

# Benchmarking the HLA typing performance of three HLA assays and seven NGS-based HLA algorithms

**Ping Liu**

Department of Oncology, the Second Xiangya Hospital of Central South University

**Aiwen Zheng**

Cancer Hospital of the University of Chinese Academy of Science

**Yu Gong**

Department of Urology, Second Affiliated Hospital, Zhejiang University College of Medicine

**Yunjie Song**

3DMed Inc. <https://orcid.org/0000-0002-9305-9161>

**Yanan Chen**

3DMed Inc.

**Yizhou Ye**

3DMed Inc.

**Xiao Liu**

3DMed Inc.

**Fugen Li**

3DMed Inc.

**Hua Dong**

3DMed Inc.

**Rui Meng** (✉ [mengruivip@163.com](mailto:mengruivip@163.com))

Tongji Medical College, Huazhong University of Science and Technology

**Hao Chen** (✉ [hao.chen1@3dmedcare.com](mailto:hao.chen1@3dmedcare.com))

3DMed Inc. <https://orcid.org/0000-0002-0201-8066>

**Minya Yao** (✉ [yminya@163.com](mailto:yminya@163.com))

The First Affiliated Hospital of Zhejiang University, School of Medicine

---

## Research article

**Keywords:** HLA assays, HLA genotyping algorithms, NGS

**Posted Date:** July 8th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-35044/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

## Background

With the great progress made recently in NGS (Next Generation Sequencing) technology, sequencing accuracy and throughput have increased, while the cost for data has decreased. Various HLA (Human Leukocyte Antigen) typing algorithms and assays have been developed and have begun to be used in clinical practice. However, there is no systematic benchmarking to evaluate the HLA typing performance of different HLA assays and algorithms. 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.

## Results

Seven HLA typing algorithms at 4- and 6-digit allele levels were compared on three different assays in terms of accuracy, read depth and read length. The algorithms HLA-HD and HISAT-genotype showed the highest accuracy at both 2- and 4-digit resolution, followed by HLAscan. We designed a capture-based HLA assay, which showed comparable or even better performance compared with WES (Whole Exome Sequencing). In the depth evaluation, the sequencing data were down-sampled from 500X to 10X based on the depth of HLA genes. We found that the minimal depth was 100X for HLA-HD and HISAT-genotype to obtain more than 90% HLA typing accuracy at the 6-digit allele level. The accuracy of all three algorithms did not change when the read length decreased from 150 bp to 76 bp.

## Conclusion

Although HISAT-genotype and HLA-HD may need more computing resources, we recommend using them for NGS-based HLA genotyping because of their higher accuracy and robustness to sequencing depth and read length. We propose that the minimal sequence depth for obtaining more than 90% HLA typing accuracy at the 6-digit allele 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.

## Background

The human leukocyte antigen (HLA), commonly referred to as major histocompatibility complex (MHC), 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[1]. 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. MHC 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-I, HLA-II and HLA-III. HLA-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 [2]. HLA-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-III gene contains approximately 75 genes, most of which are of unknown function. HLA-I (MHC I) and HLA-II (MHC 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-I genes mainly encode antigens to CD8 + T cells, and HLA-II genes mainly encode antigens to CD4 + T cells.

HLA has been widely used in bone marrow transplantation, detection of susceptibility genes in immune-related diseases and drug allergy testing. Recent studies have demonstrated that HLA typing complexity is associated with the efficacy of cancer immune checkpoint blockade (ICB) [3]. 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. 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. 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 [4]. The highly polymorphic HLA genes present unique challenges for the development of molecular approaches to genotyping 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 [5–10]. Many bioinformatics approaches have also been developed to produce HLA genotyping information from amplicon-based NGS, targeted capture including whole-exome sequencing and non-targeted whole-genome sequencing [11–18] (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 [19, 20] and predicts HLA genotypes using probabilistic models [21], 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 [21–24]. 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 sequencing coverage has a weak correlation with accuracy [21]. However, no conclusions have been made regarding several critical questions: Which HLA typing assay is more suitable in a clinical context? Are any HLA typing algorithms 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.

## Results

### HLA typing workflow

Our HLA typing workflow is outlined in Fig. 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 according to the manufacturer's protocol. For the NGS-based HLA genotyping, each sample was determined by seven software, namely seq2HLA [11], HLAmminer [12], HLAscan [15], HLA-VBSeq [16], HLA-HD [17], HLAforest [25] and HISAT-genotype [26], and default parameters were used for all software. Benchmarking HLA results of the 24 samples were produced by amplicon assay NGSgo-AmpX plus Miseq sequencing (Supplementary Table 1).

### 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 2-, 4- and 6-digit allele levels. The results were much more discordant among different algorithms than among the capture assays used. At the 2-digit allele level, six of the seven algorithms had an overall accuracy of higher than 75% no matter which assay was used (Fig. 2A). HLA-HD and HISAT-genotype had almost perfect accuracy, whereas the accuracy of HLAVBseq was much lower (the accuracy was 68%, 65% and 50% for Internal, WES and Bofuri, respectively). Among all three assays used, the overall accuracy of Bofuri was the lowest, and our internal NGS assays showed comparable or even better performance compared with WES. As the HLA resolution increased from 2- to 6-digit allele levels, the accuracy of HLA typing gradually decreased (Fig. 2B and 2C; HLA typing results for HLAmminer and HLAseq2HLA at the 6-digit allele level were not available). Only HLA-HD and HISAT-genotype showed greater than 75% accuracy at the 6-digit allele level. Thus, the combination of the HLA-HD/HISAT-genotype with our internal assay/WES showed the highest HLA typing accuracy.

