

RESEARCH

Open Access



# A computational HLA allele-typing protocol to de-noise and leverage nanopore amplicon data

Jalal Siddiqui<sup>1\*</sup>, Rohita Sinha<sup>1</sup>, James Grantham<sup>1</sup>, Ronnie LaCombe<sup>1</sup>, Judith R. Alonzo<sup>1</sup>, Scott Cowden<sup>1</sup> and Steven Kleiboeker<sup>1</sup>

## Abstract

**Background** Rapid turnaround time for a third-field resolution deceased donor human leukocyte antigen (HLA) typing is critical to improve organ transplantation outcomes. Third generation DNA sequencing platforms such as Oxford Nanopore (ONT) offer the opportunity to deliver rapid results at single nucleotide level resolution, in particular sequencing data that could be denoised computationally. Here we present a computational pipeline for up-to third-field HLA allele typing following ONT sequencing.

**Results** From a R10.3 flow cell batch of 31 samples of known HLA allele types, up to 10,000 ONT reads were aligned using BWA aligner to reference allele sequences from the IPD-IMGT/HLA database. For each gene, the top two hits to reference alleles at the third field were selected. Using our pipeline, we obtained the following percent concordance at the 1st, 2nd and 3rd field: HLA-A (98.4%, 98.4%, 98.4%), HLA-B (100%, 96.8%, 96.8%), HLA-C (100%, 98.4%, 98.4%), HLA-DPA1 (100%, 96.8%, 96.8%), HLA-DPB1 (100%, 100%, 98.4%), HLA-DQA1 (100%, 98.4%, 98.4%), HLA-DQB1 (100%, 98.4%, 98.4%), HLA-DRB1 (83.9%, 64.5%, 64.5%), HLA-DRB3 (82.6%, 73.9%, 73.9%), HLA-DRB4 (100%, 100%, 100%) and HLA-DRB5 (100%, 100%, 100%). By running our pipeline on an additional R10.3 flow cell batch of 63 samples, the following percent concordances were obtained: HLA-A (100%, 96.8%, 88.1%), HLA-B (100%, 90.5.4%, 88.1%), HLA-C (100%, 99.2%, 99.2%), HLA-DPA1 (100%, 98.4%, 97.6%), HLA-DPB1 (98.4%, 97.6%, 92.9%), HLA-DQA1 (100%, 100%, 98.4%), HLA-DQB1 (100%, 97.6%, 96.0%), HLA-DRB1 (88.9%, 68.3%, 68.3%), HLA-DRB3 (81.0%, 61.9%, 61.9%), HLA-DRB4 (100%, 97.4%, 94.7%) and HLA-DRB5 (73.3%, 66.7%, 66.7%). In addition, our pipeline demonstrated significantly improved concordance compared to publicly available pipeline HLA-LA and concordances close to Athlon2 in commercial development.

**Conclusion** Our algorithm had a > 96% concordance for non-HLA-DRB genes at 3rd field on the first batch and > 88% concordance for non-HLA-DRB genes at 3rd field and > 90% at 2nd field on the second batch tested. In addition, it out-performs HLA-LA and approaches the performance of the Athlon2. This lays groundwork for better utilizing Nanopore sequencing data for HLA typing especially in improving organ transplant outcomes.

\*Correspondence:

Jalal Siddiqui  
jalal.siddiqui@viracor.eurofinsus.com

Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

**Keywords** Human leukocyte antigen typing, Organ transplantation, Oxford nanopore technology, Long read sequencing

## Background

The success of organ transplant procedures is marred by the immune system driven rejections of allografts [1]. The UNOS data also reports a stagnancy in the short-term (within 1 year) survival rate of transplanted organs since last decade [2]. One of ways to improve the transplant survival is to better understand the immunogenicity of the transplanted organ, and the third-field resolution human leukocyte antigen (HLA) matching is certainly a way forward. The current paradigm of serotyping HLA region of the donor and the recipient, fails to truly quantify the match, leading to suboptimal organ assignments and poor allograft survival [3–7]. For a third-field resolution (SNP level) HLA matching, applications of rapid DNA sequencing technologies are reported by several groups [8–13]. Most of these works argue for the use of sequencing platforms which can not only help an SNP level data analysis but also deliver such data within a short period (3–4 h).

Short read sequencing offers a high-throughput option yet has many limitations, with the biggest being a longer sequencing time (up to 2 days) [9, 14]. Other difficulties include ambiguous alignments and gaps in read coverage. Phasing or the separation of alleles into haplotypes is also difficult. Long read sequencing technologies such as Oxford Nanopore technology (ONT) allow for the coverage of larger regions with faster turnaround time (~3 h) [10, 11, 15–18]. One caveat is that the data can be rather noisy and there is a need for bioinformatics algorithms to denoise such data to perform HLA typing.

While bioinformatics algorithms for HLA typing do exist, they have limitations such as mostly being concordant for class I alleles or requiring manual editing to achieve accurate results [10]. One of the first pipelines has been the Athlon pipeline that focused on the generation of consensus sequences of alleles to denoise long read data [10, 11, 19]. For Athlon2, previous studies have achieved concordances of over 90% (94–100%) at the 2-field resolution, however this was only for key exons on class I genes [10–12, 19]. Recently, a new version, Athlon2, has been made publicly available and has been expanded to 3rd field resolution of both class I and class II genes [20]. The Nanotyper pipeline is another approach that uses read clustering and hierarchical scoring (prioritizing key exons) to call HLA alleles. Nanotyper has achieved 100% and 92% 4th field resolution on HLA-A and HLA-B, respectively, however, it will need improvements for class II genes [10, 21, 22]. HLA-LA, an additional promising approach, utilizes the graph-based alignment approach. This graph alignment approach

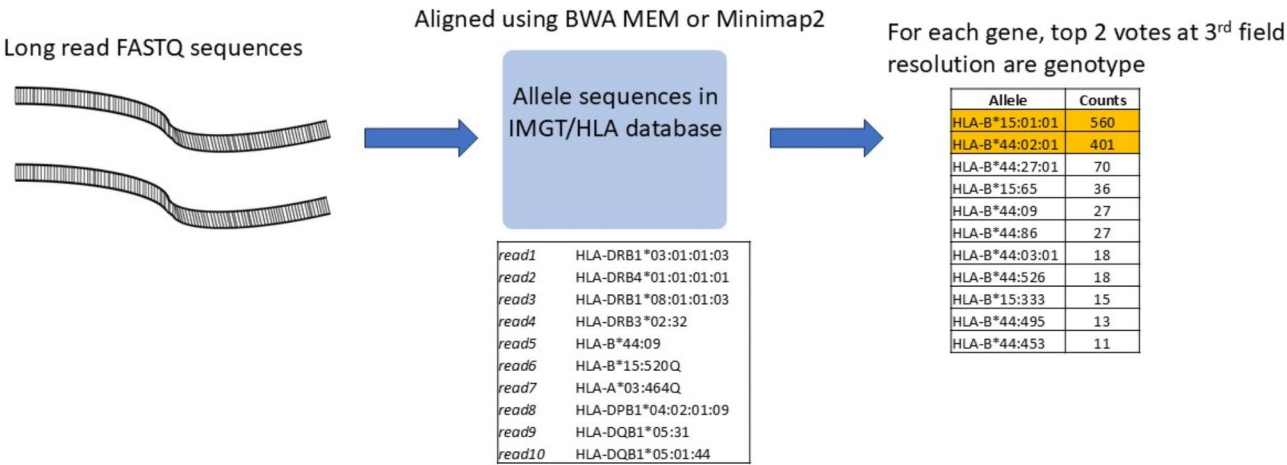
aims to identify linear alignments between reads and a population reference graph by combining known reference gene sequences into a model for genetic variation. This approach has achieved accuracies of 95–100% for long read data, however additional validation data is needed [10, 23]. While these publicly available tools do achieve considerable performance, many of the Github repositories have not been maintained for years and there can be a great deal of difficulty in installing and setting up such software especially for non-computational users. Commercial tools such as NGSengine® are widely used, however like other methods there have been concerns with homopolymers and results generally need manual inspection [10, 24–26].

