FACULTY OF ENGINEERING
AND ARCHITECTURE

Benchmark of NGS-based prediction algorithms for MHC class I and II genotyping in cancer research

Arne Claeys
Student number: 01304505

Supervisors: Prof. dr. Jimmy Van den Eynden, Prof. dr. Kathleen Marchal

Master's dissertation submitted in order to obtain the academic degree of
Master of Science in Biomedical Engineering

Academic year 2022-2023

GHENT
UNIVERSITY

VUB VRIJE
UNIVERSITEIT
BRUSSEL

Preface and copyright

Preface

As soon as the opportunity arose to follow a master's programme in bioinformatics at Ghent University, I knew it was the perfect fit for me to merge my interests in computer science and my wish to contribute to advancements in medicine. Following graduation in 2018 my thirst for knowledge remained unquenched and I decided to enrol in the Master of Biomedical Engineering.

Now four years later, I am writing a master's dissertation for the second time in my career, while simultaneously working on a PhD in the lab of Prof. Van den Eynden. As I sit down to write this preface, I realise that both stories in my life have now entered their epilogue and I am filled with a sense of accomplishment and gratitude.

I am grateful to have had the opportunity to pursue this degree, and for the support of my supervisor Prof. Van den Eynden, who graciously allowed me to finish a master's programme simultaneously with my PhD. His guidance and support were instrumental in helping me to shape and structure my ideas. Prof. Marchal provided valuable suggestions to make this research more appealing for method developers. I feel confident that the final result, which is now available as a preprint on BioRxiv (Claeys et al., 2022), would not have been possible without their extensive feedback.

Finally, I would like to express my sincere gratitude to Prof. Segers for making time on multiple occasions to find solutions for the eternal administrative challenges that come with spreading a master's programme over 5 years.

Thank you to everyone who has supported me along this journey.

Arne Claeys

29/10/2022

Permission to use on loan

The author(s) gives (give) permission to make this master dissertation available for consultation and to copy parts of this master dissertation for personal use. In all cases of other use, the copyright terms have to be respected, in particular with regard to the obligation to state explicitly the source when quoting results from this master dissertation. (29/10/2022)

Remark on the master's dissertation and the oral presentation

This master's dissertation is part of an exam. Any comments formulated by the assessment committee during the oral presentation of the master's dissertation are not included in this text.

Abstract

Background

The Human Leukocyte Antigen (HLA) genes are a group of highly polymorphic genes that are located in the Major Histocompatibility Complex (MHC) region on chromosome 6. The HLA genotype affects the presentability of tumour antigens to the immune system. While knowledge of these genotypes is of utmost importance to study differences in immune responses between cancer patients, gold standard, PCR-derived genotypes are rarely available in large Next Generation Sequencing (NGS) datasets. Therefore, a variety of methods for *in silico* NGS-based HLA genotyping have been developed, bypassing the need to determine these genotypes with separate experiments. However, there is currently no consensus on the best performing tool.

Results

We evaluated 13 MHC class I and/or class II HLA callers that are currently available for free academic use and run on either Whole Exome Sequencing (WES) or RNA sequencing data. Computational resource requirements were highly variable between these tools. Three orthogonal approaches were used to evaluate the accuracy on several large publicly available datasets: a direct benchmark using PCR-derived gold standard HLA calls, a correlation analysis with population-based allele frequencies and an analysis of the concordance between the different tools. The highest MHC-I calling accuracies were found for *Optitype* (98.0%) and *arcasHLA* (99.4%) on WES and RNA sequencing data respectively, while for MHC-II *HLA-HD* was the most accurate tool for both data types (96.2% and 99.4% on WES and RNA data respectively). We demonstrated that the combination of *Optitype*, *HLA*LA*, *Kourami* and *Polysolver* in a consensus majority voting-based metaclassifier improved the accuracy for MHC-I on WES data to 99.0% and the combination of *HLA*LA*, *HLA-HD*, *PHLAT* and *xHLA* improved the accuracy to 98.4% for MHC-II.

Conclusion

The optimal strategy for HLA genotyping from NGS data depends on the availability of either WES or RNA data, the size of the dataset and the available computational resources. If sufficient resources are available, we recommend *Optitype* and *HLA-HD* for MHC-I and MHC-II genotype calling respectively.

Benchmark of NGS-based prediction algorithms for MHC class I and II genotyping in cancer research

Arne Claeys

Supervisor(s): Jimmy Van den Eynden, Kathleen Marchal

Abstract – **The highly polymorphic Major Histocompatibility Complex (MHC) molecules are indispensable actors in the immune response to cancer. To study tumour-immune interactions and to discover new genomic predictors for immunotherapy responses, it is important to have accurate methods available that can predict MHC genotypes from large genomic NGS datasets. Numerous tools have been developed with that objective, but there is currently no consensus on the best performing algorithms. In this study, we performed an extensive benchmark of 13 different tools using data from the 1000 Genomes Project as well as The Cancer Genome Atlas (TCGA) and propose a simple method to make consensus HLA allele predictions.**

Keywords – **HLA genotyping; benchmark; tumour-immune interaction**

I. Introduction

The human Major Histocompatibility Complex (MHC) is a gene complex located on the p-arm of chromosome 6 that contains two large clusters of genes with antigen processing and presentation functions: the MHC class I and MHC class II regions [1–3].

MHC class I molecules are involved in the presentation of endogenous antigens to cytotoxic T-cells and consist of a heavy chain encoded by one of the MHC class I genes (*HLA-A*, *HLA-B* or *HLA-C)*, and a light β_2 microglobulin chain [4–6]. Their role in tumour immunity has been established for a long time [7]. Indeed, they can present neoantigens, small mutated peptides, to CD8+ T cells, resulting in an immune response and cancer cell death [8, 9].

The most frequently studied MHC class II genes include *HLA-DPA1*, *HLA-DPB1*, *HLA-DQA1*, *HLA-DQB1*, *HLA-DRA* and *HLA-DRB1*. They encode alpha/beta heterodimers that form the MHC class II protein complex. The role of these genes in anti-tumour immunity is emerging [10–12].

The peptide-binding region of HLA molecules is highly polymorphic and specific HLA alleles determine neoantigen binding and presentation to the immune system. Genotype dependent differences in HLA binding affinity could lead to differential responses to immunotherapy, as illustrated by the association that has been described between MHC-I genotypes (e.g., *HLA-B62*) and survival in immune checkpoint blockade (ICB)-treated advanced melanoma patients [13]. It is currently unclear whether MHC-II genotypes also determine responses to immunotherapy.

Such association studies require knowledge of the HLA genotype. PCR methods are currently the gold standard for this genotyping but datasets with PCR-based HLA calls are rarely available [14–16]. HLA genotyping can also be performed on Next Generation Sequencing (NGS) data. A plethora of tools has been developed for this task. *Polysolver* and *Optitype* are often recommended as the best performing tools for MHC-I genotyping [17]. For MHC-II genotyping there is currently no consensus about the best method. Several benchmarks have been performed previously [15, 17–24], but these were either not applied to MHC class II or did not include some recently published tools.

In this study, we compiled a list of 13 tools that predict HLA genotypes from NGS data and benchmarked their performance on both the 1000 genomes dataset and on an independent cell line dataset [25]. Subsequently we assessed their performance on 9162 WES and 9761 RNA sequencing files from The Cancer Genome Atlas (TCGA) by comparing the predicted allele frequencies with reference population allele frequencies. Based on these findings, we give recommendations on which tool to use for a given data type and how the outputs of multiple tools can be combined into a consensus prediction.

II. Selection of 13 HLA genotyping tools

We identified 22 available HLA genotyping tools from literature. Thirteen tools that were free for academic use, applicable on Whole Exome Sequencing (WES), Whole Genome Sequencing (WGS) or RNA-Seq data and ran on Ubuntu 20.04 were included in this study: *arcasHLA*, *HLA-HD*, *HLA-VBSeq*, *HLA*LA*, *HLAforest*, *HLAminer*, *HLAscan*, *Kourami*, *Optitype*, *PHLAT*, *Polysolver*, *seq2HLA* and *xHLA*. All 13 tools can make allele predictions for the three MHC class I genes (*HLA-A*, *HLA-B* and *HLA-C*) and 9 tools support additional calling of the MHC class II genes *HLA-DPA1*, *HLA-DPB1*, *HLA-DQA1*, *HLA-DQB1* and *HLA-DRB1*. Two methods support only a subset of the MHC class II genes: *xHLA* does not support calling *HLA-DPA1* and *HLA-DQA1*, while *PHLAT* does not support *HLA-DPA1* and *HLA-DPB1*. The tools also differ in which data types they support: 6 of them require WES data, 3 tools require RNA data and the 4 remaining tools support both data types.

III. Benchmark of the resource consumption

Firstly, the computing time and memory usage of the thirteen selected tools were measured on a random subset of 10 WES and 10 RNA sequencing files from the TCGA project (Figure 4, main text). Among the 10 WES-supporting methods *Optitype* (median 2.48 hours) and *HLA*LA* (median 1.84 hours) require the largest computing time. Apart from being computationally intensive, *HLA*LA* is also the most memory demanding WES tool (median 36.3 GiB per file). Among the 7 RNA-supporting methods, *HLA-HD* has the longest computing time per sample (median 15.0 hours). At the other end of the spectrum, the sole pseudoalignment-based tool *arcasHLA* takes only 38s per file. The most memory intensive tool is *HLA-HD*

(median memory peaks of 103.1 GiB), followed by *Optitype* (median 34.1 GiB).

IV. Accuracy on WES data

The 10 selected algorithms that are compatible with WES data were benchmarked using data from the *1000 Genomes project* [48] (Figure 5, main text). For MHC-I genes (*HLA-A*, *HLA-B*, *HLA-C*), the best accuracy was obtained with *Optitype* (98.0%), followed by *Polysolver* and *HLA*LA* (94.9% and 94.4% respectively). For MHC-II genes (*HLA-DQA1*, *HLA-DQB1* and *HLA-DRB1*), the best allele predictions were made using *HLA-HD* and *HLA*LA* (96.2% and 95.7% accuracy respectively). These were the only two methods to reach an accuracy of 90% on all tested MHC-II genes. *HLAscan* (74.2%), *HLA-VBSeq* (60.2%) and *HLAminer* (53.8%) performed considerably worse than the other tools.

V. Accuracy on RNA data

We then evaluated the 7 selected methods that support HLA calling on RNA sequencing data from the *1000 genomes project* [49] (Figure 5, main text).

A*rcasHLA* and *Optitype* had the best MHC-I allele predictions (99.4% and 99.2% accuracy, respectively), followed by *HLA-HD* (98.0%), *seq2HLA* (95.9%) and *PHLAT* (95.4%). Similar accuracies were found for MHC-II allele predictions, with *HLA-HD*, *PHLAT* and *arcasHLA* performing the best (99.4%, 98.9% and 98.1%, respectively).

VI. Validation on independent datasets

Being one of the few large sequencing datasets for which gold standard HLA genotypes for both MHC classes are available, many algorithms included in our benchmark were developed, optimized and validated using files from the *1000 genomes* project, introducing a potential bias. Therefore, we performed an indirect and independent evaluation on a large NGS dataset obtained from TCGA.

We first compared the observed allele frequencies for each tool with the expected population frequencies. We calculated how often each of the alleles was predicted by a certain tool to obtain an observed allele frequency, stratifying for Caucasian American and African American ethnicities. By comparing these frequencies to the expected allele frequencies, as derived from *Allele Frequency Net* [50], strong significant correlations were found for the WES-based tools *HLA-HD*, *HLA*LA*, *Optitype*, *Polysolver* and *xHLA* and for the RNA-based tools *Optitype*, *arcasHLA* and *PHLAT*. The correlations were considerably worse for *HLA-VBSeq*, and *HLAforest* than for the other tools (Figure 6, main text). These findings largely confirm the results of the benchmark on the 1000 genomes data.

We then calculated for each pair of tools how often their predictions are concordant (Figures S8-S11). Tools that performed poorly in the previous analyses (e.g., *HLAminer*, *HLA-VBSeq* and *HLAforest*) consistently have a low concordance with all other tools. In contrary, tools that scored high in the previous analyses (such as *Optitype*, *HLA*LA*, *arcasHLA* and *HLA-HD*) made predictions that are consistent with each other.

VII. Consensus predictions

We noted that only for a very small fraction of the samples the genotypes are wrongly typed by all tools simultaneously (median 0.79% for WES and 0.68% for RNA; Figures S12-S13). This complementarity of the tools' allele predictions opens the possibility to combine predictions of different HLA callers into a consensus prediction. We first applied a majority voting algorithm to the output of all tools, with the predicted allele pair being the one with most votes. On the WES data, this approach outperforms the predictions of each individual tool for all genes. On RNA data, where the best tools already attain accuracies over 99% by themselves, only minor improvements were made by combining the results (Figure S14).

Based on these results, we determined the minimal number of tools that must be included in the WES-based metaclassifier to produce reliable results (See *Methods*; Figure 7, main text). For the WES data, including 4 tools in the model led to a considerable improvement for all genes for both MHC classes. The best accuracies were observed when *Optitype*, *HLA*LA*, *Kourami* and *Polysolver* were combined for MHC-I predictions (99.0% accuracy) and with *HLA*LA*, *HLA-HD*, *PHLAT* and *xHLA* for MHC-II predictions (98.4% accuracy).

VIII. Conclusion

We found that *Optitype*, *Polysolver*, *HLA-HD*, *HLA*LA* and *xHLA* are all solid choices for WES-based MHC genotyping, while *Optitype*, *HLA-HD*, *arcasHLA* and *PHLAT* are the better performing tools for RNA data. On the other hand, *HLAminer*, *HLA-VBSeq* and *HLAScan* performed rather poorly in our benchmark.

The optimal strategy for HLA genotyping depends on a few factors: the availability of WES or RNA data, the size of the dataset that needs to be analysed and the available computational resources. For WES data, *Optitype* and *HLA-HD* are the best performing individual tools for MHC class I and MHC class II typing, respectively. For RNA data, the same tools are recommended when sufficient computational resources are available. However, the large resource and time consumption of *HLA-HD* on RNA data makes its usage rather impractical on large datasets. As an alternative, *arcasHLA* is recommended, which is both the fastest and more accurate tool for RNA that supports all 5 MHC class II genes.

Finally, we have demonstrated that the accuracy of the WES-based HLA genotype predictions can be improved further by combining the output of *Optitype*, *HLA*LA*, *Kourami* and *Polysolver* for MHC-I typing and combining *HLA*LA*, *HLA-HD*, *PHLAT* and *xHLA* for MHC-II typing using a majority voting rule. For RNA data a similar metaclassifier approach did not lead to a further improvement of the prediction accuracies.

Acknowledgements

The results shown here are in whole or part based upon data generated by the TCGA Research Network: https://www.cancer.gov/tcga.

References

1. Trowsdale J. Genomic structure and function in the MHC. Trends in Genetics. 1993;9:117–22.

2. Beck S, Geraghty D, Inoko H, Rowen L, Aguado B, Bahram S, et al. Complete sequence and gene map of a human major histocompatibility complex. Nature 1999 401:6756. 1999;401:921–3.

3. Horton R, Wilming L, Rand V, Lovering RC, Bruford EA, Khodiyar VK, et al. Gene map of the extended human MHC. Nature Reviews Genetics 2004 5:12. 2004;5:889–99.

4. Halenius A, Gerke C, Hengel H. Classical and non-classical MHC I molecule manipulation by human cytomegalovirus: so many targets—but how many arrows in the quiver? Cellular & Molecular Immunology 2015 12:2. 2014;12:139–53.

5. Allen RL, Hogan L. Non-Classical MHC Class I Molecules (MHC-Ib). eLS. 2013. https://doi.org/10.1002/9780470015902.A0024246.

6. Hewitt EW. The MHC class I antigen presentation pathway: strategies for viral immune evasion. Immunology. 2003;110:163.

7. Philipps C, McMillan M, Flood PM, Murphy DB, Forman J, Lancki D, et al. Identification of a unique tumor-specific antigen as a novel class I major histocompatibility molecule. Proc Natl Acad Sci U S A. 1985;82:5140–4.

8. Rooney MS, Shukla SA, Wu CJ, Getz G, Hacohen N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. Cell. 2015;160:48–61.

9. Schumacher TN, Schreiber RD. Neoantigens in cancer immunotherapy. Science (1979). 2015;348:69–74.

10. Axelrod ML, Cook RS, Johnson DB, Balko JM. Biological consequences of MHC-II expression by tumor cells in cancer. Clinical Cancer Research. 2019;25:2392–402.

11. Alspach E, Lussier DM, Miceli AP, Kizhvatov I, DuPage M, Luoma AM, et al. MHC-II neoantigens shape tumour immunity and response to immunotherapy. Nature. 2019;574:696–701.

12. Sun Z, Chen F, Meng F, Wei J, Liu B. MHC class II restricted neoantigen: A promising target in tumor immunotherapy. Cancer Lett. 2017;392:17–25.

13. 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 (1979). 2018;359:582–7.

14. Szolek A, Schubert B, Mohr C, Sturm M, Feldhahn M, Kohlbacher O. OptiType: precision HLA typing from next-generation sequencing data. Bioinformatics. 2014;30:3310–6.

15. Bauer DC, Zadoorian A, Wilson LOW, Alliance MGH, Thorne NP. Evaluation of computational programs to predict HLA genotypes from genomic sequencing data. Brief Bioinform. 2018;19:179–87.

16. Orenbuch R, Filip I, Comito D, Shaman J, Pe'Er I, Rabadan R. arcasHLA: high-resolution HLA typing from RNAseq. Bioinformatics. 2020;36:33–40.

17. Matey-Hernandez ML, Brunak S, Izarzugaza JMG. Benchmarking the HLA typing performance of Polysolver and Optitype in 50 Danish parental trios. BMC Bioinformatics. 2018;19:1–12.

18. Lee M, Seo JH, Song S, Song IH, Kim SY, Kim YA, et al. A New Human Leukocyte Antigen Typing Algorithm Combined With Currently Available Genotyping Tools Based on Next-Generation Sequencing Data and Guidelines to Select the Most Likely Human Leukocyte Antigen Genotype. Front Immunol. 2021;12:4080.

19. Li X, Zhou C, Chen K, Huang B, Liu Q, Ye H. Benchmarking HLA genotyping and clarifying HLA impact on survival in tumor immunotherapy. Mol Oncol. 2021;15:1764–82.

20. Chen J, Madireddi S, Nagarkar D, Migdal M, vander Heiden J, Chang D, et al. In silico tools for accurate HLA and KIR inference from clinical sequencing data empower immunogenetics on individual-patient and population scales. Brief Bioinform. 2021;22:1–11.

21. Yu Y, Wang K, Fahira A, Yang Q, Sun R, Li Z, et al. Systematic comparative study of computational methods for HLA typing from next-generation sequencing. HLA. 2021;97:481–92.

22. Kiyotani K, Mai TH, Nakamura Y. Comparison of exome-based HLA class I genotyping tools: identification of platform-specific genotyping errors. Journal of Human Genetics 2017 62:3. 2016;62:397–405.

23. Liu P, Yao M, Gong Y, Song Y, Chen Y, Ye Y, et al. Benchmarking the Human Leukocyte Antigen Typing Performance of Three Assays and Seven Next-Generation Sequencing-Based Algorithms. Front Immunol. 2021;12:840.

24. Yi J, Chen L, Xiao Y, Zhao Z, Su X. Investigations of sequencing data and sample type on HLA class Ia typing with different computational tools. Brief Bioinform. 2021;22:1–6.

25. Abaan OD, Polley EC, Davis SR, Zhu YJ, Bilke S, Walker RL, et al. The exomes of the NCI-60 panel: a genomic resource for cancer biology and systems pharmacology. Cancer Res. 2013;73:4372–82.

Contents

List of figures and tables

List of abbreviations

APPM: antigen processing and presentation machinery
CDC: complement-dependent cytotoxicity
cDNA: complementary DNA
ERAP: endoplasmic reticulum aminopeptidase
HLA: Human Leukocyte Antigen
HPV: human papillomavirus
ILP: Integer Linear Programming
MHC: Major Histocompatibility Complex
NGS: Next Generation Sequencing
NK: Natural-Killer
PCR: Polymerase Chain Reaction
PCR-SBT: PCR with sequencing-based typing
PCR-SSOP: PCR with sequence-specific oligonucleotide probes
PCR-SSP: PCR with sequence-specific primers
POG: partial order graph
PRG: population reference graph
RNA-seq: RNA sequencing
SMMQ: sum of mismatch qualities
SNP: single nucleotide polymorphism
TAP: transporter associated with antigen processing
TCR: T cell receptor (TCR)
WES: whole-exome sequencing
WGS: whole-genome sequencing

Chapter 1
Background

1. The Major Histocompatibility Complex

The Major Histocompatibility Complex (MHC) is a region in the vertebrate genome that contains genes with a critical role in the workings of the immune system (Horton et al., 2004). In humans, the MHC is also called the Human Leukocyte Antigen (HLA) system (Klein & Sato, 2000). Its discovery can be traced back to the middle of the 20th century, when Gorer and Snell made a series of pioneering observations in mice that demonstrated the existence of genetic factors that control the rejection of tissue transplants (Gorer, 1937; Klein, 2001; Snell, 1948). Later, a homologous system was found in humans after the discovery of the first leukocyte antigen (now known as *HLA-A*2) by Dausset (Richmond, 2009; Thorsby & Thorsby, 2009).

1.1. The human MHC: genes and protein complexes

The human MHC is located on the p-arm of chromosome 6 and contains two large clusters of genes with antigen processing and presentation functions: the MHC class I and MHC class II regions (Beck et al., 1999; Horton et al., 2004; Trowsdale, 1993). A third, less well characterized region is the MHC class III region (Figure 1). Unlike the other two MHC classes, its genes are not involved in antigen presentation (Sabbatino et al., 2020).

MHC class I region

The MHC class I genes are commonly further subdivided into *classical* and *non-classical* genes (Bjorkman & Parham, 1990; Horton et al., 2004; Shiina et al., 2009). There are three *classical* class I genes (*HLA-A*, *HLA-B*, and *HLA-C*) and several *non-classical* class I genes (among which *HLA-E*, *HLA-F*, and *HLA-G* are the most prominent members).

Classical MHC-I genes: Classical MHC-I genes are expressed in (nearly) all cells of the human body. Their gene products are incorporated in MHC class I molecules, which are protein complexes consisting of a heavy chain encoded by one of the HLA genes and a light β_2 microglobulin chain (Allen & Hogan, 2013; Halenius et al., 2014; Hewitt, 2003). These MHC-I molecules present endogenous antigens (i.e., antigens originating from within the organism) to CD8+ (cytotoxic) T cells. The latter cell type is tasked with destroying cells that are considered harmful to the host including cancer cells and cells infected by a pathogen (Andersen et al., 2006; Horton et al., 2004; Klein & Sato, 2000).

Non-classical MHC-I genes: Proteins derived from non-classical MHC-I genes also form dimers with β_2 microglobulin, but are involved in different immune related pathways, such as Natural-Killer (NK) mediated immune response (*HLA-E*) (Braud et al., 1998) and maternal-fetal immune tolerance (*HLA-G*) (Zhuang et al., 2021).

MHC class II region

Classical MHC-II genes: The most frequently studied (classical) MHC-II genes include *HLA-DPA1, HLA-DPB1, HLA-DQA1, HLA-DQB1, HLA-DRA* and *HLA-DRB1*. Contrary to MHC-I genes, the MHC-II genes are expressed only in specific cell types, including B cells, macrophages, dendritic cells, and thymic epithelial cells (Klein & Sato, 2000). The HLA genes in this region encode alpha/beta heterodimers that form MHC-II molecules that present exogenous antigens (i.e., antigens originating from outside the organism) to CD4+ T (helper) cells (Neefjes et al., 2011). When the T cell receptor (TCR) of CD4+ T cells recognizes the presented antigen, the T cell is activated and an immune

response is initiated (Johnson et al., 2021). Activation of CD4+ T cells will result in differentiation of the T cell into Th1, Th2, Th17, or Treg cells.

