










RESEARCH ARTICLE

Allele imputation for the killer cell immunoglobulin-like receptor KIR3DL1/S1

Genelle F. Harrison ^{1,2}*, Laura Ann Leaton ^{1,2}*, Erica A. Harrison³, Katherine M. Kichula ^{1,2}, Marte K. Viken ^{4,5}, Jonathan Shortt ¹, Christopher R. Gignoux ¹, Benedicte A. Lie^{4,5}, Damjan Vukcevic ^{6,7}, Stephen Leslie ^{6,7,8}, Paul J. Norman ^{1,2*}

1 Division of Biomedical Informatics and Personalized Medicine, University of Colorado, Anschutz Medical Campus, Aurora, Colorado, United States of America, **2** Department of Immunology and Microbiology, University of Colorado, Anschutz Medical Campus, Aurora, Colorado, United States of America, **3** Independent Researcher, Broomfield, Colorado, United States of America, **4** Department of Immunology, University of Oslo and Oslo University Hospital, Oslo, Norway, **5** Department of Medical Genetics, University of Oslo and Oslo University Hospital, Oslo, Norway, **6** School of Mathematics and Statistics, University of Melbourne, Parkville, Victoria, Australia, **7** Melbourne Integrative Genomics, University of Melbourne, Parkville, Victoria, Australia, **8** School of BioSciences, University of Melbourne, Parkville, Victoria, Australia

* These authors contributed equally to this work.

* paul.norman@cuanschutz.edu



OPEN ACCESS

Citation: Harrison GF, Leaton LA, Harrison EA, Kichula KM, Viken MK, Shortt J, et al. (2022) Allele imputation for the killer cell immunoglobulin-like receptor KIR3DL1/S1. *PLoS Comput Biol* 18(2): e1009059. <https://doi.org/10.1371/journal.pcbi.1009059>

Editor: Roger Dimitri Kouyos, University of Zurich, SWITZERLAND

Received: May 3, 2021

Accepted: January 10, 2022

Published: February 22, 2022

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pcbi.1009059>

Copyright: © 2022 Harrison et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All code written in support of this publication, imputation models, test data and documentation on installing and running

Abstract

Highly polymorphic interaction of KIR3DL1 and KIR3DS1 with HLA class I ligands modulates the effector functions of natural killer (NK) cells and some T cells. This genetically determined diversity affects severity of infections, immune-mediated diseases, and some cancers, and impacts the course of immunotherapies, including transplantation. KIR3DL1 is an inhibitory receptor, and KIR3DS1 is an activating receptor encoded by the *KIR3DL1/S1* gene that has more than 200 diverse and divergent alleles. Determination of *KIR3DL1/S1* genotypes for medical application is hampered by complex sequence and structural variation, requiring targeted approaches to generate and analyze high-resolution allele data. To overcome these obstacles, we developed and optimized a model for imputing *KIR3DL1/S1* alleles at high-resolution from whole-genome SNP data. We designed the model to represent a substantial component of human genetic diversity. Our Global imputation model is effective at genotyping *KIR3DL1/S1* alleles with an accuracy ranging from 88% in Africans to 97% in East Asians, with mean specificity of 99% and sensitivity of 95% for alleles >1% frequency. We used the established algorithm of the HIBAG program, in a modification named Pulling Out Natural killer cell Genomics (PONG). Because HIBAG was designed to impute *HLA* alleles also from whole-genome SNP data, PONG allows combinatorial diversity of *KIR3DL1/S1* with *HLA-A* and *-B* to be analyzed using complementary techniques on a single data source. The use of PONG thus negates the need for targeted sequencing data in very large-scale association studies where such methods might not be tractable.

are publicly available at <https://github.com/NormanLabUCD/PONG>.

Funding: This study was performed with support from National Institutes of Health of the USA, R56 AI151549 (P.J.N.) and R01 AI128775 (P.J.N.), and R01 HG010297 (C.R.G.). Australian NH and MRC, 2004262 (S.L.). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: I have read the journal's policy and the authors of this manuscript have the following competing interests: Dr. Stephen Leslie is a partner with Peptide Groove LLP. All other authors have no competing interests to declare.

Author summary

Natural killer (NK) cells are cytotoxic lymphocytes that identify and kill infected or malignant cells and guide immune responses. The effector functions of NK cells are modulated through polymorphic interactions of KIR3DL1/S1 on their surface with the human leukocyte antigens (HLA) that are found on most other cell types in the body. KIR3DL1/S1 is highly polymorphic and differentiated across human populations, affecting susceptibility and course of multiple immune-mediated diseases and their treatments. Genotyping *KIR3DL1/S1* for direct medical application or research has been encumbered by the complex sequence and structural variation, which requires targeted approaches and extensive domain expertise to generate and validate high-resolution allele calls. We therefore developed Pulling Out Natural Killer Cell Genomics (PONG) to impute *KIR3DL1/S1* alleles from whole genome SNP data, and which we implemented as an open-source R package. We assessed imputation performance using data from five broad population groups that represent a substantial portion of human genetic diversity. We can impute *KIR3DL1/S1* alleles with an accuracy ranging from 88% in Africans to 97% in East Asians. Globally, imputation of *KIR3DL1/S1* alleles having frequency >1% has a mean sensitivity of 95% and specificity of 99%. Thus, the PONG method both enables highly sensitive individual-level calling and makes large scale medical genetic studies of *KIR3DL1/S1* possible.

Introduction

The *KIR3DL1/S1* gene encodes highly polymorphic receptors that are expressed by natural killer (NK) cells and some T cells to modulate their effector functions in immunity [1, 2]. The receptors interact with HLA class I ligands that are expressed by most nucleated cells to signify their health status to the immune system [3, 4]. KIR3DL1 allotypes are inhibitory receptors, specific for subsets of highly polymorphic HLA-A and B [5, 6]. The KIR3DS1 allotypes are activating receptors, specific for non-polymorphic HLA-F and a smaller subset of HLA-A and -B [7–9]. Sequence diversity of KIR3DL1/S1 and HLA class I allotypes diversifies human immune responses to specific infections, cancers, cancer treatment and transplantation [10–19]. Accordingly, this genetically determined diversity also associates with differential susceptibility and severity for multiple immune-mediated diseases [20–26]. Although these factors render it imperative to genotype *KIR3DL1/S1* allotypes accurately for medical research and applications that include therapy decisions [27, 28], the high complexity of the genomic region presents obstacles for standard ascertainment methods [29]. The ability to impute alleles from whole-genome SNP genotype (WG-SNP) data will decrease expense and effort, and greatly increase the capacity of research or applications where knowledge of KIR3DL1/S1 and HLA class I combinatorial diversity is critical.

The *KIR* locus, on human chromosome 19, is highly divergent in sequence and structure [29]. As defined by the extensively curated ImmunoPolymorphism Database (IPD), *KIR3DL1/S1* has 220 alleles characterized (release 2.10.0: December 2020), with large numbers continuing to be discovered [30]. As observed for polymorphic *HLA*, the *KIR3DL1/S1* alleles both distinguish individuals and characterize broad ancestral human populations [31, 32]. As a likely consequence of selective pressure providing resistance to infectious diseases [33], specific combinations of KIR3DL1/S1 and HLA associate, differentially across populations, with severity of specific viral infections or autoimmune diseases [34–41]. Likewise, specific combinations of KIR3DL1/S1 with HLA class I influence cancer susceptibilities non-uniformly across populations [42]. In this regard, two key areas of human health significantly impacted by the

population differentiation of KIR3DL1/S1 combinatorial diversity with HLA are HIV research and treatment, and cancer therapy [17, 43–48]. In particular, specific combinations of KIR3DL1/S1 and HLA allotypes influence rejection and relapse rates following transplantation [49–52]. For these reasons, it is critical to establish methods for elucidating genetic variation in *KIR3DL1/S1* that can accommodate the full range of human genetic diversity.

KIR3DL1 specifically binds to subsets of HLA-A or B that carry a five amino acid motif, termed Bw4, on their external facing α 1-helix [53]. Expression of KIR3DL1 gives NK cells the ability to detect diseased cells that may have lost or altered expression of these HLA class I molecules [54, 55]. KIR3DL1 also likely serves as an immune checkpoint inhibitor for functionally mature T cells [56]. KIR3DL1 polymorphism, and polymorphism both within and outside the Bw4 motif of HLA affects the specificity and strength of the interaction [57–59]. Polymorphism also determines the expression level or signal transduction abilities of the receptor [60, 61]. KIR3DL1/S1 segregates into three ancient lineages (015, 005 and 3DS1) that have distinct expression and function phenotypes [31]. The 015 lineage comprises inhibitory receptors having high expression and high affinity for Bw4⁺HLA-B. The 005 lineage are inhibitory receptors having low expression and preferential affinity for Bw4⁺HLA-A. The 3DS1-lineage are activating receptors specific for HLA-F and some Bw4⁺HLA-B allotypes expressed by infected cells [62–64]. As defined by these phenotypes, the lineages differentially associate with distinct pathological phenotypes [15, 65–67]. Exceptions to these broad rules (e.g. 3DL1*007 belongs to the 015 lineage but has low expression [61]) contribute to a hierarchy of receptor allotype strengths and reinforce the need to genotype *KIR3DL1/S1* to high resolution [68–70].

Multiple methods are available to impute *HLA class I* genotypes with high accuracy from WG-SNP data [71–75]. We chose to adapt one of these programs so that *KIR3DL1/S1* and *HLA-A* and *B* genotypes could be imputed from the same data source, using an identical algorithm. In the current study we have adapted the HIBAG framework [75] to impute *KIR3DL1/S1* alleles, in a modification we have named Pulling Out NK cell Genomics (PONG). There are two components to the process: 1) model building that employs machine-learning to determine which combinations of SNPs correlate with known alleles, and 2) imputation that uses this model to determine *KIR3DL1/S1* allele genotypes from study cohorts [75, 76]. Construction of the imputation models required high-resolution *KIR3DL1/S1* alleles and WG-SNP data obtained from the same set of individuals. PONG thus serves as a complement to PING (Pushing Immunogenetics to the Next Generation), which can determine *KIR3DL1/S1* alleles from high throughput sequence data [77]. With a goal to create a model representing a substantial component of human genetic diversity, we compiled and rigorously tested the imputation using data from the 1,000 Genomes populations [78]. The R package PONG is freely available, as are the data sets and imputation models described in this study.