We have developed a rather simple alignment and voting-based algorithm that can successfully de-noise and perform HLA typing at high concordance with NGSengine® allele results for most genes. Unlike other approaches, we focus on alignment of long reads to the IPD-IMGT/HLA database and then aim to identify which two 3rd field allele groups have the highest number of hits (Fig. 1) [27]. The aim is not to re-construct a consensus allele sequence but to perform readily accessible third-field resolution HLA typing using publicly available long read aligners [28–31]. While this approach is rather straightforward compared to other approaches, the aim is to develop a readily accessible approach with minimal complexity that can provide third-field resolution HLA typing with reduced throughput turnaround time from both laboratory and computational processing. From our initial results, we believe that this can lay groundwork for better using long read technology for HLA typing. (Fig. 1)

## Methods

### Amplicon and sequencing chemistry

Two different batches were used for our bioinformatics study. The first batch or development batch, used for pipeline development, contained 31 samples. DNA was amplified using NGSgo® MX11-3 PCR mix (GenDx) which covers 11 HLA genes (3 class I and 8 class II) (Additional File 1: Table S1). Nanopore sequencing was performed on the GridION instrument using ONT SQK-LSK109 reagents and R10.3 flow cells [13, 27]. All HLA sequence data had been initially analyzed using NGSengine® software from GenDx (2.24.0) followed by manual inspection. The NGSengine® analysis utilized the IPD-IMGT/HLA database (3.46.0) with HLA typing performed to the 3rd field of resolution. The second batch or test batch, used for pipeline testing, contained



**Fig. 1** Overview of voting-based algorithm for HLA typing. Long FASTQ sequences aligned using BWA MEM or Minimap2 aligner. Top 2 votes at 3rd field are assigned as the genotype

63 samples. Like the first batch, DNA was amplified using NGSgo® MX11-3 PCR mix (GenDx). Nanopore sequencing was performed on the GridION instrument using ONT SQK-LSK109 reagents and R10.3 flow cells [13, 27]. Like the first batch, all HLA sequence data was analyzed using NGSengine® software (followed by manual inspection) from GenDx (2.24.0) and the IPD-IMGT/HLA database (3.46.0) with HLA typing was performed to the 3rd field of resolution. In addition, for each batch, a fastq file was made by merging up to 10,000 reads from each sample and NanoStats was used to perform a quality control analysis [32].

Extracting IPD-IMGT/HLA data

Genomic sequences from the IPD-IMGT/HLA database (3.53.0) corresponding to the 11 HLA genes (HLA-A, HLA-B, HLA-C, HLA-DRB1, HLA-DRB3, HLA-DRB4, HLA-DRB5, HLA-DPA1, HLA-DPB1, HLA-DQA1, and HLA-DQB1) were extracted via custom R scripts [27]. Furthermore, using R scripts, an allele group classification table was prepared containing for each allele, the allele group at different fields of resolution, and the allele sequence. For example, the allele HLA-A\*01:01:01:01 would belong to the 1st field group of HLA-A\*01, the 2nd field group of HLA-A\*01:01, the 3rd field group of HLA-A\*01:01:01, and the 4th field group of HLA-A\*01:01:01:01. In cases of sequences that only had a 2nd field name, the 2nd field group was the 3rd and 4th field group. From this table, a FASTA file is output and is used to develop an index for the BWA aligner and for the Minimap2 aligner.

Duplicates analysis

As amplicons from NGSgo® MX11-3 PCR mix (GenDx) did not cover all regions (Additional File 1: Table S1), we wanted to ensure that doing so would not result in a

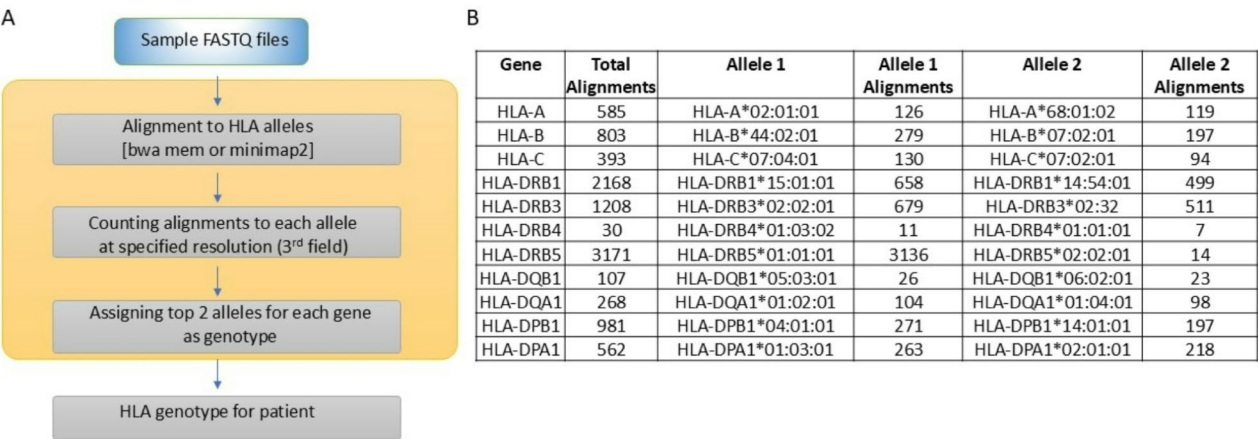
loss of information. This loss of information may arise as amplicon regions can be identical in the same region of many alleles (referred to here on as Duplicates) (Additional File 2: Figure S1). Custom scripts in R were developed to subset sequences from each gene based on regions and/or coordinates and to “duplicates”. For each gene, we created a table of the start and end regions that would be extracted. The information from this table was used via our custom scripts to subset sequences and check for duplicates. Genes that contained 2 amplicon regions had allele sequences merged by an “N” in between and these merged sequences were checked for duplicates to ensure that no duplicates were present at the 3rd field.

Alignment and voting-based algorithm

Up to 10,000 reads from the FASTQ files of ONT reads were aligned to the IPD-IMGT/HLA sequences at the 4th field using BWA-MEM aligner with “-x ont2d” mode or Minimap2 using “-ax map-ont -z 600,200” mode. Using custom R scripts, we imported the resulting SAM files into R and merged the 4th field alignments with allele group classification tables. Using the R scripts, we counted all alignments to each allele at the 3rd field resolution. For each gene, the alleles at 3rd field resolution with the top 2 counts were assigned as the genotype. (Figures 1 and 2)

Allele concordances

HLA-typing results from our previous analysis with NGSengine® software (manual inspection) were imported into R and merged with allele group tables to obtain resolution at 1st, 2nd, and 3rd fields. A match was found if a previous NGSengine® allele call was also identified by the described voting-based algorithm genotypes. In cases of alleles reported on NGSengine® to be homozygous or



**Fig. 2** Voting-based algorithm pipeline workflow and example results. Pipeline workflow for FASTQ file (A) and example results for sample (B)

for single allele calls, we only considered the allele with the top-ranked counts when determining concordances. In the case of HLA-DRB3/4/5, we only considered the unique allele calls from NGSengine®. For each gene, the percentages of matches were determined for each of the datasets.

HLA-LA and Athlon2 pipeline

The performance of our pipeline was tested against publicly available pipelines, such as HLA-LA [23] and Athlon2 [20] pipelines. The HLA-LA analysis was done as follows. Using BWA-MEM aligner with “ont2d” mode, FASTQ reads were aligned to the hg38 genome FASTA file from 1000 Genomes [33]. HLA-LA.pl was run on the resulting BAM files using the “PRG\_MHC\_GRCh38\_withIMGT” graph database. HLA genotypes were determined from the “bestguess\_G.txt” files. For the Athlon2 web-server, de-identified reads were uploaded, and allele assignment reports were downloaded. Allele concordances for both HLA-LA and Athlon2 pipelines were determined using the same method used for our voting-based algorithm.

Comparing concordances between pipelines

McNemar’s test was performed in the R environment to compare concordances to NGSengine® results between analyses from different pipelines [34]. A contingency table was prepared containing alleles concordant for pipeline 1, alleles not concordant for pipeline 1, alleles concordant for pipeline 2, and alleles not concordant for pipeline 2. McNemar’s test was performed on the contingency table and  $p < 0.05$  was deemed to be significant. McNemar’s test was performed at the overall alleles level and at the gene level.