### Computer resource consumption

All HLA programs were run on a Linux server with the maximum eight threads if possible. As expected, with the increase in panel sizes of NGS capture assays, the running times for all of the software increased (Fig. 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). In 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 4-digit allele level, followed by HLAMiner (Fig. 4A). Out of the five HLA genes, HLA-A gene was the most frequently miscalled gene, and the most discordant pattern was A\*02:07 to A\*02:01 (Fig. 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 4-digit allele level were still above 95% at 50X and higher read depths, and then they decreased gradually when the sequence depths were less than 50X (Fig. 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 4- and 6-digit allele levels (Fig. 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. 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, as in the depth research (Fig. 5C and 5D), demonstrating that the selected three HLA typing algorithms were robust to the read length.

## HLA typing performance in validation data sets

We selected another 998 Chinese population samples sequenced by the 3DMed internal developed assay. The reference HLA typing results were defined as the most concordant HLA types called by these seven algorithms. HISAT-genotype, HLA-HD and HLAScan showed higher accuracy than other algorithms again, and no obvious difference was found for the five HLA genes when these algorithms were selected (Fig. 6A and 6B), 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. 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 4-digit and 6-digit

allele levels. 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 [17]. The high accuracy of HLA-HD was more likely related to its elaborate reference database. For HISAT-genotype, it not only had 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 8-digit allele level HLA genotyping, although no reference HLA genotype was available to evaluate the accuracy.

Though NGS-based HLA typing can type HLA alleles on each homologous chromosome and can function at higher HLA resolutions, it is also limited by read length and read depth because of the highly polymorphic nature of the HLA system [21]. For example, Ka et al. [15] 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 [21]. 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 4-digit 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 6-digit level.

Though HLA genotyping accuracy was generally concordant among the three NGS assays, our internal capture-based assay showed comparable or even better 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, 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 [27]. 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.

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\*0201 and HLA-A\*0207 is the 123rd amino acid, which is either Tyr or Cys [28], making it difficult to type HLA accurately by less sensitive algorithms. Researchers have also demonstrated that HLA-A\*0207 is the most common HLA-A2 subtype among

Chinese [29], and the HLA-A0207 peptide binding repertoire is limited to a subset of the A0201 repertoire [30], so we need to pay more attention to this allele in practice when these algorithms are used.

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 [31] was not evaluated in this study because it depends on Novoalign, which requires commercial components and hence was not executable for us. Besides, all algorithms were run with their default parameters, which may not represent the best performance of the algorithms.

## Conclusions

In conclusion, the choice of algorithm and sequencing depth are very important to the accuracy of HLA typing. Among all the algorithms evaluated here, we recommend using HISAT-genotype or HLA-HD for NGS-based HLA genotyping because of their higher accuracy and robustness to sequencing depth and read length. We also propose a minimal sequence depth of 100X for obtaining more than 90% HLA typing accuracy at the 6-digit allele level.

## 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, The Netherlands). Three capture-based assays include 1) Agilent SureSelect Human All Exon V5 + UTR kits according to the Menu and paired-end sequencing (150PE) was 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 according to the Menu and paired-end sequencing (150PE) was 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%.

### NGS-based HLA genotyping algorithms

We compared seven publicly available algorithms for HLA typing: seq2HLA, HLAMiner, HLAscan, HLA-VBSeq, HLA-HD, HLAforest and HISAT-genotype. The algorithms were chosen considering their accessibility and number of citations. All algorithms were run according to their respective manuals with default parameters. For HLAscan and HLA-VBSeq analysis, 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. For the other algorithms, we used raw sequence files as an input. 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 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).

## Abbreviations

NGS: next-generation sequencing; HLA: human leukocyte antigen; MHC: major histocompatibility complex; ICB: immune checkpoint blockade; TMB: tumor mutation burden

## Declarations

### Ethics approval and consent to participate

The study protocol was reviewed and approved by the Research Ethics Committee of the First Affiliated Hospital, College of Medicine, Zhejiang University (Number: 2017KYKSD678-1). Written informed consent for genetic testing was obtained from each participant.

### Consent for publication

Written consents were obtained from each participant to publish their details.

### Availability of data and materials

The raw sequencing data (FASTQ) generated during this study are not publicly available due to the policy in China. Benchmarking HLA typing results of the 24 samples and the number of miscalled HLA genotypes used in this studies are available as supplemental data.

## **Competing interests**

All authors affiliated with 3D Medicines Inc. are current or former employees. There are no patents, products in development or marketed products to declare. No potential conflicts of interest were disclosed by the other authors.

## **Funding**

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

## **Authors' contributions**

PL, AWZ, YG, YJS and HC conceived this study. PL, AWZ, YG, RM and MYY provided samples. YJS and YNC performed the bioinformatics analyses. PL, YJS, HD, FGL, ZYY, XL and HC wrote the manuscript. All authors read and approved the final manuscript.