Non-classical MHC-II genes: Apart from these classical MHC-II genes, there are also non-classical MHC-II genes (*HLA-DMA*, *HLA-DMB*, *HLA-DOA* and *HLA-DOB*), that form alpha/beta heterodimers that regulate binding of peptides to MHC-II molecules (Mellins & Stern, 2014). *HLA-DM* is present on the surface of immature dendritic cells and B cells. It regulates the activity of MHC-II molecules by catalysing the dissociation of CLIP, a surrogate ligand needed to maintain structural integrity of MHC-II molecules during protein folding, in exchange for antigenic peptides (Arndt et al., 2000; Denzin et al., 1997; Zhong et al., 1996). *HLA-DO* is expressed in the thymic medulla, B cells and on the surface of some dendritic cells (Welsh & Sadegh-Nasseri, 2020). *HLA-DO* then modulates the activity of *HLA-DM* (Arndt et al., 2000; Denzin et al., 1997).

Genes involved in antigen processing: The MHC class II subregion also includes important genes involved in antigen processing, among which the TAP1 and TAP2 genes. These genes encode proteins of the *Transporter associated with antigen processing* (TAP) complex which transports peptides from the cytosol into the endoplasmic reticulum, where they are subsequently loaded onto MHC-I molecules (Jhunjhunwala et al., 2021).

MHC class III region

The MHC class III region, located between the MHC class I and class II regions, also contains genes with immune related functions. This region is particularly dense in genes: about 72% of the region is transcribed, with on average one gene every 10 kb (Milner, 2001; T. Xie et al., 2003). Among the genes found in this region are some genes encoding complement proteins (involved in the innate immune system), heat shock proteins, and certain cytokines (TNF, LTA and LTB) (Horton et al., 2004).

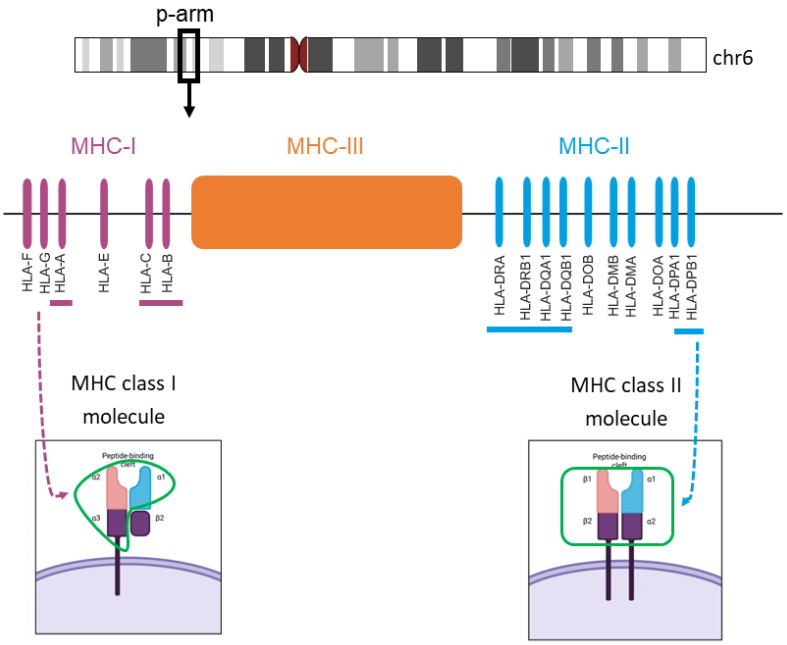

Figure 1. Gene map of the MHC region. The human MHC region are located on the p-arm of chromosome 6 and consists of 3 gene clusters: MHC-I (purple), MHC-II (blue) and MHC-III (orange). All depicted HLA genes encode components of (classical and non-classical) MHC molecules. Classical MHC genes are underlined. Created with BioRender.com

1.2. Polymorphism of the HLA genes

The MHC region constitutes one of the most polymorphic loci of the human genome, with more than 36,000 alleles recorded in the IPD-IMGT/HLA database (Robinson et al., 2015). The degree of polymorphism varies for different HLA genes. With more than 9,000 known alleles, the *HLA-B* gene is the most polymorphic HLA gene (Raghavan & Geng, 2015; Robinson et al., 2015). In contrast, *HLA-DRA* is characterized by an almost complete absence of polymorphism (Matern et al., 2020). For each gene, the largest variability is encountered in the exons that encode the peptide-binding groove of the MHC molecules (exon 2 and 3 for MHC class I, and exon 2 for MHC class II).

This large variety of HLA alleles, which are formed by point mutations, indels and recombination events, is hypothesized to have evolved as a defence of vertebrates against rapidly evolving pathogens (Fabreti-Oliveira et al., 2018; Markov & Pybus, 2015). Two evolutionary pressures are thought to be involved: heterozygote advantage and frequency-dependent selection (Sommer, 2005). First, heterozygote advantage refers to the presumed fitness advantage of individuals with two different HLA alleles at a given locus. As differences in the amino acid sequence of HLA variants lead to a distinctive repertoire of peptides that can bind efficiently, heterozygotes can target a broader range of epitopes and have a fitness advantage over individuals with two identical alleles (Markov & Pybus, 2015). Secondly, frequency-dependent selection occurs because parasite constantly adapt to become resistant to the most common HLA alleles as this allows them to spread faster through a population. In this context, carrying a rare allele is advantageous. Both selection pressures will eventually lead to a high HLA polymorphism in the population (Sommer, 2005).

1.3. HLA gene nomenclature

Different HLA alleles are named according to the Nomenclature for Factors of the HLA System (Marsh, 2022; Marsh et al., 2010). In this system allele names are composed of 1 to 4 fields, delimited by colons. The first field indicates the allele family, a classification based on serological assays (see below). The second field expresses differences in the amino acid sequence. The third and fourth fields, respectively represent synonymous variants in the coding region and variation outside the coding region (Hurley, 2021; Marsh et al., 2010).

In addition to this naming system, HLA proteins are also clustered into G and P groups, for which only polymorphism in the antigen-binding groove is considered. HLA alleles in the same G group share the same nucleotide sequence across the exons that code for the peptide binding domains, whereas alleles in the same P group share the same amino acid sequence in the peptide binding domain (Marsh et al., 2010).

1.4. Clinical role of MHC

The exact amino acid sequence in the HLA antigen-binding domain determines which peptides can bind to the MHC and therefore influences which antigens trigger an immune response (Nielsen et al., 2007). As a result, the specific combination of HLA alleles that an individual carries impacts their susceptibility to various infectious and autoimmune diseases. For example, *HLA-B27, HLA-B57* and *HLA-B51* have been associated with a longer time interval between HIV-1 infection and the onset of AIDS (Kaslow et al., 1996). There is also a strong association between certain HLA alleles and the occurrence of autoimmune diseases, such as type I diabetes, rheumatoid arthritis, psoriasis, and asthma (Hosomichi et al., 2015; Simmonds & Gough, 2009).

The clinical relevance of HLA polymorphism extends beyond establishing allele disease associations. In fact, the HLA genes were originally studied in the context of transplant rejection (Klein, 2001). Indeed, to prevent rejection of a transplanted organ, it is essential that the HLA alleles of the donor and recipient match as closely as possible, which necessitates the availability of accurate methods to determine which set of HLA alleles are carried by both subjects.

Furthermore, MHC molecules play an important role in the immune system's recognition of cancer cells by presenting small peptides, known as neoantigens, to T cells. These neoantigens are produced through the following process. Cancer is caused by mutations in the DNA of somatic cells, what results in the production of aberrant proteins. Those proteins are eventually cleaved into smaller fragments by the proteasome, and the resulting small, mutated peptides can then bind to MHC molecules that are presented to the immune system at the cell's surface. As the neoantigens are derived from mutated DNA, they are unknown to the immune system and will trigger an immune response.

The role of MHC class I molecules in this process has been established for a long time (Philipps et al., 1985): they present neoantigens to cytotoxic CD8+ T cells, resulting in cancer cell death (Rooney et al., 2015; Schumacher & Schreiber, 2015) (Figure 2).

Additionally, there is emerging evidence that MHC class II molecules are also implicated in tumour-immune interactions. This can occur through either a direct or an indirect mechanism. First, some tumours express MHC-II themselves and can directly interact with CD4+ T cells. Secondly, cancer cells can also secrete neoantigens that are taken up and presented on the MHC-II of infiltrating antigen presenting cells.

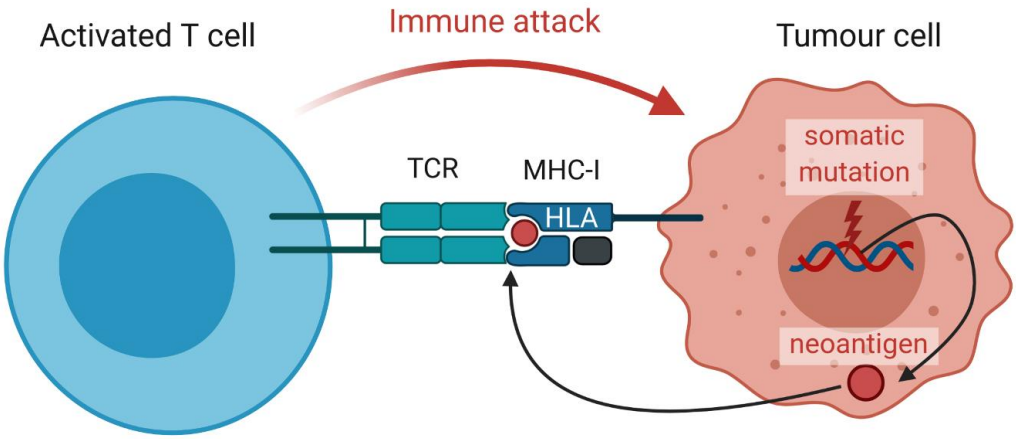

Figure 2. Presentation of neoantigens to CD8+ T cells. Neoantigens are short peptides derived from mutated DNA in tumours. When they are presented to CD8+ T cells via MHC-I molecules at the cell surface, an immune response will follow that leads to cancer cell death. Created with BioRender.com

2. Next Generation Sequencing (NGS)

Next generation sequencing (NGS) is an umbrella term for high-throughput DNA sequencing techniques that are based on the parallel processing of millions of short DNA fragments at once (Mardis, 2008). These sequencing technologies, which were preceded by older methods based on Sanger sequencing, stand out for their significantly improved speed and cost-efficiency (Metzker, 2009). The development of NGS has been indispensable for modern genomics research and opened new avenues for understanding the genetic basis of different diseases, studying the cancer genome, and personalising medical treatments (Jurgens et al., 2022; Pruis et al., 2022; Vogelstein et al., 2013).

2.1. Workflow for Next Generation Sequencing

NGS techniques exist for both DNA and RNA molecules and can be applied to either the entire genome or to a specific region of interest. A typical whole-genome sequencing (WGS) workflow consists of the following steps. First, DNA is extracted from the collected samples. Then the DNA molecules are fragmentized, amplified, and ligated to adapters, resulting in a library of short DNA chunks that can be sequenced using one of several NGS platforms (McCombie et al., 2019).

An important property of the resulting sequencing data is the *sequencing depth* or *coverage*. These terms refer to the average number of times each nucleotide is measured (Jiang et al., 2019; Sims et al., 2014). The sequencing depth influences the accuracy of various downstream analysis steps.

For many practical applications, especially those that require a high sequencing depth, it is neither economically feasible nor necessary to sequence the entire genome. To accommodate this, it is possible to use DNA capture methods prior to sequencing and focus on a subset of the genome. For example, in whole-exome sequencing (WES) protocols the exons are targeted using array-based or liquid-based hybridisation methods (Parla et al., 2011; Teer & Mullikin, 2010). Even though the exons make up only 1% of the genome, they contain the protein coding sequences, which makes these techniques appropriate to answer many research questions (S. B. Ng et al., 2009).

Additionally, there are also RNA-sequencing (RNA-seq) methods, which involve an additional step in the library preparation process where complementary DNA (cDNA) is synthesized from the mRNA. RNA-seq is a powerful tool that can be used to quantify gene expression in different biological processes (Stark et al., 2019).

2.2. Sequence assembly and alignment

NGS technologies produce a collection of short nucleotide sequences *(reads)* rather than the full genomic or transcriptomic sequences. Therefore, these steps are followed by a bioinformatics pipeline that usually starts with either sequence assembly, or mapping of the reads to a reference genome.

Sequence assembly

Sequence assembly methods aim to construct longer contiguous genomic or transcriptomic sequences *(contigs)* based on the sequenced reads (Bleidorn, 2017b).

We can distinguish two types of sequence assembly methods: *de novo* and *reference-based* assembly methods. In *de novo assembly* sequence reads are concatenated into contiguous sequences without the use of a reference genome. This form of assembly is required when a closely related genome

sequence does not yet exist. Examples of commonly used *de novo* alignment algorithms are Velvet (Zerbino & Birney, 2008), SOAPdenovo (Luo et al., 2012) and ABYSS (Jackman et al., 2017).

Reference-based assembly, on the other hand, involves mapping each read to a longer template sequence and constructing a new consensus sequence that is similar but not necessarily identical to the used template (P. C. Ng & Kirkness, 2010; Tiwary, 2022). This strategy is used by the tool SeqMap, MAQ and RMAP.

Read mapping / alignment

Read mapping (also referred to as read alignment) is the task of determining the correct position of each read relative to a reference sequence (Alser et al., 2021; Bleidorn, 2017a). There are several algorithms available for aligning NGS data, including BWA (H. Li & Durbin, 2009), Bowtie (Langmead & Salzberg, 2012), and STAR (Dobin et al., 2013).

Specifically for mapping RNA-seq reads to a reference genome, it is important to use a splice aware alignment tool (e.g., STAR) (Baruzzo et al., 2016; Williams et al., 2014). This is due to the fact that RNA-seq reads are derived from mature RNA which does no longer contain introns, while the reference genome includes both exons and introns. Splice-aware aligners are able to identify splice junctions and align reads to exons while ignoring introns. Conversely, the tools that do not consider splice junctions would need to tolerate large gaps when mapping reads that span multiple exons, which would lead to improper alignment (Baruzzo et al., 2016).

Pseudo-alignment

Base-by-base sequence alignment is a computationally intensive process that can be avoided for certain applications (Alser et al., 2021) by using pseudo-alignment methods such as Salmon (Patro et al., 2017), Sailfish (Patro et al., 2014) and Kallisto (Bray et al., 2016). These types of methods have been successfully applied to gene quantification and HLA genotyping (Corchete et al., 2020; Orenbuch, Filip, Comito, et al., 2020).

2.3. Reference genome

The human reference genome serves as a standardised version of the human genome where new sequencing data can be aligned to. The existence of a “reference genome” has several advantages. First, aligning NGS data to a reference is less computationally demanding than performing de novo assembly. Without a reference the latter step would be required for every new dataset. Secondly, it makes genome annotation, assigning genes and other features to genomic regions a one-time effort: genome annotation files are constructed for the reference genome and by aligning to this genome genes can be easily localized in new sequencing data (Frankish et al., 2021).

Different versions of the reference genome exist. The first sequence of the human genome was constructed as part of the *Human Genome Project* in 2001. The first "complete" genome was published in 2003, but it still contained multiple gaps (unlocalized sequences) in the assembly (Ballouz et al., 2019). At present, the quality of the human reference genome has improved considerably. Despite these improvements, the latest version (GRCh38) still contains "unlocalized sequences" (contigs that might be associated with a specific chromosome but cannot be ordered or oriented on that chromosome) (Assembly Terminology - Genome Reference Consortium).

Apart from the primary assembly, which consists of a single contig per chromosome, so-called *alternate loci* for a few polymorphic regions have been added to reference genome (Assembly Terminology - Genome Reference Consortium; Seal et al., 2013). These *alternate loci*, also called *alternative (ALT) contigs*, better capture the genomic diversity of a few polymorphic regions. Finally, variants of the reference genome exist, that additionally include sequences for common HLA alleles (Zheng-Bradley et al., 2017).

3. How to genotype HLA?

3.1. Serological assays

Complement-dependent cytotoxicity (CDC) assays were the standard method for HLA genotyping, until they were superseded by PCR-based typing techniques in the 1990s (Blasczyk, 2003; Gautreaux, 2017). In CDC assays, target cells are incubated with antibodies that bind to specific variants of HLA molecules. When complement proteins are added to the mixture, they initiate a series of reactions (the complement cascade), which ultimately results in lysis (breakdown) of the target cells if the HLA molecules on their surface matches the variant targeted by the antibody. By quantifying the amount of cell lysis using staining and microscopic inspection, it is possible to determine the individual's HLA type (Blasczyk, 2003; Gautreaux, 2017; Howell et al., 2010).

3.2. PCR-based technologies

Later, DNA-based HLA typing methods were developed. Compared to the now obsolete serological assays, these methods allowed for faster and more cost-effective HLA typing (Gautreaux, 2017).

The most common DNA-based genotyping methods rely on the Polymerase Chain Reaction (PCR), a laboratory technique that is commonly used to amplify specific DNA sequences. To perform PCR, the template DNA is mixed with primers (short single stranded DNA sequences that are complementary to a part of the target sequence), nucleotides, and a thermostable DNA polymerase (e.g., *Taq* polymerase). Subsequently, the DNA is amplified in a process that involves multiple cycles of heating and cooling. First, the DNA is heated to separate the two strands of the double helix (denaturation). Secondly, the mixture is cooled down to allow binding of the primers to the template (annealing). Finally, the temperature is raised again to allow the DNA polymerase to synthesize new DNA strands, starting from the primer sequences and extending in the 5' to 3' direction (extension). This three-step procedure is repeated several times to generate multiple copies of the target DNA sequence (Garibyan & Avashia, 2013; Su et al., 1996) (Figure 3).

One method to employ PCR for HLA genotyping is called PCR with sequence-specific primers (PCR-SSP) (Bunce & Passey, 2013; Shyamala & Ames, 1989). In this technique PCR is performed with multiple allele specific primers that are complementary to sequences around polymorphic sites. If the target DNA matches exactly with the primers, the primers anneal efficiently, and PCR amplification can proceed. Otherwise, when there is a mismatch between the primer and the template DNA, the primer cannot be extended by the Taq DNA polymerase. Finally gel electrophoresis is used to evaluate whether amplification has occurred and hence the corresponding allele is present (Bontadini, 2012; Sibinga et al., 2000).

Another method for molecular HLA typing is *PCR with sequence-specific oligonucleotide probes (PCR-SSOP).* Unlike PCR-SSP, which uses allele specific primers to initiate DNA amplification, PCR-SSOP starts by amplifying the entire gene using gene specific primers. Subsequently, labelled allele specific oligonucleotide probes are allowed to hybridize with the target DNA. The presence or absence of these probes can then be used to determine the HLA allele (Gautreaux, 2017; Sibinga et al., 2000).

Finally, a third technique (PCR-SBT) is based on Sanger sequencing. As is the case for PCR-SSOP, the DNA in the region of interest is first amplified using PCR with gene-specific primers. Then the

amplified DNA is sequenced using Sanger sequencing, and the resulting sequence is compared to an HLA allele sequence database to determine which HLA allele the subject (Gautreaux, 2017).

Each of the discussed PCR-based techniques has its own set of advantages and limitations. Both PCR-SSP and PCR-SSOP are incapable of identifying new alleles. In contrast, PCR-SBT involves directly sequencing the amplified DNA and can, in principle, derive the entire nucleotide sequence of the HLA genes. As a consequence, it is the only PCR-based method that allows to discover novel alleles.

However, chromosomal phase (cis/trans) ambiguities are an important shortcoming of both PCR-SBT and PCR-SSOP (S. D. Adams et al., 2004; Segawa et al., 2017). This refers to the inability of both methods to determine whether two polymorphic sequence motifs are located on the same (cis) or on different (trans) chromosomes, where both situations can correspond to different existing HLA alleles. PCR-SSP, on the other hand is capable of distinguishing these two situations (Erlich & Henry Erlich, 2012).

A common workaround to deal with phase ambiguities in PCR-SBT is to perform first a low-resolution typing using either PCR-SSP or PCR-SSO and to subsequently use suitable primers to perform amplification and Sanger sequencing for one chromosome at a time (Erlich & Henry Erlich, 2012).

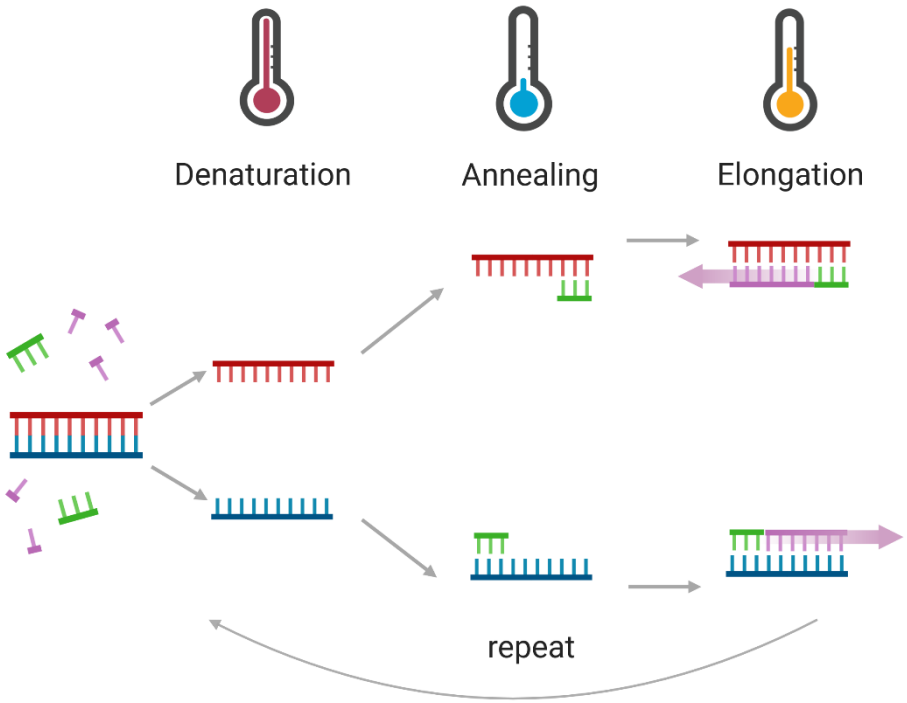

Figure 3. Principle of the Polymerase Chain Reaction. PCR is a multi-cycle process that consists of three steps: denaturation, annealing and elongation that are each performed at a different temperature. Created with BioRender.com

3.3. NGS based HLA typing on targeted sequencing data

Later, NGS based HLA typing methods were developed, which have a higher throughput and suffer less from chromosome phasing issues than the traditional PCR based typing methods (Anderson & Schrijver, 2010; Lind et al., 2010; Shiina et al., 2018). This first generation of NGS based methods involved targeted capturing of the HLA region through PCR or probe-based DNA capturing methods

prior to sequencing (Bentley et al., 2009; Gabriel et al., 2009; Hosomichi et al., 2015; Wittig et al., 2018).

In the meantime, these methods have been widely adopted by clinical laboratories to provide HLA typing for hematopoietic cell transplantation donors and patients (Edgerly & Weimer, 2018).

3.4. Need for HLA genotyping workflows on general purpose NGS data

A limitation of the first generation NGS-based genotyping methods is that they required costly library preparation protocols (Lange et al., 2014; Warren et al., 2012). These methods do not allow deriving HLA genotypes directly from typical WES, WGS, or RNA-Seq workflows that do not involve isolation and enrichment of the MHC region.

Subsequent developments in HLA genotyping introduced algorithms that bypassed the need for these steps. This evolution has an incredible potential for numerous research disciplines as it permits predicting the HLA genotypes directly from the many publicly available datasets that were constructed using these technologies.

