



High-Resolution Characterization of *KIR* Genes in a Large North American Cohort Reveals Novel Details of Structural and Sequence Diversity

Leonardo M. Amorim^{1†}, Danillo G. Augusto^{2†}, Neda Nemat-Gorgani³, Gonzalo Montero-Martin⁴, Wesley M. Marin², Hengameh Shams², Ravi Dandekar², Stacy Caillier², Peter Parham³, Marcelo A. Fernández-Viña⁴, Jorge R. Oksenberg², Paul J. Norman^{3,5} and Jill A. Hollenbach^{2*}

OPEN ACCESS

Edited by:

Jeanne E. L. Visentainer,
State University of Maringá, Brazil

Reviewed by:

Christiane Ayo,
Faculty of Medicine of São
José do Rio Preto, Brazil
Luís Cristóvão Porto,
Rio de Janeiro State University, Brazil

*Correspondence:

Jill A. Hollenbach
jill.hollenbach@ucsf.edu

[†]These authors have contributed
equally to this work and share
first authorship

Specialty section:

This article was submitted to
Molecular Innate Immunity,
a section of the journal
Frontiers in Immunology

Received: 01 March 2021

Accepted: 20 April 2021

Published: 07 May 2021

Citation:

Amorim LM, Augusto DG,
Nemat-Gorgani N, Montero-Martin G,
Marin WM, Shams H, Dandekar R,
Caillier S, Parham P,
Fernández-Viña MA, Oksenberg JR,
Norman PJ and Hollenbach JA (2021)
High-Resolution Characterization of
KIR Genes in a Large North American
Cohort Reveals Novel Details of
Structural and Sequence Diversity.
Front. Immunol. 12:674778.
doi: 10.3389/fimmu.2021.674778

¹ Programa de Pós-Graduação em Genética, Universidade Federal do Paraná, Curitiba, Brazil, ² Department of Neurology, University of California, San Francisco, CA, United States, ³ Department of Structural Biology, Stanford University, Palo Alto, CA, United States, ⁴ Histocompatibility & Immunogenetics Laboratory, Stanford Blood Center, Palo Alto, CA, United States, ⁵ Division of Biomedical Informatics and Personalized Medicine, University of Colorado, Denver, CO, United States

The *KIR* (killer-cell immunoglobulin-like receptor) region is characterized by structural variation and high sequence similarity among genes, imposing technical difficulties for analysis. We undertook the most comprehensive study to date of *KIR* genetic diversity in a large population sample, applying next-generation sequencing in 2,130 United States European-descendant individuals. Data were analyzed using our custom bioinformatics pipeline specifically designed to address technical obstacles in determining *KIR* genotypes. Precise gene copy number determination allowed us to identify a set of uncommon gene-content *KIR* haplotypes accounting for 5.2% of structural variation. In this cohort, *KIR2DL4* is the framework gene that most varies in copy number (6.5% of all individuals). We identified phased high-resolution alleles in large multi-locus insertions and also likely founder haplotypes from which they were deleted. Additionally, we observed 250 alleles at 5-digit resolution, of which 90 have frequencies $\geq 1\%$. We found sequence patterns that were consistent with the presence of novel alleles in 398 (18.7%) individuals and contextualized multiple orphan dbSNPs within the *KIR* complex. We also identified a novel *KIR2DL1* variant, Pro151Arg, and demonstrated by molecular dynamics that this substitution is predicted to affect interaction with HLA-C. No previous studies have fully explored the full range of structural and sequence variation of *KIR* as we present here. We demonstrate that pairing high-throughput sequencing with state-of-art computational tools in a large cohort permits exploration of all aspects of *KIR* variation including determination of population-level haplotype diversity, improving understanding of the *KIR* system, and providing an important reference for future studies.

Keywords: NK receptors, alleles, high resolution, NGS, sequencing, population

INTRODUCTION

A variety of membrane-bound receptors control the response of natural killer cells (NK) to infected or malignant cells (1, 2). The killer-cell immunoglobulin-like receptors (KIR) are the most polymorphic family of NK receptors and encoded by a gene family located at chromosomal region 19q13.4 (3, 4). The *KIR* genes exhibit extraordinary variation, both within populations and between them (5–7). Although *KIR* gene-content has been extensively studied in numerous populations worldwide (8–11), less is known about *KIR* allele diversity. The unusual structural variation of the *KIR* region, coupled with the numerous alleles for each *KIR* gene (12) and extensive sequence similarity within the *KIR* gene family are distinguishing characteristics of *KIR* variation. Further characterizing the *KIR* region are frequent duplications, large deletions, hybrid genes and recombinant alleles (13–17). Together, these obstacles have impeded high-resolution allelic characterization of all *KIR* genes in population studies, which has been accomplished in only a few studies (18–21).

Transduction of NK cell activating and inhibitory signals is achieved by a subset of human leukocyte antigen (HLA) class I molecules, which serve as KIR ligands (22–24). These two interacting molecule families evolve as a unique and integrated system (25–28), and combinations of KIR and HLA have been associated with numerous diseases (29–32), including autoimmune disorders (33–36), malignancies (37–39) and infections (40–43). Combinations of KIR and HLA class I also affect placentation and the success of reproduction (44–47). Therefore, high-resolution allelic analysis of *KIR* and *HLA* class I diversity across populations will be necessary to understand their evolution and lay a foundation for functional studies to determine disease mechanisms. To facilitate this progression, we have used our custom *KIR* genotyping and bioinformatics pipeline to interrogate *KIR* diversity in a sample of 2,130 US residents. These methods were explicitly designed to cope with the complexities of *KIR* alleles, gene and haplotypes. We describe the most comprehensive analysis to date of the *KIR* genes, exploring copy number variation, haplotype patterns, and novel variation not previously reported.

MATERIAL AND METHODS

Study Population

We analyzed a cohort of 2,130 unrelated healthy adult individuals previously described by Hollenbach et al., 2019 (48). All individuals self-identified as being of European descent and were resident of the United States.

KIR Genotyping

DNA samples were sequenced for all *KIR* genes, according to Norman and collaborators (49). After sequencing, raw fastq files were analyzed using our custom bioinformatics pipeline PING (Pushing Immunogenetics into the Next Generation) to obtain *KIR* gene content and allelic genotypes from next-generation

sequencing (NGS) data (49). We applied an updated version of the pipeline that precisely determines the copy number of each locus through multiple alignment and filtration steps, also accurately identifying *KIR* genotypes. The updated pipeline increased the accuracy of *KIR* genotype determination and is publicly available (50).

Haplotype Estimation

Gene-content haplotypes were identified manually, based on the precise copy number determination, the known linkage disequilibrium among *KIR* genes, and allelic information. Uncommon haplotypes were identified based on previous observations (13, 51, 52). Candidates for novel gene-content haplotypes were identified when paired with common haplotypes and observed in two or more individuals. After identifying gene-content haplotypes, we inferred the haplotypes of their *KIR* alleles using the expectation-maximization (EM) algorithm and the R package haplo.stats (<http://CRAN.R-project.org/package=haplo.stats>).

Linkage Disequilibrium (LD) Analysis

Allelic genotyping data were transformed into an Arlequin entry file (.arp) using GenA1Ex 6.5 (53). Gametic phase estimation using an EM algorithm and further pairwise linkage disequilibrium analysis were performed using Arlequin 3.5.2.2 (54).

Identification of Novel Alleles

We searched for the single nucleotide variants (SNV) in *KIR2DL1* and *KIR3DL1S1* that were identified by our software but not present in any allele listed at the Immuno Polymorphism Database (IPD)-KIR release 2.9.0 (12). For individuals carrying a candidate novel SNV in *KIR2DL1* or *KIR3DL1S1*, the respective genes were re-sequenced using the Sanger method (55) using previously described primers (25, 56).

Simulations of Molecular Dynamics

In silico, the *KIR2DL1* chain was isolated from the KIR-HLA complex (PDB ID: 1IM9) (57). To map the *rs200879366* variation on the *KIR2DL1* structure, the proline at position 151 was replaced by arginine using the Mutate plugin in Visual Molecular Dynamics (VMD) package (58). Both allotypes were solvated in separate simulation boxes using TIP3P solvent, and the ion concentration was adjusted to 150 KCl. Energy minimization was then performed on both systems for 150,000 steps. To mimic the anchorage of *KIR2DL1* to the lipid membrane, the atom of residue 200 was fixed in space. Conformational transitions of *KIR2DL1* allotypes were modeled using the NAMD software package (59) and the CHARMM36 forcefield (60) in NPT ensembles. Temperature and pressure were maintained at 310 K and 1 bar using the Langevin thermostat and Langevin piston Nose-Hoover, respectively. Periodic boundary condition in all directions and a timestep of 2fs were used. Simulations on both systems ran for 100ns. The angle between the D1 and D2 domains is the leading indicator of *KIR2DL1* conformational transition. It was obtained by aligning the corresponding atom selections and calculating

rotation and displacement at every timeframe using in-house tcl scripts. Structure visualizations were performed using VMD.

RESULTS

KIR Gene Copy Number Analysis Identified Numerous Deletions and Duplications Involving KIR3DP1, KIR2DL4, and KIR3DL1S1

We determined KIR gene copy number for all 13 KIR genes (KIR2DL1, KIR2DL23, KIR2DL4, KIR2DL5A, KIR2DL5B, KIR2DS1, KIR2DS2, KIR2DS3, KIR2DS4, KIR2DS5, KIR3DL1S1, KIR3DL2, KIR3DL3) and the two pseudogenes (KIR2DP1, and KIR3DP1). Carrier frequencies and gene frequencies, based on the direct counting of all copies for each KIR, are given in **Supplementary Table 1**.

The KIR A haplotype is defined by the presence of only one gene encoding a short-tailed activating receptor (KIR2DS4) and a fixed number of genes encoding inhibitory KIR. In contrast, KIR B haplotypes encode various combinations of activating and

inhibiting receptors (61). A total of 31.4% of individuals in the cohort of 2,130 Europeans are homozygous for the full-length KIR A haplotype ($cA01 \sim tA01/cA01 \sim tA01$, $f = 0.58$). The centromeric and telomeric portions of the KIR haplotype are flanked by the framework genes KIR3DL3-KIR3DP1 and KIR2DL4-KIR3DL2, respectively (62, 63). On analyzing these two regions separately, we observed that 94% of the centromeric diversity is explained by just three gene-content haplotypes ($cA01$, $cB02$, and $cB01$; **Figure 1A**), whereas $tA01$ and $tB01$ correspond to 93% of the telomeric haplotypes (**Figure 1B**). We also identified the presence of two novel centromeric haplotypes, and two telomeric haplotypes not previously described in Europeans. The novel $cB06$ haplotype differs from $cB01$ by lacking KIR2DP1, and $cA03$ differs from the more common $cA01$ by lacking KIR3DP1. Present in the telomeric region is $tA02$, which only differs from $tA01$ by lacking KIR2DS4. Of particular interest is a haplotype observed in two individuals that has only the KIR3DL2 framework gene in the telomeric region. The gene content and organization of 3.5% of the centromeric and 4.9% of the telomeric haplotypes could not be determined.

