



Benchmarking the Human Leukocyte Antigen Typing Performance of Three Assays and Seven Next-Generation Sequencing-Based Algorithms

Ping Liu^{1†}, Minya Yao^{2†}, Yu Gong^{3†}, Yunjie Song^{4†}, Yanan Chen⁴, Yizhou Ye⁴, Xiao Liu⁴, Fugen Li⁴, Hua Dong⁴, Rui Meng^{5*}, Hao Chen^{4*} and Aiwen Zheng^{6,7*}

OPEN ACCESS

Edited by:

Jukka Partanen,
Finnish Red Cross Blood Service,
Finland

Reviewed by:

Marco Andreani,
Bambino Gesù Children Hospital
(IRCCS), Italy
James Robinson,
Anthony Nolan Research Institute,
United Kingdom

*Correspondence:

Rui Meng
mengruvip@163.com
Hao Chen
hao.chen1@3dmedcare.com
Aiwen Zheng
zhengaw@zjcc.org.cn

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

Specialty section:

This article was submitted to
Alloimmunity and Transplantation,
a section of the journal
Frontiers in Immunology

Received: 12 January 2021

Accepted: 03 March 2021

Published: 31 March 2021

Citation:

Liu P, Yao M, Gong Y, Song Y, Chen Y,
Ye Y, Liu X, Li F, Dong H, Meng R,
Chen H and Zheng A (2021)
Benchmarking the Human Leukocyte
Antigen Typing Performance of Three
Assays and Seven Next-Generation
Sequencing-Based Algorithms.
Front. Immunol. 12:652258.
doi: 10.3389/fimmu.2021.652258

¹ Department of Oncology, The Second Xiangya Hospital of Central South University, Changsha, China, ² The First Affiliated Hospital of Zhejiang University, School of Medicine, Hangzhou, China, ³ Department of Urology, Second Affiliated Hospital, Zhejiang University College of Medicine, Hangzhou, China, ⁴ The Bioinformatics Department, 3DMed Inc., Shanghai, China, ⁵ Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China, ⁶ Cancer Hospital of the University of Chinese Academy of Sciences (Zhejiang Cancer Hospital), Hangzhou, China, ⁷ Institute of Cancer and Basic Medicine (IBMC), Chinese Academy of Sciences, Hangzhou, China

With the great progress made recently in next generation sequencing (NGS) technology, sequencing accuracy and throughput have increased, while the cost for data has decreased. Various human leukocyte antigen (HLA) typing algorithms and assays have been developed and have begun to be used in clinical practice. In this study, we compared the HLA typing performance of three HLA assays and seven NGS-based HLA algorithms and assessed the impact of sequencing depth and length on HLA typing accuracy based on 24 benchmarked samples. The algorithms HISAT-genotype and HLA-HD showed the highest accuracy at both the first field and the second field resolution, followed by HLAScan. Our internal capture-based HLA assay showed comparable performance with whole exome sequencing (WES). We found that the minimal depth was 100X for HISAT-genotype and HLA-HD to obtain more than 90% accuracy at the third field level. The top three algorithms were quite robust to the change of read length. Thus, we recommend using HISAT-genotype and HLA-HD for NGS-based HLA genotyping because of their higher accuracy and robustness to read length. We propose that a minimal sequence depth for obtaining more than 90% HLA typing accuracy at the third field level is 100X. Besides, targeting capture-based NGS HLA typing may be more suitable than WES in clinical practice due to its lower sequencing cost and higher HLA sequencing depth.

Keywords: human leukocyte antigen, accuracy, next-generation sequencing, algorithms, benchmark

INTRODUCTION

The human leukocyte antigen (HLA), commonly referred as major histocompatibility complex (MHC) which is often found in all jawed vertebrates (1), is located within a region of approximately 4 M in length on the short arm of human chromosome 6 (6p21.3), with more than 200 protein-coding genes (2). Except for identical twins, no two individuals have exactly the same HLA.

Therefore, HLA is also known as the “identity card” of the human cell. It is a marker for the mutual recognition of immune cells in different individuals. HLA gene products are expressed on different cell surfaces and play a key role in antigen presentation and immune signaling. HLA mainly includes three regions, namely HLA class I, HLA class II, and HLA class III. HLA class I genes include HLA-A, HLA-B, and HLA-C, which are distributed on almost all nucleated cell surfaces with the highest lymphocyte surface density (3). HLA class II genes include the HLA-D family, mainly HLA-DP, HLA-DQ, and HLA-DR, which are mainly distributed on the surface of professional antigen-presenting cells such as B lymphocytes, macrophages, and dendritic cells. The HLA class III gene contains approximately 75 genes, most of which are of unknown function. HLA class I and HLA class II genes are molecules that encode binding and presenting antigens, allowing cytotoxic T lymphocytes to bind to mature HLA cell surface proteins *via* antigen-binding channels. HLA class I genes mainly encode antigens to CD8+ T cells, and HLA class II genes mainly encode antigens to CD4+ T cells.

HLA has been widely used in hematopoietic stem cell transplantation (HSCT), detection of susceptibility genes in immune-related diseases, and drug allergy testing. HSCT was treated as the last resort therapeutic approach for a wide range of malignant and non-malignant diseases and suitable donor selection is determined with the utilization of HLA typing and highly similar HLA alleles improve the clinical outcome and reduce the risk of rejection (4). According to USA standards, 8/8 match for the loci HLA-A, HLA-B, HLA-C, and HLA-DRB1 is necessary for a allele-matched donor selection, and single mismatch for these regions are associated with 25% increase in post-transplant complications (5). But in most European centers the gold standard is to look for 10/10 match for HLA-A, HLA-B, HLA-C, HLA-DRB1, and HLA-DQB1 (6). The definition of “HLA matching” depends on the HLA typing resolution, mainly include: Low resolution typing or first field typing, which is equivalent to serological typing and refers to a group of alleles (alleles family); High resolution typing, or second field typing, which refers to one or a set of alleles for the same antigen binding site; Allele level typing, or all field typing, which refers to the exact nucleotide sequence of a HLA gene; Other levels of resolution, which refers to intermediate level of typing and could define specific subtypes. Currently, high resolution typing of HLA genes were recommended by National Marrow Donors Program (NMDP) (5). Thus, HLA typing at the high resolution level is of great clinical significance.

