a essas células devido a expressão aumentada de moléculas HLA que são ligantes para receptores ativadores, o que também desregula o balanço de sinais e promove ação citotóxica. Sendo assim, tanto a ativação quanto a inibição de células NK são processos importantes para o controle da resposta imune mediada por estas células (VILCHES; PARHAM, 2002; PARHAM, 2004; PEGRAM et al., 2011).

2.2 RECEPTORES SEMELHANTES A IMUNOGLOBULINA DE CÉLULAS NK (KIR)

Os receptores semelhantes a imunoglobulina das células NK (KIR) receberam este nome por terem sido primeiramente descritos nesse conjunto de células. Porém, são também encontrados em uma subpopulação de linfócitos T, as células NKT (IKEDA et al., 1997; UHRBERG et al., 2001). São receptores de membrana que possuem domínios extracelulares semelhantes a imunoglobulina, além de uma região transmembrana e um domínio intracelular efetor.

Os KIR são nomeados pelo comitê de nomenclatura HUGO (do inglês, *Human Genome Organization*) de acordo com: a quantidade de domínios semelhantes a imunoglobulina extracelulares (D), que podem ser dois ou três (2D ou 3D); a estrutura do seu domínio citoplasmático, que pode ser curto ou longo (S, do inglês *short* ou L, do inglês, *long*); ou ainda por não possuir produto proteico funcional (P, do inglês *pseudogene*) (MARSH et al., 2003).

A função ativadora ou inibidora de KIR está associada a motivos específicos baseados em tirosina em seu domínio citoplasmático (FIGURA 1). Os domínios curtos geralmente se associam a moléculas adaptadoras DAP (DAP10 ou DAP12), que contém motivos ativadores baseados em tirosina (ITAMs, do inglês *immune tirosine-based activating motifs*), responsáveis pelo desencadeamento de uma cascata de mecanismos citotóxicos. Já os domínios longos possuem em sua própria estrutura motivos inibidores baseados em tirosina (ITIM, do inglês *immune tirosine-based inibitory motif*), capazes de silenciar receptores ativadores (SMYTH et al., 2005; LANIER, 2008). A exceção é KIR2DL4, que apesar de ter domínio longo, transduz sinais inibidores e ativadores, com função ativadora preponderante (FAURE; LONG, 2002).



FIGURA 1 - CONJUNTO DE RECEPTORES KIR, DESTACANDO SEUS DOMÍNIOS EXTRACELULARES E MOTIVOS BASEADOS EM TIROSINA.

FONTE: IPD-KIR (ROBINSON et al., 2015), modificado pelo autor (2019). LEGENDA: Os domínios em vermelho são correspondentes aos ITIM e em verde são representados os domínios ITAM. Está destacado em uma caixa vermelha o receptor KIR2DL4, que possui domínio longo e atividade predominantemente ativadora.

O ITAM está presente em moléculas como o DAP12, CD3 e FcɛRI (LANIER, 2008) e contém um resíduo de aspartato carregado negativamente que ajuda na sua associação com resíduos de carga elétricas opostas, como lisina ou arginina presente em receptores KIR ativadores (FENG et al., 2005). Os KIR e as moléculas contendo ITAM formam um complexo que permite a fosforilação dos resíduos de tirosina, que por sua vez desencadeiam uma cascata ativadora de mecanismos citotóxicos. Com isso, os receptores ativadores são capazes de promover a reorganização do citoesqueleto e a liberação de grânulos contendo perforinas e granzimas, bem como a ativação da transcrição de genes de citocinas e quimiocinas e a proliferação das células NK (SMYTH et al., 2005; LANIER, 2008).

Já o ITIM é parte da cauda citoplasmática de receptores KIR inibidores e tem seus resíduos de tirosina fosforilados quando os receptores inibidores encontram seus ligantes. Com isso, há o recrutamento de uma cascata de sinalização que culmina na desfosforilação de tirosino-quinases associadas aos receptores ativadores (SMYTH et al., 2005; LANIER, 2008). Portanto, a função inibidora de KIR inibidores consiste na própria supressão dos receptores ativadores.

Não são conhecidos os ligantes de todos os KIR. Um mesmo receptor pode reconhecer mais de um ligante HLA, como é o caso de KIR3DL2, que reconhece HLA-F, HLA-B27, HLA-A3 e HLA-A11 (revisado por AUGUSTO; PETZL-ERLER, 2015). Da mesma forma, um mesmo ligante HLA pode ter mais de um receptor que o reconhece, como HLA-C2, que se liga a KIR2DL1 e KIR2DS1 (WINTER et al., 1998). Entre os KIR que não têm ligantes conhecidos estão: KIR2DL5, KIR3DL3 e KIR2DS3 (PENDE et al., 2019). Além disso, muitas das interações propostas ainda são controversas. As interações de KIR com HLA confirmadas através de estudos funcionais estão descritas na TABELA 1, revisadas por NEMAT-GORGANI et al. (2019).

GORGANI et al., 2019).						
HLA	Epítopo de HLA	KIR	Referências			
HLA-A e -B	Bw4	KIR3DL1	(GUMPERZ et al., 1995; FOLEY et al., 2008)			
HLA-A	A3	KIR3DL2	(DÖHRING et al., 1996)			
HLA-A	A11	KIR3DL2	(HANSASUTA et al., 2004)			
HLA-C	C2	KIR2DL1	(FAN et al., 1996: WINTER:			
HLA-C	C1 e C2	KIR2DL2	LONG, 1997; WINTER et al.,			
HLA-C	C1	KIR2DL3	1998; HILTON, HUGO G et al.,			
HLA-C	C2	KIR2DS1	2015);			
HLA-C	C1 (apenas <i>C*16</i>)	KIR2DS2	(MOESTA et al., 2010)			
HLA-A	A11	KIR2DS4	(GRAEF et al., 2009)			
HLA-C	C2 (apenas C*0501, C*0202 e C*0401) e C1 (apenas C*1601, C*0102 e C*1402)	KIR2DS4	(GRAEF et al., 2009)			
HLA-C	C2	KIR2DS5 (apenas os alelos 2DS5*003, *004, *005, *006, *007 e *008)	(BLOKHUIS et al., 2017);			

TABELA 1 – INTERAÇÕES CONFIRMADAS ENTRE KIR E HLA (REVISADO POR NEMAT-

2.3 GENES KIR

Genes *KIR* são encontrados em primatas símios (incluindo humanos e macacos) e em ruminantes (como carneiros, bovinos e cervos), embora a expansão tenha ocorrido de maneira independente nas duas linhagens evolutivas, a partir de diferentes genes ancestrais (GUETHLEIN; ABI-RACHED; et al., 2007; GUETHLEIN et

al., 2015). O surgimento do *MHC-C* em primatas símios (FUKAMI-KOBAYASHI et al., 2005) foi importante para a evolução de *KIR*, com a expansão dos *KIR* que contém os domínios extracelulares D1 e D2 (KIR2DL1-3, KIR2DS1-5), em que o D0 corresponde a um pseudoéxon (FIGURA 2). Esses KIR2D possuem especificidade ao MHC-C determinada por um polimorfismo na posição 44 do domínio D1 (WINTER et al., 1998; KHAKOO et al., 2000; GUETHLEIN; OLDER AGUILAR; et al., 2007). Em humanos, o surgimento de um dimorfismo na posição 80 de HLA-C, que dá origem ao epítopo C1 (resíduo de asparagina) e ao epítopo C2 (resíduo de lisina), enriqueceu a diversificação dos KIR2D (tipo I, D1 e D2), em comparação aos outros primatas. Além disso, a região genômica dos genes *KIR* em humanos se diferencia daquela de outros primatas por ter duas regiões, denominadas centromérica e telomérica. Ainda, difere pela existência de grupos de haplótipos de conteúdo gênico funcionalmente distintos (UHRBERG et al., 1997; VALIANTE et al., 1997).



FIGURA 2 – ORGANIZAÇÃO DOS GENES KIR.



LEGENDA: Genes com estruturas semelhantes foram agrupados. Os retângulos azuis representam éxons codificadores e as caixas vermelham representam os pseudoéxons 3 e o éxon 2 de *KIR3DP1*, que pode apresentar deleção. Os números sobre as caixas representam o tamanho em pares de base (pb). As correspondências entre os éxons e os domínios codificados na proteína estão representados pelas chaves na parte inferior da imagem.

A família gênica *KIR* humana contém 13 *genes* (*KIR2DL1, KIR2DL23, KIR2DL4, KIR2DL5A, KIR2DL5B, KIR2DS1, KIR2DS2, KIR2DS3, KIR2DS4, KIR2DS5, KIR3DL1S1, KIR3DL2, KIR3DL3*) e dois pseudogenes (*KIR2DP1* e *KIR3DP1*) localizados no complexo de receptores leucocitários (LCR), na região genômica 19q13.4 (WILSON et al., 2000). A diversidade de genes *KIR* é caracterizada pelo incomum polimorfismo de presença e ausência, que leva a um polimorfismo haplotípico com relação ao conteúdo de genes, por polimorfismos alélicos e ainda por

uma expressão gênica diferencial entre indivíduos de uma população e variegada entre células NK de um mesmo indivíduo (PARHAM, 2004; UHRBERG, 2005).

Como mencionado, a variação de ausência e presença de genes e pseudogenes gera uma grande diversidade de haplótipos de conteúdo gênico, que podem ser classificados em dois grandes grupos, *A* e *B* (FIGURA 3) (UHRBERG et al., 1997; VALIANTE et al., 1997). Haplótipos do grupo *A* contém um número menos variável de genes, normalmente sete genes, além de dois pseudogenes. Possuem apenas um gene ativador clássico, de cauda curta, o *KIR2DS4*, além do *KIR2DL4* que também transduz sinais ativadores (FAURE; LONG, 2002). É, portanto, um haplogrupo com maior potencial inibidor. Já os haplótipos do grupo *B* contém uma maior variedade de combinações de genes ativadores, podendo apresentar de um a seis (VALIANTE et al., 1997; UHRBERG et al., 2002; PARHAM, 2008). O haplogrupo B é definido pelo HUGO pela presença de pelo menos um dos seguintes genes: *KIR2DL5, KIR2DS1, KIR2DS2, KIR2DS3, KIR2DS5* e *KIR3DS1*; enquanto o haplogrupo A é definido pela ausência destes (MARSH et al., 2003).

Os dois haplogrupos foram encontrados em todas as populações estudadas até o momento, mas suas frequências diferem significantemente entre populações. Por exemplo, em japoneses a frequência do haplogrupo A é 80% (YAWATA et al., 2006) e em aborígenes australianos a frequência deste haplogrupo é de apenas 2% (TONEVA et al., 2001). Já foi sugerido que a presença do haplogrupo B é vantajosa, em nível populacional, para a reprodução e no combate a patógenos, já que esse é associado a maiores níveis de ativação de células NK. A maior ativação, por sua vez, aumenta o perfil citotóxico das NK em infecções e aumenta potencial de invasão do trofoblasto na gravidez (PENMAN et al., 2016). Apesar da grande variação de presença e ausência de genes observada na região KIR, os genes KIR3DL3, KIR2DL4, KIR3DL2 e ainda o pseudogene KIR3DP1 estão presentes em praticamente todos os haplótipos, sendo chamados de genes moldura. Os genes KIR3DL3 e KIR2DL4 demarcam a região centromérica do haplótipo KIR, enquanto os genes KIR3DP1 e KIR3DL2 demarcam a região telomérica (VILCHES; PARHAM, 2002) (FIGURA 4). Já foi demonstrado que a grande maioria dos haplótipos frequentes nas populações mundiais são, na verdade, diferentes combinações de um número menor de segmentos centroméricos e teloméricos (PYO et al., 2010; HOLLENBACH et al., 2012). Raramente, alguns genes podem estar duplicados e algum gene moldura não está presente (GONZÁLEZ-GALARZA et al., 2015).



FIGURA 3 - COMBINAÇÕES DE SEGMENTOS KIR CENTROMÉRICOS E KIR TELOMÉRICOS IDÊNTICOS.



LEGENDA: Organizações possíveis dos haplótipos *KIR* encontradas em diversas populações. Estão indicadas as regiões centroméricas (*Cen*) e teloméricas (*Tel*) de acordo com sua associação aos haplogrupos *A* e *B*. Cada retângulo representa um *locus KIR*. Caixas em preto indicam genes característicos dos haplogrupos *A*, e caixas em branco indicam genes característicos dos haplogrupos *B*. Caixas com preenchimento tracejado destacam os genes que podem ser encontrados em ambos os haplogrupos. Finalmente, caixas em cinza destacam os genes moldura.

A variação estrutural de presença e ausência de genes *KIR* é explicada pela alta similaridade entre suas sequências, geralmente com 80 a 90% de identidade (TRAHERNE et al., 2010), pela orientação unidirecional dos genes e pela proximidade física das sequências dos genes *KIR*, que estão espaçados por regiões intergênicas curtas (1 a 3 kb) e conservadas (MARTIN et al., 2000). Esses fatores facilitariam a ocorrência de eventos de duplicação ou deleção de genes por recombinação não reciproca (MARTIN et al., 2003). Foi demonstrado ainda que sítios de elementos repetitivos podem dirigir a recombinação dentro dessa região, como os Alu, elementos nucleares intercalados longos (LINEs) e longas repetições em tandem (LTRs) (WILSON, M J et al., 2000). Essas características aumentam a ocorrência também de fusões gênicas, que ocorrem a partir de deleções em regiões intergênicas e

subsequente justaposição de segmentos de genes diferentes. Assim, as fusões gênicas possibilitaram o surgimento de receptores KIR híbridos, como observado para KIR2DL3-KIR2DP1 e KIR2DL1-KIR2DS1. Sugere-se que a elevada ocorrência de recombinações entre genes *KIR* é um recurso evolutivo seletivamente vantajoso para evitar a evasão ou subversão de patógenos (TRAHERNE et al., 2010).

FIGURA 4 - REPRESENTAÇÃO DE DIFERENTES HAPLÓTIPOS DE PRESENÇA DE GENES *KIR*, SEPARANDO REPRESENTANTES DOS HAPLOGRUPOS A E B.



FONTE: RAJALINGAM (2012). Adaptado pelo autor (2020).

LEGENDA: Organizações possíveis dos haplótipos *KIR* encontradas em diversas populações. Estão indicadas as regiões centroméricas e teloméricas pelos colchetes superiores. Os colchetes a direita indicam o pertencimento de cada um dos haplótipos ao haplogrupo *A* ou *B*. Cada retângulo representa um *locus KIR*. Caixas em preto indicam genes moldura, caixas em cinza indicam genes de receptores inibidores e em caixas em branco estão representados genes de receptores ativadores ou pseudogenes.

Ao contrário de receptores de outros linfócitos, como células T e B, os genes de receptores de células NK não sofrem processos de rearranjo gênico em suas células somáticas para gerar diversidade entre as células de um mesmo indivíduo. Em contrapartida, possuem diversidade de expressão variegada de seus receptores na superfície que gera diversidade entre diferentes clones celulares. Sugere-se que esta

é controlada por desmetilação aleatória de promotores *KIR* durante a maturação celular. Como consequência, diferentes clones de células NK podem expressar subconjuntos diferentes de moléculas KIR em sua superfície (UHRBERG, 2005; ORR; LANIER, 2010). Além disso, os receptores KIR não silenciados apresentam diferentes níveis de expressão na superfície das células NK. É conhecido que o polimorfismo alélico de alguns genes *KIR* influencia os níveis de expressão na superfície (DUNPHY et al., 2015).

A combinação de presença de certos genes *KIR* com alelos *HLA* que codifiquem seus respectivos ligantes já foi relacionada com várias doenças, como a síndrome da imunodeficiência humana (AIDS) (MARTIN et al., 2002), infecção pelo vírus da hepatite C (KHAKOO et al., 2004), neoplasia cervical induzida pelo papilomavírus humano (CARRINGTON et al., 2005), artrite psoriática (NELSON et al., 2004), diabetes tipo I (SLIK, VAN DER et al., 2003), pré-eclâmpsia (HIBY et al., 2004), câncer (AUGUSTO, 2016) e pênfigo foliáceo (AUGUSTO et al., 2012a; AUGUSTO et al., 2015). Isso sugere que a diversidade populacional nas interações dos receptores KIR com seus ligantes são importantes na modulação da resposta das células NK.

2.4 GENE KIR2DL1

O gene *KIR2DL1* tem presença variável, sendo observado com frequências que abrangem um intervalo de 65,6% em taiwaneses (YEN et al., 2006) até 100% em algumas populações asiáticas (YAWATA et al., 2006). A frequência mais baixa desse gene em uma população brasileira, 84%, foi observada em indígenas Guarani (AUGUSTO et al., 2013a). O *KIR2DL1* tem 9 éxons, sendo que o éxon 3 é na verdade um pseudoéxon (FIGURA 5). Em outros genes, esse éxon é responsável pela codificação do domínio D0 e sua perda de função foi observada não só em *KIR2DL1* como também em outros genes *KIR*, como *KIR2DL23* e *KIR2DS1-5*. A ausência do domínio D0 combinado com o polimorfismo na posição 44 do domínio D1, que codifica uma metionina ou lisina, confere a capacidade dessas proteínas se ligarem a moléculas de HLA-C (VILCHES et al., 2000; PARHAM; GUETHLEIN, 2018).

O receptor KIR2DL1 tem como ligantes um subgrupo de moléculas de HLA-C que contém um resíduo de asparagina na posição 77 e lisina na posição 80, presente no grupo C2 (FAN et al., 2001). As moléculas HLA-C2 são reconhecidas por KIR2DL1

e também pelo seu receptor homólogo ativador KIR2DS1. Além disso, observa-se certo grau de reconhecimento cruzado de C2 por KIR2DL2 e KIR2DL3 (WINTER et al., 1998). Dentre esses três, o KIR2DL1 é o receptor que se liga com maior afinidade e especificidade ao HLA-C2, além de desencadear resposta inibidora mais pronunciada (STEWART et al., 2005; MOESTA et al., 2008; HILTON, HUGO G et al., 2015). Já foi demonstrado que o nível de expressão de KIR2DL1 altera a capacidade citotóxica das células NK (BARI et al., 2009; CHAROUDEH et al., 2012). Quando em níveis mais baixos, leva a menor inibição global e a um consequente aumento na atividade citolítica dessas células (XUE et al., 2013).

A importância funcional de *KIR2DL1* evidencia-se também pelas implicações clínicas, uma vez que o polimorfismo desse gene já foi associado a diversas doenças, como a hanseníase tuberculóide (FRANCESCHI et al., 2008), endometriose (ZHANG et al., 2006), neoplasia cervical (CARRINGTON et al., 2005), leucemia (JUNEVIK et al., 2007), pré-eclâmpsia (HIBY et al., 2004), na doença do enxerto contra hospedeiro (DECH) (LUDAJIC et al., 2009), e, ainda, a uma possível imunodeficiência devido a expressão constitutiva do receptor em toda a população de células NK (GAZIT et al., 2004), bem como a evolução clínica de transplantes (VELARDI et al., 2002; BARI et al., 2009).

FIGURA 5 - ORGANIZAÇÃO DO GENE KIR2DL1.



FONTE: (ROBINSON et al., 2015) LEGENDA: As setas indicam cada um dos éxons e o número representa a quantidade de nucleotídeos codificados por cada éxon. A caixa em vermelho representa o pseudoéxon 3.

2.5 POPULAÇÕES HUMANAS E DIVERSIDADE GENÉTICA

A evolução e a diversidade de famílias gênicas como a de *KIR* estão intrinsecamente relacionadas com a história das populações humanas. Os humanos modernos tiveram origem na África há cerca de 200 milhões de anos; porém, estimase que sua migração além da África só tenha ocorrido há cerca de 100 milhões de

anos. A partir da África, a espécie humana povoou a Ásia, Europa, Oceania e, finalmente, as Américas (CAVALLI-SFORZA et al., 1994; REYES-CENTENO, 2016). Atualmente, dentre as populações humanas modernas, as africanas em geral apresentam maior diversidade genética, complexa subestrutura populacional e baixo desequilíbrio de ligação quando comparadas a populações não africanas, de maneira correlata com a distância de migração das populações ancestrais desde a África. Portanto, os ameríndios apresentam menor diversidade genética, em torno de 60 a 70% daquela observada em africanos (ROSENBERG et al., 2002; WANG et al., 2007), menor subestruturação populacional e maior desequilíbrio de ligação (TARAZONA-SANTOS et al., 2001; GASPAR et al., 2002; TISHKOFF; VERRELLI, 2003).

2.5.1 POPULAÇÕES AMERÍNDIAS

Com relação a origem das populações ameríndias, há fortes e numerosas evidências de que essas são descendentes de populações do nordeste asiático (CAVALLI-SFORZA et al., 1988; REICH et al., 2012; POSTH et al., 2018). Essas populações teriam migrado do continente asiático para as Américas através do Estreito de Bering, em três diferentes ondas de migração. A onda de migração responsável pelo povoamento da América do Sul teria ocorrido a partir de uma população ancestral denominada paleoíndios, ou "primeiros homens" (SCHANFIELD, 1992; CAVALLI-SFORZA et al., 1994; WANG et al., 2007; REICH et al., 2012). Outras evidências sugerem que a origem dos povos da América pode ter influências também da Australasia (que inclui a Austrália, Melanésia e ilhas do Sudeste Asiático), porém essa influência é muito inferior a das populações do nordeste asiático (SKOGLUND et al., 2015). Todas as análises indicam que os ameríndios atuais são mais próximos geneticamente das populações asiáticas do que de populações de qualquer outra região geográfica.

As populações ameríndias que foram analisadas nesse estudo pertencem a três grupos étnicos principais: os Kaingang, os Guarani, e os Aché. Os Kaingang têm subsistência baseada na agricultura e sua língua nativa pertence ao grupo linguístico Jê. Essas populações vivem majoritariamente no sul do Brasil, distribuídos em mais de 30 áreas de reservas. Já os ameríndios Guarani, podem ser encontrados nas regiões sul, sudeste e central do Brasil, além de outros países da América do Sul, que

incluem Paraguai, Argentina e Bolívia. Os Guarani têm em comum, além de costumes e fenótipo, a língua Guarani, que pertence ao tronco Tupi-Guarani (TSUNETO et al., 2003). No Paraguai, os Guarani são o grupo indígena predominante, porém são encontradas também outras populações, como os Aché. Os Aché, por sua vez, são um pequeno grupo ameríndio do leste do Paraguai que vivem em pequenos grupos de 15 a 70 indivíduos e evitam visitas e miscigenação com grupos Guarani. Esse grupo difere dos Guarani em costumes; pela sua língua Aché, que também pertencente ao tronco linguístico Tupi-Guaraní; e pelo fenótipo que se caracteriza por pele, cabelos e olhos mais claros (HILL; HURTADO, 1996; TSUNETO et al., 2003). Além disso, é debatido se os Aché são um grupo Guarani do Paraguai. Porém, estudos genéticos com marcadores *HLA* (TSUNETO et al., 2003), mtDNA (SCHMITT et al., 2004) e outros marcadores autossômicos (BATTILANA et al., 2002; GASPAR et al., 2002; CALLEGARI-JACQUES et al., 2007) apontam maior proximidade entre Aché com os grupos Guarani.

As populações ameríndias continuaram isoladas ao longo de gerações, em função principalmente de fortes barreiras culturais que levaram ao isolamento reprodutivo. O fluxo gênico entre populações não-ameríndias e as populações ameríndias incluídas no presente estudo foi estimado com base nas frequências de alelos dos genes *HLA* (TSUNETO et al., 2003). Foi estimado que a população de Guarani Ñandeva possui o fluxo gênico mais elevado (14,3%), seguido das populações Kaingang (7,2%) e Guarani M'byá (3,7%). Nesse mesmo estudo, não foi encontrada evidência de fluxo gênico na população de Guarani Kaiowá, nem na população Aché.

O polimorfismo de presença e ausência de genes *KIR* foi caracterizado em quatro das seis populações de ameríndios incluídas nesse trabalho (não incluídos são os Kaingang de Ivaí e Aché) (AUGUSTO et al., 2013a). Porém, não foram descritas frequências alélicas de nenhum dos genes *KIR* nessas populações. Já foram feitos estudos a nível alélico, porém com poucos genes, em ameríndios do México (*KIR2DL1* e *KIR2DL3*), da Argentina (*KIR2DL4*) e da Venezuela (*KIR3DL1* e *3DS1*) (GONZÁLEZ-GALARZA et al., 2015), e, apenas um estudo na população de ameríndios Yucpa na Venezuela caracterizou o polimorfismo haplotípico de *KIR* (GENDZEKHADZE et al., 2009).

2.5.2 JAPONESES NO BRASIL

Atualmente, o Brasil tem a maior população de ancestralidade japonesa não residente no Japão. Segundo dados do censo demográfico brasileiro de 2010, são mais de 1,5 milhão de descendentes de japoneses no Brasil. Dentre esses, o Paraná tem a segunda maior população de descendentes de japoneses do Brasil, atrás apenas do estado de São Paulo (IBGE, 2013). Os descendentes de japoneses se organizam em comunidades em grandes centros urbanos. No Paraná, estão presentes principalmente em Londrina, Curitiba e Maringá (SAKURAI et al., 2010), mas também em maior proporção em algumas cidades menores como Assaí e Uraí (CATANIO, 2002).

A imigração japonesa ao Brasil ocorreu devido à escassez de recursos e a superpopulação que ocorria no Japão após a Restauração Meiji. No final do século XIX, muitos japoneses migraram para países como Estados Unidos, México e Peru. A migração ao Brasil começou a partir de 1895, quando o Brasil e o Japão assinaram o Tratado de Amizade, Comércio e Navegação, para suprir a demanda de mão de obra no país. Até 1941, já tinham sido registradas a entrada de aproximadamente 175 mil japoneses no Brasil. Os descendentes dos japoneses que imigraram ao Brasil têm baixa miscigenação com outros grupos étnicos (SAKURAI et al., 2010).

A diversidade alélica de *KIR* foi caracterizada em indivíduos japoneses, porém em baixa resolução (YAWATA et al., 2006). Naquele trabalho, foi visto que os japoneses têm a maior frequência mundial de haplótipos *KIR A* (80%) e também de ligantes C1 (92%). Ainda, a população de descendentes de japoneses residentes em Curitiba já foi caracterizada para a presença e ausência de genes *KIR* em estudos anteriores do nosso grupo, que mostrou que o haplótipo *KIR A* apresenta frequência de 60,8% e o ligante HLA-C1 apresentou frequência de 85,1% (AUGUSTO et al., 2016). Até o momento, não foram publicados estudos que caracterizam a diversidade dos genes *KIR* em alta resolução em indivíduos japoneses ou com ascendência japonesa. Ainda, devido às evidências de que as populações Ameríndias foram originadas de ancestrais asiáticos (REICH et al., 2012) mencionadas anteriormente, é interessante comparar a diversidade genética das populações japonesa e indígenas.

2.6 JUSTIFICATIVA

Estudos de genética de populações são de grande importância para entender a história evolutiva da espécie humana e as bases genéticas das doenças humanas. Os padrões de diversidade genética em populações humanas, considerando tanto os genes que influenciam doenças quanto outras regiões do genoma, são o resultado de uma história evolutiva complexa. A detecção de fatores evolutivos (mutações, fatores demográficos e seleção) que moldam a diversidade populacional pode auxiliar a compreender a essencialidade da função biológica dos genes analisados e, eventualmente, auxiliar na elucidação da função de regiões ainda desconhecidas. Além disso, levar em conta a estrutura genética populacional é fundamental para realizar estudos de associação a doenças complexas (multifatoriais) e no desenvolvimento de metodologias diagnósticas e terapêuticas que contemplem efetivamente todas as populações humanas.

A família gênica *KIR* evoluiu rapidamente, em comparação com outras famílias gênicas (KHAKOO et al., 2000; SAMBROOK et al., 2005; GUETHLEIN et al., 2007; TRAHERNE et al., 2010). Alguns autores propõem que a alta taxa de recombinação e o surgimento de genes híbridos seriam favorecidos por seleção positiva, pois o aumento da diversidade contribui para evitar a evasão de patógenos (TRAHERNE et al., 2010). Já foi demonstrada a ocorrência de seleção em genes *KIR* (NORMAN et al., 2007, 2013a; AUGUSTO et al., 2019) e há evidências de coevolução entre genes *KIR* e *HLA* (revisado em AUGUSTO e PETZL-ERLER, 2015). Contudo, poucos estudos exploraram a diversidade de KIR em Ameríndios (GENDZEKHADZE et al., 2006, 2009; GUTIÉRREZ-RODRÍGUEZ et al., 2006; AUGUSTO et al., 2013b; AUGUSTO; HOLLENBACH; et al., 2015)

Além de descrever a diversidade desses genes em populações sulamericanas isoladas e urbanas, a nossa hipótese é de que certas combinações funcionais de genes *KIR* e ligantes HLA possam ter sido favorecidas por seleção natural. Dentre os HLA de classe I, as moléculas HLA-C apresentam os mais baixos níveis de expressão (APPS et al., 2015), possivelmente pela sua especialização em modular a função das células NK através dos KIR2D; em contraste com HLA-A e B, que teriam se especializado na resposta de células T (OLDER AGUILAR, A. M. et al., 2010). Já foi demonstrado que o receptor inibidor KIR2DL1 reconhece o epítopo C2 de alótipos HLA-C e que isso desencadeia a resposta inibidora mais potente dentre os receptores KIR (NORMAN; PARHAM, 2005; STEWART et al., 2005; MOESTA et al., 2008). Além disso, a presença de KIR2DL1 já foi associada a diversas doenças e ao sucesso de transplantes (VELARDI et al., 2002; RAJAGOPALAN; LONG, 2005). Portanto, dentre os receptores KIR, o estudo da diversidade de KIR2DL1 e seu impacto nos níveis de expressão do receptor e nos níveis de expressão de HLA-C pode nos fornecer evidências de possíveis alvos de regulação gênica e de coevolução entre *KIR* e *HLA*.

3 OBJETIVOS

3.1 OBJETIVO GERAL

Caracterizar em alta resolução a diversidade de todos os genes *KIR* em oito populações sul-americanas (seis populações ameríndias, eurodescendentes e descendentes de japoneses), bem como a presença dos ligantes *HLA*, e avaliar o possível impacto da variação genética nos níveis de expressão gênica, em específico, de *KIR2DL1*.

3.2 OBJETIVOS ESPECÍFICOS

- Estimar as frequências alélicas de todos os genes *KIR* em amostras de três populações de ameríndios Guarani (Guarani Kaiowá – MS –, Guarani Ñandeva – MS – e Guarani M'byá – PR), duas populações de ameríndios Kaingang (provenientes de Ivaí e Rio das Cobras – PR), uma população de ameríndios Aché (oriundos do Paraguai), e duas amostras urbanas de Curitiba (PR), uma de descendentes de japoneses e outra de ascendência predominantemente de europeia.
- Estimar as frequências de moléculas HLA de classe I descritas como ligantes de KIR.
- Determinar, em alta resolução, os haplótipos KIR mais frequentes nas populações em estudo.

 Verificar o impacto funcional da diversidade de presença e ausência e da diversidade alélica de *KIR2DL1* na sua expressão gênica e de seu ligante HLA-C.

4 MATERIAL E MÉTODOS

4.1 AMOSTRAS BIOLÓGICAS

A partir de um painel de amostras do biorrepositório do Laboratório de Genética Molecular Humana (LGMH) da UFPR, foram selecionados 707 indivíduos pertencentes às oito populações sul-americanas, descritas na TABELA 2. O material biológico em estoque dessas amostras foi previamente coletado e o DNA foi extraído ao longo dos anos pelo método do fenol clorofórmio álcool isoamílico (SAMBROOK et al., 1989) e estocado a -80°C. Amostras com baixa disponibilidade ou qualidade do material biológico foram excluídas.

As populações do estudo incluem seis populações ameríndias isoladas, sendo três Guarani (Guarani Kaiowá – n=150; Guarani Ñandeva – n=81; Guarani M'bya – n=84), duas populações Kaingang provenientes de Ivaí (n=93) e de Rio das Cobras (n=84) e uma população de estabelecida no Paraguai, os Aché (n=51). Os Aché, que eram nômades, hoje vivem em reservas na região fronteiriça do Paraguai com o Brasil, a 200 quilômetros da fronteira com o Brasil e a 400 km da reserva de Rio das Cobras, onde vivem os Guarani M'bya e uma população Kaingang. A coleta de material biológico das populações ameríndias foi feita entre o final dos anos de 1980 e o começo de 1990, como descrita em trabalhos anteriores do nosso grupo (TSUNETO et al., 2003).