Materials and methods

Method overview

KIR3DL1/S1 exhibits exceptional sequence polymorphism as well as variation in gene content (Fig 1A). Here, we adapted and optimized the framework of HIBAG [75] to impute *KIR3DL1/S1* alleles, in a modification we have named Pulling Out NK cell Genomics (PONG). The development of PONG was focused on building a robust training model that could be used to impute unknown *KIR3DL1/S1* alleles from WG-SNP data across diverse global populations. Training an imputation model requires an input of individuals having known *KIR3DL1/S1* genotypes, coupled with high-density SNP data from the *KIR* region, as typically obtained through whole-genome SNP analysis (Fig 1B). We optimized the process using 1,000 Genomes individuals, because we had previously determined their *KIR3DL1/S1* alleles from sequence

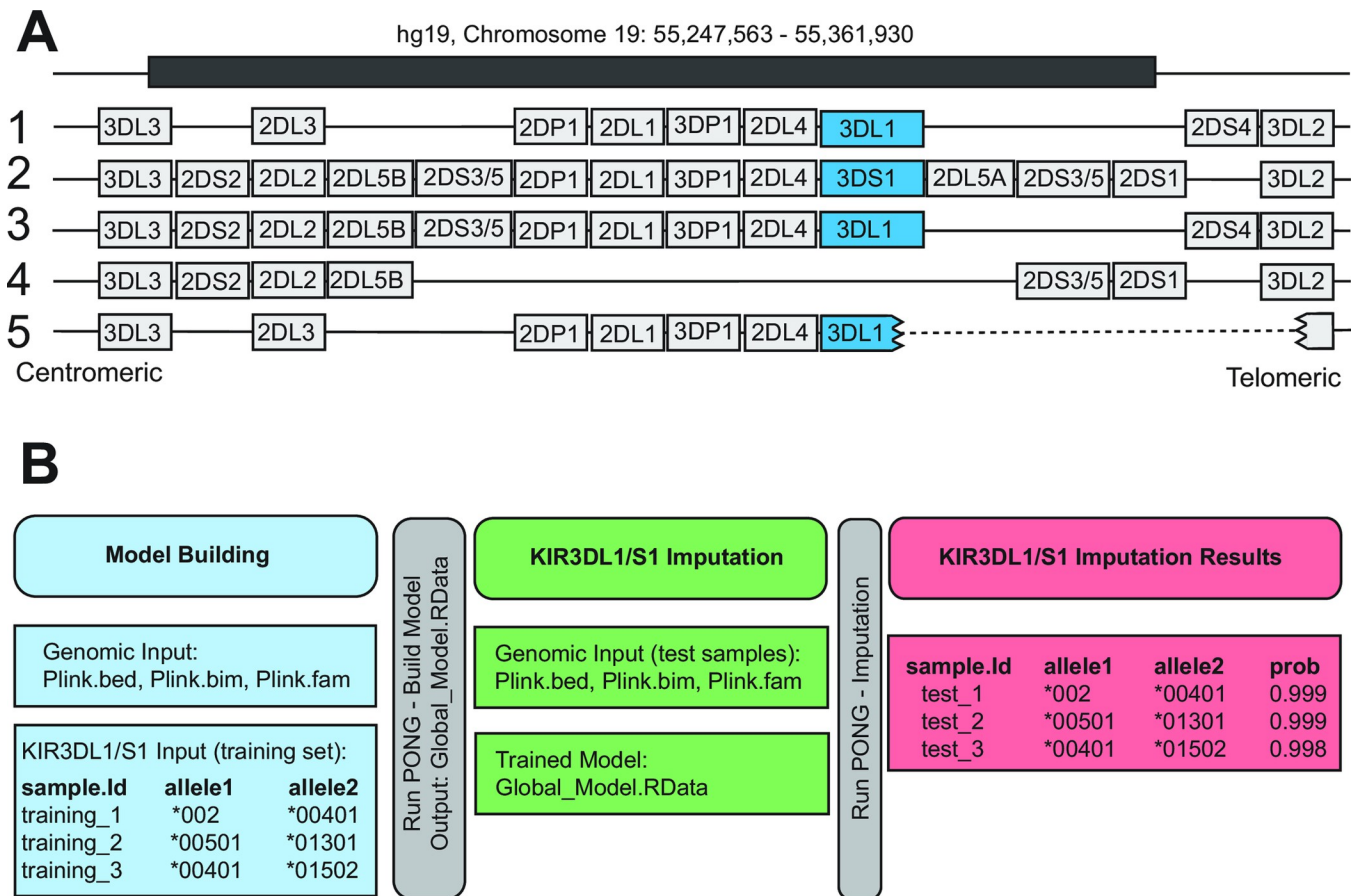


Fig 1. Genomic location of *KIR3DL1/S1* and overview of allele imputation workflow. **A.** Shows the location of the *KIR3DL1/S1* gene on five examples of *KIR* haplotypes. *KIR3DL1/S1* is shaded in blue, and other *KIR* genes are shaded grey. The *KIR3DL1/S1* gene can be absent (haplotype 4) or fused in-frame with *KIR3DL2* (haplotype 5) [92]. The human genome coordinates (build hg19) from which classifiers were drawn for imputation are given at the top. **B.** Schematic of model building, testing and output for the imputation of *KIR3DL1/S1* alleles using PONG. Shown are the required input files and their format for model building (blue) and testing (green). Red boxes give an example of the output from the imputation.

<https://doi.org/10.1371/journal.pcbi.1009059.g001>

data [79] and because high-density SNP data (Illumina Omni 2.5 chip) is available from this cohort [78]. We distributed the 1,000 Genomes individuals into the designated five major population groups: Africa and African-descent (AFR), Americas (AMR), East Asia (EAS), Europe (EUR) and South Asia (SAS). We first optimized the model building parameters using the EUR group. We randomly divided each population group into two parts, building an imputation model using the first part and testing the imputation accuracy using the second part. We then built a global model by combining all of the 1,000 genomes individuals and repeating the process. We established that we could impute *KIR3DL1/S1* alleles using data from multiple low-density SNP genotyping platforms. Finally, we tested the efficacy of PONG using an independent population having both high-resolution sequence and WG-SNP data.

Samples and genomic data

We obtained high density SNP data for chromosome 19, which contains the *KIR* genomic region (chromosome 19: 55247563–55361930, hg19), from the 1,000 Genomes Project Phase 3 individuals [78]. The data had been generated using the Illumina Omni 2.5 platform, which has 4,093 SNPs in the *KIR* genomic region [78]. To determine the *KIR3DL1/S1* alleles present

in each individual we used the Pushing Immunogenetics to the Next Generation (PING) pipeline on high-depth sequence data, as previously described [79]. In total, there were 2,082 individuals from the 1,000 Genomes data set from whom we had independently derived *KIR* sequence and chromosome 19 SNP data available (S1 Table), and these were divided into the designated five major population groups as indicated (S2 Table). We also analyzed SNP data obtained using the Infinium Immunoarray 24v2 (aka ImmunoChip) [80] from 397 Norwegians [81], from whom we determined high resolution *KIR3DL1/S1* genotypes (S3 Table) through targeted sequencing [79].

Modifications to HIBAG to impute *KIR3DL1/S1*

We modified the HIBAG package version 1.2.4. The package name and relevant C++ functions were changed from HIBAG to KIRpong to avoid any conflict when both programs are installed. We removed genome build hg18 and included hg19 and GRCh38 (hg38) instead. We maintained many of the HIBAG functions while adjusting the selected chromosome positions to target the *KIR* gene cluster on chromosome 19. We modified the 'hlaBED2Geno' function to sample chromosome 19 positions 50247563–59128983 for hg19 and 46457117–58617616 for hg38. The 'hlaLociInfo' function was updated to target the *KIR* gene cluster and was specified as 55247563–55361930 for hg19 and 54734034–54853884 for hg38. Also in this function, the name of the gene was changed to *KIR3DL1/S1*. We also created PONG–Extended Window in which the 'hlaBED2Geno' function to sample chromosome 19 positions 55,100,000–55,500,000 (hg19) that can be used with low density SNP arrays to increase the size of the genomic region from which classifiers can be established. Finally, the printout messages were changed from HIBAG and *HLA* to PONG and *KIR3DL1/S1* to avoid confusion if both programs are active. The HIBAG functions were maintained, as extensive documentation for these functions is available.

Optimization and testing of model building

The input data for model building is a text file containing the *KIR3DL1/S1* allele information, and SNP data in PLINK [82] binary format (.bed,.bim,.fam files) from the same individuals (Fig 1B). The first column of the text file contains the sample name (SampleID), the second column, *KIR3DL1/S1* allele 1 (Allele1) and the third column, allele 2 (Allele2). We optimized the model parameters using the 1,000 Genomes European populations group (EUR), comprising 353 individuals from five countries [78] and having 26 distinct *KIR3DL1/S1* alleles [79]. We chose the EUR subset to develop the filtering thresholds because this was the group having the fewest number of *KIR3DL1/S1* alleles before filtering, and because Europeans are currently the most extensively characterized of any major population groups [83–86]. We then expanded model building and testing to include populations from Africa (AFR, 558 individuals, 46 distinct *KIR3DL1/S1* alleles), Hispanic/Latino populations from the Americas (AMR, 298 individuals, 34 *KIR3DL1/S1* alleles), East Asia (EAS, 406 individuals, 28 *KIR3DL1/S1* alleles) and South Asia (SAS, 467 individuals, 30 *KIR3DL1/S1* alleles).

We randomly selected 50% of individuals from the EUR group to be used for model building. The remaining 50% of individuals were used to test the accuracy of the model. We first optimized the parameters to be used for filtering SNPs prior to model building. We compared the imputation accuracy of models built after removing SNPs with minor allele counts (MAC) < 2, or < 3, or a minor allele frequency (MAF) < 1% or < 5%. We also tested the impact of removing individuals carrying any *KIR3DL1/S1* allele having MAC < 3 in the full EUR group (model + test). Once a robust model was established for the EUR population, we expanded the model to include all populations using the pre-filtering parameters and procedures established

above. Each model was evaluated based on the time needed for model building as well as the accuracy of imputation.

Imputing *KIR3DL1/S1* alleles from low density SNP arrays

The next goal of this study was to test the accuracy of *KIR3DL1/S1* imputation using data from commonly used low density SNP arrays. From the 1,000 Genomes sequence data vcf files [78], we extracted those chromosome 19 SNPs equivalent to the ImmunoChip (4,902 SNPs, 26 in *KIR* region), Infinium (23,117 SNPs, 18 in *KIR* region) or MEGA (36,534 SNPs, 76 in *KIR* region) arrays. To increase the number of available classifiers, we built a second Global model in which we expanded the target region to chr19: 55,100,000–55,500,000 (hg19). These coordinates match those of the *KIR*IMP* program, which can be used to impute *KIR* gene content genotypes [87]. We then built and tested *KIR3DL1/S1* allele imputation models as described above. We used the Michigan imputation server [88] to supplement the low-density data with Illumina Omni 2.5 array data (1,832,506 chromosome 19 SNPs, 4,093 in *KIR* region). Input data was first phased with EAGLE [89] and an RSQ filter of 0.3 was applied to ensure imputation accuracy [90].

Imputing *KIR3DL1/S1* alleles from an independent dataset

We also tested the accuracy of *KIR3DL1/S1* imputation using DNA sequence and WG-SNP data from 397 individuals from Norway. The cohort contained 18 distinct *KIR3DL1/S1* alleles (S3 Table), 14 of which were also present in the 1,000 Genomes data set. The cohort was genotyped using the ImmunoChip array, which has 26 SNPs in the *KIR* region, 16 of which are polymorphic in the data set. To increase model classifier density, we supplemented the SNP data with that of the Illumina Omni 2.5 array using the Michigan imputation server as described above. This process produced 715 SNPs in the *KIR* region that are variable in the data set. We also tested the extended genomic window (chr19: 55,100,000–55,500,000, hg19) for model building and this produced 3,440 variable SNPs in the Norway data set.

Computational resources

All experiments were performed using a server with 512 GB 2400MHz RAM, running Ubuntu 18.04, R 3.5.1, R-server 1.1.456, and using a single core from a 2.3GHz Xeon E5-2697 CPU.

Evaluation of imputation models

Overall accuracy of a given imputation model was determined as the number of correct allele calls made per individual (0, 1 or 2) divided by 2N. Sensitivity and specificity of a given model were determined per *KIR3DL1/S1* allele. Sensitivity was measured as the percentage of individuals known to be positive for a given allele who were also called positive for that allele by imputation. Specificity was determined as the percentage of individuals known to be negative for a given allele that were also called as negative for that allele by imputation.