Of the framework genes, KIR2DL4 has the most copy number variation in our study. We observed 72 individuals (3.4%) carrying one copy of KIR2DL4 and 67 (3.1%) carrying three

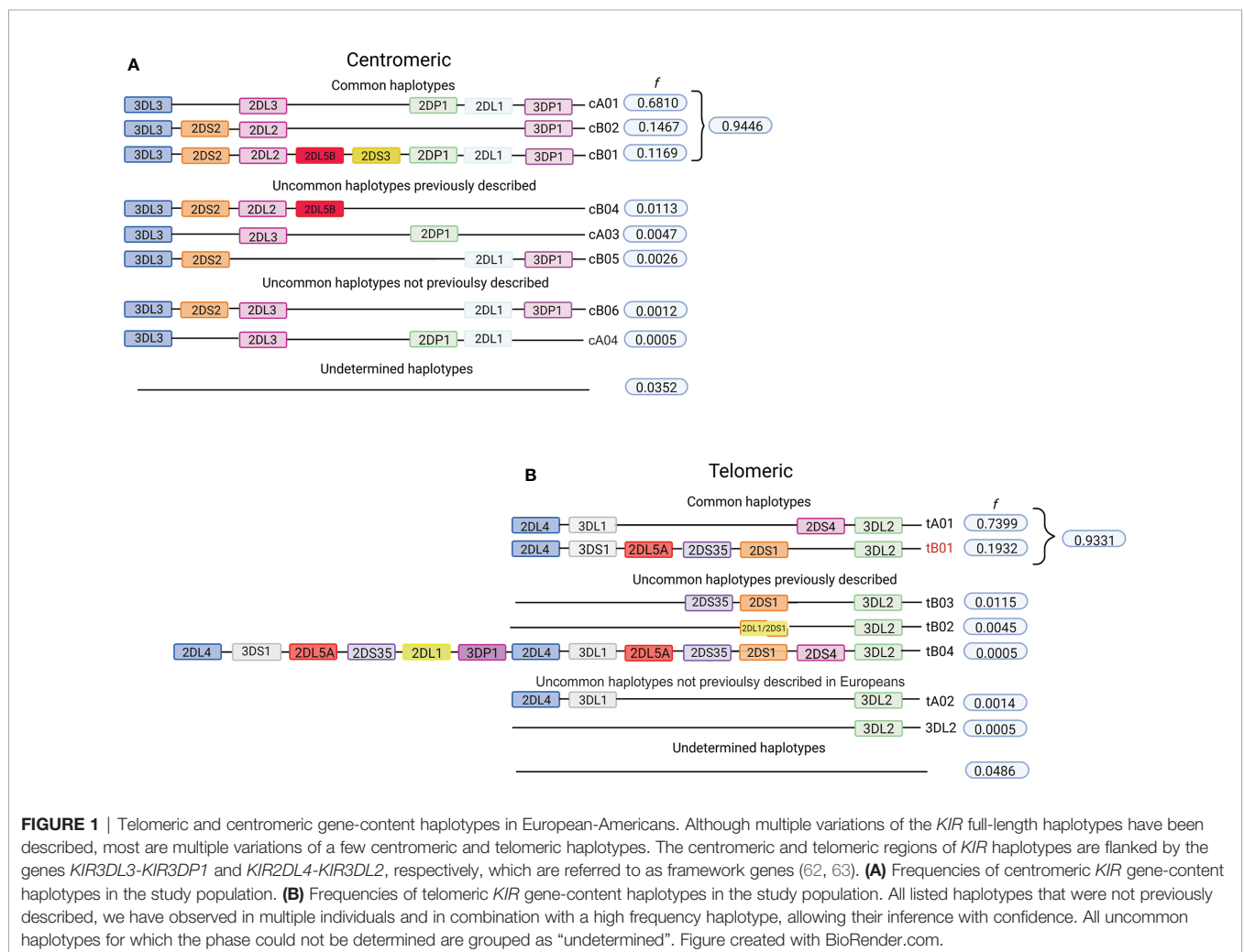
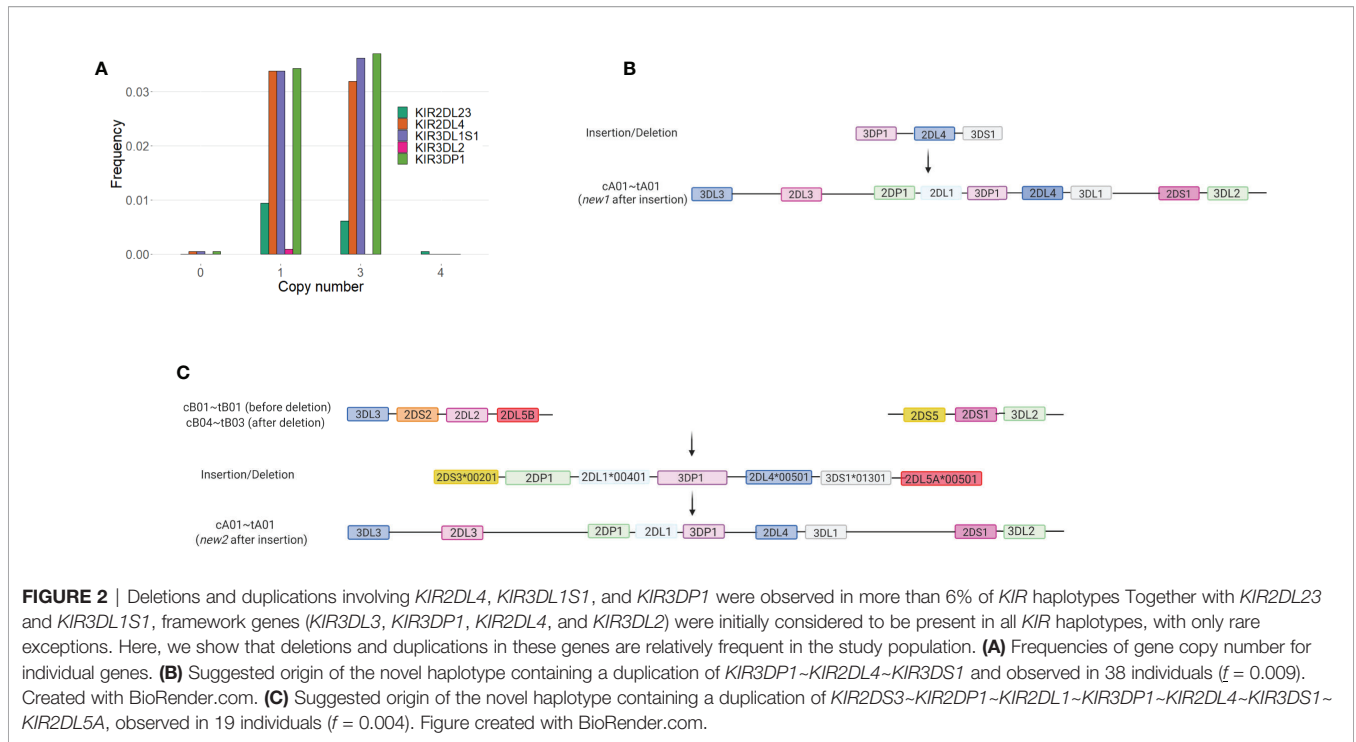


FIGURE 1 | Telomeric and centromeric gene-content haplotypes in European-Americans. Although multiple variations of the KIR full-length haplotypes have been described, most are multiple variations of a few centromeric and telomeric haplotypes. The centromeric and telomeric regions of KIR haplotypes are flanked by the genes KIR3DL3-KIR3DP1 and KIR2DL4-KIR3DL2, respectively, which are referred to as framework genes (62, 63). **(A)** Frequencies of centromeric KIR gene-content haplotypes in the study population. **(B)** Frequencies of telomeric KIR gene-content haplotypes in the study population. All listed haplotypes that were not previously described, we have observed in multiple individuals and in combination with a high frequency haplotype, allowing their inference with confidence. All uncommon haplotypes for which the phase could not be determined are grouped as “undetermined”. Figure created with BioRender.com.