Allele count ratios and sums

We sorted the BWA-MEM results from genes with unique NGSengine® allele calls of 0 to 2 calls and

determined the ratios and sums of H1 = allele with highest allele counts and H2 = alleles with second highest allele counts as well as their sums. The purpose of this is to investigate potential cutoffs for homozygous vs. heterozygous as well as whether HLA-DRB3/4/5 alleles are present or not. Unique allele calls of 0 correspond to missing HLA-DRB3/4/5 alleles, unique allele calls of 1 correspond to homozygous or only single allele present for a gene, and unique allele calls of 2 correspond to heterozygous [35, 36].

Results

Alignment and voting pipeline

Using BWA-MEM aligner and Minimapp2 aligner, alongside custom R scripts, we were able to generate a workflow for determining genotype from Nanopore reads (Fig. 2A). Each sample file would undergo alignment via ONT-based parameters using BWA-MEM or Minimapp2. Using R scripts, the resulting SAM files would be imported into R as a data frame object and reference sequences or allele IDs at 4th resolution were merged with allele group classification tables. The allele counts at 3rd resolution were summed and the top 2 counts for each gene were assigned as the genotype. This was used to determine HLA genotypes for each patient (Fig. 2B).

Duplicate analysis

Genes with full-length amplicon coverage (HLA-A, HLA-B, HLA-C, HLA-DQB1, HLA-DPA1, HLA-DQA1) had a total of 4616, 5422, 5113, 780, 318, and 402 alleles respectively and all sequences were unique to each allele with no duplicate sequences at 4th field or 3rd field (Additional File 1: Table S2).

For genes split by amplicons, duplicates were identified at 4th field and at 3rd field, however merging the amplicon-split sequences resulted in no duplicates at 3rd field (Additional File 1: Table S3). For HLA-DRB1, there were a total of 602 alleles. The HLA-DRB1 amplicon from

start to 2.5 kb contained 74 duplicates at 4th field and 60 duplicates at 3rd field. The HLA-DRB1 amplicon from exon 2 to the end contained 27 duplicates at 4th field and 1 duplicate at 3rd field. Merging the 2 HLA-DRB1 amplicon sequences resulted in 21 duplicates at 4th field and 0 duplicates at 3rd field. For DPB1, there were a total of 940 alleles. For HLA-DPB1 amplicon from start to exon 1, 11 duplicates were found at 4th field and 10 duplicates were found at 3rd field. For HLA-DBP1 amplicon from exon 2 to end, 57 duplicates were found at 4th field and 3 duplicates at 3rd field. Merging the amplicon sequences resulted in 55 duplicates at 4th field and 0 duplicates at 3rd field. For HLA-DRB3, there were 36 alleles. For HLA-DRB3 amplicon from start to 2.5 kb, there were 7 duplicates identified at 4th field and 3 duplicates identified at 3rd field. For HLA-DRB3 amplicon from exon 2 to end of gene, 2 duplicates were identified at 4th field and 0 at 3rd field. Merging the amplicon sequences resulted in 2 duplicates at 4th field and 0 at 3rd field. For HLA-DRB5, there were a total of 11 sequences. For HLA-DRB5 amplicon from start to 2.6 kb, 3 duplicates were found at 4th field and 2 at 3rd field. For HLA-DRB5 amplicon from exon 2 to end, 0 duplicates were found at 3rd field and 4th field. Merging the amplicon sequences resulted in 0 duplicates being found for the 3rd field and 4th field. HLA-DRB4 contained 28 sequences and only one amplicon region from exon 2 to exon 3. 3 duplicates were found at 4th field and 0 at 3rd field. (Additional File 1: Table S3)

The duplicates analyses demonstrated that the amplicon coverage from the NGSgo® MX11-3 PCR mix (GenDx) was sufficient to obtain up to 3rd field resolution HLA typing for the characterized 11 genes.

### NanoStats results

Using NanoStats, we were able to obtain quality control information for each batch (Additional File 3: Data S1)

[32]. For the development batch, there was a mean read quality of 13.2 and a median read quality of 14.4. 100% of reads were above Q7 cutoff, 98.5% of the reads were above the Q10 cutoff, 86.1% of reads were above the Q12 cutoff, and 34.8% of reads were above the Q15 cutoff. For the test batch, there was a mean read quality of 13 and median read quality of 14.1. 100% of reads were above Q7 cutoff, 97.8% of the reads were above the Q10 cutoff, 81.2% of reads were above the Q12 cutoff, and 32.8% of reads were above the Q15 cutoff.

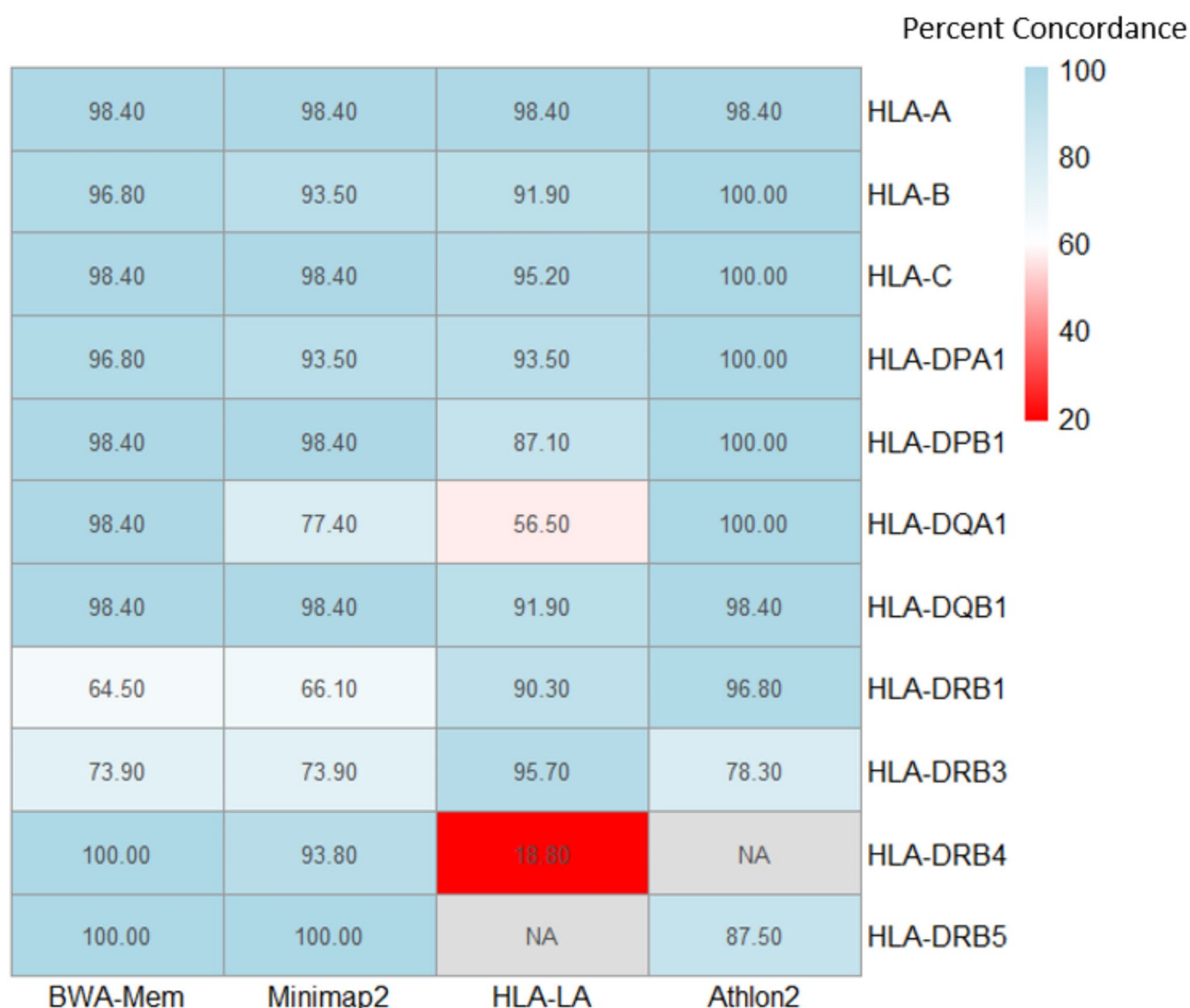
### Allele concordance resulting using development batch

