

***In silico* tools for accurate HLA and KIR inference from clinical sequencing data empower immunogenetics on individual-patient and population scales**

Jieming Chen¹, Shravan Madireddi², Deepti Nagarkar², Maciej Migdal³, Jason Vander Heiden¹, Diana Chang^{1,4}, Kiran Mukhyala¹, Suresh Selvaraj⁵, Edward E. Kadel III⁶, Matthew J. Brauer^{1,7}, Sanjeev Mariathasan⁶, Julie Hunkapiller⁴, Suchit Jhunjhunwala¹, Matthew L. Albert^{2,8}, Christian Hammer^{2,4}

Immunogenetic variation in humans is important in research, clinical diagnosis and increasingly a target for therapeutic intervention. Two highly polymorphic loci play critical roles, namely the human leukocyte antigen (HLA) system, which is the human version of the major histocompatibility complex (MHC), and the Killer-cell immunoglobulin-like receptors (KIR) that are relevant for responses of Natural Killer (NK) cells and some subsets of T cells. Their accurate classification has typically required the use of dedicated biological specimens and a combination of *in vitro* and *in silico* efforts. Increased availability of next generation sequencing data has led to the development of ancillary computational solutions. Here, we report an evaluation of recently published algorithms to computationally infer complex immunogenetic variation in the form of HLA alleles and KIR haplotypes from whole-genome or whole-exome sequencing data. For both HLA allele and KIR gene typing, we identified tools that yielded >97% overall accuracy for 4-digit HLA types, and >99% overall accuracy for KIR gene presence, suggesting the readiness of *in silico* solutions for use in clinical and high-throughput research settings.

The classical human leukocyte antigen (HLA) gene complex on chromosome 6, and the killer-cell immunoglobulin-like receptor (KIR) genes on the leukocyte receptor complex (LRC) on chromosome 19 are complex genomic loci that have been known to be difficult to genotype accurately. With the rapidly emerging treatment approaches in the fields of cancer immunotherapy^{1,2} and autoimmunity³, the accurate characterization of a patient's immunogenetic composition in the HLA and KIR regions is becoming more clinically important.

HLA proteins play an important role in presenting peptides, lipids and glycolipids derived from self, tumor or microbial antigens. They display an extreme amount of allelic polymorphism, as a result of pathogen-driven and balancing selection.⁴ Much research has shown HLA variants to be strongly associated with multiple immune and non-immune phenotypes in the fields of cancer^{5,6}, autoimmunity^{3,7}, neurodegeneration⁸, and infectious diseases^{3,7}. In the clinic, achieving matched HLA alleles between the donor and recipient is critical for organ and stem cell transplantation, such that HLA typing and matching have been integrated as part of standard clinical protocols for decades.⁹⁻¹¹ HLA typing has also become increasingly important in diagnostics and clinical practice. For example, several approved drugs carry labels indicating increased risk for adverse events for carriers of specific HLA alleles.¹²⁻¹⁴

KIR proteins are receptors for HLA class I ligands, and are predominantly expressed on natural killer (NK) cells. In contrast to most HLA genes, the genes coding for KIR display extensive copy number polymorphism¹⁵, in addition to considerable allelic variation¹⁶ for each gene. KIR have shown significant associations with disease phenotypes, mainly in the fields of infectious diseases, autoimmunity, inflammatory diseases, and cancer.¹⁷ Associations were found for both single KIR genes, and when considering their interactions with specific HLA molecules. HLA-KIR interactions were demonstrated to predict the risk of organ rejection after kidney transplantation, suggesting a clinical use case for KIR typing.^{18,19} KIR proteins are also known to be important co-determinants of NK

cell education, which is in part mediated through their interactions with different HLA molecules. Such interactions significantly define the heterogeneity of NK cell responsiveness and their sensitivity to inhibition by HLA across individuals.^{20,21} As such, KIR proteins play a critical role in the recognition of “missing-self” phenotypes in infected or tumor cells, which are typically defined by the loss or down-regulation of HLA class I cell surface expression.²²

For research purposes, genotyping arrays covering single nucleotide polymorphisms (SNPs) across the genome have been used to impute HLA and KIR types. However, they require statistical imputation methods to disentangle the complex linkage disequilibrium (LD) between SNPs and HLA or KIR types.^{23,24} These methods also rely on the availability of ancestry-specific or multi-ancestry reference panels that can be difficult to obtain, especially for populations not well represented in genomic data sets.²⁵ In clinical diagnostics, dedicated immunogenetics laboratory solutions to HLA and KIR genotyping are being continually developed.²⁶ Initial molecular typing technologies were low throughput and/or probe-based assays.²⁷ In recent years, high throughput next generation sequencing (NGS) has become increasingly affordable.²⁸ This has enabled the prevalent use of whole-genome sequencing (WGS) and whole-exome sequencing (WES) in the clinic.^{29,30} While well-validated bioinformatics pipelines have been implemented to detect millions of genetic variants from the available clinical sequences,^{31,32} they are typically employed uniformly to the entire genome or exome, and can be ineffective at particular genomic loci that are highly polymorphic, such as the HLA and KIR regions. Dedicated *in silico* typing tools that use NGS data and specifically target the HLA or KIR genes could be a cost-effective and efficient alternative to traditional laboratory HLA or KIR typing methods. While such NGS-based approaches do not require linkage disequilibrium-based statistical imputation for genotyping (since the sequencing reads directly contain the information to define e.g. the HLA allele status), they do require the use of comprehensive databases that capture the diversity and complexity of these genomic loci for alignment (as opposed to single reference genomes).