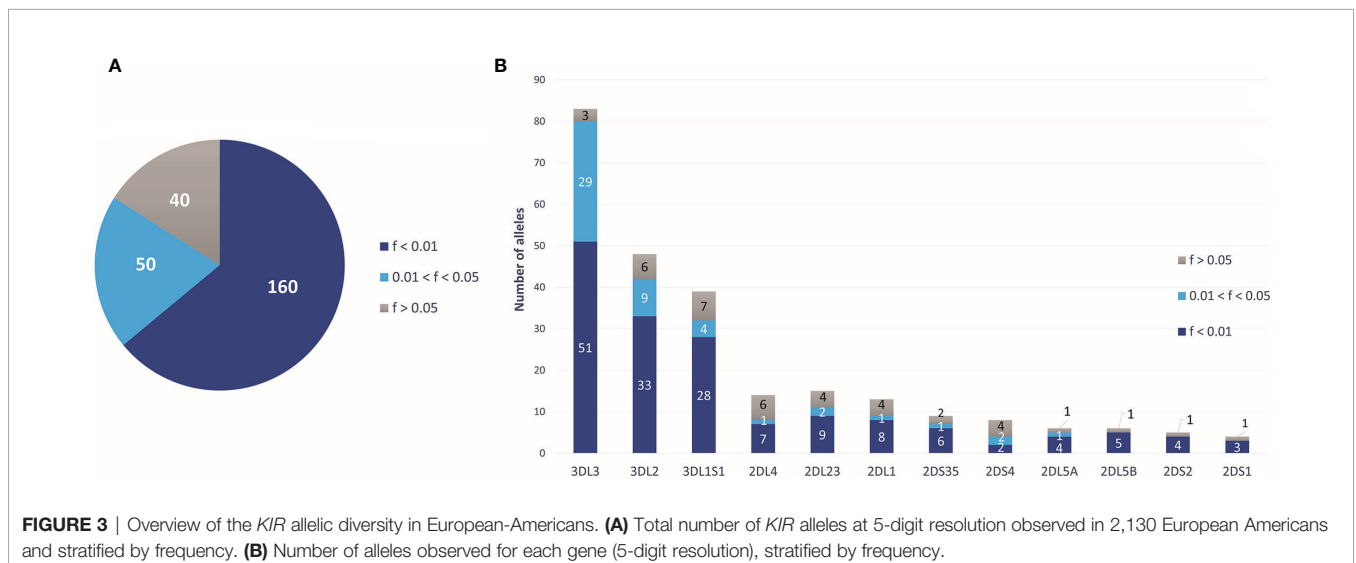


copies (**Figure 2A**). In this cohort, deletion of *KIR2DL4* is invariably accompanied by the deletion of *KIR3DL1S1*. *KIR3DP1* is deleted in 63 of the 72 (87.5%) haplotypes having the *KIR2DL4*~*KIR3DL1S1* deletion. *KIR3DL1S1* and *KIR3DP1* are duplicated in 63 out of 68 individuals carrying duplications of *KIR2DL4*. Insertion of *KIR3DP1*~*KIR2DL4*~*KIR3DS1* into a *tA01* haplotype created a novel haplotype, carried by 38 individuals ($f = 0.009$; **Figure 2B**). Insertion of the segment *KIR2DS3**00103~*KIR2DP1*~*KIR2DL1**00401~*KIR3DP1*~*KIR2DL4**00501~*KIR3DS1**01301~*KIR2DL5A**00501 also into *tA01* gave another novel haplotype observed in 19 individuals ($f = 0.004$; **Figure 2C**). In addition, we always observed the

centromeric *cB04* haplotype to be in the same gametic phase as *tB03*, and *cA03* always with *tB02*. Duplication of *KIR2DL2* present in *cB02* was observed in 7 individuals ($f = 0.002$)

Several KIR Haplotypes Are Marked by Specific KIR Alleles

All 13 *KIR* genes were genotyped to five-digit allele resolution in the study sample. We identified 250 *KIR* alleles, of which 90 (37.6%) have frequencies equal or greater than 0.01, and 40 (17%) have frequencies equal or greater than 0.05 (**Figure 3A**). *KIR3DL3* has the highest variety of alleles ($n = 83$), followed by *KIR3DL2* ($n = 48$) and *KIR3DL1S1* ($n = 39$) (**Figure 3B**).



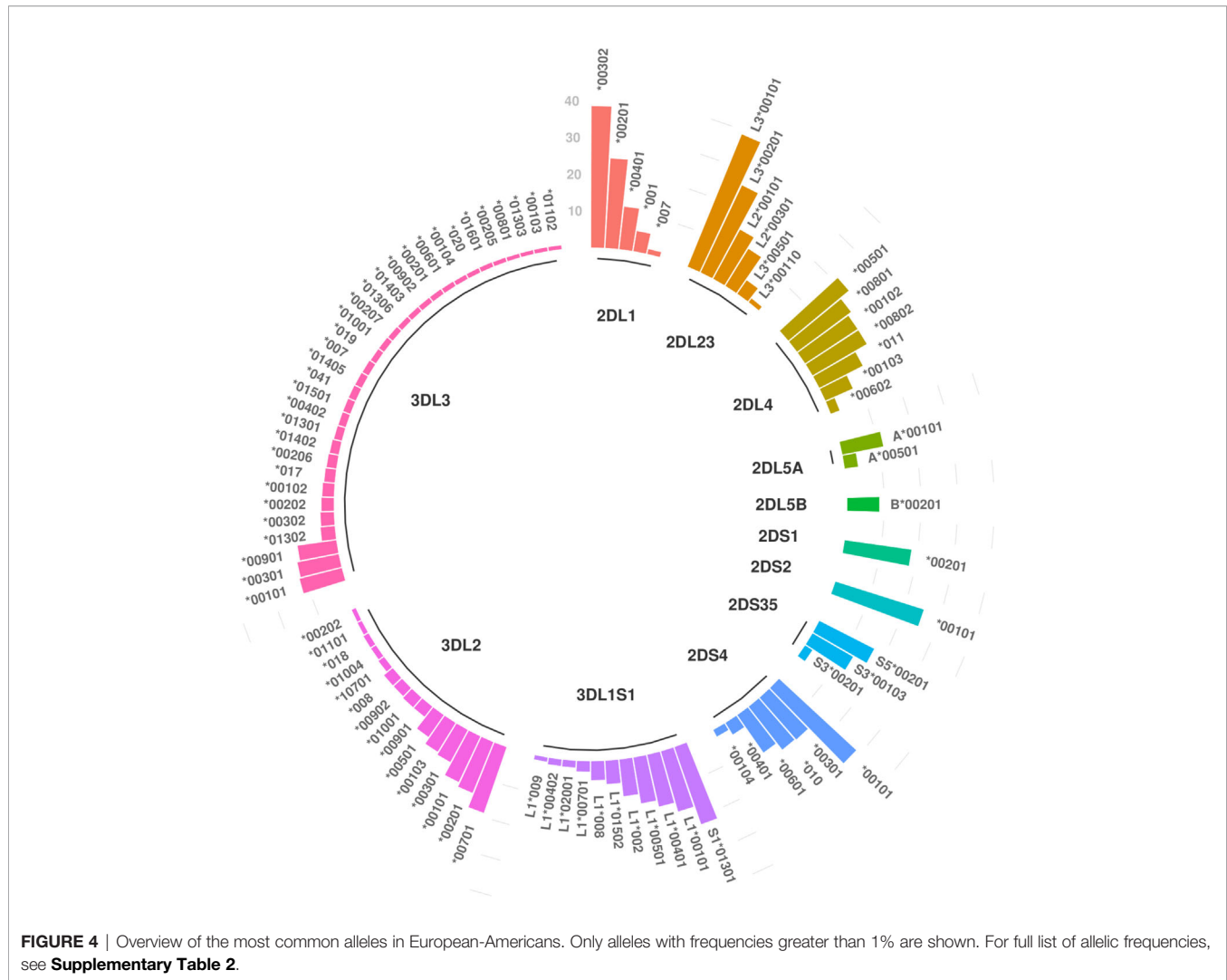
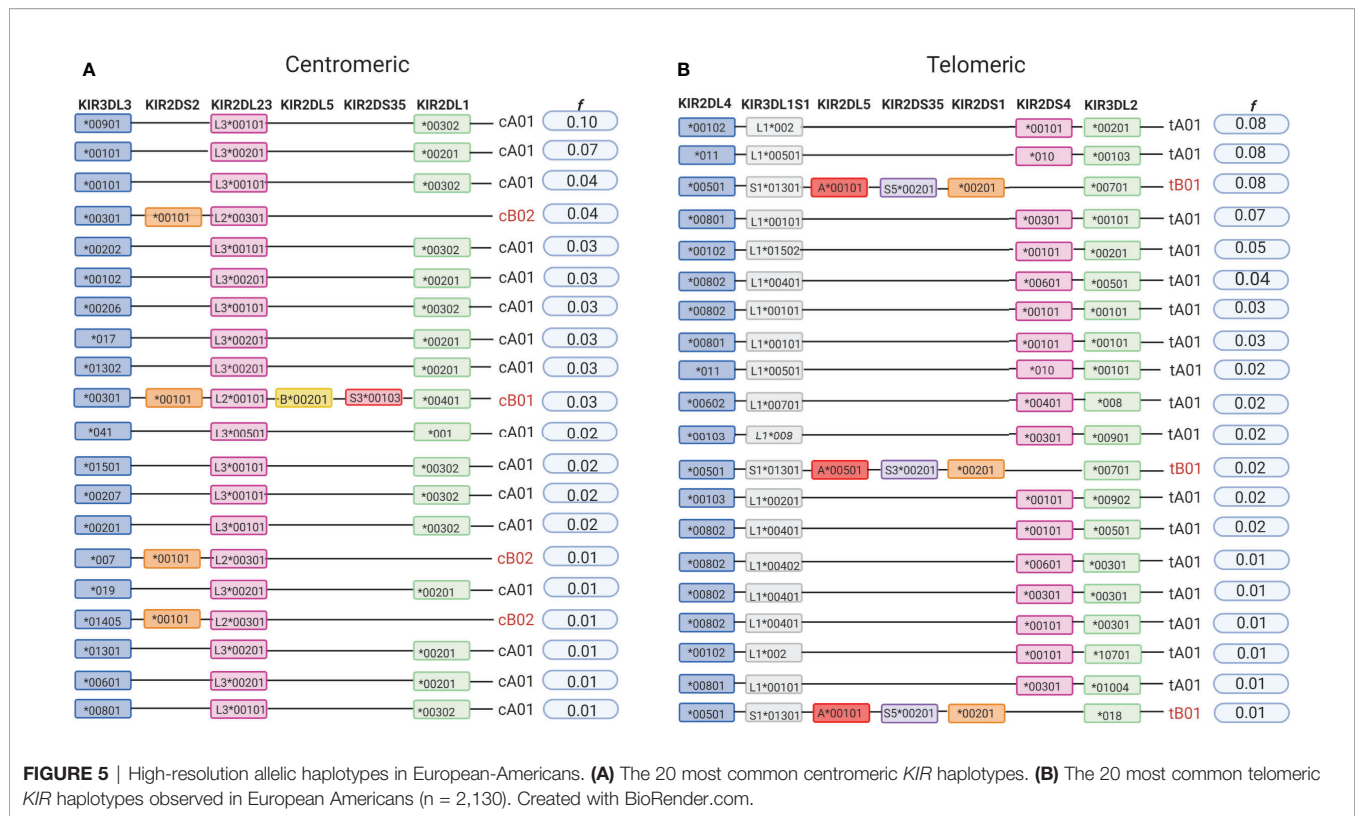


Figure 4 summarizes *KIR* allele diversity of the cohort and complete allele frequencies are given in **Supplementary Table 2**. Among the 20 most common centromeric haplotypes, 16 are *cA01* (**Figure 5A**). Similarly, 17 of the 20 most common telomeric haplotypes were *tA01* (**Figure 5B**).