Using our development batch of 31 samples, our alignment and voting-based pipeline using BWA-MEM aligner was able to obtain an overall high allele concordance with our previous results. For 1st field resolution, we obtained an overall concordance of 97.2%, for 2nd field resolution an overall concordance of 93.4%, and for 3rd field resolution, an overall concordance of 93.2%. The following percent concordances were obtained for the 1st, 2nd and 3rd field for each gene: HLA-A (98.4%, 98.4%, 98.4%), HLA- B (100%, 96.8%, 96.8%), HLA-C (100%, 98.4%, 98.4%), HLA-DPA1 (100%, 96.8%, 96.8%), HLA-DPB1 (100%, 100%, 98.4%), HLA- DQA1 (100%, 98.4%, 98.4%), HLA-DQB1 (100%, 98.4%, 98.4%), HLA-DRB1 (83.9%, 64.5%, 64.5%), HLA-DRB3 (82.6%, 73.9%, 73.9%), HLA-DRB4 (100%, 100%, 100%) and HLA-DRB5 (100%, 100%, 100%).(Additional File 3: Data S2, Table 1; Fig. 3).

Using Minimap2 aligner, our pipeline obtained an overall allele concordance of 96.1% at the 1st field, 90.8% at the 2nd field, and 90.1% at the 3rd field (Additional File 3: Data S2, Table 1; Fig. 3). The following percent concordances were obtained for the 1st, 2nd and 3rd field for each gene: HLA-A (100%, 98.4%, 98.4%), HLA- B (98.4%, 93.5%, 93.5%), HLA- C (100%, 98.4%, 98.4%), HLA-DPA1 (100%, 98.4%, 93.5%), HLA-DPB1 (98.4%, 98.4%, 98.4%),

**Table 1** Percent concordance by gene for 31 development batch samples aligned via BWA-MEM aligner and Minimap2 aligner

| Gene     | BWA Aligner              |                          |                          | Minimap2 Aligner         |                          |                          |
|----------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
|          | 1st field<br>(% success) | 2nd field<br>(% success) | 3rd field<br>(% success) | 1st field<br>(% success) | 2nd field<br>(% success) | 3rd field<br>(% success) |
| HLA-A    | 98.4%                    | 98.4%                    | 98.4%                    | 100.0%                   | 98.4%                    | 98.4%                    |
| HLA-B    | 100.0%                   | 96.8%                    | 96.8%                    | 98.4%                    | 93.5%                    | 93.5%                    |
| HLA-C    | 100.0%                   | 98.4%                    | 98.4%                    | 100.0%                   | 98.4%                    | 98.4%                    |
| HLA-DPA1 | 100.0%                   | 96.8%                    | 96.8%                    | 100.0%                   | 98.4%                    | 93.5%                    |
| HLA-DPB1 | 100.0%                   | 100.0%                   | 98.4%                    | 98.4%                    | 98.4%                    | 98.4%                    |
| HLA-DQA1 | 100.0%                   | 98.4%                    | 98.4%                    | 91.9%                    | 77.4%                    | 77.4%                    |
| HLA-DQB1 | 100.0%                   | 98.4%                    | 98.4%                    | 100.0%                   | 98.4%                    | 98.4%                    |
| HLA-DRB1 | 83.9%                    | 64.5%                    | 64.5%                    | 83.9%                    | 66.1%                    | 66.1%                    |
| HLA-DRB3 | 82.6%                    | 73.9%                    | 73.9%                    | 82.6%                    | 73.9%                    | 73.9%                    |
| HLA-DRB4 | 100.0%                   | 100.0%                   | 100.0%                   | 100.0%                   | 100.0%                   | 93.8%                    |
| HLA-DRB5 | 100.0%                   | 100.0%                   | 100.0%                   | 100.0%                   | 100.0%                   | 100.0%                   |
| Overall  | 97.2%                    | 93.4%                    | 93.2%                    | 96.1%                    | 90.8%                    | 90.1%                    |



**Fig. 3** Percent concordance for third field resolution using different methods. Heatmap showing third field resolution concordances with BWA Mem, Minimap2, HLA-LA, and Athlon2

HLA-DQA1 (91.9%, 77.4%, 77.4%), HLA-DQB1 (100%, 98.4%, 98.4%), HLA-DRB1 (83.9%, 66.1%, 66.1%), HLA-DRB3 (82.6%, 73.9%, 73.9%), HLA-DRB4 (100%, 100%, 93.8%) and HLA-DRB5 (100%, 100%, 100%).

In comparing the 2 pipelines, BWA-MEM had a higher overall concordance at all fields of resolution (1st field: 97.2% BWA-MEM vs. 96.1% Minimap2, 2nd field: 93.4% BWA-MEM vs. 90.8% Minimap2, 3rd field: 93.2% vs. 90.1% Minimap2). For non- HLA-DRB genes, Minimap2 had overall > 93% concordance at the 3rd field except for HLA-DQA1, however BWA-MEM performed much better at > 96% concordance at the 3rd field for non- HLA-DRB genes. HLA-DRB1 had a slightly better performance on Minimap2 of 66.1% at 3rd field compared to BWA-MEM at 64.5%. HLA-DRB3 performance was similar on both at 73.9% at 3rd field (Table 1). McNemar's test

was performed on 3rd field results from BWA-MEM and Minimap2 to ascertain whether the concordance differences were statistically significant. Overall, there was a higher statistically significant concordance for BWA-MEM (93.2%) compared to Minimap2 (90.1%) at  $p = 1.37E-03$ . HLA-DQA1 had a statistically significant higher concordance of 98.4% in BWA-MEM compared to Minimap2 of 77.4% ( $p = 8.74E-04$ ) (Table 2).

#### Allele concordance resulting using test batch

Using our test batch of 63 samples, our alignment and voting-based pipeline using BWA-MEM aligner was able to obtain an overall high allele concordance with our previous results. For 1st field resolution, we obtained an overall concordance of 97.5%, a 2nd field overall concordance of 92.1%, and a 3rd field overall concordance

**Table 2** Comparison of concordance between BWA-MEM and Minimap2 for development batch

| Gene     | P-Value  | BWA-MEM Concordance | Minimap2 Concordance | Difference                  |
|----------|----------|---------------------|----------------------|-----------------------------|
| HLA-A    | 1.00E+00 | 98.4%               | 98.4%                | ns                          |
| HLA-B    | 4.80E-01 | 96.8%               | 93.5%                | ns                          |
| HLA-C    | NA       | 98.4%               | 98.4%                | ns                          |
| HLA-DPA1 | 6.17E-01 | 96.8%               | 93.5%                | ns                          |
| HLA-DPB1 | NA       | 98.4%               | 98.4%                | ns                          |
| HLA-DQA1 | 8.74E-04 | 98.4%               | 77.4%                | BWA-MEM higher concordance* |
| HLA-DQB1 | NA       | 98.4%               | 98.4%                | ns                          |
| HLA-DRB1 | 1.00E+00 | 64.5%               | 66.1%                | ns                          |
| HLA-DRB3 | 1.00E+00 | 73.9%               | 73.9%                | ns                          |
| HLA-DRB4 | 1.00E+00 | 100.0%              | 93.8%                | ns                          |
| HLA-DRB5 | NA       | 100.0%              | 100.0%               | ns                          |
| Overall  | 1.37E-03 | 93.2%               | 90.1%                | BWA-MEM higher concordance* |

of 89.8%. The following percent concordances were obtained for the 1st, 2nd and 3rd field for each gene: HLA-A (100%, 96.8%, 88.1%), HLA-B (100%, 90.54%, 88.1%), HLA-C (100%, 99.2%, 99.2%), HLA-DPA1 (100%, 98.4%, 97.6%), HLA-DPB1 (98.4%, 97.6%, 92.9%), HLA-DQA1 (100%, 100%, 98.4%), HLA-DQB1 (100%, 97.6%, 96.0%), HLA-DRB1 (88.9%, 68.3%, 68.3%), HLA-DRB3 (81.0%, 61.9%, 61.9%), HLA-DRB4 (100%, 97.4%, 94.7%) and HLA-DRB5 (73.3%, 66.7%, 66.7%). (Additional File 3: Data S2, Additional File 1: Table S4, (Additional File 2: Figure S2)