¹Department of Bioinformatics and Computational Biology, Genentech, South San Francisco, CA 94080, USA. ²Department of Cancer Immunology, Genentech, South San Francisco, CA 94080, USA. ³Roche Global IT Solution Centre (RGITSC), Warsaw, Poland. ⁴Department of Human Genetics, Genentech, South San Francisco, CA 94080, USA. ⁵Biosample & Repository Management, Genentech, South San Francisco, CA 94080, USA. ⁶Department of Oncology Biomarker Development, Genentech, South San Francisco, CA 94080, USA. ⁷Current address: Maze Therapeutics, South San Francisco, CA 94080, USA. ⁸Current address: Insitro, South San Francisco, CA 94080, USA. Correspondence: hammer.christian@gene.com

Despite their biological significance and many practical advantages, the development of clinically-ready *in silico* HLA and KIR typing have thus far been largely hampered by the genetically complex and highly polymorphic nature of the two regions.^{7,33} Here, we conducted a survey of the current HLA and KIR typing capabilities for potential scaling and readiness in clinical applications. We evaluated available computational HLA inference tools by comparing the inferred HLA alleles from WES and WGS data to a gold standard dataset, which was generated using a commercial dedicated typing method. We also assessed and validated a recently published KIR method by Roe and Kuang³⁴, which can be used to infer KIR gene presence and absence from WGS data.

Methods

Generation of a gold standard HLA reference dataset

The gold standard reference dataset contains 56 patient samples. Genomic DNA was extracted from 1 ml of EDTA whole blood on the Roche MagNA Pure 96 system using the MagNA Pure 96 DNA and Viral NA Large Volume Kit. They were then sent to LabCorp (Burlington, NC, USA) for sequence-based, two-field HLA genotyping, using accepted scientific standards meeting the accreditation requirements of the American Society for Histocompatibility and Immunogenetics (ASHI) and the College of American Pathologists (CAP). HLA class I genotypes of *HLA-A*, *-B*, and *-C* were determined using a combination of long-range sequencing from PacBio (Menlo Park, CA, USA) RSII and Sanger sequencing. Class II genotypes for *HLA-DPA1*, *-DPB1*, *-DQA1*, *-DQB1*, *-DRB1*, *-DRB3*, *-DRB4*, *-DRB5* were determined using a combination of long-range sequencing from PacBio RSII, and NGS from Illumina (San Diego, CA, USA) TruSight and MiSeq technologies.

Generation of a KIR reference dataset

Genotyping of KIR alleles was performed for 72 patients using the LinkSeq KIR kit according to manufacturer's user guide (Catalog No: 5358R, One Lambda, Canoga Park, CA, USA). Briefly, the human genomic DNA was amplified and melt curves were collected on a real-time PCR instrument (QuantStudio 5 system, Thermo Fisher Scientific, Waltham, MA, USA). The data was exported to SureTyper software 6.1.2 (One Lambda) for interpretation and reporting of the genotype.

Whole-genome and whole-exome sequencing

WGS and WES were performed for the samples in the HLA and KIR reference datasets, on Illumina HiSeq instruments using paired-end reads of 150 bp. Each whole genome (WG) was sequenced at an average of 30x coverage. For WES, all samples passed QC criteria of 85% of targeted bases at 20x or greater coverage (+/- 5%). FASTQ files were generated for each WG and exome, and were used as direct inputs for each tool, where appropriate. For tools that required BAM files as inputs, the FASTQ files were processed according to a workflow built using the GATK best practices from 2015 and GATK v3.5³². Briefly, it includes read alignment using bwa 0.7.15 with GRCh38 as the reference genome, duplicate marking using Picard tools v2.9, indel re-alignment using GATK v3.5, and then base quality score recalibration using GATK 3. For bwa, except for the '-M' flag, all the default options were used. Individual-level genetic data for this study cannot be made publicly available due to consent restrictions.

Selection and configuration of HLA typing tools

For WGS and WES HLA genotyping, we specifically selected tools that (1) are recently published, (2) are easy to install and implement, (3) have the ability to work with WGS and WES data, (4) and could genotype both HLA I and II genes for potential clinical use. These included HLA*PRG:LA,³⁵ xHLA v1.2 [<https://github.com/humanlongevity/HLA>]³⁶, HLA-HD v1.2 [<https://www.genome.med.kyoto-u.ac.jp/HLA-HD>]³⁷, and HISAT2 v2.1 [<https://ccb.jhu.edu/software/hisat2/index.shtml>]³⁸. As a comparison to a more widely used tool, we also assessed results from Polysolver v1.0³⁹. Polysolver has been widely utilized in many benchmarking and research studies, and has been shown to be one of the more accurate HLA class I genotyping tools.⁴⁰⁻⁴⁴ Polysolver requires hg18 or hg19 as the human reference genome for read extraction from chromosome 6 in the BAM files; to overcome this limitation, we modified the code to include the use of BAMs that are aligned to hg38 as well. No other modifications were made to the code. For xHLA, BAMs were aligned to hg38, and then preprocessed with an additional bash script provided by the authors [<https://github.com/humanlongevity/HLA/blob/master/bin/get-reads-alt-unmap.sh>]. For HLA-HD and HISAT2, the raw FASTQs were used directly. All the tools were run using default settings, and with the HLA/IMGT databases versions that came with the respective tools. For the exact run parameters, please refer to [Supplementary Table 1](#).

Evaluation of HLA typing tools

The tools typically give two alleles per HLA gene, but we do see in occasions, albeit rare, in our study where the algorithm provided more than one pair of possible alleles, and often in descending order of significance. For any results that offer more than two alleles, we only took

the top two inferred HLA alleles; we assumed a diploid germline genome.

Each *in silico* HLA typing tool was evaluated by the number of concordant HLA alleles called by the tool when compared with the reference HLA genotypes obtained from LabCorp, for the classical HLA I and II genes. Its accuracy was defined as the quotient of the number of concordant calls and the sum of the number of concordant plus discordant calls. For uniformity, results from the tools were converted to four-digit resolution before evaluation, except for HLA*PRG:LA, which can only perform genotyping at G group resolution. A G group contains HLA alleles that have the same genomic sequence for the same binding site, while a P group contains all the HLA alleles that have the same protein sequence for the antigen binding site of the HLA binding molecule (exons 2 and 3 for HLA class I and exon 2 for class II). A minimum of P group resolution or higher (including G group, four-digit/two-field resolutions), is usually considered 'high resolution typing' and therefore clinically relevant.⁴⁵ For HLA*PRG:LA, gold standard results were first converted to G groups before evaluation. G group information was obtained from the IMGT/HLA database [http://hla.alleles.org/alleles/g_groups.html].