Results

Parameters for frequency filtering of SNP and allele data

We designed, tested and optimized a model to impute *KIR3DL1/S1* alleles from WG-SNP data using a modification to the HIBAG framework and algorithm [75]. We used SNP data from the 1,000 Genomes project [78] and *KIR3DL1/S1* genotypes that we had previously determined from sequence data from the same individuals [79]. We focused first on the EUR group, comprised of 353 individuals and having 26 distinct *KIR3DL1/S1* allele sequences, ranging from

0.14% to 20% allele frequency (Fig 2A). We randomly selected 50% of the EUR individuals to be used for model building and used the other 50% to test the accuracy of the model. With the goal of maximizing the imputation accuracy of the test dataset, while preserving computational efficiency in model building, we first determined the effect of removing low-frequency SNPs. We measured the imputation accuracy of models that were built following removal of SNPs having a minor allele count (MAC) of < 2 (1,286 SNPs remaining in the *KIR* region) or $MAC < 3$ (1,044 SNPs remaining in the *KIR* region) in the full set of 353 individuals. We also measured the accuracy following removal of SNPs with a minor allele frequency (MAF) $< 1\%$ (941 SNPs remaining in the *KIR* region) or $< 5\%$ (645 SNPs remaining in the *KIR* region) in the full set of 353 individuals. A model was also built with no filtering of the genotype data for comparison (4,089 SNPs in *KIR* region). The *KIR* region SNPs used for testing the model accuracy were not filtered, and the models took from 6–9 seconds to impute *KIR3DL1/S1* alleles from the test data set of 177 individuals.

For the purposes of this test, accuracy was determined as the percentage of correct allele calls per individual in the test set. The lowest imputation accuracy, with 91% of individuals genotyped correctly, was obtained when no filtering was used or using a $MAF < 5\%$. Models built using all other filtering parameters imputed the genotypes with 92% accuracy (Fig 2B and S4 Table). Thus, imputation accuracy was similar across all SNP frequency filtering parameters tested (Fig 2B). The model building run time ranged from 29 minutes, when SNPs were filtered at $MAF < 5\%$, to 84 minutes when no filtering was used. We selected $MAC < 3$ as this was the fastest build time (66 minutes) for models of 92% accuracy (Fig 2B). Of the 26 *KIR3DL1/S1* alleles observed in the full EUR group ($N = 353$), twelve were observed less than three times (Fig 2A). Following removal of individuals possessing at least one of these twelve infrequent alleles, 14 *KIR3DL1/S1* alleles and 339 individuals remained in the population. As above, we removed SNPs having $MAC < 3$, divided this population in half, built a model and tested it on the other half. Compared with using $MAC < 3$ alone, the time required to build this model decreased from 66 minutes to 46 minutes, and the time to run the model reduced to 5 seconds (165 individuals), whereas the imputation accuracy increased from 92% to 96% (Fig 2B). Thus, this combination of filtering parameters produced the fastest time for model building and running, with the highest accuracy for imputing *KIR3DL1/S1* genotypes. Moreover, these parameters produced the greatest imputation accuracy across all five major population groups (S5 Table). We therefore implemented these parameters ($MAC < 3$ both for SNPs and *KIR3DL1/S1* alleles) in all subsequent analyses.

We next evaluated the sensitivity and specificity of the final EUR imputation model. Of 14 alleles in the model data set, 13 were also present in the test set (Fig 2C). We observed a mean specificity of 99%. The two alleles having $< 100\%$ specificity were *KIR3DL1*00101* and *KIR3DS1*01301*, with 98% specificity, thus for every 100 individuals imputed to have either allele, two were not shown as present through sequencing. We observed a mean sensitivity of 77%. All alleles with a frequency in the EUR group greater than 1.6% were imputed with $>99\%$ sensitivity. Below 1.6% allele frequency, two alleles (*3DL1*00402* and *3DS1*049N*) were imputed with 50% sensitivity and two (absence of *KIR3DL1/S1* (*00000) and *3DL1*009*) with 0% sensitivity. *KIR3DL1*00402* and *3DS1*049N* are each distinguished from their closest (parental) alleles by a single or a doublet nucleotide substitution, respectively [30]. Accordingly, in each case these alleles were imputed as the parental allele. *KIR3DL1*009* represents a double recombination having exons 2–3 identical to *3DS1*01301* and exons 1 and 4–9 identical to *3DL1*001* [91, 92]. The haplotype that lacks *KIR3DL1/S1* represents a large-scale deletion encompassing up to seven *KIR* genes (Fig 1A), and likely has very few identifying SNPs within the *KIR* locus. Thus, we observe a clear relationship between *KIR3DL1/S1* allele frequency and accuracy of imputation, with all high-frequency alleles being imputed with high

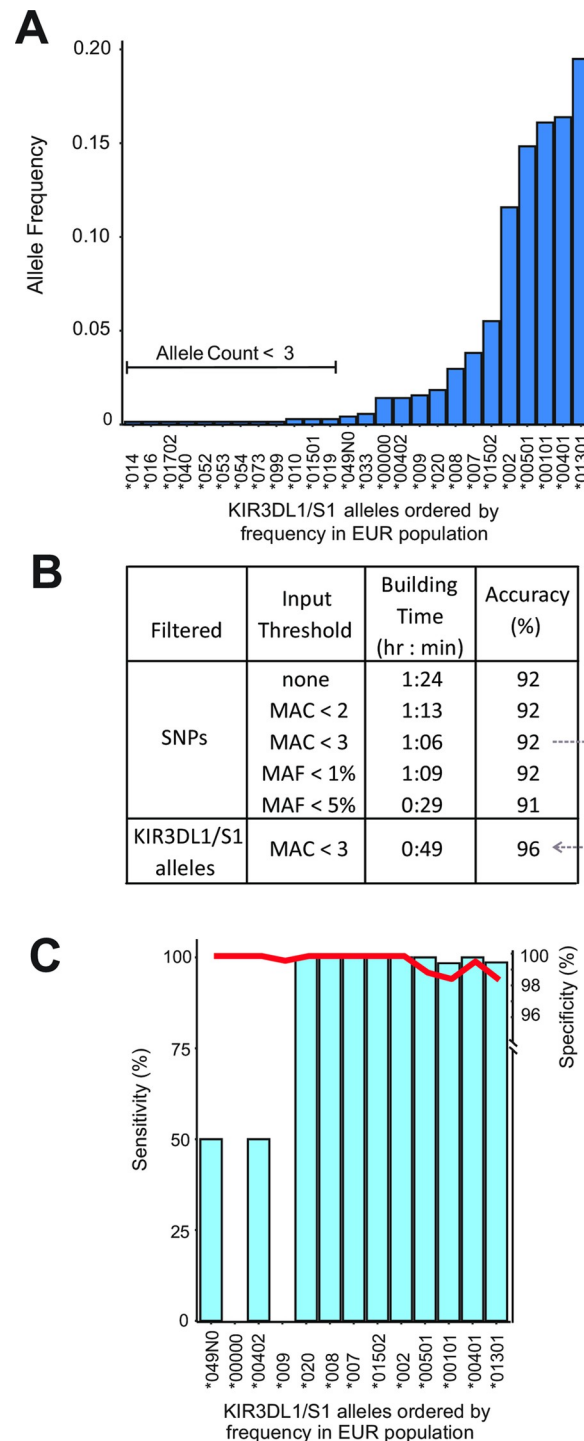


Fig 2. Optimization of KIR3DL1/S1 allele imputation using data from Europeans. **A.** Bar graph shows the KIR3DL1/S1 allele frequencies in the combined EUR population group [78] comprised of 353 individuals from Italy, Finland, United Kingdom, Spain, or Utah. The alleles were determined from short read sequence data [79]. **B.** Shown is a summary of the results obtained using models tested during optimization. From left to right are the filtered criteria (SNPs or KIR3DL1/S1 alleles), the filtering threshold values, resulting model build time, and accuracy of the imputed genotypes. Grey dotted arrow indicates that the final model was built using MAC < 3 for SNPs and for KIR3DL1/S1 alleles. **C.** Shows the imputation accuracy for each KIR3DL1/S1 allele present in the final filtered EUR data set. Blue bars indicate the sensitivity (% of times a given allele was called as present when known to be present). Red line indicates specificity (% of times a given allele was called as absent when known to be absent).

<https://doi.org/10.1371/journal.pcbi.1009059.g002>

accuracy, and those imputed with lower accuracy attributed both to their low frequency and lack of additional identifying characteristics.

Development of a trained Global model for *KIR3DL1/S1* imputation

After establishing the most robust filtering parameters for model building in the EUR population group, we expanded the analysis to the four other major population groups from the 1,000 Genomes Project (Africa—AFR, Americas—AMR, East Asia—EAS and South Asia—SAS). We also combined all five population groups to form an additional ‘Global’ group (Table 1). As above, *KIR* locus SNPs and *KIR3DL1/S1* alleles having MAC < 3 in each respective group were removed, imputation models were then built using 50% individuals, and tested on the remaining 50%. Following the filtering based on *KIR3DL1/S1* allele counts, the African population group had the highest diversity with 31 alleles and the East Asian group the lowest with 13 alleles (Fig 3A). A total of 90 distinct *KIR3DL1/S1* alleles were present in the Global group, 42 of these occurred more than twice in total and were therefore included in model building. This allele filtering process resulted in 58 of the 2,082 individuals being removed. The Global model included 1,017 individuals and took 10 weeks to build. This process also increased the number of target alleles within all the individual population groups (Fig 3A).

In testing models built within each respective population group, imputation accuracy ranged from 80% in the AFR group to 96% in the EAS group (Fig 3B and S4 Table). When using the Global model, however, imputation accuracy increased for all groups, ranging from 88% in the AFR group to 97% in EAS (Fig 3B). When the test group was comprised of individuals drawn from all five of the population groups, an accuracy of 92% was achieved. This latter finding gives an estimate of the accuracy of *KIR3DL1/S1* allele imputation for individuals of unknown genetic ancestry. Using the Global model, the imputation time ranged from 2 min 9 s for AMR (N = 146) to 4 min 33 s for SAS (N = 228), and it took 16 min 1 s to impute the Global test set of 1,007 individuals (S4 Table).

We next evaluated the specificity and sensitivity of the Global imputation model. The mean specificity across the 42 alleles was 99%, with only two alleles having a specificity below 99% (Fig 3C). The lowest specificities were observed for *3DS1*01301* at 96% and *3DL1*01502* at 98.5%. Of the individuals falsely imputed as having *3DS1*01301*, 84% were due to a *KIR3DL1/S1* deletion haplotype. This finding is consistent with the suggestion that the parental haplotype for the deletion carried *3DS1*01301* [92]. The individuals falsely imputed as having *3DL1*01502*, possessed either *3DL1*01702*, **051* or **025* (33% each). All these alleles fall into the same ancestral lineage as *3DL1*015*, and likely exhibit similar phenotypes of high expression and ligand binding [1]. In the final Global population group (2N = 4,068) there were 15 *KIR3DL1/S1* alleles with a frequency above 1% and 27 alleles with a frequency below 1%

Table 1. Number of *KIR3DL1/S1* alleles and individuals in data sets.

