PONG 2.0: Allele imputation for the killer cell immunoglobulin-like receptors

Suraju A Sadeeq[1], Laura Ann Leaton[1], Katherine M Kichula[1], Neus Font-Porterias[1], Nicholas R Pollock[1], Christopher E Collora[1], Erick C Castelli[2], Christopher R Gignoux[1], Paul J Norman[1,3]

Affiliations:[1]Department of Biomedical Informatics, University of Colorado School of Medicine, Aurora, CO 80045, USA, [2]Department of Pathology, School of Medicine, São Paulo State University (Unesp), Botucatu, Brazil. [3]Department of Immunology and Microbiology, University of Colorado School of Medicine, Aurora, CO 80045, USA.

Corresponding Author: Paul J. Norman, paul.norman@cuanschutz.edu

Abstract

Polymorphic killer cell immunoglobulin-like receptors (KIR) are vital for immune modulation. They are expressed by subsets of natural killer (NK) cells of innate immunity and T cells of adaptive immunity, and most KIR interact with highly polymorphic human leukocyte antigen (HLA) class I ligands. The genetic determinants of these interactions significantly impact infection, multiple immune-mediated diseases, cancers and immunotherapies, as well as placentation and tissue transplantation. Current methods for determining KIR genotypes require a resource-intensive sequencing and computational approach, limiting scalability.. Here, we adapted the HIBAG algorithm to develop models for imputing allele-level genotypes for the polymorphic KIR that have known interactions with HLA class I. We used SNP genotypes and matched KIR allele-level genotypes from the 1,000 Genomes Project (EUR = 187, AMR = 93, SAS = 102, AFR = 102, EAS = 102). Our imputation models demonstrate remarkable accuracy, with the lowest-performing model (KIR3DL2) achieving 99% overall accuracy, 79% mean sensitivity, and 99% mean specificity. The best-performing model (KIR2DS4) achieves 100% overall accuracy, sensitivity, and specificity across a multi-ethnic panel. We further enhanced performance for less-predictive KIR models by considering only alleles having frequencies > 1%. To facilitate immunogenetic research using biobank SNP data, we developed the imputation package, PONG2.0, for determining KIR allele-level genotypes with high accuracy.

Introduction

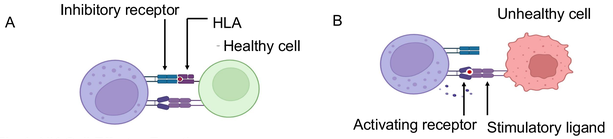

Fig 1. NK Cell Effector Function

Many Killer cell Immunoglobulin-like Receptors (KIR) interact with Human Leukocyte Antigens (HLA) class I ligands to regulate Natural Killer (NK) or T cell function, impacting both innate immunity and disease susceptibility. These interactions, mediated by activating and inhibitory KIRs (Fig. 1), balance immune defense and tolerance, influencing viral infection susceptibility, autoimmune risk, transplantation outcomes, and maternal-fetal tolerance [1–4].

Complete characterization of the combinatorial diversity and associated immunogenetic roles necessitates high-resolution genotyping for both HLA and KIR. The extreme sequence similarities among genes within the HLA or the KIR complex in tandem with exceptional polymorphism complicate accurate sequence read alignment. Whereas next-generation sequencing (NGS) along with specialized tools, enables precise imputation of gene copy numbers and allelic variants with remarkable accuracy [9–12], significant challenges remain for high scale KIR-HLA analysis. Principally, the substantial costs and extended processing times of NGS make it impractical for very large-scale studies or time-sensitive clinical applications[3,13,14]

We developed PONG2.0—an advanced multi-ancestry KIR imputation framework that achieves high accuracy while remaining computationally efficient. The PONG2.0 method was developed using a comprehensive multi-ancestry reference panel derived from the 1000 Genomes Project (1KGP) [20]. The freely available PONG2.0 R package includes pre-trained models to facilitate immunogenetic research.