We split the evaluation of the tools' accuracy into three categories: (1) HLA I, consisting of classical HLA I genes *HLA-A*, *-B*, and *-C*, (2) HLA IIa, consisting of classical HLA II genes *HLA-DPA1*, *-DPB1*, *-DQA1*, *-DQB1*, and *-DRB1*, and (3) HLA IIb, consisting of a second group of classical HLA II genes, *HLA-DRB3*, *-DRB4*, and *-DRB5* (Table 1). We created a second category (HLA IIb) for additional DR genes because only HLA-HD currently genotypes them. Additionally, since every individual carries a variable copy number of the three genes *HLA-DRB3*, *-DRB4* or *-DRB5* that is highly dependent on the *-DRB1* genotype,⁴⁶ a no-call by the computational tool is considered concordant with the gold standard results if it is not also identified. Some tools cannot genotype the full set of the classical class II genes, thus we also provided the accuracy with respect to each class II gene (Table 2).

Evaluation of KIR typing with kpi and interpretation of results

For inference of KIR gene presence or absence, we evaluated kpi [<https://github.com/drocatumn/kpi>, downloaded March 11th, 2020]. An earlier version of the software was recently presented in a preprint and did not provide a validation of its accuracy when compared to qPCR-based dedicated KIR typing methods.³⁴ Kpi requires WGS FASTQ files as input data and outputs a presence / absence call for each KIR gene, as well as possible combinations of KIR haplotypes according to a provided list of reference haplotypes [<https://github.com/drocatumn/kpi/blob/master/input/haps.txt>]. Each KIR gene can in principle be characterized by copy number and allelic variation. A KIR haplotype determines the order and presence of single KIR genes.^{47,48} However, kpi only detects presence or absence of each KIR gene, not allele status or copy number. As a result, the calls for haplotype pairs can be ambiguous (due to differences in copy number of present genes), but the presence of single KIR genes can be resolved.⁴⁹ As such, KIR typing with kpi, albeit coarse, is still useful because many associations have been reported on haplotype or gene level. Furthermore, interactions of KIRs with their HLA ligands are usually defined at the KIR gene level.^{17,22} It should be noted though, that the KIR genes do show extensive allelic polymorphism that can still have an effect on such defined interactions.^{33,50}

Inference of HLA-KIR interactions

NK cell inhibiting as well as activating KIR interactions with their HLA class I ligands were defined according to Pende *et al.*¹⁸ Briefly, some KIR interact with groups of 4-digit HLA alleles according to specific HLA amino acid residues. *HLA-B* alleles were classified as either *Bw4* or *Bw6* according to amino acid positions 77 - 83. *HLA-C* alleles were assigned C1 or C2 status based on amino acid position 80.¹⁸ Other interactions were defined between KIR and specific 2-digit or 4-digit HLA alleles (e.g. *A*03 - KIR3DL2*).

Data availability

Owing to consent restrictions, individual-level genetic data used for this study cannot be made available.

Results

High accuracy for HLA I and II typing with current gold standard WES and WGS data

We selected HLA genotyping tools to infer HLA identities of 56 patients from the EXCELS (NCT00252135)⁵¹ and AVANT (NCT00112918)⁵² clinical trials. The inferred HLA types were then compared to results from the gold standard reference dataset. Despite the limited size of our dataset, the diversity of HLA class I alleles for each HLA gene in our samples is highly comparable to the publicly available, ethnically varied and sequencing-derived 1000 Genomes/HapMap validation set generated by Ehrlich *et al.*⁵³ ([Supplementary Table 2](#)).

From the results consolidated in [Table 1](#), all the selected genotyping tools perform generally well, at an accuracy of >90% for most of the class I and II gene categories, except for xHLA on WES class I and IIa genes.

xHLA demonstrated uncharacteristically low accuracy for our WES data for both classes I and II, when compared to both reported performance³⁶ and the other tools. HISAT2 and HLA-HD performed comparably for the HLA I genes *A*, *B* and *C* for both WES and WGS data, at >98.8% accuracy. Overall, for class II genes, HLA-HD is consistently the most accurate HLA typing tool for both WES and WGS data, at >97.6% accuracy for *DRB3*, *DRB4*, and *DRB5*, and >98.6% accuracy for the *DPA1/B1*, *DQA1/B1* and *DRB1* genes. Moreover, it also provides the widest range of HLA II genes, with the ability to genotype all the classical class II genes (including *HLA-DRA*, *-DRB3*, *-DRB4* and *-DRB5*), while other tools are restricted to a subset of classical class II alleles. However, HLA-HD has lower accuracy for *DQA1* when working off WES data, compared to HISAT2 (Table 2).

While we observed similar or increased accuracy when comparing results from WGS to WES data from the same tool (Table 1), WGS and WES miscalls were not always the same. This is evident when assessing accuracy at the gene level. For example, the increase in overall accuracy of HISAT2 when using WGS data for identifying class II alleles was mainly due to a lower *HLA-DRB1* accuracy when using WES data.

We next focused on the top performant methods, HLA-HD and HISAT2. Notably, the number of miscalls in HLA I genes was too low to characterize patterns or biases in the miscalls (Table 1; maximum of 4 miscalls). For class II genes, HISAT2 showed the highest number of *HLA-DRB1* miscalls when using WES data (Table 2). By contrast, HLA-HD miscalled mostly in the class IIb genes, where it failed to discriminate the highly similar alleles of the paralogous genes, *HLA-DRB3*, *-DRB4* and *-DRB5*. There was no obvious bias in miscalls of homozygous genes in the HLA-HD and HISAT2 results (Supplementary Table 3). We note that most of HLA-HD miscalls still called an allele in the same G group as the anticipated allele. In particular, almost all the miscalls in *HLA-DQA1* (5/6 of *-DQA1* miscalls and 5/8 of total miscalls) were made when they were called as *-DQA1*03:01* and the anticipated calls were *-DQA1*03:02* and *-DQA1*03:03*; all three alleles are in the same G group. For HISAT2, many of the miscalls in class II genes were due to missing calls, i.e. calls that the tool was not able to assign an allele at all.