## **Acknowledgments**

Not applicable.

## **Author's Information**

Ping Liu, Aiwen Zheng, Yu Gong and Yunjie Song contributed equally to this work.

## **Corresponding authors**

Correspondence to Rui Meng, Hao Chen, Minya Yao.

<sup>1</sup>Department of Oncology, the Second Xiangya Hospital of Central South University, Changsha, Hunan, China;

<sup>2</sup>Cancer Hospital of the University of Chinese Academy of Sciences (Zhejiang Cancer Hospital), Hangzhou, China;

<sup>3</sup>Institute of Cancer and Basic Medicine (IBMC), Chinese Academy of Sciences, Hangzhou, China;

<sup>4</sup>Department of Urology, Second Affiliated Hospital, Zhejiang University College of Medicine, Hangzhou, China;

<sup>5</sup>3DMed Inc., Shanghai, China;

<sup>6</sup>Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China;

<sup>7</sup>The First Affiliated Hospital of Zhejiang University, School of Medicine, Hangzhou, China

## References

1. Horton R, Wilming L, Rand V, Lovering RC, Bruford EA, Khodiyar VK, Lush MJ, Povey S, Talbot CC Jr, Wright MW, et al. Gene map of the extended human MHC. *Nat Rev Genet.* 2004;5(12):889–99.
2. Trowsdale J, Knight JC. Major histocompatibility complex genomics and human disease. *Annu Rev Genomics Hum Genet.* 2013;14:301–23.
3. Chowell D, Morris LGT, Grigg CM, Weber JK, Samstein RM, Makarov V, Kuo F, Kendall SM, Requena D, Riaz N, et al. Patient HLA class I genotype influences cancer response to checkpoint blockade immunotherapy. *Science.* 2018;359(6375):582–7.
4. Efremova M, Finotello F, Rieder D, Trajanoski Z. Neoantigens Generated by Individual Mutations and Their Role in Cancer Immunity and Immunotherapy. *Front Immunol.* 2017;8:1679.
5. Lind C, Ferriola D, Mackiewicz K, Heron S, Rogers M, Slavich L, Walker R, Hsiao T, McLaughlin L, D'Arcy M, et al. Next-generation sequencing: the solution for high-resolution, unambiguous human leukocyte antigen typing. *Hum Immunol.* 2010;71(10):1033–42.
6. Erlich RL, Jia X, Anderson S, Banks E, Gao X, Carrington M, Gupta N, DePristo MA, Henn MR, Lennon NJ, et al. Next-generation sequencing for HLA typing of class I loci. *BMC Genom.* 2011;12:42.
7. Wang C, Krishnakumar S, Wilhelmy J, Babrzadeh F, Stepanyan L, Su LF, Levinson D, Fernandez-Vina MA, Davis RW, Davis MM, et al. High-throughput, high-fidelity HLA genotyping with deep sequencing. *Proc Natl Acad Sci U S A.* 2012;109(22):8676–81.
8. Danzer M, Niklas N, Stabentheiner S, Hofer K, Proll J, Stuckler C, Raml E, Polin H, Gabriel C. Rapid, scalable and highly automated HLA genotyping using next-generation sequencing: a transition from research to diagnostics. *BMC Genom.* 2013;14:221.
9. Weimer ET, Montgomery M, Petraroia 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.
10. Lange V, Bohme I, Hofmann J, Lang K, Sauter J, Schone B, Paul P, Albrecht V, Andreas JM, Baier DM, et al. Cost-efficient high-throughput HLA typing by MiSeq amplicon sequencing. *BMC Genom.* 2014;15:63.
11. Boegel S, Lower M, Schafer M, Bukur T, de Graaf J, Boisguerin V, Tureci O, Diken M, Castle JC, Sahin U. HLA typing from RNA-Seq sequence reads. *Genome Med.* 2012;4(12):102.
12. Warren RL, Choe G, Freeman DJ, Castellarin M, Munro S, Moore R, Holt RA. Derivation of HLA types from shotgun sequence datasets. *Genome Med.* 2012;4(12):95.
13. 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.
14. Huang Y, Yang J, Ying D, Zhang Y, Shotelersuk V, Hirankarn N, Sham PC, Lau YL, Yang W. HLAReporter: a tool for HLA typing from next generation sequencing data. *Genome Med.* 2015;7(1):25.