1000 Genomes Population Group	Number of Individuals in Data Set			
	All	Global <i>KIR3DL1/S1</i> MAC < 3	In Model Set	In Test Set
Africa (AFR)	558	541	272	269
Americas (AMR)	298	292	146	146
East Asia (EAS)	406	389	196	193
Europe (EUR)	353	345	174	171
South Asia (SAS)	467	457	229	228
Global	2,082	2,024	1,017	1,007

<https://doi.org/10.1371/journal.pcbi.1009059.t001>

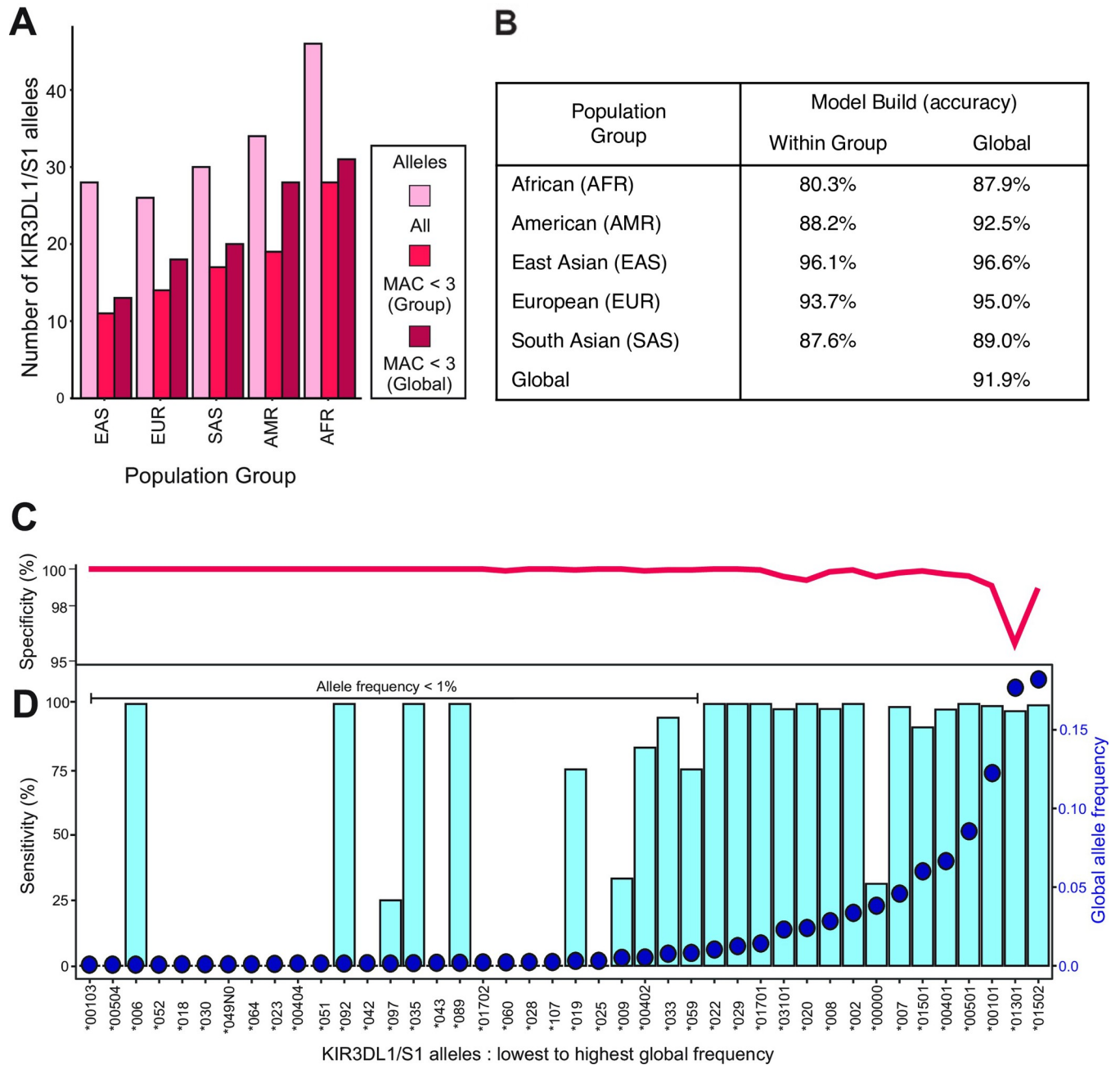


Fig 3. Accurate imputation of *KIR3DL1/S1* alleles using a Global population model. **A.** Bar graphs shows the number of *KIR3DL1/S1* alleles present in each of the five broad population groups of the 1,000 Genomes database. The bar colors indicate: (pink) the number of alleles present before filtering, (ruby) by MAC < 3 filtering, and (burgundy) by combining the five groups to form a Global population and then MAC < 3 filtering. The population groups are East Asian (EAS), European (EUR), South Asian (SAS), American (AMR) and African (AFR). **B.** Shows the imputation accuracy obtained for each of the population group and the Global models. (Within group) the model was built using 50% of the indicated group and tested on the other 50%. (Global) the model was built using 50% of all individuals and tested on the remaining 50% of the specified group. **C.** and **D.** Show the imputation efficacy for each allele present in the final Global data set. Blue bars indicate the sensitivity (% of times a given allele was called as present when known to be present). Red line indicates specificity (% of times a given allele was called as absent when known to be absent). Blue dots indicate the *KIR3DL1/S1* allele frequencies in the Global population.

<https://doi.org/10.1371/journal.pcbi.1009059.g003>

(Fig 3C). For those *KIR3DL1/S1* alleles having allele frequency below 1%, we observed a mean sensitivity of 29%. Despite a frequency of less than 1%, the alleles *006, *092, *035 and *089 were imputed with 100% sensitivity (Fig 3C). Conversely, *KIR3DL1/S1* alleles with a frequency

above 1% had a mean sensitivity of 95%. The mean sensitivity rose to 99% when the allele representing the absence of *KIR3DL1/S1* was excluded.

In total 27 *KIR3DL1/S1* alleles had a frequency less than 1% in the Global population. When the global frequency was above 1%, PONG was able to impute the alleles 91–100% of the time (Fig 3D). Thus, similar to the model built using the European population group, low frequency alleles were more likely to be incorrectly imputed than high frequency alleles. An exception to this was the allele representing the absence of *KIR3DL1/S1* (*00000) in which the frequency was 4% but PONG was only able to impute the absent allele correctly 35% of the time. Together these results show that PONG is effective for imputing common *KIR3DL1/S1* alleles across a diverse set of human populations, as well as some rare alleles. The AFR, AMR and SAS population groups had the highest number of *KIR3DL1/S1* alleles of frequency less than 1% in the test set (16, 10 and 10, respectively). Contributing to the lower accuracy of PONG in AFR and SAS, these low-frequency *KIR* alleles accounted for over 50% of those present in each group. By contrast, only three such low frequency alleles were present in the EAS group. In summary, the accuracy of PONG is affected by the frequency of *KIR3DL1/S1* alleles and is therefore less effective in more diverse human populations given a similar-sized training sample. Because we have filtered on minor allele count, the imputation accuracy in these populations will increase with the size of the model data set used.

Imputing *KIR3DL1/S1* alleles using low density genotyping datasets

We evaluated whether *KIR3DL1/S1* genotyping could be performed using data obtained from each of three commonly used low density arrays. For this test, we included the AFR and EAS groups because they represent the highest and lowest genetic diversity, respectively, of the five major population groups in the 1,000 Genomes database. We based this test on the Global model and extended the window from which classifiers are sampled to chr19: 55,100,000–55,500,000 (hg19). To ensure continuity, we used the same individuals as the Global model development described above. Using the SNP genotype data directly produced poor results for each of these chips (Table 2). Thus, we first supplemented the data by imputing Illumina Omni 2.5 array SNPs. Accuracy in the AFR population improved to 78%, 80% and 84% when the genotypes originated from Infinium, Immunochip or MEGA SNPs, respectively (Table 2). For the EAS group the accuracy was 81%, 89% and 92% for Infinium, Immunochip and MEGA arrays, respectively (Table 2). As expected, the accuracy was lower than that obtained directly from the high-density Illumina Omni 2.5 data (88% AFR and 97% EAS, Fig 3). That we observed little difference for the EUR group between low- and high-density arrays, with a minimum of 91% accuracy, may be due to original ascertainment bias in chip design.

Table 2. Accuracy of PONG across SNP arrays.

Test Set Parameters		Accuracy (% of Genotypes Imputed Correctly)			
Population	SNP Imputation ⁺	Illumina Omni 2.5	Infinium	Immunochip	MEGA
EUR	no	94%	49%	55%	58%
	yes	–	91%	91%	92%
EAS	no	96%	61%	72%	78%
	yes	–	81%	89%	92%
AFR	no	88%	28%	54%	37%
	yes	–	78%	80%	84%

Genome window chr19: 55,100,000–55,500,000 (hg19).

⁺ Model and test sets supplemented with imputed Illumina Omni 2.5 genotypes.

<https://doi.org/10.1371/journal.pcbi.1009059.t002>

Consistent with this reasoning, the MEGA chip, which was designed to reduce such bias [93], gave the greatest accuracy through this test. In summary, these results show it is tractable to use PONG with low-density arrays when higher-density array genotypes are first imputed from the starting data.

Testing PONG using an independent data set

We analyzed a cohort of 397 individuals from Norway, from whom we generated both Immunochip SNP and high-resolution *KIR3DL1/S1* allele sequence data. We observed a strong correlation between allele frequencies in Norway and the 1,000 Genomes EUR group ($r = 0.96$). In total, 18 *KIR3DL1/S1* alleles were identified in the Norwegian cohort through nucleotide sequencing. After filtering for $MAC < 3$, there were 13 *KIR3DL1/S1* alleles present. We built a model from this cohort and used it for the imputation of *KIR3DL1/S1* alleles. Imputation increased the number of variable SNPs from 26 to 715 in the *KIR* region, improving accuracy of *KIR3DL1/S1* allele imputation from 53% to 89%. We extended the window for model building to chr19: 55,100,000–55,500,000 (hg19), which contains 255 SNPs on the Immunochip. With the extended window we observed 92% accuracy, sensitivity of mode 100% and mean 75%, and specificity of mode 100%, mean 99% (Fig 4). As in previous analyses, *KIR3DL1/S1*

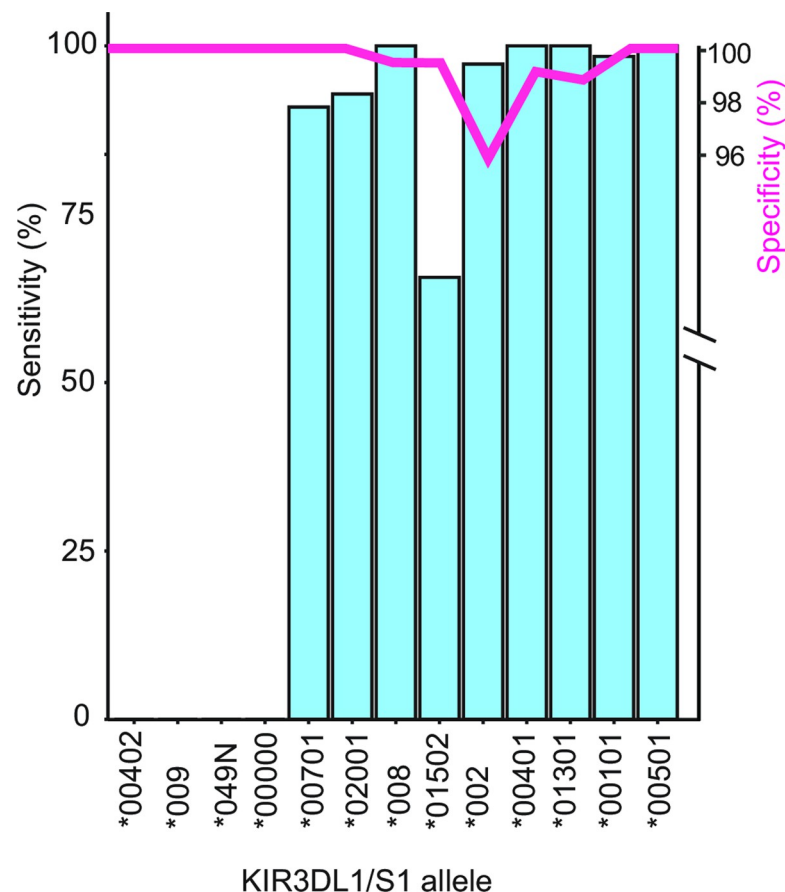


Fig 4. Accurate imputation of *KIR3DL1/S1* alleles from Immunochip SNP data. Bar graph shows the efficiency of *KIR3DL1/S1* allele imputation using a model built and tested on a cohort from Norway who also had their *KIR3DL1/S1* alleles genotyped to high resolution. Blue bars indicate the sensitivity (% of times a given allele was called as present when known to be present). Red line indicates specificity (% of times a given allele was called as absent when known to be absent).