Recent studies have demonstrated that HLA typing complexity is associated with the efficacy of cancer immune checkpoint blockade (ICB) (7). Furthermore, the combined effect of HLA class I heterozygosity and tumor mutation burden (TMB) on improved survival is greater as compared with mutation load alone (7, 8). Researchers have also sequenced the CDR3 of the hypervariable region of the T cell receptor (TCR) and found that the TCR CDR3's tumor-associated clones are significantly elevated in patients with greater heterogeneity of the HLA class of molecular sites (7). That is to say, in the treatment of ICB, the diversity of HLA molecules in patients will affect the clonal expansion of T cells against new tumor antigens and thus affect the therapeutic effect (9). The highly polymorphic HLA genes present unique challenges for the development of molecular approaches to genotype HLA alleles. According to the traditional method, both alleles of a particular HLA locus are PCR amplified and Sanger-sequenced together, resulting in multiple heterozygous positions in the electropherogram tracing. With the development of next-generation sequencing (NGS) technology, each fragment of HLA DNA is amplified and sequenced independently, dramatically reducing the phase ambiguities encountered with Sanger sequencing. Since 2009, many different approaches for HLA genotyping by the NGS method have been reported using a variety of capture strategies and sequencing platforms (10–15). While whole exome sequencing is the gold standard in some case, such as measurement of TMB in clinic, targeted next-generation sequencing panels might be ideal for HLA typing which allows us to customize probes that only include genomic regions of HLA genes, and sequence HLA gene at a much higher depth but lower input amounts than WES. Many bioinformatics approaches have also been developed to produce HLA genotyping information from amplicon-based NGS, targeted capture (e.g., whole-exome sequencing) and non-targeted whole-genome sequencing (16–23) (software used in this study are listed in **Table 1**). All these algorithms can be generally divided into two categories: alignment-based methods and assembly-based methods. The former category aligns the sequencing data to the HLA reference database IPD-IMGT/HLA (24, 25) and predicts HLA genotypes using probabilistic models (26), whereas the latter assembles reads into contigs and aligns those to the known HLA allele reference sequences. Several studies have been conducted to compare the accuracy of different software (26–30). Bauer et al. evaluated the HLA typing accuracy of five computational methods on three different data sets, finding that PHLAT has the highest accuracy, and

TABLE 1 | HLA-typing software used in this study.

Software	Resolution	Programming	Year	Journal	Cited
HLAminer	4	Perl	2012	Genome Medicine	83
seq2HLA	4	Python, R	2012	Genome Medicine	93
HLAforest	8	Perl	2013	PLOS ONE	28
HLA-VBSeq	8	Java	2015	BMC Genomics	36
HLA-HD	6	Shell	2017	Human mutation	15
HLAScan	8	Python	2017	BMC Bioinformatics	22
HISAT-genotype	8	C++, Python	2019	Nature Biotechnology	81

sequencing coverage has a weak correlation with accuracy (26). However, no conclusions have been made regarding several critical questions: Which HLA typing assay is more suitable in a clinical context? Whether HLA typing algorithms were biased towards a specific NGS assay? What are the basic sequencing requirements for accurate HLA genotyping? To answer these questions, we evaluated the performance of different combinations of HLA NGS typing assays and software using our in-house benchmarking dataset.

MATERIALS AND METHODS

Sample Preparation

A total of 24 samples were collected, and genomic DNA was extracted from white blood cell samples using a QIAamp DNA Blood Mini Kit (QIAGEN, Cat. No. 51106). DNA fragments of approximately 200 bp were selected from sheared genomic DNA for library preparation and sequencing. Another 998 Chinese patient samples were collected from Apr. 3, 2018, to Jan. 27, 2019, for HLA typing by an internally developed HLA assay.

HLA Genotyping Assays

HLA genotyping from the amplicon assay NGSgo-AmpX was used as the benchmark reference. NGSgo-AmpX consists of dedicated primer sets for the amplification of individual HLA genes, enabling the amplification of the following HLA genes: Class I: HLA-A, HLA-B, and HLAC-C; and Class II: HLA-DRB1 and HLA-DQB1 (GenDx, Utrecht, Netherlands). Three capture-based assays include 1) Agilent SureSelect Human All Exon V5+UTR kits and paired-end sequencing (150PE) strategies were carried out using standard Illumina protocols on an Illumina HiSeq X10 system (WES for short). Each sample met the average depth over 100X and capture on-target ratio >50%. 2) IDT xGen[®] Exome Research Panel kits and paired-end sequencing (150PE) strategies were carried out using standard Illumina protocols on an Illumina HiSeq X10 system (Bofuri for short). Each sample met the average depth over 100X and capture on-target ratio >60% (10 samples were not available). 3) 3DMed Inc. in-house designed and developed HLA specific probes and paired-end sequencing (150PE) was carried out using standard Illumina protocols on an Illumina HiSeq X10 system (Internal for short). Each sample met the average depth over 100X and capture on-target ratio >60%. The raw fastq files from Miseq sequencing were subsequently processed and validated by the vendor independently, and used as the benchmarked result for HLA typing.

NGS-Based HLA Genotyping Algorithms

We compared seven publicly available algorithms for HLA typing: seq2HLA (16), HLAmminer (17), HLAscan (20), HLA-VBSeq (21), HLA-HD (22), HLAforest (30), and HISAT-genotype (31). The algorithms were chosen considering their accessibility and number of citations. For HLAscan, raw sequence data were first mapped to the human reference genome UCSC hg19, and reads from chr6 of the BAM files were then generated as an input, the database file was directly

downloaded along with the program from github, and other parameters were set to default; for HLA-VBSeq, HLA v2 database and the same instruction on the website were used for HLA typing (<http://nagasakilab.csml.org/hla/>); for HISAT-genotype, we used raw sequence files as an input, and two program “hisatgenotype_extract_reads.py” and “hisatgenotype.py” was used to HLA typing; for HLAmminer, seq2HLA, HLA-HD and HLAforest, raw fastq file was used as input, and all these algorithms were run with default parameters; HLA typing accuracy was defined as the percentage of correctly identified alleles among all the reference alleles. We tested the HLA typing accuracy of all seven algorithms and selected the top three with the highest overall accuracy for our read depth and length evaluation.

Linux Server Hardware Configuration

All software were run on a Linux server (CentOS6.5, kernel version: 2.6.32-431.11.2) with the hardware configuration as follows: Intel(R) Xeon(R) CPU E5-2650 v3 @ 2.30 GHz/250 GB RAM/more than 10 TB disk space. R software was used for statistical analysis and plot creation (version: 3.6.1).