15. Ka S, Lee S, Hong J, Cho Y, Sung J, Kim HN, Kim HL, Jung J. HLAScan: genotyping of the HLA region using next-generation sequencing data. *BMC Bioinformatics*. 2017;18(1):258.
16. Nariai N, Kojima K, Saito S, Mimori T, Sato Y, Kawai Y, Yamaguchi-Kabata Y, Yasuda J, Nagasaki M. HLA-VBSeq: accurate HLA typing at full resolution from whole-genome sequencing data. *BMC Genom*. 2015;16(Suppl 2):7.
17. 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.
18. Xie C, Yeo ZX, Wong M, Piper J, Long T, Kirkness EF, Biggs WH, Bloom K, Spellman S, Vierra-Green C, 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.
19. 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):D423–31.
20. Robinson J, Barker DJ, Georgiou X, Cooper MA, Flicek P, Marsh SGE. IPD-IMGT/HLA Database. *Nucleic Acids Res*. 2020;48(D1):D948–55.
21. Bauer DC, Zadoorian A, Wilson LOW, Melbourne Genomics Health A, Thorne NP. Evaluation of computational programs to predict HLA genotypes from genomic sequencing data. *Brief Bioinform*. 2018;19(2):179–87.
22. 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 Bioinformatics*. 2018;19(1):239.
23. 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.
24. 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.
25. Kim HJ, Pourmand N. **HLA typing from RNA-seq data using hierarchical read weighting** *PLoS One* 2013, 8(6):e67885.
26. 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.
27. Kent WJ. BLAT—the BLAST-like alignment tool. *Genome Res*. 2002;12(4):656–64.
28. Shichijo S, Azuma K, Komatsu N, Ito M, Maeda Y, Ishihara Y, Itoh K. 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.
29. 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.
30. Sidney J, del Guercio MF, Southwood S, Hermanson G, Maewal A, Appella E, Sette A. The HLA-A\*0207 peptide binding repertoire is limited to a subset of the A\*0201 repertoire. *Hum Immunol*. 1997;58(1):12–20.

31. Shukla SA, Rooney MS, Rajasagi M, Tiao G, Dixon PM, Lawrence MS, Stevens J, Lane WJ, Dellagatta JL, Steelman S, et al. Comprehensive analysis of cancer-associated somatic mutations in class I HLA genes. Nat Biotechnol. 2015;33(11):1152–8.

## Tables

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

## Figures

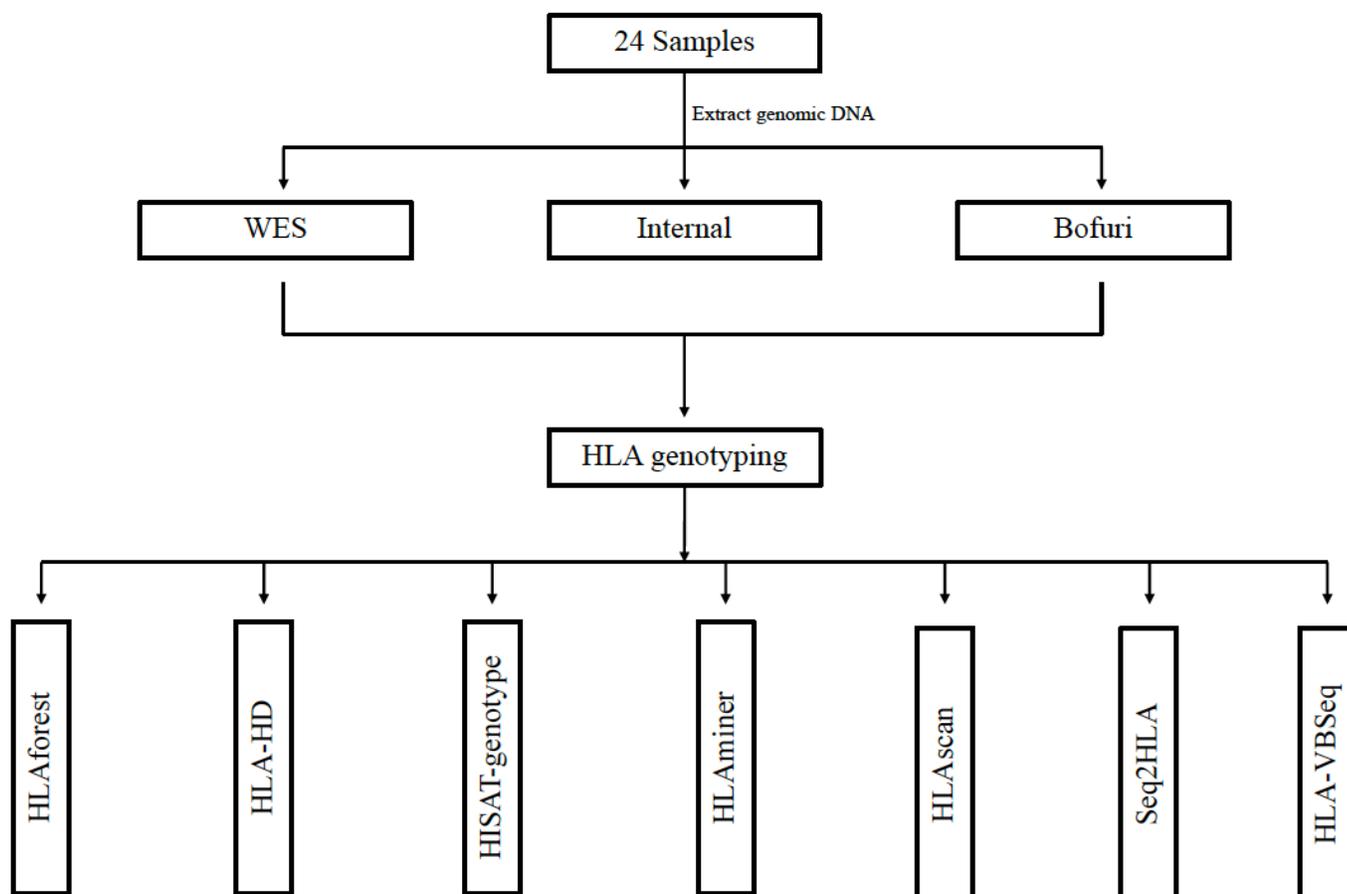


Figure 1

Workflow of HLA typing using benchmarked data sets. All HLA typing algorithms were run with default parameters.