3.5. How to predict the HLA genotypes from WES, WGS, or RNA-Seq data?

In general, these algorithms consist of the following steps. First reads (or assembled contigs) are aligned to a panel of reference HLA allele sequences (originating from the IPD-IMGT/HLA database). For this task the algorithms rely on various existing alignment tools (column *Alignment Method,* Table S1).

Then, for each HLA gene, an allele pair is selected by optimizing a certain score function. Algorithms differ in how this optimization problem is modelled exactly (e.g., Bayesian interference, Integer Linear Programming, or as a graph problem) and which variables are considered in the score function (column *Score function*, Table S1). Commonly used variables of the score function are: the consistency of the alignment of the input reads to the reference panel, base quality scores (column *PHRED score used*) and whether they use prior population frequencies (column *prior population frequencies*, Table S1).

This score function can either be defined in function of allele pairs for multiple genes at once (e.g., *Optitype*), in function of an allele pair per gene (e.g., *HLA*LA*) or in function of individual alleles separately per chromosome (e.g., *Polysolver*) (column *Jointly optimized for allele pair*, Table S1). In the first two cases, heterozygosity / homozygosity are implicitly modelled in the optimization problem. In the latter case, a separate step is typically needed to determine whether the call is homozygous or not based on numeric thresholds (e.g., for *HLA-VBSeq*).

3.6. Algorithmic description of 13 tools

The HLA typing algorithms for (general purpose) NGS data mainly differ in how they map sequencing reads to a panel of reference HLA allele sequences and the strategy they use to subsequently score candidate alleles (Bai et al., 2018; Klasberg et al., 2019) (Table S1). Below follows a description of the 13 algorithms that were selected for the benchmark (results section 1).

HLA*LA starts with a "linear alignment" step where the input reads are aligned using *BWA-MEM* to a modified reference genome (composed of GRCh38, the MHC haplotypes and IMGT/HLA genomic

sequences). These alignments are then projected onto a population reference graph (PRG) and further optimized. Finally, the HLA genotype is inferred by maximizing a likelihood function (Dilthey et al., 2019).

Kourami is a graph-guided assembly tool. As a first step, reads are extracted from the BAM file and realigned to a reference panel (using *BWA-MEM*). Similar as in *HLA*LA* these "linear alignments" are then projected onto a partial-order graph (POG) representing the known HLA alleles. During this projection step, the graph is modified to incorporate substitutions and indels that were identified during the alignment. This allows *Kourami* (as the only tool included in our benchmark) to discover new HLA alleles. Edges of the graph are weighted according to the read counts. Finally, the goal is to identify the best pair of alleles (paths in the graph) that maximizes the coverage and phasing support (H. Lee & Kingsford, 2018a, 2018b).

arcasHLA*:* First, pseudoalignment with *Kallisto* is performed to determine for each read which HLA transcripts it is compatible with. Based on this output, a built-in transcript quantification step is performed that aims to identify an attribution of reads to alleles which maximizes a likelihood function. This is performed using an iterative read re-allocation procedure (an expectation-maximization algorithm): at every step, reads are distributed to alleles with the highest abundance, while the alleles with the lowest abundances are removed from the list of candidates. When after convergence more than two alleles remain, the allele pair that explains the greatest proportion of reads is selected. Finally, either a homozygous or heterozygous call is produced based on the non-shared read counts between the top two alleles (Orenbuch, Filip, & Rabadan, 2020; Orenbuch, Filip, Comito, et al., 2020).

HLA-HD*:* First, reads are aligned to a database of reference HLA exon and intron sequences using *Bowtie2*. Reads are then assigned to candidate exons or introns based on certain filter criteria. A score is calculated per allele pair based on the number of reads that map to the corresponding sequences, considering the length of the overlap between input reads and exon sequences. The algorithm first calculates the score only based on exons in the peptide binding region and later extends to other exons. The allele pair that yields the maximum score is finally selected (Kawaguchi et al., 2017).

PHLAT starts by aligning the input reads to the human reference genome extended with various HLA allele reference sequences using *Bowtie2*. Following the alignment, candidate alleles are pre-selected using multiple filtering steps based on mapped read counts. Pairs of candidate alleles are then scored using a Bayesian likelihood model which considers the sequence consistency at Single Nucleotide Polymorphism (SNP) sites and phase consistency across adjacent SNP sites. The allele pair that best explains the observed data is selected (Bai et al., 2014, 2018).

Polysolver*:* Reads are first mapped to the reference allele sequences using *Novoalign*. *Polysolver* relies on a Bayesian classifier to select the alleles that most likely explain the observed reads. Its model incorporates the base qualities of aligned reads, the observed insert sizes and (optionally) ethnicity dependent prior probabilities. For each gene, the calls for both alleles are determined in separate steps. Once the first allele is identified, the probabilities are updated based on that information and the second allele for that locus is identified (Shukla et al., 2015).

***HLA-VBSeq**:* Reads are aligned to the reference panel using *BWA-MEM*. These alignments are then further optimized using a Bayesian framework. HLA types are subsequently called based on the expected number of reads that are mapped to each allele. A threshold on the depth of coverage is used to filter out candidate alleles. Either a heterozygous or homozygous call is outputted based on the depth of coverage of the top two ranked alleles (Nariai et al., 2015).

seq2HLA uses a two-stage approach to accomplish HLA genotyping at 4-digit resolution: genotypes are first called at 2-digit resolution before further refining them in a separate round. The algorithm starts by aligning RNA-seq reads to the reference HLA panel with *Bowtie* (v1). Then, for each HLA gene the allele group with the greatest number of reads is determined and considered to be the winner of the first round. The second allele group for that gene is then determined by removing all reads associated with the winner of the first round and repeating the previous step. Either a heterozygous or homozygous call is outputted depending on the ratio between the number of reads mapping to the winner of the second round and the median number of reads mapped to alleles in the first round (Boegel et al., 2012). Finally, the allele calls at 2-digit resolution are further refined to the 4-digit resolution by considering the amount of reads aligned to alleles within the winning group at 2-digit resolution (Boegel et al., 2014).

***Optitype**:* First, reads are mapped to a panel of reference HLA allele sequences, limited to the exons encoding the peptide binding region of MHC-I (exon 2 and 3) and the flanking introns. Partial HLA sequencing information in the reference panel was reconstructed using phylogenetic information. *Optitype* then models the scoring of HLA alleles as an Integer Linear Programming (ILP) problem that aims to find the set of MHC-I allele pairs (for all major and minor MHC-I genes) that simultaneously explain the input data the best (Szolek et al., 2014).

xHLA maps the sequencing reads to the HLA reference sequences using the *Diamond* aligner and subsequently identifies candidate alleles by applying *Optitype*'s ILP strategy using only the exons encoding the peptide binding region (for both MHC-I and MHC-II genes). This set of candidate alleles is subsequently extended to sets of alleles that explain the alignments nearly as well and then further refined using an iterative procedure. This strategy allows HLA genotyping at a finer resolution than *Optitype*. When finally two alleles remain, an additional check is performed to determine whether a homozygous or heterozygous call should be outputted by comparing the amount of reads supporting both alleles (C. Xie et al., 2017).

***HLAscan**:* After aligning reads to the reference sequences (*BWA-MEM*), *HLAscan* selects candidate alleles based on a score function that represents the distribution of aligned reads in the region of interest. Alleles are discarded based on the number of consecutive positions in the mapped HLA sequence with no read aligned to it. Out of the remaining alleles, the resulting allele pairs for each gene are determined based on which alleles have the highest read count and a check for heterozygosity (Ka et al., 2017).

***HLAforest**:* Reads are mapped to the HLA reference allele sequences using *Bowtie* (v1). For each read a tree is constructed that represents all possible mappings for that read. The first level of the tree represents the different HLA genes, each subsequent level represents a different field of the HLA nomenclature. Sum of mismatch qualities (SMMQs, based on the PHRED qualities at mismatches

between the read and reference sequence) are assigned to the leaf nodes of the tree. This score is than propagated upwards the tree where the probability value assigned to a parent node is the maximum probability of its children. These probability values are then converted into *weights*, which are distributed downwards through the tree. The final allele pair is then selected via an iterative tree pruning algorithm (Kim & Pourmand, 2013).

HLAminer supports both a *de novo* assembly-based (HPTASR) and an alignment-based (HPRA) pipeline. In the de novo assembly-based pipeline (not evaluated in this benchmark) reads are first assembled into larger contigs and are subsequently aligned to the panel of reference HLA sequences (using *BLAST*). In the alignment-based pipeline the reads are directly aligned to the reference allele sequences using *BWA*. Alleles are scored based on the contig length, depth of coverage and similarity to reference sequences of all contigs that align to it (Warren et al., 2012). *HLAminer* does not incorporate a method to impute heterozygosity / homozygosity.

4. Applications of HLA genotyping in cancer research

Being responsible for presenting neoantigens to immune cells, MHC molecules are crucial in the immune system's ability to recognize cancer cells (section 1.4). Additionally, they are highly polymorphic and which HLA alleles an individual has determines the repertoire of peptides that can be efficiently presented to the immune system (Nielsen et al., 2007). As such, someone's HLA genotype might influence cancer susceptibility and response to treatment (Chowell et al., 2018). In the previous chapter, we have discussed which HLA genotyping methods are available. Here, we will discuss how these tools are currently being applied in cancer research and give potential future directions for research in this field.

4.1. The immune system: a double-edged sword in cancer

The immune system has a complex and multifaceted role in cancer. On the one hand, our immune system has a remarkably effective ability to recognize and eradicate tumours. On the other hand, an inflammatory environment can also stimulate tumour growth. Both characteristics are considered to be *hallmarks of cancer* (Hanahan, 2022; Hanahan & Weinberg, 2011).

4.2. Immune evasion

Several mechanisms that allow tumours to escape immunogenic destruction involve genetic alterations in components of the antigen processing and presentation machinery (APPM). These alterations can result in either a defective antigen presentation, or disruptions of the pathways that produce peptides that efficiently bind to MHC-I (Jhunjhunwala et al., 2021). Additionally, cancer cells may evade destruction by Natural Killer (NK) cells through (MHC-I independent) upregulation of *HLA-E* or *HLA-G* (Borst et al., 2020; de Kruijf et al., 2010).

Three Es of immunotherapy: elimination, equilibrium and escape

Cancerous cells gradually acquire these traits in a process called *immunoediting*, which can be separated into three phases: elimination, equilibrium and escape.

In the first (elimination) phase immunogenic tumour cells are killed by the immune system. If the immune system is able to eliminate the tumour, the immunoediting process ceases and we do not advance to the next phase. Some tumour cells might survive and progress to the equilibrium phase.

During the equilibrium phase tumours are genetically unstable and under a constant pressure of the immune system that is enough to contain, but not fully eradicate the tumour. Eventually, due to natural selection, this phase results in a new population of tumour clones that is less immunogenic than its parent population.

In the final (escape) phase the tumour variants that were selected in the equilibrium phase have acquired the (epi)genetic alterations that allow them to evade immunogenic destruction. Now, tumour growth can occur in an immunologically intact environment. This allows the tumours to expand and become apparent (Dunn et al., 2004; Pasinetti et al., 2016).

Neoantigen depletion

It has been demonstrated on a mouse model that Darwinian selection in favour of tumour cells that do not express tumour-specific antigens is a possible immune evasion mechanism (Dupage et al.,

2012). According to this model, clones expressing neoantigens are expected to be eliminated over time.

For quite some time, it was believed that this process was also detected in human genomic data. Studies that seemed to confirm this quantified immunoselection using a metric that compares the observed to the expected neoantigen load (Rooney et al., 2015). However, to model the expected neoantigen load the authors did not take into account the fact that transitions between particular triplets of nucleotides are more likely to occur than others (mutational signatures) and that the presence of these triplets intrinsically influences the HLA binding affinity. After correcting for this spurious correlation, there is currently no evidence for neoantigen depletion (van den Eynden et al., 2019).

There are two possible explanations for the lack of evidence for neoantigen depletion. First, the neoantigen load in these models was calculated based on computational models that predict the binding affinity of peptides to MHC-I given the HLA genotype. It is known that only a very small fraction of the peptides that are predicted to bind to MHC-I are actually immunogenic (Wells et al., 2020). Therefore, it is possible that neoantigen depletion actually occurs, but the accuracy of these predictions is insufficient to detect it. Secondly, it is also possible that tumours develop other, more effective immune evasion mechanisms early in their development, making it no longer disadvantageous for them to produce proteins that can strongly bind to MHC.

Below, we discuss other immune evasion mechanisms involving the HLA genes.

Defective antigen processing and presentation machinery

One important category of immune evasion mechanisms involves disruption of antigen processing and presentation. This can occur either through changes in the repertoire of peptides that bind to MHC or by hampering antigen presentation itself (Jhunjhunwala et al., 2021).

The deletion of the Endoplasmic reticulum aminopeptidase (ERAP) genes and TAP gene silencing are examples of defects that alter the HLA-binding peptide repertoire. ERAP is responsible for trimming peptides to the optimal length for presentation by MHC-I molecules. Even though germline variation in ERAP is linked to cancer predisposition, it is rarely mutated in cancer (Compagnone et al., 2019; Jhunjhunwala et al., 2021; Stratikos et al., 2014). The role of TAP alterations as an immune evasion mechanism is controversial as well. TAP assists in transporting peptides from the cytosol into the endoplasmic reticulum, where they can be loaded onto MHC-I molecules. While a loss of TAP might lead to a reduction in MHC-I surface expression, other studies pointed out that it results in the presentation of cryptic antigens that increase the immunogenicity of the tumour (Garrido et al., 2019). Finally, the repertoire of peptides presented by MHC-I can also change due to somatic mutations in the peptide binding region of the HLA genes, which can impact the peptide-MHC binding affinity (Shukla et al., 2015).

Immune evasion mechanisms that relate to antigen presentation are the loss of MHC-I and B2M. Complete copy number loss of B2M results in the absence of MHC-I molecules at the cell surface. The loss of one of the HLA genes, on the other hand, results in a reduction in the diversity of peptides that can be presented to the immune system.

Differential expression of non-classical MHC-I genes

Non-classical HLA genes can also be involved in immune evasion mechanisms. Normally, the loss of MHC-I expression should result in Natural Killer (NK) cell-mediated killing of tumour cells, but cancer cells may evade this by upregulation of *HLA-E* or expression of *HLA-G* (Borst et al., 2020; de Kruijf et al., 2010). Next to its role in inhibiting NK cells, *HLA-G* inhibits the proper functioning of several other immune cell types including B cells, T cells and dendritic cells (Krijgsman et al., 2020; Zhuang et al., 2021).

Other immune evasion mechanisms: the PD-1 and CTLA-4 checkpoints

The discussion above was focussed on immune evasion mechanisms that relate to pathways where the HLA genes are directly involved. However, several other immune evasion mechanisms have been described in literature. Two important pathways that are currently therapeutically targeted are the CTLA-4 and PD-1 immune checkpoints. PD-1 is primarily expressed on the surface of activated T cells. Upon binding to its ligands PD-L1 and PD-L2, usually expressed by antigen presenting cells, it can exert inhibitory functions on T cells (Seidel et al., 2018). CTLA-4 is constitutively expressed by Treg cells but can also be expressed by other immune cells. It can inhibit the activation of T cells by binding to CD80 (also called B7-1) and CD86 (also called B7-2) on the surface of antigen presenting cells through competitive inhibition of CD28 on CD4+ or CD8+ T cells (Seidel et al., 2018).

In normal situations both pathways regulate the immune response and prevent the immune system attacking normal tissue. However, tumours can also exploit these pathways to evade immune response by expressing PD-L1 or CTLA-4 themselves (Contardi et al., 2005; Gangaev et al., 2021; Kern & Panis, 2021).

4.3. HLA genotype dependent cancer susceptibilities

Given that HLA alleles are a major determinant of peptide-MHC affinities and that MHC molecules play a key role in anti-cancer immune response, it is tempting to hypothesize that HLA genotypes may influence cancer susceptibility.

HLA genotype correlates to risk to develop pathogen induced cancer

Associations between HLA genotypes and cancer risk have been established for pathogen induced cancers. For example, certain HLA alleles have been linked to an increased risk of developing cervical cancer (caused by the Human Papillomavirus (HPV) infection), and homozygosity for the *HLA-DQB1* gene has been linked to an increased risk of hepatitis B induced hepatocellular carcinoma (Z. Liu et al., 2021; Safaeian et al., 2014).

For other cancer types, there is currently limited evidence that the HLA genotype is associated with cancer risk and further research is required. To study this further, the INDICATE initiative was founded which focusses on the presumed role of HLA genotypes in modulating cancer risk in Lynch syndrome carriers (Ahadova et al., 2022).

Genotype dependent restriction of the mutational landscape

Related to HLA genotype dependent cancer susceptibilities, two Cell papers suggested that cancer hotspot mutations are under selection depending on the underlying MHC genotype and that the mutations that are observed in a cancer patient are the mutations that are poorly presented to the immune system (Marty et al., 2017, 2018). The reasoning behind this is the following. In a patient

where a particular mutation leads to neoantigens that strongly bind to MHC, cells that have this mutation are eradicated quickly by the immune system and do not get the chance to proliferate. As a result, this mutation does not lead to cancer in that patient. Conversely, the same mutated peptides can be poorly recognized by the immune system of another patient with a different HLA genotype. In that patient, cells that with that mutation can divide nearly undisturbed and become dominant over time.

These findings seemed to contradict the conclusion of our own study where we concluded that there is currently no evidence for neoantigen depletion (van den Eynden et al., 2019). However, using a simulation we demonstrated in our lab that the observed signal was again caused by spurious correlations (due to intrinsic biochemical properties of driver mutations) rather than genotype specific immune selection (Claeys et al., 2021).

4.4. An emerging role for MHC-II in cancer immunity

As the importance of MHC-II in tumour recognition is increasingly appreciated (Alspach et al., 2019), (also see section 1.4), there are still unresolved questions about its role in cancer. While there is currently no evidence for MHC-I restricted neoantigen depletion, it remains uncertain whether this also holds true for MHC-II.

A necessary step to further investigate the role of MHC-II restricted neoantigens in cancer, is to determine which peptides can be presented to the immune system via the MHC-II complex. This requires an estimation of the binding affinity of the peptide to MHC-II, which in turn depends on the MHC-II genotype of the patient. The latter being the main topic of this thesis.

4.5. Tumour promoting immune effects

Cancer cells are not isolated entities but are embedded in a tumour microenvironment (TME) consisting of blood vessels, fibroblasts, the extracellular matrix, and immune cells. The cells of the TME and cancer cells interact with each other in a dynamic manner (Greten & Grivennikov, 2019). As such, tumours might evolve to secrete cytokines and other inflammatory mediators that create a microenvironment that is favourable for cancer cell proliferation and survival (Gómez-Valenzuela et al., 2021). Here two types of inflammatory environments are often distinguished. Whereas an acute inflammatory environment leads to anti-tumour immune responses, a chronic inflammatory environment facilitates tumour progression (Zhao et al., 2021).

This double role of the immune system depending on the composition of the tumour microenvironment might also give a different, tumour promoting role to MHC-II restricted neoantigens. As initiators of CD4+ T cell responses, MHC-II molecules also interact with Treg, Th2 and Th17 cells (Corthay, 2009; Sun et al., 2017). Therefore, MHC-II presentable neoantigens in a chronic inflammatory environment do not necessarily lead to an effective immune response against cancer cells but might under some circumstances even have a tumour promoting effect (Sun et al., 2017). Further research is needed to fully understand the intriguing role of MHC-II in cancer.

4.6. Identifying biomarkers for immune checkpoint blockade therapies

The last decade has seen the rise of immune checkpoint blockade therapies, a type of immunotherapy that stimulates the immune system to attack the tumour by inhibiting immune checkpoints (section 4.2). Multiple studies have demonstrated the success of these therapies in cancer types with a poor prognosis, such as advanced melanoma, non-small cell lung cancer, and metastatic renal cell carcinoma (Larkin et al., 2019; Motzer et al., 2019; Paz-Ares et al., 2021; Reck et al., 2016). However, responses to immunotherapy remain difficult to predict, with tumour mutational burden as one of the few available biomarkers (Cristescu et al., 2018; Havel et al., 2019).

Genotype dependent differences in HLA binding affinity could also lead to differential responses to immunotherapy, as illustrated by the association that has been described between MHC-I genotypes (e.g., *HLA-B62*) and survival in immune checkpoint blockade (ICB)-treated advanced melanoma patients (Chowell et al., 2018). It is currently unclear whether MHC-II genotypes also determine responses to immunotherapy, which is the topic of ongoing research in our lab.

Chapter 2
Aims

The research questions discussed in the previous chapter all require knowledge about the HLA genotypes. PCR based methods are currently the gold standard for this genotyping but datasets with this type of HLA calls are rarely available (Bauer et al., 2018; Orenbuch, Filip, Comito, et al., 2020; Szolek et al., 2014). As mentioned earlier, HLA genotyping can also be performed on Next Generation Sequencing (NGS) data, bypassing the need for separate wet lab experiments (section 3.5). A plethora of tools has been developed for this task. *Polysolver* and *Optitype* are often recommended as the best performing tools for MHC-I genotyping (Matey-Hernandez et al., 2018). For MHC-II genotyping there is currently no consensus about the best method. Several benchmarks have been performed previously (Bauer et al., 2018; Chen et al., 2021; Kiyotani et al., 2016; M. Lee et al., 2021; X. Li et al., 2021; P. Liu et al., 2021; Matey-Hernandez et al., 2018; Yi et al., 2021; Yu et al., 2021), but these were either not applied to MHC class II or did not include some recently published tools.

In this study, we compiled a list of 13 tools that predict HLA genotypes from NGS data and benchmarked their performance on both the 1000 genomes dataset and on an independent cell line dataset (Abaan et al., 2013). Subsequently we assessed their performance on 9162 WES and 9761 RNA sequencing files from The Cancer Genome Atlas (TCGA) by comparing the predicted allele frequencies with reference population allele frequencies. Based on these findings, we give recommendations on which tool to use for a given data type and how the outputs of multiple tools can be combined into a consensus prediction.

Chapter 3
Results

1. Selection of 13 HLA genotyping tools with variable computational resource requirements

We identified 22 available HLA genotyping tools from literature (Table 1). Thirteen tools that were free for academic use, applicable on Whole Exome Sequencing (WES), Whole Genome Sequencing (WGS) or RNA-Seq data and ran on Ubuntu 20.04 were included in this study: *arcasHLA*, *HLA-HD*, *HLA-VBSeq*, *HLA*LA*, *HLAforest*, *HLAminer*, *HLAscan*, *Kourami*, *Optitype*, *PHLAT*, *Polysolver*, *seq2HLA* and *xHLA* (Table S2). All 13 tools can make allele predictions for the three MHC class I genes (*HLA-A*, *HLA-B* and *HLA-C*) and 9 tools support additional calling of the MHC class II genes *HLA-DPA1*, *HLA-DPB1*, *HLA-DQA1*, *HLA-DQB1* and *HLA-DRB1*. Two methods support only a subset of the MHC class II genes: *xHLA* does not support calling *HLA-DPA1* and *HLA-DQA1*, while *PHLAT* does not support *HLA-DPA1* and *HLA-DPB1*. The tools also differ in which data types they support: 6 of them require WES data, 3 tools require RNA data and the 4 remaining tools support both data types (Table 1).

Firstly, the computing time and memory usage of the thirteen selected tools were measured on a random subset of 10 WES and 10 RNA sequencing files from the TCGA project (Figure 4).

Among the 10 WES-supporting methods *Optitype* (median 2.48 hours) and *HLA*LA* (median 1.84 hours) require the largest computing time. The remaining WES tools take less than 1 hour per file, with *HLAminer*, *Kourami* and *PHLAT* being the fastest (97s, 225s and 253s respectively). Apart from being computationally intensive, *HLA*LA* is also the most memory demanding WES tool (median 36.3 GiB per file). Other WES tools with a median memory consumption higher than 5 GiB are *xHLA* (median 22.9 GiB), *Kourami* (median 9.3 GiB) and *HLA-HD* (median 6.7 GiB). The relatively low memory usage of *Polysolver* makes it feasible to compensate for its long running time by processing multiple samples in parallel.