Methods and Materials

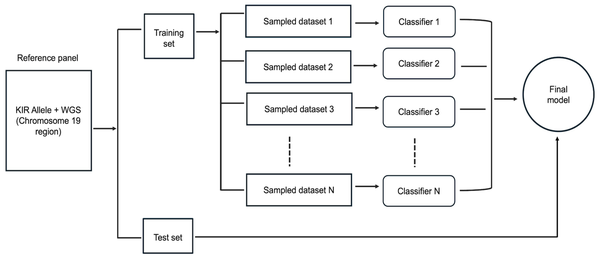

Fig 2. Development of the PONG2.0 KIR Imputation Framework.

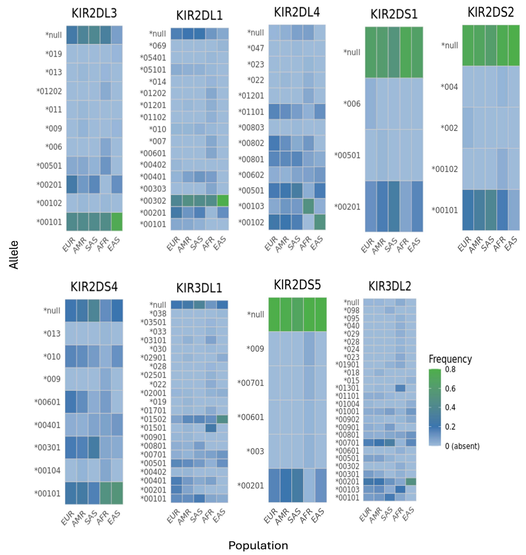

Fig 3. KIR Allele Frequency Distribution Across Populations.

Results

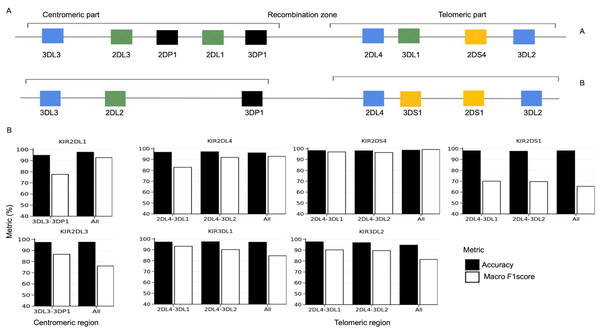

Fig 4. PONG2.0 KIR imputation performance.

(A) KIR genomic optimization strategy showing haplotype organization and gene categories. (B) Performance in centromeric and telomeric genes using accuracy (black) and macro F1 scores (white).

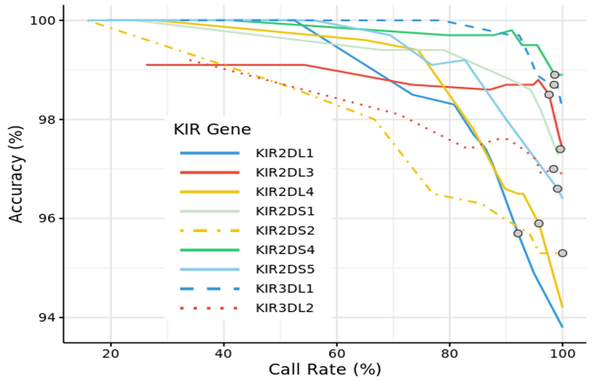

Fig 5. Precision-coverage tradeoff across probability thresholds.

Stringent thresholds (≥0.8) achieve >99% accuracy, while moderate thresholds (0.5–0.7) balance high call rates (90–95%). Grey circles (●) indicate the 0.5 threshold.

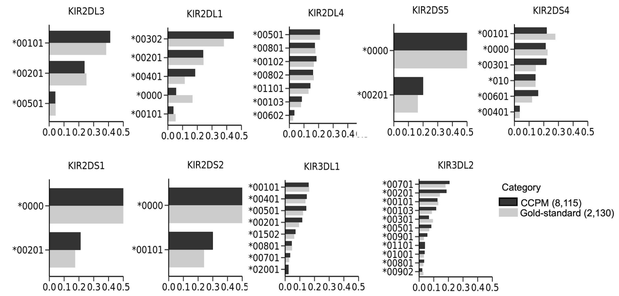

Fig 6. External validation of imputation accuracy.