Esse estudo também inclui duas populações urbanas de Curitiba, uma predominantemente eurodescendente (n=109) e uma de descendentes de japoneses (n=75). Todos os indivíduos de ascendência japonesa incluídos nesse estudo nasceram no Brasil e apresentam os dois genitores ou os quatro avós biológicos nascidos no Japão. Eles relataram não ter histórico de miscigenação com não japoneses. Dentre os eurodescendentes de Curitiba coletados para análise de expressão gênica (n=48), foram incluídos indivíduos adultos (em média 26 anos) que não relataram nenhuma doença aguda ou crônica e que residiam em Curitiba e região metropolitana.

Todos os participantes foram informados dos objetivos da pesquisa e concordaram em participar através de consentimento oral ou escrito, de acordo com a regulação vigente no momento da coleta. Esse estudo está incluído no projeto
registrado pelo número 02727412.4.00000096 e aprovado pelo Comitê de Ética em Pesquisa Humana da Universidade Federal do Paraná (CEP-UFPR) e pela Comissão Nacional de Ética em Pesquisa (Conep) sob o parecer de número 505.988 (02/01/2014).

Etnia	Detalhes	Abreviação ¹	Idade média (min - máx)	Localização no momento da coleta (TSUNETO et al, 2003)	Latitude	Longitude	Número amostral (n = 707)
Eurodescen- dentes	de Curitiba	CTBA	25,95 (20 - 64)	Curitiba/PR	-25,45	-49,22	109
Descendentes de japoneses	de Curitiba	JAP	30,62 (15 - 69)	Curitiba/PR	-25,45	-49,22	75
	Kaiowá	GKW	25,57 (12 - 66)	Limão Verde/MS	-23,12	-55,05	150
Commu				Amambai/MS	-23,06	-55,12	
Guarani	M'bya	GRC	-	Nova Laranjeiras/PR	-25,18	-52,32	84
	Ñandeva	GND	-	Amambai/MS	-23,06	-55,12	81
и ·	Ivaí	KIV	-	Ivaí/PR	-24,50	-51,67	93
Kaingang	Rio das Cobras	KRC	-	Nova Laranjeiras/PR	-25,18	-52,32	64
Aché	Paraguai	ACHE	-	Arroyo Bandera/Paraguai	-23,30	-55,50	51
				Chupa Pou/Paraguai	-24,10	-56,30	

TABELA 2 - CARACTERIZAÇÃO DAS OITO POPULAÇÕES INCLUÍDAS NO ESTUDO.

LEGENDA: 1 - Abreviação adotada no contexto deste trabalho. Na idade média, estão indicados entre parênteses as idades mínimas e máximas em cada população. O hífen indica que informações sobre a idade dos indivíduos não está disponível para aquela população.

4.2 GENOTIPAGEM DE PRESENÇA E AUSÊNCIA DO GENE KIR2DL1

Para a análise de alelos e de expressão do gene *KIR2DL1*, foram convidados 48 indivíduos eurodescendentes residentes em Curitiba. O critério de inclusão foi indivíduos com a presença de *KIR2DL1*. A presença de genes *KIR* foi caracterizada previamente na amostra de eurodescendentes de Curitiba (AUGUSTO, D. G. et al., 2012). Os indivíduos que não haviam sido previamente genotipados tiveram a presença do gene *KIR2DL1* avaliada pelo método de PCR com oligonucleotídeos para sequência específica (PCR-SSP), utilizando dois pares de oligonucleotídeos específicos para *KIR2DL1* e um par de oligonucleotídeos específicos para o gene *HLA-DRB1*, que é utilizado como controle positivo de amplificação (TABELA 3). A

análise dos produtos das reações de PCR foi feita por eletroforese em gel de agarose (2%, 80V, 60 min) banhado em TBE (Tris-Borato-EDTA), usando o corante de DNA UniSafe (*Uniscience*, São Paulo, BR), 2 µL de produto de PCR, 2 µL de tampão de carreamento (azul de bromofenol e glicerol) e o marcador de peso molecular de 1kb Scada® (*Sinapse* Inc., São Paulo, BR).

4.3 IDENTIFICAÇÃO DOS ALELOS DE KIR2DL1

A genotipagem alélica de *KIR2DL1* foi feita para os 48 indivíduos eurodescendentes que participaram do ensaio de expressão, através de sequenciamento pelo método de Sanger (SANGER et al., 1977). Dos nove éxons do gene, foram selecionados os éxons 1, 4, 5, 7 e 9, uma vez que os éxons 2, 6 e 8 e o pseudoéxon 3 são não informativos para a distinção dos principais grupos alélicos, segundo os alelos já descritos no banco de dados KIR-IPD (ROBINSON et al., 2013). Ainda, um número adicional de 258 amostras de indivíduos eurodescendentes foi sequenciado apenas para o éxon 1, que contém a variante *rs2304224*, para aumentar o poder estatístico nos testes de associação com esse SNP e a expressão de KIR2DL1 e HLA-C.

Foi utilizada a estratégia de amplificação de segmentos longos, em que foram amplificados dois segmentos, denominados A e B, a partir dos oligonucleotídeos descritos na TABELA 3. O fragmento A compreende os éxons 1 a 5 (5772 pb) e foi amplificado utilizando o *mix* enzimático GoTaq® Long PCR Master Mix (*Promega*, Fitchburg, WI). O fragmento B compreende os éxons 7 a 9 (2218 pb), e foi amplificado utilizando a enzima Taq Platinum (*Invitrogen*, CA, EUA). Os segmentos amplificados foram utilizados para sequenciamento individual de cada éxon de *KIR2DL1*, utilizando o *mix* enzimático Big Dye Terminator Cycle Sequencing Standart v3.1 (Thermo Fisher Scientific Inc., Wilmington, EUA), de acordo com as instruções do fabricante, e os oligonucleotídeos descritos na TABELA 3.

Antes da injeção do produto no sequenciador, este foi ressuspenso em 10 µL de Hi-Di Formamida (Thermo Fisher Scientific Inc., Wilmington, EUA). A leitura das sequências nucleotídicas foi realizada por eletroforese capilar no sequenciador 3500xl Genetic Analyzer (Thermo Fisher Scientific Inc., Wilmington, EUA).

			TABELA 3 - OLIGONUCL	EOTÍDEOS UTILIZADOS NESTE TRABA	ALHO.		
Reação	Gene	Par	ID do primer	Sequência (5' - 3')	Posição genômica	Tamanho do fragmento (pb)	Referências
		<	2DL1_5UTR198_Fwd	GGACACTAGGTGTCAAATTCTAGC	-198	E 770	
Amplificação pré-		K	2DL1_Ex5_5574_Rev	ACAAGCAGTGGGTCACTTGAC	5574	711'C	
sequenciamento	NIKZULI		2DL1_Int6_11767_Fwd	TCACAGGAGGACAGGTGGTT	11767		
		ם	2DL1_Int9_13985_Rev	GCTGTTGTCTCCCTAGAAGACG	13985	2,210	
		Éxon 1	2DL1_5UTR198_Fwd	GGACACTAGGTGTCAAATTCTAGC	-198	> 1,000	
		Éxon 4	2DL1_seq_Int3_3399_Fwd	GAAGGAGAGAGAGATAAGACACCAGG	3399	> 1,000	
Sequenciamento	KIRODI 1	Éxon 5	2DL1_seq_Int4_5293_Fwd	GAAGATCCTCCCTGAGGAAAC	5293	281	
		Éxon 7	2DL1_seq_Int6_12972_Fwd	GGGTGCTTGTCCTAAAGAGG	12972	> 1,000	(HILTON, HUGO G. et al., 2015)
		Éxon 9	2DL1_seq_Int7_13601_Fwd2	GGACCCAGAAGTGCCCTC	13601	384	
			KIR2DL1_17B_F	CCATCAGTCGCATGACG	3669	ç	
	KIKZUL1	1/B	KIR2DL1_17B_R	TCACTGGGAGCTGACAC	3759	90	(AUGUSTU et al., 2013a)
Genotipagem de			KIR2DL1_16_F	TGGACCAAGAGTCTGCAGGA	13656		
presença/ausência	KIR2UL1	16	KIR2DL1_16_R	TGTTGTCTCCCTAGAAGACG	14004	348	(KULKARNI et al., 2010)
	HLA-	-	DRB1_F	TGCCAAGTGGAGCACCCAA	8108	0 0 1	
	DRB1	Control	DRB1_R	GCATCTTGCTCTGTGCAGAT	8890	/82	(BUNCE et al., 1995)
			KIR2DL1_17B_F	CCATCAGTCGCATGACG	3669	ç	
	NIKZULI	9/L	KIR2DL1_17B_R	TCACTGGGAGCTGACAC	3759	90	(AUGUSTU et al., 2013a)
			2DL1_CN_F1	GTTGGTCAGATGTCATGTTTGAA	3558		
	KIR2DL1	ddPCR	2DL1_CN_R1	CCTGCCAGGTCTTGCG	3699	141	(VILCHES ET al., 2007)
			2DL1 Fluorescent Probe	AACGACACTTTGCGCCTCATT	3611		
Genotipagem de		(KIR3DL3_1C_F2	TTCTGCACAGAGGGGGGGATCA	3182		
numero de copias	KIK3DL3	2	KIR3DL3_1C_R2	GAGCCGACAACTCATAGGGTA	3346	165	(KULKAKNI et al., 2010)
		Ģ	KIR3DL3_2B_Fwd	CACTGTGGTGTCTGAAGGAC	1556	100	
	NINJULS	Q V	KIR3DL3_2B_Rev	GGAGTGTGGGTGTGAACTG	1754	20	(AUGUSTO ELal., ZUTSA)
		F	KIR3DL3_RT_63_Fwd	CTGCAATGTTGGTCAGATGTCAG	3144	100	Modificado de VILCHES et al.,
	NIRSULS	- Y	KIR3DL3_RT_63_Rev	GGGAGCYGACAACTCATAGGGTA	3348	204	(2007)
		5	HLA-C1 N80 F	AGCCAATCAGCGTCTCCGC+A	-79	L	
Genotipagem		5	HLA-C1 N80 R	GCTCTGGTTGTAGTAGCCGCGCAG+G	466	040	(AMURIN €Lal., ∠U Iδ)
alotípica de HLA-C	NLA-C	ç	HLA-C2 K80 F	CCATTGGGTGTCGGGTTCT+A	-103	C	
		Š	HLA-C2 K80 R	GCTCTGGTTGTAGTAGCCGCGCAG+T	466	ROC	(AMURIN EL AL., 2018)
LEGENDA: ((LNA) ao oliç	sequências e jonucleotíde	em negrito s o, estes est	são correspondentes aos oligonu tão indicados após o sinal de ma	cleotídeos desenhados nesse estudo. Se is (+). Int: intron. Ex: éxon. F ou Fwd: olig	e houver adiç jonucleotídec	ão de ácidos nucl direto. R ou Rev	eicos bloqueados oligonucleotídeo
reverso. PD:	pares de ba	se.					

4.4 DETERMINAÇÃO DO NÚMERO DE CÓPIAS DE KIR2DL1

Para quantificar o número de cópias de *KIR2DL1*, foram realizadas reações de PCR em tempo real (qPCR) utilizando DNA genômico. O ensaio de qPCR foi realizado em todas as amostras genotipadas para *KIR2DL1* por sequenciamento (n=48). O ensaio foi realizado em triplicata para o gene alvo (*KIR2DL1*) e o gene utilizado como referência foi o *KIR3DL3*, por ser um gene moldura com número fixo de cópias. Foi utilizado um par de oligonucleotídeos adicional de *KIR3DL3* para averiguar casos de haplótipos com número expandido de cópias (>2). O reagente utilizado foi o GoTaq® qPCR Master Mix (Promega, Fitchburg, WI), e a reação ocorreu no equipamento ViiA[™] 7 Real-Time PCR System (*Applied Biosystems*, Foster City, CA). Além disso, nos casos em que a expansão do gene (3 cópias) não foi descartada, foi realizada PCR digital em microgotas (ddPCR, do inglês, *digital droplet PCR*), para quantificação absoluta do número de cópias do gene. A ddPCR foi realizada após fragmentação do DNA com a enzima de restrição *Alul*, utilizando o gene de referência *RPP30*, na plataforma QX200[™] Droplet Digital [™] PCR System (Bio-Rad Laboratories, Hercules, CA). Os oligonucleotídeos iniciadores estão descritos na TABELA 3.

4.5 ENSAIO DE EXPRESSÃO GÊNICA DE KIR2DL1

A quantificação da expressão gênica de KIR2DL1 foi feita através do ensaio de citometria a partir de sangue fresco em uma subamostra de eurodescendentes de Curitiba (n=48). Foram coletados 8 mL de sangue periférico de cada voluntário em tubo VACUETTE® (*Greiner-Bio One*, Austria) contendo EDTA. A separação das células mononucleadas do sangue periférico (PBMC) foi feita utilizando o tubo de separação Leucosep™ (*Greiner Bio-One*, Austria) que contém uma membrana de separação seletiva e Ficoll Hypaque (*Sigma Aldrich*, Missouri, EUA). As PBMC foram quantificadas por contagem direta em câmara de Neubauer visualizada em microscópio óptico, e foram utilizadas 0,5x10⁶ células por ensaio de citometria.

As células foram marcadas com os anticorpos específicos para caracterização de células NK, o anticorpo CD56 conjugado a molécula fluorescente ficoeritrina (PE) (*Miltenyi Biotec*, Alemanha – Cat #130-100-622, RRID:AB_2658731) e o anticorpo CD3 conjugado a molécula de fluorescência ficoeritrina-cianina 5 (PE-Cy5)

(*Biolegend*, Califórnia, EUA – Cat #00328, RRID:AB_1575008). Para cada indivíduo, a marcação das células foi feita também com anti-KIR2DL1 CD158a, clone REA284, conjugado com isotiocianato de fluoresceína (FITC) (*Miltenyi Biotec*, Alemanha – Cat# 130-103-967, RRID:AB_2655323). Após incubação, as células foram lavadas com a solução fosfato-salina PBS, na concentração de 1X, e abrigadas de luz e calor em geladeira (temperatura padrão de 0 a 8 °C). As células marcadas foram analisadas no citômetro de fluxo BD FACSVerse[™] (*BD Biosciences*, San Jose, CA). Para a aquisição das células, o número mínimo de eventos por amostra foi alterado de 10.000 eventos totais para 1.000. Essa alteração é sugerida para populações celulares raras, como é o caso de células NK-KIR2DL1+ no PBMC. Foi verificado nesse estudo que os linfócitos compõem cerca de 80% do PBMC, e uma mediana de 8,71% desses são células NK. Ainda, dentre as células NK, 15,65% em mediana expressam o KIR2DL1. Isso resulta em um total de 1,09% de células NK KIR2DL1 positivas no conjunto total de PBMC. Amostras que não atingiram o número mínimo de eventos na citometria foram excluídas do estudo.

A compensação para ensaios com múltiplos fluoróforos foi feita utilizando esferas do conjunto CS&T (do inglês, Cytometer Setup & Tracking, BD Biosciences, San Jose, CA) que são marcadas com fluoróforos em concentrações conhecidas, PBMC não marcadas e PBMC marcadas com uma única fluorescência. A estratégia de gates (seleção de células com parâmetros celulares similares na FIGURA 6) foi feita a partir de um primeiro gate para selecionar a população de linfócitos, com base nos parâmetros de granulosidade (SSC - side scatter) e tamanho (FSC - forward scatter), uma vez que essas células apresentam granulosidade baixa (SSClow) e tamanho baixo (FSC^{low}). A partir da população de linfócitos, foram selecionadas no segundo gate as células NK, caracterizadas pela expressão em superfície de CD56 -PE positivas – e ausência de CD3 – PE-Cy5 negativas. A partir do subconjunto de células NK, um terceiro gate selecionou as células NK que expressam KIR2DL1 -FITC positivas. A contagem de células FITC positivas informa a porcentagem de células NK que expressam KIR2DL1 (presença do receptor no conjunto de células NK). O valor de MFI (median fluorescence intensity) dessas células permite a quantificação da expressão na superfície celular.

FIGURA 6 – ESTRATÉGIA DE *GATES* PARA QUANTIFICAR A EXPRESSÃO DE *KIR2DL1* POR CITOMETRIA.



FONTE: A AUTORA (2020)

LEGENDA: Na estratégia de *gates*, populações celulares são plotadas de acordo com diversos parâmetros, com o objetivo de isola-las das demais. (A) Na primeira imagem estão plotados o tamanho e a granulosidade das células do PBMC. Conforme a densidade de células ocupando um mesmo valor aumenta, a cor dos pontos passa de azul para vermelho. Nesse gráfico são selecionados os linfócitos, demarcados por um círculo vermelho. (B) Em seguida, estão plotadas a fluorescência de CD56 (PE-A) por CD3 (PE-Cy5-A) nos linfócitos. As células NK, que são positivas para CD56 e negativas para CD3, estão selecionadas em um quadrado verde (nomeado P4). (C) A partir do conjunto das células NK, são selecionadas as células positivas para o KIR2DL1 (FITC-A), em um quadrado vermelho nomeado P5. A comparação entre as células positivas e negativas nesse gráfico nos informa a porcentagem de células NK que expressam o KIR2DL1. (D) Quantificação da expressão nas células NK positivas para o KIR2DL1, dada pelo MFI.

Em um segundo ensaio de citometria, foi analisado um subconjunto de 30 indivíduos eurodescendentes de Curitiba. O protocolo experimental seguiu o padrão aqui descrito, porém foram modificados os anticorpos utilizados para quantificar a expressão de HLA-C. Nesse ensaio, as PBMC foram incubadas com: o anticorpo anti-HLA-C (clone DT9 sem fluorescência, BD Biosciences Cat# 566372, RRID: AB_2739715), anticorpo secundário anti-rato IgG2ab (clone X-57 conjugado ao FITC, Miltenyi Biotec Cat# 130-098-126, RRID: AB_2661521) e, por último, o anticorpo anti-CD3 (clone HIT3a conjugado ao PE/Cy5, Biolegend Cat# 300309, RRID: AB_314045). As células CD3+ formam cerca de 60% do PBMC e, ainda, todas as células nucleadas expressam o antígeno HLA-C. Portanto, devida a abundância de células positivas para o HLA-C no conjunto de células CD3+, foi definida a aquisição de 10.000 eventos por amostra.

4.6 GENOTIPAGEM DE GRUPOS ALÉLICOS DE HLA DE CLASSE I (-A, -B E -C)

Para as amostras não previamente genotipadas (BRAUN-PRADO et al., 2000), a genotipagem de grupos alélicos dos genes *HLA-C* na amostra de eurodescendentes de Curitiba foi realizada através de PCR alelo-específica seguida de curva de dissociação, utilizando oligonucleotídeos previamente descritos (HONG et al., 2011) modificados com nucleotídeos LNA (do inglês, *locked nucleic acid*). Os LNA se encontram na extremidade 3' dos oligonucleotídeos e tem a vantagem de possuir uma modificação no anel da ribose que confere maior estabilidade termal e especificidade ao oligonucleotídeo, o que permite uma melhor discriminação de substituições de bases únicas em posições variáveis. Foram utilizados na reação de PCR: Kapa 2G® fast master mix (Kapa Biosystems, África do Sul), 10 ng de DNA genômico, 20 μM de cada oligonucleotídeo alelo específico, e 10 μM de oligonucleotídeo controle para o gene da galactosylceramidase (*GALC*). A reação ocorreu por 3 minutos a 95°C, seguida de 35 ciclos alternando as temperaturas de 95°C por 15 segundos, 63°C por 15 segundos e 72°C por 1 segundo, e, finalmente, um ciclo de 72°C por 1 minuto.

4.7 SEQUENCIAMENTO DE NOVA GERAÇÃO DE TODOS OS GENES *KIR* E DE *HLA* DE CLASSE I

4.7.1 Análise da qualidade do DNA

Para o sucesso do sequenciamento de nova geração (NGS), é essencial que as amostras de DNA estejam íntegras e suficientemente livres de contaminação. Para condições, DNA foram avaliar essas as amostras de analisadas por espectrofotometria, usando o aparelho Thermo Scientific 1000 Spectrophotometer NanoDropTM (Thermo Fisher Scientific Inc., Wilmington, EUA), e por eletroforese em gel de agarose. Na espectrofotometria, o DNA é quantificado na absorbância de 260 nm, e outros contaminantes como proteínas a 280 nm. A pureza do DNA é dada pela razão das absorções de 260/280, e valores entre 1,8 e 2,2 foram considerados suficientemente puros. Para análise da integridade do DNA, foram aplicados em gel de agarose (1%) 2 µL de DNA genômico e 2 µL de solução de carreamento (azul de bromofenol + glicerol). As amostras foram submetidas a eletroforese com voltagem de 100V por 35 minutos. A eletroforese permite observar o DNA fragmentado, a partir da formação de rastros na corrida do DNA. O DNA é considerado integro caso apresente banda única, clara e definida no gel. A degradação parcial ou total, bem como baixa quantidade de DNA foram critérios de exclusão de unidades amostrais do estudo.

4.7.2 Sequenciamento de Nova Geração dos Genes KIR e HLA

O sequenciamento de nova geração para genotipagem alélica de alta resolução dos genes *KIR* e *HLA* de classe I clássicos e não clássicos foi feito na Universidade da Califórnia, São Francisco (UCSF), utilizando plataforma Illumina HiSeq 4000 (Illumina, San Diego, CA) a partir do método desenvolvido por NORMAN et al. (2016) e adaptado por Dr. Danillo Augusto, em 707 indivíduos das oito populações sul-americanas previamente descritas (detalhes da amostra na TABELA 2). Todos os genes e pseudogenes *KIR* são capturados e sequenciados por este método.

Para isso, o preparo da biblioteca foi feito inicialmente com a fragmentação enzimática do DNA em segmentos curtos utilizando o kit KAPA HyperPlus (Roche, USA). O DNA fragmentado foi ligado a adaptadores TS HT Dual Index contendo sequências identificadoras (processo conhecido como indexing, ou barcoding) do kit TruSight (Illumina, San Diego, CA). Os fragmentos ligados aos adaptadores foram selecionados pelo método de dupla seleção de tamanho utilizando microesferas magnéticas AMPure XP Beads (Beckman Coulter, USA), de forma a obter um tamanho médio de 780 pares de base. O DNA, uma vez fragmentado e indexado, foi amplificado e novamente purificado com beads magnéticas, para então ser feito o controle de qualidade. A qualidade foi verificada através de dois parâmetros: a quantificação, utilizando o kit PicoGreen® (Thermo Fisher Scientific, Waltham, MA), e a verificação do tamanho dos fragmentos, utilizando o sistema Bioanalyzer (Agilent, Santa Clara, CA). As amostras processadas foram reunidas proporcionalmente em um único tubo (individualizados pela sequência dos adaptadores) usando o manipulador de líquidos automatizado Echo ® 525 (Labcyte, San Jose, CA), formando os pools de bibliotecas para captura.

Os pools de amostras agrupadas foram concentrados por ultrafiltração, utilizando o protocolo dos kits Nextera (Illumina, San Diego, CA). Foram utilizadas 10.456 sondas biotiniladas (NORMAN et al., 2016) para capturar seletivamente e com alta afinidade fragmentos correspondentes aos genes KIR e HLA de classe I. Os fragmentos ligados às sondas foram selecionados pela ligação da biotina com SV-Magnetic Beads (Illumina, San Diego, CA). Após a captura, o produto foi amplificado (processo conhecido como enriquecimento) utilizando oligonucleotídeos específicos para a sequência dos adaptadores através do kit TruSight (Illumina, San Diego, CA). O produto foi novamente purificado com microesferas magnéticas SPRI Nextera (Illumina, San Diego, CA), para remover os reagentes da etapa de amplificação. Finalmente, a quantificação do produto capturado e amplificado foi realizada através de PCR digital em microgotas (ddPCR, do inglês, *digital droplet PCR*), para que então fosse feito o sequenciamento por síntese das bibliotecas na plataforma Illumina HiSeq 4000 (Illumina, San Diego, CA). Cada reação de sequenciamento por esse método permite sequenciar até 384 amostras de DNA simultaneamente. Sendo assim, por questões logísticas, as amostras desse trabalho foram subdivididas em 4 reações de sequenciamento.

O sequenciamento de alta cobertura enriquecido apenas para a região de genes *KIR* e *HLA* permite alta profundidade na leitura das sequências (média de 200 leituras por base). Assim, é possível obter tipagem alélica de alta resolução e identificar com alta sensibilidade novas variantes. Além disso, através do número de leituras de cada sequência gênica é possível quantificar o número de cópias dos genes *KIR*. Essa informação é essencial para a inferência de haplótipos.

4.8 ANÁLISE DOS DADOS

As análises estatísticas desse trabalho foram feitas predominantemente na linguagem de programação R (R CORE TEAM, 2019), através do programa de interface RStudio[™] (RSTUDIO TEAM, 2018) e estão detalhadas abaixo.

Para o conjunto de indivíduos que foram sequenciados com o método de Sanger, os resultados do sequenciamento foram alinhados com a sequência nucleotídica de referência através do programa Mutation Surveyor® (*SoftGenetics*, PA, USA). A inferência de alelos foi feita manualmente, com base na descrição de

alelos disponível no banco de dados IPD-KIR (ROBINSON et al., 2015). Os resultados de PCR em tempo real para determinar o número de cópias de *KIR2DL1* foram analisados de forma manual utilizando o método $2^{-\Delta\Delta CT}$. Os resultados do ensaio de citometria de fluxo foram gerados e exportados pelo software BD FacSuiteTM (*BD Biosciences*, San Jose, CA).

A partir da determinação alélica de *KIR2DL1*, foram calculadas a frequência alélica e a heterozigosidade média usando o pacote em R *genetics* (WARNES et al., 2019). O equilíbrio de Hardy-Weinberg foi testado pelo qui-quadrado de Person, usando o mesmo pacote. Com base nos 5 éxons analisados, foi possível chegar à resolução alélica de 4 dígitos em algumas das sequencias analisadas. Essas foram mantidas para a descrição das frequências alélicas de *KIR2DL1*, e posteriormente reduzidas à resolução de dois dígitos para os testes de associação com a expressão.

A normalidade dos parâmetros de expressão foi testada utilizando o teste de Kolmogorov-Smirnov, no pacote em R *nortest* (GROSS; LIGGES, 2015). O desvio da normalidade nesses parâmetros nos levou a utilizar testes não-paramétricos para realizar as associações. Para verificar se diferentes genótipos de SNPs em *KIR2DL1* estavam associados com uma diferença de expressão em HLA-C, utilizamos uma regressão logística linear com o pacote em R *stats* (R CORE TEAM, 2019), considerando sexo, genótipo do ligante HLA-C e o número de cópias de *KIR2DL1* como covariáveis. Esse teste apenas verifica a existência de diferença na distribuição de uma variável entre os grupos. Com o objetivo de verificar as diferenças cada par de genótipos (por exemplo, para um dimorfismo CA: CC *versus* AA, CC *versus* CA e AA *versus* CA), foi aplicada a análise pos-hoc de Dunn sobre o teste de Kruskal-Wallis no pacote em R *dunn.test* (DINNO, 2017).

A imputação *in silico* dos níveis de expressão de HLA-C foi feita a partir de dados publicados no estudo de APPS et al. (2013) em populações europeias e africanas. No estudo, foram caracterizados os níveis de expressão médio de 14 grupos alélicos (*C*01* ao **08, *12,* e **14* ao **18*) em células periféricas positivas para o CD3 (CD3+). Para cada indivíduo, foi identificado o nível de expressão associado a cada um dos seus alelos, e os dois valores foram somados. Nós fizemos uma correlação de Pearson entre os dados que imputamos com base no estudo de APPS et al. (2013) e a expressão *in vivo* em células CD3+, através do pacote em R *Hmisc* (JR; DUPONT, 2019). Os gráficos de expressão foram plotados utilizando os pacotes em R *base* e *beeswarm* (EKLUND, 2016; R CORE TEAM, 2019).

A partir dos dados genômicos de KIR2DL1 obtidos por NGS, o desequilíbrio de ligação entre SNPs na região promotora e codificadora foi calculado no pacote em R genetics (WARNES et al., 2019) e plotado utilizando uma versão modificada do pacote em R LDheatmap (SHIN et al., 2006). A fase dos SNPs no cromossomo foi inferida usando o programa fastPHASE (SCHEET; STEPHENS, 2006), com os parâmetros -T10 (define o número vezes em que o algoritmo EM é rodado de maneira aleatória) e -H200 (minimiza o erro por amostra). As sequências em fase foram utilizadas para estimar a ocorrência de haplótipos de homozigosidade estendida (EHH, do inglês, Extended Haplotype Homozigosity) (SABETI et al., 2002), com o pacote em R rehh (GAUTIER et al., 2017). Alelos ancestrais e derivados foram definidos de acordo com o banco de dados dbSNP (do inglês, Database of Single Nucleotide Polymorphisms). A predição in silico de sítios que se ligam a fatores de transcrição foi feita através da ferramenta LASAGNA-Search tool (LEE; HUANG, 2013), considerando os modelos de sítios de ligação a fatores de transição reportados no repositório ORegAnno (LESURF et al., 2016) e PAZAR (PORTALES-CASAMAR et al., 2007). Nesta predição, foi utilizado um valor máximo de significância de p < 0,001.

A análise dos dados do sequenciamento de nova geração para os genes KIR foi feita utilizando a pipeline PING (do inglês, Pushing Immunogenetics to the Next Generation) (NORMAN et al., 2016; MARIN et al., em preparação). Essa ferramenta realiza o alinhamento, filtragem, determinação do número de cópias e dos alelos dos genes KIR. Não foram incluídos nessa análise os pseudogenes KIR2DP1 e KIR3DP1. A pipeline possui quatro módulos principais: PING Extractor, PING GC caller, PING allele caller, e Haplo PING. O PING Extractor realiza filtragem de sequências correspondentes aos loci dos genes KIR de interesse. O PING GC caller utiliza o número de leituras de cada gene para determinar o número de cópias dos genes KIR. O módulo PING allele caller realiza a determinação alélica dos genes a partir do alinhamento da sequência codificadora com as sequências de referência disponíveis no banco de dados IPD-KIR (ROBINSON et al., 2015) e o módulo Haplo PING resolve ambiguidades a partir de desequilíbrio de ligação conhecidos entre alelos de diferentes genes. Ao final da análise, os resultados de saída do PING são integrados para produzir um único arquivo que descreve o conteúdo alélico e gênico de cada locus KIR, por indivíduo.

Estes mesmos dados de NGS foram analisados utilizando o programa HLA Explore™ (Omixon, Hungria) para obter a genotipagem de genes *HLA* de classe I, clássicos e não clássicos. O programa faz a filtragem, alinhamento e identificação de genótipos. A genotipagem é feita em alta resolução, até oito dígitos. Além disso, o programa faz a genotipagem até o 6° digito sem ocorrência de ambiguidades.

Os dados de KIR provenientes do programa PING foram conferidos manualmente, para eventual ajuste manual do número de cópias e ambiguidades. Os alelos de KIR2LD5 foram designados manualmente em KIR2DL5A e KIR2DL5B, que compreendem diferentes loci na região KIR, bem como KIR2DS3 e KIR2DS5 foram classificados em KIR2DS3 e KIR2DS5 (na região centromérica) e KIR2DS5 (na região telomérica). Esses alelos que foram previamente analisados como um só locus, podem ser designados em dois loci diferentes a partir da observação: do desequilíbrio de ligação observado nos haplótipos homozigotos em cada população; e dos haplótipos KIR em alta resolução publicados na literatura. As ambiguidades, a subclassificação dos alelos de KIR2DL5, KIR2DS3 e KIR2DS5, e a determinação de haplótipos foram resolvidas por inspeção manual e pela análise de desequilíbrio de ligação entre genes vizinhos em cada população. Para resolver as ambiguidades, também foram analisadas variantes intrônicas que não haviam sido incluídas na chamada de alelos do algoritmo PING. As análises descritivas de frequências alélicas foram feitas utilizando o pacote em R PopGenReport (ADAMACK; GRUBER, 2014). A determinação dos haplótipos foi feita manualmente por observação do desequilíbrio de ligação entre um gene e seus genes vizinhos, sucessivamente ao longo dos loci KIR e para cada um dos 707 indivíduos do estudo.