Among the 7 RNA-supporting methods, *HLA-HD* has the longest computing time per sample (median 15.0 hours). At the other end of the spectrum, the sole pseudoalignment-based tool *arcasHLA* takes only 38s per file. The most memory intensive tool is *HLA-HD* (median memory peaks of 103.1 GiB), followed by *Optitype* (median 34.1 GiB). The other RNA tools have a memory usage lower than 10 GiB. Remarkably, *HLAminer*, *PHLAT* and *HLA-HD*, which are compatible with both WES and RNA data take a longer time on RNA data (median computing time per sample is 29.4, 8.9, 6.8 times longer for *HLA-HD*, *PHLAT* and *HLAminer* respectively).

		Data type		Input filetype		HLA loci								Version
		WES	RNA	BAM	FASTQ	A	B	C	DPA1	DPB1	DQA1	DQB1	DRB1	
Included	arcasHLA (Orenbuch, et al., 2020)	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	0.2.0
	HLA-HD (Kawaguchi et al., 2017)	✓	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	1.3.0
	HLA-VBSeq (Wang et al., 2019)	✓*	✗	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓	2
	HLA*LA (Dilthey et al., 2019)	✓	✗	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓	1.0.1
	HLAforest (Kim & Pourmand, 2013)	✗	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	1
	HLAminer (Warren et al., 2012)	✓	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	1.4
	HLAscan (Ka et al., 2017)	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	2.1.4
	Kourami (H. Lee & Kingsford, 2018b)	✓*	✗	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓	0.9.6
	Optitype (Szolek et al., 2014)	✓	✓	✗	✓	✓	✓	✓	✗	✗	✗	✗	✗	1.3.5
	PHLAT (Bai et al., 2014)	✓	✓	✗	✓	✓	✓	✓	✗	✗	✓	✓	✓	1.1
	Polysolver (Rooney et al., 2015)	✓	✗	✓	✗	✓	✓	✓	✗	✗	✗	✗	✗	4
	seq2HLA (Boegel et al., 2012)	✗	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	2.3
	xHLA (C. Xie et al., 2017)	✓	✗	✓	✗	✓	✓	✓	✗	✓	✗	✓	✓	0.0.0
Not included	ALPHLARD-NT (Hayashi et al., 2019)	✓	✗	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓	
	ATHLATES (C. Liu et al., 2013)	✓	✗	✗	✓	✓	✓	✓	✗	✗	✗	✓	✓	
	HLAProfiler (Buchkovich et al., 2017)	✗	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	
	HLAreporter (Huang et al., 2015)	✓	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	
	HLAssign (Wittig et al., 2015)	✓	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	
	OncoHLA (Sverchkova et al., 2019)	✓	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	
	PolyPheMe (Abi-Rached et al., 2018)	✓	✗	✗	✓	✓	✓	✓	✗	✗	✗	✓	✓	
	SNP2HLA (Jia et al., 2013)	✗	✗	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	
	SOAP-HLA (Cao et al., 2013)	✓	✗	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓	

Table 1. Overview of evaluated tools for HLA genotyping. Checkmarks and crosses indicate which NGS methods (WES and/or RNA-Seq) and input file types (FASTQ and/or BAM) are supported and for which genes predictions can be made. The tools in the upper part of the table are benchmarked in this study. Tools in the lower part of the table did not fulfil our inclusion criteria and were not further considered. * Works preferentially with WGS instead of WES data.

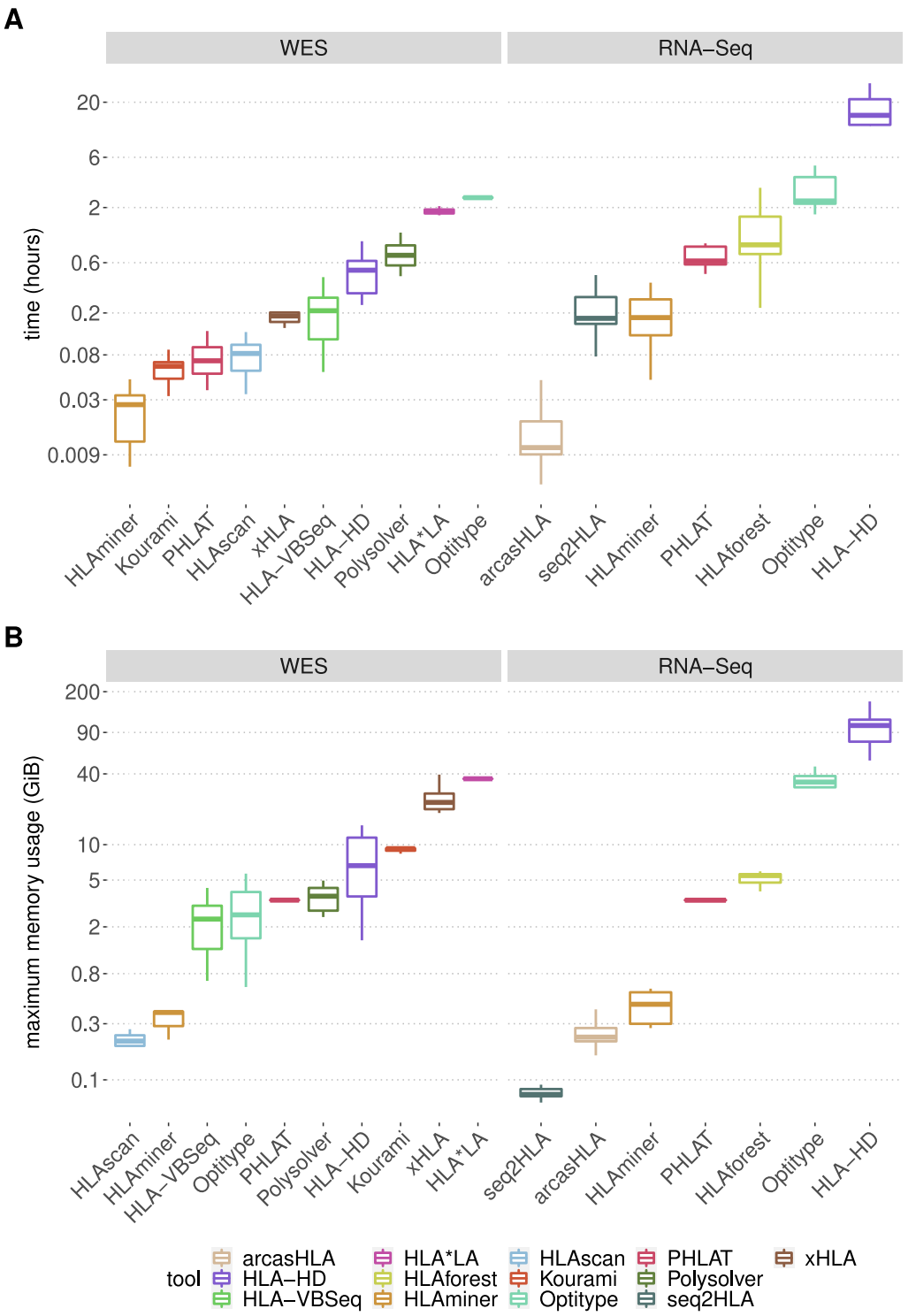

Figure 4. Computational resource consumption of the 13 selected tools. (**A-B**) Boxplots compare the resources needed by the different tools to analyse one sequencing file on a system with a single CPU core. Each tool was applied on WES and/or RNA sequencing files *(n=10)*, as indicated at the top of the figure. Different tools are represented with a different colour of the boxplot, as indicated in the legend on the right. The y-axes are displayed on a logarithmic scale. (**A**) Time consumption per sample. (**B**) Maximal memory consumption per sample.

2. HLA*LA and HLA-HD are the best performing MHC class II genotyping tools on WES data

The 10 selected algorithms that are compatible with WES data were benchmarked using data from the *1000 Genomes project* (Zheng-Bradley et al., 2017) (average HLA gene read depth = 40x +-/ 16.7). Predictions were made for *HLA-A* (n = 1012), *HLA-B* (n = 1011), *HLA-C* (n = 1010), *HLA-DQB1* (n = 1008), *HLA-DRB1* (n = 1000) and *HLA-DQA1* (n = 68) (Figure 5). *HLA-DPA1* and *HLA-DPB1* were not benchmarked due to the lack of available gold standard calls. For MHC-I genes (*HLA-A*, *HLA-B*, *HLA-C*), the best accuracy was obtained with *Optitype* (98.0%), followed by *Polysolver* and *HLA*LA* (94.9% and 94.4% respectively). For MHC-II genes (*HLA-DQA1*, *HLA-DQB1* and *HLA-DRB1*), the best allele predictions were made using *HLA-HD* and *HLA*LA* (96.2% and 95.7% accuracy respectively). These were the only two methods to reach an accuracy of 90% on all tested MHC-II genes. *HLAscan* (74.2%), *HLA-VBSeq* (60.2%) and *HLAminer* (53.8%) performed considerably worse than the other tools.

We observed large variabilities in calling accuracies between MHC class II genes (Figure 5). Overall, *HLA-DQB1* was the hardest MHC-II gene to call. Except for *PHLAT*, all tools obtained their worst MHC-II call accuracy on this gene. *HLA-DQA1*, on the other hand, was the gene with the highest calling accuracy for all tools that support it, except for *HLAminer* and *Kourami*.

Incorrect calls are either caused by wrong allele calls or a failure to make an allele call. *HLA-VBSeq* and *HLAminer* had both a high rate of incorrect and failed calls (Figures S1-S2). When *HLAscan* or *Kourami* were able to make a call, their predictions were mostly reliable (Figure S1), but these tools regularly produced no output at all (Figure S2). Miscalled samples had a significantly lower average read depth in the HLA genes than correctly called samples for most tools (Figure S3). Notably, large differences in coverage sensitivity were observed between the different tools, with *Kourami* and *HLA-VBSeq* being the most sensitive and *Optitype* being the least affected (Figure S4). An *in silico* analysis that simulated the effect of lowering coverage (to 50%, 10%, 5% and 1%) for the best performing tools suggested that the minimal average read depth to get 90% accuracy is 12.2x and 17.4x for MHC-I with Optitype and MHC-II with HLA-HD respectively (Figure S5).

Subsequently, we performed an independent benchmark using the smaller NCI-60 cell line dataset (n=58, average HLA gene read depth = 37x +/- 25.8), which largely confirmed our results (Figure S6). Additionally, this analysis indicated that the best performing MHC class II supporting tools also performed well on *HLA-DPB1*.

3. HLA-HD, PHLAT and arcasHLA are the best performing MHC class II genotyping tools on RNA data

We then evaluated the 7 selected methods that support HLA calling on RNA sequencing data from the *1000 genomes project* (Lappalainen et al., 2013) (median average HLA gene read depth = 3129x +-/ 1227) . Predictions were made for *HLA-A* (n = 373), *HLA-B* (n = 372), *HLA-C* (n = 372), *HLA-DQB1* (n = 371), *HLA-DRB1* (n = 362) and *HLA-DQA1* (n = 53) (Figure 5).

A*rcasHLA* and *Optitype* had the best MHC-I allele predictions (99.4% and 99.2% accuracy, respectively), followed by *HLA-HD* (98.0%), *seq2HLA* (95.9%) and *PHLAT* (95.4%). Similar accuracies were found for MHC-II allele predictions, with *HLA-HD*, *PHLAT* and *arcasHLA* performing the best

(99.4%, 98.9% and 98.1%, respectively). Contrary to its good prediction of MHC class I alleles, *seq2HLA* has a lower accuracy for MHC class II (87.8%). RNA-based tools were generally less affected by coverage differences than DNA-based tools, which is likely related to the higher absolute coverage of RNA-Seq as compared to WES data (Figures S3-S5).The high MHC-I accuracies of *arcasHLA* and *Optitype* were confirmed on the independent NCI-60 dataset (91.8% and 90.0%, respectively; n=58, average HLA gene read depth = 578x +/- 837). The accuracy of *HLA-HD*, *PHLAT* and *seq2HLA* was worse on the cell lines than in the benchmark on the 1000 genomes data (86.6%, 83.3% and 82.3%, respectively). As MHC-II is generally not expressed in cell lines, this benchmark was not performed for those genes.

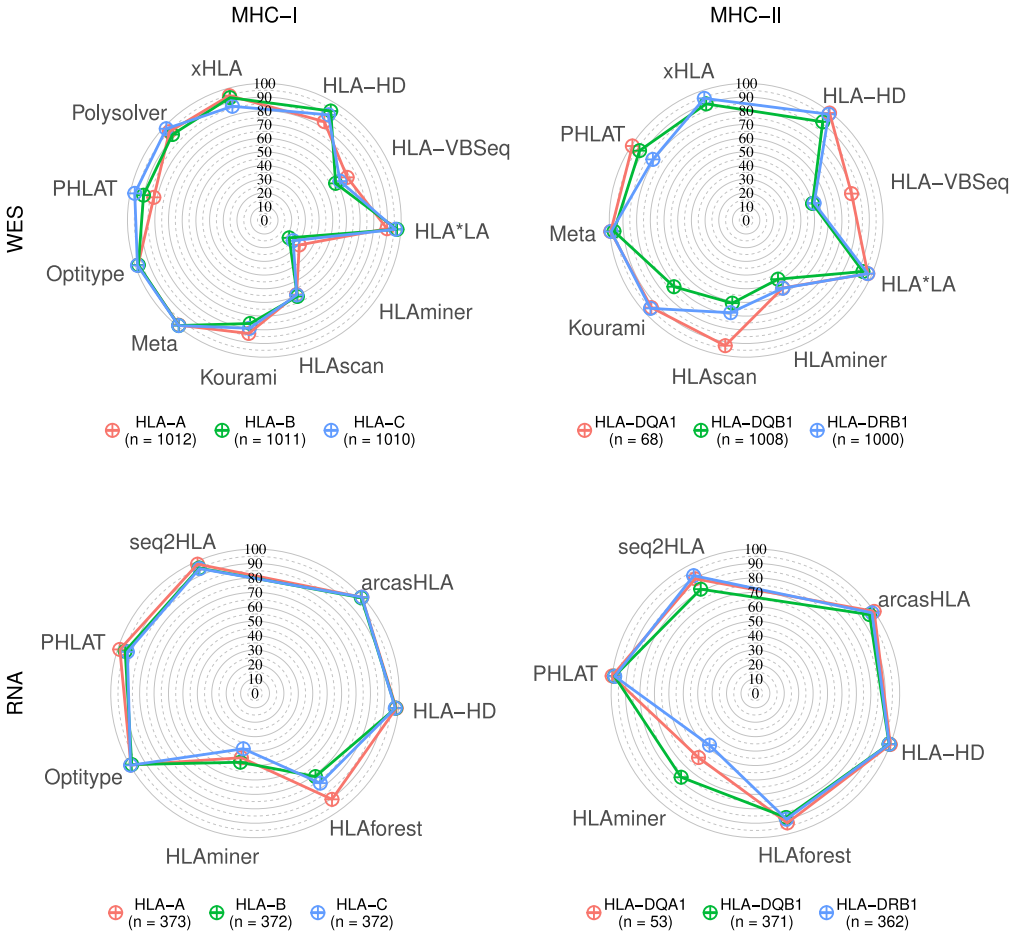

Figure 5. HLA allele prediction accuracies. Radar plots of HLA allele prediction accuracies on samples from the 1000 Genomes Project. Coloured lines represent different genes, as indicated in the legend below the plots. Corners of the radar plots correspond to the tools that were evaluated for that data type. The Meta tool corresponds to the 4-tool consensus metaclassifier.

4. Correlation and concordance analyses on large independent datasets confirm the benchmarking results

Being one of the few large sequencing datasets for which gold standard HLA genotypes for both MHC classes are available, many algorithms included in our benchmark were developed, optimized and validated using files from the *1000 genomes* project, introducing a potential bias. Additionally, no evaluation was possible for *HLA-DPA1* and *HLA-DPB1*, due to the lack of gold standard HLA calls. Therefore, we performed an indirect and independent evaluation on a large NGS dataset obtained

from TCGA (n=9162 with an average HLA read depth = 66x +/- 28.6 and n=9761 with an average HLA read depth = 3076x +/- 2775 for WES and RNA respectively).

We first compared the observed allele frequencies for each tool with the expected population frequencies. We calculated how often each of the alleles was predicted by a certain tool to obtain an observed allele frequency, stratifying for Caucasian American (n= 7935) and African American (n=938) ethnicities. By comparing these frequencies to the expected allele frequencies, as derived from *Allele Frequency Net* (Gonzalez-Galarza et al., 2020), strong significant correlations were found for the WES-based tools *HLA-HD (minimal Pearson's r* = 0.970; *P* = $1.5*10^{-5}$*)*, *HLA*LA (min. r* = 0.968; *P* = $7.6*10^{-5}$*)*, *Optitype (min. r* = 0.978; *P* = $5.5*10^{-108}$*)*, *Polysolver (min. r* = 0.976; *P* = $4.7*10^{-58}$*)* and *xHLA (min. r* = 0.978; *P* = $4.4*10^{-115}$*)* and for the RNA-based tools *Optitype (min. r* = 0.972; *P* = $6.2*10^{-47}$*)*, *arcasHLA (min. r* = 0.939; *P* = $1.2*10^{-19}$*)* and *PHLAT (min. r* = 0.937; *P* = $2.1*10^{-5}$*)*. The correlations were considerably worse for *HLA-VBSeq (worst r* = 0.867; *P* = $4.1*10^{-23}$*)*, *HLAminer (min. r* = 0.557; *P* = $1.2*10^{-8}$ and *r* = 0.593; *P* = $6*10^{-7}$*,* for WES and RNA respectively*)* and *HLAforest (minimal r* = 0.423; *P* = $1.8*10^{-3}$*)* than for the other tools (Figure 6). These findings largely confirm the results of the benchmark on the 1000 genomes data. Notably, among the well performing tools, *arcasHLA* had a worse correlation for *HLA-DRB1* in African Americans (*r* = 0.939; *P* = $1.2*10^{-19}$), which is mainly due to the discrepancy between the observed and predicted frequency of *HLA-DRB1*14:02* in this population (Figure S7).

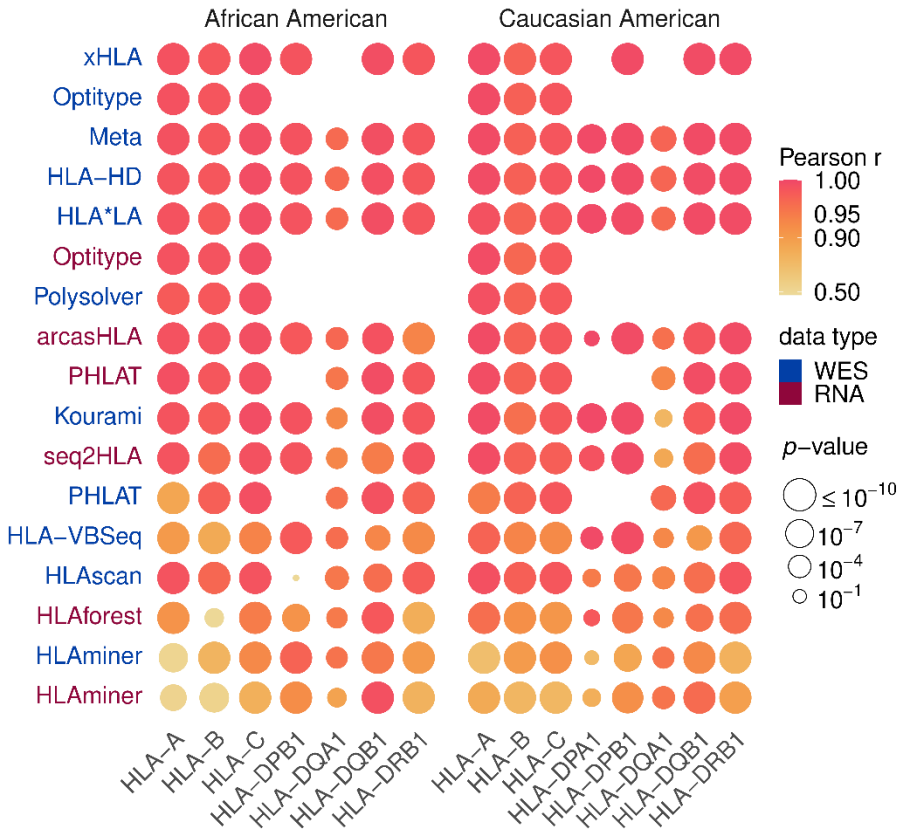

Figure 6. Correlations between observed and expected allele frequencies. Heatmap of correlations between observed allele frequencies and frequencies expected in an African American and in a Caucasian American population. Vertical axis indicates the tools, with different colours representing the data type (WES or RNA) on which the tool was applied. Rows were sorted according to the mean correlation of the tool. Size of the circles indicates the *P* value of the correlation test as indicated in legend. Absent circles indicate that the tool could not be evaluated on that gene.

We then calculated for each pair of tools how often their predictions are concordant (Figures S8-S11). Tools that performed poorly in the previous analyses (e.g., *HLAminer*, *HLA-VBSeq* and *HLAforest*) consistently have a low concordance with all other tools. In contrary, tools that scored high in the previous analyses (such as *Optitype, HLA*LA, arcasHLA* and *HLA-HD*) made predictions that are consistent with each other. Noteworthy, this is also the case for *HLA-DPA1* and *HLA-DPB1*, two genes for which no gold standard data was available, suggesting that predictions for these genes are reliable as well.

5. A consensus metaclassifier improves HLA predictions for WES data

We noted that only for a very small fraction of the samples the genotypes are wrongly typed by all tools simultaneously (median 0.79% for WES and 0.68% for RNA; Figures S12-S13). This complementarity of the tools' allele predictions opens the possibility to combine predictions of different HLA callers into a consensus prediction. We first applied a majority voting algorithm to the output of all tools, with the predicted allele pair being the one with most votes. On the WES data, this approach outperforms the predictions of each individual tool for all genes. This is best illustrated by the *HLA-DQB1* gene, where the accuracies increased from 93.2% with the best performing tool (*HLA*LA*) to 96.3% when the voting metaclassifier was used. On RNA data, where the best tools

already attain accuracies over 99% by themselves, only minor improvements were made by combining the results (Figure S14).

Based on these results, we determined the minimal number of tools that must be included in the WES-based metaclassifier to produce reliable results (See *Methods*; Figure 7). For the WES data, including 4 tools in the model led to a considerable improvement for all genes for both MHC classes. The best accuracies were observed when *Optitype*, *HLA*LA*, *Kourami* and *Polysolver* were combined for MHC-I predictions (99.0% accuracy) and with *HLA*LA*, *HLA-HD*, *PHLAT* and *xHLA* for MHC-II predictions (98.4% accuracy). Raising the number of tools further only resulted in marginal gains. Strikingly, the accuracy of the *HLA-DQB1* allele predictions even decreases when more tools were included in the model. Therefore, we suggest combining the output of 4 tools for both MHC classes.

To evaluate whether the good performance of this approach is generalizable to other datasets, we assessed the correlation between the expected allele frequencies and the allele frequencies observed using the 4-tool WES consensus predictions on the TCGA dataset and compared the results with our previous findings. The allele frequencies predicted by the metaclassifier correlated better with the expected allele frequencies than was the case for the individual tools that supported all genes of interest (Figure 6).