High accuracy for identifying KIR gene absence/presence

We used kpi to infer KIR gene presence for 824 patients with available WGS data from the AVANT trial (NCT00112918).⁵² We found that the gene carrier frequencies were very similar to those of published KIR gene carrier frequencies from an English cohort of Caucasoid ethnicity.⁵⁴ (Figure 1a). The software was unable to predict possible haplotype pairs for 3% (n=25) of cases.

We selected 72 of these 824 patients to perform qPCR-based KIR typing, based on DNA sample availability and the diversity of kpi-predicted KIR haplotypes. These patients are different from those selected for the HLA gold standard data because of limited availability of DNA. For this selection, we also included eight of the 25 patients that had yielded uninterpretable haplotype combinations using kpi.

99.2% of kpi gene presence/absence calls were correct when compared to our qPCR-based reference (range of 95.8% - 100% for the 16

tested genes, Figure 1b, Supplementary Table 4). When excluding the four framework genes *KIR3DL3*, *KIR3DP1*, *KIR2DL4*, and *KIR3DL2*, which were invariable in our dataset and are present in most common haplotypes, the accuracy for the 12 remaining genes was 99.0%. Three patients were typed with one error each (*KIR2DS3*), and one patient with six errors. The eight qPCR-typed individuals with uninterpretable kpi haplotype results had correct kpi gene calls, which could not be clearly assigned to a reference haplotype combination as provided by the software (Supplementary Table 5).

Discussion

The broad availability of NGS data, generated in a multitude of clinical scenarios, allows for the inference of disease-relevant immunogenetic variation without additional dedicated typing efforts. Hence, in order to evaluate the usefulness in a clinical setting, we were interested in a systematic comparison of the newest generation of computational HLA typing solutions, run alongside the more well-established HLA typing tool Polysolver, that is limited in only typing HLA I genes. Our analyses suggest that for both WES and WGS data, most of the tools outperform Polysolver in HLA inference, in both class I genotyping accuracy and the ability to perform inference on HLA II genes. For KIR typing using kpi, we are not aware of a published independent validation of its performance. Our evaluation demonstrates that kpi performs well at determining the absence or presence of a KIR gene, but it is not able to ascertain KIR allele or gene copy number.

Our analyses also show the breadth of class II genes that current state-of-the-art tools can infer. In particular, only HLA-HD was able to genotype the classical HLA II genes, *HLA-DRB3*, *-DRB4*, and *-DRB5*, and we were able to further examine the results with our gold standard dataset. Interestingly, we found that many of the *HLA-DRB3*, *-DRB4* and *-DRB5* miscalls can be salvaged using knowledge of the strong (and clear) linkage disequilibrium between the *HLA-DRB1* gene and its *DRB* paralogs.⁵⁵ It appears that incorporating this piece of biological information could be useful in developing tools that would like to genotype all the *DRB* genes, especially when there is high accuracy in genotyping the *DRB1* gene.

Additionally, with WGS and WES data for the same subjects, we observed that HLA inference from WGS data has yielded marginally higher accuracy compared to WES in many HLA genes (Tables 1 and 2). This possibly indicates that the addition of non-coding sequence information, or a more uniform read coverage in the HLA region, might be more relevant in these genes, especially in resolving alleles that are in the same G group, e.g. *HLA-DQA1*03:01*, *-DQA1*03:02* and *-DQA1*03:03*. It might also point to the use of bait sets in WES, which can bias the calling of some alleles; WGS does not require such baits.

Of note, the dedicated HLA typing approach (LabCorp) identified one novel *HLA-C* allele in our cohort of 56 individuals at two-digit resolution. Even though all the tools gave an estimation (i.e. it was not a missing call) and were correct at the two-digit resolution, none of the tested tools were able to identify the allele as novel. This is because the inferences are all based on aligning sequencing reads to a database of known

Method	WES Allele concordance (%)			WGS Allele concordance (%)		
	I**	IIa	IIb	I**	IIa	IIb
xHLA	156/330 (47.2)^	187/330 (56.7)^	--	327/336 (97.3)	327/336 (97.3)	--
HLA-HD	334/336 (99.4)	552/560 (98.6)	329/336 (97.9)	334/336 (99.4)	557/560 (99.5)	328/336 (97.6)
HISAT2	332/336 (98.8)	310/336 (92.3)	--	334/336 (99.4)	324/336 (96.4)	--
HLA*PRG:LA (G groups only)	333/336 (99.1)	335/336 (99.7)	--	333/336 (99.1)	335/336 (99.7)	--
Polysolver	328/336 (97.6)	--	--	330/336 (98.2)	--	--

[^]1 sample failed to run in xHLA WES.

^{**}Included miscalls for a novel *HLA-C* allele at 4-digit resolution. All tools matched the LabCorp result at 2-digit resolution correctly.

Table 1. Overall evaluation results for HLA typing using WES and WGS. HLA genes are categorized into classical class I genes (*A*, *B*, *C*), IIa genes (*DPA1*, *DPB1*, *DQA1*, *DQB1*, *DRB1*), and IIb genes (*DRB3*, *DRB4*, *DRB5*). The class IIa genes that each tool can genotype differ: HLA-HD can infer all of the above; xHLA can only infer *DPB1*, *DQB1* and *DRB1*; HISAT2 and HLA*PRG:LA only *DQA1*, *DQB1* and *DRB1*. Note that HLA*PRG:LA was evaluated based on the G group resolution. Results from the older and well-utilized Polysolver were provided as an additional source of comparison for HLA I genes.