Além disso, nós criamos um script em linguagem R para classificar os genes *HLA-A, -B* e -*C* de acordo com os epítopos relevantes para interação com KIR (SIDNEY et al., 2008). *HLA-A* foi classificado em A3 e A11 (inclui todos os alelos *A*003* e *A*011*) e Bw4 (alelos *A*023, A*024* e *A*032*). Para *HLA-B*, foi recuperada a sequência proteica de cada alelo depositado no IPD-IMGT/HLA (ROBINSON et al., 2020), e as sequências foram classificadas em Bw4 (presença do epítopo SLRNLRG nas posições 77-83) e Bw6 (presença dos epítopos NLRIALR, DLRTLLR, SLRTLLR, NLRTALR ou SLRTALR nas posições 77-83). O Bw4 foi subclassificado ainda em epítopos que contém na posição 80 o aminoácido isoleucina (Bw4I), que está associado com maior inibição pelo KIR3DL1, ou treonina (Bw4T), com menor potencial inibidor (CARR et al., 2005). Finalmente, o *HLA-C* foi classificado utilizando a sequência proteica, de maneira similar ao *HLA-B*, porém, de acordo com o dimorfismo da posição 80 em asparagina (epítopo C1) e lisina (epítopo C2).

Esses dados foram integrados com os dados de KIR para produzir a pontuação média de interações KIR-HLA em cada população, como descrito por NEMAT-GORGANI et al. (2018). Nessa pontuação, são contabilizadas as diferentes interações KIR-HLA confirmadas através de estudos funcionais. Para cada interação considerada, a presença do gene KIR e do seu ligante HLA pode assumir os valores de 0 (alelos envolvidos na interação ausentes), 1 (alelos envolvidos na interação presentes em hemizigose – apenas para KIR – ou presentes em homozigose) ou 2 (alelos envolvidos na interação presente em heterozigose). São então multiplicados o valor de presença de KIR com o valor de presença de HLA de cada par funcional para gerar a pontuação de interação, que pode apresentar os valores 0, 1, 2 e 4. Os pares funcionais considerados são: Bw4 (HLA-A) e KIR3DL1 (FOLEY et al., 2008); Bw4 (HLA-B) e KIR3DL1 (GUMPERZ et al., 1995; FOLEY et al., 2008); HLA-A*03 e KIR3DL2 (DÖHRING et al., 1996); HLA-A*11 e KIR3DL2 (HANSASUTA et al., 2004); C2 e KIR2DL1 (HILTON et al., 2015); C1 e C2 e KIR2DL2 (HILTON et al., 2015); C1 e KIR2DL3 (HILTON et al., 2015); C2 e KIR2DS1 (HILTON et al., 2015); HLA-C*16 e KIR2DS2 (MOESTA et al., 2010); HLA-A*11 e KIR2DS4 (GRAEF et al., 2009); um subconjunto de HLA-C que engloba alelos C2 – C*0501, C*0202 e C*0401 – e alelos C1 – C*1601, C*0102 e C*1402 –, que são reconhecidos por KIR2DS4 (GRAEF et al., 2009); e C2 e um subconjunto de KIR2DS5 composto pelos alelos 2DS5*003, *004, *005, *006, *007 e *008 (BLOKHUIS et al., 2017).

Descritas essas interações, um exemplo para o cálculo da pontuação é: para a interação de Bw4+HLA-A e KIR3DL1, um indivíduo com os genótipos duplo heterozigotos HLA-A*024/HLA-A*012 (Bw4/Bw4) e KIR3DL1*001/KIR3DL1*002 teria a pontuação 4, pois existem 4 diferentes interações possíveis entre KIR e HLA no indivíduo. Em contrapartida, homozigotos HLA-A*024/HLA-A*024 (Bw4/Bw4) com heterozigose KIR3DL1*001/KIR3DL1*002 produziriam a pontuação 2. Neste caso, existem só 2 interações diferentes possíveis, já que o indivíduo é homozigoto. HLA-A*024/HLA-A*024 Finalmente, duplo homozigotos (Bw4/Bw4)е KIR3DL1*001/KIR3DL1*001 gerariam a pontuação 1, uma vez que só existe uma forma de interação ocorrendo. Quando há a ausência do gene KIR3DL1 e/ou de alelos que codifiquem o epítopo Bw4 em HLA-A é atribuído a pontuação 0 para a interação Bw4 (HLA-A) e KIR3DL1, pois não existe o par funcional. Ao final, são somadas todas as interações KIR-HLA presentes em cada indivíduo, e é calculada a média da população.

1 5 RESULTADOS

2 5.1 Single nucleotide polymorphism in *KIR2DL1* is associated with HLA-C

3 expression in global populations

4 Luciana de Brito Vargas¹, Renata M. Dourado¹, Leonardo M. Amorim¹, Brenda Ho²,

5 Verónica Calonga-Solís¹, Hellen C. Issler¹, Wesley M. Marin², Marcia H. Beltrame¹,

6 Maria Luiza Petzl-Erler¹, Jill A. Hollenbach², Danillo G. Augusto^{1,2*}

- ¹Programa de Pós-Graduação em Genética, Departamento de Genética, Universidade Federal
 do Paraná, Curitiba, Brazil
- 9 ² Weill Institute for Neurosciences, Department of Neurology, University of California, San
- 10 Francisco, CA, United States of America

11 * Correspondence:

- 12 Corresponding Author
- 13 danillo@augusto.bio.br

14 Keywords: copy number variation, natural selection, extended homozygosity, linkage 15 disequilibrium, coevolution, gene expression, population genetics

16 Abstract

17 Regulation of NK cell activity is mediated through killer-cell immunoglobulin-like receptors

18 (KIR) ability to recognize human leukocyte antigen (HLA) class I molecules as ligands.

19 Interaction of KIR and HLA is implicated in viral infections, autoimmunity and reproduction

20 and there is growing evidence of the coevolution of these two independently segregating gene

21 families. By leveraging *KIR* and *HLA-C* data from 1000 Genomes consortium we observed

22 that the *KIR2DL1* variant *rs2304224T* is associated with lower expression of HLA-C in

individuals carrying the ligand HLA-C2 (p = 0.0059). Using flow cytometry, we

24 demonstrated that this variant is also associated with higher expression of KIR2DL1 on the

- 25 NK cell surface (p = 0.0002). Next, we applied next generation sequencing to analyze
- 26 *KIR2DL1* sequence variation in 109 Euro and 75 Japanese descendants. Analyzing the

extended haplotype homozygosity, we show signals of positive selection for rs4806553G and

rs687000G, which are in linkage disequilibrium with *rs2304224T*. Our results suggest that

29 lower expression of HLA-C2 ligands might be compensated for higher expression of the

30 receptor KIR2DL1 and bring new insights into the coevolution of *KIR* and *HLA*.

31 **1 Introduction**

32 The killer cell immunoglobulin-like receptor (KIR) genes on chromosome 19 encode

33 receptors that interact with a subset of human leukocyte antigen (HLA) class I molecules,

34 encoded by genes on chromosome 6, to regulate NK cell cytotoxicity cells against infected

and neoplastic cells (1–3). In fact, combinations of variants of *KIR* and *HLA* have been

- 36 repeatedly associated with autoimmune disease (4-6), cancer (7,8), viral infections (9,10), and
- 37 are also implicated in placentation and reproduction (11-14). As a result, the interaction of
- 38 KIR and HLA is relevant to fitness and survival, and therefore, candidate for evolutionary 30 studies (15)
- 39 studies (15).
- 40 KIR recognizes some subsets of HLA-A (A3, A11 and Bw4), HLA-B (Bw4 and Bw6), and all
- 41 subsets of HLA-C (C1 and C2) molecules (16). *HLA-C*, present in all humans, appears to
- 42 have had a great impact on *KIR* evolution, driving the expansion of lineage III KIR, which are
- 43 the receptor lineage that recognize HLA-C (17,18). The dimorphism in position 80 of HLA-C defines HLA C1 (80^{ASB}) and HLA C2 (80^{LVS}) and confers differential energiative KLB
- 44 defines HLA-C1 (80^{Asn}) and HLA-C2 (80^{Lys}) and confers differential specificity to KIR. 45 Among all ligands, the interaction between KIB2DL1 and III A C2 is many arbitrary to KIR.
- Among all ligands, the interaction between KIR2DL1 and HLA-C2 is responsible for the
 strongest regulatory signal and HLA-C seems to act as the main educator of NK cells (19,20).
- stongest regulatory signal and meric seems to act as the main educator of the cells (19,2
- 47 Worldwide studies demonstrate coordinated frequencies of KIR and HLA genes in
- 48 populations. In a comprehensive study consisting of 30 populations, Single et al. (21) found
- 49 that increasing frequencies of activating KIR are correlated with decreased frequencies of
- 50 their respective HLA ligands. On the other hand, Hollenbach et al. (22) showed positive
- 51 correlation between the presence of KIR2DL3 and the presence of HLA-C1 in 105 worldwide 52 normalations A strong and population of KIR parameters the latter $A_{\rm ex}$ in the strong strong that is a strong strong three strong strong three strong st
- 52 populations. A strong and negative correlation of *KIR* gene-content haplotype *A* and HLA-C2, 53 a pair which is associated with increased risk of pre-eclampsia, was found in eight populations
- from European, African and Asian ancestries (11). Moreover, there is extensive evidence of
- 55 balancing selection maintaining diversity in *KIR* genes (23–25). *KIR* and *HLA* segregate
- 56 independently and there are no reports of gametic association between these two gene
- 57 families. Here, we show that a single nucleotide polymorphism (SNP) in *KIR2DL1* is
- associated with expression levels of the KIR2DL1 receptor on the cell surface and also with
- 59 HLA-C expression. In addition, we present evidence of positive selection in linked *KIR2DL1*
- 60 variants.

61 2 Results

62 2.1 *KIR2DL1* variant *rs2304224T* is associated with lower expression levels of HLA-C

- 63 To search for possible signals of coevolution between *KIR* and *HLA*, we evaluated if variants
- 64 in inhibitory KIR that bind to HLA-C could be associated with HLA-C expression levels in
- 65 global populations. We leveraged the public sequencing information available for all
- 66 populations in the 1000 Genomes Project (1KGP) (26) and retrieved the genotypic data
- 67 available for SNPs located within *KIR2DL1* and *KIR2DL23* (*rs2304224*, *rs11673144*,
- 68 *rs12982263, rs34721508, rs35719984* and *rs35861855*) in 955 individuals of various
- 69 ancestries. We also obtained *HLA* genotyping data available for those individuals (27).
- 70 We next used previously published data of HLA-C expression levels (28) and imputed the
- expression for each *HLA-C* genotype in the 1KGP cohort. The variant rs2304224T was
- associated with lower HLA-C expression levels in individuals *HLA-C1/C2* (p = 0.0420),
- 73 HLA-C2/C2 (p = 0.0059) but not in HLA-C1/C1 individuals (p = 0.0740) (Suppl. Fig. 1). We
- replicated these results by imputing the HLA expression in an independent panel of 308

- 75 Brazilians Euro-descendants for which *HLA* genotyping data was available, and we sequenced 76 the first exon of *KIR2DL1* to genotype rs2304224 (p = 0.0107; Figure 1B).
- 77 To demonstrate that our approach to impute the HLA-C expression is predictive of the cell
- 78 surface expression in vivo, we measured the HLA-C surface levels of fresh CD3+ cells in 30
- individuals using flow cytometry and compared to the imputed values. We found a correlation $(2 1)^{-1} = (2$
- 80 of r = 0.62, p < 0.0001 (Suppl. Fig. 2).

81 2.2 *rs2304224T* is also associated with higher surface expression levels of KIR2DL1

- 82 We sought to investigate if the variant *rs2304224T* in *KIR2DL1* was associated with
- 83 KIR2DL1 surface expression. We used flow cytometry to quantify both the abundance of
- 84 KIR2DL1 on the surface of NK cells (median fluorescence intensity, MFI) as well as the
- 85 percentage of NK cells expressing KIR2DL1 on their surface (KIR2DL1⁺ NK cell), and also
- 86 interrogated if copy number variation of KIR2DL1 affects surface expression. Although
- 87 borderline, we did not find significant differences of expression levels in individuals carrying
- one copy (hemizygous) or two copies (homo- or heterozygous) of KIR2DL1 (p = 0.0594;
- 89 Suppl. Fig. 3A). However, the number of KIR2DL1+ NK cells was 2.16-fold higher in
- individuals carrying two copies (p = 0.0001) (Suppl. Fig. 3B). For all KIR2DL1 expression
- analyses, we used copy number of *KIR2DL1* as covariant in the regression model.
- 92 We observed that the allele *rs2304224T*, associated with decreased HLA-C expression, was
- also associated with 1.54-fold increase of the KIR2DL1 surface expression (p = 0.0002) and a
- 94 1.41-fold increase of KIR2DL1⁺ NK cells (p = 0.03) (Figures 1C-D). The median expression
- of each KIR allotype is shown in Suppl. Figure 4. We also observed that KIR2DL1
- 96 expression was decreased in individuals homozygous for the C2 ligand (C2/C2, p = 0.007)
- 97 (Figure 1E).

98 99 2.3 Signals of positive selection for *KIR2DL1* variants in linkage disequilibrium with *rs2304224*

- 100 We next analyzed the entire *KIR2DL1* gene in a subset of 109 Euro-descendants and 75
- 101 Japanese descendants sequenced using our custom next generation sequencing method (29).
- 102 In Euro descendants, we observed low correlation but strong linkage disequilibrium (LD)
- 103 between rs2304224 and three other variants (Sup. Fig. 5A-B). The first variant is at position -
- 104 406 upstream of the *KIR2DL1* gene (*rs4806553*, D'=0.99, r^2 = 0.18, p< 10⁻⁸). The other
- 105 variants are located within the coding region, in exon 4 (rs687000, D'=0.99, r²=0.52, $p<10^{-12}$)
- 106 and exon 7 (rs34721508, D'=0.99, r²=0.24, $p<10^{-3}$). Weaker LD was observed for the same
- 107 variants in Japanese descendants (Suppl. Fig. 5C-D). Frequencies for all SNPs in both
- 108 populations are given in Suppl. Table 1. Moreover, the frequency of HLA-C2 in our Japanese-
- 109 descendant cohort was 10.3% while in Euro-descendants it was 40.9%.
- 110 We searched for signals of population specific selection, for both Euro and Japanese
- descendants, by estimating the extended haplotype homozygosity (EHH) using rs2304224
- and also variants in significant LD with it as focal SNPs. The bifurcation patterns are
- 113 consistent with positive selection increasing frequencies of the haplotype more rapidly than
- 114 they could be broken by genetic recombination. Signals of positive selection were observed
- 115 for the derived allele *rs4806553G* in Japanese but not in Euro-descendants (Fig 2A and 2C).
- 116 Strong signals of positive selection were also observed for the derived allele *rs687000G* in
- 117 both Euro and Japanese descendants (Fig 2B and 2D).

118 **3** Discussion

119 Previous results show that cis polymorphisms associated with *HLA-C* expression do not

associate with NK cell activity (30), despite the compelling evidence that KIR-HLA are

121 coevolving as an integrated system (11,16,21,22). Here, we show evidence of coevolution of

KIR and *HLA* by identifying a variant in *KIR2DL1* that was associated with surface
 expression of the HLA-C2 ligand in worldwide populations. The allele *rs2304224T* was

associated with lower expression of imputed HLA-C surface expression in 995 individuals

from 1KGP and also in an independent cohort of 308 Brazilian Euro-descendants. The

association was only observed in individuals carrying at least one copy of the ligand HLA-C2,

127 which suggests an orchestrated and refined evolution between these two systems. Our direct

128 measurement of HLA-C expression in a subset of 30 individuals demonstrates that imputing

HLA expression based on previously published data is predictive of the expression observedon the surface of fresh blood cells.

131 It is also interesting that the same allele *rs2304224T* is associated with higher expression of

the receptor KIR2DL1 in NK cells. The SNP *rs2304224* in exon 1 causes a non-synonymous

133 substitution of value (allele G) to phenylalanine (allele T) in the signal peptide. The

134 hydrophobicity of the signal peptide can influence protein retention in the cytosol (31).

According to the Wimley-White interfacial hydrophobicity scale (32), valine has a free energy

136 of transfer of 0.07 ΔG from water to bilayer, and the free energy of phenylalanine is -1.13 ΔG .

137 The lower and negative value of phenylalanine indicates this transference is more favorable, 128 and therefore $m^{220/22} 4T$ meaning and the membrane. This could

and therefore, *rs2304224T* may increase protein availability in the membrane. This could
 explain the increased KIR2DL1 expression associated with *rs2304224T*.

140 The patterns that we observed for the expression of KIR2DL1 allotypes (Suppl. Fig. 4) are

141 consistent with previous studies (20,33–35). Our results showing that copy number of

142 *KIR2DL1* affects the quantity of KIR2DL1⁺ NK cells corroborate those by Béziat et al. (35).

143 On the other hand, the lack of significant association that we observed between *KIR2DL1*

144 copy number and the abundance of expression on the cell surface reinforces the idea that copy

number does not affect levels of KIR2DL1 as strongly as it affects the proportion of cells

expression the receptor (35). The presence of HLA-C2 was associated with lower expression

of surface KIR2DL1, according to our results and of others (33,36,37). However, differently

from the observations from Le Luduec *et al.* (36), who observed that the expression of KIR2DL1 is associated to the presence of C2 in a dose dependent manner, we found

149 KIR2DL1 is associated to the presence of C2 in a dose dependen 150 association only in individuals carrying two copies of C2.

151 We found three SNPs in LD with rs2304224 (D' = 0.99, $0.18 \le r^2 \le 0.51$). The low correlation

152 coefficient is explained by difference in the allele frequencies among them. The frequency of

153 the variant rs2304224T is 0.26 in Euro-Brazilians, while the frequency of rs4806553C is 0.67;

154 *rs687000A* is 0.57; and *rs34721508C* is 0.86. From the three variants in LD with *rs2304224*,

155 only *rs34721508*, in exon 7, has been previously associated with differential expression levels

of KIR2DL1 in transfected cell lines (34). That study showed that cells expressing allotypes

with 245^{Cys} have reduced protein stability and are more susceptible to ligand mediated

expression down-regulation in comparison to those with 245^{Arg} . Interestingly, this variant was also present in the 1KGP dataset, and we did not observe association of *rs34721508*

160 genotypes with HLA-C imputed expression levels (p = 0.28). We also demonstrated that there

is an additive effect of rs2304224T and rs34721508C on KIR2DL1 expression, which

162 indicates that each has independent effect on the expression of KIR2DL1 (Suppl. Fig. 6).

163 We applied extended haplotype homozygosity (EHH) analysis to all SNPs in LD with

- 164 rs2304224, using the next generation sequencing data that we generated for a subset of Euro
- and Japanese descendants. Homozygosity surrounding the derived allele rs4806553G is
- 166 prominent in the Japanese population, suggesting this allele has been under recent positive
- 167 selection. Japanese populations are especially interesting because they exhibit the lowest
- 168 frequency of the HLA-C2 allotype (only 8%) (38) and, accordingly, we report low frequency
- 169 of C2 also in the Brazilians of Japanese ancestry (10.3%). The low frequency of HLA-C2 170 appld hadriving the qualitien of KIR2DIL in the Japanese nonvelocion
- 170 could be driving the evolution of *KIR2DL1* in the Japanese population.
- 171 The SNP *rs4806553* is located in a predicted transcription factor (TF) binding site in
- *KIR2DL1*, according to LASAGNA-Search tool (39). The transcription factor E2F4 is
 predicted to bind to the sequence containing the allele *rs4806553G* (PAZAR 1874, combined
- affinity score = 184.89, p < 0.01), while the *rs4806553C* disrupts this site. E2F4 is a member
- of the E2F family; it represses gene expression of its target genes (40) and is expressed
- 176 mainly in the bone marrow and placenta (41). Loss of E2F4 has been associated to neonatal
- 177 lethality, especially due to dehydration and starvation (42), showing that this TF play an
- important role in reproduction. If this binding site on *KIR2DL1* is further confirmed in
- functional studies, E2F4 could be repressing *KIR2DL1* in the presence of *rs4806553G*.
- 180 Decreased KIR2DL1 expression, especially in the uterine NK (uNK) cells in the maternofetal
- 181 interface, would allow greater production of cytokine and growth factors by uNK, which has
- 182 been shown to promote spiral artery transformation and trophoblast invasion (43). Therefore,
- 183 the selection on rs4806553G could potentially be related to a reproductive advantage.
- 184 Although speculative, this is one example of the possible reasons why rs4806553G could be
- 185 implicated in KIR2DL1 expression and also being favored by positive selection in the
- 186 Japanese population.
- 187 Strong signals of positive selection were observed towards the derived allele rs687000G in 188 both the Euro and Japanese descendants. This variant is a synonymous change in exon 4, with 189 no apparent impact on expression or regulation of KIR2DL1. One hypothesis is that 190 rs687000G rose in frequency due to hitchhiking with a nearby variation that was positively 191 selected and eventually fixed. We did not observe signals of positive selection for rs2304224 and rs34721508, which strongly associate with KIR2DL1 expression levels. One possibility is 192 193 that selection could be favoring specific KIR2DL1 alleles that carry these variants. In fact, the 194 combination of rs2304224G (neutral), rs687000G (positively selected) and rs34721508C 195 (neutral) defines KIR2DL1*003, the most frequent allele across all populations worldwide 196 (44).
- 197 Coevolution of KIR and HLA is mostly driven by HLA-C (20,45), which encodes a strong 198 educator for KIR⁺ NK cells (46,47). A fine tuning mechanism of NK cell regulation through the cell-specific promoter NK-Pro (48) was recently proposed, in which expression levels of 199 200 HLA-C during NK cell education combines with expression levels and interaction strength of 201 KIR and HLA in mature NK cells to modulate their selectivity and cytotoxicity (49). 202 KIR2DL1 is the receptor with the highest affinity and avidity to HLA-C, and mediates the 203 strongest NK response (19,20,50). Therefore, it is plausible that variation in KIR2DL1 could 204 be under selection and also that KIR2DL1 and HLA-C are coevolving. Here, we show a 205 KIR2DL1 variant that is associated with lower expression of KIR2DL1 and inversely 206 associated with higher HLA-C expression in HLA-C2/C2 individuals. This could be an 207 indication that higher levels of the ligand are being compensated by lower expression of the 208 receptor. We also observed evidence of positive selection on KIR2DL1. Our data show that 209 much remains to be understood regarding the mechanisms of the KIR-HLA recognition and

- 210 evolution. They also bring insights into the evolution of these two systems and suggest that
- 211 more questions will emerge as we explore more deeply KIR-HLA diversity at high resolution.

212 4 Material and methods

213 4.1 Samples

214 A cohort of 308 individuals of predominantly European ancestry and 75 individuals of 215 Japanese ancestry from Curitiba, Brazil, were analyzed in this study. About 80% of the 216 population from Curitiba self-reported as White (51), in accordance with previous genetic 217 studies (52). For the Japanese descendants, we only included individuals who had two parents 218 or four grandparents born in Japan, with no history of admixture with non-Japanese 219 ancestries. In order to measure KIR2DL1 expression levels, we analyzed fresh blood cells 220 from a subset of 48 Euro-descendants. A subset of 30 of these individuals were included in 221 the HLA-C expression assay. Detailed information about the study design is given in Suppl. 222 Figure 7. All individuals were living in Curitiba, Brazil, at the time of blood collection. 223 Median age in the group was 26 years (ranging from 20 to 64) and the male/female ratio was 224 0.37.

For expression assays, we collected 8 mL of peripheral blood samples and isolated PBMC

226 (peripheral blood mononuclear cells) using LeucosepTM tubes (Greiner Bio-One, Austria),

227 which have a selective membrane for density-based lymphocyte separation, and Ficoll

228 Hypaque (Sigma Aldrich, MO). Isolated PBMC were counted in a Neubauer chamber under

229 an optical microscope. A total of 0.5×10^6 cells were incubated with specific antibodies for

- KIR2DL1 and HLA-C and analyzed by flow cytometry. Detailed description and gate strategyare shown in Suppl. Figure 8.
- 251 are shown in Suppl. 1 igure 6.
- All participants were informed of the research's purpose and agreed to participate through a
- 233 written informed consent, in accordance with Brazilian regulation. This study is part of the
- project registered under the number 02727412.4.00000096 and approved by the Human
- Research Ethics Committee of the Federal University of Paraná and the Brazilian National
- Commission on Research Ethics (Conep) under the number 2970200 (01/02/2014).

237 4.2 *KIR2DL1* and *HLA-C* genotyping

We initially sequenced exons 1, 4, 5, 7 and 9 to distinguish the main *KIR2DL1* allele groups using the Sanger method (53) in the 48 Euro-descendants included in the expression assay (Suppl. Figure 9). The sequences obtained were aligned with reference sequences from IPD-KIR database (44), using the software Mutation Surveyor® (SoftGenetics, PA) and identified manually. Additionally, we sequenced only the exon 1 (containing the variant *rs2304224*) in extra 260 Euro-descendant individuals to increase statistical power for the analysis of *rs2304224*.

- We applied quantitative PCR to determine copy number of *KIR2DL1* compared to *KIR3DL3*, which is present in virtually all haplotypes. *KIR2DL1* was amplified in triplicate using one set of primers and the reference gene *KIR3DL3* was amplified using other three sets of primers, each in triplicate, in a total of 12 (4 x 3) reactions per sample. The sequence of all primers used for amplification, sequencing and copy number assay, including those designed in this study as well as those described previously (54–58) are given in Supplementary Table 2.
- We also sequenced the entire *KIR2DL1* gene in 109 Euro-descendants and 75 Japanese descendants from Curitiba, Brazil. These samples were sequenced using the previously

published method for next generation sequencing of *KIR* and *HLA* genes (29) using Illuminaplatform.

255 4.3 Data analysis

256 Normality of variables was tested using Kolmogorov-Smirnov test, in R package nortest (59). 257 Difference in HLA-C expression between KIR2DL1 SNP genotypes was tested via the 258 Kruskal-Wallis test, using stats R (60). Pos-hoc analysis of Dunn was applied to Kruskal-259 Wallis results in order to identify pair-wise significant differences between genotypes, in R 260 package dunn.test (61). Median HLA-C expression by allele, as defined by Apps et al. (28), was imputed for each allele in an individual, and then summed. The imputation was 261 262 performed in all 308 Brazilians of European ancestry sequenced for rs2304224 and 1KGP 263 individuals. Correlation analysis between expected HLA-C expression in CD3⁺ cells and *in* 264 vivo HLA-C expression in CD3+ cells was calculated with R package corr. Difference in 265 KIR2DL1 expression according to copy number was tested using Mann-Whitney, in stats R (60). Association of KIR2DL1 expression with allotype and *rs2304224* was tested through 266 267 logistic regression using copy number as a covariate, also in *stats* R. Linkage disequilibrium 268 was estimated using LD function from R package genetics (62) and plotted with a modified 269 version of R package LDheatmap (63). Median expression graphs were plotted using base and 270 beeswarm R packages (60,64).

KIR2DL1 SNPs obtained from genomic sequence data were phased using fastPHASE, with
modified parameters (-T10 -H200). The phased data was used for estimation of extended
haplotype homozygosity (EHH) (65) using R package *rehh* (66). Ancestral and derived alleles
were defined according to the Database of Single Nucleotide Polymorphisms (dbSNP) (67).

275 **5** Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

278 **6** Author contributions

DGA designed the study; LV, RMD, VCS, LMA and HCI performed Sanger sequencing and
genotyping; LV, DGA and BH performed next generation sequencing; LV, RMD and DGA
performed flow cytometry analysis; LV, DGA and WM analyzed the data; MLPE, JAH and
DGA contributed with samples and/or reagents; LV and DGA drafted the manuscript. All
authors critically revised the manuscript.

284 7 Funding

285 This work was supported by grants and scholarships from Coordenação de Aperfeiçoamento

- 286 de Pessoal de Nível Superior Brasil (CAPES) Finance Code 001, Conselho Nacional de
- 287 Desenvolvimento Científico e Tecnológico (CNPq), Fundação Araucária and NIH grant
- 288 U19NS095774.

289 8 Acknowledgments

We would like to thank all the volunteers who participated in this study and all the members
of the Human Molecular Laboratory of the Federal University of Paraná, Brazil, for their

292 unconditional support.

293 9 References

1.

294

295

296 the killer cell. Eur J Immunol. 1975;5(2):117–21. 297 Waggoner SN, Reighard SD, Gyurova IE, Cranert SA, Mahl SE, Karmele EP, et al. 2. 298 Roles of natural killer cells in antiviral immunity. Curr Opin Virol. 2016;16:15–23. 299 Ciccone E, Pende D, Viale O, Di Donate C, Tripodi G, Orengo AM, et al. Evidence of 3. 300 a natural killer (NK) cell repertoire for (Allo) antigen recognition: Definition of five 301 distinct NK-determined allospecificities in humans. J Exp Med. 1992;175(3):709-718. Van der Slik AR, Koeleman BPC, Verduijn W, Bruining GJ, Roep BO, Giphart MJ. 302 4. KIR in type 1 diabetes: Disparate distribution of activating and inhibitory natural killer 303 304 cell receptors in patients versus HLA-matched control subjects. Diabetes. 305 2003;52(10):2639-42. Augusto DG, Lobo-Alves SC, Melo MF, Pereira NF, Petzl-Erler ML. Activating KIR 306 5. 307 and HLA Bw4 ligands are associated to decreased susceptibility to pemphigus 308 foliaceus, an autoimmune blistering skin disease. PLoS One. 2012;7(7). 309 6. Nelson GW, Martin MP, Gladman D, Wade J, Trowsdale J, Carrington M. Cutting 310 Edge: Heterozygote Advantage in Autoimmune Disease: Hierarchy of 311 Protection/Susceptibility Conferred by HLA and Killer Ig-Like Receptor Combinations 312 in Psoriatic Arthritis. J Immunol. 2004;173:4273-6. 313 7. Jobim MR, Jobim M, Salim PH, Portela P, Jobim LF, Leistner-Segal S, et al. Analysis 314 of KIR gene frequencies and HLA class I genotypes in breast cancer and control group. 315 Hum Immunol. 2013;74(9):1130-3. Middleton D, Diler AS, Meenagh A, Sleator C, Gourraud PA. Killer immunoglobulin-316 8. 317 like receptors (KIR2DL2 and/or KIR2DS2) in presence of their ligand (HLA-C1 group) protect against chronic myeloid leukaemia. Tissue Antigens. 2009;73(6):553-318 319 60. 320 9. Martin MP, Qi Y, Gao X, Yamada E, Martin JN, Pereyra F, et al. Innate partnership of HLA-B and KIR3DL1 subtypes against HIV-1. Nat Genet. 2007;39(6):733-740. 321 322 Khakoo SI, Thio CL, Martin MP, Brooks CR, Gao X, Astemborski J, et al. HLA and 10. NK cell inhibitory receptor genes in resolving hepatitis C virus infection. Science. 323 324 2004;305(5685):872-4. 325 11. Hiby SE, Walker JJ, O'Shaughnessy KM, Redman CWG, Carrington M, Trowsdale J, 326 et al. Combinations of Maternal KIR and Fetal HLA-C Genes Influence the Risk of 327 Preeclampsia and Reproductive Success. J Exp Med. 2004;200(8):957–965. 328 12. Trowsdale J, Moffett A. NK receptor interactions with MHC class I molecules in 329 pregnancy. Semin Immunol. 2008;20(6):317-20. 330 13. Nakimuli A, Chazara O, Hiby SE, Farrell L, Tukwasibwe S, Jayaraman J, et al. A KIR 331 B centromeric region present in Africans but not Europeans protects pregnant women 332 from pre-eclampsia. Proc Natl Acad Sci. 2015;112(3). 333 14. Bulmer JN, Lash GE. The role of uterine NK cells in normal reproduction and 334 reproductive disorders. Adv Exp Med Biol. 2015;868:95–126. 335 15. Parham P. MHC class I molecules and KIRS in human history, health and survival. Nat 336 Rev Immunol. 2005;5(3):201-14. 337 16. Augusto DG, Petzl-Erler ML. KIR and HLA under pressure: evidences of coevolution 338 across worldwide populations. Hum Genet. 2015;134(9):929–40. 339 17. Older Aguilar AM, Guethlein LA, Adams EJ, Abi-Rached L, Moesta AK, Parham P. 340 Coevolution of Killer Cell Ig-Like Receptors with HLA-C To Become the Major 341 Variable Regulators of Human NK Cells. J Immunol. 2010;185(7):4238-51.