This is the first study to describe high-resolution (5-digit) *KIR* haplotypes for all functional *KIR* genes in a large population sample. Analyzing this large number of individuals gave us sufficient power to fully explore the patterns of LD and identify alleles that are exclusively or predominately associated with specific haplotypes (**Figure 6**). For example, the alleles *KIR3DL2**00701 and *018 were observed solely in *tB01* haplotypes, whereas *KIR3DL2**00103, *00201, *00501, *00901, *00101, and *008 were observed only in *tA01*. Similarly, *KIR3DL3**00301 and *00402 are characteristic of *CenB*, whereas *00901 and *00101 are exclusive to *CenA* haplotypes. With few exceptions, *2DL4**00501 is the only *KIR2DL4* allele found in *tB01* haplotypes (99.4%), being in complete LD with *KIR3DS1**013. Additionally, a few low-frequency alleles are associated with specific, uncommon haplotypes; for example,

*KIR3DL2**034 ($f = 0.002$) is present only in *tB03*, and *KIR2DL5B**00801 ($f = 0.004$), is present only in *cB04* ($f = 0.01$). Furthermore, for the first time we describe multiple high-resolution allelic configurations of the full-length *cB04*-*tB03* haplotype (**Supplementary Table 3**).

As well as the haplotypic associations, we observed many instances of strong LD among specific sets of *KIR* alleles (**Figure 7**). In summary, *KIR2DL4**00501, *KIR3DS1**01301, *KIR2DS1**00201, *KIR2DL5A**00101, *KIR2DS5**00201, and *KIR3DL2**00701 are frequently observed together. Many other *KIR2DL4* alleles are in strong LD with specific *KIR3DL1* alleles. Specific examples are *KIR2DL4**00801 with *KIR3DL1**00101 ($D' = 0.99$, $r^2 = 0.85$); *KIR2DL4**011 with *KIR3DL1**00501 ($D' = 0.95$, $r^2 = 0.85$); *KIR2DL4**00802 with *KIR3DL1**00401 ($D' = 0.95$, $r^2 = 0.73$); *KIR2DL4**00602 with *KIR3DL1**00701 ($D' = 0.91$, $r^2 = 0.68$); and *KIR2DL4**00103 with *KIR3DL1**008 ($D' = 1$, $r^2 = 0.6$). In the centromeric region, *KIR2DL5B**00201 and *KIR2DS3**00103 are always observed together. *KIR2DL1**00401 is in strong LD with *KIR2DS3**00103 ($D' = 0.98$ and $r^2 = 0.83$) and *KIR2DL5B**00201 ($D' = 0.95$ and $r^2 = 0.82$), whereas



*KIR2DL3*00201* is associated with *KIR2DL1*00201* ($D' = 0.97$, $r^2 = 0.89$). A full list of LD values for pairs of *KIR* alleles is given in **Supplementary Table 4**.

Numerous Novel *KIR* Variants Are Present in the Cohort of European Americans

In this analysis of 2,130 individuals, we identified 398 individuals (18.7%) carrying at least one *KIR* recombinant allele that do not correspond to any sequences deposited in the *KIR* database. We define as recombinant allele those that are characterized by different phasing combinations of previously known variable sites. These observations are likely to represent the presence of new alleles that were not present or not detected in previous studies of *KIR* variation. *KIR3DL1S1* accounts for 33% of the observations of candidate new alleles, corresponding to a total allelic frequency of 0.04 at this locus. A large proportion of individuals carrying possible novel alleles were also observed for other *KIR2D* and *KIR3D* genes (**Table 1**).

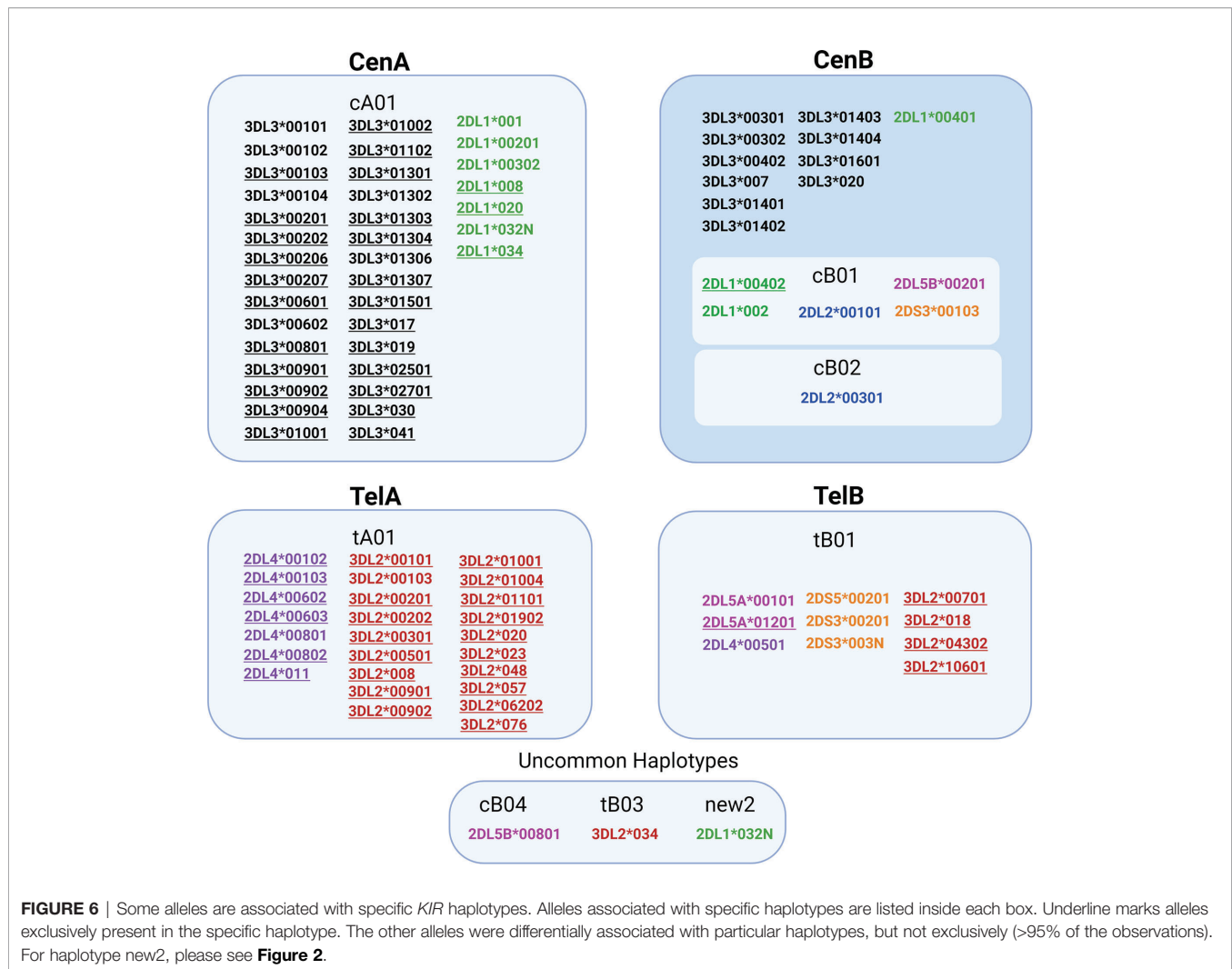
In addition to these candidate novel *KIR* recombinant alleles, we also used our software to identify possible novel SNVs. In some cases, these SNVs may have been reported in the dbSNP database (64), but they were not associated with any *KIR* allele sequence deposited in the IPD-*KIR* database release 2.9.0 (12). Therefore, in the context of *KIR*, these SNVs would be contributing to novel alleles that each differ from a known *KIR* allele by a single nucleotide substitution. To confirm the sequences of novel variants, we used the Sanger method to re-sequence individuals carrying any possible novel SNV in *KIR2DL1* and *KIR3DL1S1*. While confirmation of all novel

variants was out of scope for the current project, we selected these loci as exemplars for this work due to substantial previous work examining their structure and function (25, 56, 61, 65–71), including the availability of crystal for molecular modeling structures (57, 72–76). In future work we will continue to explore novel variants that were detected at other loci during this study.

For *KIR2DL1*, 8 of 30 variants were confirmed by Sanger sequencing, while 10 of 32 variants were similarly confirmed for *KIR3DL1S1*. Most of the SNVs in *KIR3DL1S1* were observed in only a single individual, except for two synonymous variants, *rs754894112* and *rs1462310393* (**Table 2**). In contrast, the majority of confirmed novel variants in *KIR2DL1* were observed in several individuals (**Table 3**). Interestingly, 14 out of the 18 confirmed variants in these two genes were non-synonymous substitutions, with functional effects ranging from conservative to radical according to the Grantham scale of physicochemical distances between amino acids (77).

Simulation of Molecular Dynamics Predicts That Dimorphism in Codon 151 of *KIR2DL1* Affects Binding to HLA Class I

To explore the functional differences of *KIR2DL1* alleles that differ by a single nucleotide, we simulated and compared the molecular dynamics simulations of *KIR2DL1* allotypes that differ by the non-conservative substitution of proline to arginine at position 151. Underlying this difference is *rs200879366**C>G. Different conformations were sampled during the simulation trajectory, so that each time step features an individual



conformer. The angle change mediated the transition between free and HLA-bound states of KIR2DL1 between the Ig domains (D1 and D2) that eventually affected the HLA binding region (**Figure 8A**). Within 100ns of simulation, the angle between the D1 and D2 domains was decreased by 10° in the wildtype (Pro) but was not perturbed in the mutant (Arg) (**Figure 8B**). The conformational transition appears to be mediated by a network of interactions spanning from Met44 to Arg 151. In the wildtype, Met44 is released from Pro185, allowing the angle between D1 and D2 to decrease (**Figure 8C**). By contrast, the angle is increased in the mutant. This enables Arg151 to form a salt bridge with Asp135, leading to an interaction between Met136 and Pro185 (**Figure 8D**) that allows Pro185 to more strongly associate with Met44. This set of localized rearrangements in KIR2DL1 is likely essential for its stable binding to HLA class I.

