Remarkably low KIR and HLA diversity in Amerindians reveals signatures of strong purifying selection shaping the centromeric KIR region

Luciana de Brito Vargas¹, Marcia H. Beltrame¹, Brenda Ho², Wesley M. Marin², Ravi Dandekar², Gonzalo Montero-Martín³, Marcelo A. Fernández-Viña³, A. Magdalena Hurtado⁴, Kim R. Hill⁴, Luiza T. Tsuneto⁵, Mara H. Hutz⁶, Francisco M. Salzano⁶,†, Maria Luiza Petzl-Erler¹, Jill A. Hollenbach²,⁷, Danillo G. Augusto¹,²*

¹Programa de Pós-Graduação em Genética, Departamento de Genética, Universidade Federal do Paraná, Curitiba, PR 81531-980, Brazil; ²Weill Institute for Neurosciences, Department of Neurology, University of California, San Francisco, San Francisco, CA 94158, USA; ³Department of Pathology, Stanford University School of Medicine, Palo Alto, CA 94304, USA; ⁴School of Human Evolution and Social Change, Arizona State University, Tempe, AZ 85287, USA; ⁵Departamento de Análises Clínicas, Universidade Estadual de Maringá, Maringá, PR 87020-900, Brazil; ⁶Departamento de Genética, Instituto de Biociências, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS 91501-970, Brazil; ⁷Department of Epidemiology and Biostatistics, University of California, San Francisco, CA 94158, USA.

† Deceased September 28, 2018.

*Corresponding author:
Danillo G. Augusto, PhD
University of California, San Francisco
675 Nelson Rising Lane, Room 240
San Francisco CA 94158, USA

E-mail: danillo@augusto.bio.br

Keywords: killer-cell immunoglobulin-like receptor, high-resolution, population, evolution, human leukocyte antigen.
Abstract

The killer-cell immunoglobulin-like receptors (KIR) recognize human leukocyte antigen (HLA) molecules to regulate the cytotoxic and inflammatory responses of natural killer cells. KIR genes are encoded by a rapidly evolving gene family on chromosome 19 and present an unusual variation of presence and absence of genes and high allelic diversity. Although many studies have associated KIR polymorphism with susceptibility to several diseases over the last decades, the high-resolution allele-level haplotypes have only recently started to be described in populations. Here, we use a highly innovative custom next-generation sequencing method that provides a state-of-art characterization of KIR and HLA diversity in 706 individuals from eight unique South American populations: five Amerindian populations from Brazil (three Guarani and two Kaingang); one Amerindian population from Paraguay (Aché); and two urban populations from Southern Brazil (European and Japanese descendants from Curitiba). For the first time, we describe complete high-resolution KIR haplotypes in South American populations, exploring copy number, linkage disequilibrium, and KIR-HLA interactions. We show that all Amerindians analyzed to date exhibit the lowest numbers of KIR-HLA interactions among all described worldwide populations, and that 83-97% of their KIR-HLA interactions rely on a few HLA-C molecules. Using multiple approaches, we found signatures of strong purifying selection on the KIR centromeric region, which codes for the strongest NK cell educator receptors, possibly driven by the limited HLA diversity in these populations. Our study expands the current knowledge of KIR genetic diversity in populations to understand KIR-HLA coevolution and its impact on human health and survival.

Introduction

Natural killer (NK) cells are cytotoxic lymphocytes that were first discovered due to their ability to spontaneously kill tumor cells in vitro without prior sensitization (Herberman and Holden 1978) and later recognized as critical components of the first line of defense against tumor and infected cells (Morvan and Lanier 2016; Flórez-Álvarez et al. 2018). Among the receptors that control NK cell cytotoxicity is the killer-cell immunoglobulin-like
receptor (KIR) family, which recognizes human leukocyte antigen (HLA) molecules as primary ligands (Lanier and Phillips 1995; Moretta et al. 1996).

KIR molecules are encoded by a highly polymorphic gene family located on the chromosome region 19q13.4, characterized by an uncommon and complex structural variation of presence and absence of genes (Wende et al. 1999; Wilson et al. 2000). The homology and high sequence similarity among the 13 KIR loci contribute to the occurrence of non-reciprocal recombination (Wilson et al. 2000; Martin et al. 2003), which generates duplication and deletion of entire genes or groups of genes and the formation of hybrid genes and alleles (Martin et al. 2003; Norman et al. 2009; Traherne et al. 2010; Roe et al. 2017). The KIR region is known for its complexity and rapid evolution (Khakoo et al. 2000; Sambrook et al. 2005; Guethlein et al. 2007).

The presence and absence of genes generate a wide diversity of KIR gene-content haplotypes, classified into two groups: A and B (Uhrberg et al. 2002). Both KIR A and B haplotypes are present in all studied human populations; however, their frequencies vary significantly. For example, haplotype A is present in approximately 80% of Japanese individuals (Yawata et al. 2006), but only 2% of Australian aborigines (Toneva et al. 2001). Despite a large number of known haplotypes among human populations (Hsu et al. 2002; Uhrberg et al. 2002; Martin et al. 2008; Pyo et al. 2013; Roe et al. 2017), most haplotypes are the result of different combinations of a smaller set of centromeric (flanked by KIR3DL3 and KIR3DP1) and telomeric segments (flanked by KIR2DL4 and KIR3DL2) (Pyo et al. 2010; Hollenbach et al. 2012). This feature is most likely driven by a hotspot facilitating the recombination of telomeric and centromeric segments (Traherne et al. 2010).

In contrast with KIR genes, which emerged 23-1.7 million years ago (mya) (Wilson et al. 2000; Pyo et al. 2010; Parham and Guethlein 2018), the HLA constitutes an evolutionary older gene family that arose 49-22 mya (Piontkivska and Nei 2003; Fukami-Kobayashi et al. 2005). HLA genes are located within the major histocompatibility complex (MHC) on chromosome 6 (Horton et al. 2004) and are the most polymorphic genes in the human genome (Hill 1999). The interaction of KIR and HLA is critical for NK cell education during the early stages of maturation (Kärre et al. 1986; Kim et al. 2005), for regulating NK cell cytotoxicity (Smyth et al. 2005; Lanier 2008), and reproduction (Hiby et al. 2004;
Xiong et al. 2013; Blokhuis et al. 2017). In addition, combinations of KIR-HLA have been associated with numerous diseases (Martin et al. 2002; Van der Slik et al. 2003; Khakoo et al. 2004; Nelson et al. 2004; Carrington et al. 2005; Augusto et al. 2012; Augusto 2016; Anderson et al. 2020), and there is growing evidence that these two families coevolve as a unique system (Gendzekhadze et al. 2009; Norman et al. 2013a; Augusto and Petzl-Erler 2015; Moffett and Colucci 2015; Vargas et al. 2020).

Despite the relevance of KIR for disease and survival, the complexity of the structural variation of haplotypes and the high sequence similarity among genes impose technical difficulties on their study. Although over a thousand KIR alleles have been deposited in the IPD (ImmunoPolymorphism Database)-KIR (Robinson et al. 2015), the distribution of these alleles in global populations is poorly known. To date, the vast majority of population genetics studies have only analyzed KIR at the gene-content level and sometimes in combination with a few HLA ligands (González-Galarza et al. 2015). The study of allelic diversity at high resolution for all KIR genes is still restricted to a few populations (Norman et al. 2016; Nemat-Gorgani et al. 2018; Nemat-Gorgani et al. 2019; Alicata et al. 2020; Solloch et al. 2020; Tao et al. 2020; Amorim et al. 2021; Deng et al. 2021).

Our previous work and that of others have shown a limited number of KIR gene-content haplotypes in Amerindians (Gutiérrez-Rodríguez et al. 2006; Flores et al. 2007; Single et al. 2007; Augusto et al. 2015). However, KIR allelic diversity has only been described for a single Amerindian population, the Yucpa from Venezuela, with high-resolution genotyping restricted to a few loci (Gendzekhadze et al. 2009). Evidence has shown that Eastern Amerindians, such as those living in Brazil, bear even lower genetic diversity than Western Amerindians (Wang et al. 2007). Additionally, previous remarkable studies identified several high-frequency HLA alleles in Brazilian Amerindians never found in Amerindians from North America or any other global population (Belich et al. 1992; Watkins et al. 1992; Parham et al. 1997).

Here, we present the first high-resolution characterization of KIR allelic variation in seven South American populations from Brazil and one from Paraguay. We analyzed six isolated Amerindian populations from Guarani, Kaingang, and Aché ethnicities, and also individuals of European and Japanese ancestries living in Southern Brazil. We deliver
the first study to analyze the in-depth KIR diversity at high-resolution in South American populations, identifying signatures of purifying selection on the centromeric KIR region, possibly driven by the reduced diversity of HLA ligands.

**Results**

**Amerindians from our study exhibit the lowest KIR diversity among all worldwide populations analyzed to date**

We define as a KIR allele each unique DNA sequence at a particular locus. To directly compare the allelic diversity of the study populations, we randomly selected 50 individuals from each study population. We observed a remarkably low diversity of KIR alleles in the Amerindian study populations, averaging 53±9 alleles at 5-digit resolution per population. The number of KIR alleles was especially small in Aché (ACHE; 41 alleles per n=50), Kaingang from Rio das Cobras (KRC, 46 alleles per n=50), and Guarani Mbya (GRC; 48 alleles per n=50). In sharp contrast, we observed 123 and 89 alleles per n=50 in the two Brazilian urban populations, European and Japanese descendants, respectively. Individually, KIR3DL3 and KIR3DL2 were the two genes with the highest number of alleles (Figure 1A). The complete list of allelic frequencies is described in Supplementary Table 1.

Multiple alleles observed in our Amerindian samples were not found in the two urban populations. Six of these alleles observed exclusively in Amerindians were found only in Kaingang from Ivaí (KIV), and six were observed only in Guarani Kaiowá (GKW) (Figure 1B; Supplementary Table 2). Overall, the six Amerindian populations shared an average of 76.64% of their KIR alleles. A high proportion of alleles is shared between Amerindians and Japanese descendants (66.93%), and there is reduced proximity between Euro-descendants and Amerindians, only 60.11% (p < 0.001; Supplementary Table 3).