Kiessling R, Klein E, Pross H, Wigzell H. "Natural" killer cells in the mouse. II.

Cytotoxic cells with specificity for mouse Moloney leukemia cells. Characteristics of

342 18. Parham P, Guethlein LA. Genetics of Natural Killer Cells in Human Health, Disease, 343 and Survival. Annu Rev Immunol. 2018;36(1):519-48. 344 19. Stewart CA, Laugier-Anfossi F, Vely F, Saulquin X, Riedmuller J, Tisserant A, et al. 345 Recognition of peptide-MHC class I complexes by activating killer immunoglobulin-346 like receptors. Proc Natl Acad Sci. 2005;102(37):13224-13229. 347 20. Hilton HG, Guethlein LA, Goyos A, Nemat-Gorgani N, Bushnell DA, Norman PJ, et 348 al. Polymorphic HLA-C Receptors Balance the Functional Characteristics of KIR 349 Haplotypes. J Immunol. 2015;195(7):3160-3170. Single RM, Martin MP, Gao X, Meyer D, Yeager M, Kidd JR, et al. Global diversity 350 21. 351 and evidence for coevolution of KIR and HLA. Nat Genet. 2007;39(9):1114-9. 352 Hollenbach JA, Augusto DG, Alaez C, Bubnova L, Fae I, Fischer G, et al. 16th IHIW: 22. 353 Population Global Distribution of Killer Immunoglobulin-like Receptor (KIR) and 354 Ligands. Int J Immunogenet. 2013;40(1):39-45. 355 23. Augusto DG, Norman PJ, Dandekar R, Hollenbach JA. Fluctuating and geographically specific selection characterize rapid evolution of the human Kir region. Front Immunol. 356 357 2019;10:989. 358 24. Gendzekhadze K, Norman PJ, Abi-Rached L, Layrisse Z, Parham P. High KIR 359 diversity in Amerindians is maintained using few gene-content haplotypes. 360 Immunogenetics. 2006;58(5-6):474-80. 361 25. Nemat-Gorgani N, Edinur HA, Hollenbach JA, Traherne JA, Dunn PPJ, Chambers GK, 362 et al. KIR diversity in Maori and Polynesians: populations in which HLA-B is not a 363 significant KIR ligand. Immunogenetics. 2014;66(11):597-611. 364 26. The 1000 Genomes Project Consortium. A global reference for human genetic 365 variation. Nature. 2015;526(7571):68-74. 366 27. Gourraud PA, Khankhanian P, Cereb N, Yang SY, Feolo M, Maiers M, et al. HLA 367 diversity in the 1000 genomes dataset. PLoS One. 2014;9(7):e97282. 368 28. Apps R, Qi Y, Carlson JM, Chen H, Gao X, Thomas R, et al. Influence of HLA-C 369 expression level on HIV control. Science (80-). 2013;340(6128):87-91. 370 29. Norman PJ, Hollenbach JA, Nemat-Gorgani N, Marin WM, Norberg SJ, Ashouri E, et 371 al. Defining KIR and HLA Class I Genotypes at Highest Resolution via High-372 Throughput Sequencing. Am J Hum Genet. 2016;99(2):375-391. 373 30. Charoudeh HN, Schmied L, Gonzalez A, Terszowski G, Czaja K, Schmitter K, et al. 374 Quantity of HLA-C surface expression and licensing of KIR2DL+ natural killer cells. 375 Immunogenetics. 2012;64(10):739-45. 376 31. Zhang L, Leng Q, Mixson AJ. Alteration in the IL-2 signal peptide affects secretion of 377 proteins in vitro and in vivo. J Gene Med. 2005;7(3):354-65. 378 32. Wimley WC, White SH. Experimentally determined hydrophobicity scale for proteins 379 at membrane interfaces. Nat Struct Biol. 1996;3(10):842-8. 380 33. Dunphy SE, Guinan KJ, Chorcora CN, Jayaraman J, Traherne J a, Trowsdale J, et al. 381 2DL1, 2DL2 and 2DL3 all contribute to KIR phenotype variability on human NK cells. 382 Genes Immun. 2015;16(5):301-10. 383 34. Bari R, Bell T, Leung WH, Vong QP, Chan WK, Das Gupta N, et al. Significant 384 functional heterogeneity among KIR2DL1 alleles and a pivotal role of arginine245. 385 Blood. 2009;114(25):5182-90. 386 35. Béziat V, Traherne JA, Liu LL, Jayaraman J, Enqvist M, Larsson S, et al. Influence of KIR gene copy number on natural killer cell education. Blood. 2013;121(23):4703-7. 387 Le Luduec JB, Boudreau JE, Freiberg JC, Hsu KC. Novel approach to cell surface 388 36. 389 discrimination between KIR2DL1 subtypes and KIR2DS1 identifies hierarchies in NK 390 repertoire, education, and tolerance. Front Immunol. 2019;10:734. 391 37. He Y, Tao S, Ying Y, He J, Zhu F, Lv H. Allelic polymorphism, mRNA and antigen

392		expression of KIR2DL1 in the Chinese Han population. Hum Immunol.
393		2014;75(3):245–9.
394	38.	Yawata M, Yawata N, Draghi M, Little A-M, Partheniou F, Parham P. Roles for HLA
395		and KIR polymorphisms in natural killer cell repertoire selection and modulation of
396		effector function. J Exp Med. 2006;203(3):633-45.
397	39.	Lee C, Huang CH. LASAGNA-search: An integrated web tool for transcription factor
398		binding site search and visualization. Biotechniques. 2013:54(3):141–53.
399	40.	Trimarchi JM, Lees JA, Sibling rivalry in the E2F family. Nat Rev Mol Cell Biol.
400		2002:3(1):11–20.
401	41.	Fagerberg L. Hallstrom BM, Oksvold P, Kampf C, Diureinovic D, Odeberg J, et al.
402		Analysis of the human tissue-specific expression by genome-wide integration of
403		transcriptomics and antibody-based proteomics. Mol Cell Proteomics. 2014:13(2):397–
404		406
405	42	Humbert PO Rogers C. Ganiatsas S. Landsberg RL. Trimarchi IM. Dandanani S. et al.
406	12.	F2F4 is essential for normal erythrocyte maturation and neonatal viability. Mol Cell
407		2000.6(2).281-91
408	43	Faas MM de Vos P. Uterine NK cells and macronhages in pregnancy. Placenta
100	чу.	2017.56.44 52
407	11	Robinson I Halliwell IA Havburst ID Flicek P Parham P Marsh SGF The IPD and
411		IMGT/HI & database: Allele variant databases Nucleic Acids Res 2015:43:D423_
412		D431
413	45	Nemat-Gorgani N Hilton HG Henn BM Lin M Gignoux CR Myrick IW et al
414	ч.Э.	Different Selected Mechanisms Attenuated the Inhibitory Interaction of KIR2DI 1 with
-1 /15		$C_2 + HI A_C$ in Two Indigenous Human Populations in Southern Africa. I Immunol
416		22 + 112A = 0 in 1 wo indigenous framan ropulations in Southern Africa : 5 initiation. $2018 \cdot 200(8) \cdot 2640 = 55$
417	46	David G Diaoud 7 Willem C Legrand N Rettman P Gagne K et al Large Spectrum
418	40.	of HI A-C Recognition by Killer Ig-like Recentor (KIR)2DI 2 and KIR2DI 3 and
410		Restricted C1 Specificity of KIR2DS2: Dominant Impact of KIR2DI 2/KIR2DS2 on
420		KIR2D NK Cell Repertoire Formation 1 Immunol 2013:191(9):4778–88
421	47	Horowitz A Diaoud Z Nemat-Gorgani N Blokhuis I Hilton HG Béziat V et al
422	17.	Class I HI A haplotypes form two schools that educate NK cells in different ways. Sci
423		Immunol 2016:1(3):eaag1672
474	48	Li H. Ivarsson MA. Walker-Sperling VF. Subleski I. Johnson IK. Wright PW. et al.
425	40.	Identification of an elaborate NK-specific system regulating HI A-C expression PL oS
426		Genet 2018:14(1):e1007163
420	49	Goodson-Gregg FL Krenel SA Anderson SK Tuning of human NK cells by
427	ч у ,	endogenous HI A-C expression Immunogenetics 2020:72(4):205_15
420 //20	50	Moesta AK Norman PI Vawata M Vawata N Gleimer M Parham P Synergistic
420	50.	nolymorphism at two positions distal to the ligand-binding site makes KIR2DL2 a
431		stronger recentor for HI A-C than KIR2DI 3 J Immunol 2008:180(6):3969-79
431 /32	51	IBGE Censo 2010 Atlas censo demografico 2013
л <u>э</u> 2 Лзз	51. 52	Broun Drodo K Vigira Mion AI Forch Pareira N Culni I Detzl Erler MI HI A class
тэ <u>э</u> Лзл	52.	I nolymorphism as characterized by PCP SSOP in a Brazilian exogamic nopulation
тэт //35		Tissue Antigens 2000:56(5):417-27
436	53	Sanger F. Nicklen S. Coulson AR DNA sequencing with chain-terminating inhibitors
437	55.	Proc Natl A cad Sci 1977.74(12):5463_7
438	54	Hilton HG Norman PI Nemat-Gorgani N Govos A Hollenbach IA Henn BM et al
430	л.	Loss and Gain of Natural Killer Cell Recentor Function in an African Hunter Catherer
4 <u>4</u> 0		Population PLoS Genet 2015.11(8).e1005430
<u>4</u> <u>4</u> 1	55	Augusto DG Piovezan BZ Tsuneto I T Callegari-Jacques SM Datzl Erlar MI VID
1-11	55.	ragasto DO, i lovezan DZ, i suncto L1, Canegari-Jacques Bivi, i etzi-Ener WE. KIK

442 Gene Content in Amerindians Indicates Influence of Demographic Factors. PLoS One. 443 2013: 444 56. Kulkarni S, Martin MP, Carrington M. KIR genotyping by multiplex PCR-SSP. 445 Methods Mol Biol. 2010;612:365-375. 446 57. Bunce M, O'Neill CM, Barnardo MCNM, Krausa P, Browning MJ, Morris PJ, et al. 447 Phototyping: comprehensive DNA typing for HLA-A, B, C, DRB1, DRB3, DRB4, 448 DRB5 & DQB1 by PCR with 144 primer mixes utilizing sequence-specific primers 449 (PCR-SSP). Tissue Antigens. 1995;46(5):355-67. 450 Vilches C, Castaño J, Gómez-Lozano N, Estefanía E. Facilitation of KIR genotyping 58. 451 by a PCR-SSP method that amplifies short DNA fragments. Tissue Antigens. 452 2007;70(15):415-22. 453 59. Gross J, Ligges U. nortest: Tests for Normality [Internet]. 2015. Available from: 454 https://cran.r-project.org/package=nortest 455 60. R Core Team. R: A language and environment for statistical computing [Internet]. 456 Vienna, Austria.: R Foundation for Statistical Computing; 2019. Available from: 457 https://www.r-project.org/. 458 61. Dinno A. dunn.test: Dunn's Test of Multiple Comparisons Using Rank Sums [Internet]. 459 2017. Available from: https://cran.r-project.org/package=dunn.test Warnes G, Gorjanc G, Leisch F, Man. M. genetics: Population Genetics [Internet]. 460 62. 461 2019. Available from: https://cran.r-project.org/package=genetics 462 63. Shin J-H, Blay S, Graham J, McNeney B. LDheatmap : An R Function for Graphical 463 Display of Pairwise Linkage Disequilibria Between Single Nucleotide Polymorphisms 464 [Internet]. Vol. 16, Journal of Statistical Software. 2006. Available from: 465 http://stat.sfu.ca/statgen/research/ldheatmap.html 466 Eklund A. Beeswarm: The Bee Swarm Plot, an Alternative to Stripchart. [Internet]. R 64. 467 package version 0.2.0. 2015. Available from: http://cran.r-468 project.org/package=beeswarm 469 Sabeti PC, Reich DE, Higgins JM, Levine HZP, Richter DJ, Schaffner SF, et al. 65. 470 Detecting recent positive selection in the human genome from haplotype structure. 471 Nature. 2002;419(6909):832-7. 472 66. Gautier M, Klassmann A, Vitalis R. rehh 2.0: a reimplementation of the R package 473 rehh to detect positive selection from haplotype structure. Mol Ecol Resour. 474 2017;17(1):78-90. 475 67. Sherry ST, Ward M, Kholodov M, Baker J, Phan L, Smigielski EM, et al. dbSNP : the 476 NCBI database of genetic variation. 2001;29(1):308–11.







493

494 Figure 2. Extended haplotype homozygosity (EHH) in KIR2DL1. The extended 495 homozygosity analysis is based on the premise that advantageous alleles increase in frequency 496 at a higher pace than the local recombination rate breaks down the haplotypes in which these 497 alleles are located. Therefore, alleles marking regions with elevated extended homozygosity 498 are possibly under recent positive selection. Here we identify extended haplotypes 499 surrounding KIR2DL1 variants rs4806553 and rs687000. The possible haplotypes of 500 rs4806553 and rs687000 in relation to rs2304224 are represented at the top of the image. The 501 continuous line represents the most common configuration between two variants, and the 502 dashed line represent less frequent configurations. On the left of each haplotype, arrows 503 indicate higher or lower expression of KIR2DL1 and HLA-C, as associated with rs2304224 504 alleles G or T. A representation of the genomic organization of KIR2DL1 with the indicated location of the three variants is represented above. (A) EHH graph of decay in homozygosity 505 506 (left) and furcation plot (right) for rs4806553 in Euro-Brazilians. The graph shows little to no 507 difference between ancestral rs4806553C (blue) and derived rs4806553G (red) alleles in 508 Euro-Brazilians. (C) EHH graph of decay in homozygosity (left) and furcation plot (right) for 509 rs4806553 in Japanese. In Japanese, elevated homozygosity is associated with derived allele 510 rs4806553G (red). (B) EHH graph of decay in homozygosity (left) and furcation plot (right) 511 for rs687000 in Brazilians with European ancestry and (D) Brazilians with Japanese ancestry. 512 Elevated homozygosity associated with derived allele rs687000G (red) is consistent with the 513 selective sweep model, in which recent positive selection sweeps the diversity on nearby *loci*. 514 Vertical dotted lines indicate the position of the core SNP. The thickness of each branch in the 515 furcation plot is determined by haplotype frequency.

Supplementary Material

1 Supplementary Figures and Tables

1.1 Supplementary Figures



Supplementary Figure 1. Associations of rs2304224 and HLA-C expression levels in 1KGP samples. (A) Non-significant differences between G/G and T/T genotypes of rs2304224 were observed in donors carrying two copies of HLA-C1 (p=0.0740). (B) Allele rs2304224T is associated to HLA-C expression in both heterozygous C1/C2 (p=0.0420) and (C) homozygous C2/C2 (p=0.0059). Each dot in the graphs represents one individual. Median values are shown in horizontal lines and p-values in the top right corner. HLA data were previously published (27) for the samples of the 1000 Genomes Project (1KGP). Median HLA-C expression by allele, as defined by Apps et al. (28), was imputed for each individual genotype.



Supplementary Figure 2. Correlation fit between HLA-C expression in fresh CD3+ cells compared to imputed HLA-C expression levels also from CD3+ cells, as described in Apps et al. (28).



Supplementary Figure 3. The effect of gene copy number on KIR2DL1 expression.

KIR2DL1 copy number was not associated with (A) abundance of KIR2DL1 on the surface of NK cells (p=0.0594) but was associated with (B) percentage of NK cells positive for KIR2DL1(p=0.0001). Gene copy number is indicated in the x axis: 1 copy (hemizygosity) or 2 copies (homozygosity or heterozygosity). Each dot in the graphs represents one individual. Median values are shown in horizontal lines and p-values in the top right corner.



Supplementary Figure 4. Differential expression of common KIR2DL1 allotypes. *KIR2DL1* allelic diversity is associated with (A) abundance of KIR2DL1 on the surface of NK cells ($p < 10^{-5}$) and to (B) percentage of NK cells positive for KIR2DL1 (p = 0.0028). Each dot in the graphs represents one individual, which is plotted twice (once for each allele). Red dots indicate hemizygosity for *KIR2DL1*. Median values are shown in horizontal line. *p*-values in the top right corner of each graph indicate the significance of differences in expression among all allele groups given by the Kruskal-Wallis test. Post-hoc Dunn test shows that there is no significant difference within *Cen A* alleles (*002 and *003) or within *Cen B* alleles (*004 and *006) (p > 0.05), but there is a difference between *Cen A* and *Cen B* alleles in all pairwise comparisons ($p < 10^{-2}$).



Supplementary Figure 5. Linkage disequilibrium (LD) between *KIR2DL1*

polymorphisms. (A) LD spanning 700 bp in the 5' untranslated region (UTR) up to rs2304224, and (B) LD across *KIR2DL1* coding region in Euro-Brazilians (n=109). (C) LD spanning the 5' untranslated region (UTR) and rs2304224, and (D) LD across *KIR2DL1* coding region in Japanese (n=75). Cell color corresponds to the correlation coefficient (r²) between markers, and normalized LD coefficient (D') is given by the number inside each cell. Asterisks and their respective *rs* ID accession numbers highlight all SNPs in significant linkage disequilibrium with rs2304224 (p<0.01).



Supplementary Figure 6. Additive effect of rs2304224T and rs34721508C on KIR2DL1 expression. (A) KIR2DL1 surface expression ($p < 10^{-4}$) and (B) the percentage of NK cells KIR2DL1 positive (p = 0.0043). Each dot in the graphs represents one individual, which is plotted twice (once for each allele). Red dots indicate hemizygosity for *KIR2DL1*. The score of 2 in the *x* axis represents individuals who have both rs2304224T and rs34721508C; while 1 indicates those who have only rs2304224T or rs34721508C; and the score 0 is given when both variants are not present. Median values are shown in horizontal line. *p*-values in the top right corner of each graph indicate the significance of differences in expression among all allele groups given by the Kruskal-Wallis test.



Supplementary Figure 7. Study design. All samples were collected in Curitiba, in the south of Brazil.



Supplementary Figure 8. Gating strategy for flow cytometry analysis. For KIR2DL1 expression, a total of 0.5x10⁶ cells were incubated with the following antibodies: anti-CD56 (REA196 clone PE conjugated, Miltenvi Biotec Cat# 130-100-622, RRID:AB 2658731), anti-CD3 (HIT3a clone PE-Cy5 conjugated, Biolegend Cat# 300328, RRID:AB 1575008) and the specific anti-KIR2DL1 antibody (REA clone FITC conjugated, Miltenyi Biotec Cat# 130-103-967, RRID:AB 2655323). Cells were washed with PBS 1X solution after incubation and then run through BD FacsVerseTM cytometer. The cytometer was calibrated and compensated using CS&T beads (BD Biosciences Cat# 650621). Gating and cell populations were defined using the equipment's software BD FacSuiteTM (BD Biosciences, CA). (A) First, lymphocytes were selected from the PBMC pool based on size (FSC-A) and granularity (SSC-A). Each dot represents one cell, as the density of cells occupying the same position in the graph increases, the color changes from blue to red. (B) Then, we selected lymphocytes which were CD56 positive (PE-A) and CD3 negative (Pe-Cy5-A), corresponding to NK cells. (C) From the NK cell pool, cells positive for KIR2DL1 (FITC-A) were selected. The count of positive and negative cells in this stage determine the percentage of cells expressing KIR2DL1. (D) The expression of KIR2DL1 is given using the median fluorescence intensity (MFI) for each individual. A number of 1,000 target nuclei events were acquired per sample, given the rare occurrence of NK-KIR2DL1+ cells in PBMC (lymphocytes are present in ~80% of PBMC, ~8.7% of those are NK cells, and a median of 15.65% of NK cells are positive for KIR2DL1, which makes a total of 1.09% NK-KIR2DL1⁺ cells in the PBMC pool). A similar approach was used to measure HLA-C expression levels in a subset of 30 individuals using anti-HLA-C antibody (DT9 clone with no fluorescence, BD Biosciences Cat# 566372, RRID: AB 2739715), with the secondary antibody anti-mouse IgG2ab (X-57 clone FITC conjugated, Miltenyi Biotec Cat# 130-098-126, RRID:AB 2661521), and a third antibody anti-CD3 (HIT3a Clone PE/Cy5 conjugated, Biolegend Cat# 300309, RRID:AB 314045). The acquisition was modified to 10,000 target data events acquired per sample, given the high abundancy of cells containing the HLA-C receptor in the CD3+ pool.


Supplementary Figure 9. Strategy used to sequence KIR2DL1 using Sanger. We sequenced exons 1, 4, 5, 7 and 9, which distinguish between the main *KIR2DL1* allele groups. Two long amplification reactions were performed. The first segment comprised exons 1 to 6 (5,564 bp) and was amplified using GoTaq® Long PCR Master Mix (Promega, WI) with 1 cycle at 94°C for 3 minutes, 35 cycles (94°C for 20 seconds, 59°C for 30 seconds, 68°C for 5 minutes) and 1 cycle at 68°C for 10 minutes. The second amplification included exons 7 to 9 (2,218 bp) and it was amplified using Taq Platinum (Invitrogen, CA). The reaction conditions were 1 cycle at 95°C for 5 minutes, 35 cycles (95°C for 20 seconds, 61°C for 30 seconds and 68°C for 5 minutes) and 1 cycle at 68°C for 5 minutes. The amplification products were purified with Exonuclease I (Fermentas, MA) and alkaline phosphatase (Thermo Fisher Scientific, MA) and subsequently used for the five individual sequencing reactions of exons using master mix Big Dye Terminator Cycle Sequencing Standard v3.1 (Life Technologies), for 1 cycle at 95°C for 10 seconds and 35 cycles (50°C for 5 seconds and 60°C for 4 minutes) and 1 cycle at 68°C for 10 minutes. The sequencing product was purified by precipitation and resuspended in 10 µL of Hi-Di Formamide (Life Technologies, CA). Capillary electrophoresis was performed using 3500xl Genetic Analyzer (Life Technologies, CA).

1.2 Supplementary Tables

	rs480	6553	rs2304	224	rs682	7000	rs347.	21508
CDS position	-40)6	13		14	4	796	
Protein codon	non-co	non-coding -		-17 (sig. peptide)		7	245	
Protein substitution	non-co	oding	V>F		synonymous		R>	>C
Allele	C^*	G	G^*	Т	A*	G	C^*	Т
Euro-descendants	0.67	0.33	0.74	0.26	0.57	0.43	0.86	0.14
Japanese	0.44	0.56	0.89	0.11	0.23	0.77	0.99	0.01
	rs480	6553	rs2304	rs2304224		rs687000		21508
High expressing alloty	/pes							
KIR2DL1*002	C		Т		A	1	С	
KIR2DL1*003	G of	r C	G		C	ć	C	
Low expressing alloty	pes							
KIR2DL1*004	G	G			А		Т	
KIR2DL1*006	G	ŕ	G		А		Т	

Supplementary Table 1. *KIR2DL1* variants and frequencies. Allele frequencies of rs2304224 and linked SNPs in Euro-Brazilians (n=109) and Japanese (n=75) populations, as well as their corresponding alleles in common KIR2DL1 allotypes. CDS: coding sequence (DNA). Asterisks indicate the ancestral allele of each SNP. Frequencies corresponding to the alleles which appear to be under positive selection (rs4806553G and rs687000G) are indicated in red.

GGACACTAGGTGTCAAATTCTAGC	
Pre-sequencing KIDODIA ACAAGCAGTGGGTCACTTGAC	
amplification TCACAGGAGGACAGGTGGTT	
^D GCTGTTGTCTCCCTAGAAGACG	
Exon 1 GGACACTAGGTGTCAAATTCTAGC	
Exon 4 GAAGGAGAGAGAGAGACACCAGG	
Sequencing KIR2DL1 Exon 5 GAAGATCCTCCCTGAGGAAAC	
Exon 7 GGGTGCTTGTCCTAAAGAGG (54)	
Exon 9 GGACCCAGAAGTGCCCTC	
CCATCAGTCGCATGACG	
1/B TCACTGGGAGCTGACAC (55)	
Presence/absence TGGACCAAGAGTCTGCAGGA	
typing ¹⁶ TGTTGTCTCCCTAGAAGACG ⁽⁵⁶⁾	
HLA- TGCCAAGTGGAGCACCCAA	
DRB1 GCATCTTGCTCTGTGCAGAT (57)	
CCATCAGTCGCATGACG	
TCACTGGGAGCTGACAC (55)	
KIR2DL1 GTTGGTCAGATGTCATGTTTGAA	
ddPCR CCTGCCAGGTCTTGCG (58)	
AACGACACTTTGCGCCTCATT	
TTCTGCACAGAGAGGGGATCA	
Copy number 1C GAGCCGACAACTCATAGGGTA (56)	
CACTGTGGTGTCTGAAGGAC	
KIR3DL3 ^{2B} GGAGTGTGGGTGTGAACTG ⁽⁵⁵⁾	
CTGCAATGTTGGTCAGATGTCAG	
RT GGGAGCYGACAACTCATAGGGTA Modified from	m (58)
GCTCTGGTTGTAGTAGCCGCGCAG+T	- /

Supplementary Table 2. Primers used in *KIR2DL1* analysis. Sequences in bold were designed in this study.

1 5.2 *KIR* allelic characterization in Southeastern Amerindians and Brazilian

2 **urban populations**

Luciana de Brito Vargas¹, Brenda Ho², Wesley M. Marin², Marcia H. Beltrame¹, Maria Luiza Petzl-Erler¹, Jill A. Hollenbach², Danillo G. Augusto^{1,2*}

- 5 ¹Laboratório de Genética Molecular Humana, Departamento de Genética, Universidade
- 6 Federal do Paraná, Curitiba, Brazil
- ⁷ ²Hollenbach Lab, Weill Institute for Neurosciences, Department of Neurology, University of
- 8 California, San Francisco, CA, United States of America

9 * Correspondence:

- 10 Dr. Danillo Gardenal Augusto
- 11 Danillo.Augusto@ucsf.edu

12 Abstract

- 13 Natural killer (NK) cells trigger cytotoxic and inflammatory responses to viral infections and
- 14 neoplastic cells. NK cells are regulated by killer-cell immunoglobulin-like receptors (KIR),
- 15 encoded by a rapidly evolving gene family on 19q13.4. These receptors present an unusual
- 16 presence and absence polymorphism, with distinctive activating and inhibitory haplotype
- profiles and high allelic diversity. Many studies have associated *KIR* polymorphism with
 medical conditions. However, the interpretation of that relationship is complex and often
- 19 controversial, especially regarding KIR's role in transplantations. Only recently, high-
- resolution allele-level haplotypes have been described in populations, especially those of
- 21 European and African descent. Here, we use an innovative Next Generation sequencing
- 22 method to describe *KIR* diversity in 707 individuals from eight South American populations,
- 23 including six isolated Amerindian populations (Guarani Kaiowá, Guarani Ñandeva, Guarani
- 24 M'bya, Kaingang from Ivaí, Kaingang from Rio das Cobras, and the Aché from Paraguay,
- 25 n=523), as well as two urban populations (Euro-Brazilians, n=109, and Japanese descendants,
- 26 n=75). We report 167 *KIR* alleles in Brazilians of European ancestry and 96 alleles in
- 27 Brazilians of Japanese ancestry, but only an average of 62 KIR alleles were found in
- 28 Amerindians, which reflects the reduced genetic diversity of these populations. Nonetheless,
- 29 19 Amerindian alleles were not found in the urban Brazilian populations. This study expands
- 30 the knowledge of *KIR* genetic diversity in populations and contributes to understanding the
- 31 impact of *KIR* variability in human health.

Keywords: killer-cell immunoglobulin-like receptors, high-resolution, evolution, HLA ligand.

34 **Declarations**

35 Funding

- 36 This work was supported by scholarship and grants from Coordenação de Aperfeiçoamento
- 37 de Pessoal de Nível Superior Brasil (CAPES) Finance Code 001, Conselho Nacional de
- 38 Desenvolvimento Científico e Tecnológico (CNPq), Fundação Araucária, and NIH grant
- 39 U19NS095774.

40 Conflicts of interest/Competing interests (include appropriate disclosures)

- 41 The authors declare that the research was conducted in the absence of any commercial or
- 42 financial relationships that could be construed as a potential conflict of interest.

43 Ethics approval (include appropriate approvals or waivers)

- 44 Ethical approval was given by the Human Research Ethics Committee of the Federal
- 45 University of Paraná and the Brazilian National Human Research Ethics Committee (Conep),
- 46 under the certificate number 02727412.4.0000.0096.

47 Consent to participate (include appropriate statements)

All participants were informed about the research purpose and given written or oral consent toparticipate in the study, in accordance to the Brazilian Federal laws.

50 Authors' contributions

- 51 DGA designed the study; LV, DGA and BH sequenced DNA samples; WM processed KIR
- 52 raw data and DGA processed *HLA* raw data; LV performed statistical analysis of the data;
- 53 DGA and JH contributed with reagents; LV and DGA drafted the manuscript. All authors
- 54 contributed to manuscript revision, read and approved the submitted version.

55 1 Introduction

Natural killer (NK) are cytotoxic lymphocytes that were first discovered due to its ability to 56 57 spontaneously kill tumor cells in vitro without prior sensitization (HERBERMAN; HOLDEN, 58 1978) and later recognized as critical components of the first line of defense against tumor and infected cells (MORVAN; LANIER, 2016; FLÓREZ-ÁLVAREZ et al., 2018). Among 59 60 the receptors that control NK cell cytotoxicity are the killer-cell immunoglobulin-like 61 receptors (KIR), which recognize human leukocyte antigen (HLA) molecules as ligands 62 (LANIER; PHILLIPS, 1995; MORETTA et al., 1996). Nomenclature of KIR is based on the 63 structure of the receptors, which may present two or three extracellular domains (2D or 3D) 64 and short or long cytoplasmatic tails (S or L) (LONG et al., 1996).

65 KIR are encoded by a highly polymorphic gene family located on the chromosome region 66 19q13.4, characterized by a complex structural variation of presence and absence of genes (WENDE et al., 1999; WILSON, M. J. et al., 2000). The homology and high sequence 67

similarity among the 13 KIR loci contribute to the occurrence of non-reciprocal recombination 68

69 (WILSON et al., 2000; MARTIN et al., 2003), which generates duplication and deletion of

70 genes, as well as the formation of hybrid genes and alleles, in a family known by their rapid evolution (KHAKOO et al., 2000; SAMBROOK et al., 2005; GUETHLEIN et al., 2007;

71

72 TRAHERNE et al., 2010).

73 The presence and absence of *KIR* genes generate a wide diversity of gene-content haplotypes

74 which were previously classified in two groups: A and B (UHRBERG et al., 2002).

75 Haplotypes from group B include at least one of the following: KIR2DL2, KIR2DL5A or

76 KIR2DL5B, KIR3DS1, KIR2DS1, KIR2DS2, KIR2DS3 and KIR2DS5. On the other hand, A

77 haplotypes generally exhibit seven fixed genes and are characterized by the lack of all the 78

- genes that characterize group B (GONZALEZ-GALARZA et al., 2020). Despite the large 79
- number of described haplotypes among human populations, the majority of haplotypes found 80 are the result of different combinations of a smaller number of centromeric and telomeric

81 segments (PYO et al., 2010; HOLLENBACH et al., 2012). This feature is most likely driven

82 by a hotspot that facilitates the recombination of segments (TRAHERNE et al., 2010). Both

KIR A and B haplogroups are present in all studied human populations, however, their 83

- 84 frequencies vary significantly. As an example, haplotype A is present in approximately 80%
- 85 of Japanese individuals (YAWATA et al., 2006), but in only 2% of Australian aborigines
- 86 (TONEVA et al., 2001).