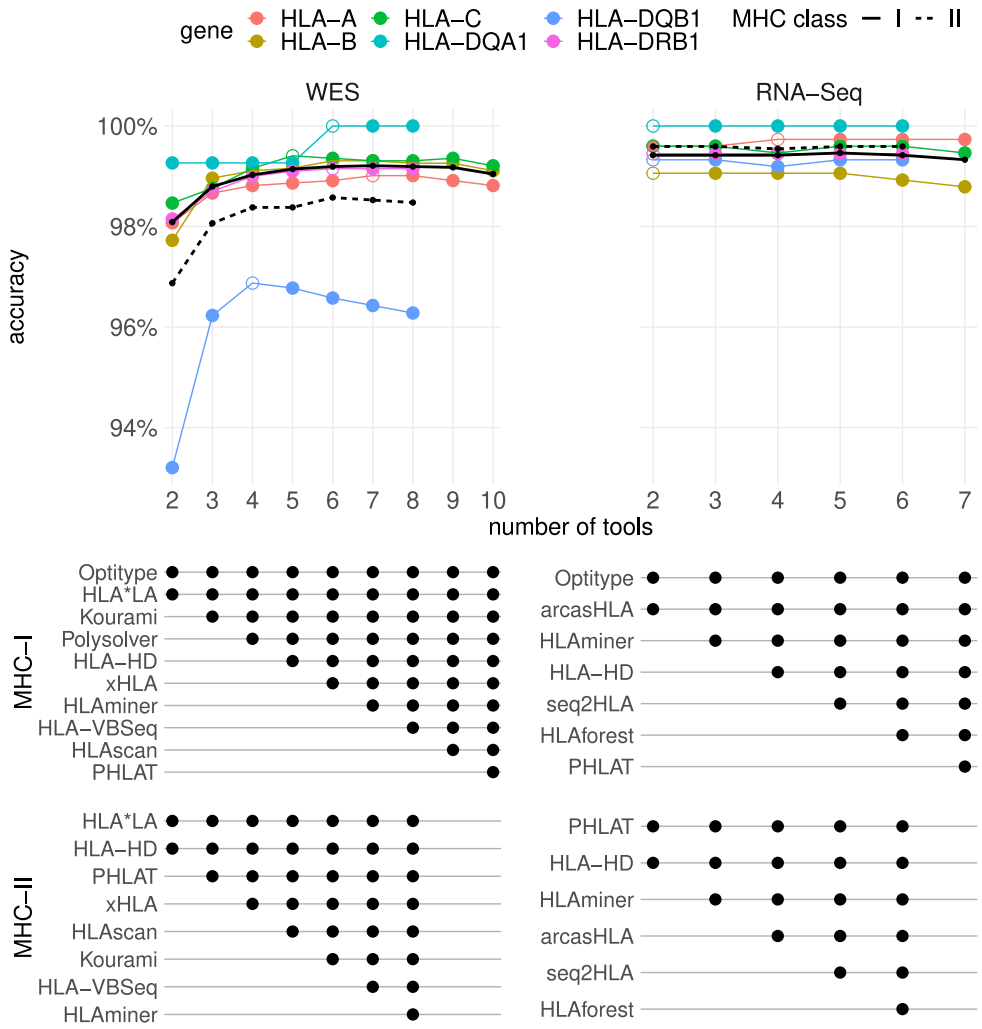

Figure 7. Accuracies of meta-prediction models with an increasing number of included tools. Tools were added one by one to the consensus metaclassifier model. At each step, the prediction accuracies of the best performing metaclassifier model for a given number of tools were plotted at the top of the figure. Unfilled markers are placed at the smallest number of tools where the maximal accuracy was obtained for that gene. Black lines indicate the average accuracy of the consensus predictions for the two MHC classes (averaged over all genes of that class). The table below the plot indicates which tools were selected in each model for a given number of tools.

Chapter 4
Discussion

Rapid technological advancements in NGS have resulted in the generation of numerous publicly available DNA and RNA sequencing datasets. These data have been critical for understanding the genomic basis of human carcinogenesis (Vogelstein et al., 2013). In the field of immuno-oncology, genomic data have also been used to study immune selection (Claeys et al., 2021; van den Eynden et al., 2019) and, additionally, the availability of corresponding clinical data opens possibilities for studying HLA-dependent cancer susceptibility or even differences in clinical ICB responses between cancer patients (Chowell et al., 2018; D. Liu et al., 2019; Naranbhai et al., 2022; Riaz et al., 2017). However, this requires that the HLA genotype for each subject can be accurately determined. An ever-increasing number of NGS-based HLA typing software applications have been developed. In this study, we benchmarked the performance of 13 publicly available tools. To our knowledge, this is the most extensive benchmark of MHC genotyping tools that has been performed so far (Table S3).

First, we evaluated the tools by comparing their output to genotypes derived from a PCR-based approach. While PCR methods are the gold standard for HLA typing, they have limitations that could lead to ambiguous typing results (S. D. Adams et al., 2004). Furthermore, inconsistencies have been reported across PCR-based HLA typing datasets that are available for the 1000 genomes samples (Bauer-Mehren et al., 2011) which could have affected our benchmarking results. Therefore, we also used 2 other, indirect approaches to assess the performance of the different tools.

Both a concordance analysis between the tools' predictions and a correlation analysis between predicted and expected allele frequencies confirmed our benchmarking results. To avoid biasing the results of this correlation analysis, we disabled ethnicity-specific allele frequencies for the algorithms that support this (i.e., *arcasHLA* and *Polysolver*). However, in the case of *arcasHLA*, when no specific ethnicity is specified, it uses prior frequencies that depend on the prevalence of the alleles in the entire human population, possibly hindering its ability to call alleles that are uncommon in the specified population. This is illustrated by the worse correlation between observed and expected allele frequencies of *arcasHLA* for *HLA-DRB1* in the African American population, due to an overestimation of the frequency of the rare *HLA-DRB1*14:02* allele.

The benchmarking of DNA-based tools was limited to WES data in our study. This likely explains the worse performance and strong coverage sensitivity of both *Kourami* and *HLA-VBSeq*, which are algorithms that were primarily developed to be applied on (high-coverage) Whole Genome Sequencing data (H. Lee & Kingsford, 2018b; Nariai et al., 2015).

We found that *Optitype, Polysolver, HLA-HD, HLA*LA* and *xHLA* are all solid choices for WES-based MHC genotyping, while *Optitype*, *HLA-HD*, *arcasHLA* and *PHLAT* are the better performing tools for RNA data. On the other hand, *HLAminer*, *HLA-VBSeq* and *HLAScan* performed rather poorly in our benchmark. Similar trends were observed in previous independent benchmarking studies (Bauer et al., 2018; Chen et al., 2021; Kiyotani et al., 2016; M. Lee et al., 2021; P. Liu et al., 2021; Yi et al., 2021; Yu et al., 2021) that focused on a subset of tools and/or genes (Table S3), with the exception of *xHLA* where we obtained considerably higher accuracies on WES data than reported in a study by Chen et al. (Chen et al., 2021).

The optimal strategy for HLA genotyping depends on a few factors: the availability of WES or RNA data, the size of the dataset that needs to be analysed and the available computational resources.

Additionally, MHC class II typing based on RNA data is only feasible on sequencing data derived from MHC-II expressing cells. For WES data, *Optitype* and *HLA-HD* are the best performing individual tools for MHC class I and MHC class II typing, respectively. For RNA data, the same tools are recommended when sufficient computational resources are available. However, the large resource and time consumption of *HLA-HD* on RNA data makes its usage rather impractical on large datasets. As an alternative, *arcasHLA* is recommended, which is both the fastest and more accurate tool for RNA that supports all 5 MHC class II genes. Finally, we have demonstrated that the accuracy of the WES-based HLA genotype predictions can be improved further by combining the output of *Optitype*, *HLA*LA*, *Kourami* and *Polysolver* for MHC-I typing and combining *HLA*LA*, *HLA-HD*, *PHLAT* and *xHLA* for MHC-II typing using a majority voting rule. The drawback of this metaclassifier approach is that it vastly increases the computational requirements, implying it is only a realistic option if sufficient resources are available or the sample size is relatively small. For RNA data a similar metaclassifier approach did not lead to a further improvement of the prediction accuracies.

Conclusion

Our extensive benchmark demonstrated that the optimal strategy for HLA genotyping from NGS data depends on the availability of either DNA or RNA sequencing data, the size of the dataset and the available computational resources. If sufficient resources are available, we recommend *Optitype* and *HLA-HD* for MHC-I and MHC-II genotype calling respectively.

Chapter 5
Methods

1. Selection of tools

A list of existing HLA genotyping tools for NGS data was compiled from literature between October and December 2020. The tools that fulfilled the following criteria were selected for further analysis: the tool should be free for academic use, support WES and/or RNA sequencing data, should not require enrichment of the HLA region before sequencing and should be a Linux command line tool that we could successfully run on our system. When the authors provided instructions on how to update the IPD-IMGT/HLA database used by their tool, this database was updated to version 3.43. This was the case for three tools: *HLA-HD*, *HLAminer* and *Kourami*.

2. Next-generation sequencing datasets for benchmark

Slices of the 1012 CRAM files of WES data from the *1000 Genomes on GRCh38* dataset (Zheng-Bradley et al., 2017) that were used for the benchmark on WES data were obtained from the *International Genome Sample Resource* using the *samtools view* command (version 1.12). The following contigs were included in the download: the MHC region on the primary assembly (chr6:28509970-33480727), all 525 contigs starting with *HLA-* and all unmapped reads. The sliced BAM files for the RNA benchmark were obtained from the *Geuvadis* (Lappalainen et al., 2013) RNA-Seq dataset (part of the *1000 genomes* project) via *ArrayExpress* (accession number *E-GEUV-1*). All reads mapped to the MHC region and the unmapped reads were included in the download. Sequencing data from NCI-60 cell lines were obtained from the *Sequence Read Archive* with accession numbers *SRP150855* (WES) (Abaan et al., 2013) and *SRP133178* (RNA) (Reinhold et al., 2019). The NCI-60 sequencing data was realigned according to the same alignment pipeline used by the *1000 Genomes on GRCh38* dataset (Zheng-Bradley et al., 2017): reads were aligned to the complete GRCh38 reference genome, including ALT contigs and HLA sequences, using an alternative scaffold-aware version of BWA-MEM. As done in the same 1000 genomes alignment pipeline, PCR-introduced duplicates were marked using the *markduplicates* function in BioBamBam (version 2.0.182). Aligned sequences of Whole Exome Sequencing (WES) and RNA sequencing experiments from *The Cancer Genome Atlas (TCGA)* were downloaded in BAM format from the *Genomic Data Commons (GDC)* portal. All 9162 available BAM files of Blood Derived normal WES samples were selected. For RNA-Seq, all 9762 RNA-Seq samples that were derived from primary tumours and were aligned using the "STAR 2-Pass" workflow, were selected. Reads mapped to the MHC region of chromosome 6 (chr6:28509970-33480727) and unmapped reads were extracted from the BAM files and downloaded following the instructions that are described in the GDC API. For the RNA-Seq samples one file failed to download after multiple attempts. The resulting dataset consists of 9162 blood-derived normal WES samples and 9761 primary tumour RNA-Seq samples from 33 available cancer types. The most resource intensive RNA tools were applied on a subset of the TCGA dataset. *Optitype* was applied on 2226 RNA files, *HLAforest* on 2900 files and *HLA-HD* was not applied on the TCGA data.

3. Calculating the coverage of sequencing data and assessing its influence on accuracy

For all downloaded whole-exome and RNA sequencing files, the average read depth in each of the exons of the HLA genes (*HLA-A, HLA-B, HLA-C, HLA-DPA1, HLA-DPB1, HLA-DQA1, HLA-DQB1* and *HLA-DRB1*) was determined using *Mosdepth (version 0.2.9)* (Pedersen & Quinlan, 2018). To assess the influence of coverage on the HLA typing accuracy, we first calculated the *average HLA read depth*

by averaging the read depth in the most polymorphic region of the HLA genes (i.e., the exons encoding the peptide binding region: exons 2 and 3 for MHC-I and exon 2 for MHC-II). The average HLA read depths for genes and samples that were correctly predicted (both alleles correct) were then compared with the average HLA read depths that correspond to incorrect predictions using a Wilcoxon rank sum test. Subsequently a logistic regression model was fitted that relates the average HLA read depths for a gene and sample with the correctness of the corresponding allele pair prediction. Then, we performed an *in silico* analysis to simulate the effect of lowering coverage. 100 WES and 100 RNA sequencing files were randomly selected. From each of these files subsampled BAM files were derived that contain respectively 100%, 50%, 10%, 5%, 1% of the reads of the original file (using the *samtools view* command, version 1.12). To obtain an absolute read depth for these samples, we multiplied the average HLA read depth by the fraction of the reads that was retained. The minimum read depth required to obtain an accuracy of 90% was then calculated by linearly interpolating the results of this analysis.

4. Gold standard HLA typing data

Gold standard PCR-based HLA calls for the samples from the *1000 genomes on GRCh38* dataset were provided by three earlier studies (Gourraud et al., 2014). The HLA genotypes from these datasets were merged. Where the calls did not agree, the calls by Gourraud et al. (Gourraud et al., 2014) were preferred. For the NCI-60 cell lines, PCR-based HLA genotypes were provided in a study by Adams et al (S. Adams et al., 2005). For both reference datasets alleles were mapped to the corresponding G-groups, as defined by IPD-IMGT (http://hla.alleles.org/alleles/g_groups.html), and trimmed to the second-field resolution.

5. HLA allele predictions

All 13 selected tools were run on the sliced BAM files following the guidelines of the authors. For tools requiring FASTQ input files, a FASTQ file was extracted from the sliced BAM files using *samtools fastq*. For *HLAScan*, which supports input files in either file format, the input was provided in BAM format. For tools that allowed to specify a list of loci that should be called: *HLA-A, HLA-B, HLA-C, HLA-DPA1, HLA-DPB1, HLA-DQA1, HLA-DQB1* and *HLA-DRB1* were chosen. *Kourami* was run with the *-a* (additional loci) parameter to call the *HLA-DPA1* and *HLA-DPB1* genes. In rare cases, this led to a crash of the tool and *Kourami* was run again without the *-a* parameter. For *HLAminer* only the HPRA mode was evaluated. *xHLA*, *Polysolver* and *HLA-VBSeq* were not compatible with BAM files that are aligned to a reference genome build that includes alternative contigs. For these tools, an additional realignment step was performed before the tool was executed. Input data for *xHLA* and *Polysolver* were realigned to a GRCh38 build that excludes alternative contigs. The input data for *HLA-VBSeq* was realigned to GRCh37. All allele predictions were mapped to the corresponding G-groups and trimmed at second-field resolution.

6. Measuring the resource consumption

The running time and memory consumption required by the tools were measured for a random subset of 10 WES and 10 RNA sequencing files from the TCGA project. Each tool was executed in a separate Docker container (version 19.03.3) that was allocated a single CPU core. When the package provided a parameter to specify the number of threads, this was set to 1. Per file, the memory usage of the Docker container was monitored using the *docker stats* command. The running time was

calculated as the time interval between the start and the end of the tool, excluding the time to start the Docker container. Pre-processing steps related to realignment to a different genome build (as required for *xHLA*, *Polysolver* and *HLA-VBSeq*) were not included in the resource consumption assessment. For HLA-HD the analysis of a single sample did not complete successfully as the required amount of memory exceeded what we have available on our system.

7. Performance metric

For each sample, two allele predictions were made. An allele prediction was labelled "correct" when it was listed as one of the two alleles in the gold standard for that patient. When a tool made a homozygous prediction, while the gold standard was heterozygous, at most one of the two predictions was labelled "correct" for that sample. The accuracy of the predictions is then defined as the proportion of all correctly predicted alleles divided by twice the number of samples. Samples where the gold standard was missing for a particular gene were ignored for that gene.

8. Population frequency data

Lists of expected HLA allele frequencies for an African American and for a Caucasian American population were constructed based on 18 different studies in the *Allele Frequency Net* (Gonzalez-Galarza et al., 2020) database (Table S4). The studies were selected based on the following criteria. First, we required that the study was conducted on a *Black* or *Caucasoid* population from the United States. This was not possible for *HLA-DPA1* where no HLA allele frequencies were available for these ethnicities. As a substitute, the allele frequencies of three European populations (French, Swedish and Basques) were used to approximate the allele frequencies for this gene in Caucasian Americans. As a second requirement, the HLA calls should be determined by a PCR-based method. Thirdly, the *Allele Frequency Net* database should have assigned a gold label (i.e., allele frequency sums to 1, sample size of study > 50, and at least 2-field resolution) to the study for the gene of interest. Lastly, it was required that the subjects included in the selected studies were healthy subjects (i.e., selected for an anthropological study, blood donors, bone marrow registry or controls for a disease study). Allele frequencies from different studies were combined by taking the average frequency, weighted according to the study's sample size. All alleles were mapped to the corresponding G-groups and trimmed at second-field resolution.

9. Correlation between expected and observed allele frequencies

For all tools and for each supported data type, the number of times that each allele was called was counted. This count was divided by the total number of samples to obtain the "observed allele frequency". The Pearson correlation was calculated between observed allele frequencies and the allele frequencies that were expected based on the *Allele Frequency Net* database.

10. Concordance of predictions among different tools

Per gene, the concordance of the predictions between each pair of tools was assessed by counting the number of allele pair predictions made by the first tool that were also made by the second tool (for the same sample and gene). Samples where one of both tools did not make a prediction were not considered. This analysis was performed on the 1000 genomes and TCGA dataset.

11. Consensus HLA predictions

A majority voting rule was used to determine the most likely HLA genotype for each sample. For each gene of interest, we selected the pair of alleles that has been predicted the most frequently for that sample (i.e., outputted by the highest number of tools). When ties occurred (i.e., multiple allele pairs had equal numbers of predictions), priority was given to the allele pair that was predicted by the tool with the best individual performance for that gene.

12. Selecting a minimum number of tools to make consensus HLA predictions

The minimal set of tools that must be included in the majority voting scheme to make reliable consensus predictions was determined using an iterative procedure. Initially, two tools were selected for the model: the tool that performed the best in the benchmark on the 1000 genomes data and the one that best complements that tool. The latter tool was defined as the tool that most often made a correct prediction (for both alleles) on the samples that were wrongly predicted by the best performing tool. Additional tools were added to this initial model with k = 2 tools in a stepwise manner. At each step, a model with $k + 1$ tools was obtained by adding one additional tool to the model with k tools. To determine which additional tool would be the most suitable choice, we evaluated all unselected tools and added the tool to the model that led to the largest increase (or the smallest decrease) in accuracy. This procedure was repeated until we obtained a model where all tools were selected.

13. Hardware and software environment

Analyses were performed on Ubuntu 20.04 on a Dell EMC PowerEdge R940xa server with 4 Intel Xeon Gold 6240 CPUs (2.60 GHz), each with 18 physical CPU cores, and 376 GiB RAM installed.

14. Data processing and statistical analysis

Data processing and statistical analyses were performed using R (version 4.0).

15. Code availability

The code that underlies this thesis is available on GitHub at
https://github.com/CCGGlab/mhc_genotyping/

Chapter 6
Annexes

Annex A - Ethical considerations

Ethical aspects directly related to the work done in the thesis

All analyses that were performed for this study are entirely based on publicly available cancer data from the 1000 Genomes Project (https://www.internationalgenome.org/1000-genomes-summary#g1k_data_reuse) and The Cancer Genome Atlas (TCGA) (https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga/history/policies). During these projects the participants have given their informed consent for sample collection and to store de-identified data derived from these samples in genomic databases.

Our research was performed entirely in agreement with the TCGA and 1000 genomes data usage policies and the relevant data protection legislation. All data has been stored on a protected server and is only accessible by the members of our lab group involved in this research. We have respected the right to anonymity of the participants and made no attempt to identify patients.

Further, we have designed our benchmark to be as inclusive as technically possible. We considered that the frequency of HLA alleles differs between populations and that existing tools might be biased towards accurate calling of alleles that are common in one particular ethnicity. To accommodate this, we assessed the performance of the tools in our TCGA based correlation analysis on both Caucasian and African Americans. Due to statistical necessities, we did not perform this analysis on other populations.

Reflection about the potential (future) impact of study results

This thesis is a part of a larger effort in our lab to understand how the immune system shapes the cancer genome. To do this, we require approaches that allow to derive the HLA genotype from existing cancer genomics datasets.

Hence, our work might lead to important new insights into the interactions between the immune system and tumours that could help to optimize the response to immunotherapy. Furthermore, the ability to predict in advance whether immunotherapy will be effective could help to lower the economic burden on the healthcare system.

As discussed in the introduction (section 1.4), HLA genotyping tools have a broader applicability outside cancer research. HLA genotyping workflows based on amplicon sequencing are currently applied in clinical laboratories to identify matching donors and recipients of hematopoietic stem cell transplantation. Given the enormous potential of WGS and WES technologies in healthcare programs (c.f. the UK 100,000 Genomes Project), the tools discussed in this thesis might eventually be integrated in such programs and allow to identify potential organ or stem cell donors with limited additional costs.

Scientific integrity

Great care has been taken to ensure the scientific integrity of this work. Sources that were consulted during the literature study have been collected using the *Mendeley* reference manager and cited where appropriate.

In our lab we put strong emphasis on transparent and reproducible research. The code that underlies this thesis is available on GitHub at https://github.com/CCGGlab/mhc_genotyping/ and includes

scripts that allow the reader to recreate the exact Conda containers that the tools and downstream analysis scripts were run in. To ensure the validity and reproducibility of our results, we additionally reran all of our scripts on a subset of 100 samples from each dataset before submission.

A master's dissertation is always a collaborative effort, with important contributions from my supervisors who provided feedback throughout the project. The results of this thesis are currently under review in *BMC Bioinformatics* and we have also incorporated the highly valuable suggestions of the reviewers into the text.

Annex B - Supplementary materials

	Alignment method	Score function			Tool-specific steps
		Jointly optimized for allele pair	PHRED score used	Prior population frequencies	
HLA*LA	BWA-MEM	yes	yes	no	Graph-based optimization of alignment
Kourami	BWA-MEM	yes	yes	no	Graph-guided assembly (post-alignment)
arcasHLA	Kallisto (pseudo-alignment)	hybrid	no	yes	Pseudo-alignment with Kallisto, followed by an iterative (EM) algorithm for transcript quantification
HLA-HD	Bowtie2	yes	no	yes*	Optimizes a score per allele pair which considers the overlap length between reads and exon sequences.
PHLAT	Bowtie2	yes	yes	yes*	Bayesian framework with likelihood calculated based on sequence consistency at SNP sites and phase consistency across adjacent SNP sites
Polysolver	Novoalign	no	yes	optional	Bayesian framework. First determines the first allele, then identifies the second allele in a separate round with updated probabilities.
HLA-VBSeq	BWA-MEM	no	no	no	First optimizes read alignments using a Bayesian framework. Then HLA types are inferred based on depth of coverage per allele. Alleles that do not pass coverage threshold are filtered out.
seq2HLA	Bowtie	no	no	no	Reads associated with first found allele removed before determining second allele. First genotyping at 2-digit level, before extended to 4-digit level.
Optitype	RazerS3	yes	no	no	Uses ILP to find the set of MHC-I allele pairs that simultaneously explain the input data the best. Does not consider rare alleles (not present in Allele Frequency Net)
xHLA	DIAMOND	hybrid	no	no	Identifies candidate alleles by applying Optitype's ILP strategy on the PBR exons with subsequent iterative refinement steps.
HLAscan	BWA-MEM	no	no	no	Alleles discarded based on number of consecutive positions with no read aligned to it
HLAforest	Bowtie	no	yes	no	SMMQ scores are propagated throughout trees that represent all possible alignments per read.
HLAminer	BWA-backtrack	no	no	no	Supports assembling reads into longer contigs before comparing them to known allele sequences†.