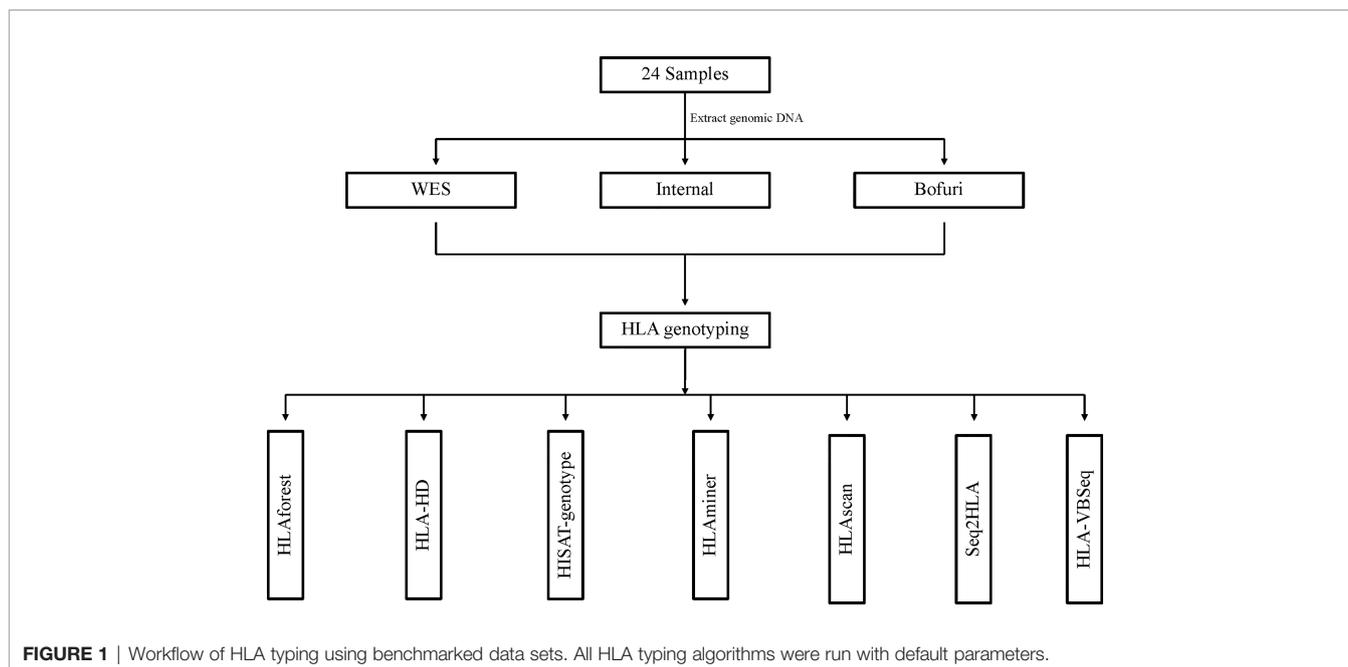
RESULTS

HLA Typing Workflow

Our HLA typing workflow is outlined in **Figure 1**, including DNA isolation, library preparation, high-throughput sequencing, and bioinformatics analysis. Three HLA typing NGS assays—whole-exome sequence (WES), IDT xGen[®] Exome Research Panel (Bofuri), and 3DMed internal panel (Internal)—were selected to generate benchmarked HLA sequencing libraries. Genomic DNA of 24 samples was collected, and then libraries were prepared and sequenced using PE150bp on an Illumina HiSeq X10 system. For the NGS-based HLA genotyping, each sample was determined by seven software, namely seq2HLA, HLAmminer, HLAscan, HLA-VBSeq, HLA-HD, HLAforest, and HISAT-genotype, and default parameters were used for all software. Benchmarking HLA results of the 24 samples (**Supplementary Table 1**) were produced by amplicon assay NGSgo-AmpX plus Miseq sequencing.

HLA Typing Accuracy for All Assay-Software Combinations

As a preliminary screening, we first compared the HLA typing accuracy of all possible assay-software combinations at the first, second, and third field levels. The results were much more discordant among different algorithms than among the capture assays used. At the first field level, six of the seven algorithms had an overall accuracy of higher than 75% no matter which assay was used (**Figure 2A**). HLA-HD and HISAT-genotype had almost perfect accuracy, whereas the accuracy of HLA-VBSeq was lower (the accuracy was 68, 65, and 50% for Internal, WES, and Bofuri, respectively). As the HLA resolution increased from the first field to the second field levels, the accuracy of HLA typing gradually decreased (**Figures 2B, C**; HLA typing results for



HLAminer and HLAscan at the third field level were not available). Only HLA-HD and HISAT-genotype showed greater than 75% accuracy at the third field level. Among all three assays used, the overall accuracy of Bofuri was the lowest, and our internal NGS assays showed comparable performance compared with WES when other algorithms than HISAT-genotype and HLA-HD were used. On the other hand, HISAT-genotype and HLA-HD are quite robust to the assays and resolutions. Above all, our research showed that the choice of the HLA typing algorithms contribute most to the accuracy and target-captured panel could match performance of WES.

Computer Resource Consumption

All HLA programs were run on a Linux server with eight threads if possible. As expected, with the increase in panel sizes of NGS capture assays, the running time for all the software increased (Figure 3). Unsurprisingly, the running time for WES increased exponentially compared with the other two assays (median running time: WES, 77 min; Internal, 4 min; Bofuri, 3 min). For the other two assays, the most time-consuming algorithms were HLA-HD and HISAT-genotype.

Discordant HLA Typing Patterns Across Algorithms

We investigated the specific patterns of discordance in each algorithm. Among all the algorithms, HLA-VBSeq had the highest number of miscalled HLA typing at the second field level, followed by HLAminer (Figure 4A). Out of the five HLA genes, HLA-A gene was the most frequently miscalled gene, and the most discordant pattern was HLA-A*02:07 to HLA-A*02:01 (Figure 4B). Each algorithm had biases on ratios of miscalled HLA typing within specific serological allele groups. For example, 81% (57 out of 70) HLA-A miscalled errors observed

in HLAforest were within the same serological allele group, whereas the ratio decreased to less than 15% for HLAscan, HLA-HD, and HISAT-genotypes (Supplementary Table 2).

The Impact of Sequence Depth and Length on HLA Typing Accuracy

Based on the above evaluations, we focused on the three algorithms with the highest accuracies, that is, HISAT-genotype, HLA-HD, and HLAscan, to investigate the impact of read length and read depth on HLA typing.

Regarding the depth evaluation, when the sequencing data of Bofuri were down-sampled from 700X to 10X, the accuracies of HLA-HD and HISAT-genotype at the second field level were still above 95% at 50X, and then they decreased gradually when the sequence depths were less than 50X (Figure 5A). The overall accuracy of HLAscan was lower than the other two algorithms. The required sequence depth for HLA-HD and HISAT-genotype to get more than 90% HLA typing accuracy was above 100X at both the second and the third field levels (Figure 5B).