87 HLA genes, on the other hand, are located within the major histocompatibility complex (MHC) on chromosome 6 (HORTON et al., 2004) and are considered the most polymorphic 88

89 genes of the human genome (HILL, 1999). The interaction of KIR and HLA molecules is

90 critical for NK cell education during early stages of maturation (KÄRRE et al., 1986; KIM et

91 al., 2005), regulation of NK cell cytotoxicity (SMYTH et al., 2005; LANIER, 2008), and

92 reproduction (HIBY et al., 2004; XIONG et al., 2013; BLOKHUIS et al., 2017). In addition,

93 combinations of KIR-HLA have been associated with numerous diseases (MARTIN et al.,

94 2002; SLIK, VAN DER et al., 2003; HIBY et al., 2004; KHAKOO et al., 2004; NELSON et

- 95 al., 2004; CARRINGTON et al., 2005; AUGUSTO, DANILLO G. et al., 2012; AUGUSTO, 96 2016).
- 97 Not all HLA serve as ligands for KIR, which only recognizes specific motifs in certain HLA

98 allotypes. The HLA-C allotypes are divided in two groups, C1 and C2, which can contain

99 either asparagine or lysine at amino acid position 80 of the HLA-C polypeptide, respectively. 100 HLA-C molecules carrying C1 group serve as ligand for KIR2DL2, KIR2DL3 (HILTON et

- al., 2015), KIR2DS2 (MOESTA et al., 2010) and KIR2DS4 (GRAEF et al., 2009), while C2
- 102 group is recognized by KIR2DL1, KIR2DL2, KIR2DS1 (HILTON et al., 2015), KIR2DS4
- 103 (GRAEF et al., 2009) and KIR2DS5 (BLOKHUIS et al., 2017). HLA-A and HLA-B
- 104 molecules that exhibit the Bw4 epitope are recognized by KIR3DL1 (GUMPERZ et al., 1995; 105 FOLEX et al. 2008) Management of All all the second and a second se
- FOLEY et al., 2008). Moreover, the HLA-A3 and A11 allotypes are recognized by KIR3DL2
 (DÖHRING et al., 1996; HANSASUTA et al., 2004) and KIR2DS4 (GRAEF et al., 2009).
- Because both KIR and HLA gene families are highly polymorphic and segregate
- independently, the number of KIR-HLA interactions as well as the strength and affinity of
- 109 binding and signaling differ among individuals and populations.

110 The complexity of the structural variation combined with sequence similarity among genes 111 impose technical difficulties to the study of KIR. The vast majority of the population genetics 112 studies to date only analyzed KIR at gene-content level and sometimes in combination of a 113 few HLA ligands (GONZÁLEZ-GALARZA et al., 2015). Despite the fact that over a 114 thousand KIR alleles have been deposited in the IPD-KIR database (ROBINSON et al., 2015), 115 the distribution of these variants in global populations is poorly known. The study of allelic 116 diversity at high resolution for all KIR genes is still restricted to a few European (NORMAN 117 et al., 2016; SOLLOCH et al., 2020) and African populations (NEMAT-GORGANI et al., 118 2019). In Amerindians, allelic diversity was described only for the Yucpa from Venezuela, 119 but high resolution was restricted to a few loci (GENDZEKHADZE et al., 2009). In that 120 study, the low diversity in the KIR and HLA system found in the Yucpa allowed the authors to suggest that Yucpa retained the lowest essential diversity for the survival of human 121 122 populations, highlighting the importance of studying isolated Amerindian populations. 123 Evidence shows, however, that eastern Amerindians, such as the ones living in the Brazilian 124 region, bear lower genetic diversity than western Amerindians (WANG et al., 2007). Therefore, we hypothesize that southeastern Amerindians may have an even lower and 125 126 differentiated KIR diversity to the western Amerindians previously described. Moreover, we

- aim to compare these Amerindian populations to Brazilians with European and Japanese
- ancestries, which may provide insights into the extent of gene flow between Amerindians and

non-Amerindians, and to the genetic proximity of the *KIR* and *HLA* system in Amerindian

130 and Asian populations.

131 Here, we deliver the first high-resolution characterization of *KIR* allelic variation in eight

- 132 South American populations, seven from Brazil and one from Paraguay. We analyzed six
- 133 isolated Amerindian populations from Guarani, Kaingang and Aché groups, as well as
- 134 descendants of European and Japanese living in Southern Brazil. We identified the most
- 135 frequent allele-level haplotypes in each population and analyzed the *KIR* variation in the
- 136 context of their HLA ligands.
- 137 **2 Results**

138 2.1 Amerindians exhibit reduced diversity of *KIR* alleles

A total of 209 *KIR* variants at allelic level were observed across the study populations. In

Amerindians, we observed 122 *KIR* variants, averaging 62 per population. Nineteen of the

141 variants observed in Amerindians were not observed in the urban populations (Figure 1). Six

142 of these Amerindian variants have been only found in the Kaingang from Ivaí, and five were

- exclusive to Guarani Kaiowá. In sharp contrast, we found 167 *KIR* variants solely in the urban
- 144 population of predominantly European ancestry and 96 in the Japanese descendants from

145 Curitiba. *KIR3DL3* and *KIR3DL2* were the two genes with the highest number of alleles 146 (Figure 2). Detailed allelic frequencies for each population are described in Figures 3 and 4

Overall, Amerindians share an average of 64.2% of *KIR* alleles with Euro-Brazilians, ranging
from 59.2 to 69.9%. A higher proportion of alleles, 71.2%, is shared between Amerindians
and Japanese descendants, ranging from 65.6% to 75.7% (Table 1). Considering only the

150 Amerindian populations, 82.4% of *KIR* variants are shared among them. However, the two

151 Kaingang populations share a higher proportion of variants than the three Guarani populations

152 (90.6% and 82.8%, respectively). Interestingly, the Aché has the lowest proportion of shared

alleles with Euro-Brazilians, only 59.2%, and *KIR* allelic diversity in this population is more similar to the Guaranis (80.1%) than to the Kaingang populations (76.3%).

- 155 We estimated genetic differentiation in *KIR* genes between populations using F_{ST} analysis
- 156 (Fig. 5). The higher overall differentiation values were observed for KIR2DL1 (median $F_{ST} =$
- 157 0.10), followed by *KIR2DL23* (median $F_{ST} = 0.09$) (Figure 5C). In fact, we observed striking

differences in the presence and absence polymorphism of *KIR2DL1*, which is present in

159 60.1% of haplotypes in the of Guarani M'bya to 99.99% in the Aché (Figure 3).

160 The most divergent Amerindian population was the Aché (median $F_{ST} = 0.05$, Figure 5B).

161 This observation is supported in the principal component analysis (PCA), which distributes

162 populations according to the overall heterogeneity across their allele frequencies. The first

163 two principal components (PC1 and PC2) explain 42.6% of the variation in *KIR* and show that

164 *KIR* diversity in the Aché is indeed distinct from the other Amerindian groups (Figure 6).

165 Although Aché differ significantly from all other Amerindians, their *KIR* diversity is closer to 166 that found in Guarani Kaiowá ($F_{ST} = 0.01$) in comparison to other Guarani ($F_{ST} = 0.02$ and

that found in Guarani Kaiowá ($F_{ST} = 0.01$) in comparison to other Guarani ($F_{ST} = 0.02$ and 0.07) and also Kaingang ($F_{ST} = 0.05$ and 0.06). Interestingly, *KIR* diversity in Guarani M'Byá

168 is more closely related to the Kaingang than to the other Guarani populations (Figure 5D and

169 Table 1).

170 The first two components of PCA analysis distinguish the groups of Amerindians, Japanese

and Euro-Brazilians. The third component, which explains 7.7% of KIR diversity, groups the

172 Japanese with Amerindians. The fourth component (3.7%) separated the Amerindians in two

groups: one containing the Aché, Guarani Kaiowá and Guarani Ñandeva; and a second group
containing Guarani M'bya, Kaingang from Ivaí and Kaingang from Rio das Cobras (Suppl.

175 Fig. 1).

176 We also tested the neutrality of *KIR* genes in the eight South American populations. The rates

177 of non-synonymous to synonymous substitutions (dN/dS) in KIR genes were according to the

178 expected under a neutral model for most locus. Only the KIR3DL3 locus in the Aché

population was significantly departing from neutrality (p = 0.049), with a negative overall

180 dN/dS value of -1.99, which may indicate purifying selection (Suppl. Table 1). Notably, the

181 D1 domain and the portion containing the stem, transmembrane and cytoplasmatic domains

182 presented negative dN/dS values (-0.73 and -0.12, respectively), while domains D0 and D2

183 had positive dN/dS values (1.48 and 1.71, respectively). When analyzed separately, all

184 KIR3DL3 domains were neutral (p > 0.05).

185 2.3 Amerindians maintain limited diversity of *KIR* haplotypes

186 Considering only the *KIR* haplotypes observed in at least three individuals, we observed 44

telomeric and 37 centromeric haplotypes in the study populations. As expected, increased

188 diversity of haplotypes was observed in populations of European and Japanese descent, with 14

and 7 centromeric haplotypes, and 15 and 6 telomeric haplotypes, respectively. Among
 Amerindians, the highest number of haplotypes was observed among the three Guarani
 populations, with up to 12 telomeric and 9 centromeric haplotypes. The two Kaingang and the
 Aché exhibited five or less centromeric and six or less telomeric haplotypes. The most common
 haplotypes in all populations are shown in Figure 7.

In regards to the low frequency haplotypes, we observed 121 centromeric 127 telomeric haplotypes that were found in less than three individuals (Supplemental Table 2 and 3). These haplotypes are usually variations of the most common ones and differ by the allelic polymorphism of a few *KIR* genes. Allelic variation of *KIR3DL3* was responsible for around 30% of the less common centromeric haplotypes. On the telomeric region, the allelic variation on *KIR3DL1* differentiated most of the less frequent haplotypes.

- 200 Remarkably, we observed that the Aché have an extremely high frequency of centromeric A
- haplotypes (*CenA*, f=0.99) and lower frequency of telomeric *A* haplotypes (*TelA*, f=0.62),
- which result in 58.0% frequency of haplotype A (CenA + TelA) (Figure 7C-E). The Guarani
- 203 Kaiowá and Guarani Ñandeva also exhibit high frequencies of Cen A (f=0.89 and 0.84,
- 204 respectively) and lower frequencies of *Tel A* (f=0.50 and 0.60, respectively). High frequency
- of *Cen A* was also observed in the Japanese (f=0.88), however, they exhibited *Tel A* (f=0.79)
- 206 more frequently than Amerindians, and, therefore, the highest frequency of the *KIR A* 207 hard true in this starts (6, 0.75)
- 207 haplotype in this study (f=0.75).

208 2.2 2 Fewer functional KIR-HLA interactions were observed in the isolated populations

209 We analyzed the presence of KIR ligands to provide a detailed analysis of KIR-HLA 210 interactions in the study populations. This is also the first report of the frequencies of HLA ligands in Aché, in which we show the lack of HLA-A3 and A11 ligands, similarly to what 211 212 has been observed in Guarani Kaiowá, Guarani M'bya and Kaingang from Ivaí (Table 2). 213 Presence of these ligands at low frequencies was observed in Guarani Ñandeva (1.0%) and 214 Kaingang from Rio das Cobras (4.0%). Interestingly, Aché also exhibit low frequency of Bw4 215 (f=0.05), which results in most of their KIR-HLA interactions being dependent on HLA-C1 216 and HLA-C2. In average, Amerindians exhibited only ~3 to 5 KIR-HLA interactions (Table 217 3), and 95% of individuals presented between 1 and 7 interactions (Suppl. Fig. 2). In 218 Japanese, we found and average of 6 KIR-HLA interactions with 95% of individuals 219 presenting between 2 and 11 interactions, and in the Euro-Brazilians we found an average of 220 7.65, with 95% of individuals having from 4 to 14 interactions. HLA-C was responsible for 221 53.9% of the observed KIR-HLA interactions in Japanese, 58.3% in Euro-Brazilians, and 222 83.0% to 97.2% in Amerindians. The HLA-C allotype frequencies, however, differed 223 significantly among Amerindians (Figure 8). Most of the KIR-HLA interactions observed in

all study populations (>60%) are inhibitory.

225 **3** Discussion

We observed lower *KIR* allelic diversity in Amerindians, an average of only 62 variants per population, in comparison to Japanese (96) and European (167). The highest *KIR* allelic

diversity to a single population sample has been reported in Sub-Saharan Africans from

229 Ghana, with 175 alleles described (NORMAN et al., 2013a). Amerindian populations are

- 230 unique due to the demographic history of migrations of their Asian ancestors via the Behring
- 231 Strait, their complex dispersal along the American continent, as well as the persisting genetic
- isolation during the last five centuries (PETZL-ERLER et al., 1993; TSUNETO et al., 2003).
- 233 Amerindian demographic history is especially interesting because they underwent strong

- bottleneck and founder effects (AMOS; HOFFMAN, 2010; O'FALLON; FEHREN-
- 235 SCHMITZ, 2011), and they also have been subject to greater genetic drift effects due to their
- 236 mostly small population sizes (CONROY et al., 2000; TARAZONA-SANTOS et al., 2001),
- all of which contribute to their low genomic diversity (DEGIORGIO et al., 2009). Reduced
- 238 KIR gene-content diversity was previously observed in Amerindians (AUGUSTO et al.,
- 239 2013a). The reduced diversity of the KIR and HLA interactions in Amerindians may provide
- 240 insights in the essentiality of *KIR* genes and ligands for the survival of populations, as has
- been suggested for the Yucpa indigenous from Venezuela (GENDZEKHADZE et al., 2009).
- 242 The Kaingang from Ivaí exhibited six variants that were not found in the other study
- 243 populations (Figure 1). Interestingly, four of them form the haplotype KIR3DL3*01406~
- 244 KIR2DL2*00602~ KIR2DL5B*00601~KIR2DS5*004. This haplotype has been described in
- the Ga-Adangbe from Ghana (NORMAN et al., 2013b) and in the KhoeSan from Southern
- Africa (NEMAT-GORGANI et al., 2018). Similarly, the Guarani Kaiowá present five variants observed only in this population. Three of those define the haplotype
- 248 KIR2DL4*022~KIR3DL1*024N~KIR2DS4*00104. Of these, KIR3DL1*024N encodes a
- truncated KIR3DL1 protein and was reported in the Ga-Adangbe from Ghana (NORMAN et
- al., 2007). In addition, KIR2DS4*00104 was also observed in African Americans (HOU et al.,
- 251 2009). Therefore, we suggest that some of the 19 variants that were found in the Amerindians
- but not observed in two urban populations of European and Japanese descent may be result of
- 253 gene flow with African descendants that migrated to Brazil centuries ago.
- 254 *KIR* frequencies were similar among Amerindians, with the exception of the Aché. The origin
- 255 of the Aché groups currently living in Paraguay is still not completely clear, but genetic
- evidence indicates that they are closer to Guarani than to other Amerindian populations
- 257 (BATTILANA et al., 2002; GASPAR et al., 2002; TSUNETO et al., 2003; SCHMITT et al.,
- 258 2004; CALLEGARI-JACQUES et al., 2007). We observed that *KIR* diversity in Aché
- corroborates their closer relationship with Guarani. However, Aché exhibited the greatest *KIR*
- differentiation among Amerindians, especially in comparison with the Kaingang from Ivaí
- and Rio das Cobras and Guarani M'bya ($F_{ST} > 0.04$). The lowest proportion of shared alleles were also observed between the Aché and these populations (72.7 to 78.0%).
- 263 According to principal component analysis and differentiation statistics, *KIR* diversity in the
- Aché, Guarani Kaiowá and Guarani M'bya is differentiated. Despite they share low
- 265 frequencies of A3, A11 and Bw4, the epitope C1 occurs at lower frequency in the Aché
- 266 (f=0.55) than in both the Guarani Kaiowá and Guarani M'bya (f=0.64 and 0.85),
- 267 respectively). Moreover, the Aché present a remarkably high frequency of Cen A haplotypes
- 268 (f=0.99) which contains *KIR2DL3*, a weaker C1 receptor (ANFOSSI et al., 2006), and
- usually also KIR2DL1*00302 (f=0.92), a strong C2 receptor (HILTON et al., 2015), which
- 270 seem to compensate the low frequency of C1 in this population. In the Aché, the rate of
- 271 synonymous substitutions was significantly higher than the rate of non-synonymous
- substitutions for *KIR3DL3* (p = 0.049), which is suggestive of purifying selection. Even
- though the ligand for KIR3DL3 is not yet known, the high degree of polymorphism along
- with the conservation of certain portions of its sequence suggest that this receptor plays an important role in KIR function, especially in reproduction (LEATON et al., 2019). The
- domain most likely to contribute to this negative value is the D1 domain, which is predicted
- to have important functions in the receptor clustering to form the immunological synapsis of
- 278 NK cells (BOYINGTON et al., 2001).
- 279 Due to the common origin of the Guarani populations, we would expect *KIR* diversity to be 280 more similar within Guarani populations and different from Kaingang. However, Guarani

M'bya are more similar to Kaingang than to other Guarani, an observation corroborated by
the proportion of shared alleles, differentiation statistics and principal component analysis.
Both Guarani M'bya and Kaingang live in the Rio das Cobras indigenous land, which spans
19.000 hectares in Paraná (INSTITUTO SOCIOAMBIENTAL, 2020). Although they occupy
the same area, these populations live in different villages and remained genetically isolated
due to strong cultural and linguistic barriers. The estimated gene flow between Guarani

- 287 M'bya and Kaingang, based on HLA-DRB1 frequencies is of only 5% (TSUNETO et al.,
- 2003). Therefore, the similarity in *KIR* could be stochastic, driven by genetic drift, as it has
- been shown that demographic factors apparently impact *KIR* diversity in Amerindians to a
- 290 greater extent than natural selection (AUGUSTO et al., 2013b). Alternatively, we speculate
- that the similarity between these two populations could be explained by local selective
- pressures, rather than identity by descent. Further analysis comparing the diversity in *KIR*
- 293 with whole-genome diversity would be informative to understand the nature of this similarity.

294 Frequencies of KIR ligands have already been reported in Brazilians of European ancestry 295 (AUGUSTO et al., 2012), Japanese ancestry (AUGUSTO et al., 2016) and Brazilian 296 Amerindians (PARHAM et al., 1997; AUGUSTO et al., 2013a), but this is study is the first to 297 describe their frequencies for Aché. The absence or low frequencies HLA-A and HLA-B 298 ligands (absence of HLA-A3, HLA-A11, and low frequency of HLA-Bw4) occur in three 299 Amerindian populations: Aché, Guarani Kaiowá and Guarani M'bya. Absence or low 300 frequencies of HLA-A3, HLA-A11 and HLA-Bw4 means that most of the KIR-HLA 301 interactions in these populations rely on the recognition of HLA-C1 and HLA-C2 by KIR2D. 302 In addition, the Amerindian population Yucpa exhibited 2.9 KIR-HLA interactions, which led 303 the authors to suggest that this number may define the minimum diversity necessary to 304 maintain NK function (GENDZEKHADZE et al., 2009). Here, using the same approach, we 305 observed even lower interactions for the Kaingang from Rio das Cobras and the Guarani 306 Kaiowá, 2.76 and 2.75 respectively. It has been suggested that western Amerindians, such as 307 the Yucpa, have higher genetic diversity and a lower level of population structure than eastern 308 Amerindians, such as the southern Brazilian Amerindian populations (WANG et al., 2007). 309 Up to 97% of the KIR-HLA interactions found in these populations rely on the recognition of 310 HLA-C by KIR2DL1, KIR2DL2 and KIR2DL3. These interactions are known to be the 311 strongest HLA-mediated educators of NK cells (MOESTA et al., 2008; HILTON et al., 2015). 312 Altogether, our results support the suggestion that HLA-C specialized and evolved to become 313 primarily KIR ligands while HLA-A and HLA-B kept their primary function as T cell 314 receptor ligands (OLDER AGUILAR et al., 2010).

315 Interestingly, we report that Brazilians of European ancestry have 7.65 interactions per individual, which is higher than the average of 6.5 interactions found in European populations 316 (NEMAT-GORGANI et al., 2019). Euro-Brazilians from Curitiba are an admixed population 317 318 with a high proportion of European ancestry, close to 80% (IBGE, 2013). However, even 319 individuals who are phenotypically white have approximately 6% African and/or Amerindian 320 genetic background (BRAUN-PRADO et al., 2000). This admixture may explain why they 321 have higher genetic diversity and more interactions than the 6.5 KIR-HLA interactions per 322 individual reported in Europeans (NEMAT-GORGANI et al., 2019).

- 323 In this study, we make an outstanding contribution by reporting more than 150 centromeric
- and 170 telomeric high-resolution allele-level haplotypes in South American populations. In
- 325 most cases, these haplotypes are variations of a smaller number of conserved, more frequent,
- haplotypes. Because our method allows for 5-digit characterization of alleles, single
- 327 nucleotide polymorphisms further diversify haplotypes in the populations. This

microvariation provides insights into the linkage disequilibrium pattern between the *KIR* genes and enhance the ability to predict *KIR* haplotypes from different types of data.

330 4 Material and methods

331 4.1 Samples

332 We analyzed 707 samples from eight different populations (Table 4). They include: Euro-333 descendants from Curitiba (CTBA, n=109), Japanese descendants from Curitiba (JAP, n=75), 334 Guarani Kaiowá (GKW, n=150), Guarani Ñandeva (GND, n=81), Guarani M'bya (GRC, n=84), Kaingang from Ivaí (KIV, n=93), Kaingang from Rio das Cobras (KRC, n=64), and 335 Aché (ACHE, n=51). Sampling of these populations is detailed in (TSUNETO et al., 2003) 336 337 and Table 4, and sampling locations are indicated in Figure 9. DNA samples were extracted 338 throughout the years using the Phenol-Chloroform-Isoamyl Alcohol method (SAMBROOK et 339 al., 1989) or salting-out method (LAHIRI; NURNBERGER, 1991), and stored at -80°C. All 340 participants were informed about the research purpose and given written or oral consent to 341 participate in the study. In accordance to the Brazilian Federal laws, ethical approval was 342 given by the Brazilian National Human Research Ethics Committee (CONEP) (CAAE 343 02727412.4.0000.0096).

344 The Amerindian population samples were collected between 1988 and 2000, as previously 345 described in TSUNETO et al. (2003). The individuals were part of genetically isolated groups 346 living in indigenous lands in the Brazilians states of Paraná and Mato Grosso do Sul, and in 347 the bordering country Paraguay. The Guarani are closely related groups, and it is suggested they subdivided in the three populations here studied (Guarani Ñandeva, Guarani Kaiowá and 348 349 Guarani M'bya) around 1,800 years ago, while the Kaingang, which speak a Jê language are 350 suggested to be very closely related groups which split approximately 200 years ago 351 (MARRERO et al., 2007). The Aché groups currently living in Paraguay present cultural and 352 genetic similarity to both Guarani and Kaingang groups. However, evidence based on 353 autosomal and sexual genetic markers suggest that they are closer to Guarani groups 354 (BATTILANA et al., 2002; GASPAR et al., 2002; TSUNETO et al., 2003; SCHMITT et al., 355 2004; CALLEGARI-JACQUES et al., 2007).

356 The urban populations were collected in the city of Curitiba, capital of Paraná State and one 357 of the largest cities in Brazil, with over 2 million inhabitants (IBGE, 2013). The south of 358 Brazil, which was inhabited mainly by descendants of Amerindians and Africans during the 359 colonization period, received large influx of European immigrants during the XIX and XX 360 centuries (KEHDY et al., 2015). In the city of Curitiba, most people have Portuguese, 361 German, Italian, Polish, and Ukrainian background, among others. According to the last 362 Brazilian population census in 2010, Curitiba's population self-reported its ethnicity as 78.8% 363 white, 16.7% admixed, 3% black, 1.4% Asian and 0.2% of Native Americans (IBGE, 2013). 364 Contributing to the Curitiba's genetic diversity are also Japanese communities. During 365 Japan's crisis after the Meiji restauration in the late XIX century, Japanese people immigrated 366 heavily to the United States, Peru, Mexico and Brazil. Nowadays, Brazil has the largest 367 population of Japanese outside of Japan, with over 1.5 million people living mainly in the states of São Paulo and Paraná (SAKURAI et al., 2010; IBGE, 2013). Individuals from 368 369 Japanese ancestry in this study were born in Curitiba, Brazil and reported no admixture with 370 non-Japanese. Moreover, these individuals reported that both parents and/or four grandparents 371 were born in Japan.

77

372 4.2 KIR and HLA sequencing

373 Next Generation Sequencing targeted for KIR and HLA class I genes was performed 374 according to our custom method (NORMAN et al., 2016) in the 707 individuals described 375 earlier. DNA was enzymatically fragmentated using KAPA HyperPlus kit (Roche, USA) and 376 barcoded with TruSight kit (Illumina, San Diego, CA). Dual size selection was performed 377 with AMPure XP Beads (Beckman Coulter, USA) to obtain fragments with an average size of 378 780 bp. Quality control was done using PicoGreen® kit (Thermo Fisher Scientific, Waltham, 379 MA) for product quantification and Bioanalyzer (Agilent, Santa Clara, CA) to determine 380 quantity and size of fragments. Pooling was done in the automated liquid handling Echo ® 381 525(Labcyte, San Jose, CA). Capture was performed with 10,456 biotinylated probes 382 designed by NORMAN et al. (2016) to capture fragments corresponding to all KIR and HLA 383 class I loci. These were purified using SV-Magnetic Beads (Illumina, San Diego, CA). 384 Enrichment was performed with TruSight kit (Illumina, San Diego, CA) and captured with 385 SPRI Nextera magnetic beads (Illumina, San Diego, CA). Sequencing was performed in the 386 platform Illumina HiSeq 4000 (Illumina, San Diego, CA).

387 4.3 Data analysis

Sequence filtering, alignment and genotype calling of *KIR* genes were made using PING
pipeline (NORMAN et al., 2016; MARIN et al., in preparation). Genotype and copy number
were obtained for all thirteen *KIR* genes. *KIR* data was manually curated for copy number

determination and resolution of ambiguities. *KIR2DL5* alleles were manually classified in

392 *KIR2DL5A* and *KIR2DL5B*, which correspond to different *loci* in the *KIR* region. Similarly,

393 KIR2DS3 and KIR2DS5 were classified in the loci containing KIR2DS3 and KIR2DS5 alleles

in the centromeric region, and the *loci* that contains only *KIR2DS5* alleles and is mapped to

the telomeric region. Observation of linkage disequilibrium between neighboring loci was

396 used to resolve ambiguities, for allelic classification of the above-mentioned loci, and for the 397 resolution of *KIR* haplotypes. Intronic variants were also used for resolution of ambiguities.

- resolution of *KIR* haplotypes. Intronic variants were also used for resolution of ambiguities.
- 398 The HLA class I genes were sequenced simultaneously with KIR. The corresponding sequence
- 399 filtering, alignment and genotype calling of HLA genes were done using HLA ExploreTM
- 400 (Omixon, Hungary), which determines 6-digit genotypes unambiguously.

401 Allele frequencies, as well as the proportion of shared KIR alleles among populations, were 402 calculated and plot using a customized version of *PopGenReport* R package (ADAMACK; 403 GRUBER, 2014). Intersecting sets of alleles were identified using intersect from base R and 404 plotted using upset from Package UpSetR (CONWAY et al., 2017). FST was calculated using 405 pegas R package (WEIR; COCKERHAM, 1984; PARADIS, 2010), and the neighbor-joining 406 (NJ) tree was estimated in R package ape (PARADIS; SCHLIEP, 2019). Neutrality tests of 407 non-synonymous (dN) to synonymous (dS) ratio (NEI; GOJOBORI, 1986) were performed in 408 MEGA X software (KUMAR et al., 2018). Principal components were calculated and plotted 409 using the ade4 R package (DRAY; DUFOUR, 2007).

A custom R script was developed to classify *HLA-A*, *HLA-B* and *HLA-C* according to the
relevant epitopes for KIR interaction (SIDNEY et al., 2008). *HLA-A* was classified in A3 and
A11 epitopes, which include alleles of the *A*003* and *A*011* lineages, and the epitope Bw4,
which included allele lineages *A*023*, *A*024* and *A*032*. For *HLA-B*, we recovered protein

- 414 sequences of each allele in the IPD-IMGT/HLA database (ROBINSON et al., 2020), and the
- 415 sequences were classified in Bw4 (presence of the SLRNLRG epitope in positions 77-83) and

416 Bw6 (presence of either one of NLRIALR, DLRTLLR, SLRTLLR, NLRTALR, SLRTALR

- 417 epitopes in positions 77-83). Bw4 was subclassified in epitopes containing an isoleucine
- 418 residue at position 80 (Bw4I) which confers greater inhibition upon interaction with
- 419 KIR3DL1 than threonine at position 80 (Bw4T) (CARR et al., 2005). Finally, *HLA-C* was
- 420 classified in a similar manner considering the asparagine (C1 epitope) or lysine (C2 epitope)421 in position 80.

422 HLA and KIR data were integrated to generate the medium score of KIR-HLA interactions as 423 described (NEMAT-GORGANI et al., 2018). Considered in this score are the interactions 424 confirmed in functional assays. They include the following pairs: Bw4 (HLA-A) and 425 KIR3DL1 (FOLEY et al., 2008); HLA-Bw4 and KIR3DL1 (GUMPERZ et al., 1995; FOLEY 426 et al., 2008); HLA-A*03 and KIR3DL2 (DÖHRING et al., 1996); HLA-A*11 and KIR3DL2 427 (HANSASUTA et al., 2004); HLA-C2 e KIR2DL1 (HILTON et al., 2015); HLA-C and 428 KIR2DL2 (HILTON et al., 2015); HLA-C1 and KIR2DL3 (HILTON et al., 2015); HLA-C2 and KIR2DS1 (HILTON et al., 2015); HLA-C*16 and KIR2DS2 (MOESTA et al., 2010); 429 430 HLA-A*11 and KIR2DS4 (GRAEF et al., 2009); a subset of HLA-C which includes C2 431 (alleles C*05:01, C*02:02 and C*04:01) and C1 (alleles C*16:01, C*01:02 and C*14:02), 432 recognized by KIR2DS4 (GRAEF et al., 2009); and, finally, HLA-C2 and a subset of 433 KIR2DS5 receptors (encoded by alleles 2DS5*003, *004, *005, *006, *007 and *008) 434 (BLOKHUIS et al., 2017). For each individual, each interaction is scored in the following 435 manner: the number of HLA alleles involved in the interaction is multiplied by the number of 436 KIR alleles involved in the interaction; the scores can be 0, 1, 2 and 4. It is worth noticing that 437 if there are two alleles which can participate in the interaction and they are in homozygosity, 438 they are given the score of 1. Alternatively, if those two alleles are in heterozygosity, they are 439 given the score of 2. For each individual, the score of every interaction is summed. Then, this 440 information is used to calculate the average number of KIR-HLA interactions in the 441 population.

442 **5** Acknowledgments

We would like to thank all the participants in this study and all colleagues from the HumanMolecular Laboratory of the Federal University of Paraná, Brazil.