Using Minimap2 aligner, we obtained an overall concordance of 96.0% for 1st field, 88.5% for 2nd field, and 86.4% for 3rd field. %. The following percent concordances were obtained for the 1st, 2nd and 3rd field for each gene: HLA-A (99.2%, 86.5%, 84.9%), HLA-B (98.4%, 86.5%, 85.7%), HLA-C (99.2%, 95.2%, 93.7%), HLA-DPA1 (100%, 98.4%, 97.6%), HLA-DPB1 (97.6%, 97.6%, 95.2%), HLA-DQA1 (94.4%, 82.5%, 77.8%), HLA-DQB1 (99.2%, 99.2%, 96.8%), HLA-DRB1 (88.9%, 69.0%, 69.0%), HLA-DRB3 (73.8%, 66.7%, 66.7%), HLA-DRB4 (100%, 97.4%, 84.2%) and HLA-DRB5 (73.3%, 66.7%, 66.7%). (Additional File 1: Table S4, (Additional File 2: Figure S2)

To compare BWA-MEM and Minimap2 concordances, we performed McNemar's test between the BWA-MEM and Minimap2 concordance results at the 3rd field. Overall, there was a statistically significantly higher overall concordance in BWA-MEM (89.8% overall concordance) compared to Minimap2 (86.4% overall concordance) at  $p = 2.51E-05$ . From the results, HLA-C (BWA-MEM 99.2%, Minimap2 93.7%,  $p = 2.33E-02$ ) and HLA-DQA1 (BWA-MEM 98.4%, Minimap2 77.8%,  $p = 2.31E-06$ ) had significantly higher concordances in BWA-MEM. (Additional File 1: Table S5) Due to BWA-MEM pipeline having an overall higher statistically significant concordance compared to Minimap2 across both datasets as well as 2 genes with significantly higher concordance, we decided to focus on the BWA-MEM aligner-based pipeline when making comparisons to other pipelines.

### Comparing performance of our pipeline with other publicly available tools

Using the HLA-LA pipeline, HLA genotypes were obtained for all genes except HLA-DRB5. With Athlon2, HLA genotypes were obtained for all genes except HLA-DRB4. For the development batch, the HLA-LA pipeline resulted in overall concordances of 99.1% (1st field), 89.5% (2nd field), and 86.4% (3rd field). The following percent concordances were obtained for the 1st, 2nd and 3rd field for each gene: HLA-A (100%, 100.0%, 98.4%), HLA-B (100%, 98.44%, 91.1%), HLA-C (100%, 98.4%, 95.2%), HLA-DPA1 (100%, 93.5%, 93.5%), HLA-DPB1 (91.9%, 91.9%, 87.1%), HLA-DQA1 (100%, 67.7%, 56.5%), HLA-DQB1 (100%, 91.9%, 91.9%), HLA-DRB1 (100.0%, 90.3%, 90.3%), HLA-DRB3 (100.0%, 95.7%, 95.7%), and HLA-DRB4 (100%, 18.8%, 18.8%). For non-HLA-DRB genes, the HLA-LA concordances were >80% at 3rd field except for HLA-DQA1 at 56.5%. HLA-DRB1 and HLA-DRB3 concordance were above 90%, however HLA-DRB4 had low concordance at 3rd field of 18.8%. (Table 3, Additional File 3: Data S2, Fig. 3)

Athlon2 concordances were at 98.5% (1st field), 98.3% (2nd field), and 98.1% (3rd field). The following percent concordances were obtained for the 1st, 2nd and 3rd field for each gene: HLA-A (100%, 100.0%, 98.4%), HLA-B (100%, 100.0%, 100.0%), HLA-C (100%, 100.0%, 100.0%), HLA-DPA1 (100%, 100.0%, 100.0%), HLA-DPB1 (100.0%, 100.0%, 100.0%), HLA-DQA1 (100.0%, 100.0%, 100.0%), HLA-DQB1 (100%, 98.4%, 98.4%), HLA-DRB1 (96.8%, 96.8%, 96.8%), HLA-DRB3 (78.3%, 78.3%, 78.3%), and HLA-DRB5 (87.5%, 87.5%, 87.5%). For non-HLA-DRB genes, Athlon2 had >98% concordances for all non-HLA-DRB genes with a 96.8% concordance for HLA-DRB1 and a 78.3% for HLA-DRB3. (Table 3, Additional File 3: Data S2, Fig. 3)

Using McNemar's test, we compared the concordances of HLA-LA and Athlon2 that of BWA aligner voting based pipeline. In comparing BWA-MEM results to the HLA-LA pipeline, there was a higher statistically

**Table 3** Percent concordance by gene for 31 development batch samples aligned via HLA-LA pipeline and Athlon2 pipeline

| Gene     | HLA-LA pipeline          |                          |                          | Athlon2 Pipeline         |                          |                          |
|----------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
|          | 1st field<br>(% success) | 2nd field<br>(% success) | 3rd field<br>(% success) | 1st field<br>(% success) | 2nd field<br>(% success) | 3rd field<br>(% success) |
| HLA-A    | 100.0%                   | 100.0%                   | 98.4%                    | 100.0%                   | 100.0%                   | 98.4%                    |
| HLA-B    | 100.0%                   | 98.4%                    | 91.9%                    | 100.0%                   | 100.0%                   | 100.0%                   |
| HLA-C    | 100.0%                   | 98.4%                    | 95.2%                    | 100.0%                   | 100.0%                   | 100.0%                   |
| HLA-DPA1 | 100.0%                   | 93.5%                    | 93.5%                    | 100.0%                   | 100.0%                   | 100.0%                   |
| HLA-DPB1 | 91.9%                    | 91.9%                    | 87.1%                    | 100.0%                   | 100.0%                   | 100.0%                   |
| HLA-DQA1 | 100.0%                   | 67.7%                    | 56.5%                    | 100.0%                   | 100.0%                   | 100.0%                   |
| HLA-DQB1 | 100.0%                   | 91.9%                    | 91.9%                    | 100.0%                   | 98.4%                    | 98.4%                    |
| HLA-DRB1 | 100.0%                   | 90.3%                    | 90.3%                    | 96.8%                    | 96.8%                    | 96.8%                    |
| HLA-DRB3 | 100.0%                   | 95.7%                    | 95.7%                    | 78.3%                    | 78.3%                    | 78.3%                    |
| HLA-DRB4 | 100.0%                   | 18.8%                    | 18.8%                    | 0.0%                     | 0.0%                     | 0.0%                     |
| HLA-DRB5 | N/A                      | N/A                      | N/A                      | 87.5%                    | 87.5%                    | 87.5%                    |
| Overall  | 99.1%                    | 89.5%                    | 86.4%                    | 98.5%                    | 98.3%                    | 98.1%                    |

**Table 4** Comparison of concordance between BWA-MEM and HLA-LA for development batch

| Gene     | P-Value  | BWA-MEM Concordance | HLA-LA Concordance | Difference                  |
|----------|----------|---------------------|--------------------|-----------------------------|
| HLA-A    | 1.00E+00 | 98.4%               | 98.4%              | ns                          |
| HLA-B    | 3.71E-01 | 96.8%               | 91.9%              | ns                          |
| HLA-C    | 6.17E-01 | 98.4%               | 95.2%              | ns                          |
| HLA-DPA1 | 6.17E-01 | 96.8%               | 93.5%              | ns                          |
| HLA-DPB1 | 4.55E-02 | 98.4%               | 87.1%              | BWA-MEM higher concordance* |
| HLA-DQA1 | 9.44E-07 | 98.4%               | 56.5%              | BWA-MEM higher concordance* |
| HLA-DQB1 | 1.34E-01 | 98.4%               | 91.9%              | ns                          |
| HLA-DRB1 | 2.20E-03 | 64.5%               | 90.3%              | HLA-LA higher concordance*  |
| HLA-DRB3 | 7.36E-02 | 73.9%               | 95.7%              | ns                          |
| HLA-DRB4 | 8.74E-04 | 100.0%              | 18.8%              | BWA-MEM higher concordance* |
| Overall  | 3.54E-04 | 93.1%               | 86.4%              | BWA-MEM higher concordance* |