Table S1. Main algorithmic characteristics of the 13 selected HLA genotyping algorithms.
Algorithms differ in how reads are aligned to the HLA allele reference sequences (column A*lignment method*) and how they subsequently score candidate alleles (column *Score function*). Commonly used variables of the score function are: base quality scores (column *PHRED score used*) and whether they use prior population frequencies (column *Prior population frequencies*). The score function can either be jointly optimized for allele pairs (column *Jointly optimized for allele pair*) or be defined on a single allele at once. The tools that have the value "hybrid" in this column combine both types of scoring functions. The column *tool-specific steps* describes additional particularities of the tools.
* only used for breaking-ties in case of ambiguities, † Only in Targeted Assembly of Shotgun Reads (HPTASR) mode, which was not evaluated in this paper

	(Freely) available for academic use	FASTQ or BAM input files from WGS, WES and/or RNA-Seq	Running on Ubuntu 20.04
ALPHLARD(-NT)	✗		
ATHLATES			✗
HLAProfiler			✗
HLAreporter			✗
HLAssign		✗†	✗*
OncoHLA	✗		
PolyPheMe	✗		
SNP2HLA		✗	
SOAP-HLA			✗

Table S2. Overview of tools that were not benchmarked in our study and the reason for their exclusion.
Excluded tools were either not freely available for academic use, do not use FASTQ or BAM input files from WGS, WES and/or RNA-Seq experiments (e.g., enrichment of the HLA region prior to sequencing is needed) or we were not able to get them running on Ubuntu 20.04.
* Latest version of HLAssign is a Windows GUI tool; † HLAssign is developed for targeted sequencing. Whole genome or exome data is supported but requires manual interpretation.

		Lee			Kiyotani			Bauer		Yu		Liu*	Yi*
		A	B	C	A	B	C	class I	class I+II	class I	class II	class I+II	class I
HLA*LA	DNA	81,6	83,8	48,3									
HLA-HD	DNA	81,7	75,0	73,3								*100,0*	
HLAminer	DNA							26,0	26,0	*31,8*	*56,8*	*68,0*	
HLAscan	DNA	54,7	41,5	10,3								*78,0*	
HLA-VBSeq	DNA							86,0	68,0	60,3	*54,1*	*35,0*	
Kourami	DNA	57,2	53,9	58,0									
Optitype	DNA	91,9	85,4	89,7	97,3	96,6	97,7	98,0		88,8			99,0
PHLAT	DNA	60,7	67,9	73,0	79,1	85,1	92,8	88,0	73,0				94,0
Polysolver	DNA				93,4	92,5	96,1						96,0
xHLA	DNA												
arcasHLA	RNA									*96,5*	*92,7*		
HLA-HD	RNA												
HLAforest	RNA									*81,8*	*88,7*		
HLAminer	RNA							20,0	20,0	*37,7*	*50,6*		
Optitype	RNA							99,0		54,9			96,4
PHLAT	RNA							96,0	81,0				84,5
seq2HLA	RNA							95,0	79,0	*96,5*	*88,7*		91,1

Table S3. Comparison of our results with 7 other independent benchmark studies.
The allele prediction accuracies obtained in our benchmark on the 1000 genomes data are compared with 7 other independent benchmark papers. Each score was assigned a colour code: red under 50%, yellow between 50% and 80% and green above 80%. A grey box indicates that a tool was not evaluated for the corresponding gene in that study. The second row of the table indicates the gene or MHC class, where *class I + II* corresponds to the overall (combined) accuracy for MHC class I and class II. For the Lee, Kiyotani, Bauer and Yu studies the tools were evaluated on data from the 1000 genomes project. The studies by Liu, Yi and Chen used a different in house dataset. DNA data always refers to WES data.

		Chen*								Our study							
		A	B	C	**class I**	DQA1	DQB1	DRB1	**class II**	A	B	C	**class I**	DQA1	DQB1	DRB1	**class II**
HLA*LA	DNA	98,2	100,0	99,1	**99,1**	100,0	99,1	100,0	**99,7**	89,9	97,2	96,2	**94,4**	97,1	93,2	97,0	**95,7**
HLA-HD	DNA	100,0	99,1	99,1	**99,4**	94,6	100,0	100,0	**98,6**	84,3	93,4	90,1	**89,3**	99,3	91,1	98,2	**96,2**
HLAminer	DNA									31,3	22,2	25,8	**26,4**	56,2	48,9	56,4	**53,8**
HLAscan	DNA									59,1	60,7	59,5	**59,8**	92,9	61,3	68,4	**74,2**
HLA-VBSeq	DNA									68,2	58,7	62,2	**63,0**	79,4	50,0	51,3	**60,2**
Kourami	DNA									83,4	76,1	79,6	**79,7**	94,1	71,6	95,0	**86,9**
Optitype	DNA									98,0	97,6	98,4	**98,0**				
PHLAT	DNA									82,4	90,1	96,7	**89,7**	99,3	93,0	81,3	**91,2**
Polysolver	DNA	100,0	95,5	97,3	**97,6**					94,9	91,8	98,0	**94,9**				
xHLA	DNA	54,5	41,8	45,5	**47,2**		48,2	58,2	**56,7**	94,6	93,0	86,5	**91,4**		89,9	94,2	**92,0**
arcasHLA	RNA									99,6	99,1	99,6	**99,4**	100,0	96,1	98,2	**98,1**
HLA-HD	RNA									98,4	98,0	97,7	**98,0**	100,0	99,1	99,2	**99,4**
HLAforest	RNA									90,8	71,2	76,9	**79,6**	92,5	88,5	89,8	**90,3**
HLAminer	RNA									45,6	48,8	39,2	**44,5**	59,4	77,6	47,9	**61,7**
Optitype	RNA									99,3	98,7	99,6	**99,2**				
PHLAT	RNA									98,5	94,6	93,0	**95,4**	100,0	98,4	98,2	**98,9**
seq2HLA	RNA									98,0	95,0	94,6	**95,9**	89,6	81,5	92,1	**87,8**

Table S3 (continued)

Ethnic origin	Gene	AFN population ID	AFN population name
Black	*HLA-A*	1480	USA African American
Black	*HLA-A*	1620	USA African American Bethesda
Black	*HLA-A*	2223	USA African American pop 3
Black	*HLA-A*	2419	USA African American pop 4
Black	*HLA-B*	1480	USA African American
Black	*HLA-B*	2223	USA African American pop 3
Black	*HLA-B*	2419	USA African American pop 4
Black	*HLA-C*	1480	USA African American
Black	*HLA-C*	2223	USA African American pop 3
Black	*HLA-C*	2419	USA African American pop 4
Black	*HLA-DPB1*	2779	USA African American pop 7
Black	*HLA-DQA1*	1620	USA African American Bethesda
Black	*HLA-DQB1*	2419	USA African American pop 4
Black	*HLA-DQB1*	2779	USA African American pop 7
Black	*HLA-DRB1*	1511	USA Colorado Univ Cord Blood Bank African American
Black	*HLA-DRB1*	1620	USA African American Bethesda
Black	*HLA-DRB1*	2223	USA African American pop 3
Black	*HLA-DRB1*	2419	USA African American pop 4
Black	*HLA-DRB1*	2779	USA African American pop 7
Caucasoid	*HLA-A*	1359	USA San Antonio Caucasian
Caucasoid	*HLA-A*	1479	USA Caucasian pop 2
Caucasoid	*HLA-A*	1619	USA Caucasian Bethesda
Caucasoid	*HLA-A*	2570	USA Eastern European
Caucasoid	*HLA-B*	1359	USA San Antonio Caucasian
Caucasoid	*HLA-B*	1479	USA Caucasian pop 2
Caucasoid	*HLA-B*	1895	USA Philadelphia Caucasian
Caucasoid	*HLA-B*	2570	USA Eastern European
Caucasoid	*HLA-C*	1359	USA San Antonio Caucasian
Caucasoid	*HLA-C*	1479	USA Caucasian pop 2
Caucasoid	*HLA-C*	1619	USA Caucasian Bethesda
Caucasoid	*HLA-C*	1895	USA Philadelphia Caucasian
Caucasoid	*HLA-DPA1*	1279	France Ceph
Caucasoid	*HLA-DPA1*	1401	Spain Navarre Basques
Caucasoid	*HLA-DPA1*	2531	Sweden pop 2
Caucasoid	*HLA-DPB1*	2780	USA Caucasian pop 5
Caucasoid	*HLA-DQA1*	1619	USA Caucasian Bethesda
Caucasoid	*HLA-DQB1*	1359	USA San Antonio Caucasian
Caucasoid	*HLA-DQB1*	1895	USA Philadelphia Caucasian
Caucasoid	*HLA-DQB1*	2780	USA Caucasian pop 5
Caucasoid	*HLA-DRB1*	1359	USA San Antonio Caucasian
Caucasoid	*HLA-DRB1*	1513	USA Colorado Univ Cord Blood Bank Caucasian

Caucasoid	*HLA-DRB1*	1586	USA Caucasian Houston
Caucasoid	*HLA-DRB1*	1588	USA Caucasian Pittsburgh
Caucasoid	*HLA-DRB1*	1619	USA Caucasian Bethesda
Caucasoid	*HLA-DRB1*	1895	USA Philadelphia Caucasian
Caucasoid	*HLA-DRB1*	2570	USA Eastern European
Caucasoid	*HLA-DRB1*	2780	USA Caucasian pop 5

Table S4. Overview of studies from the Allele Frequency Net (AFN) database which were used to compile the list of expected HLA allele frequencies.
For the listed ethnicities (column *Ethnic origin*) and genes (column *Gene*), we indicate which AFN "populations" (datasets) were used to calculate the corresponding expected allele frequencies. For each dataset, the names (column *AFN population name*) and internal IDs (column *AFN population ID*) are given.

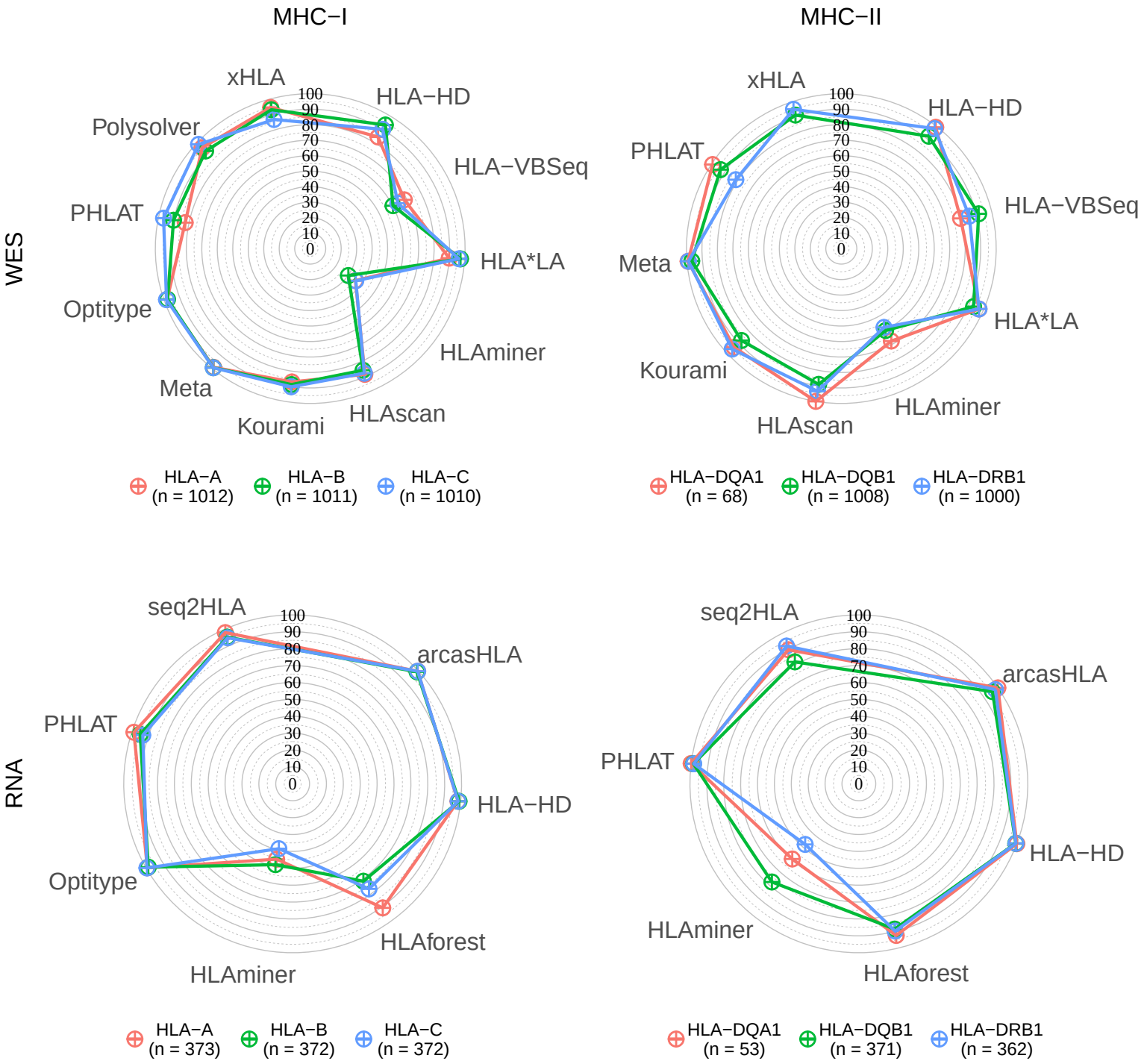

Figure S1: Fraction of correct allele predictions (1000 genomes)
Radar plots depicting the fraction of correct allele predictions relative to the total number of alleles for which the algorithm was able to make a prediction on the 1000 genomes dataset. Coloured lines represent different genes, as indicated in the legend below the plots. Corners of the radar plots correspond to the tools that were evaluated for that data type. The Meta tools correspond to the 4-tools metaclassifiers.

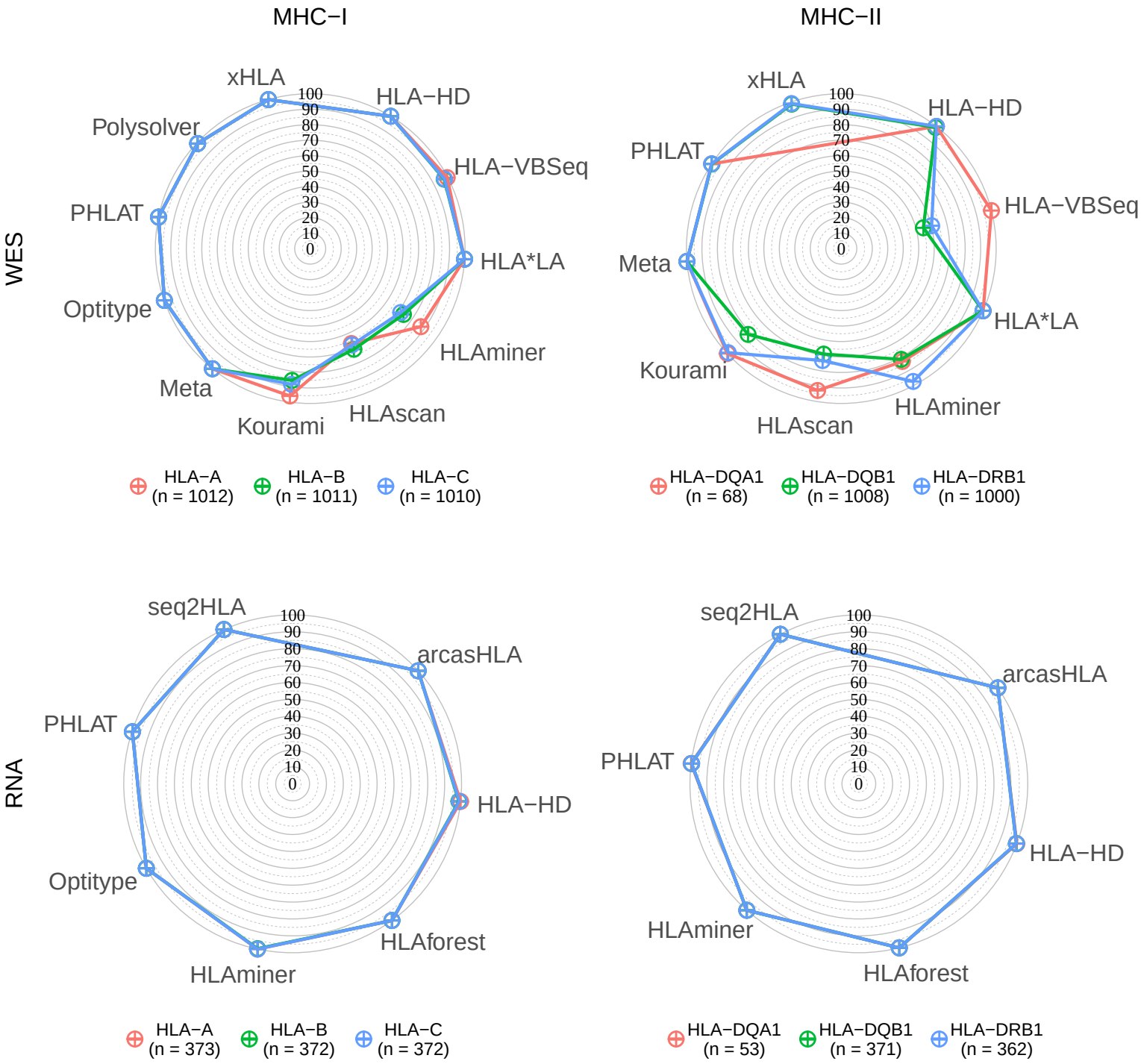

Figure S2: Fraction of successful allele predictions (1000 genomes)
Radar plots depicting the fraction of alleles for which the tool was able to make a prediction on the 1000 genomes dataset. Coloured lines represent different genes, as indicated in the legend below the plots. Corners of the radar plots correspond to the tools that were evaluated for that data type. The Meta tools correspond to the 4-tools metaclassifiers.

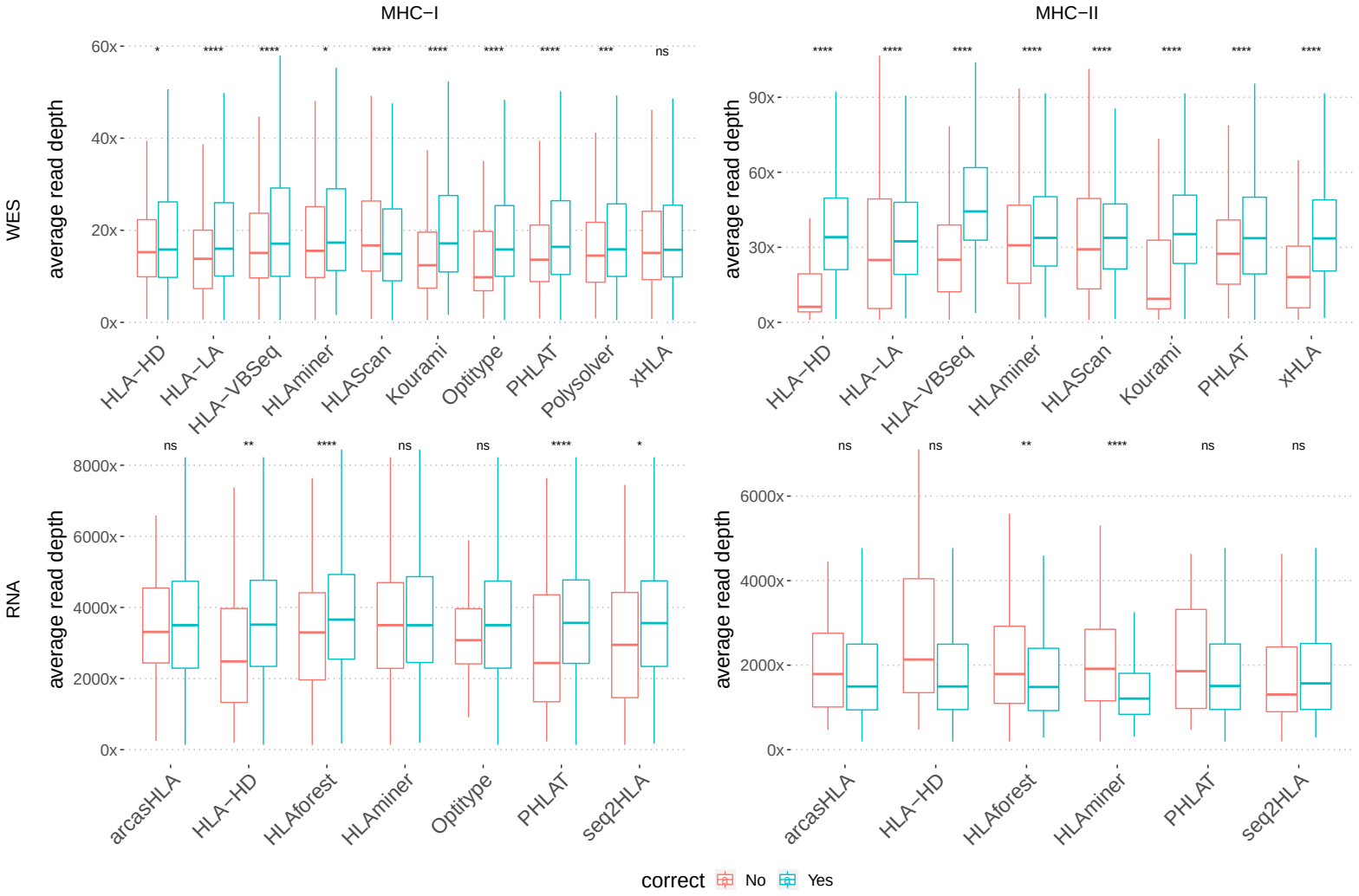

Figure S3: Comparison between the average HLA read depth for correct and incorrect predictions Boxplot comparing the average HLA read depth in samples and genes that were either correctly (cyan) or incorrectly (red) predicted. The y-axis indicates the coverage in these exons. The x-axis indicates the different tools.

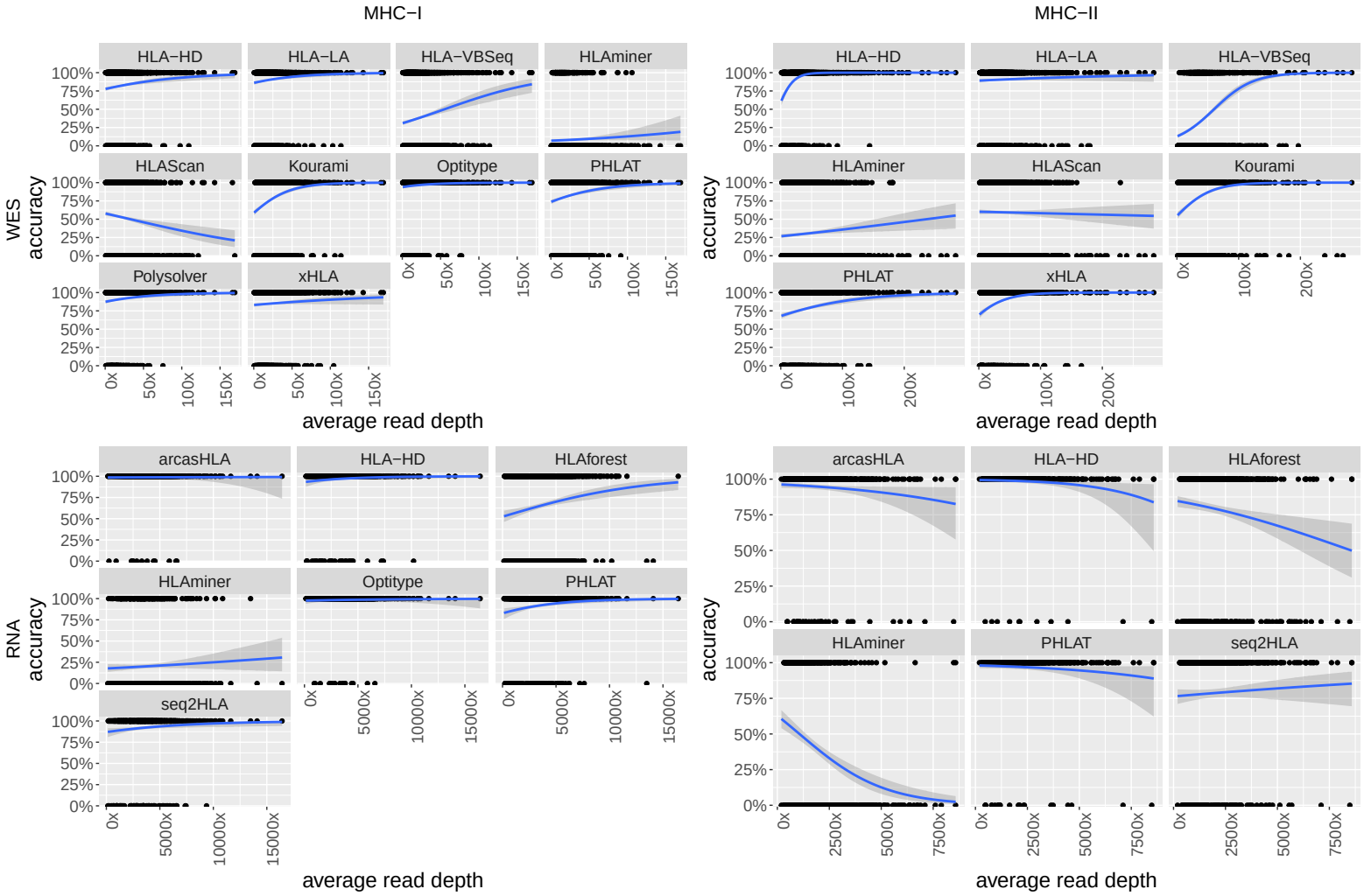

Figure S4: Logistic regression between average HLA read depth and the accuracy of the allele predictions

Logistic regression model that relates the average HLA read depth with the correctness of the allele pair prediction. The x-axis indicates the coverage. The y-axis indicates the probability that a prediction is correct for a sample with the corresponding coverage.