<https://doi.org/10.1371/journal.pcbi.1009059.g004>

alleles having frequency $> 1\%$ have greater imputation accuracy than those $< 1\%$ (Fig 4). In this analysis, the mean sensitivity of alleles with a frequency of $< 1\%$ was 33%. By contrast, *KIR3DL1/S1* alleles with a frequency $> 1\%$ had a mean sensitivity of 85%. We note that, consistent with the 1,000 Genome EUR test above, imputation of SNPs in the extended window prior to model building did not increase the accuracy in this cohort. These experiments further demonstrate that high resolution *KIR3DL1/S1* genotypes can be imputed from low-density SNP arrays, and with similar accuracy to high-density arrays.

Running PONG

The PONG program is installed using the command line and opened as a library in R (R version 2.14.0–4.0.0.) [94]. PONG can be run using WG-SNP data mapped either to hg19 or hg38. The HIBAG imputation algorithm does not require data to be phased [75]. The Global model (hg19) and the model built with the EUR group (hg38) are available for download. Other models will be added as they become available. Using our Global model, we estimate that 1,000 individuals could be imputed every 15 minutes using a single core on a laboratory server, such as the one we have used. Users can also create their own models when WG-SNP data and *KIR3DL1/S1* allele genotypes are available, and modify the data input and filtering parameters, as described in the tutorial.

Discussion

Knowledge of *KIR3DL1/S1* diversity can help predict the course of specific infections, immune-mediated diseases, and their therapies [95–101]. However, by virtue of the polymorphic and structural complexity at the locus, it is often excluded from genome-wide association studies. The primary goal of this study was to develop a model trained to impute *KIR3DL1/S1* alleles rapidly from WG-SNP data encompassing a wide range of human genetic diversity. We built imputation models using high-density WG-SNP data [78] and high-resolution *KIR3DL1/S1* allele calls [79] from the five broadly defined 1,000 Genomes population groups, and then built a model for the Global group. To achieve these goals, we adapted the coding framework and algorithm from HIBAG [75] in a modification that we have named PONG. We determined that the imputation models are most effective when both the WG-SNP data and *KIR3DL1/S1* alleles have been filtered to remove alleles that occur infrequently. The former filter to reduce the model building run time and the latter to increase imputation accuracy. The resulting range of imputation accuracies of the final Global model was 88% for Africans and South Asians, to 97% for East Asians. The 1,000 Genomes WG-SNP data has a dense set of genotypes, including 1,832,506 SNPs from chromosome 19 [78]. Other genotype chips used for disease association studies have less dense sets of SNPs, including the Infinium Immunoarray, which targets markers associated with autoimmune disease and inflammatory disorders [80]. Because *KIR3DL1/S1* diversity is associated with development or severity of multiple autoimmune diseases [20, 22, 23, 25, 102], we tested the accuracy of imputation using results generated from this genotyping chip and achieved similar imputation accuracy as achieved from the high-density array.

Although PONG is effective in imputing *KIR3DL1/S1* alleles, there are a few limitations to this program that we are optimistic will improve over time. As observed for *HLA* [103], we found a negative correlation between the accuracy of PONG and the diversity of *KIR3DL1/S1* alleles in a population. For example, the African population group we studied has the highest number of distinct *KIR3DL1/S1* alleles as well as the highest number with an allele frequency below 1%. We also observed a high incidence of alleles at less than 1% in the South Asian group, as well as a higher frequency of haplotypes having duplicated *KIR3DL1/S1* compared

with Africans [92]. The result is that imputation accuracy is lowest in Africans and South Asians. Conversely, the East Asian group has the lowest number of *KIR3DL1/S1* alleles of allele frequency below 1%, and the highest imputation accuracy. Therefore, PONG is most effective at imputing the most frequent alleles. Finally, *KIR3DL1/S1* allelic diversity is under-characterized for many non-European origin populations. Any *KIR3DL1/S1* alleles newly identified in these populations will need to be incorporated into future models to increase imputation accuracy. The imputation accuracy of the model will improve over time as more immunogenetic studies of *KIR3DL1/S1* are conducted, thus expanding our sample set for building more robust and diverse models. Given that the model is open source, and that PONG has a model building function available, this can be achieved both by the developers and users.

PONG is less accurate imputing the absence of the *KIR3DL1/S1* gene (which we designated *00000), and as we were only able to impute this null *KIR3DL1/S1* allele at an accuracy of 35% using the Global model. One solution to improve the accuracy of imputing the 'KIR3DL1/S1 absent' allele may be to couple the use of PONG with the program KIR*IMP. KIR*IMP is targeted to *KIR* gene content diversity and can impute the presence or absence of *KIR3DL1/S1* with an accuracy above 90% [87]. KIR*IMP does not require installation and is run through a web interface, via data upload. Accurate sequencing and assembly can be challenging for highly polymorphic and structurally diverse regions of the genome [104]. Both these phenomena are characteristics of the *KIR* locus [29]. Therefore, PONG relies on high quality WG-SNP data with robust quality control measures implemented in SNP calling pipelines. New techniques to improve the identification of structural variation are being created, including long-range optical mapping, which uses the optical signal strength from each SNP genotype to identify deletions and duplications [105]. Together, an increased sampling of individuals having rare *KIR3DL1/S1* alleles and better characterization of structural variation from WG-SNP data will likely improve the imputation accuracy of PONG. PONG is open source and gives users the ability to publish reference models that do not contain information about individuals, nor is the upload of genotype data to any public server required. These features both support reproducibility of research and increase efficiency of the imputation as the models are expanded.

Highly polymorphic interactions of KIR3DL1/S1 with HLA-A and -B modulate the critical functions of NK cells in immunity, which include the destruction of infected or cancerous cells [2]. Combinatorial diversity of KIR3DL1/S1 with HLA-A and -B allotypes thus affects the susceptibility and course of multiple immune-mediated diseases. Several methods are available to impute *HLA* alleles [72–75], but large-scale genetic studies often exclude analysis of *KIR3DL1/S1* due to the exceptional polymorphism and structural diversity of the genomic region. A secondary goal of this study was thus to produce imputation models that could be used in conjunction with existing models to impute the combinatorial diversity of KIR3DL1/S1 and *HLA* allotypes. By comparison with *KIR3DL1/S1*, the mean imputation accuracy for HIBAG across seven *HLA* genes was 81.2% in African populations and 91.1% in East Asians [75]. In African populations, *HLA-DPB1* had the lowest imputation accuracy at 74.2% and *HLA-A* had the highest observed accuracy at 92.4%. The corresponding imputation accuracies of these *HLA* genes in East Asians were 89.8% and 92.1% respectively. Whereas HIBAG did not originally include South Asians, subsequent studies in a population from India found an imputation accuracy of 88% for *HLA-B* rising to 94% for other *HLA* genes [90]. The mean accuracy of *KIR3DL1/S1* allele imputation described herein is equivalent, and often better than that obtained for *HLA class I* and *II* using the same underlying algorithm. We therefore propose that using this algorithm to impute both *KIR3DL1/S1* and *HLA-A* and *-B* genotypes from WG-SNP data presents a considerable advantage over other approaches. This approach is particularly applicable for studies of Biobank data, where targeted sequencing of *KIR3DL1/S1* and

HLA-A and *-B* from many thousands of individuals is not currently tractable. Utilizing our pre-built models, PONG can be implemented to make genetic association studies of *KIR3DL1/S1* in combination with *HLA-A* and *-B* possible at very large scale.

Due to their established associations with multiple human ailments including infectious disease, cancer and autoimmunity [1], we focused this study on determining the allotypes of *KIR3DL1/S1*. Included among those with the greatest imputation accuracy are common allotypes having altered or no expression, specificity-determining polymorphism, and an activating variant [2]. Any of these allotypes can have dramatic consequence for the function of NK cells. Thus, knowledge of *KIR3DL1/S1* genotype, in combination with *HLA-A* and *-B* genotypes, is often sufficient to understand the disease mechanism or move the field forward in other ways [15, 51, 102, 106]. In other cases, it will be necessary to form a complete picture of KIR interactions with HLA-C in addition to *-A* and *-B*, by generating a genotype from the entire *KIR* locus [29]. Validation of PONG to impute *KIR3DL1/S1* is a robust proof of concept that can be built upon to impute entire *KIR* region genotypes. Expanding to other *KIR* genes will be challenging as the genomic region contains substantial structural variation and, unlike *HLA*, classifiers are sampled from the entire locus and flanking parts, rather than individual genes. We conducted rigorous testing to determine the efficacy and limits of HIBAG in imputing *KIR3DL1/S1* genotypes. This was a first but important step before expanding the concept in future work to impute complete *KIR* haplotypes.

Supporting information

S1 Table. *KIR3DL1/S1* Genotypes of 1,000 Genomes Individuals. This table contains the *KIR3DL1/S1* alleles corresponding to each of the individuals present in the 1000 Genomes database that were used in this study. The population designation and sample ID is present for each.

(XLSX)

S2 Table. Summary of the 1000 Genomes populations utilized in this study. This table shows the number of individuals from each sub-population used to develop PONG, arranged by population group.

(XLSX)

S3 Table. *KIR3DL1/S1* Genotypes of Norwegian Individuals. This table shows the *KIR3DL1/S1* genotypes used for testing the independent data set from Norway.

(XLSX)

S4 Table. Parameters and Output statistics of Imputation Models. This table provides a summary of the results obtained through each experiment conducted to develop the optimal parameters for PONG model building, as well as testing the accuracy. The filtering parameters, populations used for both model building and testing, genotype array, accuracy, and time for model building and imputation are all available in this table.

(XLSX)

S5 Table. Filtering parameter optimization for Imputation Models. This table shows the percent accuracy of PONG for each of the main population groups as a function of distinct SNP and *KIR3DL1/S1* filtering parameters.

(XLSX)

Author Contributions

Conceptualization: Damjan Vukcevic, Stephen Leslie, Paul J. Norman.

Data curation: Genelle F. Harrison, Laura Ann Leaton, Katherine M. Kichula, Marte K. Viken, Paul J. Norman.

Formal analysis: Genelle F. Harrison, Laura Ann Leaton.

Funding acquisition: Paul J. Norman.

Investigation: Genelle F. Harrison, Laura Ann Leaton, Paul J. Norman.

Methodology: Genelle F. Harrison, Laura Ann Leaton, Jonathan Shortt, Christopher R. Gignoux, Paul J. Norman.

Project administration: Paul J. Norman.

Resources: Marte K. Viken, Christopher R. Gignoux, Benedicte A. Lie, Paul J. Norman.

Software: Genelle F. Harrison, Laura Ann Leaton, Erica A. Harrison, Jonathan Shortt.

Supervision: Paul J. Norman.

Validation: Genelle F. Harrison, Laura Ann Leaton, Erica A. Harrison.