According to the KIR official nomenclature, 3-digit resolution means the identification of the DNA sequence of each allele that allows the distinction only of the substitutions that change the protein sequence (Robinson et al. 2015). We define as an allotype every unique protein encoded by a KIR allele. To maximize the comparison of KIR diversity and include multiple global populations (Jones et al. 2006; Gendzekhadze et al. 2009;
Vierra-Green et al. 2012; Norman et al. 2013b; Nemat-Gorgani et al. 2014; Nemat-Gorgani et al. 2018; Nemat-Gorgani et al. 2019; Alicata et al. 2020; Tao et al. 2020; Amorim et al. 2021; Deng et al. 2021), we compared the study populations with others using KIR alleles at the 3-digit resolution, which define the allotypes, considering only those with frequencies greater than or equal to 1%. With only 24, KRC has the lowest number of common (f ≥ 1%) KIR 3-digit alleles ever reported. The average number of common KIR alleles at 3-digit resolution in Amerindians was 32±5, while the other worldwide populations had an average of 73±13 (Figure 2). We also estimated the allele richness for each population (Figure 2; Supplementary Table 4). We show a reduced KIR allele richness in the Amerindians compared to worldwide populations, also showing that Aché exhibited the lowest allelic richness to date. In addition, we found a strong correlation (r = 0.97; p = 1.5x10^{-18}) between number of common alleles and allele richness (Supplementary Figure 1A) and no correlation between sample sizes and number of common alleles (Supplementary Figure 1B).

The study populations exhibit an overall great population differentiation (Supplementary Figure 2). The genes responsible for most of the differentiation were KIR2DL1 (median $F_{ST} = 0.11$), KIR2DL23 (median $F_{ST} = 0.10$), and KIR3DL1S1 (median $F_{ST} = 0.09$; Supplementary Figure 2C). We generated bidimensional plots using the principal component analysis (PCA) to visualize the overall difference in allele frequencies across all KIR loci in populations. Populations formed three clusters (Figure 3): in red, all Africans; in yellow, East Asians (including BrJAP) and Oceanic populations; in blue, all Middle Eastern and Euro-descendant populations. The Amerindians did not form a cluster despite being separated from all other populations. However, when the KIR centromeric and telomeric were plotted separately, we observed greater proximity for Amerindian groups in telomeric than centromeric KIR genes (Supplementary Figure 3).

**Remarkably high frequencies of cA01 haplotypes in Amerindians: carrier frequency reaches 100% in Aché.**

For the first time, we report the frequencies of centromeric and telomeric haplotypes for Amerindians at both gene-content and high-resolution allelic levels (Figure 4A-B; Tables S5 and S6). We also provide pairwise linkage disequilibrium (LD) between high-resolution alleles in Supplementary Figures 4-11 and Supplementary Table 7).
frequencies of CenA haplotypes were remarkably high in the Aché and two Guarani groups (Figure 4D). Specifically, cA01 was present in 100% of ACHE individuals and exhibited frequencies ranging from 81.1 to 88.7% in GND (Guarani Ñandeva), GKW, and BrJAP (Brazilian Japanese) (Supplementary Table 8). High frequencies of the Cen A haplotype were also observed in Asians, ranging from 76.7 to 93.2% (Yawata et al. 2006; Tao et al. 2020; Deng et al. 2021). In contrast, Cen A frequencies in other worldwide populations range from 29.0 to 69.4% in Africans (Nemat-Gorgani et al. 2018; Nemat-Gorgani et al. 2019) and 65 to 69.0% in Iranians and European descendants (Vierra-Green et al. 2012; Alicata et al. 2020). Interestingly, cB02 was highly frequent in GRC, carried by 39.3% of the population. As expected, greater diversity of haplotypes was observed in populations of European and Japanese ancestries, with 30 and 15 centromeric and 25 and 22 telomeric haplotypes at the gene-content level, respectively.

The most common gene content haplotype in all populations is cA01~tA01, followed by cA01~tB01 or cB02~tB01 (Supplementary Table 8). We identified haplotypes carrying large structural deletions or duplications involving multiple loci. The haplotype cB02~tB01-del6, with the deletion of KIR3DP1~ KIR2DL4~ KIR3DS1, was found in six populations, with frequencies greater than 2% in Japanese descendants and 1% in Kaingang Ivaí (KIV). On the other hand, the genes deleted in del6 are duplicated in the expanded haplotype cA01~tA01-ins4.

Considering the common KIR haplotypes at high resolution found in at least three individuals, we observed 46 telomeric and 37 centromeric haplotypes at the allelic level. Among the less frequent allele-level haplotypes, we observed 119 centromeric 141 telomeric haplotypes found in less than three individuals, and only 0.67% of haplotypes were not resolved. The low-frequency haplotypes are usually variations of the most common ones, differing by only a few alleles in some specific genes. Allelic variation of KIR3DL3 was responsible for around 30% of the less common centromeric haplotypes. This fact is likely the result of a recombination hotspot in KIR3DL3 exon 5, which increases sequence variation in the gene and leads to lower LD with the remainder of the KIR centromeric region (Jones et al. 2006; Abi-Rached et al. 2010). On the telomeric region, the allelic variation on KIR3DL1S1 differentiated most of the less frequent
haplotypes. Overall, we found greater diversity on telomeric when compared to centromeric haplotypes for all Amerindian populations.

**HLA-C ligands almost exclusively mediate the KIR regulation in all Amerindians studied to date**

We provide a detailed analysis of the combinatorial diversity of KIR-HLA interactions in the study populations (Table 1). Our study is the first to report the frequencies of HLA ligands in Aché. Like the other Amerindians in this study and the literature (Johnson et al. 1978; Petzl-Erler et al. 1993; Fernández-Viña et al. 1997; Parham et al. 1997; Layrisse et al. 2001; García-Ortiz et al. 2006), Aché also exhibited the complete lack of HLA-A3 and A11 ligands (Supplementary Table 9). Interestingly, the Aché also exhibited a low frequency of Bw4 ($f=0.05$), resulting in most of their KIR-HLA interactions being dependent on HLA-C.

On average, Amerindians in this study exhibited only ~3 to 5 KIR-HLA interactions per individual (Table 1). In contrast, the Japanese had an average of 5.8 KIR-HLA interactions. In the European-Brazilians, we found an average of 6.9. Among all interactions, HLA-C was responsible for 51.3% of the observed KIR-HLA interactions in Japanese and 57.8% in Euro-Brazilians. However, HLA-C has a remarkably high contribution to KIR-HLA interactions in Amerindians, ranging from 82.6% to 97.1% (Figure 5; Table 1; and Supplementary Table 10). Moreover, more than 60% of the KIR-HLA interactions observed in all study populations are inhibitory, with KIR2DL23 accounting for most of the considered interactions, ranging from 17.8% in Euro-Brazilians to 41.4% in Guarani Kaiowá.

**Reduced HLA diversity imposes purifying selection on the centromeric KIR region**

Based on the observation of a small number of HLA ligands in the Amerindians, we hypothesized that reduced HLA diversity could be shaping the KIR diversity in these populations. To test this hypothesis, we first calculated the difference of synonymous to nonsynonymous substitution ($d_N - d_S$) for all KIR loci (Figure 6A). We observed a significant reduced nonsynonymous rate in the centromeric region (Cen $d_N - d_S = -1.2$) in comparison to the telomeric region (Tel $d_N - d_S = 0.29$) when we analyzed all study populations ($p = 6\times10^{-5}$). Individually, the deviation of neutrality in the KIR centromeric
region was significant in the Aché (d_N - d_S = -2.28, p = 0.02), the population with the most
limited diversity of HLA allotypes and with 97.1% of the KIR-HLA interactions mediated
by HLA-C alone. The negative values in the centromeric KIR indicate an excess of
synonymous substitutions in lieu of nonsynonymous substitutions, which could indicate
purifying selection.

Next, we applied the Ewens-Watterson test of neutrality. While the d_N - d_S test informs
selection on a deeper timescale (Bamshad and Wooding 2003; Nielsen 2005; Mugal et
al. 2014), the Ewens-Watterson test is especially suited for detecting recent selection
events by comparing the observed homozygosity (F_{obs}) with the expected homozygosity
(F_{exp}) (Nielsen 2001). Our Ewens-Watterson test results corroborate the possibility of
selection limiting the diversity in centromeric KIR genes in Amerindians. The
homozygosity was pronounced (F_{obs} > F_{exp}) in several centromeric KIR genes in
Amerindians, and a strikingly different pattern (p = 4x10^{-6}) was observed in the
neighboring telomeric region (Figure 6B). In the Aché, significantly increased
homozygosity was observed in five of the six centromeric genes (p > 0.975). The Ewens-Watterson patterns corroborate the d_N - d_S deviations and suggest selective pressures
on the centromeric KIR portion but not on the telomeric.

To further our analysis and eliminate the confounder effects of stochastic and
demographic factors, we retrieved data from 678 genomic microsatellites (Msat) markers
previously described in the Aché, Guarani, and Kaingang (Wang et al. 2007). We
analyzed genome-wide diversity and population differentiation and compared them to
those found in KIR. We observed that the overall differentiation in the Msat markers
(mean FST = 0.0873) among these populations is higher than in KIR (Figure 6C). These
significant differences (1.6 x 10^{-4} < p < 1.1 x 10^{-7}) would not be expected if the KIR
region was under neutral evolution. Next, we analyzed the heterozygosity rates across
the centromeric and telomeric KIR regions, comparing them to the heterozygosity of the
genomic Msat. We found that the KIR centromeric heterozygosity is much lower than the
genomic Msat heterozygosity in Amerindians (p = 7x10^{-4}), while the telomeric KIR region
exhibits heterozygosity rates similar to the neutral genomic markers (Fig. 6D). These
results show a significantly different heterozygosity pattern between the centromeric and
telomeric regions (p = 4.7x10^{-4}). Altogether, our observations indicate deviations from
neutrality that cannot be fully explained by stochastic factors, pointing to a strong stabilizing selection specifically on the centromeric KIR region.

**Discussion**

The uniqueness of the South American Amerindians is due to their singular demographic history of migration from Asian ancestral populations via the Behring Strait (Skoglund et al. 2015), followed by a complex dispersion along the American continent (Reich et al. 2012; Castro e Silva et al. 2020), as well as remaining genetically isolated during the last five centuries (Petzl-Erler et al. 1993; Tsuneto et al. 2003). The Amerindian demographic history is especially interesting because of the intense bottleneck and founder effects (Amos and Hoffman 2010; O’Fallon and Fehren-Schmitz 2011) and other stochastic effects intensified by their mostly small population sizes (Conroy et al. 2000; Tarazona-Santos et al. 2001). Together, these events contributed to the low genomic diversity currently observed in Amerindian populations (DeGiorgio et al. 2009). Moreover, the unique genetic diversity of Amerindians may partially have emerged after the ancestral migration. A remarkable example is the episodic evolution in Amerindians from South America that replaced a large part of the ancestral HLA alleles with novel sets of population-specific HLA alleles never observed elsewhere (Belich et al. 1992; Watkins et al. 1992; Parham et al. 1997). Although describing such unique isolated populations is highly relevant, the genetic characterization of underrepresented admixed populations in South America is equally essential for anthropological genetics. Studies such as the one we present here lay the foundation for understanding the normal and pathologic human genetic variation and may contribute to creating personalized medicine solutions applicable to populations typically neglected in large genetic studies.