445 6 References

- 446 ADAMACK, A. T.; GRUBER, B. PopGenReport: Simplifying basic population genetic
- 447 analyses in R. **Methods in Ecology and Evolution**, 2014.
- 448 AMOS, W.; HOFFMAN, J. I. Evidence that two main bottleneck events shaped modern
- 449 human genetic diversity. **Proceedings of the Royal Society B: Biological Sciences**, 2010.
- 450 ANFOSSI, N.; ANDRÉ, P.; GUIA, S.; et al. Human NK Cell Education by Inhibitory
- 451 Receptors for MHC Class I. Immunity, 2006.
- 452 AUGUSTO, D. G. The impact of KIR polymorphism on the risk of developing cancer: Not as453 strong as imagined? Frontiers in Genetics, 2016.
- 454 AUGUSTO, D. G.; AMORIM, L. M.; FARIAS, T. D. J.; PETZL-ERLER, M. L. KIR and
- 455 HLA genotyping of Japanese descendants from Curitiba, a city of predominantly European
- 456 ancestry from Southern Brazil. Human Immunology, 2016.
- 457 AUGUSTO, DANILLO G.; LOBO-ALVES, S. C.; MELO, M. F.; PEREIRA, N. F.; PETZL-
- 458 ERLER, M. L. Activating KIR and HLA Bw4 ligands are associated to decreased

- 459 susceptibility to pemphigus foliaceus, an autoimmune blistering skin disease. PLoS ONE, v.
 460 7. n. 7, 2012.
- 461 AUGUSTO, D. G.; PIOVEZAN, B. Z.; TSUNETO, L. T.; CALLEGARI-JACQUES, S. M.;
- 462 PETZL-ERLER, M. L. KIR Gene Content in Amerindians Indicates Influence of
 463 Demographic Factors. PLoS ONE, 2013a.
- 464 AUGUSTO, D. G.; PIOVEZAN, B. Z.; TSUNETO, L. T.; CALLEGARI-JACQUES, S. M.;
- 465 PETZL-ERLER, M. L. KIR Gene Content in Amerindians Indicates Influence of
- 466 Demographic Factors. **PLoS ONE**, v. 8, n. 2, p. e56755, 2013b.
- 467 AUGUSTO, D. G.; ZEHNDER-ALVES, L.; PINCERATI, M. R.; et al. Diversity of the KIR
 468 gene cluster in an urban Brazilian population. Immunogenetics, v. 64, n. 2, p. 143–152,
 469 2012.
- 470 BATTILANA, J.; BONATTO, S. L.; FREITAS, L. B.; et al. Alu insertions versus blood
- group plus protein genetic variability in four Amerindian populations. Annals of Human
 Biology, 2002.
- 473 BLOKHUIS, J. H.; HILTON, H. G.; GUETHLEIN, L. A.; et al. KIR2DS5 allotypes that
- 474 recognize the C2 epitope of HLA-C are common among Africans and absent from Europeans.
 475 Immunity Inflammation and Disease, 2017.
- 476 BOYINGTON, J. C.; BROOKS, A. G.; SUN, P. D. Structure of killer cell immunoglobulin-
- 476 BOTHNOTON, J. C., BROOKS, A. G., SON, T. D. Structure of Kinef cent initialogiobulin477 like receptors and their recognition of the class I MHC molecules. Immunological Reviews,
 478 v. 181, n. 1, p. 66–78, 2001.
- 479 BRAUN-PRADO, K.; VIEIRA MION, A. L.; FARAH PEREIRA, N.; CULPI, L.; PETZL-
- 480 ERLER, M. L. HLA class I polymorphism, as characterised by PCR-SSOP, in a Brazilian 481 exogamic population. **Tissue Antigens**, v. 56, n. 5, p. 417–427, 2000.
- 482 CALLEGARI-JACQUES, S. M.; CROSSETTI, S. G.; KOHLRAUSCH, F. B.; et al. The β-
- 483 globin gene cluster distribution revisited Patterns in native American populations. American
 484 Journal of Physical Anthropology, 2007.
- 485 CARR, W. H.; PANDO, M. J.; PARHAM, P. KIR3DL1 Polymorphisms That Affect NK Cell
 486 Inhibition by HLA-Bw4 Ligand. The Journal of Immunology, 2005.
- 487 CARRINGTON, M.; WANG, S.; MARTIN, M. P.; et al. Hierarchy of resistance to cervical
- 488 neoplasia mediated by combinations of killer immunoglobulin-like receptor and human
 489 leukocyte antigen loci. The Journal of Experimental Medicine, 2005.
- 490 CONROY, G. C.; WEBER, G. W.; SEIDLER, H.; et al. Genetic structure of Quechua-
- 491 speakers of the Central Andes and geographic patterns of gene frequencies in South
- 492 Amerindian populations. American Journal of Physical Anthropology, 2000.
- 493 CONWAY, J. R.; LEX, A.; GEHLENBORG, N. UpSetR: An R package for the visualization 494 of intersecting sets and their properties. **Bioinformatics**, 2017.
- 495 DEGIORGIO, M.; JAKOBSSON, M.; ROSENBERG, N. A. Explaining worldwide patterns
- 496 of human genetic variation using a coalescent-based serial founder model of migration
- 497 outward from Africa. Proceedings of the National Academy of Sciences of the United
 498 States of America. 2009.
- 498 States of America, 2009.
- 499 DÖHRING, C.; SCHEIDEGGER, D.; SAMARIDIS, J.; CELLA, M.; COLONNA, M. A
- human killer inhibitory receptor specific for HLA-A1,2. Journal of immunology
 (Baltimore, Md.: 1950), 1996.
- 502 DRAY, S.; DUFOUR, A. B. The ade4 package: Implementing the duality diagram for
- 503 ecologists. Journal of Statistical Software, 2007.

- 504 FLÓREZ-ÁLVAREZ, L.; HERNANDEZ, J. C.; ZAPATA, W. NK cells in HIV-1 infection:
- 505 From basic science to vaccine strategies. **Frontiers in Immunology**, 2018.
- 506 FOLEY, B. A.; SANTIS, D. DE; BEELEN, E. VAN; et al. The reactivity of Bw4+ HLA-B 507 and HLA-A alleles with kir3dll: Implications for patient and donor suitability for
- 508 haploidentical stem cell transplantations. **Blood**, 2008.
- 509 GASPAR, P. A.; HUTZ, M. H.; SALZANO, F. M.; et al. Polymorphisms of CYP1A1,
- 510 CYP2E1, GSTM1, GSTT1, and TP53 genes in Amerindians. American Journal of Physical
 511 Anthropology, 2002.
- 512 GENDZEKHADZE, K.; NORMAN, P. J.; ABI-RACHED, L.; et al. Co-evolution of
- 513 KIR2DL3 with HLA-C in a human population retaining minimal essential diversity of KIR
- and HLA class I ligands. Proceedings of the National Academy of Sciences of the United
 States of America, 2009.
- 516 GONZALEZ-GALARZA, F. F.; MCCABE, A.; SANTOS, E. J. M. DOS; et al. Allele
- 517 frequency net database (AFND) 2020 update: gold-standard data classification, open access
- 518 genotype data and new query tools. Nucleic acids research, v. 48, n. D1, p. D783–D788,
- 519 2020.
- 520 GONZÁLEZ-GALARZA, F. F.; TAKESHITA, L. Y. C.; SANTOS, E. J. M.; et al. Allele
- 521 frequency net 2015 update: New features for HLA epitopes, KIR and disease and HLA
- adverse drug reaction associations. Nucleic Acids Research, v. 43, n. D1, p. D784–D788,
 2015.
- 524 GRAEF, T.; MOESTA, A. K.; NORMAN, P. J.; et al. KIR2DS4 is a product of gene
- conversion with KIR3DL2 that introduced specificity for HLA-A*11 while diminishing
 avidity for HLA-C. Journal of Experimental Medicine, 2009.
- 527 GUETHLEIN, L. A.; OLDER AGUILAR, A. M.; ABI-RACHED, L.; PARHAM, P.
- 528 Evolution of Killer Cell Ig-Like Receptor (KIR) Genes: Definition of an Orangutan KIR
- 529 Haplotype Reveals Expansion of Lineage III KIR Associated with the Emergence of MHC-C
- 530 . The Journal of Immunology, 2007.
- 531 GUMPERZ, J. E.; LITWIN, V.; PHILLIPS, J. H.; LANIER, L. L.; PARHAM, P. The Bw4
- public epitope of HLA-B molecules confers reactivity with natural killer cell clones that
 express NKB1, a putative HLA receptor. Journal of Experimental Medicine, 1995.
- 533 express NKB1, a putative HLA receptor. Journal of Experimental Medicine, 1995.
- HANSASUTA, P.; DONG, T.; THANANCHAI, H.; et al. Recognition of HLA-A3 and HLAA11 by KIR3DL2 is peptide-specific. European Journal of Immunology, 2004.
- 536 HERBERMAN, R. B.; HOLDEN, H. T. Natural cell-mediated immunity. Advances in
- 537 Cancer Research, 1978.
- 538 HIBY, S. E.; WALKER, J. J.; O'SHAUGHNESSY, K. M.; et al. Combinations of Maternal
- 539 KIR and Fetal HLA-C Genes Influence the Risk of Preeclampsia and Reproductive Success.
- 540 **The Journal of Experimental Medicine**, v. 200, n. 8, p. 957–965, 2004.
- 541 HILL, A. V. S. Defence by diversity. Nature, 1999.
- 542 HILTON, H. G.; GUETHLEIN, L. A.; GOYOS, A.; et al. Polymorphic HLA-C Receptors
- 543 Balance the Functional Characteristics of KIR Haplotypes. Journal of immunology, v. 195, 544 n. 7, p. 3160–3170, 2015.
- 545 HOLLENBACH, J. A.; NOCEDAL, I.; LADNER, M. B.; SINGLE, R. M.;
- 546 TRACHTENBERG, E. A. Killer cell immunoglobulin-like receptor (KIR) gene content
- 547 variation in the HGDP-CEPH populations. **Immunogenetics**, 2012.
- 548 HORTON, R.; WILMING, L.; RAND, V.; et al. Gene map of the extended human MHC.

- 549 Nature Reviews Genetics, 2004.
- 550 HOU, L. H.; CHEN, M.; JIANG, B.; et al. In contrast to other stimulatory natural killer cell
- 551 immunoglobulin-like receptor loci, several KIR2DS5 alleles predominate in African
- 552 Americans. Human Immunology, 2009.
- 553 IBGE. Censo 2010. 2013.
- 554 INSTITUTO SOCIOAMBIENTAL. Indigenous Lands in Brazil. Disponível em:
- 555 https://terrasindigenas.org.br/. Acesso em: 14/3/2020.
- 556 KÄRRE, K.; LJUNGGREN, H. G.; PIONTEK, G.; KIESSLING, R. Selective rejection of H-
- 557 2-deficient lymphoma variants suggests alternative immune defence strategy. Nature, v. 319,
- 558 n. 6055, p. 675–678, 1986. Disponível em:
- 559 <http://www.nature.com/doifinder/10.1038/319675a0>. .
- 560 KEHDY, F. S. G.; GOUVEIA, M. H.; MACHADO, M.; et al. Origin and dynamics of
- admixture in Brazilians and its effect on the pattern of deleterious mutations. **Proceedings of**
- the National Academy of Sciences of the United States of America, 2015.
- 563 KHAKOO, S. I.; RAJALINGAM, R.; SHUM, B. P.; et al. Rapid evolution of NK cell
 564 receptor systems demonstrated by comparison of chimpanzees and humans. Immunity, 2000.
- 565 KHAKOO, S. I.; THIO, C. L.; MARTIN, M. P.; et al. HLA and NK cell inhibitory receptor
- genes in resolving hepatitis C virus infection. Science (New York, N.Y.), v. 305, n. 5685, p.
 872–874, 2004.
- 568 KIM, S.; POURSINE-LAURENT, J.; TRUSCOTT, S. M.; et al. Licensing of natural killer
- cells by host major histocompatibility complex class I molecules. **Nature**, v. 436, n. 7051, p.
- 570 709–713, 2005. Disponível em: http://www.nature.com/doifinder/10.1038/nature03847>.
- 571 KUMAR, S.; STECHER, G.; LI, M.; KNYAZ, C.; TAMURA, K. MEGA X: Molecular
- 572 evolutionary genetics analysis across computing platforms. **Molecular Biology and**
- 573 **Evolution**, v. 35, n. 6, p. 1547–1549, 2018.
- 574 LAHIRI, D. K.; NURNBERGER, J. I. A rapid non-enzymatic method for the preparation of
- 575 HMW DNA from blood for RFLP studies. Nucleic acids research, v. 19, n. 19, p. 5444,
 576 1991.
- 577 LANIER, L. L. Up on the tightrope: natural killer cell activation and inhibition. **Nature**
- 578 **Immunology**, v. 9, n. 5, p. 495–502, 2008. Disponível em:
- 579 <http://www.nature.com/doifinder/10.1038/ni1581>...
- LANIER, L. L.; PHILLIPS, J. H. NK cell recognition of major histocompatibility complex
 class I molecules. Seminars in Immunology, v. 7, n. 2, p. 75–82, 1995.
- 582 LEATON, L. A.; SHORTT, J.; KICHULA, K. M.; et al. Conservation, extensive
- heterozygosity, and convergence of signaling potential all indicate a critical role for KIR3DL3
 in higher primates. Frontiers in Immunology, v. 10, p. 24, 2019.
- 585 LONG, E. O.; COLONNA, M.; LANIER, L. L. Inhibitory MHC class I receptors on NK and
- 586 T cells: a standard nomenclature. Immunology today, 1996.
- 587 MARRERO, A. R.; SILVA-JUNIOR, W. A; BRAVI, C. M.; et al. Demographic and
- 588 evolutionary trajectories of the Guarani and Kaingang natives of Brazil. American journal of
- 589 **physical anthropology**, v. 132, n. 2, p. 301–310, 2007.
- 590 MARTIN, M. P.; BASHIROVA, A.; TRAHERNE, J.; TROWSDALE, J.; CARRINGTON,
- 591 M. Cutting edge: expansion of the KIR locus by unequal crossing over. Journal of
- 592 immunology (Baltimore, Md. : 1950), v. 171, n. 5, p. 2192–2195, 2003.

593 MARTIN, M. P.; GAO, X.; LEE, J. H.; et al. Epistatic interaction between KIR3DS1 and 594 HLA-B delays the progression to AIDS. **Nature Genetics**, 2002.

595 MOESTA, A. K.; GRAEF, T.; ABI-RACHED, L.; et al. Humans Differ from Other Hominids

in Lacking an Activating NK Cell Receptor That Recognizes the C1 Epitope of MHC Class I.
 The Journal of Immunology, 2010.

- 598 MOESTA, A. K.; NORMAN, P. J.; YAWATA, M.; et al. Synergistic polymorphism at two
- 599 positions distal to the ligand-binding site makes KIR2DL2 a stronger receptor for HLA-C
- than KIR2DL3. Journal of immunology (Baltimore, Md.: 1950), v. 180, n. 6, p. 3969–79,
 2008.
- MORETTA, A.; BOTTINO, C.; VITALE, M.; et al. RECEPTORS FOR HLA CLASS-I
- MOLECULES IN HUMAN NATURAL KILLER CELLS. Annual Review of Immunology,1996.
- MORVAN, M. G.; LANIER, L. L. NK cells and cancer: You can teach innate cells new
 tricks. Nature Reviews Cancer, 2016.
- NEI, M. Analysis of gene diversity in subdivided populations. Proceedings of the National
 Academy of Sciences of the United States of America, 1973.
- NEI, M.; GOJOBORI, T. Simple methods for estimating the numbers of synonymous and
- 610 nonsynonymous nucleotide substitutions. **Molecular Biology and Evolution**, v. 3, n. 5, p.
- 611 418–26, 1986.
- 612 NELSON, G. W.; MARTIN, M. P.; GLADMAN, D.; et al. Cutting Edge: Heterozygote
- 613 Advantage in Autoimmune Disease: Hierarchy of Protection/Susceptibility Conferred by
- 614 HLA and Killer Ig-Like Receptor Combinations in Psoriatic Arthritis. The Journal of
- 615 **Immunology**, v. 173, p. 4273–4276, 2004.
- 616 NEMAT-GORGANI, N.; GUETHLEIN, L. A.; HENN, B. M.; et al. Diversity of KIR, HLA
- 617 Class I, and Their Interactions in Seven Populations of Sub-Saharan Africans. **The Journal of** 618 **Immunology**, v. 202, n. 9, p. 2636–2647, 2019.
- 619 NEMAT-GORGANI, N.; HILTON, H. G.; HENN, B. M.; et al. Different Selected
- 620 Mechanisms Attenuated the Inhibitory Interaction of KIR2DL1 with C2 + HLA-C in Two
- Indigenous Human Populations in Southern Africa . The Journal of Immunology, v. 200, n.
 8, p. 2640–2655, 2018.
- 623 NORMAN, P. J.; ABI-RACHED, L.; GENDZEKHADZE, K.; et al. Unusual selection on the
- 624 KIR3DL1/S1 natural killer cell receptor in Africans. Nature Genetics, v. 39, n. 9, p. 1092–9,
- 625 2007.
- 626 NORMAN, P. J.; HOLLENBACH, J. A.; NEMAT-GORGANI, N.; et al. Co-evolution of
- 627 Human Leukocyte Antigen (HLA) Class I Ligands with Killer-Cell Immunoglobulin-Like
- 628 Receptors (KIR) in a Genetically Diverse Population of Sub-Saharan Africans. **PLoS**
- 629 **Genetics**, 2013a.
- 630 NORMAN, P. J.; HOLLENBACH, J. A.; NEMAT-GORGANI, N.; et al. Co-evolution of
- 631 Human Leukocyte Antigen (HLA) Class I Ligands with Killer-Cell Immunoglobulin-Like
- 632 Receptors (KIR) in a Genetically Diverse Population of Sub-Saharan Africans. **PLoS**
- 633 **Genetics**, 2013b.
- NORMAN, P. J.; HOLLENBACH, J. A.; NEMAT-GORGANI, N.; et al. Defining KIR and
- 635 HLA Class I Genotypes at Highest Resolution via High-Throughput Sequencing. American
- 636 **Journal of Human Genetics**, v. 99, n. 2, p. 375–391, 2016.
- 637 O'FALLON, B. D.; FEHREN-SCHMITZ, L. Native Americans experienced a strong

- 638 population bottleneck coincident with European contact. Proceedings of the National
- 639 Academy of Sciences of the United States of America, 2011.
- 640 OLDER AGUILAR, A. M.; GUETHLEIN, L. A.; ADAMS, E. J.; et al. Coevolution of Killer
- 641 Cell Ig-Like Receptors with HLA-C To Become the Major Variable Regulators of Human NK
- 642 Cells. **The Journal of Immunology**, v. 185, n. 7, p. 4238–4251, 2010.
- PARADIS, E. Pegas: An R package for population genetics with an integrated-modular
 approach. Bioinformatics, 2010.
- PARADIS, E.; SCHLIEP, K. Ape 5.0: An environment for modern phylogenetics and
 evolutionary analyses in R. Bioinformatics, 2019.
- 647 PARHAM, P.; ARNETT, K. L.; ADAMS, E. J.; et al. Episodic evolution and turnover of
- 648 HLA-B in the indigenous human populations of the Americas. Tissue Antigens, 1997.
- 649 PETZL-ERLER, M. L.; LUZ, R.; SOTOMAIOR, V. S. The HLA polymorphsm of two
- distinctive South-American Indian tribes: The Kaingang and the Guarani. Tissue Antigens,1993.
- 652 PYO, C. W.; GUETHLEIN, L. A.; VU, Q.; et al. Different patterns of evolution in the
- 653 centromeric and telomeric regions of group A and B haplotypes of the human killer cell Ig-654 like receptor locus. PLoS ONE, 2010.
- ROBINSON, J.; BARKER, D. J.; GEORGIOU, X.; et al. IPD-IMGT/HLA Database. Nucleic
 acids research, 2020.
- ROBINSON, J.; HALLIWELL, J. A.; HAYHURST, J. D.; et al. The IPD and IMGT/HLA
 database: Allele variant databases. Nucleic Acids Research, v. 43, p. D423–D431, 2015.
- 659 SAITOU, N.; NEI, M. The neighbor-joining method: a new method for reconstructing
 660 phylogenetic trees. Molecular Biology and Evolution, v. 4, n. 4, p. 406–425, 1987.
- 661 SAKURAI, C.; ASARI, A. Y.; BELTRÃO, K. I.; et al. Resistência & integração 100 anos
 662 de imigração japonesa no Brasil. Rio de Janiero: IBGE, 2010.
- 663 SAMBROOK, J.; FRITSCH, E. F.; MANIATIS, T. Molecular cloning: a laboratory manual.
 664 Cold Spring Harbor Laboratory Press, 1989.
- 665 SAMBROOK, J. G.; BASHIROVA, A.; PALMER, S.; et al. Single haplotype analysis
- demonstrates rapid evolution of the killer immunoglobulin-like receptor (KIR) loci in
 primates. Genome Research, 2005.
- SCHMITT, R.; BONATTO, S. L.; FREITAS, L. B.; et al. Extremely limited mitochondrial
 DNA variability among the Aché Natives of Paraguay. Annals of Human Biology, 2004.
- 670 SIDNEY, J.; PETERS, B.; FRAHM, N.; BRANDER, C.; SETTE, A. HLA class I supertypes:
 671 A revised and updated classification. BMC Immunology, 2008.
- 672 SLIK, A. R. VAN DER; KOELEMAN, B. P. C.; VERDUIJN, W.; et al. KIR in type 1
- diabetes: Disparate distribution of activating and inhibitory natural killer cell receptors in patients versus HLA-matched control subjects. **Diabetes**, v. 52, n. 10, p. 2639–42, 2003.
- 675 SMYTH, M. J.; CRETNEY, E.; KELLY, J. M.; et al. Activation of NK cell cytotoxicity.
- 676 **Molecular Immunology**, 2005.
- 677 SOLLOCH, U. V; SCHEFZYK, D.; SCHÄFER, G.; MASSALSKI, C.; KOHLER, M.
- 678 Estimation of German KIR Allele Group Haplotype Frequencies. Frontiers in Immunology,
 679 v. 11, 2020.
- 680 TAJIMA, F. Statistical method for testing the neutral mutation hypothesis by DNA
- 681 polymorphism. Genetics, v. 123, n. 3, p. 585–95, 1989.

- 682 TARAZONA-SANTOS, E.; CARVALHO-SILVA, D. R.; PETTENER, D.; et al. Genetic
- 683 differentiation in South Amerindians is related to environmental and cultural diversity:
- Evidence from the Y chromosome. American Journal of Human Genetics, 2001.

TONEVA, M.; LEPAGE, V.; LAFAY, G.; et al. Genomic diversity of natural killer cell

- receptor genes in three populations. Tissue antigens, v. 57, n. 4, p. 358–62, 2001. Disponível
 em: http://www.ncbi.nlm.nih.gov/pubmed/11380947>.
- TRAHERNE, J. A.; MARTIN, M.; WARD, R.; et al. Mechanisms of copy number variation
 and hybrid gene formation in the KIR immune gene complex. Human Molecular Genetics,
 2010.
- TSUNETO, L. T.; PROBST, C. M.; HUTZ, M. H.; et al. HLA class II diversity in seven
- Amerindian populations. Clues about the origins of the Ach?? Tissue Antigens, v. 62, n. 6, p.
 512–526, 2003.
- 694 UHRBERG, M.; PARHAM, P.; WERNET, P. Definition of gene content for nine common
- group B haplotypes of the Caucasoid population: KIR haplotypes contain between seven and
 eleven KIR genes. Immunogenetics, v. 54, n. 4, p. 221–229, 2002.
- WANG, S.; LEWIS, C. M.; JAKOBSSON, M.; et al. Genetic variation and population
 structure in Native Americans. PLoS Genetics, 2007.
- WEIR, B. S.; COCKERHAM, C. C. Estimating F-Statistics for the Analysis of PopulationStructure. Evolution, 1984.
- WENDE, H.; COLONNA, M.; ZIEGLER, A.; VOLZ, A. Organization of the leukocyte
 receptor cluster (LRC) on human Chromosome 19q13.4. Mammalian Genome, v. 10, n. 2, p.
 154–160, 1999.
- WILSON, M. J.; TORKAR, M.; HAUDE, A.; et al. Plasticity in the organization and
- sequences of human KIR/ILT gene families. Proceedings of the National Academy of
 Sciences, 2000.
- 707 WILSON, M J; TORKAR, M.; HAUDE, A.; et al. Plasticity in the organization and
- sequences of human KIR/ILT gene families. Proceedings of the National Academy of
- 709 Sciences of the United States of America, v. 97, n. 9, p. 4778–83, 2000. Disponível em:
- 710 <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=18309&tool=pmcentrez&rendert 711 ype=abstract>. .
- 712 WRIGHT, S. Evolution and the Genetics of Populations: Variability within and among
- 713 natural populations. Chicago, 1978.
- 714 XIONG, S.; SHARKEY, A. M.; KENNEDY, P. R.; et al. Maternal uterine NK cell-activating
- 715 receptor KIR2DS1 enhances placentation. Journal of Clinical Investigation, v. 123, n. 10, p.
- 716 4264–4272, 2013.
- 717 YAWATA, M.; YAWATA, N.; DRAGHI, M.; et al. Roles for HLA and KIR polymorphisms
- in natural killer cell repertoire selection and modulation of effector function. **The Journal of**
- 719 **Experimental Medicine**, v. 203, n. 3, p. 633–645, 2006.



722 Figure 1. Number of KIR alleles which intersect in the 8 populations studied. Filled circles 723 indicate the populations in which a set of alleles was found, and the vertical bar plot above 724 indicate the size of the set of alleles. e.g.: 60 KIR alleles were found exclusively in the 725 population of Brazilians with European ancestry (CTBA). The number of alleles found in each population are indicated in the horizontal bar plot. Sets of alleles found exclusively in 726 727 Amerindians are highlighted in green, and in blue are the sets of alleles found in only in 728 Japanese or in Japanese and Amerindians. CTBA: Brazilians of European ancestry. JAP: 729 Brazilians of Japanese ancestry. ACHE: Aché. GKW: Guarani Kaiowá. GND: Guarani 730 Nandeva. GRC: Guarani M'bya. KIV: Kaingang from Ivaí. KRC: Kaingang from Rio das

731 Cobras.



732
733 Figure 2. Number of alleles found in the different populations for each *KIR* gene. The total
734 number of alleles per gene is indicated in the bottom axis. CTBA: Brazilians of European
735 ancestry. JAP: Brazilians of Japanese ancestry. ACHE: Aché. GKW: Guarani Kaiowá. GND:
736 Guarani Ñandeva. GRC: Guarani M'bya. KIV: Kaingang from Ivaí. KRC: Kaingang from
737 Rio das Cobras.



738

Figure 3. Allele frequencies of centromeric *KIR* genes (*KIR3DL3 ~ KIR2DL1*). Each graph
displays allele frequencies of a given *locus* in the eight populations studied. CTBA: Brazilians
of European ancestry. JAP: Brazilians of Japanese ancestry. ACHE: Aché. GKW: Guarani
Kaiowá. GND: Guarani Ñandeva. GRC: Guarani M'bya. KIV: Kaingang from Ivaí. KRC:

743 Kaingang from Rio das Cobras. Colors indicate increasing frequency (from yellow to red).

Relative frequencies are indicated inside each cell and empty cells correspond to alleles not

found in the population.





Figure 4. Allele frequencies of telomeric *KIR* genes (*KIR2DL4 ~ KIR3DL2*). Each graph
displays allele frequencies of a given *locus* in the eight populations studied. CTBA: Brazilians

of European ancestry. JAP: Brazilians of Japanese ancestry. ACHE: Aché. GKW: Guarani
Kaiowá. GND: Guarani Ñandeva. GRC: Guarani M'bya. KIV: Kaingang from Ivaí. KRC:
Kaingang from Rio das Cobras. Colors indicate increasing frequency (from yellow to red).
Relative frequencies are indicated inside each cell and empty cells correspond to alleles not
found in the population.



754

755 Figure 5. (A) Gray scale table of pairwise F_{ST} differentiation values (NEI, 1973). Values can indicate low (<0.05), medium (0.05 to 0.15) or high (>0.15) differentiation between 756 757 populations (WRIGHT, 1978). (B) Box plot showing the distribution of differentiation values 758 found for each population among the pairwise F_{ST} values. The line marks the median F_{ST} 759 value and the upper and bottom limits of the box mark the superior and inferior quartiles, 760 respectively. Whiskers represent minimum and maximum F_{ST} values found. (C) Average F_{ST} 761 differentiation in each KIR locus. The continuous line indicates the average F_{ST} across all 762 populations and the dotted line indicates the median F_{ST} . (D) Genetic distance tree of F_{ST} values estimated by the neighbor-joining (NJ) method (SAITOU; NEI, 1987). CTBA: 763 764 Brazilians of European ancestry. JAP: Brazilians of Japanese ancestry. ACHE: Aché. GKW: 765 Guarani Kaiowá. GND: Guarani Ñandeva. GRC: Guarani M'bya. KIV: Kaingang from Ivaí. 766 KRC: Kaingang from Rio das Cobras.



767

Principal component 1 (PC1) - 37.3%

768 Figure 6. Principal components analysis (PCA) of KIR variation. In PCA analysis, data containing allele frequencies of all *KIR* genes is reassigned to new variables which better 769 770 explain the distribution observed, which are called principal components. Each dot represents 771 the frequency of one allele, colored according to the population bearing that frequency. 772 Ellipsis show the distribution of KIR variance in each population, and rectangles sum the 773 diversity in that population. The first principal component (x axis) is plotted against the 774 second principal component (y axis), together they explain 42.6% of KIR variation. Relative 775 contribution of each component to overall variance (Eingenvalues) is given at the bottom left 776 graph, where each component correspond to a column, in order of percentage of variation 777 explained, from left to right (PC1, PC2, PC3...). CTBA: Brazilians of European ancestry. 778 JAP: Brazilians of Japanese ancestry. ACHE: Aché. GKW: Guarani Kaiowá. GND: Guarani 779 Ñandeva. GRC: Guarani M'bya. KIV: Kaingang from Ivaí. KRC: Kaingang from Rio das 780 Cobras.

	KIR3DL3	KIR2DS2	KIR2DL23	KIR2DL5B	KIR2DS35	KIR2DL1	СТВА	JAP	ACHE	GKW	GND	GRC	κιν	KRC
1	*00902		3*00101			•00302 — Cen A	3	14	14	75	35	25	48	38
2	*00901		3*00101			•00302 — Cen A	26	14	42	71	20	15	11	20
3	*00402	*00101	2*00301			Cen B2	5	3		29	21	64	47	41
4	*00802		3*00101			•00302 — Cen A	5	13	33	46	31	23	2	1
5	*01001		3*00101			*00302 — Cen A	1	23		3		28	43	18
6	*01002		3*009			•00302 — Cen A	1			35	14		3	1
7	*01002		3*00101			•00302 — Cen A		20			1	5	12	2
8	*01001		3*00201			•00202 — Cen A	17			2	10	1		
9	*01001		3*009			•00302 — Cen A				15	5			
10	*048		3*00101			*00302 — Cen A				10	7			
11	*00301	*00101	2*00301			Cen B2	8				1			1

В

Α

Telomeric KIR

		KIR2DL4	KIR3DL1S1	KIR2DL5A	KIR2DS5	KIR2DS1	KIR2DS4	KIR3DL2		JAP	СТВА	ACHE	GKW	GND	GRC	κιν	KRC
1	-	*00102	L1*01502				*00101	*00201	- Tel A	32	16	15	34	11	8	67	43
2	-	*00102	L1*02901				*00101	*00201	- Tel A	1		34	93	51	25	3	5
3	-	*00501	S1*01301	A*00101	5*00201	*00201		*00701	- Tel B	3	18	1	23	18	33	44	36
4	-	*00501	S1*01301	A*00101	5*00201	*00501		*00701	- Tel B		1	21	54	19		9	18
5	-	*011	L1*00501				*010	*01001	- Tel A	5	6		2	7	2	18	13
6	-	*00801	L1*00101				*00301	*00101	- Tel A	1	18		4	7			1
7	-	*00801	L1*00101				*00101	*00101	Tel A	2	10		8	10			
8	-	*00501	S1*01301	A*00103	5*00201	*00201		*00701	Tel B	1	1		1	7	4	6	6
9	-	*00102	L1*002				*00101	*00201	Tel A		25			1			
10	-	*00501	S1*01301	A*00501	3*00201	*00201		*00701	Tel B	3	6		3		13		
11	-	*00102	L1*01502				*00101	*00103	- Tel A	1		8			9	1	
12	-	*00501	S1*01301	A*00103	5*00201	*00201		*00201	- Tel B				13	3	2		
13	-	*00501	S1*01301	A*00101	5*00201	*00201		*00201	– Tel B	3		7	1		3		
14	-	*00501	S1*01301	A*00103	5*00201	*00201		*015	– Tel B				12	1			
15	-	*00602	L1*00701				*00401	*008	- Tel A	6	5			1			
16	-	*00802	L1*00401				*00601	*00501	– Tel A		11						
17	-	*00501	51*01301	A*00103	5*00201	*00201		*01001	_ Tel B					3	8		
18	-	*00102	L1*01501				*00101	*00103	_ Tel A						11		
19	-	*00103	L1*01502				*00101	*00201	_ Tel A	7	1		1				1
20	-	*00501	51*01301	A*00101	5*00201	*00502		*00701	- Tel B			4	5	1			
21	-	*00501	51*01301	A*01201	5*00201	*00201		*007101	– Tel B					1	6	2	1
		C				П						F					



781

Figure 7. Common *KIR* haplotypes in South Americans. Shown are the haplotypes observed in at least 10 individuals. (A) Centromeric and (B) telomeric haplotypes are shown. At the end of the haplotype, whether the haplotype belongs to *KIR A* or *B* is indicated. On the right side of each haplotype is the number of times each haplotype was found in the populations. (C) Bar plot showing the proportion of *KIR A* and *B* haplotypes in the eight South American populations. These are subclassified in (D) centromeric (*Cen A* and *Cen B*) and (E) telomeric

788 (*Tel A* and *Tel B*) regions. CTBA: Brazilians of European ancestry. JAP: Brazilians of

789 Japanese ancestry. ACHE: Aché. GKW: Guarani Kaiowá. GND: Guarani Ñandeva. GRC:

Guarani M'bya. KIV: Kaingang from Ivaí. KRC: Kaingang from Rio das Cobras. Refer to

791Supplementary Tables 2 and 3 for a list of all haplotypes found.