Regarding the read length evaluation, we manually generated paired-end 100 bp (PE100) and paired-end 75 bp (PE75) sequence data based on paired-end 150 bp (PE150) using an in-house pipeline which trimmed the sequence from both sides randomly. When the read length decreased from PE150 to PE100 and PE75, the overall HLA typing accuracy was quite similar for each algorithm, except that HLAscan had lower accuracy (Figures 5C, D), which demonstrated that the selected three HLA typing algorithms were robust to the read length.

HLA Typing Performance in Validation Datasets

We selected another 998 Chinese population samples sequenced by the 3DMed internal developed assay. The reference HLA

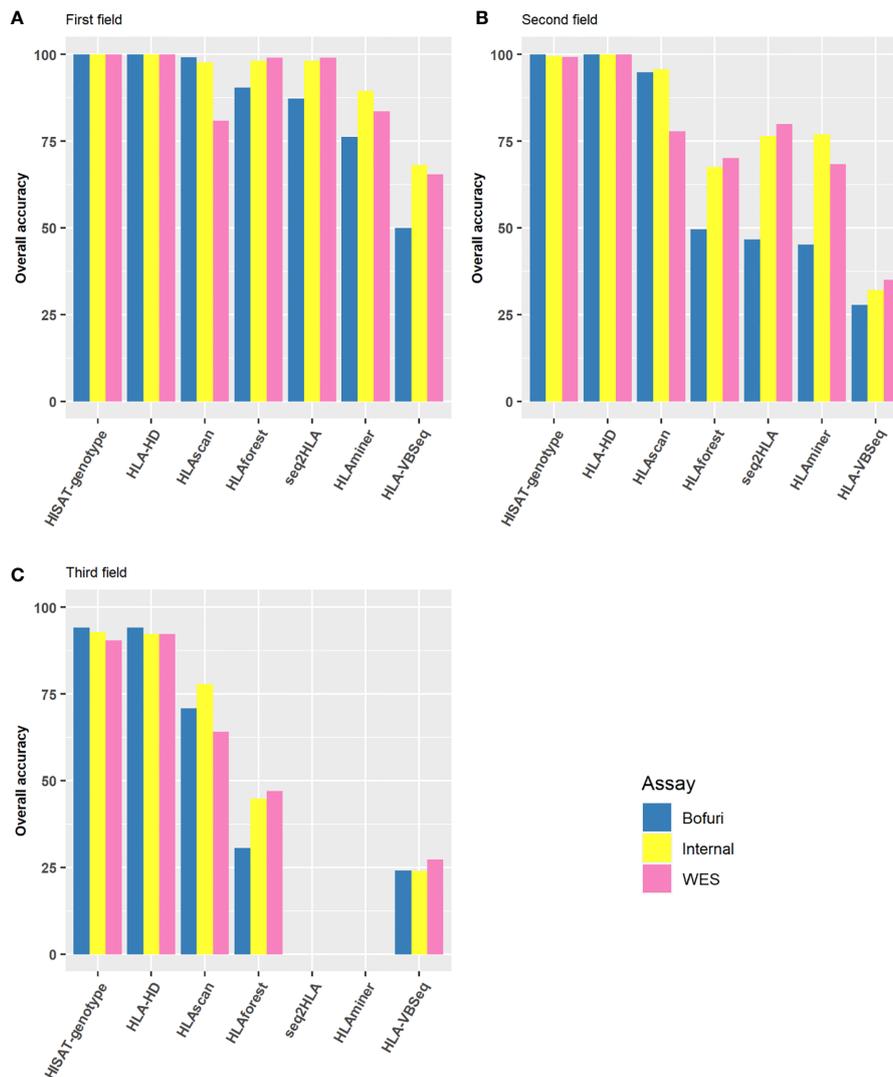


FIGURE 2 | Performance of HLA typing algorithms and the three different HLA assays. Accuracy of HLA alleles typed at **(A)** the first field level; **(B)** the second field level; **(C)** the third field level based on the seven algorithms and three capture assays. Accuracy was calculated by the fraction of total number alleles that were correctly typed.

typing results were defined as the most concordant HLA types called by these seven algorithms. To verify this approach, we compared the most concordant HLA genotypes predicted by different algorithms of the 24 benchmark samples with their reference HLA types at the second field level, and found relatively good consistency between them (The overall accuracy is 0.974, 0.965, 0.970 for internal, WES, and Bofuri respectively), which demonstrated the feasibility of this strategy. For the 998 validation samples, HISAT-genotype, HLA-HD, and HLA-scan showed higher accuracy than other algorithms again, and no obvious difference was found for the five HLA genes when HISAT-genotype, HLA-scan, and HLA-HD were selected (**Figures 6A, B**), reaffirming our comparison results on HLA typing accuracy.

DISCUSSION

In this study, we performed a benchmarked analysis of HLA typing based on seven algorithms and three capture-based sequencing methods. As we stressed, the aim of this study is to identify the best technical combination of NGS-based HLA genotyping in clinical context, rather than to evaluate different algorithms. The algorithms which performed not well in this study could be the best choice in a different application context. We found that the choice of NGS-based HLA typing algorithm and the sequencing depth contributed most to the overall HLA typing accuracy. Among the seven algorithms tested, HLA-HD and HISAT-genotype displayed the highest overall accuracies at both the second field and the third field resolutions, which is of

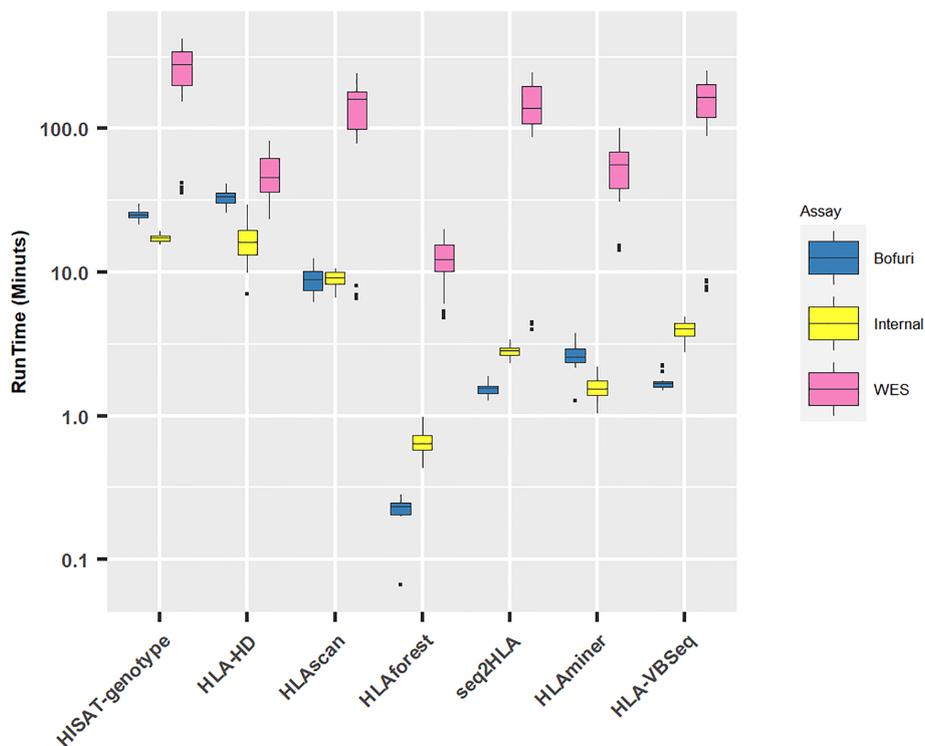


FIGURE 3 | Running time for different HLA typing software. Y axis is plotted in log10 scale.