HLA II gene	xHLA (%)		HLA-HD (%)		HISAT2 (%)		HLA*PRG:LA (%)	
	WES [^]	WGS	WES	WGS	WES	WGS	WES	WGS
<i>DPA1</i>	--	--	112/112 (100)	112/112 (100)	--	--	--	--
<i>DPB1</i>	70/110 (63.6)	109/112 (97.3)	110/112 (98.2)	110/112 (98.2)	--	--	--	--
<i>DQA1</i>	--	--	106/112 (94.6)	111/112 (99.1)	111/112 (99.1)	110/112 (98.2)	112/112 (100)	112/112 (100)
<i>DQB1</i>	53/110 (48.2)	107/112 (95.5)	112/112 (100)	112/112 (100)	111/112 (99.1)	110/112 (98.2)	111/112 (99.1)	112/112 (100)
<i>DRB1</i>	64/110 (58.2)	111/112 (99.1)	112/112 (100)	112/112 (100)	88/112 (78.6)	104/112 (92.9)	112/112 (100)	111/112 (99.1)
<i>DRB3</i>	--	--	110/112 (98.2)	111/112 (99.1)	--	--	--	--
<i>DRB4</i>	--	--	110/112 (98.2)	108/112 (96.4)	--	--	--	--
<i>DRB5</i>	--	--	109/112 (97.3)	109/112 (97.3)	--	--	--	--

[^]1 sample failed to run in xHLA WES.

Table 2. Evaluation results by HLA II genes for HLA typing using WES and WGS. Note that HLA*PRG:LA was evaluated based on the G group resolution, not at 4-digit resolution.

HLA alleles. This might not be highly important for large-scale genetic association approaches, but might be relevant in a clinical setting focused on individual patients, especially for ethnicities that have thus far been underrepresented with regard to genome sequencing, let alone HLA typing.^{56,57}

The choice and accuracy of a given HLA method might depend on read length and sequencing coverage, which are factors that are not included in the current study. A recent comparison of HLA-HD and xHLA for use with target capture methods or amplification sequencing suggested that HLA-HD might decrease in sensitivity at read lengths below 150 base pairs (paired-end).⁵⁸

We noticed that documentation for most of the HLA typing tools tested is mainly centered on the final inference, but not the auxiliary output files. The latter set of files typically contain the scores for all the candidates used for inference. While accuracy is important, tool documentation in a clinical setting is also imperative to better understand the tool and its outputs, so that best practices can be developed in the clinic for different contexts.

As for the KIR typing efforts, *kpi* was shown to predict KIR gene presence/absence at >99% accuracy overall, and at >95% for each gene. Six out of nine errors were found in a single individual. This was likely due to a sample swap, and the remaining three miscalls were

all for *KIR2DS3* in different patients. In this case, all other genes were inferred at 100% accuracy. However, since *kpi* detects gene presence/absence and does not perform an estimation of copy number, it assigns one or more possible haplotype combinations, resulting in considerable ambiguity. Thus, we recommend to analyze *kpi* results at the level of individual KIR genes, if possible. It is likely that the 25 uninterpretable haplotype pairs are due to carriers of rare haplotypes not present in the reference, which would prevent an assignment of possible reference haplotype combinations.

Notably, *kpi* requires WGS data and does not consider allelic variation within KIR genes. This is a significant limitation, since allotypes for a given KIR gene can be functionally different,⁵⁹ and also have differential binding capacities to their predicted HLA ligands.^{50,60,61} Allele-level typing would be desirable. The only software presently available that we are aware of that provides this level of granularity was not designed to work with NGS data in a high-throughput setting.⁶²

In conclusion, our survey for both high-resolution four-digit clinically-relevant HLA typing and inference of KIR gene presence from NGS data (of conventional read length and coverage) indicated that recently published software tools can yield very high accuracy (>97% for HLA alleles and >95% for KIR genes, respectively) that may be suitable not only for research us, but also for the clinic. For comparison, the 2019 Standards for Accredited Laboratories issued by the American Society

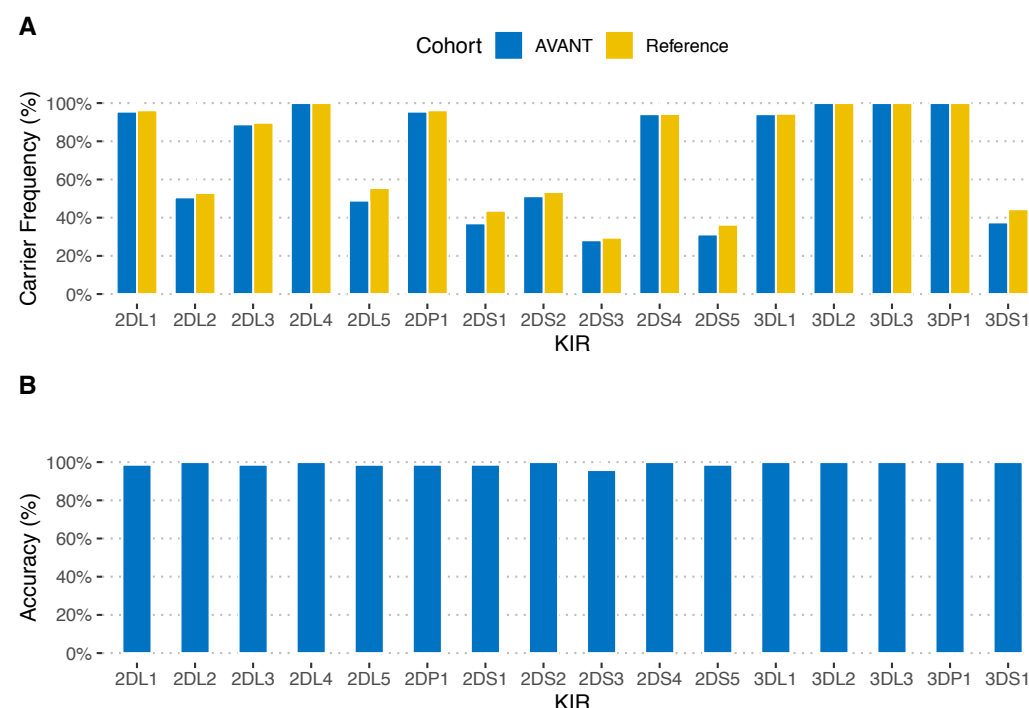


Figure 1. KIR gene carrier frequencies and accuracy of *kpi* typing

(A) KIR gene presence for AVANT patients (N=824) was inferred from *kpi* haplotype predictions, and compared to published frequencies for an English reference cohort (N=584). (B) For 72 AVANT patients typed with *kpi*, KIR typing with a qPCR-based method (LinkSeq) was performed to assess typing accuracy.