We found a limited number of KIR alleles in the Amerindians compared to the two urban populations. Aiming to compare the KIR diversity in Amerindians with global populations and due to the scarcity of high-resolution data available, we compared the study population with worldwide populations at 3-digit resolution with allelic frequencies greater than or equal to 1%. We did not include rare alleles (f < 1%) to consider only the most representative alleles of each population and reduce the bias of unbalanced sample sizes. We have shown a strong correlation of allele richness, a diversity measure that explicitly accounts for different sample sizes (Hurlbert 1971; el Mousadik and Petit 1996),
with the number of common alleles. This correlation indicates that the number of 3-digit KIR alleles with frequencies greater or equal to 1% is predictive of KIR allele richness in populations with unbalanced sample sizes. Previously, the Yucpa from Colombia were regarded as exhibiting the lowest KIR diversity, with 25 KIR common alleles (Gendzekhadze et al. 2009). We observed similarly low diversity in the Brazilian Amerindians, with Kaingang of Rio das Cobras exhibiting an even smaller number of KIR alleles at 3-digit resolution (n = 24). Looking specifically to the allele richness, Aché was the lowest diverse population (26.3), followed by Yucpa (28.9). In sharp contrast, the highest KIR diversity to date was observed in two African KhoeSan populations, Nama and Khomani, with 100 and 98 KIR alleles, respectively, and allele richness 78.3 and 78.9, respectively (Nemat-Gorgani et al. 2018). In the European-descent sample from our study, we found the KIR diversity was comparable to the observed in European descendants from the United States (Vierra-Green et al. 2012; Amorim et al. 2021). Our study is also the first to describe the allelic diversity of all KIR genes in individuals of Japanese ancestry. Our data allow us to conclude that the multiple Amerindian populations analyzed so far are those with the lowest KIR diversity, in agreement with previous studies limited to KIR gene-content level (Gendzekhadze et al. 2006; Augusto et al. 2013a; Augusto et al. 2015).

Nevertheless, we found 21 alleles in our Amerindian samples that were not present in the two urban Brazilian populations of European and Japanese ancestries. In the Kaingang from Ivaí alone, we observed six alleles not found in the other study populations. Four alleles not found in the two urban populations form the centromeric haplotype KIR3DL3*01406~ KIR2DL2*00602~ KIR2DL5B*00601~ KIR2DS5*004. This haplotype was previously described in the Ga-Adangbe from Ghana (Norman et al. 2013b) and the KhoeSan from Southern Africa (Nemat-Gorgani et al. 2018). Similarly, the Guarani Kaiowá presented three alleles that form the haplotype KIR2DL4*022~ KIR3DL1*024N~ KIR2DS4*00104, reported in the Ga-Adangbe from Ghana (Norman et al. 2007) in addition to KIR2DS4*00104, which was reported in African Americans (Hou et al. 2009). Interestingly, the allele KIR2DL5A*01201, first described in one African American individual (Hou et al. 2009), was found in all the study Amerindians (except the Aché), exhibiting frequencies from 0.6% to 4.2%. Our data suggest that some of the current alleles in the Amerindian populations may result from gene flow from African
populations centuries ago. In fact, genomic-wide analysis indicates that the beginning of
the admixture of Amerindians with Europeans and Africans started several generations
ago when Brazil was still a Portuguese colony (Kehdy et al. 2015; Castro e Silva et al.
2020). However, our population differentiation and principal component analyses
corroborate previous studies that show that despite a low level of gene flow,
Amerindians from this study remained genetically isolated due to strong cultural barriers
(Petzl-Erler et al. 1993; Tsuneto et al. 2003).

Amerindian populations did not group in the PCA analysis, consistent with the intense
demographic effects they experienced (Amos and Hoffman 2010; O’Fallon and Fehren-
Schmitz 2011). For example, the three Guarani groups included in this study diverged
despite sharing a common ancestor (Marrero et al. 2007); the differentiation among
Guarani groups has also been seen for HLA and other immune markers (Tsuneto et al.
2003; Calonga-Solís et al. 2019). Therefore, the KIR differentiation observed in
Amerindians is consistent with their history of intense genetic drift and is even more
pronounced in Yucpa and Aché.

The description of high-resolution (5-digit) KIR haplotypes for all functional KIR genes
and pairwise LD analysis were only achieved by one study that analyzed Euro-
Americans (Amorim et al. 2021). We observed population-specific patterns of allelic
pairwise LD, which highlights the uniqueness of the study populations. Our current study
contributes to revealing the still unknown LD patterns of KIR alleles in global
populations. We found more variation in copy number arrangements, including deletions,
duplications, and hybrid gene formation in telomeric haplotypes than we found in
centromeric haplotypes. This observation is consistent with previous suggestions that
there is a greater selective advantage in the diversification of gene content (presence
and absence of genes) and large structural variations in the telomeric region (Jiang et al.
2012). Among the uncommon haplotypes, we observed deletion of the framework gene
KIR3DL2 in the haplotype cA03~tB07 in Euro-Brazilians and Japanese descendants ($f =
1.38$ and $0.69\%$, respectively), along with the deletion of $\text{KIR2DL1}~\text{KIR3DP1}~
\text{KIR2DL4}~\text{KIR3DS1}~\text{KIR2DL5}~\text{KIR2DS35}$. This pattern possibly represents a fusion of
$\text{KIR2DL1}$ and $\text{KIR2DS1}$ with the deletion of the framework gene $\text{KIR3DL2}$ (Pyo et al.
2013).
In contrast, we found a remarkably low diversity in the Amerindian centromeric KIR haplotypes. For example, in Aché, the gene-content haplotype cA01 was observed in 100% of the individuals, with a minor allelic variation. Interestingly, a study analyzing single nucleotide polymorphisms within genes of the innate immune system (Lindenau et al. 2016) reported reduced heterozygosity in Aché regardless of the similar genomic microsatellites heterozygosity compared to Guarani and Kaingang (Wang et al. 2007).

Because the centromeric KIR region encodes the strongest educators for NK cells (Stewart et al. 2005; Hilton et al. 2015), we hypothesize that the reduced diversity in this region is a consequence of selective pressure imposed by the limited HLA alleles in these populations. We and others have already reported frequencies of the KIR ligands in Brazilians of European ancestry (Augusto et al. 2012), Japanese ancestry (Augusto et al. 2016), and some Brazilian Amerindians (Parham et al. 1997; Augusto et al. 2013b), while we have characterized Aché here for the first time. We observed a limited number of KIR-HLA interactions per individual in the Amerindians from our study, particularly the Guarani Mbya, which has the lowest score ever reported for a human population, 2.86. Most striking is the scarcity of HLA-A and HLA-B ligands, which causes 85 to 97% of all KIR-HLA interactions in all studied Amerindians to rely exclusively on HLA-C. These results support the suggestion that HLA-C specialized and evolved to become primarily KIR ligands while HLA-A and HLA-B kept their primary function as T cell receptor ligands (Older Aguilar et al. 2010; Augusto et al. 2015).

Gendzekhadze and colleagues (Gendzekhadze et al. 2009) earlier suggested that the limited diversity of KIR-HLA in Amerindians may be the minimum necessary for human survival. We further hypothesize that the reduced KIR-HLA interactions might have selected specific sets of KIR centromeric variants in these populations. We observed an excess of synonymous to nonsynonymous substitutions in the centromeric region and significantly increased homozygosity in many centromeric loci, both indicating purifying selection. Conversely, this was not observed in the telomeric portion, and the differences in synonymous and nonsynonymous nucleotide substitution rates in the telomeric region were similar to the observed in the majority of other protein-coding genes (Kryazhimskiy and Plotkin 2008; Dasmeh et al. 2014). Although significant, both dN - dS and Ewens-Watterson analyses may not alone overrule the occurrence of genetic drift and other stochastic factors. To neutralize these confounding variables, we compared the
differentiation and homozygosity of the KIR region with 678 neutral genomic markers (Wang et al. 2007). The observation that Amerindians are less differentiated for KIR than for the neutral genomic markers is an additional signature of selection. Lastly, we found that the homozygosity in the centromeric KIR region is much lower than the homozygosity in the telomeric region compared to genomic homozygosity rates. Once again, these last two observations are robust signatures of purifying selection in centromeric KIR but not in telomeric KIR. Our observations corroborate previous findings showing a geographically specific selection on the KIR complex, particularly on the centromeric region (Yawata et al. 2006; Augusto et al. 2019; Deng et al. 2021). Thus, we suggest that the limited number of HLA-KIR interactions, represented mainly by few HLA-C ligands, impose an intense population-specific stabilizing pressure on the KIR centromeric region to maintain the minimal inhibitory signals required for NK education and the consequent survival of these populations. Population-specific selection of KIR-HLA combinations driven by distinct sets of HLA alleles could also contribute to the greater differentiation observed for the centromeric KIR in our Amerindians compared to the telomeric.

In conclusion, we provide a comprehensive and novel characterization of KIR at high-resolution in unique populations, also in the context of their HLA ligands. For the first time, we report high-resolution allele-level haplotypes in South American populations, including six Amerindian and two Southern Brazilian urban populations. Importantly, we provide compelling evidence of purifying selection on KIR centromeric haplotypes, possibly driven by the reduced diversity of HLA alleles. Genetic characterization of such unique populations is of paramount interest to understand the evolutionary constraints and impacts of reduced diversity in human populations. This study significantly contributes to understanding high-resolution allelic and haplotypic KIR variation in global populations, bringing new insights into linkage disequilibrium patterns among KIR alleles and enhancing our ability to identify KIR haplotypes.
Materials and Methods

Characterization of the study populations

All participants were informed about the research purpose and given written or oral consent to participate in the study, according to the local law and regulation at the time of sample collection. This study was approved by the Human Research Ethics Committee of the Federal University of Paraná and the Brazilian National Human Research Ethics Committee (CONEP), protocol number CAAE 02727412.4.0000.0096, under the Brazilian Federal laws. We analyzed 706 individuals from eight different populations (Figure 7). They include Euro-descendants from Curitiba (CTBA, n=109), Japanese descendants from Curitiba (BrJAP, n=74), Guarani Kaiowá (GKW, n=150), Guarani Ñandeva (GND, n=81), Guarani Mbya (GRC, n=84), Kaingang from Ivaí (KIV, n=93), Kaingang from Rio das Cobras (KRC, n=64), and the Aché (ACHE, n=51). As described previously, Amerindian sample collection occurred between the late 1980s and early 1990s (Tsuneto et al. 2003). Detailed information of the study populations is given in Supplementary Table 11 and Figure 7.