great importance in clinic, especially in HSCT where a matched donor was found mainly by HLA typing. When no 10/10 (European standard) or 8/8 (USA standard) matched donor is found, looking for 9/10 or 7/8 matched donor is necessary for HSCT (6). Thus, the NGS based HLA typing accuracy for HSCT should be no less than 90% at the high-resolution level. In this case, HISAT-genotype and HLA-HD may be a good choice. HLA-HD constructed an extensive dictionary of HLA alleles and calculated a score based on weighted read counts to select the most suitable pair of alleles (22). The high accuracy of HLA-HD was more likely related to its elaborate reference database. For HISAT-genotype, it had not only higher HLA typing accuracy but also could be used in CYP (cytochrome P450) typing and V(D)J [variable (V), diversity (D), and joining (J) recombination] typing, which have broad clinical applications. Besides, it could provide fourth field level HLA genotyping, although no reference HLA genotype was available to evaluate the accuracy. A recent study showed that the HLA matched status changed in 29% of pairs after ultra-high resolution (UHR) HLA typing using Pacific Biosciences Single Molecule Real-Time sequencing (Menlo Park, CA, USA), and had significant improved clinic benefit (32), demonstrating that allele level typing (or all field typing) using NGS or other technologies might be an important trend in the field of HSCT. Thus, HISAT-genotype may be a good choice in the HSCT field because of its high accuracy and high resolution.

NGS-based HLA typing can type HLA alleles on each homologous chromosome and can function at higher HLA resolutions, but it is also limited by read length and read depth because of the highly polymorphic nature of the HLA system (26). For example, Ka et al. (20) found that read depth is a critical factor for successful HLA typing by HLAScan and recommended a coverage depth over 90X to ensure 100% predictive accuracy for clinical use, whereas in another accuracy evaluation study of five HLA typing methods, only a weak Pearson correlation between HLA typing accuracy and coverage was found (26). In this study, we evaluated the impact of read depth and read length on the HLA typing accuracy of three algorithms, and the result showed that HLA typing accuracy decreased gradually when the sequence depth was down-sampled from 700X to 10X regardless of which algorithm was used, demonstrating that read depth was a critical factor for accurate HLA typing. To achieve more than 90% HLA typing accuracy at the second field level, the minimal read depth was 50X for the three algorithms used, whereas 100X read depth was needed for HLA-HD and HISAT-genotype to obtain 90% overall accuracy at the third field level.

Though HLA genotyping accuracy was generally concordant among the three NGS assays, our internal capture-based assay showed comparable performance compared with WES, no matter which algorithms were used. Our internal assay designed exon probes of 10 HLA genes (HLA-A, HLA-B, HLA-C, HLA-DPA1, HLA-DPB1, HLA-DQA1, HLA-DQB1,

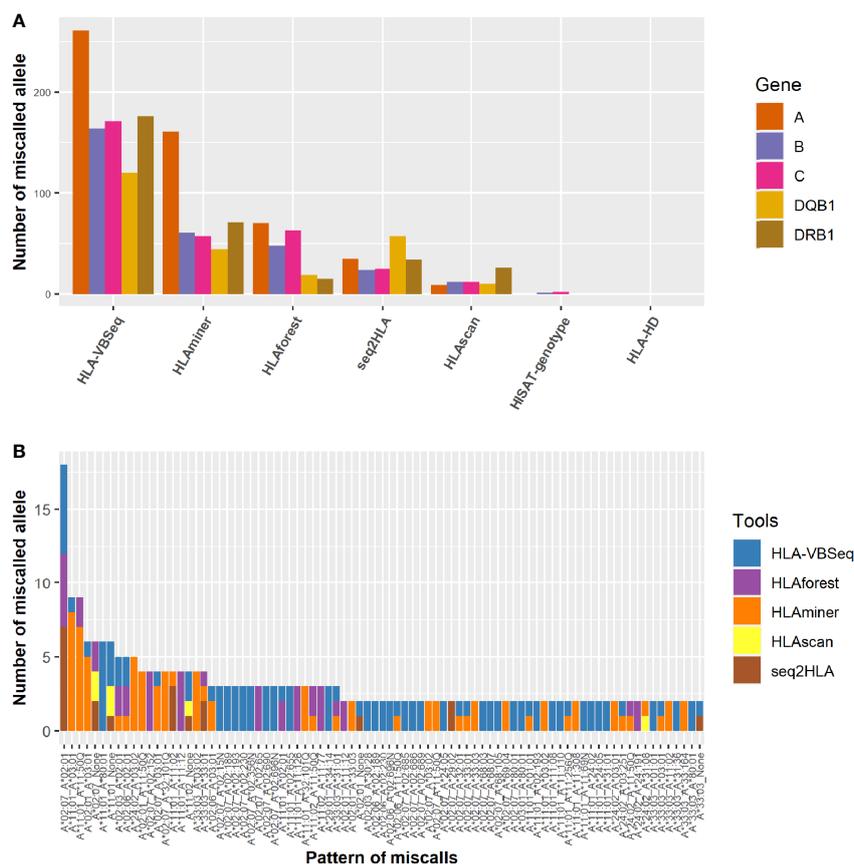


FIGURE 4 | Distribution of the pattern of genotyping errors in HLA-A genes. **(A)** The number of miscalled alleles by each algorithm grouped by the HLA genes. **(B)** The pattern of discordant HLA-A alleles at the second field level. None, not determined by the algorithms.

HLA-DRA, HLA-DRB1, and HLA-DRB5). The geographic range is the union of the coding regions of all possible transcripts of the gene. Design rationale included but were not limited to the following (1): For exons longer than the length of the probe, the target area is completely covered by overlapped probes, and the overlaps are larger than 60 nt (2). Each probe was aligned to the whole genome by BLAT (33). The total score was calculated based on the number of hits. The higher the score, the worse the probe specificity. Probes with scores greater than 2 were not considered (3). Probes were not considered in regions of homologous repeats (e.g., SINE, LINE, LTR, etc.). A well-designed probe may improve probe specificity and HLA exon coverage, thus contributing to the accuracy of NGS-based HLA genotyping. Though the accuracy of HLA typing was similar between WES and capture-based assay, capture-based assay is more cost-effective than WES since it only sequence HLA gene. Besides, the sequencing and data analysis speed of capture-based assay is much faster, which shortens the overall turnaround time and more feasible in clinic.