DISCUSSION

The general configuration of *KIR* gene-content haplotypes was first described two decades ago, when it was observed that four

framework genes separate two distinct sub-clusters of genes (7, 24, 62, 78, 79). *KIR3DL3* and *KIR3DP1* were seen to delimit the centromeric region, while *KIR2DL4* and *KIR3DL2* delimit the telomeric region of the *KIR* gene family. The other two genes that are present in most *KIR* haplotypes are *KIR2DL23* and *KIR3DL1S1* (6). Although deletions and duplications of these genes have been previously reported (15–17, 80–82), technical limitations have precluded direct copy number determination of all *KIR* in large-scale population studies. We show here that large structural deletions and duplications involving the framework genes are relatively frequent in European-descendant individuals. For instance, more than 6% of individuals carry a deletion or duplication of *KIR2DL4*. Similar to haplotype variants described in the literature (13, 51, 52), all gene-content *KIR* haplotypes lacking *KIR2DL4* also lacked *KIR3DL1S1* (*tB02* and *tB03*). These haplotypes have been described for other European-Americans (52), while an extensive study of Europeans from Germany did not seek to analyze novel gene-content haplotypes (83). Observation of *KIR2DL4*~*KIR3DL1S1* deletions at high-frequencies in Africans (18, 25) raises the possibility that these variant *KIR* haplotypes originated prior to the modern human

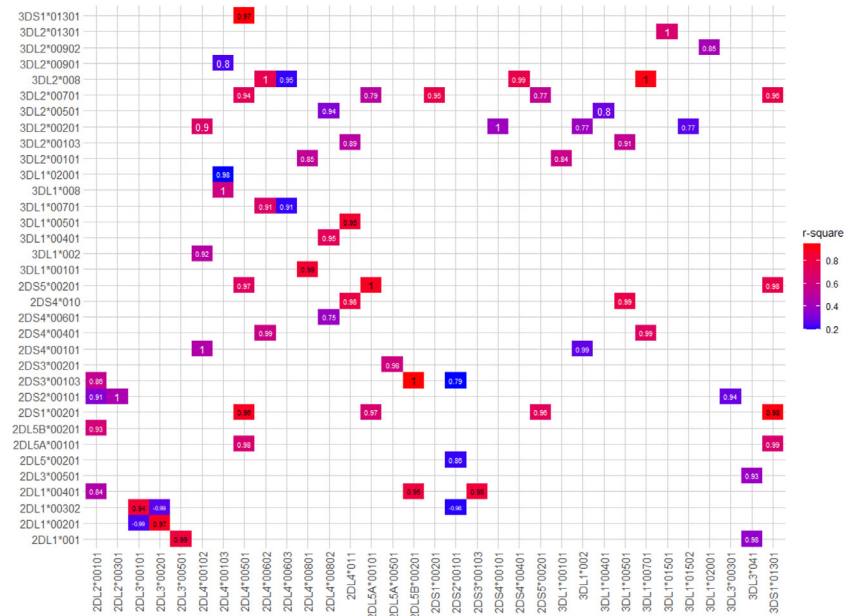


FIGURE 7 | Strong linkage disequilibrium (LD) between *KIR* alleles. Boxes represent the pairs of *KIR* alleles with the strongest LD in the present study. *D'* values are written inside of each box and color scale represent r^2 values. Only pairs with $r^2 > 0.2$, $D' > 0.7$ were shown. The p -value of all pairs was $< 10^{-5}$.

migration out of Africa, and that they might also be present in most worldwide populations.

Interestingly, haplotypes carrying *KIR2DL4* duplications also have duplications of *KIR3DL1S1* and *KIR3DP1*. Based on our observations, including the fact that *cB04~tB03* are always in phase, we propose that a single deletion of *KIR2DS3*00103~KIR2DP1~KIR2DL1*00401~KIR3DP1~KIR2DL4*00501~KIR3DS1*01301~KIR2DL5A*00501* from the haplotype *cB01~tB01* originated the *cB04~tB03* haplotype. This seven-locus fragment was possibly inserted into *cA01~tA01*, which would explain the novel full-length haplotype that we identified in multiple individuals. The large cohort that we analyzed allowed us to phase the alleles of the seven-locus indel at high-resolution, providing a unique opportunity to infer the origin of these haplotypic variants.

Previous studies described *KIR* haplotypes at lower genotyping resolution and in smaller sample sizes. For example, Vierra-Green et al. (52) described *KIR* haplotypes at 3-digit resolution in 506

Euro-Americans while Hou et al. (20) analyzed most *KIR* genes at higher resolution but in a small cohort. Here, we present the first study to show 5-digit allelic haplotypes of all *KIR* genes for a large sample of the European-descent U.S population. Notably, the most common centromeric haplotype in our study cohort, *KIR3DL3*00901~KIR2DL3*00101~KIR2DL1*00302* ($f = 0.10$), is also the most common in four African populations (Datoonga, $f = 0.11$; Baka, $f = 0.15$; Dogon, $f = 0.18$; and Fulani, $f = 0.14$) (18). This haplotype is likely the same reported as the most frequent ($f = 0.23$) in a smaller European American cohort that was not analyzed for all *KIR* genes at high resolution (20). Interestingly, two of the low-frequency telomeric haplotypes present in our sample (*tA02* and *KIR3DL2*) have not been reported in other European populations, but were observed in African populations from Mali (*3DL2*, $f = 0.01$), Democratic Republic of Congo (*tA02*, $f = 0.08$) and Tanzania (*3DL2*, $f = 0.01$; *tA02*, $f = 0.01$) (18). A limitation of determining haplotypes without family segregation studies or confirmation by long-range sequencing is the impossibility to identify unknown haplotype structures or to precisely infer those haplotypes observed in lower frequencies. For this reason, we were not able to confidently identify the less common haplotypes, therefore, presenting data only for the most common ones. However, our large population sample coupled with the curated high-quality data allow us to identify the haplotypic diversity that represent most of the *KIR* diversity in Europeans.

Our well-powered analysis of LD across the *KIR* region shows that some alleles are clearly associated with specific structural haplotypes. Because *KIR2DL4*00501* and *KIR3DL2*00701* are present in *tB01* and are associated with other *tB01*-associated alleles, such as *KIR3DS1*01301*, it was possible to verify that the

TABLE 1 | Large proportion of individuals carrying possible novel alleles.

Locus	n	f
<i>KIR3DL1S1</i>	154	7.23%
<i>KIR3DL2</i>	57	2.68%
<i>KIR2DL5</i>	56	2.63%
<i>KIR2DL4</i>	53	2.49%
<i>KIR3DL3</i>	48	2.25%
<i>KIR2DL1</i>	43	2.02%
<i>KIR2DL23</i>	31	1.46%
<i>KIR2DS3</i>	24	1.13%
<i>KIR2DS5</i>	10	0.47%
<i>KIR2DS4</i>	5	0.23%

n, absolute number individuals carrying potential at least one possible novel allele; *f*, relative frequency of individuals carrying possible novel alleles.

TABLE 2 | Ten confirmed novel single nucleotide variants in *KIR3DL1S1*.

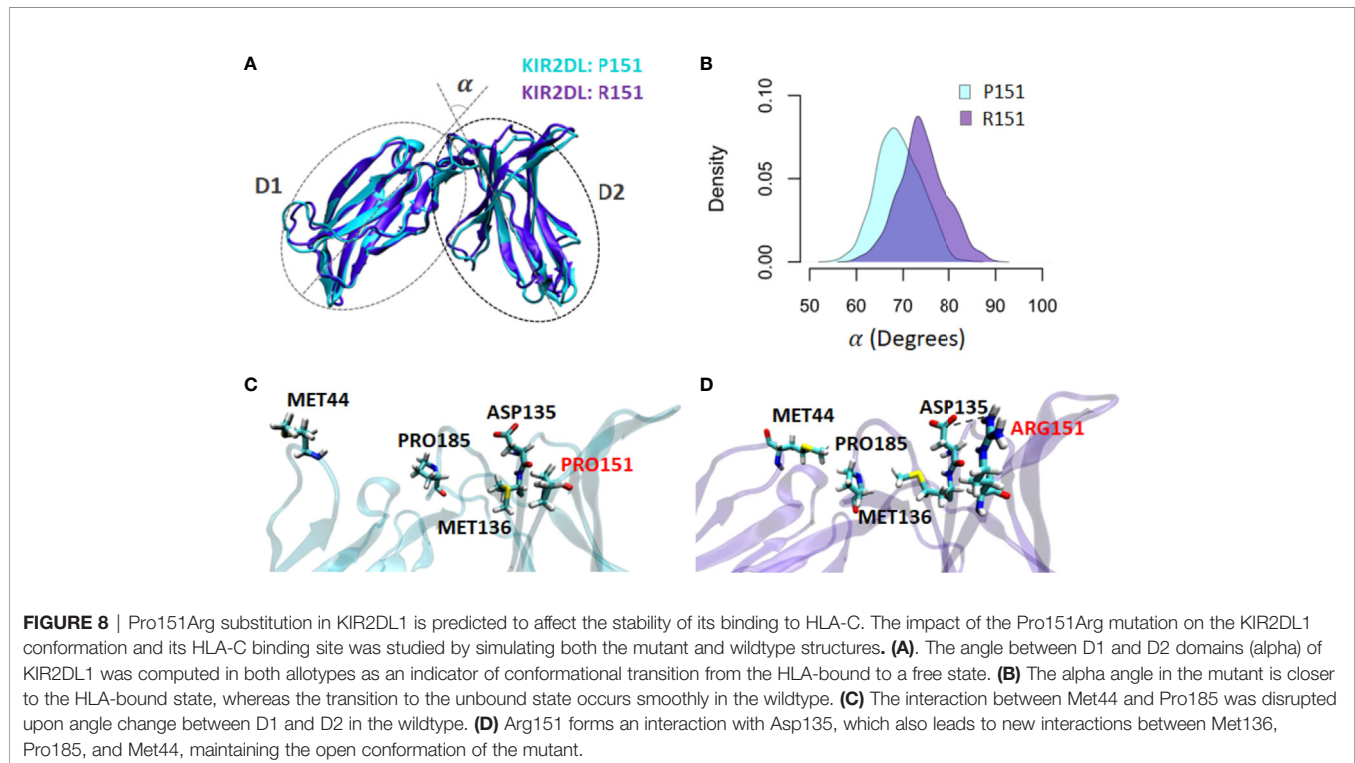
rsID	n	Exon	Change	Mature protein	Amino acid	Grantham
rs771871523	1	3	G>T	12	Ala>Ser	99
rs1272096635	1	4	C>A	107	Ala>Asp	126
rs769147743	1	4	T>C	127	Ile>Thr	89
rs200893904	1	4	C>T	171	Thr>Ile	89
rs1182774591	1	4	G>C	187	Ala>Pro	27
new SNP	1	5	G>C	249	Arg>Pro	103
rs754894112	12	5	C>T	266	synonymous	
rs1462310393	2	5	C>A	275	synonymous	
new SNP	1	7	C>T	319	His>Tyr	83
rs984592565	1	9	G>A	392	Arg>His	29

These variants have been reported and assigned rsID in the dbSNP database (64), but not present in any KIR allele deposited in the IPD-KIR database release 2.9.0 (12), therefore representing novel KIR alleles differing from others by only one nucleotide position. New SNP represent variants not previously assigned rsID in the dbSNP database.