All Amerindian individuals were part of genetically isolated groups living in indigenous lands in the Brazilian states of Paraná, and Mato Grosso do Sul, and in the bordering country Paraguay. The Guarani populations are closely related groups and speak dialects of a common language, Guarani, from the Tupi-Guarani linguistic family. They are subdivided into Guarani Ñandeva, Guarani Kaiowá, and Guarani Mbya, which diverged around 1,800 years ago. In contrast, the Kaingang populations speak a language belonging to the Jê family and were suggested to have split only approximately 200 years ago (Marrero et al. 2007). The Aché, also known as Guayaki, live in Paraguay and also speak a language belonging to the Tupi-Guarani family, sharing some cultural similarities with Guarani and Kaingang groups. However, autosomal and sexual genetic markers suggest they are closer to Guarani groups than Kaingang (Battilana et al. 2002; Gaspar et al. 2002; Tsuneto et al. 2003; Schmitt et al. 2004; Callegari-Jacques et al. 2007).

The urban populations were collected in Curitiba, Paraná State, one of Brazil’s largest cities, with over 2 million inhabitants. Southern Brazil was initially inhabited by
Amerindians and later by Africans brought enslaved by the Portuguese during the colonization period. Later, this region received a large influx of European immigrants during the XIX and XX centuries (Santos 2002; Kehdy et al. 2015; Pena et al. 2020). The population of Curitiba has predominantly European ancestry, with mainly Portuguese, German, Italian, Polish, and Ukrainian backgrounds, among others. According to the Brazilian population census in 2010 (IBGE 2013), 78.8% of Curitiba’s population self-declared as Euro-descendants, 16.7% admixed, 3% Afro-descendants, 1.4% Asian, and 0.2% Native American.

During Japan’s crisis after the Meiji Restoration in the late XIX century, there was a large immigration of Japanese people to Brazil, the United States, Peru, and Mexico (Sakurai et al. 2010). Currently, Brazil hosts the largest Japanese population outside Japan, with over 1.5 million Japanese descendants living mainly in São Paulo and Paraná States (IBGE 2013). Paraná hosts the second-largest Japanese community in Brazil, comprised of over thirty thousand Japanese or Japanese descendants, which remain relatively isolated. All Japanese-descendant individuals included in this study were born in Curitiba, Paraná, and reported that all four grandparents were born in Japan without known admixture with non-Japanese ancestries.

**KIR and HLA sequencing**

DNA was extracted using the phenol-chloroform-isoamyl alcohol method (Sambrook et al. 1989) or the salting-out method (Lahiri and Nurnberger 1991) and stored at -80°C. DNA was enzymatically fragmented using the KAPA HyperPlus kit (Roche, USA) and barcoded with unique adaptors. Dual size selection was performed with AMPure XP magnetic beads (Beckman Coulter, USA) to obtain fragments with an average size of 780 bp. Quality control was performed using the PicoGreen® kit (Thermo Fisher Scientific, Waltham, MA) for product quantification and Bioanalyzer (Agilent, Santa Clara, CA) to determine the quantity and size of fragments. Pooling was performed with the automated liquid handler Echo® 525 (Labcyte, San Jose, CA). The enrichment of the targeted regions was performed with the Nextera kit (Illumina, San Diego, CA) using 10,456 biotinylated probes designed by Norman et al. (2016) to capture fragments corresponding to all KIR and HLA class I loci. After this step, the fragments were purified.
and amplified. Sequencing was performed using Illumina HiSeq 4000 (Illumina, San Diego, CA) 150 bp paired-end protocol.

**Data analysis**

Sequence filtering, alignment, and genotype calling of KIR genes were made using an updated version of the PING bioinformatic pipeline (Norman et al. 2016; Marin et al. 2021). Genotype and copy number were obtained for all KIR loci, except the two pseudogenes for which we only determined copy number. After processing with PING, KIR data were manually curated for the resolution of ambiguities. For HLA-A, HLA-B, and HLA-C genotyping, we processed raw FASTQ files with HLA Explore™ (Omixon, Hungary), which determined unambiguous calls for 3-field resolution genotypes.

Allele frequencies, allelic richness, and the proportion of shared KIR alleles among populations were calculated and plotted using a custom version of the PopGenReport R package (Adamack and Gruber 2014). Intersecting sets of alleles were identified using `intersect` from base R and plotted using `upset` from R package UpSetR (Conway et al. 2017). Linkage disequilibrium between the multiallelic KIR loci was performed on Arlequin version 3.5.2 (Excoffier and Lischer 2010). Population pairwise F\textsubscript{ST} was calculated with PopGenReport and locus-specific F\textsubscript{ST} with pegas R packages (Weir and Cockerham 1984; Paradis 2010; Adamack and Gruber 2014). The neighbor-joining (NJ) tree was estimated in the R package ape (Paradis and Schliep 2019). The significance of the genetic differentiation between populations was tested using the exact test of population differentiation (Raymond and Rouset 1995; Goudet et al. 1996) on Arlequin version 3.5.2 (Excoffier and Lischer 2010). Principal components analysis was performed and plotted using the ade4 R package (Dray and Dufour 2007). Maps were plotted using the R packages maps (Becker et al. 2018) and geobr (Pereira and Gonçalves 2020). Neutrality tests of nonsynonymous (d\textsubscript{N}) to synonymous (d\textsubscript{S}) substitutions (Nei and Gojobori 1986) were performed in MEGA X software (Kumar et al. 2018). The statistical significance of the difference was tested using the bootstrap method with 1,000 replicates and performing the two-tailed Z-test of selection, in which $Z = (d\textsubscript{N} - d\textsubscript{S}) / \sqrt{\text{Var}(d\textsubscript{S}) + \text{Var}(d\textsubscript{N})}$ (Kumar et al. 2018). The Ewens-Watterson test was performed with 1,000 replicates on Arlequin version 3.5.2 (Excoffier and Lischer 2010), with Slatkin's probability correction for multiallelic data (Ewens 1972; Watterson 2000).
1978; Slatkin 1994; Slatkin 1996). For this test, significant deviation towards
homozygosity is indicated by probability values close to one, and heterozygote excess is
indicated with values close to zero. Therefore, we established the significant thresholds
as $p < 0.025$ or $p > 0.975$, considering a cumulative error of 5% for the Ewens-Watterson
test. Heterozygosity values were obtained using R package adegenet (Jombart 2008).

We defined KIR centromeric and telomeric haplotypes at the gene-content level
according to the nomenclature previously described (Pyo et al. 2010; Vierra-Green et al.
2012). Furthermore, we included the nomenclature proposed in Pyo et al. (2013) (Pyo et
al. 2013) to describe unusual haplotype patterns (e.g., $tB01$-$d6$, which represents
KIR3DP1- KIR2DL4- KIR3DS1 deletion). The identification of centromeric and telomeric
haplotypes, including the allele-level KIR haplotype determination, were performed
manually based on known patterns (Vierra-Green et al. 2012; Pyo et al. 2013; Roe et al.
2017).

HLA-A, HLA-B, and HLA-C alleles were classified according to their encoded relevant
epitopes for KIR interaction. HLA and KIR data were integrated to generate individual
interaction scores described previously (Nemat-Gorgani et al. 2018). We considered the
following pairs: Bw4 (HLA-A) and KIR3DL1 (Foley et al. 2008); Bw4 (HLA-B) and
KIR3DL1 (Gumperz et al. 1995; Foley et al. 2008); HLA-A*03 and KIR3DL2 (Döhring et
al. 1996); HLA-A*11 and KIR3DL2 (Hansasuta et al. 2004); HLA-C2 e KIR2DL1 (Hilton
et al. 2015); HLA-C1 and 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); HLA-A*11 and KIR2DS4 (Graef et al. 2009); a subset of HLA-C which includes
C2 (alleles HLA-C*05:01, *02:02 and *04:01) and C1 (alleles HLA-C*16:01, *01:02 and
*14:02), recognized by KIR2DS4 (Graef et al. 2009); and, finally, HLA-C2 and a subset
of KIR2DS5 receptors (encoded by alleles KIR2DS5*003, *004, *005, *006, *007 and
*008) (Blokhuis et al. 2017).

Acknowledgments

This work was supported by Coordenação de Aperfeiçoamento de Pessoal de Nível
Superior – Brasil (CAPES) – Finance Code 001; Programa de Apoio a Núcleos de
Excelência – Fundação Araucária de Apoio ao Desenvolvimento Científico e
Tecnológico do Paraná (PRONEX-FA - Agreement 116/2018 - Protocol 50530); Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq); and National Institutes of Health (grant number U19NS095774). We would also like to thank all the participants in this study and all colleagues from the Human Molecular Genetics Laboratory of the Federal University of Paraná, Brazil.

**Author Contributions**

DGA designed the study; LV, DGA, BH, and GM-M sequenced DNA samples; WM and RD processed KIR raw data and DGA processed HLA raw data; LV performed statistical analysis; DGA, MF-V, and JH contributed with reagents; MLP-E, AMH, KRH, LTT, MHH, and FMS contributed with DNA samples; LV and DGA drafted the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

**Data availability**

The data underlying this article are available in the article itself and in its online supplementary material.

**Competing Interest Statement**

The authors declare no competing interest.
References

Abi-Rached L, Moesta AK, Rajalingam R, Guethlein LA, Parham P. 2010. Human-specific evolution and adaptation led to major qualitative differences in the variable receptors of human and chimpanzee natural killer cells. *PLoS Genetics*.

Adamack AT, Gruber B. 2014. PopGenReport: Simplifying basic population genetic analyses in R. *Methods in Ecology and Evolution*.

Alicata C, Ashouri E, Nemat-Gorgani N, Guethlein LA, Marin WM, Tao S, Moretta L, Hollenbach JA, Trowsdale J, Traherne JA, et al. 2020. KIR Variation in Iranians Combines High Haplotype and Allotype Diversity With an Abundance of Functional Inhibitory Receptors. *Frontiers in Immunology* 11:556.

Amorim LM, Augusto DG, Nemat-Gorgani N, Montero-Martin G, Marin WM, Shams H, Dandekar R, Caillier S, Parham P, Fernández-Viña MA, et al. 2021. High-Resolution Characterization of KIR Genes in a Large North American Cohort Reveals Novel Details of Structural and Sequence Diversity. *Frontiers in Immunology* 12:674778.