HLA-C

792

Figure 8. Frequencies of *HLA-C* lineages. Alleles with frequencies greater than 5% are above

the dotted horizontal line. CTBA: Brazilians of European ancestry. JAP: Brazilians of

Japanese ancestry. ACHE: Aché. GKW: Guarani Kaiowá. GND: Guarani Ñandeva. GRC:
 Guarani M'bya. KIV: Kaingang from Ivaí. KRC: Kaingang from Rio das Cobras.





Figure 9. Map displaying portion of South America with the sampling locations of the eight

populations included in this study indicated. CTBA: Brazilians of European ancestry. JAP:

800 Brazilians of Japanese ancestry. ACHE: Aché. GKW: Guarani Kaiowá. GND: Guarani

- 801 Ñandeva. GRC: Guarani M'bya. KIV: Kaingang from Ivaí. KRC: Kaingang from Rio das
- 802 Cobras.

	СТВА	JAP	GKW	GND	GRC	KIV	KRC	ACHE
СТВА	1							
JAP	0.7004	1						
GKW	0.6235	0.7148	1					
GND	0.6990	0.7579	0.8920	1				
GRC	0.6294	0.6564	0.8019	0.7890	1			
KIV	0.6723	0.7521	0.8224	0.8536	0.8425	1		
KRC	0.6383	0.6959	0.8323	0.8195	0.8743	0.9065	1	
ACHE	0.5918	0.6968	0.8530	0.8226	0.7270	0.7798	0.7462	1

Table 1. Proportion of shared *KIR* alleles among populations. These values are based on the

minimum frequency of each allele between two populations, the minor frequencies for all

alleles of a given *locus* are then summed, and this sum is averaged across all *KIR locus*.

806 CTBA: Brazilians of European ancestry. JAP: Brazilians of Japanese ancestry. ACHE: Aché.

807 GKW: Guarani Kaiowá. GND: Guarani Ñandeva. GRC: Guarani M'bya. KIV: Kaingang

808 from Ivaí. KRC: Kaingang from Rio das Cobras.

KIR Ligand	CTBA	JAP	ACHE	GKW	GND	GRC	KIV	KRC
A3/A11	0.17	0.14			0.01			0.04
Bw4	0.68	0.58	0.05	0.12	0.22	0.08	0.34	0.32
Bw4I	0.59	0.43	0.05	0.12	0.20	0.08	0.34	0.29
Bw4T	0.08	0.15			0.02			0.03
C1	0.58	0.89	0.55	0.64	0.59	0.85	0.52	0.59
C2	0.42	0.11	0.45	0.36	0.41	0.15	0.48	0.41

809 Table 2. Frequencies of *HLA* alleles that code for KIR ligands (i.e. not carrier frequency).

810 Bw4 is further subdivided in Bw4I (isoleucine at position 80) and Bw4T (threonine at

811 position 80), because this residue confers differential affinity of the epitope to distinct KIR.

812 Blank cells indicate the epitope was not found in the corresponding population. CTBA:

813 Brazilians of European ancestry. JAP: Brazilians of Japanese ancestry. ACHE: Aché. GKW:

814 Guarani Kaiowá. GND: Guarani Ñandeva. GRC: Guarani M'bya. KIV: Kaingang from Ivaí.

815 KRC: Kaingang from Rio das Cobras.

HLA									
epitope	KIR	CTBA	JAP	GRC	ACHE	GND	KIV	KRC	GKW
Bw4	KIR3DL1	184	132	34	66	33	33	4	8
A3	KIR3DL2	63	28	7		4	1		
A11	KIR3DL2	63	28	7		4	1		
A11	KIR2DS4	38	20	5		2			
C2	KIR2DL1	114	20	42	66	75	97	37	14
C1C2	KIR2DL2*	112	24	71	79	44	50	2	82
C1	KIR2DL3*	141	153	67	79	114	209	42	96
C2	KIR2DS1	41	4	38	47	48	82	25	19
C16	KIR2DS2	6					1		
C2	KIR2DS5*								
HLA-C *	KIR2DS4	72	42	41	84	29	32	31	12
Activating (%)		0.188	0.146	0.269	0.311	0.224	0.227	0.397	0.134
Inhibitory (%)		0.812	0.854	0.731	0.689	0.776	0.773	0.603	0.866
Mean int	eractions	7.65	6.09	4.88	4.53	4.36	3.37	2.76	2.75

816 Table 3. Number of functional interactions between KIR and HLA in South American

817 populations. Blank cells indicate that the interaction was not found in the corresponding

818 population. Asterisks indicate that only a subset of the molecules are considered to be

819 involved in the interaction, these are detailed in methods. The percentage of interactions

found in the population corresponding to activating or inhibitory interaction is also shown.

821 Mean interactions are given in the bottom line and indicate the average of functional pairs per

822 individual in each population. CTBA: Brazilians of European ancestry. JAP: Brazilians of

823 Japanese ancestry. ACHE: Aché. GKW: Guarani Kaiowá. GND: Guarani Ñandeva. GRC:

824 Guarani M'bya. KIV: Kaingang from Ivaí. KRC: Kaingang from Rio das Cobras.

Population		Primary language	Mean age (interval)	Municipality and state at the time of averiguation ¹	Latitu de	Longitu de	Sample size (n = 707)	
Euro- Brazilians	from Curitiba	Portugues e	25.95 (20 - 64)	Curitiba/PR	-25.45	-49.22	109	
	Kaiowá	Guarani	25.57 (12 - 66)	Limão Verde/MS Amambai/MS	-23.12 -23.06	-55.05 -55.12	150	
Guarani	M'bya	Guarani	-	Nova Laranjeiras/PR	-25.18	-52.32	84	
	Ñandeva	Guarani	-	Amambai/MS	-23.06	-55.12	81	
	Ivaí	Jê	-	Ivaí/PR	-24.5	-51.67	93	
Kaingang	Rio das Cobras	Jê	-	Nova Laranjeiras/PR	-25.18	-52.32	64	
Ashá	Doroguov	Achó		Arroyo Bandera/Paraguay	-23.3	-55.5	51	
Acne	Falaguay	Ache	-	Chupa Pou/Paraguay	-24.1	-56.3	51	
Japanese descendants	from Curitiba	Portugues e	30.62 (15 - 69)	Curitiba/PR	-25.45	-49.22	75	

825 Table 4. Characterization of the eight South American populations included in this study.

11 Supplementary material



Supplementary Figure 1. Four principal components of *KIR* variation. The four principal components are plotted against each other, from PC1 to PC4. Contribution of each component (*Eingenvalues*) to overall *KIR* variation is given in the diagonal axis. Each dot represents the frequency of one allele, colored according to the population bearing that frequency. Ellipsis show the distribution of *KIR* variance in each population, and rectangles sum the diversity in that population. CTBA: Brazilians of European ancestry. JAP: Brazilians of Japanese ancestry. ACHE: Aché. GKW: Guarani Kaiowá. GND: Guarani Ñandeva. GRC: Guarani M'bya. KIV: Kaingang from Ivaí. KRC: Kaingang from Rio das Cobras.



Supplementary Figure 2. Number of KIR and HLA interactions per individuals. The y axis indicates the number of interactions present in the same individual, values range from 1 to 15 different interactions possible. The x axis indicates the 8 populations included in this study. Each cell contain the number of individuals in a given population which present the same number of KIR and HLA interactions, this number is also indicated through the color scale, ranging from light yellow (0 individuals) to dark red (>40 individuals). CTBA: Brazilians of European ancestry. JAP: Brazilians of Japanese ancestry. GKW: Guarani Kaiowá. GND: Guarani Ñandeva. GRC: Guarani M'bya. KIV: Kaingang from Ivaí. KRC: Kaingang from Rio das Cobras. ACHE: Aché.

dN/dS	СТВА	JAP	GKW	GND	GRC	KIV	KRC	ACHE
KIR3DL3	ns	ns	ns	ns	ns	ns	ns	-1.99 (<i>p</i> =0.049)
KIR2DS2	ns	-	-	-	-	-	-	-
KIR2DL23	ns	ns	ns	ns	ns	ns	ns	ns
KIR2DL5B	ns	ns	ns	ns	ns	ns	ns	ns
KIR2DS35	ns	ns	ns	ns	ns	-	-	-
KIR2DL1	ns	ns	ns	ns	ns	ns	ns	ns
KIR2DL4	ns	ns	ns	ns	ns	ns	ns	ns
KIR3DL1S1	ns	ns	ns	ns	ns	ns	ns	ns
KIR2DL5A	ns	ns	ns	ns	ns	ns	ns	ns
KIR2DS5	-	-	-	-	-	-	-	-
KIR2DS1	ns	ns	ns	ns	ns	ns	ns	-
KIR2DS4	ns	ns	ns	ns	ns	ns	ns	ns
KIR3DL2	ns	ns	ns	ns	ns	ns	ns	ns

Supplementary Table 1. Codon-based test of neutrality across all allele pairs in a given *KIR* locus. This test is based on the ratio of non-synonymous (dN) to synonymous (dS) substitutions. Significance values were based on 500 replicates based on the bootstrap method. ns: non-significant departure from expected neutrality values. When significant, the dN/dS ratio is given, followed by its p-value in parenthesis. CTBA: Brazilians of European ancestry. JAP: Brazilians of Japanese ancestry. GKW: Guarani Kaiowá. GND: Guarani Ñandeva. GRC: Guarani M'bya. KIV: Kaingang from Ivaí. KRC: Kaingang from Rio das Cobras. ACHE: Aché.

Supplementary Table 2. All centromeric *KIR* haplotypes (*KIR3DL3* ~ *KIR2DL1*) found in South American populations. On the right side of each haplotype is whether the haplotype belongs to *KIR A* or *B* and the number of times the haplotype was found in each population. CTBA: Brazilians of European ancestry. JAP: Brazilians of Japanese ancestry. ACHE: Aché. GKW: Guarani Kaiowá. GND: Guarani Ñandeva. GRC: Guarani M'bya. KIV: Kaingang from Ivaí. KRC: Kaingang from Rio das Cobras.

Supplementary Table 3. All telomeric *KIR* haplotypes (*KIR2DL4* ~ *KIR3DL2*) found in South American populations. On the right side of each haplotype is whether the haplotype belongs to *KIR A* or *B* and the number of times the haplotype was found in each population. CTBA: Brazilians of European ancestry. JAP: Brazilians of Japanese ancestry. ACHE: Aché. GKW: Guarani Kaiowá. GND: Guarani Ñandeva. GRC: Guarani M'bya. KIV: Kaingang from Ivaí. KRC: Kaingang from Rio das Cobras.

6 CONSIDERAÇÕES FINAIS

Neste trabalho, nós caracterizamos a diversidade de *KIR* em oito populações sul-americanas. Assim como é observado para o restante do genoma (DEGIORGIO et al., 2009), as populações ameríndias apresentaram menor diversidade em *KIR* e em interações KIR e HLA quando comparadas a outras populações. Foram encontrados cerca de 69 alelos *KIR* por população, predominantemente ligantes HLA-C. Além disso, populações ameríndias tem baixas frequências dos ligantes A3, A11 e Bw4. Nós reforçamos a sugestão de GENDZEKHADZE et al. (2009) de que a diversidade mantida nos indígenas representa o mínimo necessário para manter as funções das células NK em populações humanas. Portanto, esses achados reforçam as evidências de que HLA-C tenha se especializado na função de receptor de NK, em contraste a HLA-A e -B terem mantido sua principal função na imunidade adaptativa de células B e T. Encontramos ainda indícios que fortalecem as observações de coevolução entre HLA-C e o receptor que o reconhece mais fortemente, o KIR2DL1.

As principais conclusões desse trabalho foram:

- Encontramos 209 alelos entre os 13 genes KIR em populações sul-americanas, organizados em 37 haplótipos centroméricos e 44 haplótipos teloméricos;
- A menor diversidade em Ameríndios é concentrada em interações com o HLA-C, o que evidencia que este é o principal ligante KIR;
- O polimorfismo de *KIR2DL1* apresenta sinais de coevolução com o HLA-C, pois:
 - A variante *rs2304224* se relaciona com a expressão do receptor KIR2DL1 e do seu ligante HLA-C, apenas em indivíduos que possuem o ligante funcional C2;
 - A variante *rs2304224* está em desequilíbrio de ligação com duas outras variantes, *rs4806553G* e *rs687000G*, que apresentam sinais de seleção positiva.

Nosso estudo representa um grande avanço na área de imunogenética. Apesar da limitação da ausência dos pseudogenes *KIR2DP1* e *KIR3DP1*, a caracterização dos haplótipos de *KIR* em máxima resolução apresentada nesse trabalho poderá ser usada como referência em estudos futuros. A análise alélica de *KIR* é ainda incipiente devido à grande dificuldade de sua análise e, como evidenciado nesse trabalho, esta não é uma tarefa fácil. Esperamos que os resultados aqui apresentados possam servir
de base para a compreensão da complexa e fascinante diversidade de genes KIR, fornecendo base para trabalhos evolutivos e também na interpretação de como KIR e HLA afetam a reprodução humana e a suscetibilidade a doenças.

REFERÊNCIAS

ADAMACK, A. T.; GRUBER, B. PopGenReport: Simplifying basic population genetic analyses in R. **Methods in Ecology and Evolution**, 2014.

AMORIM, L. M.; TONG, H. VAN; HOAN, N. X.; et al. KIR-HLA distribution in a Vietnamese population from Hanoi. **Human Immunology**, v. 79, n. 2, p. 93–100, 2018. Disponível em:

<https://www.sciencedirect.com/science/article/pii/S0198885917305529?via%3Dihub >...

AMOS, W.; HOFFMAN, J. I. Evidence that two main bottleneck events shaped modern human genetic diversity. **Proceedings of the Royal Society B: Biological Sciences**, 2010.

ANFOSSI, N.; ANDRÉ, P.; GUIA, S.; et al. Human NK Cell Education by Inhibitory Receptors for MHC Class I. **Immunity**, 2006.

APPS, R.; MENG, Z.; PRETE, G. Q. DEL; et al. Relative expression levels of the HLA class-I proteins in normal and HIV-infected cells. **Journal of immunology** (**Baltimore, Md. : 1950**), v. 194, n. 8, p. 3594–600, 2015. Disponível em: http://www.ncbi.nlm.nih.gov/pubmed/25754738%5Cnhttp://www.pubmedcentral.nih.gov/pubmed/25754738%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4390493>...

APPS, R.; QI, Y.; CARLSON, J. M.; et al. Influence of HLA-C expression level on HIV control. **Science**, 2013.

AUGUSTO, D. G. The impact of KIR polymorphism on the risk of developing cancer: Not as strong as imagined? **Frontiers in Genetics**, 2016.

AUGUSTO, D. G.; AMORIM, L. M.; FARIAS, T. D. J.; PETZL-ERLER, M. L. KIR and HLA genotyping of Japanese descendants from Curitiba, a city of predominantly European ancestry from Southern Brazil. **Human Immunology**, 2016.

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. **Human Immunology**, 2015.

AUGUSTO, DANILLO G.; LOBO-ALVES, S. C.; MELO, M. F.; PEREIRA, N. F.; PETZL-ERLER, M. L. Activating KIR and HLA Bw4 ligands are associated to decreased susceptibility to pemphigus foliaceus, an autoimmune blistering skin disease. **PLoS ONE**, 2012a.

AUGUSTO, DANILLO G.; LOBO-ALVES, S. C.; MELO, M. F.; PEREIRA, N. F.; PETZL-ERLER, M. L. 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, 2012b.

AUGUSTO, D. G.; NORMAN, P. J.; DANDEKAR, R.; HOLLENBACH, J. A.

Fluctuating and geographically specific selection characterize rapid evolution of the human Kir region. **Frontiers in Immunology**, 2019.

AUGUSTO, D. G.; O'CONNOR, G. M.; LOBO-ALVES, S. C.; et al. Pemphigus is associated with KIR3DL2 expression levels and provides evidence that KIR3DL2 may bind HLA-A3 and A11 in vivo. **European Journal of Immunology**, 2015.

AUGUSTO, D. G.; PETZL-ERLER, M. L. KIR and HLA under pressure: evidences of coevolution across worldwide populations. **Human Genetics**, 2015.

AUGUSTO, D. G.; PIOVEZAN, B. Z.; TSUNETO, L. T.; CALLEGARI-JACQUES, S. M.; PETZL-ERLER, M. L. KIR Gene Content in Amerindians Indicates Influence of Demographic Factors. **PLoS ONE**, 2013a.

AUGUSTO, D. G.; PIOVEZAN, B. Z.; TSUNETO, L. T.; CALLEGARI-JACQUES, S. M.; PETZL-ERLER, M. L. KIR Gene Content in Amerindians Indicates Influence of Demographic Factors. **PLoS ONE**, 2013b.

AUGUSTO, D. G.; ZEHNDER-ALVES, L.; PINCERATI, M. R.; et al. Diversity of the KIR gene cluster in an urban Brazilian population. **Immunogenetics**, v. 64, n. 2, p. 143–152, 2012.

BARI, R.; BELL, T.; LEUNG, W. H.; et al. Significant functional heterogeneity among KIR2DL1 alleles and a pivotal role of arginine245. **Blood**, 2009.

BATTILANA, J.; BONATTO, S. L.; FREITAS, L. B.; et al. Alu insertions versus blood group plus protein genetic variability in four Amerindian populations. **Annals of Human Biology**, 2002.

BENNETT, E. M.; BENNINK, J. R.; YEWDELL, J. W.; BRODSKY, F. M. Cutting edge: adenovirus E19 has two mechanisms for affecting class I MHC expression. **Journal of immunology (Baltimore, Md. : 1950)**, 1999.

BÉZIAT, V.; TRAHERNE, J. A.; LIU, L. L.; et al. Influence of KIR gene copy number on natural killer cell education. **Blood**, v. 121, n. 23, p. 4703–4707, 2013.

BIRON, C. A. Activation and function of natural killer cell responses during viral infections. **Current opinion in immunology**, v. 9, n. 1, p. 24–34, 1997.

BLOKHUIS, J. H.; HILTON, H. G.; GUETHLEIN, L. A.; et al. KIR2DS5 allotypes that recognize the C2 epitope of HLA-C are common among Africans and absent from Europeans. **Immunity Inflammation and Disease**, 2017.

BRAUN-PRADO, K.; VIEIRA MION, A. L.; FARAH PEREIRA, N.; CULPI, L.; PETZL-ERLER, M. L. HLA class I polymorphism, as characterised by PCR-SSOP, in a Brazilian exogamic population. **Tissue Antigens**, 2000.

BULMER, J. N.; LASH, G. E. The role of uterine NK cells in normal reproduction and reproductive disorders. **Advances in Experimental Medicine and Biology**, 2015.

BUNCE, M.; O'NEILL, C. M.; BARNARDO, M. C. N. M.; et al. Phototyping: comprehensive DNA typing for HLA-A, B, C, DRB1, DRB3, DRB4, DRB5 & DQB1 by

PCR with 144 primer mixes utilizing sequence-specific primers (PCR-SSP). **Tissue Antigens**, 1995.

CALLEGARI-JACQUES, S. M.; CROSSETTI, S. G.; KOHLRAUSCH, F. B.; et al. The β -globin gene cluster distribution revisited - Patterns in native American populations. **American Journal of Physical Anthropology**, 2007.

CARR, W. H.; PANDO, M. J.; PARHAM, P. KIR3DL1 Polymorphisms That Affect NK Cell Inhibition by HLA-Bw4 Ligand. **The Journal of Immunology**, 2005.

CARRINGTON, M.; WANG, S.; MARTIN, M. P.; et al. Hierarchy of resistance to cervical neoplasia mediated by combinations of killer immunoglobulin-like receptor and human leukocyte antigen loci. **The Journal of Experimental Medicine**, 2005.

CATANIO, P. A. G. O caminho que o dekassegui sonhou (dekassegui no yumêji) : cultura e subjetividade no movimento dekassegui. Londrina: FAPESP, 2002.

CAVALLI-SFORZA, L. L.; CAVALLI-SFORZA, L.; MENOZZI, P.; PIAZZA, A. **The History and Geography of Human Genes**. Princeton, New Jersey: Princeton University Press, 1994.

CAVALLI-SFORZA, L. L.; PIAZZA, A.; MENOZZI, P.; MOUNTAIN, J. Reconstruction of human evolution: Bringing together genetic, archaeological, and linguistic data. **Proceedings of the National Academy of Sciences of the United States of America**, 1988.

CHAROUDEH, H. N.; SCHMIED, L.; GONZALEZ, A.; et al. Quantity of HLA-C surface expression and licensing of KIR2DL+ natural killer cells. **Immunogenetics**, 2012.

CICCONE, E.; PENDE, D.; VIALE, O.; et al. Evidence of a natural killer (NK) cell repertoire for (Allo) antigen recognition: Definition of five distinct NK-determined allospecificities in humans. **Journal of Experimental Medicine**, 1992.

CONROY, G. C.; WEBER, G. W.; SEIDLER, H.; et al. Genetic structure of Quechuaspeakers of the Central Andes and geographic patterns of gene frequencies in South Amerindian populations. **American Journal of Physical Anthropology**, 2000.

CONWAY, J. R.; LEX, A.; GEHLENBORG, N. UpSetR: An R package for the visualization of intersecting sets and their properties. **Bioinformatics**, 2017.

COOLEY, S.; WEISDORF, D. J.; GUETHLEIN, L. A.; et al. Donor selection for natural killer cell receptor genes leads to superior survival after unrelated transplantation for acute myelogenous leukemia. **Blood**, v. 116, n. 14, p. 2411–2419, 2010.

DAVID, G.; DJAOUD, Z.; WILLEM, C.; et al. Large Spectrum of HLA-C Recognition by Killer Ig–like Receptor (KIR)2DL2 and KIR2DL3 and Restricted C1 Specificity of KIR2DS2: Dominant Impact of KIR2DL2/KIR2DS2 on KIR2D NK Cell Repertoire Formation. **The Journal of Immunology**, 2013.

DEGIORGIO, M.; JAKOBSSON, M.; ROSENBERG, N. A. Explaining worldwide

patterns of human genetic variation using a coalescent-based serial founder model of migration outward from Africa. **Proceedings of the National Academy of Sciences of the United States of America**, 2009.

DINNO, A. dunn.test: Dunn's Test of Multiple Comparisons Using Rank Sums., 2017.

DÖHRING, C.; SCHEIDEGGER, D.; SAMARIDIS, J.; CELLA, M.; COLONNA, M. A human killer inhibitory receptor specific for HLA-A1,2. **Journal of immunology (Baltimore, Md.: 1950)**, 1996.

DOKUN, A. O.; KIM, S.; SMITH, H. R. C.; et al. Specific and nonspecific NK cell activation during virus infection. **Nature Immunology**, 2001.

DRAY, S.; DUFOUR, A. B. The ade4 package: Implementing the duality diagram for ecologists. **Journal of Statistical Software**, 2007.

DUNPHY, S. E.; GUINAN, K. J.; CHORCORA, C. N.; et al. 2DL1, 2DL2 and 2DL3 all contribute to KIR phenotype variability on human NK cells. **Genes and Immunity**, , n. March, p. 1–10, 2015. Disponível em: http://www.ncbi.nlm.nih.gov/pubmed/25950617>.

EKLUND, A. beeswarm: The Bee Swarm Plot, an Alternative to Stripchart. , 2016. Disponível em: https://cran.r-project.org/package=beeswarm.

FAAS, M. M.; VOS, P. DE. Uterine NK cells and macrophages in pregnancy. **Placenta**, 2017.

FAGERBERG, L.; HALLSTROM, B. M.; OKSVOLD, P.; et al. Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. **Molecular and Cellular Proteomics**, 2014.

FAN, Q. R.; GARBOCZI, D. N.; WINTER, C. C.; et al. Direct binding of a soluble natural killer cell inhibitory receptor to a soluble human leukocyte antigen-Cw4 class I major histocompatibility complex molecule. **Proceedings of the National Academy of Sciences of the United States of America**, 1996.

FAN, Q. R.; LONG, E. O.; WILEY, D. C. Crystal structure of the human natural killer cell inhibitory receptor KIR2DL1-HLA-Cw4 complex. **Nature immunology**, 2001.

FAURE, M.; LONG, E. O. KIR2DL4 (CD158d), an NK cell-activating receptor with inhibitory potential. **Journal of immunology**, v. 168, n. 12, p. 6208–6214, 2002.

FENG, J.; GARRITY, D.; CALL, M. E.; MOFFETT, H.; WUCHERPFENNIG, K. W. Convergence on a distinctive assembly mechanism by unrelated families of activating immune receptors. **Immunity**, v. 22, n. 4, p. 427–438, 2005.

FLÓREZ-ÁLVAREZ, L.; HERNANDEZ, J. C.; ZAPATA, W. NK cells in HIV-1 infection: From basic science to vaccine strategies. **Frontiers in Immunology**, 2018.

FOLEY, B. A.; SANTIS, D. DE; BEELEN, E. VAN; et al. The reactivity of Bw4+ HLA-B and HLA-A alleles with kir3dll: Implications for patient and donor suitability for haploidentical stem cell transplantations. Blood, 2008.

FRANCESCHI, D. S. A.; MAZINI, P. S.; RUDNICK, C. C. C.; et al. Association between killer-cell immunoglobulin-like receptor genotypes and leprosy in Brazil. **Tissue Antigens**, 2008.

FUKAMI-KOBAYASHI, K.; SHIINA, T.; ANZAI, T.; et al. Genomic evolution of MHC class I region in primates. **Proceedings of the National Academy of Sciences of the United States of America**, 2005.

GARRIDO, F.; RUIZ-CABELLO, F.; CABRERA, T.; et al. Implications for immunosurveillance of altered HLA class I phenotypes in human tumours. **Immunology Today**, 1997.

GASPAR, P. A.; HUTZ, M. H.; SALZANO, F. M.; et al. Polymorphisms of CYP1A1, CYP2E1, GSTM1, GSTT1, and TP53 genes in Amerindians. **American Journal of Physical Anthropology**, 2002.

GAUTIER, M.; KLASSMANN, A.; VITALIS, R. rehh 2.0: a reimplementation of the R package rehh to detect positive selection from haplotype structure. Molecular Ecology Resources. **Anais...**, 2017.

GAZIT, R.; GARTY, B. Z.; MONSELISE, Y.; et al. Expression of KIR2DL1 on the entire NK cell population: A possible novel immunodeficiency syndrome [1]. **Blood**, 2004.

GENDZEKHADZE, K.; NORMAN, P. J.; ABI-RACHED, L.; et al. Co-evolution of KIR2DL3 with HLA-C in a human population retaining minimal essential diversity of KIR and HLA class I ligands. **Proceedings of the National Academy of Sciences of the United States of America**, 2009.

GENDZEKHADZE, K.; NORMAN, P. J.; ABI-RACHED, L.; LAYRISSE, Z.; PARHAM, P. High KIR diversity in Amerindians is maintained using few gene-content haplotypes. **Immunogenetics**, 2006.

GODFREY, D.; MACDONALD, H. R.; KRONENBERG, M.; SMYTH, M. J.; KAER, L. VAN. NKT Cells: What's in a Name? **Nature Reviews Immunology**, v. 4, n. 3, p. 231–7, 2004.

GONZÁLEZ-GALARZA, F. F.; TAKESHITA, L. Y. C.; SANTOS, E. J. M.; et al. Allele frequency net 2015 update: New features for HLA epitopes, KIR and disease and HLA adverse drug reaction associations. **Nucleic Acids Research**, v. 43, n. D1, p. D784–D788, 2015.

GOURRAUD, P. A.; KHANKHANIAN, P.; CEREB, N.; et al. HLA diversity in the 1000 genomes dataset. **PLoS ONE**, 2014.

GRAEF, T.; MOESTA, A. K.; NORMAN, P. J.; et al. KIR2DS4 is a product of gene conversion with KIR3DL2 that introduced specificity for HLA-A*11 while diminishing avidity for HLA-C. **Journal of Experimental Medicine**, 2009.

GRÉGOIRE, C.; CHASSON, L.; LUCI, C.; et al. The trafficking of natural killer cells.

108

Immunological Reviews, 2007.

GROSS, J.; LIGGES, U. nortest: Tests for Normality., 2015.

GUETHLEIN, L. A.; ABI-RACHED, L.; HAMMOND, J. A.; PARHAM, P. The expanded cattle KIR genes are orthologous to the conserved single-copy KIR3DX1 gene of primates. **Immunogenetics**, 2007.

GUETHLEIN, L. A.; NORMAN, P. J.; HILTON, H. H. G.; PARHAM, P. Co-evolution of MHC class I and variable NK cell receptors in placental mammals. **Immunological Reviews**, v. 267, n. 1, 2015.

GUETHLEIN, L. A.; OLDER AGUILAR, A. M.; ABI-RACHED, L.; PARHAM, P. Evolution of Killer Cell Ig-Like Receptor (KIR) Genes: Definition of an Orangutan KIR Haplotype Reveals Expansion of Lineage III KIR Associated with the Emergence of MHC-C. **The Journal of Immunology**, 2007.

GUINAN, K. J.; CUNNINGHAM, R. T.; MEENAGH, A; et al. Receptor systems controlling natural killer cell function are genetically stratified in Europe. **Genes and immunity**, v. 11, n. 1, p. 67–78, 2010. Disponível em: http://www.ncbi.nlm.nih.gov/pubmed/19693090>.

GUMPERZ, J. E.; LITWIN, V.; PHILLIPS, J. H.; LANIER, L. L.; PARHAM, P. The Bw4 public epitope of HLA-B molecules confers reactivity with natural killer cell clones that express NKB1, a putative HLA receptor. **Journal of Experimental Medicine**, 1995.

GUTIÉRREZ-RODRÍGUEZ, M. E.; SANDOVAL-RAMÍREZ, L.; DÍAZ-FLORES, M.; et al. KIR Gene in Ethnic and Mestizo Populations from Mexico. **Human Immunology**, v. 67, n. 1–2, p. 85–93, 2006.

HANSASUTA, P.; DONG, T.; THANANCHAI, H.; et al. Recognition of HLA-A3 and HLA-A11 by KIR3DL2 is peptide-specific. **European Journal of Immunology**, 2004.

HE, Y.; TAO, S.; YING, Y.; et al. Allelic polymorphism, mRNA and antigen expression of KIR2DL1 in the Chinese Han population. **Human Immunology**, 2014.

HEIDENREICH, S.; KRÖGER, N. Reduction of relapse after unrelated donor stem cell transplantation by KIR-Based graft selection. **Frontiers in Immunology**, 2017.

HERBERMAN, R. B.; HOLDEN, H. T. Natural cell-mediated immunity. **Advances in Cancer Research**, 1978.

HERBERMAN, R.; ORTALDO, J. Natural killer cells: their roles in defenses against disease. **Science**, v. 214, n. 4516, p. 24–30, 1981.

HIBY, S. E.; WALKER, J. J.; O'SHAUGHNESSY, K. M.; et al. Combinations of Maternal KIR and Fetal HLA-C Genes Influence the Risk of Preeclampsia and Reproductive Success. **The Journal of Experimental Medicine**, 2004.

HILL, A. V. S. Defence by diversity. Nature, 1999.

HILL, K.; HURTADO, A. M. Ache Life History: The Ecology and Demography of a

Foraging People. 1996.

HILTON, HUGO G; GUETHLEIN, L. A.; GOYOS, A.; et al. Polymorphic HLA-C Receptors Balance the Functional Characteristics of KIR Haplotypes. **Journal of immunology**, , n. 195, 2015.