**Table 5** Comparison of concordance between BWA-MEM and Athlon2 for development batch

| Gene     | P-Value  | BWA-MEM Concordance | Athlon2 Concordance | Difference                  |
|----------|----------|---------------------|---------------------|-----------------------------|
| HLA-A    | 1.00E+00 | 98.4%               | 98.4%               | ns                          |
| HLA-B    | 4.80E-01 | 96.8%               | 100.0%              | ns                          |
| HLA-C    | 1.00E+00 | 98.4%               | 100.0%              | ns                          |
| HLA-DPA1 | 4.80E-01 | 96.8%               | 100.0%              | ns                          |
| HLA-DPB1 | 1.00E+00 | 98.4%               | 100.0%              | ns                          |
| HLA-DQA1 | 1.00E+00 | 98.4%               | 100.0%              | ns                          |
| HLA-DQB1 | NA       | 98.4%               | 98.4%               | ns                          |
| HLA-DRB1 | 2.15E-05 | 64.5%               | 96.8%               | Athlon2 higher concordance* |
| HLA-DRB3 | 1.00E+00 | 73.9%               | 78.3%               | ns                          |
| HLA-DRB5 | 1.00E+00 | 100.0%              | 87.5%               | ns                          |
| Overall  | 3.14E-05 | 93.0%               | 98.1%               | Athlon2 higher concordance* |

significant concordance in BWA-MEM (93.1%) compared to HLA-LA (86.4%) at  $p = 3.54E-06$ . HLA-DPB1 (98.4% BWA-MEM, 87.1% HLA-LA,  $p = 4.55E-02$ ), HLA-DQA1 (98.4% BWA-MEM, 56.5% HLA-LA,  $p = 9.44E-07$ ), and HLA-DRB4 (100.0% BWA-MEM, 18.8% HLA-LA,  $p = 8.74E-04$ ) were found to have statistically significant higher concordances in BWA-MEM aligner. HLA-DRB1 (64.5% BWA-MEM, 90.3% HLA-LA,  $p = 2.20E-03$ ) was found to have a statistically significant concordance in HLA-LA aligner. HLA-DRB3 had a higher concordance

in HLA-LA (95.7%) compared to BWA-MEM (73.9%), however that difference was not significant ( $p = 7.36E-02$ ) (Table 4). In comparing BWA-MEM's results to that of the Athlon2 pipeline, there was a higher statistically significant concordance in Athlon2 (98.1%) compared to BWA-MEM (93.0%) at  $p = 3.14E-05$ . The only statistically significant difference at the gene level was HLA-DRB1 concordance being higher in Athlon2 (64.5% BWA-MEM, 96.8%,  $p = 2.15E-05$ ). (Table 5)

For the test batch (second batch), the HLA-LA pipeline resulted in overall concordances of 98.3% (1st field), 87.5% (2nd field), and 81.0% (3rd field). The following percent concordances were obtained for the 1st, 2nd and 3rd field for each gene: HLA-A (97.6%, 95.2%, 88.1%), HLA-B (100%, 100.0%, 88.1%), HLA-C (99.2%, 96.8%, 88.9%), HLA-DPA1 (100%, 96.0%, 94.4%), HLA-DPB1 (93.7%, 93.7%, 85.7%), HLA-DQA1 (100%, 63.5%, 47.6%), HLA-DQB1 (97.6%, 82.5%, 81.7%), HLA-DRB1 (98.4%, 87.3%, 85.7%), HLA-DRB3 (97.6%, 97.6%, 92.9%), and HLA-DRB4 (97.4%, 26.3%, 26.3%). For non- HLA-DRB genes, the HLA-LA concordances were > 80% at 3rd field except for HLA-DQA1 at 47.6%. HLA-DRB1 concordance was at 85.7% and HLA-DRB3 concordance was at 92.9%. Athlon2 overall concordances were at 97.0% (1st field), 96.6% (2nd field), and 94.8% (3rd field). The following percent concordances were obtained for the 1st, 2nd and 3rd field for each gene: HLA-A (97.6%, 97.6%, 94.4%), HLA-B (100%, 100.0%, 95.2%), HLA-C (99.2%, 99.2%, 99.2%), HLA-DPA1 (100%, 99.2%, 99.2%), HLA-DPB1 (100.0%, 100.0%, 96.0%), HLA-DQA1 (100.0%, 100.0%, 99.2%), HLA-DQB1 (97.6%, 96.0%, 94.4%), HLA-DRB1 (97.6%, 96.8%, 96.0%), HLA-DRB3 (69.0%, 69.0%, 69.0%), and HLA-DRB5 (40.0%, 40.0%, 40.0%). Athlon2 had > 94% concordances for all non-HLA-DRB genes with a 96.0% concordance for HLA-DRB1 and a 69.0% concordance for HLA-DRB3. (Additional File 1: Table S6, Additional File 3: Data S2, (Additional File 2: Figure S2)

In comparing overall BWA-MEM results to the HLA-LA pipeline, BWA-MEM had a higher statistically significant concordance of 90.1% compared to 81.0% for HLA-LA ( $p=5.29\text{E-}11$ ). HLA-C (99.2% BWA-MEM, 88.9% HLA-LA,  $p=8.74\text{E-}04$ ), HLA-DPB1 (92.9% BWA-MEM, 85.7% HLA-LA,  $p=2.65\text{E-}02$ ), HLA-DQA1 (98.4% BWA-MEM, 47.6% HLA-LA,  $p=3.41\text{E-}15$ ), HLA-DQB1 (BWA-MEM 96.0%, 81.7% HLA-LA,  $p=5.20\text{E-}04$ ), and HLA-DRB4 (BWA-MEM 94.7%, HLA-LA 26.3%,  $p=2.31\text{E-}06$ ) were found to have statistically significant higher concordances in BWA-MEM aligner. HLA-DRB1 (68.3% BWA-MEM, 85.7% HLA-LA,  $p=6.58\text{E-}04$ ) and HLA-DRB3 (61.9% BWA-MEM, 92.9%, HLA-LA,  $p=8.74\text{E-}04$ ) were found to have a statistically significant concordance in HLA-LA aligner. (Additional File 1: Table S7). In comparing overall BWA-MEM results to the HLA-LA pipeline, Athlon2 has a higher statistically significant concordance of 94.8% compared to 89.6% for BWA-MEM ( $p=2.76\text{E-}08$ ). HLA-A (88.1% BWA-MEM, 94.4% Athlon2,  $p=2.69\text{E-}02$ ), HLA-B (88.1% BWA-MEM, 95.2% Athlon2,  $p=7.66\text{E-}03$ ), and HLA-DRB1 (68.3% BWA-MEM, 96.0% Athlon2,  $p=2.28\text{E-}08$ ) were found to have statistically significant higher concordances in Athlon2 aligner. (Additional File 1: Table S8)

### Comparing counts of top 2 alleles

As a preliminary study to incorporate homozygosity and heterozygosity, we studied the ratios of the allele counts of the top-ranked allele (H1) and the allele counts of the second top-ranked allele (H2). Using the development batch, for genes with unique allele calls of 1 (homozygous) in the samples, the H2/H1 ratio was 0.040 to 0.209 for non-HLA-DRB genes. For HLA-DRB3 and HLA-DRB5, the H2/H1 ratio was 0.681 and 0.658 respectively. For samples with unique allele calls of 2 (heterozygous), the H2/H1 ratio was 0.442 to 0.789. In cases of missing HLA-DRB3/4/5 alleles, mean H2 + H1 was 42 and under. For other alleles, the mean H2 + H1 was 203 and above. Using the test batch, for genes with unique allele calls of 1 (homozygous) in the samples, the H2/H1 ratio was 0.080 to 0.180 for non-HLA-DRB genes. For HLA-DRB1, HLA-DRB3 and HLA-DRB5, the H2/H1 ratio was 0.493, 0.576, and 0.757 respectively. For samples with unique allele calls of 2 (heterozygous), the H2/H1 ratio was 0.514 to 0.857. In cases of missing HLA-DRB3/4/5 alleles, mean H2 + H1 was 57 and under. For other alleles, the mean H2 + H1 was 147 and above. (Additional File 3: Data S3)