for Histocompatibility and Immunogenetics only requires a minimum of 80% concordance with another CLIA-certified ASHI-accredited laboratory to be deemed satisfactory in clinical testing.⁶³ It is noteworthy that WGS and WES continue to become less expensive, thereby presenting an alternative even in scenarios that focus only on HLA or KIR typing. However, a foreseeable hurdle is the process of obtaining regulatory approval for computational tools for HLA and KIR typing in the clinic, either as a stand-alone device, or as part of a pipeline. Such a process could be tricky as it can be highly dependent on the context of how the tool is being applied in the clinic. There are pros and cons for each tool. Other considerations in the choice of method that we did not explore in this study and might merit investigation in the future, include the characteristics of the NGS datasets at hand, such as the read length and read coverage, which can affect accuracy and thus cause deviations from what is shown in the present report.

Finally, we would like to further emphasize that computational tools can generate HLA and KIR information in a high-throughput manner on large cohorts of patients with clinical sequencing. Furthermore, the time and logistical challenges and risks associated with acquisition, preparation and shipping of valuable clinical specimens to perform a separate genotyping would be greatly reduced. In a clinical setting, the HLA and KIR results from these tools can then be applied directly to detect immunogenetic biomarkers that might be relevant for treatment decisions, or to predict the likelihood of adverse events for a given treatment of choice.⁶⁴ HLA typing is also a requirement in the context of individualized cancer treatment strategies, including immunization efforts and neoadjuvant-directed T-cell therapies.^{65,66} Neoepitope prediction requires highly accurate HLA types in order to maximize the likelihood of an immunogenic anti-tumor response.⁶⁷ KIR genes are emerging biomarkers in several disease areas, including cancer immunology,⁶⁸ and should ideally be investigated in the context of their interactions with their HLA ligands. Having both HLA and KIR information will also allow stratification of patients according to their individual and biologically relevant HLA-KIR interactions (Supplementary Figure 1).²² As more computational tools for HLA and KIR typing are likely being developed in the future, they should be continuously evaluated so that they can fulfill a greater role in assessing clinical genomes.

Acknowledgments

The authors would like to thank the author of kpi, David Roe, for helpful discussions. We are also thankful to Sean Kelley for support of this study.

References

- Chen DS, Mellman I. Oncology meets immunology: the cancer-immunity cycle. *Immunity*. 2013;39(1):1-10. doi:10.1016/j.immuni.2013.07.012
- Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. *Nat Rev Cancer*. 2012;12(4):252-264. doi:10.1038/nrc3239
- Matzaraki V, Kumar V, Wijmenga C, Zhernakova A. The MHC locus and genetic susceptibility to autoimmune and infectious diseases. *Genome Biol*. 2017;18(1):76. doi:10.1186/s13059-017-1207-1
- Meyer D, C Aguiar VR, Bitarello BD, C Brandt DY, Nunes K. A genomic perspective on HLA evolution. *Immunogenetics*. 2018;70(1):5-27. doi:10.1007/s00251-017-1017-3
- Kaneko K, Ishigami S, Kijima Y, et al. Clinical implication of HLA class I expression in breast cancer. *BMC Cancer*. 2011;11:454. doi:10.1186/1471-2407-11-454
- Aptsiauri N, Cabrera T, Mendez R, Garcia-Lora A, Ruiz-Cabello F, Garrido F. Role of altered expression of HLA class I molecules in cancer progression. *Adv Exp Med Biol*. 2007;601:123-131. doi:10.1007/978-0-387-72005-0_13
- Trowsdale J. The MHC, disease and selection. *Immunol Lett*. 2011;137(1-2):1-8. doi:10.1016/j.imlet.2011.01.002
- O'Keefe GM, Nguyen VT, Benveniste EN. Regulation and function of class II major histocompatibility complex, CD40, and B7 expression in macrophages and microglia: Implications in neurological diseases. *J Neurovirol*. 2002;8(6):496-512. doi:10.1080/13550280290100941
- Zachary AA, Leffell MS. HLA Mismatching Strategies for Solid Organ Transplantation - A Balancing Act. *Front Immunol*. 2016;7:575. doi:10.3389/fimmu.2016.00575
- Mahdi BM. A glow of HLA typing in organ transplantation. *Clin Transl Med*. 2013;2(1):6. doi:10.1186/2001-1326-2-6
- Sheldon S, Poulton K. HLA typing and its influence on organ transplantation. *Methods Mol Biol*. 2006;333:157-174. doi:10.1385/1-59745-049-9:157
- U.S. Food and Drug Administration. Table of Pharmacogenomic Biomarkers in Drug Labeling. 2018. <https://www.fda.gov/media/107901/download>.
- Dean L. Carbamazepine Therapy and HLA Genotype. In: V P, H M, W R, et al., eds.