HILTON, HUGO G.; NORMAN, P. J.; NEMAT-GORGANI, N.; et al. Loss and Gain of Natural Killer Cell Receptor Function in an African Hunter-Gatherer Population. **PLoS Genetics**, 2015.

HILTON, H. G.; PARHAM, P. Missing or altered self: human NK cell receptors that recognize HLA-C. **Immunogenetics**, 2017.

HOLLENBACH, J. A.; AUGUSTO, D. G.; ALAEZ, C.; et al. 16th IHIW: Population Global Distribution of Killer Immunoglobulin-like Receptor (KIR) and Ligands. International Journal of Immunogenetics. **Anais...**, 2013.

HOLLENBACH, J. A.; NOCEDAL, I.; LADNER, M. B.; SINGLE, R. M.; TRACHTENBERG, E. A. Killer cell immunoglobulin-like receptor (KIR) gene content variation in the HGDP-CEPH populations. **Immunogenetics**, 2012.

HONG, H. A.; LOUBSER, A. S.; ASSIS ROSA, D. DE; et al. Killer-cell immunoglobulin-like receptor genotyping and HLA killer-cell immunoglobulin-like receptor-ligand identification by real-time polymerase chain reaction. **Tissue Antigens**, 2011.

HOROWITZ, A.; DJAOUD, Z.; NEMAT-GORGANI, N.; et al. Class I HLA haplotypes form two schools that educate NK cells in different ways. **Science Immunology**, 2016.

HORTON, R.; WILMING, L.; RAND, V.; et al. Gene map of the extended human MHC. **Nature Reviews Genetics**, 2004.

HUMBERT, P. O.; ROGERS, C.; GANIATSAS, S.; et al. E2F4 is essential for normal erythrocyte maturation and neonatal viability. **Molecular Cell**, 2000.

IBGE. Censo 2010. 2013.

IKEDA, H.; LETHÉ, B.; LEHMANN, F.; et al. Characterization of an antigen that is recognized on a melanoma showing partial HLA loss by CTL expressing an NK inhibitory receptor. **Immunity**, v. 6, n. 2, p. 199–208, 1997.

INSTITUTO SOCIOAMBIENTAL. Indigenous Lands in Brazil. Disponível em: https://terrasindigenas.org.br/. Acesso em: 14/3/2020.

ISHIDO, S.; WANG, C.; LEE, B.-S.; COHEN, G. B.; JUNG, J. U. Downregulation of Major Histocompatibility Complex Class I Molecules by Kaposi's Sarcoma-Associated Herpesvirus K3 and K5 Proteins. **Journal of Virology**, 2000.

JOBIM, M. R.; JOBIM, M.; SALIM, P. H.; et al. Analysis of KIR gene frequencies and HLA class I genotypes in breast cancer and control group. **Human Immunology**, 2013.

JONES, T. R.; HANSON, L. K.; SUN, L.; et al. Multiple independent loci within the human cytomegalovirus unique short region down-regulate expression of major histocompatibility complex class I heavy chains. **Journal of virology**, 1995.

JR, F. E. H.; DUPONT, C. Hmisc: Harrell Miscellaneous. , 2019. Disponível em: https://cran.r-project.org/package=Hmisc.

JUNEVIK, K.; WERLENIUS, O.; HASSELBLOM, S.; et al. The expression of NK cell inhibitory receptors on cytotoxic T cells in B-cell chronic lymphocytic leukaemia (B-CLL). **Annals of Hematology**, 2007.

KÄMMERER, U.; MARZUSCH, K.; KRÖBER, S.; et al. A subset of CD56+ large granular lymphocytes in first-trimester human decidua are proliferating cells. **Fertility and Sterility**, v. 71, n. 1, p. 74–79, 1999.

KANEKO, K.; ISHIGAMI, S.; KIJIMA, Y.; et al. Clinical implication of HLA class I expression in breast cancer. **BMC Cancer**, 2011.

KÄRRE, K.; LJUNGGREN, H. G.; PIONTEK, G.; KIESSLING, R. Selective rejection of H–2-deficient lymphoma variants suggests alternative immune defence strategy. **Nature**, v. 319, n. 6055, p. 675–678, 1986. Disponível em: <http://www.nature.com/doifinder/10.1038/319675a0>. .

KEHDY, F. S. G.; GOUVEIA, M. H.; MACHADO, M.; et al. Origin and dynamics of admixture in Brazilians and its effect on the pattern of deleterious mutations. **Proceedings of the National Academy of Sciences of the United States of America**, 2015.

KHAKOO, S. I.; CARRINGTON, M. KIR and disease: A model system or system of models? **Immunological Reviews**, 2006.

KHAKOO, S. I.; RAJALINGAM, R.; SHUM, B. P.; et al. Rapid evolution of NK cell receptor systems demonstrated by comparison of chimpanzees and humans. **Immunity**, 2000.

KHAKOO, S. I.; THIO, C. L.; MARTIN, M. P.; et al. HLA and NK cell inhibitory receptor genes in resolving hepatitis C virus infection. **Science (New York, N.Y.)**, 2004.

KIESSLING, R.; KLEIN, E.; PROSS, H.; WIGZELL, H. "Natural" killer cells in the mouse. II. Cytotoxic cells with specificity for mouse Moloney leukemia cells. Characteristics of the killer cell. **European Journal of Immunology**, 1975.

KIM, S.; POURSINE-LAURENT, J.; TRUSCOTT, S. M.; et al. Licensing of natural killer cells by host major histocompatibility complex class I molecules. **Nature**, v. 436, n. 7051, p. 709–713, 2005. Disponível em: http://www.nature.com/doifinder/10.1038/nature03847>.

KING, A.; LOKE, Y. W. On the nature and function of human uterine granular lymphocytes. **Immunology Today**, 1991.

KULKARNI, S.; MARTIN, M. P.; CARRINGTON, M. KIR genotyping by multiplex

PCR-SSP. Methods in molecular biology (Clifton, N.J.), 2010.

LAHIRI, D. K.; NURNBERGER, J. I. A rapid non-enzymatic method for the preparation of HMW DNA from blood for RFLP studies. **Nucleic acids research**, v. 19, n. 19, p. 5444, 1991.

LANIER, L. L. NK CELL RECEPTORS. Annual Review of Immunology, 1998.

LANIER, L. L. Up on the tightrope: natural killer cell activation and inhibition. **Nature Immunology**, v. 9, n. 5, p. 495–502, 2008. Disponível em: <http://www.nature.com/doifinder/10.1038/ni1581>..

LANIER, L. L.; PHILLIPS, J. H. NK cell recognition of major histocompatibility complex class I molecules. **Seminars in Immunology**, v. 7, n. 2, p. 75–82, 1995.

LANIER, L. L.; TESTI, R.; BINDL, J.; PHILLIPS, J. H. Identity of Leu-19 (CD56) leukocyte differentiation antigen and neural cell adhesion molecule. **The Journal of experimental medicine**, v. 169, n. 6, p. 2233–8, 1989. Disponível em: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2189344&tool=pmcentre z&rendertype=abstract>. .

LEE, C.; HUANG, C. H. LASAGNA-search: An integrated web tool for transcription factor binding site search and visualization. **BioTechniques**, 2013.

LESURF, R.; COTTO, K. C.; WANG, G.; et al. ORegAnno 3.0: A community-driven resource for curated regulatory annotation. **Nucleic Acids Research**, 2016.

LJUNGGREN, H. G.; KÄRRE, K. In search of the "missing self": MHC molecules and NK cell recognition. **Immunology Today**, 1990.

LONG, E. O.; COLONNA, M.; LANIER, L. L. Inhibitory MHC class I receptors on NK and T cells: a standard nomenclature. **Immunology today**, 1996.

LUDAJIC, K.; BALAVARCA, Y.; BICKEBÖLLER, H.; et al. KIR genes and KIR ligands affect occurrence of acute GVHD after unrelated, 12/12 HLA matched, hematopoietic stem cell transplantation. **Bone Marrow Transplantation**, 2009.

LUDUEC, J. B. LE; BOUDREAU, J. E.; FREIBERG, J. C.; HSU, K. C. Novel approach to cell surface discrimination between KIR2DL1 subtypes and KIR2DS1 identifies hierarchies in NK repertoire, education, and tolerance. **Frontiers in Immunology**, 2019.

MARRERO, A. R.; SILVA-JUNIOR, W. A; BRAVI, C. M.; et al. Demographic and evolutionary trajectories of the Guarani and Kaingang natives of Brazil. **American journal of physical anthropology**, v. 132, n. 2, p. 301–310, 2007.

MARSH, S. G. E.; PARHAM, P.; DUPONT, B.; et al. Killer-cell immunoglobulin-like receptor (KIR) nomenclature report, 2002. **Tissue Antigens**, 2003.

MARTIN, A M.; FREITAS, E. M.; WITT, C. S.; CHRISTIANSEN, F. T. The genomic organization and evolution of the natural killer immunoglobulin-like receptor (KIR) gene cluster. **Immunogenetics**, 2000.

MARTIN, M. P.; BASHIROVA, A.; TRAHERNE, J.; TROWSDALE, J.; CARRINGTON, M. Cutting edge: expansion of the KIR locus by unequal crossing over. **Journal of immunology (Baltimore, Md. : 1950)**, v. 171, n. 5, p. 2192–2195, 2003.

MARTIN, M. P.; GAO, X.; LEE, J. H.; et al. Epistatic interaction between KIR3DS1 and HLA-B delays the progression to AIDS. **Nature Genetics**, 2002.

MARTIN, M. P.; QI, Y.; GAO, X.; et al. Innate partnership of HLA-B and KIR3DL1 subtypes against HIV-1. **Nature Genetics**, 2007.

MIDDLETON, D.; DILER, A. S.; MEENAGH, A.; SLEATOR, C.; GOURRAUD, P. A. Killer immunoglobulin-like receptors (KIR2DL2 and/or KIR2DS2) in presence of their ligand (HLA-C1 group) protect against chronic myeloid leukaemia. **Tissue Antigens**, 2009.

MOESTA, A. K.; GRAEF, T.; ABI-RACHED, L.; et al. Humans Differ from Other Hominids in Lacking an Activating NK Cell Receptor That Recognizes the C1 Epitope of MHC Class I. **The Journal of Immunology**, 2010.

MOESTA, A. K.; NORMAN, P. J.; YAWATA, M.; et al. Synergistic polymorphism at two positions distal to the ligand-binding site makes KIR2DL2 a stronger receptor for HLA-C than KIR2DL3. **Journal of immunology (Baltimore, Md. : 1950)**, v. 180, n. 6, p. 3969–79, 2008. Disponível em:

<http://www.ncbi.nlm.nih.gov/pubmed/18322206>...

MORETTA, A.; BOTTINO, C.; VITALE, M.; et al. RECEPTORS FOR HLA CLASS-I MOLECULES IN HUMAN NATURAL KILLER CELLS. **Annual Review of Immunology**, 1996.

MORVAN, M. G.; LANIER, L. L. NK cells and cancer: You can teach innate cells new tricks. **Nature Reviews Cancer**, 2016.

NAKIMULI, A.; CHAZARA, O.; HIBY, S. E.; et al. A KIR B centromeric region present in Africans but not Europeans protects pregnant women from pre-eclampsia . **Proceedings of the National Academy of Sciences**, v. 112, n. 3, 2015.

NEI, M. Analysis of gene diversity in subdivided populations. **Proceedings of the National Academy of Sciences of the United States of America**, 1973.

NELSON, G. W.; MARTIN, M. P.; GLADMAN, D.; et al. Cutting Edge: Heterozygote Advantage in Autoimmune Disease: Hierarchy of Protection/Susceptibility Conferred by HLA and Killer Ig-Like Receptor Combinations in Psoriatic Arthritis. **The Journal of Immunology**, 2004.

NEMAT-GORGANI, N.; EDINUR, H. A.; HOLLENBACH, J. A.; et al. KIR diversity in Māori and Polynesians: populations in which HLA-B is not a significant KIR ligand. **Immunogenetics**, v. 66, n. 11, 2014.

NEMAT-GORGANI, N.; GUETHLEIN, L. A.; HENN, B. M.; et al. Diversity of KIR, HLA Class I, and Their Interactions in Seven Populations of Sub-Saharan Africans. **The Journal of Immunology**, v. 202, n. 9, p. 2636–2647, 2019.

NEMAT-GORGANI, N.; HILTON, H. G.; HENN, B. M.; et al. Different Selected Mechanisms Attenuated the Inhibitory Interaction of KIR2DL1 with C2 + HLA-C in Two Indigenous Human Populations in Southern Africa . **The Journal of Immunology**, 2018.

NORMAN, P. J.; ABI-RACHED, L.; GENDZEKHADZE, K.; et al. Unusual selection on the KIR3DL1/S1 natural killer cell receptor in Africans. **Nature Genetics**, 2007.

NORMAN, P. J.; HOLLENBACH, J. A.; NEMAT-GORGANI, N.; 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, 2013a.

NORMAN, P. J.; HOLLENBACH, J. A.; NEMAT-GORGANI, N.; 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**, 2013b.

NORMAN, P. J.; HOLLENBACH, J. A.; NEMAT-GORGANI, N.; et al. Defining KIR and HLA Class I Genotypes at Highest Resolution via High-Throughput Sequencing. **American Journal of Human Genetics**, 2016.

NORMAN, P. J.; PARHAM, P. Complex interactions: the immunogenetics of human leukocyte antigen and killer cell immunoglobulin-like receptors. **Semin Hematol**, v. 42, n. 2, p. 65–75, 2005.

O'FALLON, B. D.; FEHREN-SCHMITZ, L. Native Americans experienced a strong population bottleneck coincident with European contact. **Proceedings of the National Academy of Sciences of the United States of America**, 2011.

O'LEARY, J. G.; GOODARZI, M.; DRAYTON, D. L.; ANDRIAN, U. H. VON. T celland B cell-independent adaptive immunity mediated by natural killer cells. **Nature Immunology**, 2006.

OLDER AGUILAR, A. M.; GUETHLEIN, L. A.; ADAMS, E. J.; et al. Coevolution of Killer Cell Ig-Like Receptors with HLA-C To Become the Major Variable Regulators of Human NK Cells. **The Journal of Immunology**, v. 185, n. 7, p. 4238–4251, 2010. Disponível em: http://www.jimmunol.org/cgi/doi/10.4049/jimmunol.1001494>.

OLDER AGUILAR, ANASTAZIA M.; GUETHLEIN, L. A.; ADAMS, E. J.; et al. Coevolution of Killer Cell Ig-Like Receptors with HLA-C To Become the Major Variable Regulators of Human NK Cells. **The Journal of Immunology**, 2010.

ORR, M. T.; LANIER, L. L. Natural Killer Cell Education and Tolerance. Cell, 2010.

PARADIS, E. Pegas: An R package for population genetics with an integratedmodular approach. **Bioinformatics**, 2010.

PARADIS, E.; SCHLIEP, K. Ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. **Bioinformatics**, 2019.

PARHAM, P. Killer cell immunoglobulin-like receptor diversity: Balancing signals in

the natural killer cell response. Immunology Letters. Anais... v. 92, p.11–13, 2004.

PARHAM, P. MHC class I molecules and KIRS in human history, health and survival. **Nature Reviews Immunology**, 2005.

PARHAM, P. The genetic and evolutionary balances in human NK cell receptor diversity. **Seminars in Immunology**, v. 20, n. 6, p. 311–316, 2008.

PARHAM, P.; ARNETT, K. L.; ADAMS, E. J.; et al. Episodic evolution and turnover of HLA-B in the indigenous human populations of the Americas. **Tissue Antigens**, 1997.

PARHAM, P.; GUETHLEIN, L. A. Genetics of Natural Killer Cells in Human Health, Disease, and Survival. **Annual Review of Immunology**, v. 36, n. January, 2018.

PAUST, S.; GILL, H. S.; WANG, B. Z.; et al. Critical role for the chemokine receptor CXCR6 in NK cell-mediated antigen-specific memory of haptens and viruses. **Nature Immunology**, 2010.

PEGRAM, H. J.; ANDREWS, D. M.; SMYTH, M. J.; DARCY, P. K.; KERSHAW, M. H. Activating and inhibitory receptors of natural killer cells. **Immunology and Cell Biology**, v. 89, n. 2, p. 216–224, 2011. Disponível em: http://www.nature.com/doifinder/10.1038/icb.2010.78>.

PENDE, D.; FALCO, M.; VITALE, M.; et al. Killer Ig-like receptors (KIRs): Their role in NK cell modulation and developments leading to their clinical exploitation. **Frontiers in Immunology**, v. 10, p. 1179, 2019.

PENMAN, B. S.; MOFFETT, A.; CHAZARA, O.; GUPTA, S.; PARHAM, P. Reproduction, infection and killer-cell immunoglobulin-like receptor haplotype evolution. **Immunogenetics**, 2016.

PETZL-ERLER, M. L.; LUZ, R.; SOTOMAIOR, V. S. The HLA polymorphsm of two distinctive South-American Indian tribes: The Kaingang and the Guarani. **Tissue Antigens**, 1993.

PORTALES-CASAMAR, E.; KIROV, S.; LIM, J.; et al. PAZAR: A framework for collection and dissemination of cis-regulatory sequence annotation. **Genome Biology**, 2007.

POSTH, C.; NAKATSUKA, N.; LAZARIDIS, I.; et al. Reconstructing the Deep Population History of Central and South America. **Cell**, 2018.

PYO, C. W.; GUETHLEIN, L. A.; VU, Q.; 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**, 2010.

R CORE TEAM. R: A language and environment for statistical computing. , 2019. Vienna, Austria.: R Foundation for Statistical Computing. Disponível em: <https://www.r-project.org/.>. .

RAJAGOPALAN, S.; LONG, E. O. Understanding how combinations of HLA and KIR

genes influence disease. **The Journal of Experimental Medicine**, v. 201, n. 7, p. 1025–1029, 2005. Disponível em: http://www.jem.org/lookup/doi/10.1084/jem.20050499>.

RAJALINGAM, R.; KRAUSA, P.; SHILLING, H. G.; et al. Distinctive KIR and HLA diversity in a panel of north Indian Hindus. **Immunogenetics**, v. 53, n. 12, p. 1009–1019, 2002.

REICH, D.; PATTERSON, N.; CAMPBELL, D.; et al. Reconstructing Native American population history. **Nature**, v. 488, n. 7411, p. 370–374, 2012. Disponível em: http://www.nature.com/doifinder/10.1038/nature11258>.

REYES-CENTENO, H. Out of Africa and into Asia: Fossil and genetic evidence on modern human origins and dispersals. **Quaternary International**, 2016.

ROBERTSON, M. J.; RITZ, J. Biology and Clinical Relevance of Human Natural Killer Cells. **The Journal of The American Society of Hematology**, v. 76, n. 12, p. 2421–2438, 1990.

ROBINSON, J.; BARKER, D. J.; GEORGIOU, X.; et al. IPD-IMGT/HLA Database. **Nucleic acids research**, 2020.

ROBINSON, J.; HALLIWELL, J. A.; HAYHURST, J. D.; et al. The IPD and IMGT/HLA database: Allele variant databases. **Nucleic Acids Research**, 2015.

ROBINSON, J.; HALLIWELL, J. A.; MCWILLIAM, H.; LOPEZ, R.; MARSH, S. G. E. IPD - The Immuno Polymorphism Database. **Nucleic Acids Research**, v. 41, n. D1, 2013.

ROSENBERG, N. A.; PRITCHARD, J. K.; WEBER, J. L.; et al. Genetic structure of human populations. **Science**, 2002.

RSTUDIO TEAM. RStudio: Integrated Development for R., 2018. Boston, MA: RStudio, Inc. Disponível em: http://www.rstudio.com/.

SABETI, P. C.; REICH, D. E.; HIGGINS, J. M.; et al. Detecting recent positive selection in the human genome from haplotype structure. **Nature**, 2002.

SAITOU, N.; NEI, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. **Molecular Biology and Evolution**, v. 4, n. 4, p. 406–425, 1987.

SAKURAI, C.; ASARI, A. Y.; BELTRÃO, K. I.; et al. **Resistência & integração 100** anos de imigração japonesa no Brasil. Rio de Janiero: IBGE, 2010.

SAMBROOK, J.; FRITSCH, E. F.; MANIATIS, T. Molecular cloning: a laboratory manual. **Cold Spring Harbor Laboratory Press**, 1989.

SAMBROOK, J. G.; BASHIROVA, A.; PALMER, S.; et al. Single haplotype analysis demonstrates rapid evolution of the killer immunoglobulin-like receptor (KIR) loci in primates. **Genome Research**, 2005.

SANGER, F.; NICKLEN, S.; COULSON, A. R. DNA sequencing with chain-

terminating inhibitors. Proceedings of the National Academy of Sciences, 1977.

SCHANFIELD, M. S. Immunoglobulin allotypes (GM and KM) indicate multiple founding populations of Native Americans: evidence of at least four migrations to the New World. **Human biology**, v. 64, n. 3, p. 381–402, 1992.

SCHEET, P.; STEPHENS, M. A fast and flexible statistical model for large-scale population genotype data: Applications to inferring missing genotypes and haplotypic phase. **American Journal of Human Genetics**, 2006.

SCHMITT, R.; BONATTO, S. L.; FREITAS, L. B.; et al. Extremely limited mitochondrial DNA variability among the Aché Natives of Paraguay. **Annals of Human Biology**, 2004.

SCHWARTZ, O.; MARÉCHAL, V.; GALL, S. LE; LEMONNIER, F.; HEARD, J. M. Endocytosis of major histocompatibility complex class I molecules is induced by the HIV-1 Nef protein. **Nature Medicine**, 1996.

SHILLING, H. G.; GUETHLEIN, L. A.; CHENG, N. W.; et al. Allelic Polymorphism Synergizes with Variable Gene Content to Individualize Human KIR Genotype. **The Journal of Immunology**, 2002.

SHIN, J.-H.; BLAY, S.; GRAHAM, J.; MCNENEY, B. LDheatmap : An R Function for Graphical Display of Pairwise Linkage Disequilibria Between Single Nucleotide Polymorphisms . **Journal of Statistical Software**, 2006.

SIDNEY, J.; PETERS, B.; FRAHM, N.; BRANDER, C.; SETTE, A. HLA class I supertypes: A revised and updated classification. **BMC Immunology**, 2008.

SINGLE, R. M.; MARTIN, M. P.; GAO, X.; et al. Global diversity and evidence for coevolution of KIR and HLA. **Nature Genetics**, 2007.

SKOGLUND, P.; MALLICK, S.; BORTOLINI, M. C.; et al. Genetic evidence for two founding populations of the Americas. **Nature**, 2015. Disponível em: http://www.nature.com/doifinder/10.1038/nature14895>.

SLIK, A. R. VAN DER; KOELEMAN, B. P. C.; VERDUIJN, W.; et al. KIR in type 1 diabetes: Disparate distribution of activating and inhibitory natural killer cell receptors in patients versus HLA-matched control subjects. **Diabetes**, 2003.

SMYTH, M. J.; CRETNEY, E.; KELLY, J. M.; et al. Activation of NK cell cytotoxicity. **Molecular Immunology**, 2005.

SOLLOCH, U. V; SCHEFZYK, D.; SCHÄFER, G.; MASSALSKI, C.; KOHLER, M. Estimation of German KIR Allele Group Haplotype Frequencies. **Frontiers in Immunology**, v. 11, 2020.

STEWART, C. A.; LAUGIER-ANFOSSI, F.; VELY, F.; et al. Recognition of peptide-MHC class I complexes by activating killer immunoglobulin-like receptors. **Proceedings of the National Academy of Sciences**, 2005.

STORKUS, W.; HOWELL, D.; SALTER, R.; DAWSON, J.; CRESSWELL, P. NK

Susceptibility Varies Inversely With Target Cell Class I HLA Antigen Expression. **Journal of Immunology**, v. 138, n. 6, p. 1657–9, 1987.

SUN, J. C.; BEILKE, J. N.; LANIER, L. L. Adaptive immune features of natural killer cells. **Nature**, v. 457, n. 7229, p. 557–561, 2009. Disponível em: http://www.nature.com/doifinder/10.1038/nature07665>.

TARAZONA-SANTOS, E.; CARVALHO-SILVA, D. R.; PETTENER, D.; et al. Genetic differentiation in South Amerindians is related to environmental and cultural diversity: Evidence from the Y chromosome. **American Journal of Human Genetics**, 2001.

THE 1000 GENOMES PROJECT CONSORTIUM. A global reference for human genetic variation. **Nature**, 2015.

TISHKOFF, S. A.; VERRELLI, B. C. Patterns of human genetic diversity: Implications for Human Evolutionary History and Disease. **Annual Review of Genomics and Human Genetics**, 2003.

TONEVA, M.; LEPAGE, V.; LAFAY, G.; et al. Genomic diversity of natural killer cell receptor genes in three populations. **Tissue antigens**, v. 57, n. 4, p. 358–62, 2001. Disponível em: http://www.ncbi.nlm.nih.gov/pubmed/11380947>.

TRAHERNE, J. A.; MARTIN, M.; WARD, R.; et al. Mechanisms of copy number variation and hybrid gene formation in the KIR immune gene complex. **Human Molecular Genetics**, 2010.

TRIMARCHI, J. M.; LEES, J. A. Sibling rivalry in the E2F family. **Nature Reviews Molecular Cell Biology**, 2002.

TROWSDALE, J.; MOFFETT, A. NK receptor interactions with MHC class I molecules in pregnancy. **Seminars in Immunology**, 2008.

TSUNETO, L. T.; PROBST, C. M.; HUTZ, M. H.; et al. HLA class II diversity in seven Amerindian populations. Clues about the origins of the Ach?? **Tissue Antigens**, v. 62, n. 6, p. 512–526, 2003.

UHRBERG, M. Shaping the human NK cell repertoire: An epigenetic glance at KIR gene regulation. **Molecular Immunology**, 2005.

UHRBERG, M.; PARHAM, P.; WERNET, P. Definition of gene content for nine common group B haplotypes of the Caucasoid population: KIR haplotypes contain between seven and eleven KIR genes. **Immunogenetics**, v. 54, n. 4, p. 221–229, 2002.

UHRBERG, M.; VALIANTE, N. M.; SHUM, B. P.; et al. Human diversity in killer cell inhibitory receptor genes. **Immunity**, 1997.

UHRBERG, M.; VALIANTE, N. M.; YOUNG, N. T.; et al. The Repertoire of Killer Cell Ig-Like Receptor and CD94:NKG2A Receptors in T Cells: Clones Sharing Identical TCR Rearrangement Express Highly Diverse Killer Cell Ig-Like Receptor Patterns. **The Journal of Immunology**, v. 166, n. 6, p. 3923–3932, 2001. Disponível em: http://www.jimmunol.org/cgi/doi/10.4049/jimmunol.166.6.3923>. ULLAH, M. A.; HILL, G. R.; TEY, S. K. Functional reconstitution of natural killer cells in allogeneic hematopoietic stem cell transplantation. **Frontiers in Immunology**, 2016.

VALIANTE, N. M.; UHRBERG, M.; SHILLING, H. G.; et al. Functionally and structurally distinct NK cell receptor repertoires in the peripheral blood of two human donors. **Immunity**, v. 7, n. 6, p. 739–751, 1997.

VELARDI, A.; RUGGERI, L.; ALESSANDRO; MORETTA; MORETTA, L. NK cells: A lesson from mismatched hematopoietic transplantation. **Trends in Immunology**, 2002.

VILCHES, C.; CASTAÑO, J.; GÓMEZ-LOZANO, N.; ESTEFANÍA, E. Facilitation of KIR genotyping by a PCR-SSP method that amplifies short DNA fragments. **Tissue Antigens**, 2007.

VILCHES, C.; PANDO, M. J.; PARHAM, P. Genes encoding human killer-cell Ig-like receptors with D1 and D2 extracellular domains all contain untranslated pseudoexons encoding a third Ig-like domain. **Immunogenetics**, v. 51, n. 8–9, p. 639–646, 2000.

VILCHES, C.; PARHAM, P. KIR: Diverse, Rapidly Evolving Receptors of Innate and Adaptive Immunity. **Annual Review of Immunology**, v. 20, n. 1, p. 217–251, 2002. Disponível em:

http://www.annualreviews.org/doi/abs/10.1146/annurev.immunol.20.092501.134942

VIVIER, E.; UGOLINI, S.; BLAISE, D.; CHABANNON, C.; BROSSAY, L. Targeting natural killer cells and natural killer T cells in cancer. **Nature Reviews Immunology**, 2012.

WAGGONER, S. N.; REIGHARD, S. D.; GYUROVA, I. E.; et al. Roles of natural killer cells in antiviral immunity. **Current Opinion in Virology**, 2016.

WANG, S.; LEWIS, C. M.; JAKOBSSON, M.; et al. Genetic variation and population structure in Native Americans. **PLoS Genetics**, 2007.

WARNES, G.; GORJANC, G.; LEISCH, F.; MAN., M. genetics: Population Genetics., 2019.

WEIR, B. S.; COCKERHAM, C. C. Estimating F-Statistics for the Analysis of Population Structure. **Evolution**, 1984.

WENDE, H.; COLONNA, M.; ZIEGLER, A.; VOLZ, A. Organization of the leukocyte receptor cluster (LRC) on human Chromosome 19q13.4. **Mammalian Genome**, v. 10, n. 2, p. 154–160, 1999.

WILSON, M J; TORKAR, M.; HAUDE, A.; et al. Plasticity in the organization and sequences of human KIR/ILT gene families. **Proceedings of the National Academy of Sciences of the United States of America**, v. 97, n. 9, p. 4778–83, 2000. Disponível em:

<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=18309&tool=pmcentrez&r

endertype=abstract>..

WILSON, M. J.; TORKAR, M.; HAUDE, A.; et al. Plasticity in the organization and sequences of human KIR/ILT gene families. **Proceedings of the National Academy of Sciences**, 2000.

WIMLEY, W. C.; WHITE, S. H. Experimentally determined hydrophobicity scale for proteins at membrane interfaces. **Nature Structural Biology**, 1996.

WINTER, C. C.; GUMPERZ, J. E.; PARHAM, P.; LONG, E. O.; WAGTMANN, N. Direct binding and functional transfer of NK cell inhibitory receptors reveal novel patterns of HLA-C allotype recognition. **Journal of immunology (Baltimore, Md. : 1950)**, v. 161, n. 2, p. 571–7, 1998. Disponível em: http://www.jimmunol.org/content/161/2/571.full.

WINTER, C. C.; LONG, E. O. A single amino acid in the p58 killer cell inhibitory receptor controls the ability of natural killer cells to discriminate between the two groups of HLA-C allotypes. **Journal of immunology (Baltimore, Md.: 1950)**, 1997.

WRIGHT, S. Evolution and the Genetics of Populations: Variability within and among natural populations. Chicago, 1978.

XIONG, S.; SHARKEY, A. M.; KENNEDY, P. R.; et al. Maternal uterine NK cellactivating receptor KIR2DS1 enhances placentation. **Journal of Clinical Investigation**, v. 123, n. 10, p. 4264–4272, 2013.

XUE, T. Y.; HE, Y. Z.; ZHANG, J. J.; et al. KIR2DL1-HLA signaling pathway: Notable inhibition in the cytotoxicity of allo-NK cell against KG1A cell. **Clinical Laboratory**, v. 59, n. 5–6, p. 613–619, 2013.

YAWATA, M.; YAWATA, N.; DRAGHI, M.; et al. Roles for HLA and KIR polymorphisms in natural killer cell repertoire selection and modulation of effector function. **The Journal of Experimental Medicine**, v. 203, n. 3, p. 633–645, 2006. Disponível em: http://www.jem.org/lookup/doi/10.1084/jem.20051884>.

YEN, J.; LIN, C.; TSAI, W.; et al. Killer cell immunoglobulin-like receptor gene s repertoire in rheumatoid arthritis. **Scandinavian Journal of Rheumatology**, v. 35, n. 2, p. 124–127, 2006.

ZHANG, C.; MAEDA, N.; IZUMIYA, C.; et al. Killer immunoglobulin-like receptor and human leukocyte antigen expression as immunodiagnostic parameters for pelvic endometriosis. **American Journal of Reproductive Immunology**, 2006.

ZHANG, L.; LENG, Q.; MIXSON, A. J. Alteration in the IL-2 signal peptide affects secretion of proteins in vitro and in vivo. **Journal of Gene Medicine**, 2005.