Amos W, Hoffman JI. 2010. Evidence that two main bottleneck events shaped modern human genetic diversity. *Proceedings of the Royal Society B: Biological Sciences*.

Anderson KM, Augusto DG, Dandekar R, Shams H, Zhao C, Yusufali T, Montero-Martín G, Marin WM, Nemat-Gorgani N, Creary LE, et al. 2020. Killer Cell Immunoglobulin-like Receptor Variants Are Associated with Protection from Symptoms Associated with More Severe Course in Parkinson Disease. *The Journal of Immunology* 205:1323–1330.

Augusto DG. 2016. The impact of KIR polymorphism on the risk of developing cancer: Not as strong as imagined? *Frontiers in Genetics*.

Augusto DG, Amorim LM, Farias TDJ, Petzl-Erler ML. 2016. KIR and HLA genotyping of Japanese descendants from Curitiba, a city of predominantly European ancestry from Southern Brazil. *Human Immunology*.

Augusto DG, Hollenbach JA, Petzl-Erler ML. 2015. A deep look at KIR-HLA in Amerindians: Comprehensive meta-analysis reveals limited diversity of KIR haplotypes. *Human Immunology*.

Augusto DG, Lobo-Alves SC, Melo MF, Pereira NF, Petzl-Erler ML. 2012. Activating KIR and HLA Bw4 ligands are associated to decreased susceptibility to pemphigus foliaceus, an autoimmune blistering skin disease. *PLoS ONE* 7.

Augusto DG, Norman PJ, Dandekar R, Hollenbach JA. 2019. Fluctuating and geographically specific selection characterize rapid evolution of the human Kir region. *Frontiers in Immunology* 10:989.

Augusto DG, Petzl-Erler ML. 2015. KIR and HLA under pressure: evidences of coevolution across worldwide populations. *Human Genetics* 134:929–940.
Augusto DG, Piovezan BZ, Tsuneto LT, Callegari-Jacques SM, Petzl-Erler ML. 2013a. KIR Gene Content in Amerindians Indicates Influence of Demographic Factors. *PLoS ONE* 8:e56755.

Augusto DG, Piovezan BZ, Tsuneto LT, Callegari-Jacques SM, Petzl-Erler ML. 2013b. KIR Gene Content in Amerindians Indicates Influence of Demographic Factors. *PLoS ONE*.

Augusto DG, Zehnder-Alves L, Pincerati MR, Martin MP, Carrington M, Petzl-Erler ML. 2012. Diversity of the KIR gene cluster in an urban Brazilian population. *Immunogenetics* 64:143–152.

Bamshad M, Wooding SP. 2003. Signatures of natural selection in the human genome. *Nature Reviews Genetics* 4:99–111.

Battilana J, Bonatto SL, Freitas LB, Hutz MH, Weimer TA, Callegari-Jacques SM, Batzer MA, Hill K, Hurtado AM, Tsuneto LT, et al. 2002. Alu insertions versus blood group plus protein genetic variability in four Amerindian populations. *Annals of Human Biology*.

Becker RA, Wilks AR, Brownrigg R, Minka TP, Deckmyn A. 2018. maps: Draw Geographical Maps.

Belich MP, Madrigal JA, Hildebrand WH, Zemmour J, Williams RC, Luz R, Petzl-Erler ML, Parham P. 1992. Unusual HLA-B alleles in two tribes of Brazilian Indians. *Nature*.

Blokhuis JH, Hilton HG, Guethlein LA, Norman PJ, Nemat-Gorgani N, Nakimuli A, Chazara O, Moffett A, Parham P. 2017. KIR2DSS allotypes that recognize the C2 epitope of HLA-C are common among Africans and absent from Europeans. *Immunity Inflammation and Disease*.

Callegari-Jacques SM, Crossetti SG, Kohlrausch FB, Salzano FM, Tsuneto LT, Petzl-Erler ML, Hill K, Hurtado AM, Hutz MH. 2007. The β-globin gene cluster distribution revisited - Patterns in native American populations. *American Journal of Physical Anthropology*.

Calonga-Solís V, Malheiros D, Beltrame MH, de Brito Vargas L, Dourado RM, Issler HC, Wassem R, Petzl-Erler ML, Augusto DG. 2019. Unveiling the diversity of Immunoglobulin Heavy Constant Gamma (IGHG) gene segments in Brazilian populations reveals 28 novel alleles and evidence of gene conversion and natural selection. *Frontiers in Immunology* 10.

Carrington M, Wang S, Martin MP, Gao X, Schifman M, Cheng J, Herrera R, Rodriguez AC, Kurman R, Mortel R, et al. 2005. Hierarchy of resistance to cervical neoplasia mediated by combinations of killer immunoglobulin-like receptor and human leukocyte antigen loci. *The Journal of Experimental Medicine*.

Castro e Silva MA, Nunes K, Lemes RB, Mas-Sandoval À, Amorim CEG, Krieger JE, Mill JG, Salzano FM, Bortolini MC, da Costa Pereira A, et al. 2020. Genomic insight into the origins and dispersal of the Brazilian coastal natives. *Proceedings of the National Academy of Sciences of the United States of America*.

Conroy GC, Weber GW, Seidler H, Recheis W, Nedden D Zur, Mariam JH. 2000. Genetic structure of Quechua-speakers of the Central Andes and geographic patterns of gene frequencies in South Amerindian populations. *American Journal of Physical Anthropology*.

Conway JR, Lex A, Gehlenborg N. 2017. UpSetR: An R package for the visualization of intersecting sets and their properties. *Bioinformatics*. 21
Dasmeh P, Serohijos AWR, Kepp KP, Shakhnovich EI. 2014. The influence of selection for protein stability on dN/dS estimations. *Genome Biology and Evolution* 6:2956–2967.

DeGiorgio M, Jakobsson M, Rosenberg NA. 2009. 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*.

Deng Z, Zhen J, Harrison GF, Zhang G, Chen R, Sun G, Yu Q, Nemat-gorgani N, Guethlein LA, He L, et al. 2021. Adaptive Admixture of HLA class I Allotypes Enhanced Genetically Determined Strength of Natural Killer Cells in East Asians. *Molecular Biology and Evolution*:msab053.

Döhring C, Scheidegger D, Samaridis J, Cella M, Colonna M. 1996. A human killer inhibitory receptor specific for HLA-A1,2. *Journal of immunology (Baltimore, Md. : 1950)*.

Dray S, Dufour AB. 2007. The ade4 package: Implementing the duality diagram for ecologists. *Journal of Statistical Software*.

Ewens WJ. 1972. The sampling theory of selectively neutral alleles. *Theoretical Population Biology* 3:87–112.

Excoffier L, Lischer HEL. 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* 10:564–567.

Fernández-Viña MA, Lázaro AM, Marcos CY, Nulf C, Raimondi E, Haas EJ, Stastny P. 1997. Dissimilar evolution of B-locus versus A-locus and class II loci of the HLA region in South American Indian tribes. *Tissue Antigens* 50.

Flores AC, Marcos CY, Paladino N, Capucchio M, Theiler G, Arruvito L, Pardo R, Habegger A, Williams F, Middleton D, et al. 2007. KIR genes polymorphism in Argentinean Caucasoid and Amerindian populations. *Tissue Antigens* 69:568–576.

Flórez-Álvarez L, Hernandez JC, Zapata W. 2018. NK cells in HIV-1 infection: From basic science to vaccine strategies. *Frontiers in Immunology*.

Foley BA, De Santis D, Van Beelen E, Lathbury LJ, Christiansen FT, Witt CS. 2008. The reactivity of Bw4+ HLA-B and HLA-A alleles with KIR3dDll: Implications for patient and donor suitability for haploidentical stem cell transplantations. *Blood*.

Fukami-Kobayashi K, Shiina T, Anzai T, Sano K, Yamazaki M, Inoko H, Tateno Y. 2005. Genomic evolution of MHC class I region in primates. *Proceedings of the National Academy of Sciences of the United States of America* 102:9230–9234.

García-Ortiz JE, Sandoval-Ramírez L, Rangel-Villalobos H, Maldonado-Torres H, Cox S, García-Sepúlveda CA, Figuera LE, Marsh SGE, Little AM, Madrigal JA, et al. 2006. High-resolution molecular characterization of the HLA class I and class II in the Tarahumara Amerindian population. *Tissue Antigens* 68.

Gaspar PA, Hutz MH, Salzano FM, Hill K, Hurtado AM, Petzl-Erler ML, Tsuneto LT, Weimer TA. 2002. Polymorphisms of CYP1A1, CYP2E1, GSTM1, GSTT1, and TP53 genes in Amerindians. *American Journal of Physical Anthropology*.
Gendzekhadze K, Norman PJ, Abi-Rached L, Graef T, Moesta AK, Layrisse Z, Parham P. 2009. Coevolution 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.*

Gendzekhadze K, Norman PJ, Abi-Rached L, Layrisse Z, Parham P. 2006. High KIR diversity in Amerindians is maintained using few gene-content haplotypes. *Immunogenetics* 58:474–480.

González-Galarza FF, Takeshita LYC, Santos EJM, Kempson F, Maia MHT, Da Silva ALS, Teles E Silva AL, Ghattaoraya GS, Alfirevic A, Jones AR, et al. 2015. Allele frequency net 2015 update: New features for HLA epitopes, KIR and disease and HLA adverse drug reaction associations. *Nucleic Acids Research* 43:D784–D788.

Goudet J, Raymond M, De Meeüs T, Rousset F. 1996. Testing differentiation in diploid populations. *Genetics* 144:1933–1940.

Graef T, Moesta AK, Norman PJ, Abi-Rached L, Vago L, Older Aguilar AM, Gleimer M, Hammond JA, Guethlein LA, Bushnell DA, et al. 2009. 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.*

Guethlein LA, Abi-Rached L, Hammond JA, Parham P. 2007. The expanded cattle KIR genes are orthologous to the conserved single-copy KIR3DX1 gene of primates. *Immunogenetics.*

Gumperz JE, Litwin V, Phillips JH, Lanier LL, Parham P. 1995. 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.*

Gutiérrez-Rodríguez ME, Sandoval-Ramírez L, Díaz-Flores M, Marsh SGE, Valladares-Salgado A, Madrigal JA, Mejía-Arangure JM, García CA, Huerta-Zepeda A, Ibarra-Cortés B, et al. 2006. KIR Gene in Ethnic and Mestizo Populations from Mexico. *Human Immunology* 67:85–93.