TABLE 3 | Eight confirmed novel single nucleotide variants in *KIR2DL1*.

rsID	n	Exon	Change	Mature protein	Amino acid	Grantham
rs201225013	7	1	A>G	signal peptide	Met>Val	21
rs148427642	4	1	T>G	signal peptide	Cys>Gly	159
rs749640662	13	5	G>A	120	Ala>Thr	58
rs200879366	8	5	C>G	151	Pro>Arg	103
rs749653872	3	5	C>T	157	synonymous	
rs570412759	16	7	G>A	245	Arg>Hist	29
rs201527316	1	9	G>A	296	Arg>Hist	29
rs778821930	2	9	C>A	309	synonymous	

These variants have been reported and assigned rsID in the dbSNP database (64), but not present in any KIR allele deposited in the IPD-KIR database release 2.9.0 (12), therefore representing novel KIR alleles differing from others by only one nucleotide position.



insertion of fragments containing *KIR3DS1* occurred on the *cA01~tA01* haplotype. In some cases, specific alleles also associate with uncommon haplotypes, which may be used as markers for these unique haplotypes. These examples highlight

how our detailed LD information for high-resolution *KIR* sequencing constitutes a significant resource, yielding valuable information that will facilitate the comprehension and identification of *KIR* haplotypes in future studies.

With the development of robust high-resolution *KIR* typing methods, the discovery of novel *KIR* variants has become more achievable. However, short read misalignment is a major confounding factor in discovering novel variants, particularly those differing from known variants by only one nucleotide. Because the identification of possible novel SNVs is overall not likely in comparison to the possibility of being artifacts due to misalignment of highly similar sequence reads, our pipeline is initially set to exclude them from the primary analysis but flag them for further detailed inspection. This is a limitation intrinsic to short-sequence data analysis in the context of the *KIR* sequence homology, and rare novel variants might be missed. Aiming for an overview of the extent of this methodological limitation, we applied Sanger method to re-sequence all possible novel variants in two highly polymorphic genes, regardless of their frequencies. Although we confirmed approximately one-third of the possible novel SNVs for *KIR3DL1S1* and *KIR2DL1*, most of these variants were observed in only one or a few individuals. In other words, even though our pipeline can possibly miss some of the novel variation, the new SNVs do not represent significant overall distortions in our dataset. In fact, our method has been proven to show an overall high performance for determining accurate genotypes (median 96.5%), and only 1% to 3% of unresolved genotypes (50).

Low frequent variants may, however, be of particular interest especially if they may cause impact on the receptors' function. It is remarkable that most confirmed novel variants cause moderate to radical non-synonymous substitutions, ultimately leading to functional protein variation. To provide new insights to the functional significance of describing novel variants, we focused on variant *rs200879366**G, which was previously reported with a frequency of 0.01 in the Finnish population (84) but had not previously been associated with a specific *KIR* allele. Although residue 151 is in the D2 domain it does not make direct contact with the HLA-C ligand. Nevertheless, polymorphism at the neighboring residue, position 154, has been implicated in differential avidity for the HLA-C ligand (85). Our prediction shows that Asp135, which is directly engaged in HLA binding, forms a bond with Arg151, allowing us to speculate that *rs200879366**G may result in reduced binding to HLA. This example demonstrates the likelihood that many other functional variants will be identified as we interrogate *KIR* allelic diversity in worldwide populations.

According to the recently updated IPD-KIR database (Release 2.10.0, 16 December 2020), the variant *rs200879366**G marks two unconfirmed alleles, *KIR2DL1**044 and *KIR2DL1**046, corroborating our findings. These unconfirmed allele sequences were freshly submitted by the same group that genotyped *KIR* in over a million European samples from the DKMS donor registry (86). Although a remarkable effort in its scale and importance to the field, that study targeted specific exons of each *KIR*, resulting in a 3-digit resolution genotyping with substantial ambiguities. In contrast, our study sought to analyze all *KIR* exons and introns of each gene (5-digit resolution). While smaller than the DKMS study by orders of magnitude, our study is nevertheless the largest sample to-date to comprehensively

analyze all aspects of *KIR* variation at this resolution, including copy number, allele-haplotype associations, pairwise LD, and functional consequences of novel variation.

For decades, most *KIR* studies in populations were limited to analyzing the presence and absence of genes (5, 7, 9, 62, 87–89). The study of *KIR* gene content laid the basis of the field and suggested that *KIR* diversity and plasticity were and may still be ahead of our technical capabilities. Here, we aimed to set new ground for exploring *KIR* diversity by providing the first large-scale study to deeply analyze copy number variation and high-resolution allelic variation of all genes in a large population sample from the United States. Our results show a large proportion of multi-locus deletions and duplications of genes that were until recently considered rare, in addition to unusual gene-content haplotypes and a high frequency of novel alleles. We argue that as we continue to interrogate *KIR* at high-resolution, we will continue to uncover more layers of this region's complexity, discovering frequent novel variants with functional relevance that have been previously missed due to technical limitations.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**. Further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

JH, PN, and DA designed the study. NN-G, GM-M, and DA performed NGS and Sanger sequencing. LA and DA analyzed the data. WM and RD contributed with bioinformatics analysis. HS performed molecular dynamics simulations. SC managed and organized the samples. PP, MF-V, JO, PN, and JH contributed with samples and/or reagents. LA, DA, and JH drafted the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by grants U19NS095774 and U01 AI090905 from the U.S. National Institutes of Health (NIH). The UCSF DNA biorepository is supported by Si-2001-35701 from the National Multiple Sclerosis Society.

ACKNOWLEDGMENTS

LA acknowledges the scholarship from the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001 and the Doctorate Dissertation Research Award scholarship from Fulbright Program.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2021.674778/full#supplementary-material>

Supplementary Table 1 | KIR gene-content frequencies in European-Americans.

Supplementary Table 2 | High-resolution allelic frequencies of the KIR genes in European-Americans.

Supplementary Table 3 | Multiple configurations of the haplotype *cB04~tB03* in European-Americans.

Supplementary Table 4 | Pairwise linkage disequilibrium between KIR alleles. Only pairs with $r^2 > 0.2$ and $D' > 0.7$ are shown.