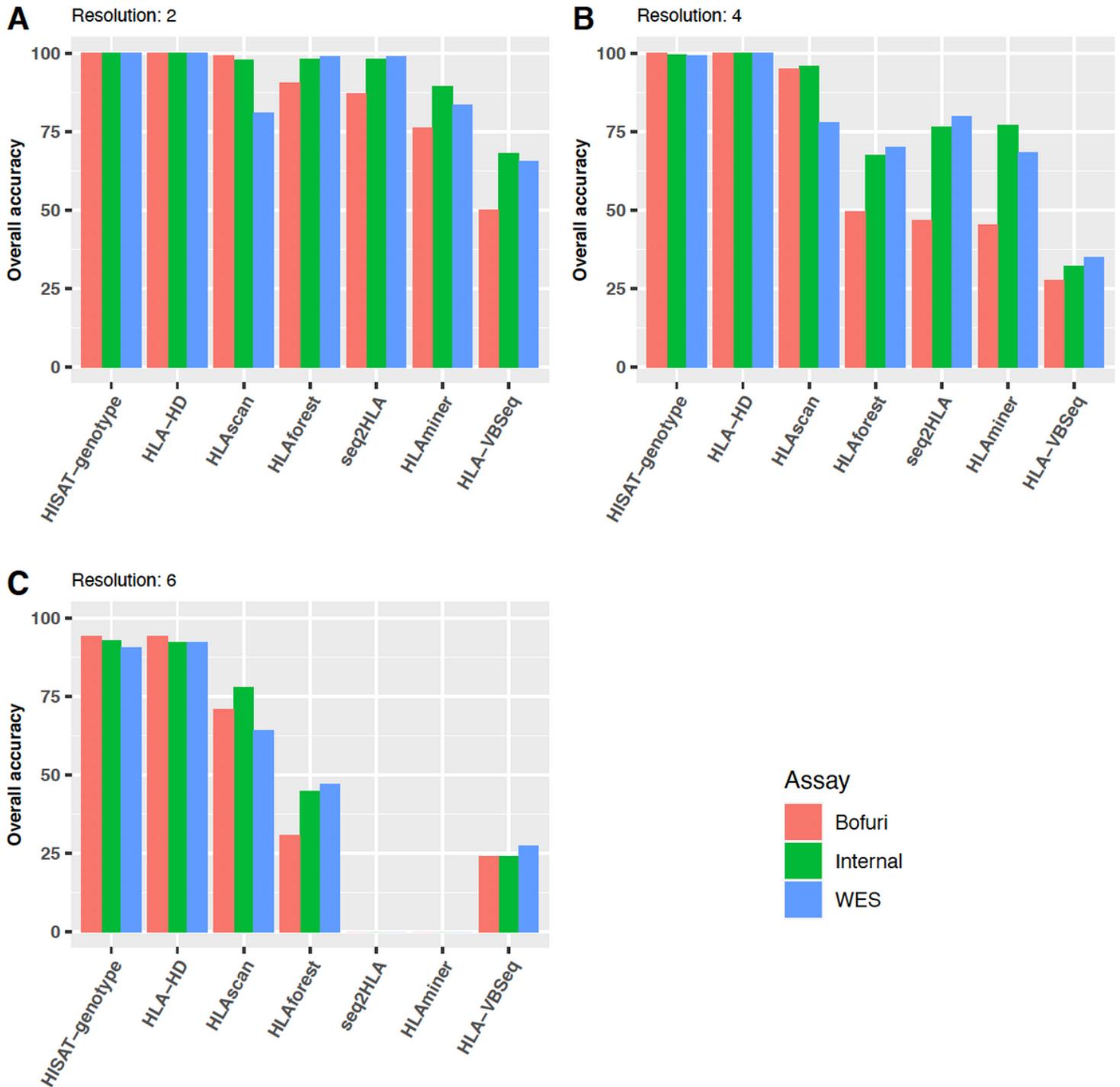


Figure 2

Performance of HLA typing algorithms and the three different HLA assays. Accuracy of HLA alleles typed at A) 2-digit allele level; B) 4-digit allele level; C) 6-digit allele level based on the seven algorithms and

three capture assays. Accuracy was calculated by the fraction of total number alleles that were correctly typed.