Hansasuta P, Dong T, Thananchai H, Weekes M, Willberg C, Aldemir H, Rowland-Jones S, Braud VM. 2004. Recognition of HLA-A3 and HLA-A11 by KIR3DL2 is peptide-specific. *European Journal of Immunology.*

Herberman RB, Holden HT. 1978. Natural cell-mediated immunity. *Advances in Cancer Research.*

Hiby SE, Walker JJ, O'Shaughnessy KM, Redman CWG, Carrington M, Trowsdale J, Moffett A. 2004. Combinations of Maternal KIR and Fetal HLA-C Genes Influence the Risk of Preeclampsia and Reproductive Success. *The Journal of Experimental Medicine* 200:957–965.

Hill AVS. 1999. Defence by diversity. *Nature.*

Hilton HG, Guethlein LA, Goyos A, Nemat-Gorgani N, Bushnell DA, Norman PJ, Parham P. 2015. Polymorphic HLA-C Receptors Balance the Functional Characteristics of KIR Haplotypes. *Journal of Immunology* 195:3160–3170.
Hollenbach JA, Nocedal I, Ladner MB, Single RM, Trachtenberg EA. 2012. Killer cell immunoglobulin-like receptor (KIR) gene content variation in the HGDP-CEPH populations. *Immunogenetics*.

Horton R, Wilming L, Rand V, Lovering RC, Bruford EA, Khodiyar VK, Lush MJ, Povey S, Talbot CC, Wright MW, et al. 2004. Gene map of the extended human MHC. *Nature Reviews Genetics*.

Hou LH, Chen M, Jiang B, Kariyawasam K, Ng J, Hurley CK. 2009. In contrast to other stimulatory natural killer cell immunoglobulin-like receptor loci, several KIR2DS5 alleles predominate in African Americans. *Human Immunology* 70:733–737.

Hsu KC, Liu X-R, Selvakumar A, Mickelson E, O’Reilly RJ, Dupont B. 2002. Killer Ig-Like Receptor Haplotype Analysis by Gene Content: Evidence for Genomic Diversity with a Minimum of Six Basic Framework Haplotypes, Each with Multiple Subsets. *The Journal of Immunology* 169:5118–5129.

Hurlbert SH. 1971. The Nonconcept of Species Diversity: A Critique and Alternative Parameters. *Ecology* 52.

IBGE. 2013. Censo 2010.

Jiang W, Johnson C, Jayaraman J, Simecek N, Noble J, Moffatt MF, Cookson WO, Trowsdale J, Traherne JA. 2012. Copy number variation leads to considerable diversity for B but not A haplotypes of the human KIR genes encoding NK cell receptors. *Genome Research* 20:1845–1854.

Johnson AH, Noreen H, Spees EK, Villalobos H, Serrano H, Amos DB, Yunis EJ. 1978. The Distribution of HLA Antigens in the Motilones Indians of Venezuela. *Tissue Antigens* 12:163–169.

Jombart T. 2008. Adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics* 24:1403–1405.

Jones DC, Hiby SE, Moffett A, Trowsdale J, Young NT. 2006. Nature of allelic sequence polymorphism at the KIR3DL3 locus. *Immunogenetics* 58:614–627.

Kärre K, Ljunggren HG, Piontek G, Kiessling R. 1986. Selective rejection of H–2-deficient lymphoma variants suggests alternative immune defence strategy. *Nature* 319:675–678.

Kehdy FSG, Gouveia MH, Machado M, Magalhães WCS, Horimoto AR, Horta BL, Moreira RG, Leal TP, Sciar MO, Soares-Souza GB, et al. 2015. 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*.

Khakoo SI, Rajalingam R, Shum BP, Weidenbach K, Flodin L, Muir DG, Canavez F, Cooper SL, Valiante NM, Lanier LL, et al. 2000. Rapid evolution of NK cell receptor systems demonstrated by comparison of chimpanzees and humans. *Immunity*.

Khakoo SI, Thio CL, Martin MP, Brooks CR, Gao X, Astemborski J, Cheng J, Goedert JJ, Vlahov D, Hilgartner M, et al. 2004. HLA and NK cell inhibitory receptor genes in resolving hepatitis C virus infection. *Science (New York, N.Y.)* 305:872–874.
Kim S, Poursine-Laurent J, Truscott SM, Lybarger L, Song Y-J, Yang L, French AR, Sunwoo JB, Lemieux S, Hansen TH, et al. 2005. Licensing of natural killer cells by host major histocompatibility complex class I molecules. *Nature* 436:709–713.

Kryazhimskiy S, Plotkin JB. 2008. The population genetics of dN/dS. *PLoS Genetics* 4:e1000304.

Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution* 35:1547–1549.

Lahiri DK, Nurnberger JL. 1991. A rapid non-enzymatic method for the preparation of HMW DNA from blood for RFLP studies. *Nucleic acids research* 19:5444.

Lanier LL. 2008. Up on the tightrope: natural killer cell activation and inhibition. *Nature Immunology* 9:495–502.

Lanier LL, Phillips JH. 1995. NK cell recognition of major histocompatibility complex class I molecules. *Seminars in Immunology* 7:75–82.

Layrisse Z, Guedez Y, Domínguez E, Paz N, Montagnani S, Matos M, Herrera F, Ogando V, Balbas O, Rodríguez-Larralde A. 2001. Extended HLA haplotypes in a Carib Amerindian population: The Yucpa of the Perija Range. In: Human Immunology. Vol. 62.

Lindenau JDR, Salzano FM, Hurtado AM, Hill KR, Petzl-Erler ML, Tsuneto LT, Hutz MH. 2016. Variability of innate immune system genes in Native American populations - Relationship with history and epidemiology. *American Journal of Physical Anthropology* 159:722–728.

Marin WM, Dandekar R, Augusto DG, Yusufali T, Heyn B, Hofmann J, Lange V, Sauter J, Norman PJ, Hollenbach JA. 2021. High-throughput interpretation of killer-cell immunoglobulin-like receptor short-read sequencing data with PING. *PLoS Computational Biology* 17:e1008904.

Marrero AR, Silva-Junior WA, Bravi CM, Hutz MH, Petzl-Erler ML, Ruiz-Linares A, Salzano FM, Bortolini MC. 2007. Demographic and evolutionary trajectories of the Guarani and Kaingang natives of Brazil. *American journal of physical anthropology* 132:301–310.

Martin MP, Bashirova A, Traherne J, Trowsdale J, Carrington M. 2003. Cutting edge: expansion of the KIR locus by unequal crossing over. *Journal of immunology (Baltimore, Md. : 1950)* 171:2192–2195.

Martin MP, Gao X, Lee JH, Nelson GW, Detels R, Goedert JJ, Buchbinder S, Hoots K, Vlahov D, Trowsdale J, et al. 2002. Epistatic interaction between KIR3DS1 and HLA-B delays the progression to AIDS. *Nature Genetics*.

Martin MP, Single RM, Wilson MJ, Trowsdale J, Carrington M. 2008. KIR haplotypes defined by segregation analysis in 59 Centre d’Etude Polymorphisme Humain (CEPH) families. *Immunogenetics* 60:767–774.

Moesta AK, Graef T, Abi-Rached I, Older Aguilar AM, Guethlein LA, Parham P. 2010. 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*.
Moffett A, Colucci F. 2015. Coevolution of NK receptors and HLA ligands in humans is driven by reproduction. *Immunological Reviews* 267.

Moretta A, Bottino C, Vitale M, Pende D, Biassoni R, Mingari MC, Moretta L. 1996. RECEPTORS FOR HLA CLASS-I MOLECULES IN HUMAN NATURAL KILLER CELLS. *Annual Review of Immunology*.

Morvan MG, Lanier LL. 2016. NK cells and cancer: You can teach innate cells new tricks. *Nature Reviews Cancer*.

el Mousadik A, Petit RJ. 1996. High level of genetic differentiation for allelic richness among populations of the argan tree [Argania spinosa (L.) Skeels] endemic to Morocco. *Theoretical and Applied Genetics* 92.

Mugal CF, Wolf JBW, Kaj I. 2014. Why time matters: Codon evolution and the temporal dynamics of dN/dS. *Molecular Biology and Evolution* 31:212–231.

Nei M, Gojobori T. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Molecular Biology and Evolution* 3:418–426.

Nelson GW, Martin MP, Gladman D, Wade J, Trowsdale J, Carrington M. 2004. 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* 173:4273–4276.

Nemat-Gorgani N, Edinur HA, Hollenbach JA, Traherne JA, Dunn PJP, Chambers GK, Parham P, Norman PJ. 2014. KIR diversity in Māori and Polynesians: populations in which HLA-B is not a significant KIR ligand. *Immunogenetics* 66:597–611.

Nemat-Gorgani N, Guethlein LA, Henn BM, Norberg SJ, Chiaroni J, Sikora M, Quintana-Murci L, Mountain JL, Norman PJ, Parham P. 2019. Diversity of KIR, HLA Class I, and Their Interactions in Seven Populations of Sub-Saharan Africans. *The Journal of Immunology* 202:2636–2647.

Nemat-Gorgani N, Hilton HG, Henn BM, Lin M, Gignoux CR, Myrick JW, Werely CJ, Granka JM, Möller M, Hoal EG, et al. 2018. 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* 200:2640–2655.

Nielsen R. 2001. Statistical tests of neutrality at the age of genomics. *Heredity* 86:641–647.

Nielsen R. 2005. Molecular signatures of natural selection. *Annual Review of Genetics* 39:197–218.

Norman PJ, Abi-Rached L, Gendzekhadze K, Hammond JA, Moesta AK, Sharma D, Graef T, McQueen KL, Guethlein LA, Carrington CVF, et al. 2009. Meiotic recombination generates rich diversity in NK cell receptor genes, alleles, and haplotypes. *Genome Research* 19:757–769.
Norman PJ, Abi-Rached L, Gendzekhadze K, Korbel D, Gleimer M, Rowley D, Bruno D, Carrington CVF, Chandanayingyong D, Chang YH, et al. 2007. Unusual selection on the KIR3DL1/S1 natural killer cell receptor in Africans. *Nature Genetics* 39:1092–1099.

Norman PJ, Hollenbach JA, Nemat-Gorgani N, Guethlein LA, Hilton HG, Pando MJ, Koram KA, Riley EM, Abi-Rached L, Parham P. 2013a. Coevolution 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*.

Norman PJ, Hollenbach JA, Nemat-Gorgani N, Guethlein LA, Hilton HG, Pando MJ, Koram KA, Riley EM, Abi-Rached L, Parham P. 2013b. Coevolution 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* 9.