Concordance of allele frequencies between imputed (CCPM, dark grey) and gold-standard (Amorim et al., 2021, light grey) data across KIR genes.

Discussion

Our findings demonstrate that optimal KIR imputation requires a nuanced, region-specific approach that accounts for the distinct genomic architectures of centromeric and telomeric gene clusters[42,43]. The success of our imputation method has important implications for clinical and population genetic studies, particularly for diseases involving specific KIR-HLA interactions, as it enables more reliable detection of KIR alleles across the highly polymorphic KIR region. Future work should explore ancestry-specific optimizations, as our current approach, while robust for multi-ancestral populations, may require adjustment for additional haplotype backgrounds. Similarly, this established framework can be extended beyond KIR to other challenging loci having similar genomic architectural features. Our validation results demonstrate that PONG2.0 achieves highly accurate KIR allele frequency estimation across independent datasets, with gene-specific concordance of (R^2=0.815-0.999).

Conclusions

We present PONG2.0, a multi-ancestry imputation framework, that enables comprehensive and accurate KIR gene imputation through innovative region-specific optimization. By systematically accounting for the unique genomic architecture of the KIR locus during model development, our approach achieves improved performance. The implementation includes pre-trained multi-ancestry models distributed as an open-source R package, providing researchers with an efficient solution for high-resolution KIR genotyping directly from SNP array data. This resource effectively addresses the major cost and technical bottlenecks associated with traditional KIR genotyping techniques. The availability of PONG2.0 will significantly advance immunogenetics research by enabling large-scale investigations into KIR-disease associations, population-specific diversity patterns, and clinical applications in transplantation medicine and immunotherapy development.

Funding: NIH/NIAID R01AI158410, Australia NH and MRC, 2004262.

Contact

Paul Norman
Department of Biomedical Informatics
University of Colorado Anschutz Medical Campus
Email:paul.norman@cuanschutz.edu
Website: https://medschool.cuanschutz.edu/dbmi/lab-pages/norman-lab

References

1. Pollock, N. R., Harrison, G. F. & Norman, P. J. Immunogenomics of Killer Cell Immunoglobulin-Like Receptor (KIR) and HLA Class I: Coevolution and Consequences for Human Health. *Journal of Allergy and Clinical Immunology: In Practice* **10**, (2022).
2. Wasilewska, A. *et al.* Immunological Aspects of Infertility—The Role of KIR Receptors and HLA-C Antigen. *Cells* vol. 13 Preprint at https://doi.org/10.3390/cells13010059 (2024).
3. Li, L., Kolk, M., Fernandez-Viña, M., de Lima, M. & Otegbeye, F. Interrogating the impact of KIR ligand mismatch in engraftment following HLA-disparate stem cell transplantation. *Bone Marrow Transplant* **55**, (2020).
4. Parham, P. MHC class I molecules and KIRS in human history, health and survival. *Nature Reviews Immunology* vol. 5 Preprint at https://doi.org/10.1038/nri1570 (2005).
5. Augusto, D. G. & Petzl-Erler, M. L. KIR and HLA under pressure: evidences of coevolution across worldwide populations. *Human Genetics* vol. 134 Preprint at https://doi.org/10.1007/s00439-015-1579-9 (2015).
6. Jiang, Y. *et al.* KIR3DS1/L1 and HLA-Bw4-80I are associated with HIV disease progression among HIV typical progressors and long-term nonprogressors. *BMC Infect Dis* **13**, (2013).
7. Trefny, M. P. *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* **25**, (2019).
8. Goedert, J. J. *et al.* Risk of classic kaposi sarcoma with combinations of killer immunoglobulin-like receptor and human leukocyte antigen loci: A population-based case-control study. *Journal of Infectious Diseases* **213**, (2016).