### Discussion

While ONT sequencing data is still error prone, computational pipelines to denoise the data and perform third-field resolution HLA typing seem to mature rapidly. In our current work we not only developed a new bioinformatics protocol to perform the HLA typing, but we also compared its performance with other publicly available tools. The results for our BWA-MEM voting-based pipeline are promising with > 96% at 3rd field concordance for non-HLA-DRB genes in our development dataset and > 88% at the 3rd field and > 90% for 2nd field in our test dataset. Our results had a greater and statistically significant overall concordance compared to HLA-LA (development batch BWA-MEM voting based method 93.1% vs. HLA-LA 86.4%, test batch BWA-MEM voting based method 90.1% vs. 81.0% HLA-LA) with differences of 6.7% in development batch and 9.1% in test batch. The major exceptions were in the genotype concordances for HLA-DRB1 and HLA-DRB3 with both being stronger in HLA-LA. Athlon2 overall had a greater concordance (development batch BWA-MEM voting based method 93.0% vs. Athlon2 98.1%, test batch BWA voting based method 89.6% vs. Athlon2 94.8%) with differences of 5.1 and 5.2% in development and test batch respectively. HLA-DRB1 concordance was stronger in both development and test datasets in Athlon2. In the test dataset, genes HLA-A and HLA-B were more concordant in Athlon2. In general, there was an overall lower concordance across the different algorithms used for the test dataset compared to the development dataset. From the NanoStats results, it did appear that the development

dataset had slightly better quality and this may account for the lower concordance in the test dataset.

The main area of improvement for our algorithm is the HLA-DRB genes performance. HLA-DRB1 had <70% concordance at 3rd field using the voting-based algorithm. A previous study using full-length amplicons also reported a low HLA-DRB1 concordance (<70%) [37]. It is important to note that HLA-DRB1 is a highly polymorphic gene and can be linked to alleles of other genes [38, 39]. In addition, the HLA-DRB genes themselves have sequence similarity due to evolutionary relationships [39]. HLA-DRB3 had concordances ranging from 61 to 73% across the aligners. HLA-DRB4 had >90% 3rd field resolution concordances across the datasets and aligners. HLA-DRB5 had >90% concordances for the development batch but <70% for the test batch. Especially for HLA-DRB1 and HLA-DRB3, there will be a need to study adaptations to the algorithm that can improve such performance.

While other pipelines did perform better on HLA-DRB1 and in the case of HLA-LA on HLA-DRB3, other HLA-DRB concordances were weak. The HLA-LA pipeline has stronger concordances for both HLA-DRB1 (>85%) and HLA-DRB3 (>90%). Athlon2 had stronger concordance on HLA-DRB1 (~96%) however <80% performance for HLA-DRB3 that was not significantly different from the BWA-MEM method. HLA-DRB4 had low concordances <30% at 3rd field on HLA-LA and HLA-DRB5 concordance using Athlon2 on the test dataset (40%) was weaker than the BWA-MEM method (66.7%) though not significantly different. One potential explanation for the lower performance of HLA-DRB genes may result from amplicon designs that split the HLA-DRB genes into 2 regions. Many of these amplicon regions can be identical across alleles, making it difficult to distinguish them. Of note, the higher BWA-MEM HLA-DRB4 concordance (>90%) may be in part due to the use of a single amplicon. One future modification can be considering HLA-DRB linkages however this may exclude rare associations [35, 36].

In addition to weaker HLA-DRB performance, there are other limitations. While other algorithms do have alignment and mapping steps, much of their aim is reconstructing the allele sequences using de novo assembly [10, 11, 19, 20]. Others incorporate reads clustering and graph-based alignment algorithms [21–23, 29]. In any case, highly conclusive results may still be difficult due to noisy Nanopore data even if these additional steps are incorporated. Our aim is quantifying and sorting the top alleles by their alignment hits. Unfortunately, ambiguities may arise especially if reads can align to multiple alleles due to amplicon design especially with limited and/or split coverage. Additional interpretation of alignment results may be needed to finalize alignments especially

in cases of ambiguity. Currently, we are focused on identifying whether the NGS Engine® results contain a match with our top 2 alleles. We will need to further investigate cutoffs for homozygous and heterozygous alleles. For non-HLA-DRB genes, the ratio of H2/H1 was >0.4 for heterozygous alleles and <0.2 for homozygous alleles. This however did not hold for HLA-DRB genes and that remains inconclusive. We did also notice that missing HLA-DRB3/4/5 alleles had low counts of 42 or below compared to other alleles with counts above 140 (up to 10,000 reads). Another limitation is that the current algorithm does not take into account the possibility certain alleles are being preferentially amplified over others. By taking this into account in future versions, we can better determine cutoffs for assigning alleles. Another current limitation is that the algorithm relies solely on alignment to sequences within the IPD-IMGT/HLA database. One limitation of this is that there are partial or incomplete sequences. In addition, at a given locus, there can be variations in allele length. While aligners tuned for long reads can capture distinctive nucleotide changes, this may still result in biases or error especially with noisy data. In the future, we will modify our algorithm to account for this.

In the future, additional modifications to our initial algorithm can further improve performance, especially for HLA-DRB genes. These include filtering sequences by quality, mismatches, and weighting the score of different alignments (e.g. multiple mapping reads). Presently, we have focused on BWA-MEM and Minimap2 aligners. While BWA-MEM aligner gave an overall better performance, the overall difference was less than 5%. Future studies can focus on studying additional long read aligners. In addition, integrating other pipelines such as HLA-LA into our algorithm may also help improve performance. As an example, HLA-LA had overall 1st field performance of >98% and had superior performance for HLA-DRB1 and HLA-DRB3. HLA-LA can be used to determine 1st field allele calls and the voting-based method can hone in on the 3rd field. Of course, incorporating other pipelines should be done in a careful manner to not increase the computational complexity and processing time beyond what is needed.

Despite the limitations, this algorithm shows promising results with concordances above HLA-LA and approaching Athlon2. A major strength is that it promotes the use of widely used and publicly available aligners with configurations for long-read sequences. This algorithm can be extended to other long read aligners. It also may serve as an important tool to assist in amplicon design (via simulations). Of note, due to its simplicity of using publicly available aligners, this method is relatively accessible to the wider bioinformatics community and can provide reduced computational processing time, something that will be crucial in deceased donor typing. Additionally, in

conjunction with other methods, it may serve as a valuable tool to better utilize Nanopore sequencing technology for assisting transplant patients.

#### Abbreviations

HLA Human Leukocyte Antigen

ONT Oxford Nanopore Technology

H1 Allele with highest counts for a gene in a sample

H2 Allele with second highest counts for a gene in a sample

#### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-025-11547-4>.

Supplementary Material 1

Supplementary Material 2

Supplementary Material 3

#### Acknowledgements

The authors would like to thank the team members at Eurofins Viracor Clinical Diagnostics for their help and support throughout the study process.

#### Author contributions

J.S. wrote the main draft of the manuscript. All authors edited and gave their comments of the manuscript. R.S. and S.K. came up with the study design. J.S. and J.G. assisted with the study design. J.S. developed and conducted the analyses. J.G., R.L., J.A., and S.C. designed and performed the experiments.

#### Funding

The study was internally funded by Eurofins Viracor Clinical Diagnostics.

#### Data availability

The datasets and codes for this study are proprietary and owned by Eurofins Viracor. They are available by contacting Jalal Siddiqui ([jalal.siddiqui@viracor.eurofinsus.com](mailto:jalal.siddiqui@viracor.eurofinsus.com)) or Rohita Sinha ([rohita.sinha@viracor.eurofinsus.com](mailto:rohita.sinha@viracor.eurofinsus.com)) upon reasonable request, subject to the following conditions: 1) The requester must provide a clear explanation of the intended use of the data. 2) The requester must sign a non-disclosure agreement (NDA) to protect the proprietary nature of the data. 3) The data will only be shared for research purposes and not for commercial purposes. 4) Eurofins Viracor reserves the right to deny requests for data sharing if the intended use is deemed incompatible with our company's interests.

#### Declarations

##### Ethics approval and consent to participate

Not applicable.

##### Consent for publication

Not applicable.