Norman PJ, Hollenbach JA, Nemat-Gorgani N, Marin WM, Norberg SJ, Ashouri E, Jayaraman J, Wroblewski EE, Trowsdale J, Rajalingam R, et al. 2016. Defining KIR and HLA Class I Genotypes at Highest Resolution via High-Throughput Sequencing. *American Journal of Human Genetics* 99:375–391.

O'Fallon BD, Fehren-Schmitz L. 2011. Native Americans experienced a strong population bottleneck coincident with European contact. *Proceedings of the National Academy of Sciences of the United States of America*.

Older Aguilar AM, Guethlein LA, Adams EJ, Abi-Rached L, Moesta AK, Parham P. 2010. Coevolution of Killer Cell Ig-Like Receptors with HLA-C To Become the Major Variable Regulators of Human NK Cells. *The Journal of Immunology* 185:4238–4251.

Paradis E. 2010. Pegas: An R package for population genetics with an integrated-modular approach. *Bioinformatics*.

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

Parham P, Arnett KL, Adams EJ, Little AM, Tees K, Barber LD, Marsh SGE, Ohta T, Markow T, Petzl-Erler ML. 1997. Episodic evolution and turnover of HLA-B in the indigenous human populations of the Americas. *Tissue Antigens* 50:219–232.

Parham P, Guethlein LA. 2018. Genetics of Natural Killer Cells in Human Health, Disease, and Survival. *Annual Review of Immunology* 36:519–548.

Pena SDJ, Santos FR, Tarazona-Santos E. 2020. Genetic admixture in Brazil. *American Journal of Medical Genetics* 184:928–938.

Pereira RHM, Gonçalves CN. 2020. geobr: Loads Shapefiles of Official Spatial Data Sets of Brazil.

Petzl-Erler ML, Luz R, Sotomaior VS. 1993. The HLA polymorphism of two distinctive South-American Indian tribes: The Kaingang and the Guaraní. *Tissue Antigens* 41:227–237.

Piontkivska H, Nei M. 2003. Birth-and-death evolution in primate MHC class I genes: Divergence time estimates. *Molecular Biology and Evolution* 20:601–609.
Pyo CW, Guethlein LA, Vu Q, Wang R, Abi-Rached L, Norman PJ, Marsh SGE, Miller JS, Parham P, Geraghty DE. 2010. 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 5:e15115.

Pyo CW, Wang R, Vu Q, Cereb N, Yang SY, Duh FM, Wolinsky S, Martin MP, Carrington M, Geraghty DE. 2013. Recombinant structures expand and contract inter and intragenic diversification at the KIR locus. BMC Genomics 14:89.

Raymond M, Rousset F. 1995. An exact test for population differentiation. Evolution 49:1280–1283.

Reich D, Patterson N, Campbell D, Tandon A, Mazieres S, Ray N, Parra M V., Rojas W, Duque C, Mesa N, et al. 2012. Reconstructing Native American population history. Nature 488:370–374.

Robinson J, Halliwell JA, Hayhurst JD, Flicek P, Parham P, Marsh SGE. 2015. The IPD and IMGT/HLA database: Allele variant databases. Nucleic Acids Research 43:D423–D431.

Roe D, Vierra-Green C, Pyo CW, Eng K, Hall R, Kuang R, Spellman S, Ranade S, Geraghty DE, Maiers M. 2017. Revealing complete complex KIR haplotypes phased by long-read sequencing technology. Genes and Immunity 18:127–134.

Sakurai C, Asari AY, Beltrão KI, Kodama K, Kawamura L, Oliveira LAP, Bassanezi MSSB, Ninomiya M, Schpun MR, Pereira N de OM, et al. 2010. Resistência & integração 100 anos de imigração japonesa no Brasil. Rio de Janiero: IBGE.

Sambrook J, Fritsch EF, Maniatis T. 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press:626.

Santos SA dos. 2002. Historical Roots of the Whitening of Brazil. Latin American Perspectives 29:61–82.

Schmitt R, Bonatto SL, Freitas LB, Muschner VC, Hill K, Hurtado AM, Salzano FM. 2004. Extremely limited mitochondrial DNA variability among the Aché Natives of Paraguay. Annals of Human Biology.

Single RM, Martin MP, Gao X, Meyer D, Yeager M, Kidd JR, Kidd KK, Carrington M. 2007. Global diversity and evidence for coevolution of KIR and HLA. Nature Genetics 39:1114–1119.

Skoglund P, Mallick S, Bortolini MC, Chennagiri N, Hünemeier T, Petzl-Erler ML, Salzano FM, Patterson N, Reich D. 2015. Genetic evidence for two founding populations of the Americas. Nature.

Slatkin M. 1994. An exact test for neutrality based on the Ewens sampling distribution. Genetical Research 64:71–74.
Slatkin M. 1996. A correction to the exact test based on the Ewens sampling distribution. *Genetical Research* 68:259–260.

Van der Slik AR, Koeleman BPC, Verduijn W, Bruining GJ, Roep BO, Giphart MJ. 2003. KIR in type 1 diabetes: Disparate distribution of activating and inhibitory natural killer cell receptors in patients versus HLA-matched control subjects. *Diabetes* 52:2639–2642.

Smyth MJ, Cretney E, Kelly JM, Westwood JA, Street SEA, Yagita H, Takeda K, Dommelen SLH Van, Degli-Esposti MA, Hayakawa Y. 2005. Activation of NK cell cytotoxicity. *Molecular Immunology* 42:501–510.

Solloch U V, Schefzyk D, Schäfer G, Massalski C, Kohler M. 2020. Estimation of German KIR Allele Group Haplotype Frequencies. *Frontiers in Immunology* 11.

Stewart CA, Laugier-Anfossi F, Vely F, Saulquin X, Riedmuller J, Tisserant A, Gauthier L, Romagne F, Ferracci G, Arosa FA, et al. 2005. Recognition of peptide-MHC class I complexes by activating killer immunoglobulin-like receptors. *Proceedings of the National Academy of Sciences* 102:13224–13229.

Tao S, Kichula KM, Harrison GF, Farias TDJ, Palmer WH, Leaton LA, Hajar CGN, Zefarina Z, Edinur HA, Zhu F, et al. 2020. The combinatorial diversity of KIR and HLA class I allotypes in Peninsular Malaysia. *Immunology*.

Tarazona-Santos E, Carvalho-Silva DR, Pettener D, Luiselli D, De Stefano GF, Labarga CM, Rickards O, Tyler-Smith C, Pena SDJ, Santos FR. 2001. Genetic differentiation in South Amerindians is related to environmental and cultural diversity: Evidence from the Y chromosome. *American Journal of Human Genetics*.

Toneva M, Lepage V, Lafay G, Dulphy N, Busson M, Lester S, Vu-Trin a, Michaylova a, Naumova E, McCluskey J, et al. 2001. Genomic diversity of natural killer cell receptor genes in three populations. *Tissue antigens* 57:358–362.

Traherne JA, Martin M, Ward R, Ohashi M, Pellett F, Gladman D, Middleton D, Carrington M, Trowsdale J. 2010. Mechanisms of copy number variation and hybrid gene formation in the KIR immune gene complex. *Human Molecular Genetics*.

Tsuneto LT, Probst CM, Hutz MH, Salzano FM, Rodriguez-Delfin LA, Zago MA, Hill K, Hurtado AM, Ribeiro-Dos-Santos AKC, Petzl-Erler ML. 2003. HLA class II diversity in seven Amerindian populations. Clues about the origins of the Ach?? *Tissue Antigens* 62:512–526.

Uhrberg M, Parham P, Wernet P. 2002. Definition of gene content for nine common group B haplotypes of the Caucasoid population: KIR haplotypes contain between seven and eleven KIR genes. *Immunogenetics* 54:221–229.

Vargas L de B, Dourado RM, Amorim LM, Ho B, Calonga-Solís V, Issler HC, Marin WM, Beltrame MH, Petzl-Erler ML, Hollenbach JA, et al. 2020. Single Nucleotide Polymorphism in KIR2DL1 Is Associated With HLA-C Expression in Global Populations. *Frontiers in Immunology* 11:1–8.
Vierra-Green C, Roe D, Hou L, Hurley CK, Rajalingam R, Reed E, Lebedeva T, Yu N, Stewart M, Noreen H, et al. 2012. Allele-Level Haplotype Frequencies and Pairwise Linkage Disequilibrium for 14 KIR Loci in 506 European-American Individuals. *PLoS ONE* 7:1–10.

Wang S, Lewis CM, Jakobsson M, Ramachandran S, Ray N, Bedoya G, Rojas W, Parra M v., Molina JA, Gallo C, et al. 2007. Genetic variation and population structure in Native Americans. *PLoS Genetics* 3:2049–2067.

Watkins DI, McAdam SN, Liu X, Strang CR, Milford EL, Levine CG, Garber TL, Dogon AL, Lord CI, Ghim SH, et al. 1992. New recombinant HLA-B alleles in a tribe of South American Amerindians indicate rapid evolution of MHC class I loci. *Nature* 357:329–333.

Watterson GA. 1978. The homozygosity test of neutrality. *Genetics* 88:405–417.

Weir BS, Cockerham CC. 1984. Estimating F-Statistics for the Analysis of Population Structure. *Evolution*.

Wende H, Colonna M, Ziegler A, Volz A. 1999. Organization of the leukocyte receptor cluster (LRC) on human Chromosome 19q13.4. *Mammalian Genome* 10:154–160.

Wilson MJ, Torkar M, Haude A, Milne S, Jones T, Sheer D, Beck S, Trowsdale J. 2000. 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* 97:4778–4783.

Xiong S, Sharkey AM, Kennedy PR, Gardner L, Farrell LE, Chazara O, Bauer J, Hiby SE, Colucci F, Moffett A. 2013. Maternal uterine NK cell-activating receptor KIR2DS1 enhances placentation. *Journal of Clinical Investigation* 123:4264–4272.

Yawata M, Yawata N, Draghi M, Little A-M, Partheniou F, Parham P. 2006. Roles for HLA and KIR polymorphisms in natural killer cell repertoire selection and modulation of effector function. *The Journal of Experimental Medicine* 203:633–645.