REFERENCES

- Martinet L, Smyth MJ. Balancing Natural Killer Cell Activation Through Paired Receptors. *Nat Rev Immunol* (2015) 15:243–54. doi: 10.1038/nri3799
- Lanier LL. NK Cell Receptors. *Annu Rev Immunol* (1998) 16:359–93. doi: 10.1146/annurev.immunol.16.1.359
- Wende H, Colonna M, Ziegler A, Volz A. Organization of the Leukocyte Receptor Cluster (LRC) on Human Chromosome 19q13.4. *Mamm Genome* (1999) 10:154–60. doi: 10.1007/s003359900961
- Wilson MJ, Torkar M, Trowsdale J. Genomic Organization of a Human Killer Cell Inhibitory Receptor Gene. *Tissue Antigens* (1997) 49:574–9. doi: 10.1111/j.1399-0039.1997.tb02804.x
- Norman PJ, Chandanayingyong D, Vaughan RW, Verity DH, Stephens HAF. Distribution of Natural Killer Cell Immunoglobulin-Like Receptor Sequences in Three Ethnic Groups. *Immunogenetics* (2001) 52:195–205. doi: 10.1007/s002510000281
- Uhrberg M, Valiante NM, Shum BP, Shilling HG, Lienert-Weidenbach K, Corliss B, et al. Human Diversity in Killer Cell Inhibitory Receptor Genes. *Immunity* (1997) 7:753–63. doi: 10.1016/S1074-7613(00)80394-5
- Hsu KC, Liu X-R, Selvakumar A, Mickelson E, O'Reilly RJ, Dupont B. Killer Ig-Like Receptor Haplotype Analysis by Gene Content: Evidence for Genomic Diversity With a Minimum of Six Basic Framework Haplotypes, Each With Multiple Subsets. *J Immunol* (2002) 169:5118–29. doi: 10.4049/jimmunol.169.9.5118
- Gonzalez-Galarza FF, McCabe A, Santos EJMD, Jones J, Takeshita L, Ortega-Rivera ND, et al. Allele Frequency Net Database (AFND) 2020 Update: Gold-Standard Data Classification, Open Access Genotype Data and New Query Tools. *Nucleic Acids Res* (2020) 48:D783–8. doi: 10.1093/nar/gkz1029
- Hollenbach JA, Nosedal I, Ladner MB, Single RM, Trachtenberg EA. Killer Cell Immunoglobulin-Like Receptor (KIR) Gene Content Variation in the HGDP-CEPH Populations. *Immunogenetics* (2012) 64:719–37. doi: 10.1007/s00251-012-0629-x
- Single RM, Martin MP, Gao X, Meyer D, Yeager M, Kidd JR, et al. Global Diversity and Evidence for Coevolution of KIR and HLA. *Nat Genet* (2007) 39:1114–9. doi: 10.1038/ng2077
- Misra MK, Augusto DG, Martin GM, Nemat-Gorgani N, Sauter J, Hofmann JA, et al. Report From the Killer-cell Immunoglobulin-Like Receptors (KIR) Component of the 17th International HLA and Immunogenetics Workshop. *Hum Immunol* (2018) 79:825–33. doi: 10.1016/j.humimm.2018.10.003
- Robinson J, Barker DJ, Georgiou X, Cooper MA, Flicke P, Marsh SG. Ipd-IMGT/HLA Database. *Nucleic Acids Res* (2020) 48:D948–55. doi: 10.1093/nar/gkz950
- Roe D, Vierra-Green C, Pyo C-W, Eng K, Hall R, Kuang R, et al. Revealing Complete Complex KIR Haplotypes Phased by Long-Read Sequencing Technology. *Genes Immun* (2017) 18:127–34. doi: 10.1038/gene.2017.10
- Traherne JA, Martin M, Ward R, Ohashi M, Pellett F, Gladman D, et al. Mechanisms of Copy Number Variation and Hybrid Gene Formation in the KIR Immune Gene Complex. *Hum Mol Genet* (2010) 19:737–51. doi: 10.1093/hmg/ddp538
- Martin MP, Bashirova A, Traherne J, Trowsdale J, Carrington M. Cutting Edge: Expansion of the KIR Locus by Unequal Crossing Over. *J Immunol* (2003) 171:2192. doi: 10.4049/jimmunol.171.5.2192
- Norman PJ, Abi-Rached L, Gendzekhadze K, Hammond JA, Moesta AK, Sharma D, et al. Meiotic Recombination Generates Rich Diversity in NK Cell Receptor Genes, Alleles, and Haplotypes. *Genome Res* (2009) 19:757–69. doi: 10.1101/gr.085738.108
- Pyo C-W, Wang R, Vu Q, Cereb N, Yang SY, Duh F-M, et al. Recombinant Structures Expand and Contract Inter and Intragenic Diversification at the KIR Locus. *BMC Genomics* (2013) 14:89. doi: 10.1186/1471-2164-14-89
- Nemat-Gorgani N, Guethlein LA, Henn BM, Norberg SJ, Chiaroni J, Sikora M, et al. Diversity of KIR, HLA Class I, and Their Interactions in Seven Populations of Sub-Saharan Africans. *J Immunol* (2019) 202:2636–47. doi: 10.4049/jimmunol.1801586
- Alicata C, Ashouri E, Nemat-Gorgani N, Guethlein LA, Marin WM, Tao S, et al. Kir Variation in Iranians Combines High Haplotype and Allotype Diversity With an Abundance of Functional Inhibitory Receptors. *Front Immunol* (2020) 11:556. doi: 10.3389/fimmu.2020.00556
- Hou L, Chen M, Ng J, Hurley CK. Conserved KIR Allele-Level Haplotypes are Altered by Microvariation in Individuals With European Ancestry. *Genes Immun* (2012) 13:47–58. doi: 10.1038/gene.2011.52
- Tao S, Kichula KM, Harrison GF, Farias TDJ, Palmer WH, Leaton LA, et al. The Combinatorial Diversity of KIR and HLA Class I Allotypes in Peninsular Malaysia. *Immunology* (2020) 162:389–404. doi: 10.1111/imm.13289
- Long EO, Rajagopalan S. HLA Class I Recognition by Killer Cell Ig-like Receptors. *Semin Immunol* (2000) 12:101–8. doi: 10.1006/smim.2000.0212
- Norman PJ, Parham P. Complex Interactions: The Immunogenetics of Human Leukocyte Antigen and Killer Cell Immunoglobulin-Like Receptors. *Semin Hematol* (2005) 42:65–75. doi: 10.1053/j.seminhematol.2005.01.007
- Trowsdale J, Barten R, Haude A, Stewart CA, Beck S, Wilson MJ. The Genomic Context of Natural Killer Receptor Extended Gene Families: Genomics of NK Receptor Gene Families. *Immunol Rev* (2001) 181:20–38. doi: 10.1034/j.1600-065X.2001.1810102.x
- Norman PJ, Hollenbach JA, Nemat-Gorgani N, Guethlein LA, Hilton HG, Pando MJ, 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 Genet* (2013) 9:e1003938. doi: 10.1371/journal.pgen.1003938
- Augusto DG, Petzl-Erler ML. KIR and HLA Under Pressure: Evidences of Coevolution Across Worldwide Populations. *Hum Genet* (2015) 134:929–40. doi: 10.1007/s00439-015-1579-9
- Gendzekhadze K, Norman PJ, Abi-Rached L, Graef T, Moesta AK, Layrisse Z, et al. Co-Evolution of KIR2DL3 With HLA-C in a Human Population Retaining Minimal Essential Diversity of KIR and HLA Class I Ligands. *PNAS* (2009) 106:18692–7. doi: 10.1073/pnas.0906051106
- Moffett A, Colucci F. Co-Evolution of NK Receptors and HLA Ligands in Humans is Driven by Reproduction. *Immunol Rev* (2015) 267:283–97. doi: 10.1111/imr.12323
- Kulkarni S, Martin MP, Carrington M. The Yin and Yang of HLA and KIR in Human Disease. *Semin Immunol* (2008) 20:343–52. doi: 10.1016/j.smim.2008.06.003
- Williams AP, Bateman AR, Khakoo SI. Hanging in the Balance. KIR and Their Role in Disease. *Mol Interv* (2005) 5:226–40. doi: 10.1124/mi.5.4.6
- Boudreau JE, Hsu KC. Natural Killer Cell Education in Human Health and Disease. *Curr Opin Immunol* (2018) 50:102–11. doi: 10.1016/j.coi.2017.11.003
- Augusto DG. The Impact of KIR Polymorphism on the Risk of Developing Cancer: Not as Strong as Imagined? *Front Genet* (2016) 7:121. doi: 10.3389/fgene.2016.00121
- Anderson KM, Augusto DG, Dandekar R, Shams H, Zhao C, Yusufali T, et al. Killer Cell Immunoglobulin-Like Receptor Variants Are Associated With Protection From Symptoms Associated With More Severe Course in Parkinson Disease. *J Immunol* (2020) 205:1323–30. doi: 10.4049/jimmunol.2000144
- Augusto DG, Lobo-Alves SC, Melo MF, Pereira NF, Petzl-Erler ML. Activating KIR and HLA Bw4 Ligands are Associated to Decreased Susceptibility to Pemphigus Foliaceus, an Autoimmune Blistering Skin Disease. *PLoS One* (2012) 7:e39991. doi: 10.1371/journal.pone.0039991
- Suzuki Y, Hamamoto Y, Ogasawara Y, Ishikawa K, Yoshikawa Y, Sasazuki T, et al. Genetic Polymorphisms of Killer Cell Immunoglobulin-Like Receptors

- are Associated With Susceptibility to Psoriasis Vulgaris. *J Invest Dermatol* (2004) 122:1133–6. doi: 10.1111/j.0022-202X.2004.22517.x
36. Hollenbach JA, Pando MJ, Caillier SJ, Gourraud P-A, Oksenberg JR. The Killer Immunoglobulin-Like Receptor KIR3DL1 in Combination With HLA-Bw4 is Protective Against Multiple Sclerosis in African Americans. *Genes Immun* (2016) 17:199–202. doi: 10.1038/gene.2016.5
 37. Kim H-J, Choi H-B, Jang J-P, Baek I-C, Choi E-J, Park M, et al. HLA-Cw Polymorphism and Killer Cell Immunoglobulin-Like Receptor (KIR) Gene Analysis in Korean Colorectal Cancer Patients. *Int J Surg* (2014) 12:815–20. doi: 10.1016/j.ijssu.2014.06.012
 38. Middleton D, Vilchez JR, Cabrera T, Meenagh A, Williams F, Halfpenny I, et al. Analysis of KIR Gene Frequencies in HLA Class I Characterised Bladder, Colorectal and Laryngeal Tumours. *Tissue Antigens* (2007) 69:220–6. doi: 10.1111/j.1399-0039.2006.00792.x
 39. Al Omar S, Middleton D, Marshall E, Porter D, Xinarianos G, Raji O, et al. Associations Between Genes for Killer Immunoglobulin-Like Receptors and Their Ligands in Patients With Solid Tumors. *Hum Immunol* (2010) 71:976–81. doi: 10.1016/j.humimm.2010.06.019
 40. Auer ED, Tong HV, Amorim LM, Malheiros D, Hoan NX, Issler HC, et al. Natural Killer Cell Receptor Variants and Chronic Hepatitis B Virus Infection in the Vietnamese Population. *Int J Infect Dis* (2020) 96:541–7. doi: 10.1016/j.ijid.2020.05.033
 41. 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) 39:733–40. doi: 10.1038/ng2035
 42. Podhorzer A, Dirchwolf M, Machicote A, Belen S, Montal S, Paz S, et al. The Clinical Features of Patients With Chronic Hepatitis C Virus Infections are Associated With Killer Cell Immunoglobulin-Like Receptor Genes and Their Expression on the Surface of Natural Killer Cells. *Front Immunol* (2017) 8:1912. doi: 10.3389/fimmu.2017.01912
 43. Alves HV, de Moraes AG, Pepineli AC, Tiyo BT, de Lima Neto QA, Santos T da S, et al. The Impact of KIR/HLA Genes on the Risk of Developing Multibacillary Leprosy. *PLoS Negl Trop Dis* (2019) 13:e0007696. doi: 10.1371/journal.pntd.0007696
 44. Trowsdale J, Moffett A. NK Receptor Interactions With MHC Class I Molecules in Pregnancy. *Semin Immunol* (2008) 20:317–20. doi: 10.1016/j.smim.2008.06.002
 45. Moffett A, Chazara O, Colucci F, Johnson MH. Variation of Maternal KIR and Fetal HLA-C Genes in Reproductive Failure: Too Early for Clinical Intervention. *Reprod BioMed Online* (2016) 33:763–9. doi: 10.1016/j.rbmo.2016.08.019
 46. Colucci F. The Role of KIR and HLA Interactions in Pregnancy Complications. *Immunogenetics* (2017) 69:557–65. doi: 10.1007/s00251-017-1003-9
 47. Hiby SE, Apps R, Sharkey AM, Farrell LE, Gardner L, Mulder A, et al. Maternal Activating KIRs Protect Against Human Reproductive Failure Mediated by Fetal HLA-C2. *J Clin Invest* (2010) 120:4102–10. doi: 10.1172/JCI43998
 48. Hollenbach JA, Norman PJ, Creary LE, Damotte V, Montero-Martin G, Caillier S, et al. A Specific Amino Acid Motif of HLA-DRB1 Mediates Risk and Interacts With Smoking History in Parkinson's Disease. *Proc Natl Acad Sci USA* (2019) 116:7419–24. doi: 10.1073/pnas.1821778116
 49. Norman PJ, Hollenbach JA, Nemat-Gorgani N, Marin WM, Norberg SJ, Ashouri E, et al. Defining KIR and HLA Class I Genotypes At Highest Resolution Via High-Throughput Sequencing. *Am J Hum Genet* (2016) 99:375–91. doi: 10.1016/j.ajhg.2016.06.023
 50. Marin W, Dandekar R, Augusto DG, Yusufali T, Heyn B, Hofmann J, et al. High-Throughput Interpretation of Killer-cell Immunoglobulin-Like Receptor Short-Read Sequencing Data With PING. *bioRxiv [Preprint]* (2021). doi: 10.1101/2021.03.24.436770
 51. Pyo C-W, Guethlein LA, Vu Q, Wang R, Abi-Rached L, Norman PJ, 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) 5:e15115. doi: 10.1371/journal.pone.0015115
 52. Vierra-Green C, Roe D, Hou L, Hurley CK, Rajalingam R, Reed E, et al. Allele-Level Haplotype Frequencies and Pairwise Linkage Disequilibrium for 14 KIR Loci in 506 European-American Individuals. *PLoS One* (2012) 7:e47491. doi: 10.1371/journal.pone.0047491
 53. Peakall R, Smouse PE. GenAIE5 6.5: Genetic Analysis in Excel. Population Genetic Software for Teaching and Research—an Update. *Bioinformatics* (2012) 28:2537–9. doi: 10.1093/bioinformatics/bts460
 54. Excoffier L, Lischer HEL. Arlequin Suite Ver 3.5: A New Series of Programs to Perform Population Genetics Analyses Under Linux and Windows. *Mol Ecol Resour* (2010) 10:564–7. doi: 10.1111/j.1755-0998.2010.02847.x
 55. Sanger F, Nicklen S, Coulson AR. DNA Sequencing With Chain-Terminating Inhibitors. *PNAS* (1977) 74:5463–7. doi: 10.1073/pnas.74.12.5463
 56. Vargas L de B, Dourado RM, Amorim LM, Ho B, Calonga-Solis V, Issler HC, et al. Single Nucleotide Polymorphism in KIR2DL1 is Associated With HLA-C Expression in Global Populations. *Front Immunol* (2020) 11:1881. doi: 10.3389/fimmu.2020.01881
 57. Fan QR, Long EO, Wiley DC. Crystal Structure of the Human Natural Killer Cell Inhibitory Receptor KIR2DL1-HLA-Cw4 Complex. *Nat Immunol* (2001) 2:452–60. doi: 10.1038/87766
 58. Grant BJ, Rodrigues APC, ElSawy KM, McCammon JA, Caves LSD. Bio3d: An R Package for the Comparative Analysis of Protein Structures. *Bioinformatics* (2006) 22:2695–6. doi: 10.1093/bioinformatics/btl461
 59. Phillips JC, Hardy DJ, Maia JDC, Stone JE, Ribeiro JV, Bernardi RC, et al. Scalable Molecular Dynamics on CPU and GPU Architectures With NAMM. *J Chem Phys* (2020) 153:044130. doi: 10.1063/5.0004475
 60. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of Simple Potential Functions for Simulating Liquid Water. *J Chem Phys* (1983) 79:926–35. doi: 10.1063/1.445869
 61. Pugh J, Nemat-Gorgani N, Djaoud Z, Guethlein LA, Norman PJ, Parham P. In Vitro Education of Human Natural Killer Cells by KIR3DL1. *Life Sci Alliance* (2019) 2:e201900434. doi: 10.26508/lsa.201900434
 62. Wilson MJ, Torkar M, Haude A, Milne S, Jones T, Sheer D, et al. Plasticity in the Organization and Sequences of Human KIR/ILT Gene Families. *Proc Natl Acad Sci USA* (2000) 97:4778. doi: 10.1073/pnas.080588597
 63. Vilches C, Parham P. KIR: Diverse, Rapidly Evolving Receptors of Innate and Adaptive Immunity. *Annu Rev Immunol* (2002) 20:217–51. doi: 10.1146/annurev.immunol.20.092501.134942
 64. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, et al. dbSNP: The NCBI Database of Genetic Variation. *Nucleic Acids Res* (2001) 29:308–11. doi: 10.1093/nar/29.1.308
 65. Stewart CA, Laugier-Anfossi F, Vély F, Saulquin X, Riedmuller J, Tisserant A, et al. Recognition of Peptide-MHC Class I Complexes by Activating Killer Immunoglobulin-Like Receptors. *PNAS* (2005) 102:13224–9. doi: 10.1073/pnas.0503594102
 66. Hilton PG, Guethlein LA, Goyas A, Nemat-Gorgani N, Bushnell DA, Norman PJ, et al. Polymorphic HLA-C Receptors Balance the Functional Characteristics of KIR Haplotypes. *J Immunol* (2015) 195:3160–70. doi: 10.4049/jimmunol.1501358
 67. Bari R, Bell T, Leung W-H, Vong QP, Chan WK, Das Gupta N, et al. Significant Functional Heterogeneity Among KIR2DL1 Alleles and a Pivotal Role of Arginine245. *Blood* (2009) 114:5182–90. doi: 10.1182/blood-2009-07-231977
 68. 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) 121:4703–7. doi: 10.1182/blood-2012-10-461442
 69. Le Lueduc J-B, Boudreau JE, Freiberg JC, Hsu KC. Novel Approach to Cell Surface Discrimination Between KIR2DL1 Subtypes and KIR2DS1 Identifies Hierarchies in NK Repertoire, Education, and Tolerance. *Front Immunol* (2019) 10:734. doi: 10.3389/fimmu.2019.00734
 70. Norman PJ, Abi-Rached L, Gendzekhadze K, Korbel D, Gleimer M, Rowley D, et al. Unusual Selection on the KIR3DL1/S1 Natural Killer Cell Receptor in Africans. *Nat Genet* (2007) 39:1092–9. doi: 10.1038/ng2111
 71. Augusto DG, Norman PJ, Dandekar R, Hollenbach JA. Fluctuating and Geographically Specific Selection Characterize Rapid Evolution of the Human KIR Region. *Front Immunol* (2019) 10:989. doi: 10.3389/fimmu.2019.00989
 72. Vivian JP, Duncan RC, Berry R, O'Connor GM, Reid HH, Beddoe T, et al. Killer Cell Immunoglobulin-Like Receptor 3DL1-Mediated Recognition of Human Leukocyte Antigen B. *Nature* (2011) 479:401–5. doi: 10.1038/nature10517
 73. Pymm P, Illing PT, Ramarathnam SH, O'Connor GM, Hughes VA, Hitchen C, et al. MHC-I Peptides Get Out of the Groove and Enable a Novel Mechanism of HIV-1 Escape. *Nat Struct Mol Biol* (2017) 24:387–94. doi: 10.1038/nsmb.3381