Different algorithms showed different miscall patterns, with HLA-A*02:07 to HLA-A*02:01 being the most widely miscalled allele by HLAforest, seq2HLA, and HLA-VBSeq. It has been

reported that the only difference in the peptide sequence between HLA-A*02:01 and HLA-A*02:07 is the 123rd amino acid, which is either Tyr or Cys (34), making it difficult to type HLA accurately by less sensitive algorithms. Researchers have also demonstrated that HLA-A*02:07 is the most common HLA-A2 subtype among Chinese (35), and the HLA-A*02:07 peptide binding repertoire is limited to a subset of the HLA-A*02:01 repertoire (36), so we need to pay more attention to this allele in practice when these algorithms are used. Except for HLA-A*02:07 allele, HLA-A*11:01 allele had the second highest frequency of miscall for HLA-A gene family. We found that HLAforest was more prone to miscall HLA-A*02:07 allele, while HLAminer had a higher miscall frequency for HLA-A*11:01 in our benchmarked samples. As for HLA-B gene, HLA-B*13:01 is the most frequently miscalled alleles by HLA-VBSeq and HLAforest, while HLA-B*58:01 is inclined to be miscalled by HLAminer and Seq2HLA. As for HLA-C gene, HLA-C*03:02 and HLA-C*03:03 is inclined to be miscalled by HLAminer and Seq2HLA, while HLA-C*01:02 are more frequently miscalled by HLAforest and HLA-VBSeq (the top two miscall patterns for each gene are summarized in **Supplementary Table 3**). These miscall patterns

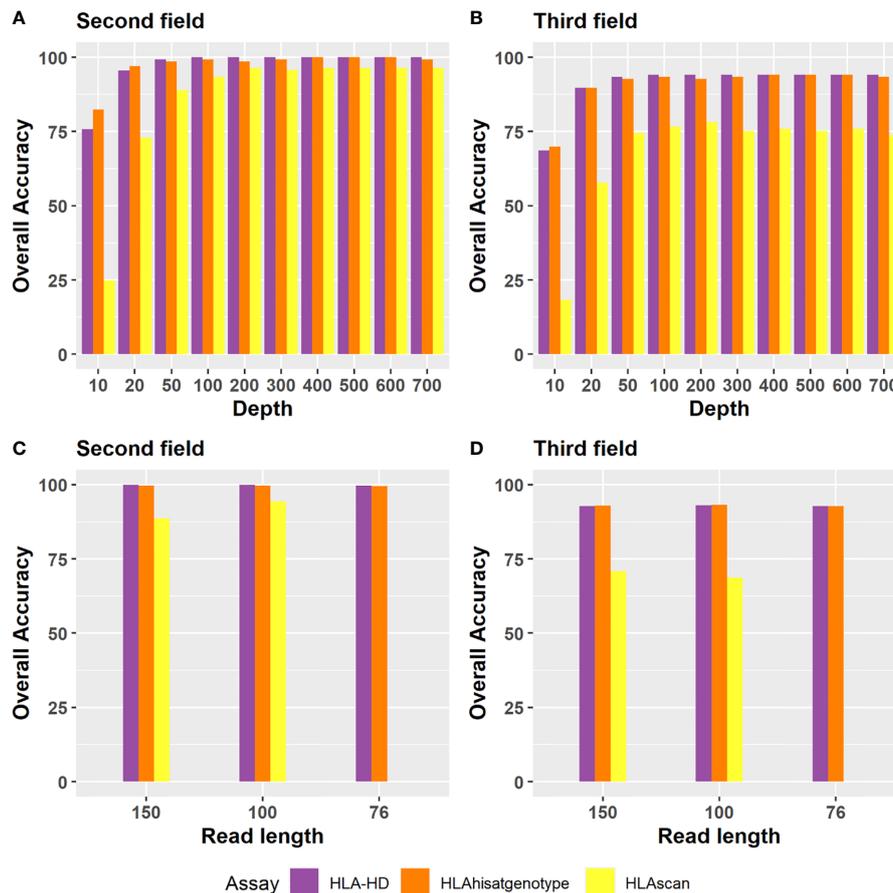


FIGURE 5 | Accuracy of the three tools for HLA typing at the second field or the third field resolution for different depths and read lengths. Depth evaluation at **(A)** the second field level; **(B)** the third field level. For sequence depth evaluation, alignment files of the 24 Bofuri samples were down-sampled from 700X to 10X based on the raw depths of HLA genes. **(C, D)** are the overall HLA typing accuracy at the second field and the third field level, respectively, while the read length decreased from 150 bp to 76 bp.

demonstrated that each algorithm had specific systematical bias, which need to be taken into account when developing more accurate algorithm in future.

One of the drawbacks of this study was that only seven HLA typing algorithms (which were selected considering the ease of use of the software and the number of citations of the corresponding articles) were used in this benchmarking evaluation. For example, Polysolver (37) is not evaluated in this study because it depend on Novoalign, which requires commercial components and is also not executable for us because of the incompatible Linux version. Besides, it is reported that the concordance of HLA typing by the current gold standard methods (PCR-based) is only 84%, reflecting the inaccuracy of the laboratory methods as well as inter-laboratory variability (26). We used NGSgo-AmpX as our benchmarked assay, which is a Research Use Only (RUO) and the only one CE-marked IVD product when our study started, and yielded almost 100% homology results compared to Sanger sequencing (38). In addition, seq2HLA and HLAforest are originally used for RNA-seq based HLA typing, they perform

best on RNAseq data as the datatype they were designed for, whereas the performance in WES/WGS data decreased significantly (26). Finally, all algorithms were run with their default parameters or the default script without any modification, which may not represent the best performance of the algorithms and could affect the accuracy of HLA typing, it may be the case for HLA-VBseq. There is one more thing that should not be overlooked, the default HLA reference sequence file used in our analysis was provided by the software itself and may be derived from different version of IPD-IMGT/HLA reference database, which may compromise the performance of these software. According to the recommendations of American Society for Histocompatibility and Immunogenetics (ASHI) in regards to NGS based HLA typing, the database which is used for HLA typing should be updated at least every 12 months with the most recent version of the IPD-IMGT/HLA database, and for accredited laboratories, documents should record which version of the IPD-IMGT/HLA or other appropriate nucleotide sequence database was used for allele interpretation. Thus, beside HLA