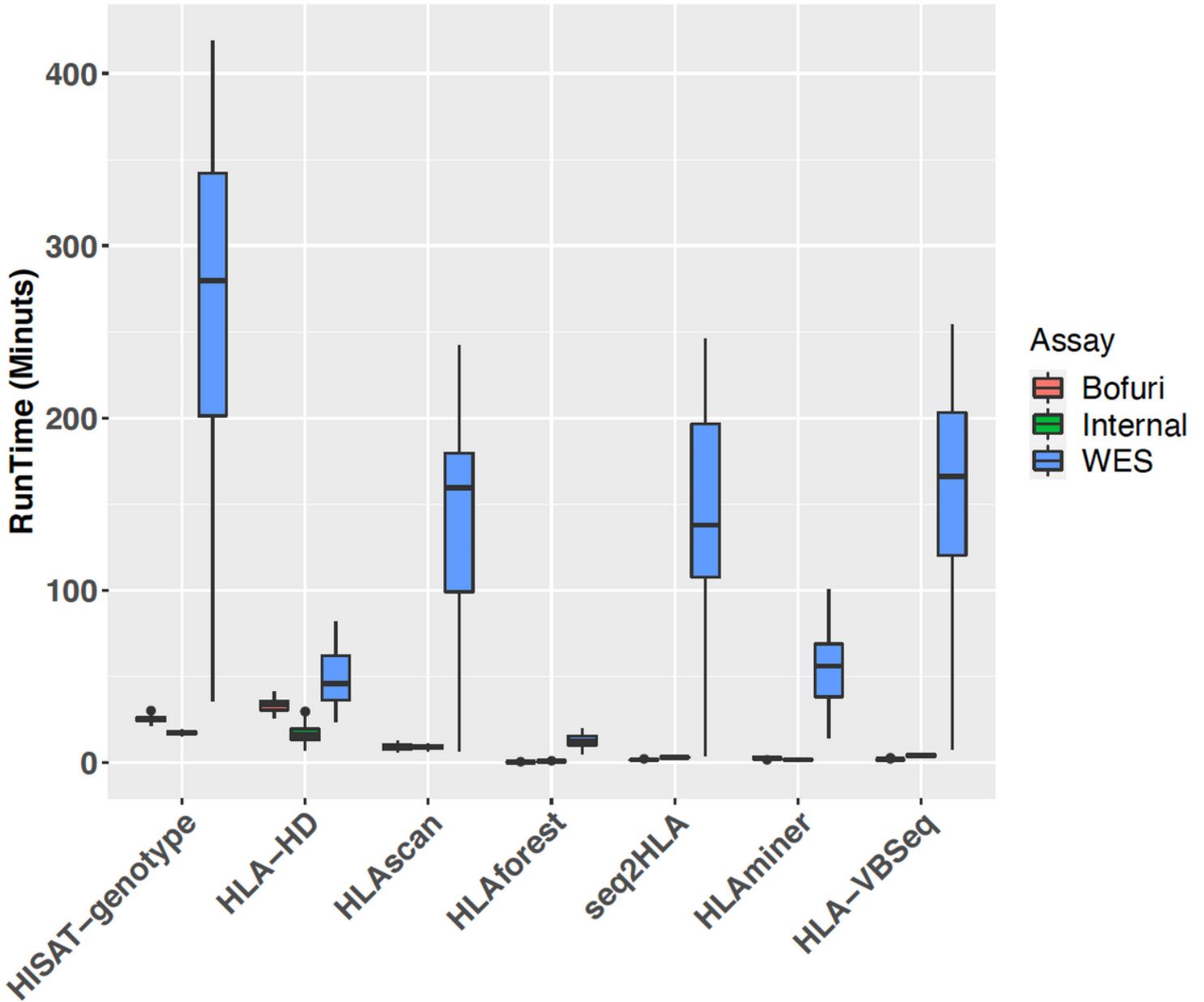
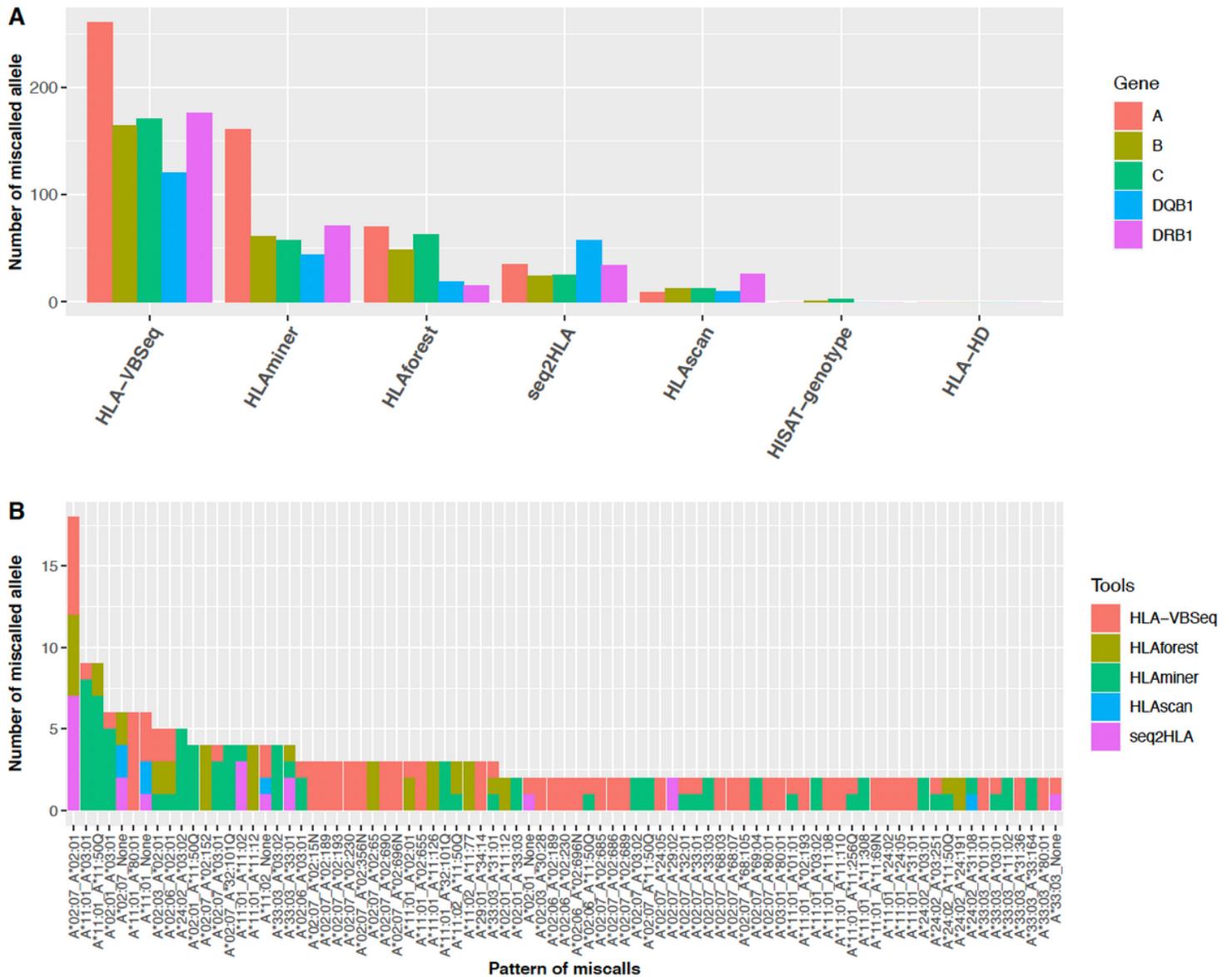