Yawata M, Yawata N, Draghi M, Little AM, Partheniou F, Parham P. 2006. Roles for HLA and KIR polymorphisms in natural killer cell repertoire selection and modulation of effector function. *Journal of Experimental Medicine* 203:633–645.
Figure 1. Overview of the KIR diversity in the study populations. A) The bar plot shows the number of high-resolution (5-digit) alleles found for each KIR gene in the study populations in random subsets of n=50. B) Allele sharing among study
populations. The UpSetR plot is an alternative to the Venn diagram and allows better visualization when comparing multiple sets (populations). The vertical bars on the upper portion represent the absolute number of shared alleles by a unique intersect of populations. On the bottom, filled circles indicate the populations of each intersect that share the number of alleles represented by the vertical bars. e.g.: the first vertical bar indicates that 59 KIR alleles were found exclusively in CTBA (only filled circle), while the second vertical bar indicates that 21 alleles are shared by all populations (all circles are filled). The horizontal bars, in the left, represent the total number of alleles found in each population. Sets of alleles found or shared exclusively in Amerindians are highlighted in green, and in blue are the sets of alleles found or shared only in Japanese or in Japanese and Amerindians. The study populations include Brazilians of European ancestry (CTBA); Brazilians of Japanese ancestry (BrJAP); Aché (ACHE); Guarani Kaiowá (GKW); Guarani Ñandeva (GND); Guarani Mbya (GRC); Kaingang from Ivaí (KIV); and Kaingang from Rio das Cobras (KRC).
Figure 2. Amerindians exhibit the lowest KIR diversity among worldwide populations. Tables display the number of KIR alleles at 3-digit resolution with frequencies greater or equal to 1% (common alleles, CA) and allele richness (AR) in global populations. Red boxes indicate analyzed in the present study. The complete dataset is given in Supplementary Table 4. Raw data from worldwide populations were obtained from previous studies (Gendzekhadze et al. 2009; Vierra-Green et al. 2012; Norman et al. 2013b; Nemat-Gorgani et al. 2014; Nemat-Gorgani et al. 2018; Nemat-Gorgani et al. 2019; Alicata et al. 2020; Tao et al. 2020; Amorim et al. 2021; Deng et al. 2021).
Figure 3. Principal components analysis (PCA) of KIR variation. In PCA analysis, data containing allele frequencies of all KIR genes are reassigned to new variables, called principal components, that summarize the total distribution of the variation. The PCA analysis included worldwide populations studied at the allele level for all KIR genes. To maximize our analysis, we included studies that described KIR variation at 3-digit allelic resolution. The rectangles resume KIR diversity in each population. The two first components plotted here explain 57.39% of KIR variation. The dashed circles highlight the similarity of KIR diversity shared by a group of populations; yellow: Southeast Asians; red: Africans; blue: European-descendants and Middle-Easterns. The study populations include Brazilians of European ancestry (CTBA); Brazilians of Japanese ancestry (BrJAP); Aché (ACHE); Guarani Kaiowá (GKW); Guarani Ñandeva (GND); Guarani Mbya (GRC); Kaingang from Ivaí (KIV); and Kaingang from Rio das Cobras (KRC). Data from previously analyzed populations are found in (Gendzekhadze et al. 2009; Vierra-Green et al. 2012; Norman et al. 2013b; Nemat-Gorgani et al. 2014; Nemat-
Gorgani et al. 2018; Nemat-Gorgani et al. 2019; Alicata et al. 2020; Tao et al. 2020; Amorim et al. 2021; Deng et al. 2021).
Figure 4. Common high-resolution KIR allelic haplotypes in the study populations. 

A) Centromeric and B) telomeric haplotypes are shown. The classification of each haplotype is indicated on the right side of each haplotype, followed by a box containing the frequencies.
their relative frequencies. C) Bar plot shows the proportion of KIR A and B haplotypes in the eight South American study populations. These are subclassified according to the gene content of D) centromeric (Cen A and Cen B) and E) telomeric (Tel A and Tel B) regions. The study populations include Brazilians of European ancestry (CTBA); Brazilians of Japanese ancestry (BrJAP); Aché (ACHE); Guarani Kaiowá (GKW); Guarani Ñandeva (GND); Guarani Mbya (GRC); Kaingang from Ivaí (KIV); and Kaingang from Rio das Cobras (KRC). Refer to Supplementary Tables 5 and 6 for the lists of all haplotypes.
Figure 5. HLA-C ligands represent 85-97% of the total HLA ligands in the Amerindians from this study. Vertical bars represent the average of functional KIR-HLA interactions per individual in worldwide populations. Color code represents the proportion of these interactions that are represented by different HLA subgroups.

*Populations from this study (in red) include Guarani Mbya (GRC); Aché (ACHE); Guarani Kaiowá (GKW); Guarani Ñandeva (GND); Kaingang from Rio das Cobras (KRC); Kaingang from Ivaí (KIV); Brazilians of Japanese ancestry (BrJAP); and Brazilians of European ancestry (CTBA). Data from previously analyzed populations are found in (Yawata, et al. 2006; Gendzekhadze et al. 2009; Vierra-Green et al. 2012; Norman et al. 2013b; Nemat-Gorgani et al. 2014; Nemat-Gorgani et al. 2018; Tao et al. 2020; Amorim et al. 2021; Deng et al. 2021).
Figure 6. Signatures of purifying selection in the centromeric KIR region. A) Analysis of nonsynonymous (d_N) and synonymous (d_S) substitutions in KIR genes. The rates of nonsynonymous substitutions are significantly reduced in the centromeric KIR region compared to the telomeric KIR (Cen d_N-d_S < Tel d_N-d_S; p = 6.0x10^-5). The difference was calculated across the coding sequences of all KIR genes within the centromeric (KIR3DL3~KIR2DL1) and telomeric (KIR2DL4~KIR3DL2) regions. Positive values (d_N > d_S) indicate positive selection and negative values (d_N < d_S) indicate purifying selection. The study populations include Brazilians of European ancestry (CTBA); Brazilians of Japanese ancestry (BrJAP); Aché (ACHE); Guarani Kaiowá (GKW); Guarani Ñandeva (GND); Guarani Mbya (GRC); Kaingang from Ivaí (KIV); and Kaingang from Rio das Cobras (KRC). B) Ewens-Watterson test of neutrality in all KIR genes. The color scale represents the differences between the observed homozygosity (F_obs) and expected homozygosity (F_exp). Asterisk indicates significative p-values (p < 0.025 or p > 0.975). Shades of red in the centromeric region, combined with significant p values, indicate that homozygosity was observed at higher rates than expected under neutral evolution. The homozygosity rates were overall significantly reduced in the centromeric KIR region compared to the telomeric (Cen F_obs-F_exp < Tel F_obs-F_exp; p = 4.0x10^-6). C) The fixation index (F_ST) analysis of the KIR allelic diversity in Amerindians compared to 678 genomic microsatellite markers (Msat) shows that differentiation is significantly reduced in KIR. D) Compared to genomic markers, the heterozygosity in the centromeric KIR region is reduced in Aché, Guarani, and Kaingang (p = 7x10^-4), while the telomeric has similar heterozygosity rates to the genomic microsatellite markers (p = 0.12). The differences in heterozygosity rates between centromeric and telomeric KIR are highly significant (Cen H_obs < Tel H_obs; p = 4.7x10^-4).
Figure 7. Characterization of the study populations. Map of the geographic locations of the eight populations included in this study. The territory in white corresponds to Brazil, light gray highlights Paraguay, and other countries are indicated in dark gray. Sample sizes are shown inside parenthesis. The study populations include Brazilians of European ancestry (CTBA); Brazilians of Japanese ancestry (BrJAP); Aché (ACHE); Guarani Kaiowá (GKW); Guarani Ñandeva (GND); Guarani Mbya (GRC); Kaingang from Ivaí (KIV); and Kaingang from Rio das Cobras (KRC). *The BrJAP population is comprised of individuals whose all ancestors were born in Japan, without known admixture with non-Japanese. For detailed information, please see Supplementary Table 1.
Table 1. Functional KIR-HLA interactions in the study populations. Interaction values are given as the percentage of individuals that present that functional interaction in a population. Blank cells indicate that the interaction was not found in a population. Asterisks indicate that only a subset of the molecules is considered in the interaction, as detailed in the methods section. The proportion of activating and inhibitory interactions in each population is also shown. Mean KIR-HLA indicates how many functional pairs one individual of each population has on average. The minimum and the maximum numbers of interactions observed in a single individual are shown in parenthesis. CTBA: Brazilians of European ancestry. BrJAP: Brazilians of Japanese ancestry. ACHE: Aché. GKW: Guarani Kaiowá. GND: Guarani Ñandeva. GRC: Guarani Mbya. KIV: Kaingang from Ivaí. KRC: Kaingang from Rio das Cobras.

| HLA epitope | KIR          | CTBA | BrJAP | KIV | KRC | GND | GKW | ACHE | GRC |
|-------------|--------------|------|-------|-----|-----|-----|-----|------|-----|
| Bw4 (HLA-A) | KIR3DL1      | 0.09 | 0.17  | 0.02| 0.01| 0.05| 0.01| 0.02 |     |
| Bw4 (HLA-B) | KIR3DL1      | 0.16 | 0.16  | 0.13| 0.10| 0.04| 0.05| 0.01 | 0.04|
|             | KIR3DL2      | 0.06 | 0.01  |     | 0.01|     |     |      |     |
|             | KIR3DL2      | 0.03 | 0.07  |     | 0.02| 0.01| 0.00|      |     |
| A11         | KIR2DS4      | 0.02 | 0.05  |     | 0.01| 0.01|     |      |     |
| C2          | KIR2DL1      | 0.15 | 0.04  | 0.16| 0.14| 0.21| 0.19| 0.26 | 0.06|
| C1C2        | KIR2DL2*     | 0.15 | 0.04  | 0.19| 0.23| 0.12| 0.10| 0.01 | 0.36|
|             | KIR2DL3*     | 0.18 | 0.36  | 0.19| 0.22| 0.32| 0.41| 0.30 | 0.41|
| C2          | KIR2DS1      | 0.06 | 0.01  | 0.11| 0.13| 0.14| 0.16| 0.17 | 0.09|
|             | KIR2DS2      | 0.01 |      |     |     |     |     |      |     |
|             | KIR2DS5*     | 0.10 | 0.09  | 0.20| 0.13| 0.09| 0.06| 0.22 | 0.05|
|             | KIR2DS4      |      |       |     |     |     |     |      |     |
| Activating (%) |     | 0.18 | 0.15  | 0.31| 0.27| 0.23| 0.23| 0.40 | 0.14|
| Inhibitory (%)  |     | 0.82 | 0.85  | 0.69| 0.73| 0.77| 0.77| 0.60 | 0.86|

Mean KIR-HLA interactions per individual (min-max)

|                | CTBA | BrJAP | KIV | KRC | GND | GKW | ACHE | GRC |
|----------------|------|-------|-----|-----|-----|-----|------|-----|
| (min-max)      | (1-13)| (2-12)| (1-10)| (1-10)| (1-12)| (1-7)| (1-7)| (1-8) |