Visualization: Genelle F. Harrison, Paul J. Norman.

Writing – original draft: Genelle F. Harrison, Laura Ann Leaton, Paul J. Norman.

Writing – review & editing: Genelle F. Harrison, Laura Ann Leaton, Erica A. Harrison, Katherine M. Kichula, Marte K. Viken, Jonathan Shortt, Christopher R. Gignoux, Benedicte A. Lie, Damjan Vukcevic, Stephen Leslie, Paul J. Norman.

References

1. Parham P, Norman PJ, Abi-Rached L, Guethlein LA. Variable NK cell receptors exemplified by human KIR3DL1/S1. *The Journal of Immunology*. 2011; 187(1):11–9. <https://doi.org/10.4049/jimmunol.0902332> PMID: 21690332
2. O'Connor GM, McVicar D. The yin-yang of KIR3DL1/S1: molecular mechanisms and cellular function. *Critical reviews in immunology*. 2013; 33(3):203–18. <https://doi.org/10.1615/critrevimmunol.2013007409> PMID: 23756244
3. Quatrini L, Chiesa MD, Sivori S, Mingari MC, Pende D, Moretta L. Human NK cells, their receptors and function. *European journal of immunology*. 2021 Apr 26. <https://doi.org/10.1002/eji.202049028> PMID: 33899224
4. Hammer Q, Rückert T, Romagnani C. Natural killer cell specificity for viral infections. *Nature immunology*. 2018; 19(8):800–8. <https://doi.org/10.1038/s41590-018-0163-6> PMID: 30026479
5. Litwin V, Gumperz J, Parham P, Phillips JH, Lanier LL. NKB1: a natural killer cell receptor involved in the recognition of polymorphic HLA-B molecules. *The Journal of experimental medicine*. 1994 Aug 1; 180(2):537–43. <https://doi.org/10.1084/jem.180.2.537> PMID: 8046332
6. Colonna M, Samaridis J. Cloning of immunoglobulin-superfamily members associated with HLA-C and HLA-B recognition by human natural killer cells. *Science (New York, NY)*. 1995 Apr 21; 268(5209):405–8. <https://doi.org/10.1126/science.7716543> PMID: 7716543
7. Garcia-Beltran WF, Hölzemer A, Martrus G, Chung AW, Pacheco Y, Simoneau CR, et al. Open conformers of HLA-F are high-affinity ligands of the activating NK-cell receptor KIR3DS1. *Nat Immunol*. 2016 Sep; 17(9):1067–74. <https://doi.org/10.1038/ni.3513> PMID: 27455421
8. Kiani Z, Bruneau J, Geraghty DE, Bernard NF. HLA-F on Autologous HIV-Infected Cells Activates Primary NK Cells Expressing the Activating Killer Immunoglobulin-Like Receptor KIR3DS1. *Journal of virology*. 2019 Sep 15; 93(18). <https://doi.org/10.1128/JVI.00933-19> PMID: 31270222
9. O'Connor GM, Vivian JP, Gostick E, Pymm P, Lafont BA, Price DA, et al. Peptide-Dependent Recognition of HLA-B*57:01 by KIR3DS1. *Journal of virology*. 2015 May; 89(10):5213–21. <https://doi.org/10.1128/JVI.03586-14> PMID: 25740999
10. Ahlenstiel G, Martin MP, Gao X, Carrington M, Rehermann B. Distinct KIR/HLA compound genotypes affect the kinetics of human antiviral natural killer cell responses. *The Journal of clinical investigation*. 2008 Mar; 118(3):1017–26. <https://doi.org/10.1172/JCI32400> PMID: 18246204

11. Digitale JC, Callaway PC, Martin M, Nelson G, Viard M, Rek J, et al. Inhibitory KIR ligands are associated with higher *P. falciparum* parasite prevalence. *The Journal of infectious diseases*. 2020 Nov 9.
12. Forlenza CJ, Boudreau JE, Zheng J, Le Ludeuc JB, Chamberlain E, Heller G, et al. KIR3DL1 Allelic Polymorphism and HLA-B Epitopes Modulate Response to Anti-GD2 Monoclonal Antibody in Patients With Neuroblastoma. *Journal of clinical oncology: official journal of the American Society of Clinical Oncology*. 2016 Jul 20; 34(21):2443–51. <https://doi.org/10.1200/JCO.2015.64.9558> PMID: 27069083
13. López-Vázquez A, Rodrigo L, Martínez-Borra J, Pérez R, Rodríguez M, Fdez-Morera JL, et al. Protective effect of the HLA-Bw4180 epitope and the killer cell immunoglobulin-like receptor 3DS1 gene against the development of hepatocellular carcinoma in patients with hepatitis C virus infection. *The Journal of infectious diseases*. 2005 Jul 1; 192(1):162–5. <https://doi.org/10.1086/430351> PMID: 15942906
14. MacFarlane AWt, Jillab M, Smith MR, Alpaugh RK, Cole ME, Litwin S, et al. NK cell dysfunction in chronic lymphocytic leukemia is associated with loss of the mature cells expressing inhibitory killer cell Ig-like receptors. *Oncoimmunology*. 2017; 6(7):e1330235. <https://doi.org/10.1080/2162402X.2017.1330235> PMID: 28811973
15. 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. *Nature Genetics*. 2007 2007/06/01; 39(6):733–40. <https://doi.org/10.1038/ng2035> PMID: 17496894
16. Ruggeri L, Mancusi A, Urbani E, Velardi A. Identifying NK Alloreactive Donors for Haploidentical Hematopoietic Stem Cell Transplantation. *Methods in molecular biology (Clifton, NJ)*. 2016; 1393:141–5.
17. Trefny MP, Rothschild SI, Uhlenbrock F, Rieder D, Kasenda B, Stanczak MA, et al. A Variant of a Killer Cell Immunoglobulin-like Receptor Is Associated with Resistance to PD-1 Blockade in Lung Cancer. *Clinical cancer research: an official journal of the American Association for Cancer Research*. 2019 May 15; 25(10):3026–34. <https://doi.org/10.1158/1078-0432.CCR-18-3041> PMID: 30765392
18. Greenland JR, Sun H, Calabrese D, Chong T, Singer JP, Kukreja J, et al. HLA Mismatching Favoring Host-Versus-Graft NK Cell Activity Via KIR3DL1 Is Associated With Improved Outcomes Following Lung Transplantation. *American journal of transplantation: official journal of the American Society of Transplantation and the American Society of Transplant Surgeons*. 2017 Aug; 17(8):2192–9. <https://doi.org/10.1111/ajt.14295> PMID: 28375571
19. van der Ploeg K, Le Ludeuc JB, Stevenson PA, Park S, Gooley TA, Petersdorf EW, et al. HLA-A alleles influencing NK cell function impact AML relapse following allogeneic hematopoietic cell transplantation. *Blood advances*. 2020 Oct 13; 4(19):4955–64. <https://doi.org/10.1182/bloodadvances.2020002086> PMID: 33049053
20. Ahn RS, Moslehi H, Martin MP, Abad-Santos M, Bowcock AM, Carrington M, et al. Inhibitory KIR3DL1 alleles are associated with psoriasis. *The British journal of dermatology*. 2016 Feb; 174(2):449–51. <https://doi.org/10.1111/bjd.14081> PMID: 26286807
21. 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 Jul 24. <https://doi.org/10.4049/jimmunol.2000144> PMID: 32709660
22. 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(7):e39991. <https://doi.org/10.1371/journal.pone.0039991> PMID: 22768326
23. Lorentzen AR, Karlsen TH, Olsson M, Smestad C, Mero IL, Woldseth B, et al. Killer immunoglobulin-like receptor ligand HLA-Bw4 protects against multiple sclerosis. *Annals of neurology*. 2009 Jun; 65(6):658–66. <https://doi.org/10.1002/ana.21695> PMID: 19630074
24. Petrushkin H, Norman PJ, Lougee E, Parham P, Wallace GR, Stanford MR, et al. KIR3DL1/S1 Allotypes Contribute Differentially to the Development of Behçet Disease. *The Journal of Immunology*. 2019; 203(6):1629–35. <https://doi.org/10.4049/jimmunol.1801178> PMID: 31405953
25. Vendelbosch S, Heslinga SC, John M, van Leeuwen K, Geissler J, de Boer M, et al. Study on the protective effect of the KIR3DL1 gene in ankylosing spondylitis. *Arthritis & rheumatology (Hoboken, NJ)*. 2015 Nov; 67(11):2957–65.
26. Yawata N, Shirane M, Woon K, Lim X, Tanaka H, Kawano YI, et al. Molecular Signatures of Natural Killer Cells in CMV-Associated Anterior Uveitis, A New Type of CMV-Induced Disease in Immunocompetent Individuals. *International journal of molecular sciences*. 2021 Mar 31; 22(7). <https://doi.org/10.3390/ijms22073623> PMID: 33807229
27. Erbe AK, Wang W, Carmichael L, Hoefges A, Grzywacz B, Reville PK, et al. Follicular lymphoma patients with KIR2DL2 and KIR3DL1 and their ligands (HLA-C1 and HLA-Bw4) show improved