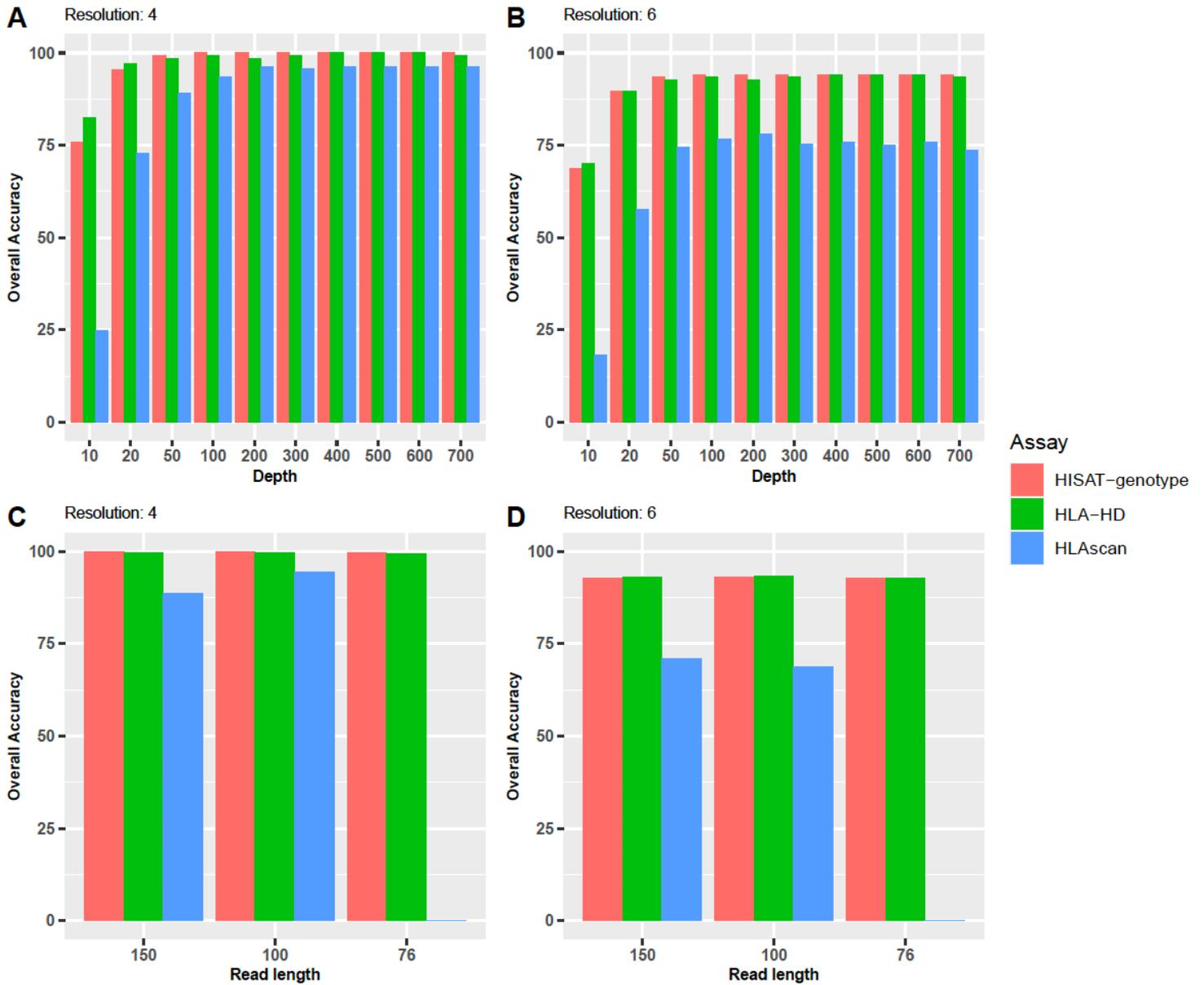
Figure 3

Running time for different HLA typing software.