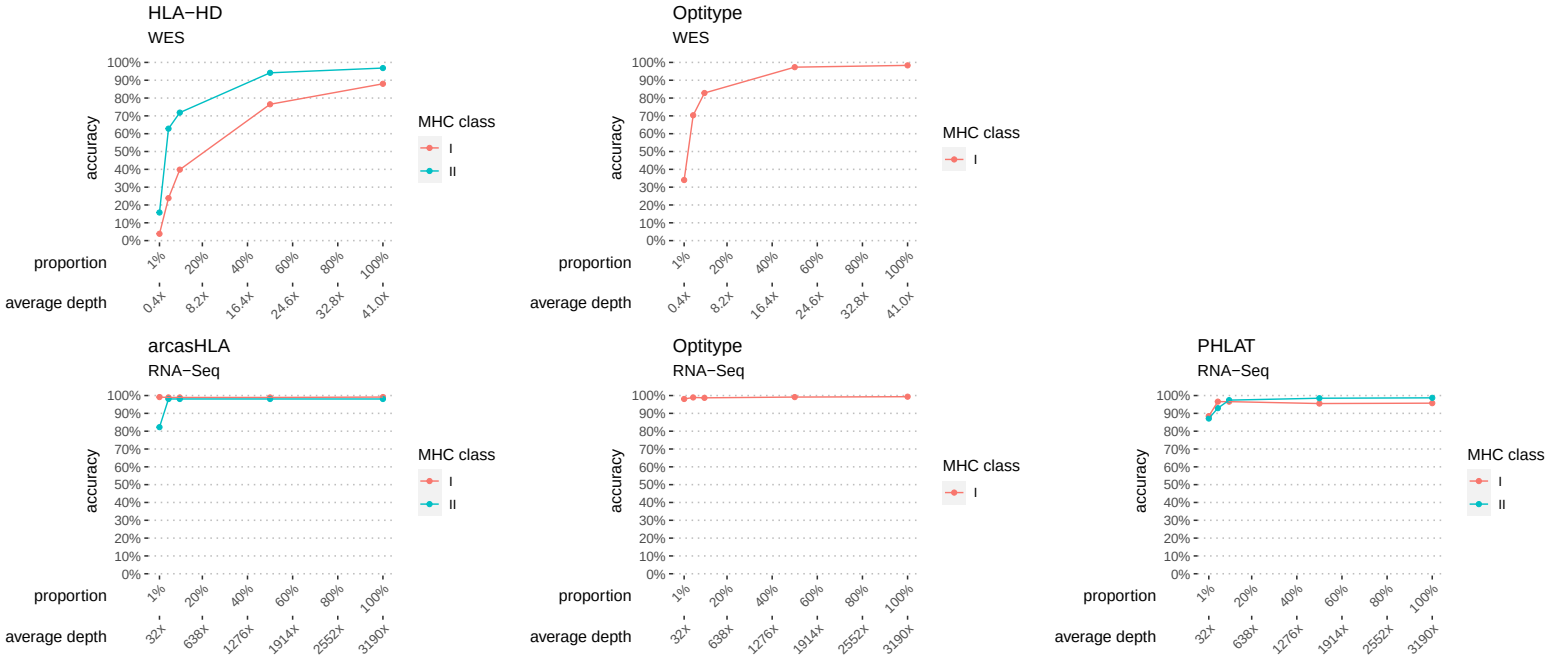

Figure S5: Accuracy of HLA allele predictions in subsampled sequencing files for the recommended tools

Scatter plot that displays for 100 randomly selected WES and 100 randomly selected RNA sequencing files which accuracy was be obtained when a given proportion of the reads was retained. The line type indicates the average accuracy of the allele predictions for a certain MHC class. The colour of the lines indicates the data type (red for WES and cyan for RNA-Seq).

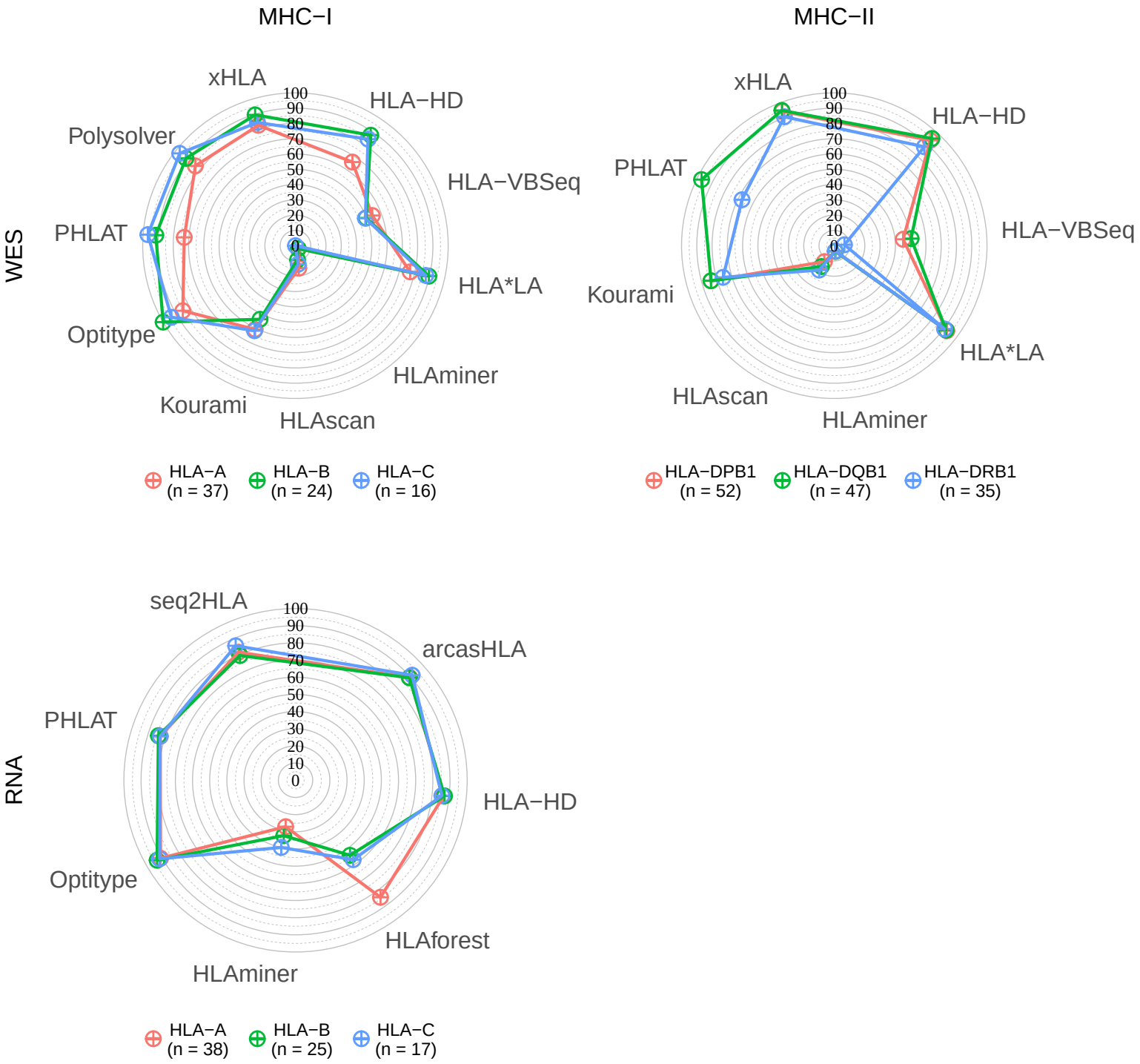

Figure S6: HLA allele prediction accuracies on NCI-60 cell lines
Radar plots of HLA allele prediction accuracies on data from NCI-60 cell lines. Coloured lines represent different genes, as indicated in the legend below the plots. Corners of the radar plots correspond to the tools that were evaluated for that data type.

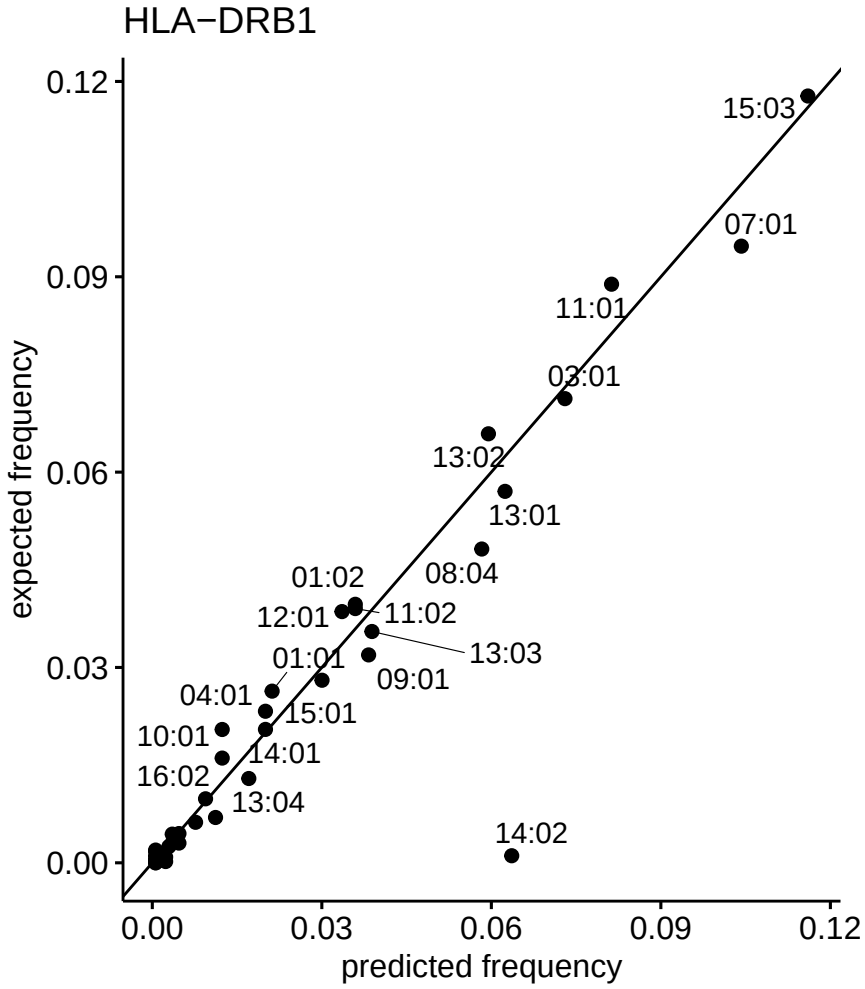

Figure S7: Expected frequency of HLA-DRB1 alleles in an African American population vs frequencies predicted by arcasHLA

Scatter plot that compares the allele frequency as predicted by arcasHLA (x-axis) with the expected allele frequencies based on data from Allele Frequency Net (y-axis).

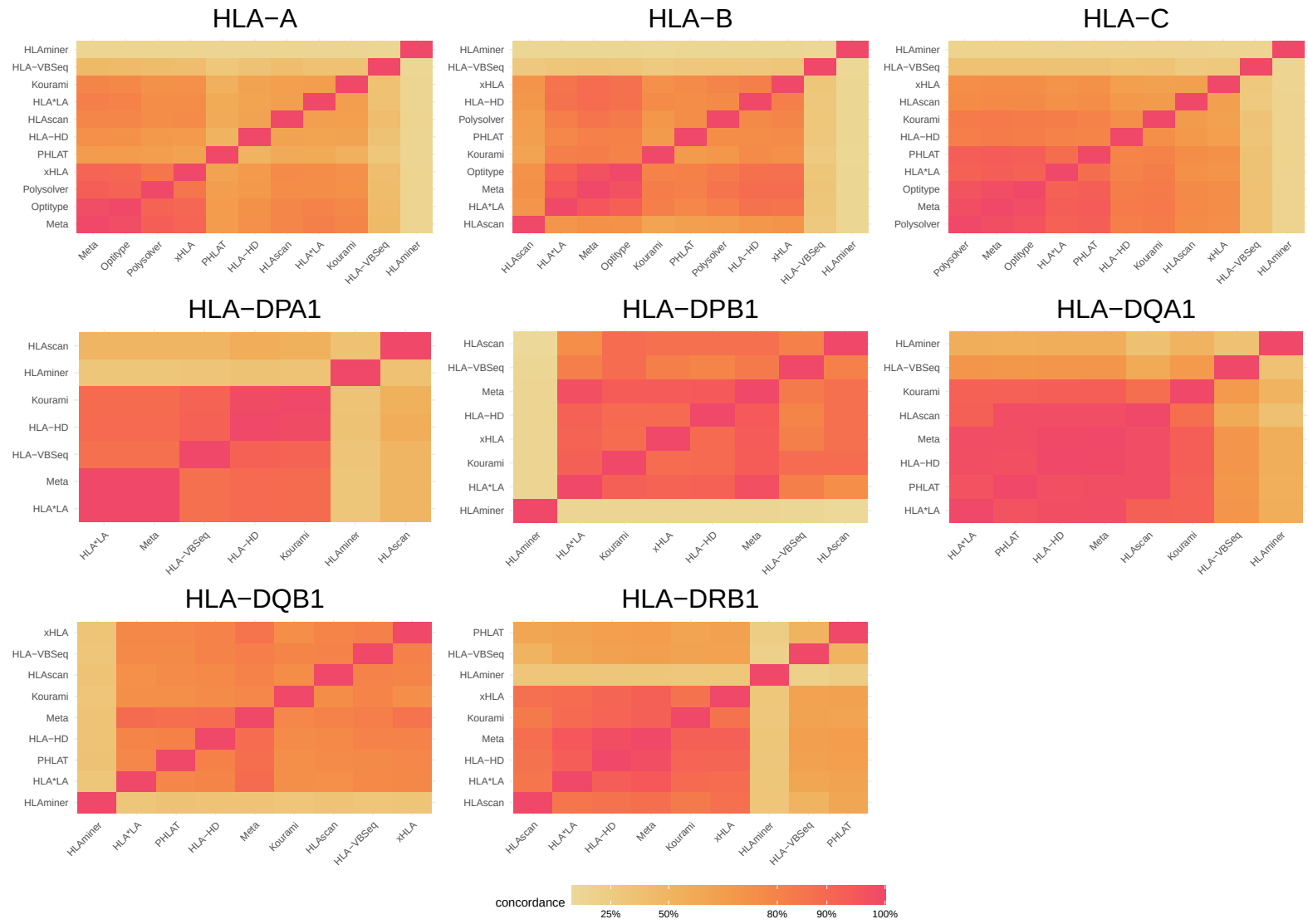

Figure S8: Concordance of HLA calls between each pair of tools on DNA data (1000 genomes) Heatmaps representing the concordance of the HLA calls between each pair of tools, applied on the 1000 genomes DNA data. Hierarchical clustering was applied on the tools. The Meta tool corresponds to the 4-tool consensus metaclassifier.

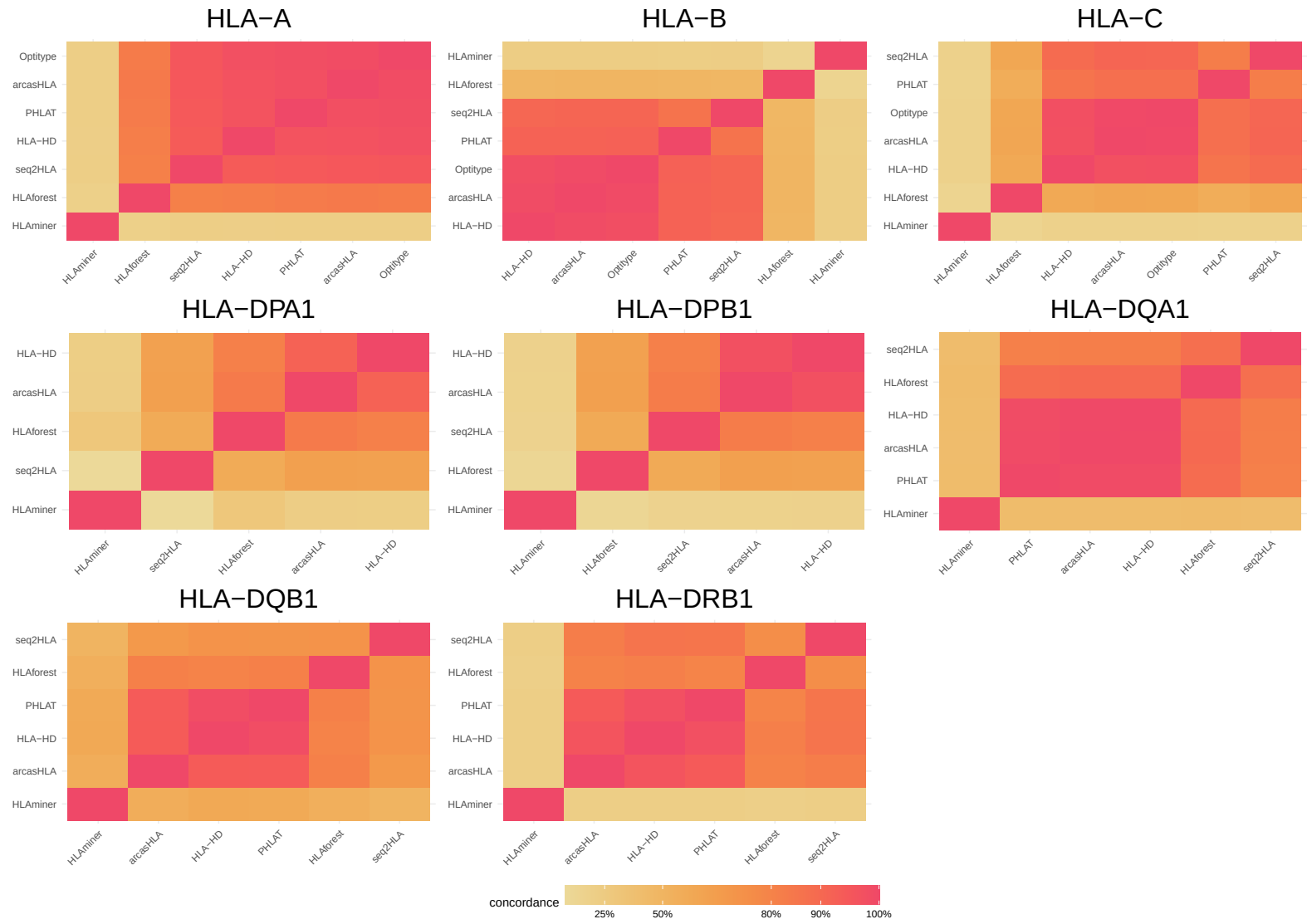

Figure S9: Concordance of HLA calls between each pair of tools on RNA data (1000 genomes) Heatmaps representing the concordance of the HLA calls between each pair of tools, applied on the 1000 genomes RNA data. Hierarchical clustering was applied on the tools.

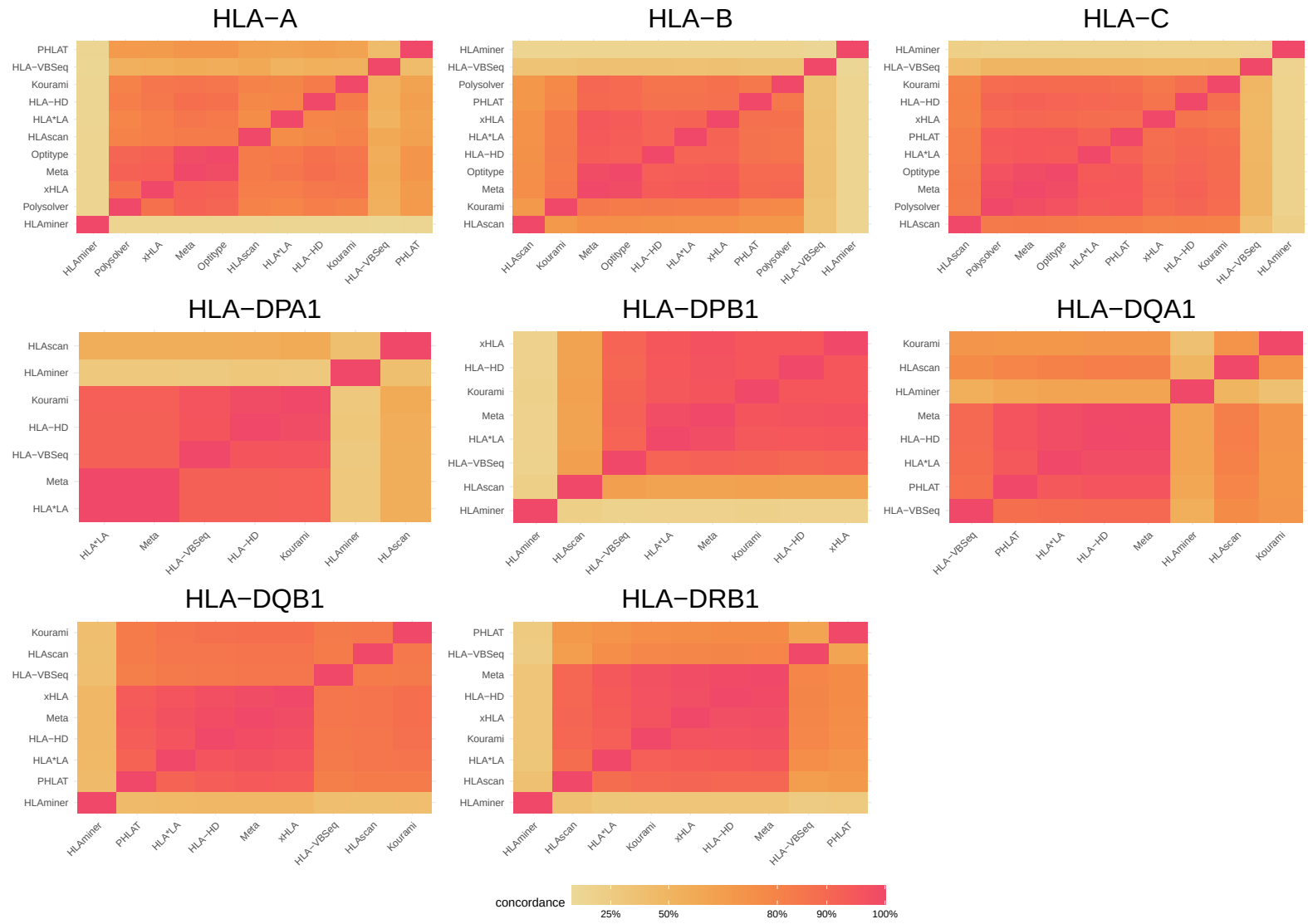

Figure S10: Concordance of HLA calls between each pair of tools on DNA data (TCGA) Heatmaps representing the concordance of the HLA calls between each pair of tools, applied on the TCGA DNA data. Hierarchical clustering was applied on the tools. The Meta tool corresponds to the 4-tool consensus metaclassifier.