- outcome when receiving rituximab. *Journal for immunotherapy of cancer*. 2019 Mar 12; 7(1):70. <https://doi.org/10.1186/s40425-019-0538-8> PMID: 30871628
28. Shaffer BC, Hsu KC. Selection of allogeneic hematopoietic cell transplant donors to optimize natural killer cell alloreactivity. *Seminars in hematology*. 2020 Oct; 57(4):167–74. <https://doi.org/10.1053/j.seminhematol.2020.10.005> PMID: 33256909
 29. Béziat V, Hilton HG, Norman PJ, Traherne JA. Deciphering the killer-cell immunoglobulin-like receptor system at super-resolution for natural killer and T-cell biology. *Immunology*. 2017; 150(3):248–64. <https://doi.org/10.1111/imm.12684> PMID: 27779741
 30. Robinson J, Halliwell JA, Hayhurst JD, Flicek P, Parham P, Marsh SG. The IPD and IMGT/HLA database: allele variant databases. *Nucleic Acids Res*. 2015 Jan; 43(Database issue):D423–31. <https://doi.org/10.1093/nar/gku1161> PMID: 25414341
 31. 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 Sep; 39(9):1092–9. <https://doi.org/10.1038/ng2111> PMID: 17694054
 32. Nunes K, Maia MHT, Dos Santos EJM, Dos Santos SEB, Guerreiro JF, Petzl-Erler ML, et al. How natural selection shapes genetic differentiation in the MHC region: A case study with Native Americans. *Human immunology*. 2021 Mar 31. <https://doi.org/10.1016/j.humimm.2021.03.005> PMID: 33812704
 33. Deng Z, Zhen J, Harrison GF, Zhang G, Chen R, Sun G, et al. Adaptive Admixture of HLA class I Alleotypes Enhanced Genetically Determined Strength of Natural Killer Cells in East Asians. *Molecular biology and evolution*. 2021 Feb 22.
 34. Koyro TF, Kraus E, Lunemann S, Hölzemer A, Wulf S, Jung J, et al. Upregulation of HLA-F expression by BK polyomavirus infection induces immune recognition by KIR3DS1-positive natural killer cells. *Kidney international*. 2020 Dec 23. <https://doi.org/10.1016/j.kint.2020.12.014> PMID: 33359499
 35. Townsley E, O'Connor G, Cosgrove C, Woda M, Co M, Thomas SJ, et al. Interaction of a dengue virus NS1-derived peptide with the inhibitory receptor KIR3DL1 on natural killer cells. *Clinical and experimental immunology*. 2016 Mar; 183(3):419–30. <https://doi.org/10.1111/cei.12722> PMID: 26439909
 36. Beltrame LM, Sell AM, Moliterno RA, Clementino SL, Cardozo DM, Dalalio MM, et al. Influence of KIR genes and their HLA ligands in susceptibility to dengue in a population from southern Brazil. *Tissue antigens*. 2013 Dec; 82(6):397–404. <https://doi.org/10.1111/tan.12256> PMID: 24498996
 37. Carrington M, Wang S, Martin MP, Gao X, Schiffman M, Cheng J, et al. Hierarchy of resistance to cervical neoplasia mediated by combinations of killer immunoglobulin-like receptor and human leukocyte antigen loci. *The Journal of experimental medicine*. 2005 Apr 4; 201(7):1069–75. <https://doi.org/10.1084/jem.20042158> PMID: 15809352
 38. Luo M, Czarnecki C, Nebroski M, Kimani J, Bernard N, Plummer FA. KIR3DL1 alleles and their epistatic interactions with human leukocyte antigen class I influence resistance and susceptibility to HIV-1 acquisition in the Pumwani sex worker cohort. *AIDS (London, England)*. 2018 Apr 24; 32(7):841–50. <https://doi.org/10.1097/QAD.0000000000001735> PMID: 29280757
 39. Hollenbach JA, Pando MJ, Caillier SJ, Gourraud PA, Oksenberg JR. The killer immunoglobulin-like receptor KIR3DL1 in combination with HLA-Bw4 is protective against multiple sclerosis in African Americans. *Genes Immun*. 2016 Apr; 17(3):199–202. <https://doi.org/10.1038/gene.2016.5> PMID: 26866467
 40. Saito H, Hirayama A, Umemura T, Joshita S, Mukawa K, Suga T, et al. Association between KIR-HLA combination and ulcerative colitis and Crohn's disease in a Japanese population. *PLoS One*. 2018; 13(4):e0195778. <https://doi.org/10.1371/journal.pone.0195778> PMID: 29649328
 41. Umemura T, Joshita S, Saito H, Yoshizawa K, Norman GL, Tanaka E, et al. KIR/HLA genotypes confer susceptibility and progression in patients with autoimmune hepatitis. *JHEP reports: innovation in hepatology*. 2019 Nov; 1(5):353–60. <https://doi.org/10.1016/j.jhepr.2019.09.003> PMID: 32039386
 42. Deng Z, Zhao J, Cai S, Qi Y, Yu Q, Martin MP, et al. Natural Killer Cells Offer Differential Protection From Leukemia in Chinese Southern Han. *Frontiers in immunology*. 2019; 10:1646. <https://doi.org/10.3389/fimmu.2019.01646> PMID: 31379844
 43. Fadda L, O'Connor GM, Kumar S, Piechocka-Trocha A, Gardiner CM, Carrington M, et al. Common HIV-1 peptide variants mediate differential binding of KIR3DL1 to HLA-Bw4 molecules. *Journal of virology*. 2011 Jun; 85(12):5970–4. <https://doi.org/10.1128/JVI.00412-11> PMID: 21471246
 44. Gooneratne SL, Richard J, Lee WS, Finzi A, Kent SJ, Parsons MS. Slaying the Trojan horse: natural killer cells exhibit robust anti-HIV-1 antibody-dependent activation and cytolysis against allogeneic T cells. *Journal of virology*. 2015 Jan; 89(1):97–109. <https://doi.org/10.1128/JVI.02461-14> PMID: 25320293
 45. Song R, Lisovsky I, Lebouché B, Routy JP, Bruneau J, Bernard NF. HIV protective KIR3DL1/S1-HLA-B genotypes influence NK cell-mediated inhibition of HIV replication in autologous CD4 targets. *PLoS*

- pathogens. 2014 Jan; 10(1):e1003867. <https://doi.org/10.1371/journal.ppat.1003867> PMID: 24453969
46. Koehler RN, Alter G, Tovanabutra S, Saathoff E, Arroyo MA, Walsh AM, et al. Natural killer cell-mediated innate sieve effect on HIV-1: the impact of KIR/HLA polymorphism on HIV-1 subtype-specific acquisition in east Africa. *The Journal of infectious diseases*. 2013 Oct 15; 208(8):1250–4. <https://doi.org/10.1093/infdis/jit349> PMID: 23922366
 47. Ayello J, Hochberg J, Flower A, Chu Y, Baxi LV, Quish W, et al. Genetically re-engineered K562 cells significantly expand and functionally activate cord blood natural killer cells: Potential for adoptive cellular immunotherapy. *Experimental hematology*. 2017; 46:38–47. <https://doi.org/10.1016/j.exphem.2016.10.003> PMID: 27765614
 48. Makanga DR, Jullien M, David G, Legrand N, Willem C, Dubreuil L, et al. Low number of KIR ligands in lymphoma patients favors a good rituximab-dependent NK cell response. *Oncoimmunology*. 2021; 10(1):1936392. <https://doi.org/10.1080/2162402X.2021.1936392> PMID: 34178429
 49. Prakash S, Sarangi AN, Alam S, Sonawane A, Sharma RK, Agrawal S. Putative role of KIR3DL1/3DS1 alleles and HLA-Bw4 ligands with end stage renal disease and long term renal allograft survival. *Gene*. 2017 Dec 30; 637:219–29. <https://doi.org/10.1016/j.gene.2017.09.033> PMID: 28942035
 50. Stern M, Hadaya K, Hönger G, Martin PY, Steiger J, Hess C, et al. Telomeric rather than centromeric activating KIR genes protect from cytomegalovirus infection after kidney transplantation. *American journal of transplantation: official journal of the American Society of Transplantation and the American Society of Transplant Surgeons*. 2011 Jun; 11(6):1302–7. <https://doi.org/10.1111/j.1600-6143.2011.03516.x> PMID: 21486386
 51. Venstrom JM, Gooley TA, Spellman S, Pring J, Malkki M, Dupont B, et al. Donor activating KIR3DS1 is associated with decreased acute GVHD in unrelated allogeneic hematopoietic stem cell transplantation. *Blood*. 2010 Apr 15; 115(15):3162–5. <https://doi.org/10.1182/blood-2009-08-236943> PMID: 20124216
 52. Foley BA, De Santis D, Van Beelen E, Lathbury LJ, Christiansen FT, Witt CS. The reactivity of Bw4+ HLA-B and HLA-A alleles with KIR3DL1: implications for patient and donor suitability for haploidentical stem cell transplantations. *Blood*. 2008 Jul 15; 112(2):435–43. <https://doi.org/10.1182/blood-2008-01-132902> PMID: 18385451
 53. 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 Oct 23; 479(7373):401–5. <https://doi.org/10.1038/nature10517> PMID: 22020283
 54. Boudreau JE, Hsu KC. Natural killer cell education in human health and disease. *Current opinion in immunology*. 2018; 50:102–11. <https://doi.org/10.1016/j.coi.2017.11.003> PMID: 29413815
 55. Malmberg KJ, Sohlberg E, Goodridge JP, Ljunggren HG. Immune selection during tumor checkpoint inhibition therapy paves way for NK-cell "missing self" recognition. *Immunogenetics*. 2017 Aug; 69(8–9):547–56. <https://doi.org/10.1007/s00251-017-1011-9> PMID: 28699110
 56. Parsons MS, Zipperlen K, Gallant M, Grant M. Killer cell immunoglobulin-like receptor 3DL1 licenses CD16-mediated effector functions of natural killer cells. *Journal of leukocyte biology*. 2010 Nov; 88(5):905–12. <https://doi.org/10.1189/jlb.1009687> PMID: 20664023
 57. Saunders PM, MacLachlan BJ, Widjaja J, Wong SC, Oates CVL, Rossjohn J, et al. The Role of the HLA Class I $\alpha 2$ Helix in Determining Ligand Hierarchy for the Killer Cell Ig-like Receptor 3DL1. *J Immunol*. 2021 Feb 15; 206(4):849–60. <https://doi.org/10.4049/jimmunol.2001109> PMID: 33441440
 58. Saunders PM, Pymm P, Pietra G, Hughes VA, Hitchen C, O'Connor GM, et al. Killer cell immunoglobulin-like receptor 3DL1 polymorphism defines distinct hierarchies of HLA class I recognition. *The Journal of experimental medicine*. 2016 May 2; 213(5):791–807. <https://doi.org/10.1084/jem.20152023> PMID: 27045007
 59. Kim S, Sunwoo JB, Yang L, Choi T, Song YJ, French AR, et al. HLA alleles determine differences in human natural killer cell responsiveness and potency. *Proceedings of the National Academy of Sciences of the United States of America*. 2008 Feb 26; 105(8):3053–8. <https://doi.org/10.1073/pnas.0712229105> PMID: 18287063
 60. Wirth T, Hildebrand F, Allix-Béguec C, Wölbeling F, Kubica T, Kremer K, et al. Origin, spread and demography of the *Mycobacterium tuberculosis* complex. *PLoS pathogens*. 2008; 4(9):e1000160. <https://doi.org/10.1371/journal.ppat.1000160> PMID: 18802459
 61. Yawata M, Yawata N, Draghi M, Little AM, Partheniou F, Parham P. Roles for HLA and KIR polymorphisms in natural killer cell repertoire selection and modulation of effector function. *The Journal of experimental medicine*. 2006 Mar 20; 203(3):633–45. <https://doi.org/10.1084/jem.20051884> PMID: 16533882
 62. Carlomagno S, Falco M, Bono M, Alicata C, Garbarino L, Mazzocco M, et al. KIR3DS1-Mediated Recognition of HLA- \ast B51: Modulation of KIR3DS1 Responsiveness by Self HLA-B Allotypes and Effect on