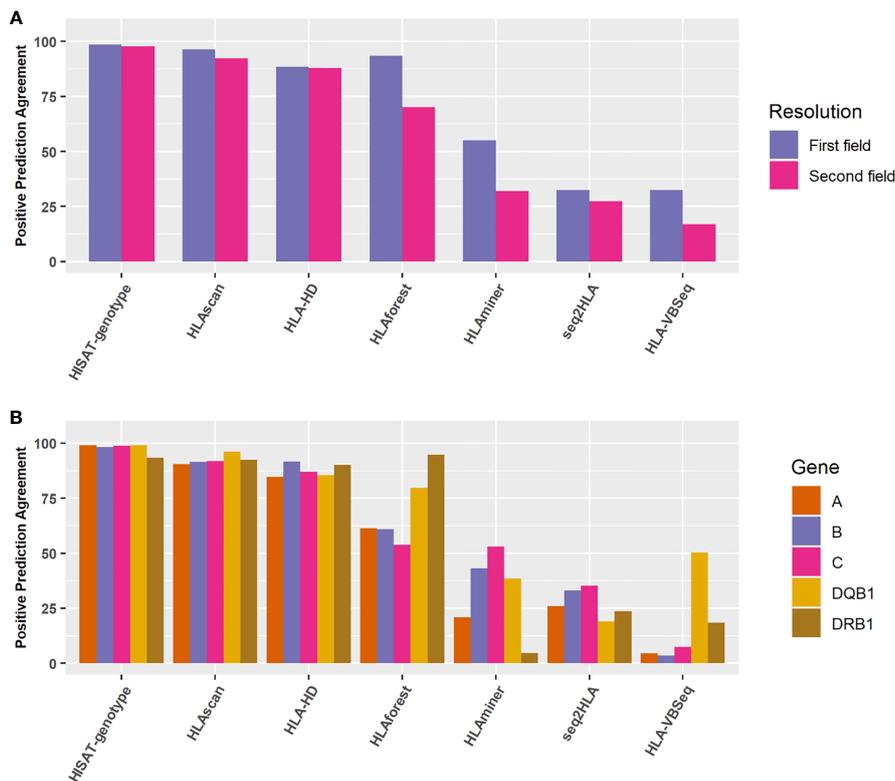


FIGURE 6 | Positive prediction agreement of the seven algorithms for HLA typing at different resolutions and genes. **(A)** Agreement for the seven algorithms at the first field or the second field levels. **(B)** Agreement for the seven algorithms for different HLA genes at the second field level.

typing accuracy and resolution, selecting a suitable software that have good version control and could regularly update the HLA reference database is of great importance and should never be ignored in clinic.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because human genetic resource data should not be publicly available due to the Regulations on Management of Human Genetic Resources (Guo Ling No.717). Requests to access the datasets should be directed to HC, hao.chen1@3dmedcare.com.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Research Ethics Committee of the First Affiliated Hospital, College of Medicine, Zhejiang University (2017KYKSD678-1). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

PL, AZ, YG, YS, and HC conceived this study. PL, AZ, YG, RM, and MY provided samples. YS and YC performed the bioinformatics analyses. PL, YS, HD, FL, YY, XL, and HC wrote the manuscript. AZ provide valuable suggestions and comments during the interactive review. All authors contributed to the article and approved the submitted version.

FUNDING

This work was funded by the National Key Research and Development Program of China (2017YFC0908500).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2021.652258/full#supplementary-material>