- Medical Genetics Summaries [Internet]. Bethesda (MD): National Center for Biotechnology Information (US); 2015. <https://www.ncbi.nlm.nih.gov/books/NBK321445/>.
- Dean L. Phenotypic Therapy and HLA-B*15:02 and CYP2C9 Genotypes. In: Medical Genetics Summaries [Internet]. Bethesda (MD): National Center for Biotechnology Information (US); 2016.
- Wilson MJ, Torkar M, Haude A, et al. Plasticity in the organization and sequences of human KIR/ILT gene families. *Proc Natl Acad Sci U S A*. 2000;97(9):4778-4783. doi:10.1073/pnas.080588597
- Wagner I, Scheffzyk D, Pruschke J, 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
- Kulkarni S, Martin MP, Carrington M. The Yin and Yang of HLA and KIR in human disease. *Semin Immunol*. 2008;20(6):343-352. doi:10.1016/j.smim.2008.06.003
- Littera R, Piredda G, Argiolas D, et al. KIR and their HLA Class I ligands: Two more pieces completing the puzzle of chronic rejection and graft loss in kidney transplantation. *PLoS One*. 2017;12(7):e0180831. doi:10.1371/journal.pone.0180831
- Jafari D, Nafar M, Yekaninejad MS, et al. Investigation of Killer Immunoglobulin-like Receptor (KIR) and HLA Genotypes to Predict the Occurrence of Acute Allograft Rejection after Kidney Transplantation. *Iran J Allergy Asthma Immunol*. 2017;16(3):245-255. <http://www.ncbi.nlm.nih.gov/pubmed/28732438>.
- Boudreau JE, Hsu KC. Natural Killer Cell Education and the Response to Infection and Cancer Therapy: Stay Tuned. *Trends Immunol*. 2018;39(3):222-239. doi:10.1016/j.it.2017.12.001
- Horowitz A, Djaoud Z, Nemat-Gorgani N, et al. Class I HLA haplotypes form two schools that educate NK cells in different ways. *Sci Immunol*. 2016;1(3). doi:10.1126/sciimmunol.aag1672
- Pende D, Falco M, Vitale M, et al. Killer Ig-Like Receptors (KIRs): Their Role in NK Cell Modulation and Developments Leading to Their Clinical Exploitation. *Front Immunol*. 2019;10:1179. doi:10.3389/fimmu.2019.01179
- Karnes JH, Shaffer CM, Bastarache L, et al. Comparison of HLA allelic imputation programs. *PLoS One*. 2017;12(2):e0172444. doi:10.1371/journal.pone.0172444
- Vukcevic D, Traherne JA, Naess S, et al. Imputation of KIR Types from SNP Variation Data. *Am J Hum Genet*. 2015;97(4):593-607. doi:10.1016/j.ajhg.2015.09.005
- Sirugo G, Williams SM, Tishkoff SA. The Missing Diversity in Human Genetic Studies. *Cell*. 2019;177(1):26-31. doi:10.1016/j.cell.2019.02.048
- Christiansen FT, Tait BD, eds. *Immunogenetics*. Vol 882. Totowa, NJ: Humana Press; 2012. doi:10.1007/978-1-61779-842-9
- Edgerly CH, Weimer ET. The Past, Present, and Future of HLA Typing in Transplantation. *Methods Mol Biol*. 2018;1802:1-10. doi:10.1007/978-1-4939-8546-3_1
- Check Hayden E. Technology: The \$1,000 genome. *Nature*. 2014;507(7492):294-295. doi:10.1038/507294a
- Shevchenko Y, Bale S. Clinical Versus Research Sequencing. *Cold Spring Harb Perspect Med*. 2016;6(11). doi:10.1101/cshperspect.a025809
- Thiffault I, Farrow E, Zellmer L, et al. Clinical genome sequencing in an unbiased pediatric cohort. *Genet Med*. 2019;21(2):303-310. doi:10.1038/s41436-018-0075-8
- Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009;25(16):2078-2079. doi:10.1093/bioinformatics/btp352
- DePristo MA, Banks E, Poplin R, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet*. 2011;43(5):491-498. doi:10.1038/ng.806
- Middleton D, Gonzelez F. The extensive polymorphism of KIR genes. *Immunology*. 2010;129(1):8-19. doi:10.1111/j.1365-2567.2009.03208.x
- Roe D, Kuang R. Predicting KIR Structural Haplotypes with Novel Sequence Signatures from Short-read Whole Genome Sequencing. *bioRxiv*. 2019. doi:10.1101/541938
- Dilthey AT, Mentzer AJ, Carapito R, et al. HLA*LA-HLA typing from linearly projected graph alignments. *Bioinformatics*. 2019;35(21):4394-4396. doi:10.1093/bioinformatics/btz235
- Xie C, Yeo ZX, Wong M, et al. Fast and accurate HLA typing from short-read next-generation sequencing data with xHLA. *Proc Natl Acad Sci U S A*. 2017;114(30):8059-8064. doi:10.1073/pnas.1707945114
- Kawaguchi S, Higasa K, Shimizu M, Yamada R, Matsuda F. HLA-HD: An accurate HLA typing algorithm for next-generation sequencing data. *Hum Mutat*. 2017;38(7):788-797. doi:10.1002/humu.23230
- Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol*. 2019;37(8):907-915. doi:10.1038/s41587-019-0201-4
- Shukla SA, Rooney MS, Rajasagi M, et al. Comprehensive analysis of cancer-associated somatic mutations in class I HLA genes. *Nat Biotechnol*. 2015;33(11):1152-1158. doi:10.1038/nbt.3344
- Bauer DC, Zadoorian A, Wilson LOW, Melbourne Genomics Health Alliance, Thorne NP. Evaluation of computational programs to predict HLA genotypes from genomic sequencing data. *Brief Bioinform*. 2018;19(2):179-187. doi:10.1093/bib/bbw097
- Kiyotani K, Mai TH, Nakamura Y. Comparison of exome-based HLA class I genotyping tools: identification of platform-specific genotyping errors. *J Hum Genet*. 2017;62(3):397-405. doi:10.1038/jhg.2016.141
- Larjo A, Eveleigh R, Kilpeläinen E, et al. Accuracy of Programs for the Determination of Human Leukocyte Antigen Alleles from Next-Generation Sequencing Data. *Front Immunol*. 2017;8:1815. doi:10.3389/fimmu.2017.01815
- Matey-Hernandez ML, Danish Pan Genome Consortium, Brunak S, Izarzugaza JMG. Benchmarking the HLA typing performance of Polysolver and Optitype in 50 Danish

parental trios. *BMC Bioinformatics*. 2018;19(1):239. doi:10.1186/s12859-018-2239-6