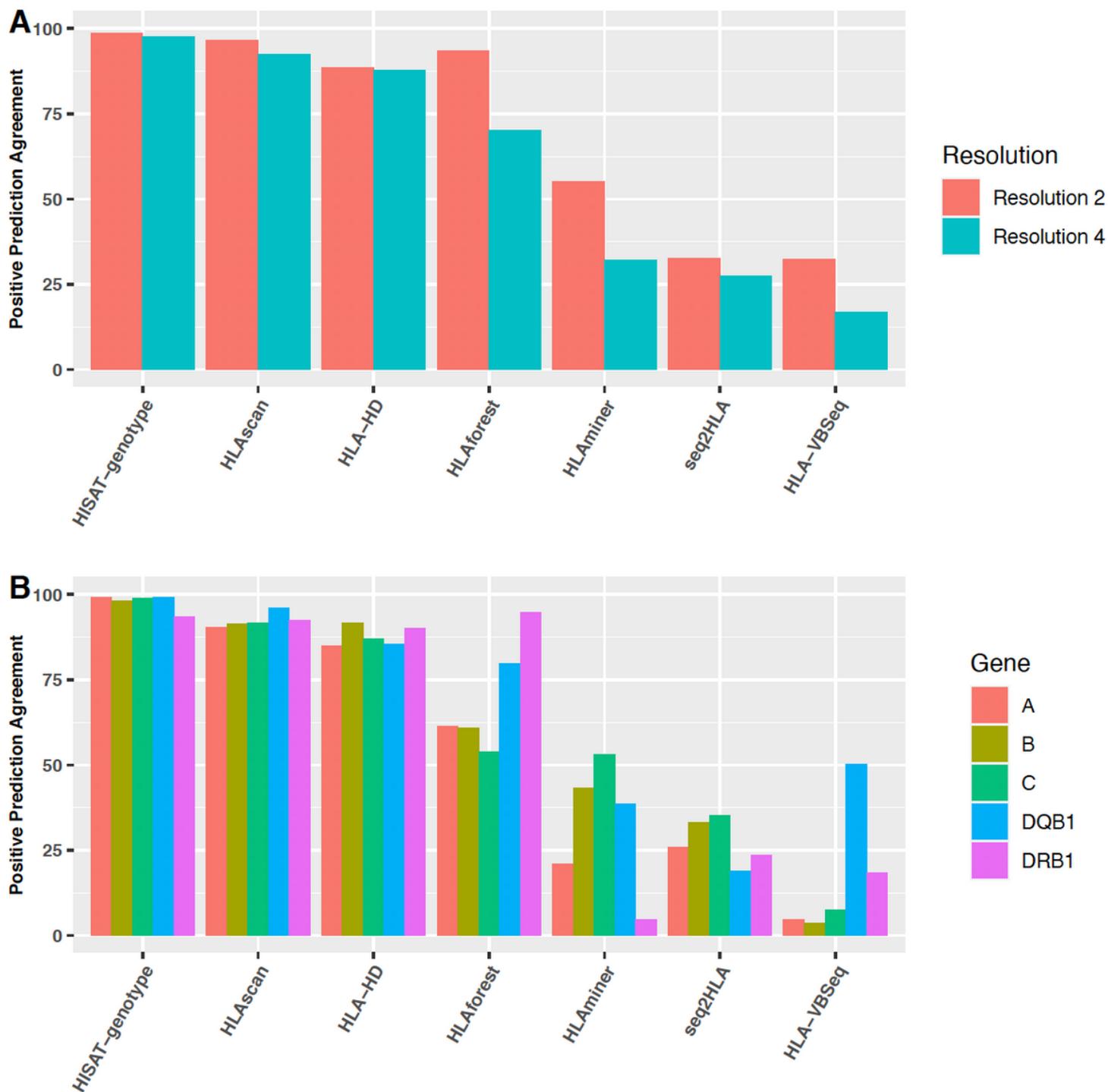
**Figure 4**

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 4-digit allele level. None: not determined by the algorithms.



**Figure 5**

Accuracy of the three tools for HLA typing at 4-digit allele or 6-digit allele resolution for different depths and read lengths. Depth evaluation at A) 4-digit allele level; B) 6-digit allele 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) and D) are the overall HLA typing accuracy at 4-digit allele and 6-digit allele level, respectively, while the read length decreased from 150 bp to 76 bp.



**Figure 6**

Positive prediction agreement of the seven algorithms for HLA typing at different resolutions and genes. A) Agreement for the seven algorithms at 2- or 4-digit allele levels. B) Agreement for the seven algorithms for different HLA genes at 4-digit allele level.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementaryfiles.xlsx](#)