REFERENCES

- Kaufman J. Unfinished Business: Evolution of the MHC and the Adaptive Immune System of Jawed Vertebrates. *Annu Rev Immunol* (2018) 36:383–409. doi: 10.1146/annurev-immunol-051116-052450
- Horton R, Wilming L, Rand V, Lovering RC, Bruford EA, Khodiyar VK, et al. Gene map of the extended human MHC. *Nat Rev Genet* (2004) 5(12):889–99. doi: 10.1038/nrg1489
- Trowsdale J, Knight JC. Major histocompatibility complex genomics and human disease. *Annu Rev Genomics Hum Genet* (2013) 14:301–23. doi: 10.1146/annurev-genom-091212-153455
- Morishima Y, Sasazuki T, Inoko H, Juji T, Akaza T, Yamamoto K, et al. The clinical significance of human leukocyte antigen (HLA) allele compatibility in patients receiving a marrow transplant from serologically HLA-A, HLA-B, and HLA-DR matched unrelated donors. *Blood* (2002) 99(11):4200–6. doi: 10.1182/blood.v99.11.4200
- Horowitz MM. High-resolution typing for unrelated donor transplantation: how far do we go? *Best Pract Res Clin Haematol* (2009) 22(4):537–41. doi: 10.1016/j.beha.2009.09.006
- Tiercy JM. How to select the best available related or unrelated donor of hematopoietic stem cells? *Haematologica* (2016) 101(6):680–7. doi: 10.3324/haematol.2015.141119
- Chowell D, Morris LGT, Grigg CM, Weber JK, Samstein RM, Makarov V, et al. Patient HLA class I genotype influences cancer response to checkpoint blockade immunotherapy. *Science* (2018) 359(6375):582–7. doi: 10.1126/science.aao4572
- Shim JH, Kim HS, Cha H, Kim S, Kim TM, Anagnostou V, et al. HLA-corrected tumor mutation burden and homologous recombination deficiency for the prediction of response to PD-(L)1 blockade in advanced non-small-cell lung cancer patients. *Ann Oncol* (2020) 31(7):902–11. doi: 10.1016/j.annonc.2020.04.004
- Li B, Li T, Pignon JC, Wang B, Wang J, Shukla SA, et al. Landscape of tumor-infiltrating T cell repertoire of human cancers. *Nat Genet* (2016) 48(7):725–32. doi: 10.1038/ng.3581
- Lind C, Ferriola D, Mackiewicz K, Heron S, Rogers M, Slavich L, et al. Next-generation sequencing: the solution for high-resolution, unambiguous human leukocyte antigen typing. *Hum Immunol* (2010) 71(10):1033–42. doi: 10.1016/j.humimm.2010.06.016
- Erllich RL, Jia X, Anderson S, Banks E, Gao X, Carrington M, et al. Next-generation sequencing for HLA typing of class I loci. *BMC Genomics* (2011) 12:42. doi: 10.1186/1471-2164-12-42
- Wang C, Krishnakumar S, Wilhelm J, Babrzadeh F, Stepanyan L, Su LF, et al. High-throughput, high-fidelity HLA genotyping with deep sequencing. *Proc Natl Acad Sci U S A* (2012) 109(22):8676–81. doi: 10.1073/pnas.1206614109
- Danzer M, Niklas N, Stabentheiner S, Hofer K, Proll J, Stuckler C, et al. Rapid, scalable and highly automated HLA genotyping using next-generation sequencing: a transition from research to diagnostics. *BMC Genomics* (2013) 14:221. doi: 10.1186/1471-2164-14-221
- Weimer ET, Montgomery M, Petrarolo R, Crawford J, Schmitz JL. Performance Characteristics and Validation of Next-Generation Sequencing for Human Leucocyte Antigen Typing. *J Mol Diagn* (2016) 18(5):668–75. doi: 10.1016/j.jmoldx.2016.03.009
- Lange V, Bohme I, Hofmann J, Lang K, Sauter J, Schone B, et al. Cost-efficient high-throughput HLA typing by MiSeq amplicon sequencing. *BMC Genomics* (2014) 15:63. doi: 10.1186/1471-2164-15-63
- Boegel S, Lower M, Schafer M, Bukur T, de Graaf J, Boisguerin V, et al. HLA typing from RNA-Seq sequence reads. *Genome Med* (2012) 4(12):102. doi: 10.1186/gm403
- Warren RL, Choe G, Freeman DJ, Castellarin M, Munro S, Moore R, et al. Derivation of HLA types from shotgun sequence datasets. *Genome Med* (2012) 4(12):95. doi: 10.1186/gm396
- Szolek A, Schubert B, Mohr C, Sturm M, Feldhahn M, Kohlbacher O. OptiType: precision HLA typing from next-generation sequencing data. *Bioinformatics* (2014) 30(23):3310–6. doi: 10.1093/bioinformatics/btu548
- Huang Y, Yang J, Ying D, Zhang Y, Shotelersuk V, Hirankarn N, et al. HLAreporter: a tool for HLA typing from next generation sequencing data. *Genome Med* (2015) 7(1):25. doi: 10.1186/s13073-015-0145-3
- Ka S, Lee S, Hong J, Cho Y, Sung J, Kim HN, et al. HLAscan: genotyping of the HLA region using next-generation sequencing data. *BMC Bioinf* (2017) 18(1):258. doi: 10.1186/s12859-017-1671-3
- Nariai N, Kojima K, Saito S, Mimori T, Sato Y, Kawai Y, et al. HLA-VBSeq: accurate HLA typing at full resolution from whole-genome sequencing data. *BMC Genomics* (2015) 16 Suppl 2:S7. doi: 10.1186/1471-2164-16-S2-S7
- 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):788–97. doi: 10.1002/humu.23230
- Xie C, Yeo ZX, Wong M, Piper J, Long T, Kirkness EF, et al. Fast and accurate HLA typing from short-read next-generation sequence data with xHLA. *Proc Natl Acad Sci U S A* (2017) 114(30):8059–64. doi: 10.1073/pnas.1707945114
- 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) 43(Database issue):D423–431. doi: 10.1093/nar/gku1161
- Robinson J, Barker DJ, Georgiou X, Cooper MA, Flicek P, Marsh SGE. IPD-IMGT/HLA Database. *Nucleic Acids Res* (2020) 48(D1):D948–55. doi: 10.1093/nar/gkz950
- Bauer DC, Zadoorian A, Wilson LOWMelbourne Genomics Health A. Thorne NP. Evaluation of computational programs to predict HLA genotypes from genomic sequencing data. *Brief Bioinform* (2018) 19(2):179–87. doi: 10.1093/bib/bbw097
- Matey-Hernandez ML, Danish Pan Genome C, Brunak S, Izarzugaza JMG. Benchmarking the HLA typing performance of Polysolver and Optitype in 50 Danish parental trios. *BMC Bioinf* (2018) 19(1):239. doi: 10.1186/s12859-018-2239-6
- 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
- Sverchkova A, Anzar I, Stratford R, Clancy T. Improved HLA typing of Class I and Class II alleles from next-generation sequencing data. *HLA* (2019) 94(6):504–13. doi: 10.1111/tan.13685
- Kim HJ, Pourmand N. HLA typing from RNA-seq data using hierarchical read weighting [corrected]. *PloS One* (2013) 8(6):e67885. doi: 10.1371/journal.pone.0067885
- 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):907–15. doi: 10.1038/s41587-019-0201-4
- Mayor NP, Hayhurst JD, Turner TR, Szydlo RM, Shaw BE, Bulititude WP, et al. Recipients Receiving Better HLA-Matched Hematopoietic Cell Transplantation Grafts, Uncovered by a Novel HLA Typing Method, Have Superior Survival: A Retrospective Study. *Biol Blood Marrow Transplant* (2019) 25(3):443–50. doi: 10.1016/j.bbmt.2018.12.768
- Kent WJ. BLAT—the BLAST-like alignment tool. *Genome Res* (2002) 12(4):656–64. doi: 10.1101/gr.229202
- Shichijo S, Azuma K, Komatsu N, Ito M, Maeda Y, Ishihara Y, et al. Two proliferation-related proteins, TYMS and PGK1, could be new cytotoxic T lymphocyte-directed tumor-associated antigens of HLA-A2+ colon cancer. *Clin Cancer Res* (2004) 10(17):5828–36. doi: 10.1158/1078-0432.CCR-04-0350
- Shieh DC, Lin DT, Yang BS, Kuan HL, Kao KJ. High frequency of HLA-A*0207 subtype in Chinese population. *Transfusion* (1996) 36(9):818–21. doi: 10.1046/j.1537-2995.1996.36996420761.x
- Sidney J, del Guercio MF, Southwood S, Hermanson G, Maewal A, Appella E, et al. The HLA-A*0207 peptide binding repertoire is limited to a subset of the A*0201 repertoire. *Hum Immunol* (1997) 58(1):12–20. doi: 10.1016/s0198-8859(97)00206-1
- Shukla SA, Rooney MS, Rajasagi M, Tiao G, Dixon PM, Lawrence MS, et al. Comprehensive analysis of cancer-associated somatic mutations in class I HLA genes. *Nat Biotechnol* (2015) 33(11):1152–8. doi: 10.1038/nbt.3344
- Glotov O, Romanova O, Eismont Y, Sarana A, Scherbak S, Kuzmich E, et al. Comparative analysis of NGS and Sanger sequencing methods for

HLA typing at a Russian university clinic. *Cell Ther Transplantation* (2018) 7:72–82. doi: 10.18620/ctt-1866-8836-2018-7-4-72-82

Conflict of Interest: All authors affiliated with 3D Medicines Inc. are current or former employees. There are no patents, products in development or marketed products to declare. The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Liu, Yao, Gong, Song, Chen, Ye, Liu, Li, Dong, Meng, Chen and Zheng. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.