44. McGranahan N, Rosenthal R, Hiley CT, et al. Allele-Specific HLA Loss and Immune Escape in Lung Cancer Evolution. *Cell*. 2017;171(6):1259-1271.e11. doi:10.1016/j.cell.2017.10.001

45. Nunes E, Heslop H, Fernandez-Vina M, et al. Definitions of histocompatibility typing terms. *Blood*. 2011;118(23):e180-3. doi:10.1182/blood-2011-05-353490

46. Fernando MMA, Vyse TJ. Major Histocompatibility Complex Class II. In: *Systemic Lupus Erythematosus*. 5th ed. Elsevier; 2011:3-19. doi:10.1016/B978-0-12-374994-9.10001-4

47. Vierra-Green C, Roe D, Hou L, 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. doi:10.1371/journal.pone.0047491

48. Pyo C-W, Guethlein LA, Vu Q, et al. Different patterns of evolution in the centromeric and telomeric regions of group A and B haplotypes of the human killer cell Ig-like receptor locus. *PLoS One*. 2010;5(12):e15115. doi:10.1371/journal.pone.0015115

49. Roe D, Kuang R. kpi haplotype file. bioRxiv. https://github.com/droecatumn/kpi/blob/master/input/all_haps_v6.txt. Published 2019.

50. Frazier WR, Steiner N, Hou L, Dakshanamurthy S, Hurley CK. Allelic variation in KIR2DL3 generates a KIR2DL2-like receptor with increased binding to its HLA-C ligand. *J Immunol*. 2013;190(12):6198-6208. doi:10.4049/jimmunol.1300464

51. Long AA, Fish JE, Rahmaoui A, et al. Baseline characteristics of patients enrolled in EXCELS: a cohort study. *Ann Allergy Asthma Immunol*. 2009;103(3):212-219. doi:10.1016/S1081-1206(10)60184-6

52. de Gramont A, Van Cutsem E, Schmoll H-J, et al. Bevacizumab plus oxaliplatin-based chemotherapy as adjuvant treatment for colon cancer (AVANT): a phase 3 randomised controlled trial. *Lancet Oncol*. 2012;13(12):1225-1233. doi:10.1016/S1470-2045(12)70509-0

53. Erlich RL, Jia X, Anderson S, et al. Next-generation sequencing for HLA typing of class I loci. *BMC Genomics*. 2011;12:42. doi:10.1186/1471-2164-12-42

54. Hiby SE, Ashrafi-Bonab M, Farrell L, et al. Distribution of killer cell immunoglobulin-like receptors (KIR) and their HLA-C ligands in two Iranian populations. *Immunogenetics*. 2010;62(2):65-73. doi:10.1007/s00251-009-0408-5

55. Steiman AJ. *Systemic Lupus Erythematosus*, 5th Edition. *J Rheumatol*. 2011;38(8):1814-1814. doi:10.3899/jrheum.110551

56. Bustamante CD, Burchard EG, De la Vega FM. Genomics for the world. *Nature*. 2011;475(7355):163-165. doi:10.1038/475163a

57. Pidala J, Kim J, Schell M, et al. Race/ethnicity affects the probability of finding an HLA-A, -B, -C and -DRB1 allele-matched unrelated donor and likelihood of subsequent transplant utilization. *Bone Marrow Transplant*. 2013;48(3):346-350. doi:10.1038/bmt.2012.150

58. Meng R, Zhu T, Gong Y, et al. Abstract 5116: Benchmarking the HLA typing performance of three HLA assay and two software. In: *Bioinformatics, Convergence Science, and Systems Biology*. American Association for Cancer Research; 2019:5116-5116. doi:10.1158/1538-7445.AM2019-5116

59. Steiner NK, Dakshanamurthy S, VandenBussche CJ, Hurley CK. Extracellular domain alterations impact surface expression of stimulatory natural killer cell receptor KIR2DS5. *Immunogenetics*. 2008;60(11):655-667. doi:10.1007/s00251-008-0322-2

60. Blokhuis JH, Hilton HG, Guethlein LA, et al. KIR2DS5 allotypes that recognize the C2 epitope of HLA-C are common among Africans and absent from Europeans. *Immunity, Inflamm Dis*. 2017;5(4):461-468. doi:10.1002/iid3.178

61. 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;112(2):435-443. doi:10.1182/blood-2008-01-132902

62. Norman PJ, Hollenbach JA, Nemat-Gorgani N, et al. Defining KIR and HLA Class I Genotypes at Highest Resolution via High-Throughput Sequencing. *Am J Hum Genet*. 2016;99(2):375-391. doi:10.1016/j.ajhg.2016.06.023

63. American Society for Histocompatibility and Immunogenetics. Proposed 2019 ASHI Standards with Guidance. https://cdn.ymaws.com/www.ashi-hla.org/resource/resmgr/docs/standards/2019/clean_2019_0715_-_2019_ashi_.pdf. Published 2019.

64. Khan Z, Hammer C, Guardino E, Chandler GS, Albert ML. Mechanisms of immune-related adverse events associated with immune checkpoint blockade: using germline genetics to develop a personalized approach. *Genome Med*. 2019;11(1):39. doi:10.1186/s13073-019-0652-8

65. Drake CG. Prostate cancer as a model for tumour immunotherapy. *Nat Rev Immunol*. 2010;10(8):580-593. doi:10.1038/nri2817

66. Capietto A-H, Jhunjhunwala S, Delamarre L. Characterizing neoantigens for personalized cancer immunotherapy. *Curr Opin Immunol*. 2017;46:58-65. doi:10.1016/j.coi.2017.04.007

67. The problem with neoantigen prediction. *Nat Biotechnol*. 2017;35(2):97. doi:10.1038/nbt.3800

68. Trefny MP, Rothschild SI, Uhlenbrock E, et al. A Variant of a Killer Cell Immunoglobulin-like Receptor Is Associated with Resistance to PD-1 Blockade in Lung Cancer. *Clin Cancer Res*. 2019;25(10):3026-3034. doi:10.1158/1078-0432.CCR-18-3041