- NK Cell Licensing. *Frontiers in immunology*. 2017; 8:581. <https://doi.org/10.3389/fimmu.2017.00581> PMID: 28603523
63. O'Connor GM, Yamada E, Rampersaud A, Thomas R, Carrington M, McVicar DW. Analysis of binding of KIR3DS1*014 to HLA suggests distinct evolutionary history of KIR3DS1. *J Immunol*. 2011 Sep 1; 187(5):2162–71. <https://doi.org/10.4049/jimmunol.1002906> PMID: 21804024
 64. Thananchai H, Gillespie G, Martin MP, Bashirova A, Yawata N, Yawata M, et al. Cutting Edge: Allele-specific and peptide-dependent interactions between KIR3DL1 and HLA-A and HLA-B. *J Immunol*. 2007 Jan 1; 178(1):33–7. <https://doi.org/10.4049/jimmunol.178.1.33> PMID: 17182537
 65. Körner C, Altfeld M. Role of KIR3DS1 in human diseases. *Frontiers in immunology*. 2012; 3:326. <https://doi.org/10.3389/fimmu.2012.00326> PMID: 23125843
 66. Kanya P, Boulet S, Tsoukas CM, Routy JP, Thomas R, Côté P, et al. Receptor-ligand requirements for increased NK cell polyfunctional potential in slow progressors infected with HIV-1 coexpressing KIR3DL1*h/*y and HLA-B*57. *Journal of virology*. 2011 Jun; 85(12):5949–60. <https://doi.org/10.1128/JVI.02652-10> PMID: 21471235
 67. Berinstein J, Pollock R, Pellett F, Thavaneswaran A, Chandran V, Gladman DD. Association of variably expressed KIR3dl1 alleles with psoriatic disease. *Clinical rheumatology*. 2017 Oct; 36(10):2261–6. <https://doi.org/10.1007/s10067-017-3784-5> PMID: 28801811
 68. Boudreau JE, Giglio F, Gooley TA, Stevenson PA, Le Ludeuc JB, Shaffer BC, et al. KIR3DL1/HLA-B Subtypes Govern Acute Myelogenous Leukemia Relapse After Hematopoietic Cell Transplantation. *Journal of clinical oncology: official journal of the American Society of Clinical Oncology*. 2017 Jul 10; 35(20):2268–78. <https://doi.org/10.1200/JCO.2016.70.7059> PMID: 28520526
 69. Boudreau JE, Mulrooney TJ, Le Ludie JB, Barker E, Hsu KC. KIR3DL1 and HLA-B Density and Binding Calibrate NK Education and Response to HIV. *J Immunol*. 2016 Apr 15; 196(8):3398–410. <https://doi.org/10.4049/jimmunol.1502469> PMID: 26962229
 70. Martin MP, Naranbhai V, Shea PR, Qi Y, Ramsuran V, Vince N, et al. Killer cell immunoglobulin-like receptor 3DL1 variation modifies HLA-B*57 protection against HIV-1. *The Journal of clinical investigation*. 2018 May 1; 128(5):1903–12. <https://doi.org/10.1172/JCI98463> PMID: 29461980
 71. Leslie S, Donnelly P, McVean G. A statistical method for predicting classical HLA alleles from SNP data. *American journal of human genetics*. 2008 Jan; 82(1):48–56. <https://doi.org/10.1016/j.ajhg.2007.09.001> PMID: 18179884
 72. Li SS, Wang H, Smith A, Zhang B, Zhang XC, Schoch G, et al. Predicting multiallelic genes using unphased and flanking single nucleotide polymorphisms. *Genetic epidemiology*. 2011 Feb; 35(2):85–92. <https://doi.org/10.1002/gepi.20549> PMID: 21254215
 73. Dilthey A, Leslie S, Moutsianas L, Shen J, Cox C, Nelson MR, et al. Multi-population classical HLA type imputation. *PLoS computational biology*. 2013; 9(2):e1002877. <https://doi.org/10.1371/journal.pcbi.1002877> PMID: 23459081
 74. Jia X, Han B, Onengut-Gumuscu S, Chen WM, Concannon PJ, Rich SS, et al. Imputing amino acid polymorphisms in human leukocyte antigens. *PLoS One*. 2013; 8(6):e64683. <https://doi.org/10.1371/journal.pone.0064683> PMID: 23762245
 75. Zheng X, Shen J, Cox C, Wakefield JC, Ehm MG, Nelson MR, et al. HIBAG—HLA genotype imputation with attribute bagging. *Pharmacogenomics J*. 2014 Apr; 14(2):192–200. <https://doi.org/10.1038/tpj.2013.18> PMID: 23712092
 76. Breiman L. Heuristics of instability and stabilization in model selection. *Annals of Statistics*. 1996; 24(6):2350–83.
 77. Marin WM, 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. *PLoS computational biology*. 2021 Aug; 17(8):e1008904. <https://doi.org/10.1371/journal.pcbi.1008904> PMID: 34339413
 78. Auton A, Abecasis GR, Altshuler DM, Durbin RM, Abecasis GR, Bentley DR, et al. A global reference for human genetic variation. *Nature*. 2015 2015/10/01; 526(7571):68–74. <https://doi.org/10.1038/nature15393> PMID: 26432245
 79. 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. *The American Journal of Human Genetics*. 2016; 99(2):375–91. <https://doi.org/10.1016/j.ajhg.2016.06.023> PMID: 27486779
 80. Trynka G, Hunt KA, Bockett NA, Romanos J, Mistry V, Szperl A, et al. Dense genotyping identifies and localizes multiple common and rare variant association signals in celiac disease. *Nature genetics*. 2011; 43(12):1193. <https://doi.org/10.1038/ng.998> PMID: 22057235

81. Lande A, Fluge Ø, Strand EB, Flåm ST, Sosa DD, Mella O, et al. Human Leukocyte Antigen alleles associated with Myalgic Encephalomyelitis/Chronic Fatigue Syndrome (ME/CFS). *Scientific reports*. 2020 Mar 24; 10(1):5267. <https://doi.org/10.1038/s41598-020-62157-x> PMID: 32210306
82. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *American journal of human genetics*. 2007 Sep; 81(3):559–75. <https://doi.org/10.1086/519795> PMID: 17701901
83. Amorim LM, Augusto DG, Nemat-Gorgani N, Montero-Martin G, Marin WM, Shams H, et al. High-Resolution Characterization of KIR Genes in a Large North American Cohort Reveals Novel Details of Structural and Sequence Diversity. *Front Immunol*. 2021; 12:674778. <https://doi.org/10.3389/fimmu.2021.674778> PMID: 34025673
84. 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(11):e47491. <https://doi.org/10.1371/journal.pone.0047491> PMID: 23139747
85. Vierra-Green C, Roe D, Jayaraman J, Trowsdale J, Traherne J, Kuang R, et al. Estimating KIR Haplotype Frequencies on a Cohort of 10,000 Individuals: A Comprehensive Study on Population Variations, Typing Resolutions, and Reference Haplotypes. *PLoS One*. 2016; 11(10):e0163973. <https://doi.org/10.1371/journal.pone.0163973> PMID: 27723813
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. *Frontiers in immunology*. 2018; 9:2843. <https://doi.org/10.3389/fimmu.2018.02843> PMID: 30564239
87. Vukcevic D, Traherne JA, Næss S, Ellinghaus E, Kamatani Y, Dilthey A, et al. Imputation of KIR Types from SNP Variation Data. *American journal of human genetics*. 2015 Oct 1; 97(4):593–607. <https://doi.org/10.1016/j.ajhg.2015.09.005> PMID: 26430804
88. Taliun D, Harris DN, Kessler MD, Carlson J, Szpiech ZA, Torres R, et al. Sequencing of 53,831 diverse genomes from the NHLBI TOPMed Program. *BioRxiv*. 2019:563866.
89. Loh P-R, Palamara PF, Price AL. Fast and accurate long-range phasing in a UK Biobank cohort. *Nature genetics*. 2016; 48(7):811–6. <https://doi.org/10.1038/ng.3571> PMID: 27270109
90. Degenhardt F, Wendorff M, Wittig M, Ellinghaus E, Datta LW, Schembri J, et al. Construction and benchmarking of a multi-ethnic reference panel for the imputation of HLA class I and II alleles. *Human molecular genetics*. 2019; 28(12):2078–92. <https://doi.org/10.1093/hmg/ddy443> PMID: 30590525
91. 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 Oct; 56(4):313–26. <https://doi.org/10.1034/j.1399-0039.2000.560403.x> PMID: 11098931
92. 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 research*. 2009 May; 19(5):757–69. <https://doi.org/10.1101/gr.085738.108> PMID: 19411600
93. Wojcik GL, Graff M, Nishimura KK, Tao R, Haessler J, Gignoux CR, et al. Genetic analyses of diverse populations improves discovery for complex traits. *Nature*. 2019 Jun; 570(7762):514–8. <https://doi.org/10.1038/s41586-019-1310-4> PMID: 31217584
94. Team RC. R: A language and environment for statistical computing. 2013.
95. Caocci G, Martino B, Greco M, Abruzzese E, Trawinska MM, Lai S, et al. Killer immunoglobulin-like receptors can predict TKI treatment-free remission in chronic myeloid leukemia patients. *Exp Hematol*. 2015 Dec; 43(12):1015–8.e1. <https://doi.org/10.1016/j.exphem.2015.08.004> PMID: 26306453
96. Erbe AK, Wang W, Reville PK, Carmichael L, Kim K, Mendonca EA, et al. HLA-Bw4-I-80 Isoform Differentially Influences Clinical Outcome As Compared to HLA-Bw4-T-80 and HLA-A-Bw4 Isoforms in Rituximab or Dinutuximab-Based Cancer Immunotherapy. *Frontiers in immunology*. 2017; 8:675. <https://doi.org/10.3389/fimmu.2017.00675> PMID: 28659916
97. Gabriel IH, Sergeant R, Szydlo R, Apperley JF, DeLavallade H, Alsuliman A, et al. Interaction between KIR3DS1 and HLA-Bw4 predicts for progression-free survival after autologous stem cell transplantation in patients with multiple myeloma. *Blood*. 2010 Sep 23; 116(12):2033–9. <https://doi.org/10.1182/blood-2010-03-273706> PMID: 20562327
98. Martin MP, Gao X, Lee JH, Nelson GW, Detels R, Goedert JJ, et al. Epistatic interaction between KIR3DS1 and HLA-B delays the progression to AIDS. *Nat Genet*. 2002 Aug; 31(4):429–34. <https://doi.org/10.1038/ng934> PMID: 12134147
99. Alicata C, Pende D, Meazza R, Canevali P, Loiacono F, Bertaina A, et al. Hematopoietic stem cell transplantation: Improving alloreactive Bw4 donor selection by genotyping codon 86 of KIR3DL1/S1. *European journal of immunology*. 2016 Jun; 46(6):1511–7. <https://doi.org/10.1002/eji.201546236> PMID: 26990677

100. Gagne K, Busson M, Bignon JD, Balère-Appert ML, Loiseau P, Dormoy A, et al. Donor KIR3DL1/3DS1 gene and recipient Bw4 KIR ligand as prognostic markers for outcome in unrelated hematopoietic stem cell transplantation. *Biology of blood and marrow transplantation: journal of the American Society for Blood and Marrow Transplantation*. 2009 Nov; 15(11):1366–75. <https://doi.org/10.1016/j.bbmt.2009.06.015> PMID: 19822295
101. Ureshino H, Shindo T, Kojima H, Kusunoki Y, Miyazaki Y, Tanaka H, et al. Allelic Polymorphisms of KIRs and HLAs Predict Favorable Responses to Tyrosine Kinase Inhibitors in CML. *Cancer immunology research*. 2018 Jun; 6(6):745–54. <https://doi.org/10.1158/2326-6066.CIR-17-0462> PMID: 29695383
102. Díaz-Peña R, Vidal-Castiñeira JR, Alonso-Arias R, Suarez-Alvarez B, Vicario JL, Solana R, et al. Association of the KIR3DS1* 013 and KIR3DL1* 004 alleles with susceptibility to ankylosing spondylitis. *Arthritis & Rheumatism*. 2010; 62(4):1000–6. <https://doi.org/10.1002/art.27332> PMID: 20131260
103. Pappas DJ, Lizee A, Paunic V, Beutner KR, Motyer A, Vukcevic D, et al. Significant variation between SNP-based HLA imputations in diverse populations: the last mile is the hardest. *Pharmacogenomics J*. 2018 May 22; 18(3):367–76. <https://doi.org/10.1038/tpj.2017.7> PMID: 28440342
104. Vollger MR, Dishuck PC, Sorensen M, Welch AE, Dang V, Dougherty ML, et al. Long-read sequence and assembly of segmental duplications. *Nature methods*. 2019; 16(1):88–94. <https://doi.org/10.1038/s41592-018-0236-3> PMID: 30559433
105. Mostovoy Y, Yilmaz F, Chow SK, Chu C, Lin C, Geiger EA, et al. Genome mapping resolves structural variation within segmental duplications associated with microdeletion/microduplication syndromes. *bioRxiv*. 2020.
106. Shaffer BC, Le Ludec JB, Park S, Devlin S, Archer A, Davis E, et al. Prospective KIR genotype evaluation of hematopoietic cell donors is feasible with potential to benefit patients with AML. *Blood advances*. 2021 Apr 13; 5(7):2003–11. <https://doi.org/10.1182/bloodadvances.2020002701> PMID: 33843984