##### Competing interests

The authors are employed by Eurofins Viracor Clinical Diagnostics.

##### Author details

<sup>1</sup>Eurofins Viracor Clinical Diagnostics, 18000 W 99th St, Lenexa, KS 66219, United States of America

Received: 7 March 2024 / Accepted: 28 March 2025

Published online: 08 April 2025

#### References

- Ingulli E. Mechanism of cellular rejection in transplantation. *Pediatr Nephrol*. 2010;25:61–74.

- Rana A, Godfrey EL. Outcomes in solid-organ transplantation: success and stagnation. *Tex Heart Inst J*. 2019;46(1):75–6.
- 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*. 2017;114(30):8059–64.
- Mahdi BM. A glow of HLA typing in organ transplantation. *Clin Translational Med*. 2013;2:1–5.
- Kishore A, Petrek M. Next-generation sequencing based HLA typing: Deciphering Immunogenetic aspects of sarcoidosis. *Front Genet*. 2018;9:503.
- Huang Y, Dinh A, Heron S, Gasiewski A, Kneib C, Mehler H, et al. Assessing the utilization of high-resolution 2-field HLA typing in solid organ transplantation. *Am J Transplant*. 2019;19(7):1955–63.
- Vogiatzi P. Some considerations on the current debate about typing resolution in solid organ transplantation. *Transplantation Res*. 2016;5(1):1–6.
- Hosomichi K, Shiina T, Tajima A, Inoue I. The impact of next-generation sequencing technologies on HLA research. *J Hum Genet*. 2015;60(11):665–73.
- Danzer M, Niklas N, Stabenheimer S, Hofer K, Pröll J, Stückler C, et al. Rapid, scalable and highly automated HLA genotyping using next-generation sequencing: a transition from research to diagnostics. *BMC Genomics*. 2013;14(1):1–11.
- Liu C. A long road/read to rapid high-resolution HLA typing: the nanopore perspective. *Hum Immunol*. 2021;82(7):488–95.
- Liu C, Xiao F, Hoisington-Lopez J, Lang K, Quenzel P, Duffy B, et al. Accurate typing of human leukocyte antigen class I genes by Oxford nanopore sequencing. *J Mol Diagn*. 2018;20(4):428–35.
- Montgomery MC, Liu C, Petraroia R, Weimer ET. Using nanopore whole-transcriptome sequencing for human leukocyte antigen genotyping and correlating donor human leukocyte antigen expression with flow cytometric crossmatch results. *J Mol Diagn*. 2020;22(1):101–10.
- Mosbrugger TL, Dinou A, Duke JL, Ferriola D, Mehler H, Pagkrati I, et al. Utilizing nanopore sequencing technology for the rapid and comprehensive characterization of eleven HLA loci; addressing the need for deceased donor expedited HLA typing. *Hum Immunol*. 2020;81(8):413–22.
- Bravo-Egana V, Sanders H, Chitnis N. New challenges, new opportunities: next generation sequencing and its place in the advancement of HLA typing. *Hum Immunol*. 2021;82(7):478–87.
- Jain M. Nanopore sequencing updates using Q20+ and R10.4: Oxford Nanopore Technologies; 2021 [Available from: <https://nanoporetech.com/resource-centre/video/ncm21/nanopore-sequencing-updates-using-q20-and-r104>]
- Jain M. Human genome assembly and analysis using R10.4.1, Kit 14, and duplex data: Oxford Nanopore Technologies; 2022 [Available from: <https://nanoporetech.com/resource-centre/video/lc22/human-genome-assembly-and-analysis-using-r10-4-1-kit-14-and-duplex-data>]
- Sereika M, Kirkegaard RH, Karst SM, Michaelsen TY, Sørensen EA, Wollenberg RD, et al. Oxford nanopore R10. 4 long-read sequencing enables the generation of near-finished bacterial genomes from pure cultures and metagenomes without short-read or reference polishing. *Nat Methods*. 2022;19(7):823–6.
- van Deutekom HW, Kooter R, Geerlings J, Ruzius F-P, Meulenbergh P, Surendranath V, et al. P177 NGS typing results using Oxford nanopore sequencing: can minion data be reliably used for HLA typing? *Hum Immunol*. 2017;78:190.
- Liu C, Berry R. Rapid high-resolution typing of class I HLA genes by nanopore sequencing. *Bioinf Cancer Immunotherapy: Methods Protocols*. 2020:93–9.
- Liu C. *Athlon2*. 2020.
- Klasberg S, Putke K, Surendranath V, Schmidt A, Lange V, Schöfl G. P084 typing in the third generation: A HLA typing approach for nanopore sequencing data. *Hum Immunol*. 2019;80:116.
- Klasberg S, Schmidt AH, Lange V, Schöfl G. DR2S: an integrated algorithm providing reference-grade haplotype sequences from heterozygous samples. *BMC Bioinformatics*. 2021;22(1):1–15.
- Dilthey AT, Mentzer AJ, Carapito R, Cutland C, Cereb N, Madhi SA, et al. HLA\* LA—HLA typing from linearly projected graph alignments. *Bioinformatics*. 2019;35(21):4394–6.
- Kooter R, Ruzius FP, Penning MT, Mulder W, Rozemuller EH. 157-P: NGSengine: THE ULTIMATE TOOL FOR NGS HLA TYPING. *Hum Immunol*. 2013;74:156.
- Roquemuller EH, Penning M, Mulder W. 117-P: NGSengine: THE ULTIMATE TOOL FOR NGS HLA-TYPING. *Hum Immunol*. 2012;73:123.
- Stelet VN, Cita RF, Romero M, Mendes MF, Binato R. P054 using NGSengine® data analysis software to analyze third party NGS HLA data. *Hum Immunol*. 2019;80:94.
- Barker DJ, Maccari G, Georgiou X, Cooper MA, Flicek P, Robinson J, et al. The IPD-IMGT/HLA database. *Nucleic Acids Res*. 2023;51(D1):D1053–60.

28. Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*. 2009;25(14):1754–60.
29. Li H, Durbin R. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics*. 2010;26(5):589–95.
30. Li H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*. 2018;34(18):3094–100.
31. Becht C, Schmidt J, Blessing F, Wenzel F. Comparative analysis of alignment tools for application on nanopore sequencing data. *Curr Dir Biomedical Eng*. 2021;7(2):831–4.
32. De Coster W, D'hert S, Schultz DT, Cruts M, Van Broeckhoven C. *Bioinformatics*. 2018;34(15):2666–9. NanoPack: visualizing and processing long-read sequencing data.
33. 20150309.GRCh38\_full\_analysis\_set\_plus\_decoy\_hla [Internet]. 2015. Available from: [https://ftp.1000genomes.ebi.ac.uk/vol1/ftp/technical/reference/GRCh38\\_reference\\_genome/GRCh38\\_full\\_analysis\\_set\\_plus\\_decoy\\_hla.fa](https://ftp.1000genomes.ebi.ac.uk/vol1/ftp/technical/reference/GRCh38_reference_genome/GRCh38_full_analysis_set_plus_decoy_hla.fa)
34. Sundjaja JH, Shrestha R, Krishan K. McNemar And Mann-Whitney U Tests. 2020.
35. Hader SL, Hodge TW, Buchacz KA, Bray RA, Padian NS, Rausa A, et al. Discordance at human leukocyte antigen-DRB3 and protection from human immunodeficiency virus type 1 transmission. *J Infect Dis*. 2002;185(12):1729–35.
36. Tsamadou C, Engelhardt D, Platzbecker U, Sala E, Valerius T, Wagner-Drouet E, et al. HLA-DRB3/4/5 matching improves outcome of unrelated hematopoietic stem cell transplantation. *Front Immunol*. 2021;12:771449.
37. Johansson S, Juhos S, Redin D, Ahmadian A, Käller M. Comprehensive haplotyping of the HLA gene family using nanopore sequencing. 2018.
38. Bergström TF, Josefsson A, Erlich HA, Gyllenstein U. Recent origin of HLA-DRB1 alleles and implications for human evolution. *Nat Genet*. 1998;18(3):237–42.
39. Doxiadis GG, Hoof I, de Groot N, Bontrop RE. Evolution of HLA-DRB genes. *Mol Biol Evol*. 2012;29(12):3843–53.

## Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.