74. Saunders PM, MacLachlan BJ, Pymm P, Illing PT, Deng Y, Wong SC, et al. The Molecular Basis of How Buried Human Leukocyte Antigen Polymorphism Modulates Natural Killer Cell Function. *PNAS* (2020) 117:11636–47. doi: 10.1073/pnas.1920570117
75. Fan QR, Mosyak L, Winter CC, Wagtmann N, Long EO, Wiley DC. Structure of the Inhibitory Receptor for Human Natural Killer Cells Resembles Haematopoietic Receptors. *Nature* (1997) 389:96–100. doi: 10.1038/38028
76. Stewart-Jones GBE, di Gleria K, Kollnberger S, McMichael AJ, Jones EY, Bowness P. Crystal Structures and KIR3DL1 Recognition of Three Immunodominant Viral Peptides Complexed to HLA-B*2705. *Eur J Immunol* (2005) 35:341–51. doi: 10.1002/eji.200425724
77. Grantham R. Amino Acid Difference Formula to Help Explain Protein Evolution. *Science* (1974) 185:862–4. doi: 10.1126/science.185.4154.862
78. Vilches C, Rajalingam R, Uhrberg M, Gardiner CM, Young NT, Parham P. KIR2DL5, a Novel Killer-Cell Receptor With a D0-D2 Configuration of Ig-Like Domains. *J Immunol* (2000) 164:5797–804. doi: 10.4049/jimmunol.164.11.5797
79. Barten R, Torkar M, Haude A, Trowsdale J, Wilson MJ. Divergent and Convergent Evolution of NK-cell Receptors. *Trends Immunol* (2001) 22:52–7. doi: 10.1016/S1471-4906(00)01802-0
80. Gómez-Lozano N, Estefania E, Williams F, Halfpenny I, Middleton D, Solís R, et al. The Silent KIR3DP1 Gene (CD158c) is Transcribed and Might Encode a Secreted Receptor in a Minority of Humans, in Whom the KIR3DP1, KIR2DL4 and KIR3DL1/KIR3DS1 Genes are Duplicated. *Eur J Immunol* (2005) 35:16–24. doi: 10.1002/eji.200425493
81. Norman PJ, Carrington CVF, Byng M, Maxwell LD, Curran MD, Stephens HAF, et al. Natural Killer Cell Immunoglobulin-Like Receptor (KIR) Locus Profiles in African and South Asian Populations. *Genes Immun* (2002) 3:86–95. doi: 10.1038/sj.gene.6363836
82. Williams F, Maxwell LD, Halfpenny IA, Meenagh A, Sleator C, Curran MD, et al. Multiple Copies of KIR 3DL/S1 and KIR 2DL4 Genes Identified in a Number of Individuals. *Hum Immunol* (2003) 64:729–32. doi: 10.1016/s0198-8859(03)00089-2
83. Solloch UV, Schefzyk D, Schäfer G, Massalski C, Kohler M, Pruschke J, et al. Estimation of German Kir Allele Group Haplotype Frequencies. *Front Immunol* (2020) 11:429. doi: 10.3389/fimmu.2020.00429
84. Auton A, Abecasis GR, Altshuler DM, Durbin RM, Abecasis GR, Bentley DR, et al. A Global Reference for Human Genetic Variation. *Nature* (2015) 526:68–74. doi: 10.1038/nature15393
85. Hilton HG, Vago L, Older Aguilar A, Moesta AK, Graef T, Abi Rached L, et al. Mutation At Positively Selected Positions in the Binding Site for HLA-C Shows KIR2DL1 is a More Refined But Less Adaptable NK Cell Receptor Than KIR2DL3. *J Immunol* (2012) 189:1418–30. doi: 10.4049/jimmunol.1100431
86. Wagner I, Schefzyk D, Pruschke J, Schöfl G, Schöne B, Gruber N, et al. Allele-Level KIR Genotyping of More Than a Million Samples: Workflow, Algorithm, and Observations. *Front Immunol* (2018) 9:2843. doi: 10.3389/fimmu.2018.02843
87. Crum KA, Logue SE, Curran MD, Middleton D. Development of a PCR-SSOP Approach Capable of Defining the Natural Killer Cell Inhibitory Receptor (KIR) Gene Sequence Repertoires. *Tissue Antigens* (2000) 56:313–26. doi: 10.1034/j.1399-0039.2000.560403.x
88. Hollenbach JA, Augusto DG, Alaez C, Bubnova L, Fae I, Fischer G, et al. 16 (Th) IHIW: Population Global Distribution of Killer Immunoglobulin-Like Receptor (KIR) and Ligands. *Int J Immunogenet* (2013) 40:39–45. doi: 10.1111/iji.12028
89. Augusto DG, Piovezan BZ, Tsuneto LT, Callegari-Jacques SM, Petzl-Erler ML. KIR Gene Content in Amerindians Indicates Influence of Demographic Factors. *PloS One* (2013) 8:e56755. doi: 10.1371/journal.pone.0056755

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.

Copyright © 2021 Amorim, Augusto, Nemat-Gorgani, Montero-Martin, Marin, Shams, Dandekar, Caillier, Parham, Fernández-Viña, Oksenberg, Norman and Hollenbach. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.