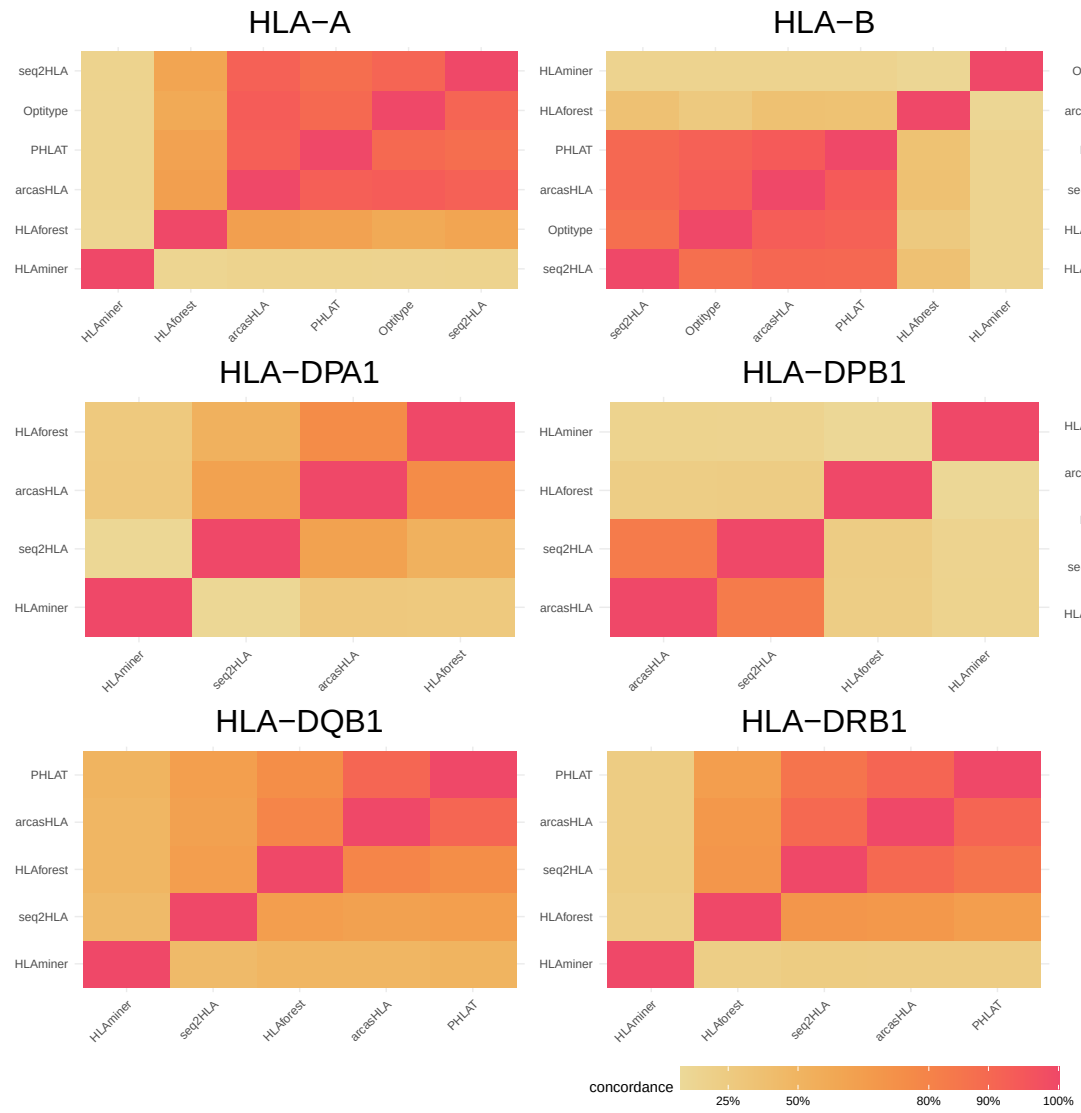

Figure S11: Concordance of HLA calls between each pair of tools on RNA data (TCGA) Heatmaps representing the concordance of the HLA calls between each pair of tools, applied on the TCGA RNA data. Hierarchical clustering was applied on the tools.

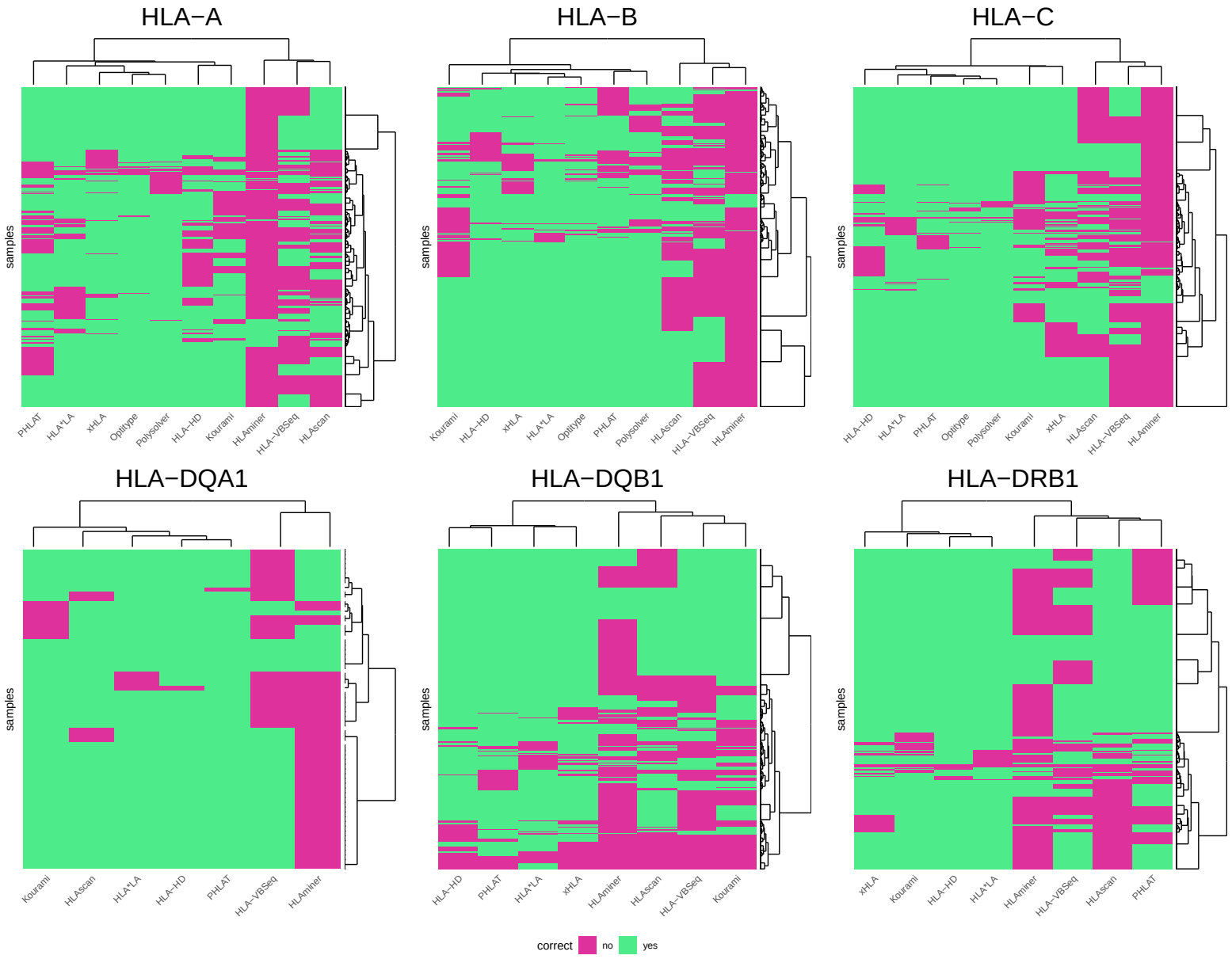

Figure S12: Correctness of predictions on DNA data
Heatmap indicating correctness of predictions on DNA data for each sample (rows) and tool (columns). Hierarchical clustering was applied on tools and samples. Dendrogram for the tools is shown on top of the plots. Dendrogram for the samples is shown right of the plots.

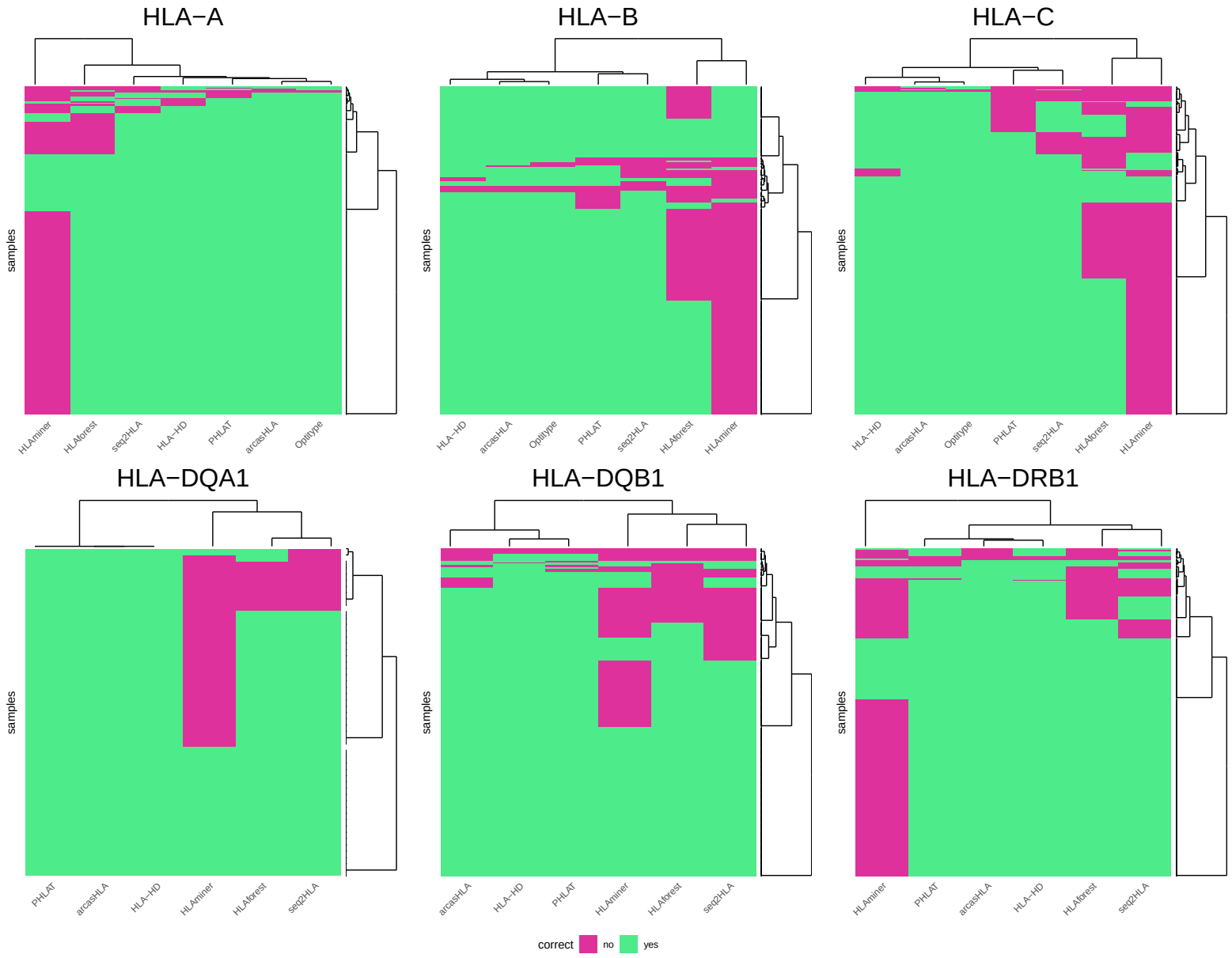

Figure S13: Correctness of predictions on RNA data
Heatmap indicating correctness of predictions on RNA data for each sample (rows) and tool (columns). Hierarchical clustering was applied on tools and samples. Dendrogram for the tools is shown on top of the plots. Dendrogram for the samples is shown right of the plots.

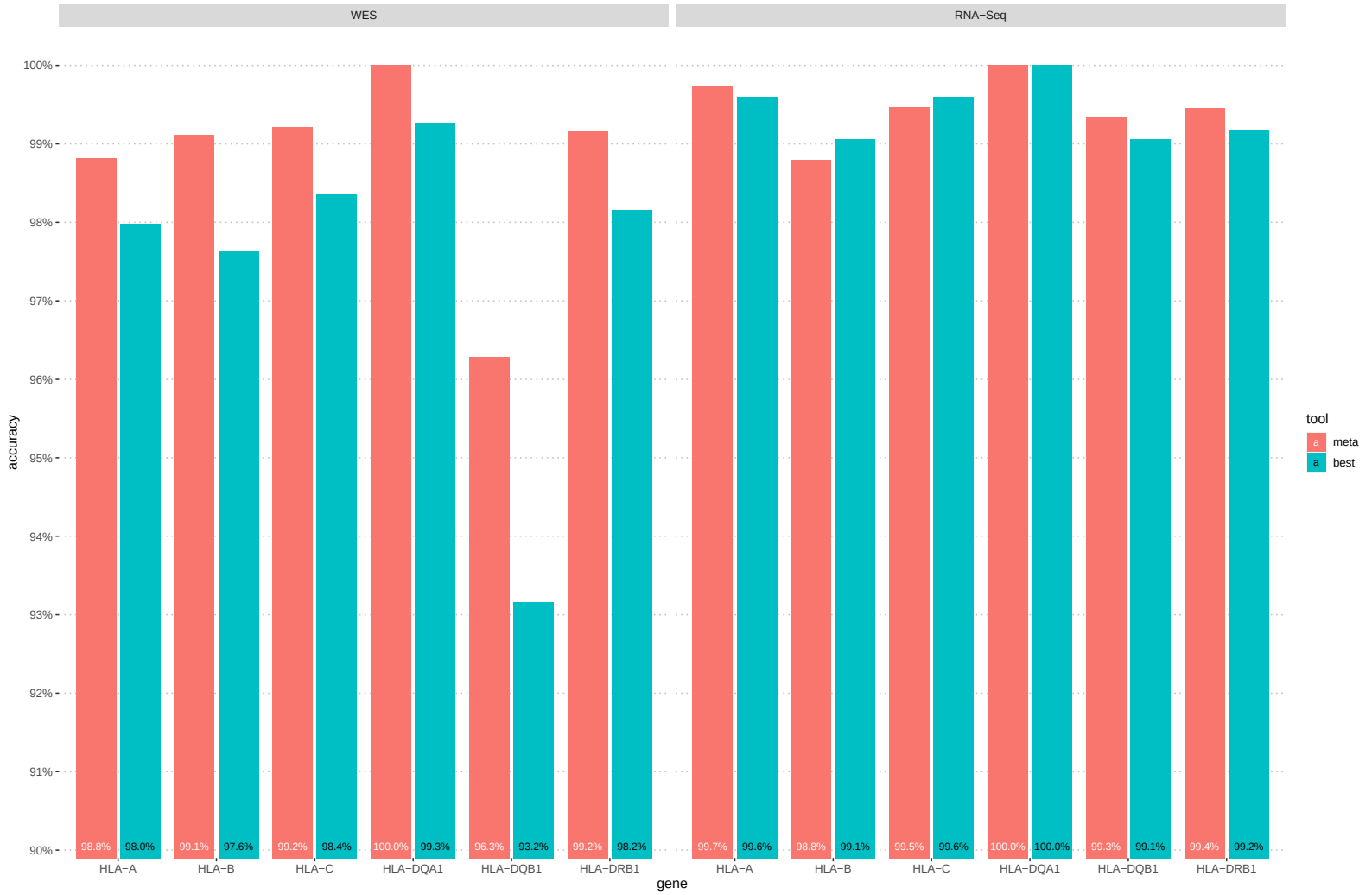

Figure S14: Comparison of accuracies of all-tool metaclassifier with best performing individual tool per gene

Barplots comparing the accuracy of the best tool for each gene and data type to the accuracy of a classifier that chooses an HLA genotype from the output of all tools that support that data type and gene based on a majority voting rule. Bars in a red correspond to the accuracies of the voting classifier. Bars in blue correspond to the accuracies of the best individual tool for that gene.

Chapter 7
References

Abaan, O. D., Polley, E. C., Davis, S. R., Zhu, Y. J., Bilke, S., Walker, R. L., Pineda, M., Gindin, Y., Jiang, Y., Reinhold, W. C., Holbeck, S. L., Simon, R. M., Doroshow, J. H., Pommier, Y., & Meltzer, P. S. (2013). The exomes of the NCI-60 panel: a genomic resource for cancer biology and systems pharmacology. *Cancer Research*, *73*(14), 4372–4382. https://doi.org/10.1158/0008-5472.CAN-12-3342

Abi-Rached, L., Gouret, P., Yeh, J. H., Cristofaro, J. di, Pontarotti, P., Picard, C., & Paganini, J. (2018). Immune diversity sheds light on missing variation in worldwide genetic diversity panels. *PLOS ONE*, *13*(10), e0206512. https://doi.org/10.1371/JOURNAL.PONE.0206512

Adams, S. D., Barracchini, K. C., Chen, D., Robbins, F., Wang, L., Larsen, P., Luhm, R., & Stroncek, D. F. (2004). *Ambiguous allele combinations in HLA Class I and Class II sequence-based typing: when precise nucleotide sequencing leads to imprecise allele identification*. https://doi.org/10.1186/1479-5876-2-30

Adams, S., Robbins, F. M., Chen, D., Wagage, D., Holbeck, S. L., Morse, H. C., Stroncek, D., & Marincola, F. M. (2005). HLA class I and II genotype of the NCI-60 cell lines. *Journal of Translational Medicine*, *3*, 11. https://doi.org/10.1186/1479-5876-3-11

Ahadova, A., Witt, J., Haupt, S., Gallon, R., Hüneburg, R., Nattermann, J., ten Broeke, S., Bohaumilitzky, L., Hernandez-Sanchez, A., Santibanez-Koref, M., Jackson, M. S., Ahtiainen, M., Pylvänäinen, K., Andini, K., Grolmusz, V. K., Möslein, G., Dominguez-Valentin, M., Møller, P., Fürst, D., … Kloor, M. (2022). Is HLA type a possible cancer risk modifier in Lynch syndrome? *International Journal of Cancer*, *17*. https://doi.org/10.1002/IJC.34312

Allen, R. L., & Hogan, L. (2013). Non-Classical MHC Class I Molecules (MHC-Ib). *ELS*. https://doi.org/10.1002/9780470015902.A0024246

Alser, M., Rotman, J., Deshpande, D., Taraszka, K., Shi, H., Baykal, P. I., Yang, H. T., Xue, V., Knyazev, S., Singer, B. D., Balliu, B., Koslicki, D., Skums, P., Zelikovsky, A., Alkan, C., Mutlu, O., & Mangul, S. (2021). Technology dictates algorithms: recent developments in read alignment. *Genome Biology*, *22*(1), 249. https://doi.org/10.1186/s13059-021-02443-7

Alspach, E., Lussier, D. M., Miceli, A. P., Kizhvatov, I., DuPage, M., Luoma, A. M., Meng, W., Lichti, C. F., Esaulova, E., Vomund, A. N., Runci, D., Ward, J. P., Gubin, M. M., Medrano, R. F. V., Arthur, C. D., White, J. M., Sheehan, K. C. F., Chen, A., Wucherpfennig, K. W., … Schreiber, R. D. (2019). MHC-II neoantigens shape tumour immunity and response to immunotherapy. *Nature*, *574*(7780), 696–701. https://doi.org/10.1038/s41586-019-1671-8

Andersen, M. H., Schrama, D., thor Straten, P., & Becker, J. C. (2006). Cytotoxic T Cells. *Journal of Investigative Dermatology*, *126*(1), 32–41. https://doi.org/10.1038/sj.jid.5700001

Anderson, M. W., & Schrijver, I. (2010). Next Generation DNA Sequencing and the Future of Genomic Medicine. *Genes 2010, Vol. 1, Pages 38-69*, *1*(1), 38–69. https://doi.org/10.3390/GENES1010038

Arndt, S. O., Vogt, A. B., Markovic-Plese, S., Martin, R., Moldenhauer, G., Wölpl, A., Sun, Y., Schadendorf, D., Hämmerling, G. J., & Kropshofer, H. (2000). Functional HLA-DM on the surface of B cells and immature dendritic cells. *The EMBO Journal*, *19*(6), 1241–1251. https://doi.org/10.1093/emboj/19.6.1241

Assembly Terminology - Genome Reference Consortium. (n.d.). Retrieved 5 January 2023, from https://www.ncbi.nlm.nih.gov/grc/help/definitions/

Bai, Y., Ni, M., Cooper, B., Wei, Y., & Fury, W. (2014). Inference of high resolution HLA types using genome-wide RNA or DNA sequencing reads. *BMC Genomics*, *15*(1), 1–16. https://doi.org/10.1186/1471-2164-15-325/FIGURES/3

Bai, Y., Wang, D., & Fury, W. (2018). PHLAT: Inference of high-resolution HLA types from RNA and whole exome sequencing. *Methods in Molecular Biology*, *1802*, 193–201. https://doi.org/10.1007/978-1-4939-8546-3_13

Baruzzo, G., Hayer, K. E., Kim, E. J., di Camillo, B., Fitzgerald, G. A., & Grant, G. R. (2016). Simulation-based comprehensive benchmarking of RNA-seq aligners. *Nature Methods 2016 14:2*, *14*(2), 135–139. https://doi.org/10.1038/nmeth.4106

Bauer, D. C., Zadoorian, A., Wilson, L. O. W., Alliance, M. G. H., & Thorne, N. P. (2018). Evaluation of computational programs to predict HLA genotypes from genomic sequencing data. *Briefings in Bioinformatics*, *19*(2), 179–187. https://doi.org/10.1093/BIB/BBW097

Bauer-Mehren, A., Bundschus, M., Rautschka, M., Mayer, M. A., Sanz, F., & Furlong, L. I. (2011). Gene-disease network analysis reveals functional modules in mendelian, complex and environmental diseases. *PLoS ONE*, *6*(6), e20284. https://doi.org/10.1371/journal.pone.0020284

Beck, S., Geraghty, D., Inoko, H., Rowen, L., Aguado, B., Bahram, S., Campbell, R. D., Forbes, S. A., Guillaudeux, T., Hood, L., Horton, R., Janer, M., Jasoni, C., Madan, A., Milne, S., Neville, M., Oka, A., Qin, S., Ribas-Despuig, G., … Yamazaki, M. (1999). Complete sequence and gene map of a human major histocompatibility complex. *Nature 1999 401:6756*, *401*(6756), 921–923. https://doi.org/10.1038/44853

Bentley, G., Higuchi, R., Hoglund, B., Goodridge, D., Sayer, D., Trachtenberg, E. A., & Erlich, H. A. (2009). High-resolution, high-throughput HLA genotyping by next-generation sequencing. *Tissue Antigens*, *74*(5), 393. https://doi.org/10.1111/J.1399-0039.2009.01345.X

Bjorkman, P. J., & Parham, P. (1990). Structure, function, and diversity of class I Major Histocompatibility Complex molecules. *Annual Review of Biochemistry*, *59*(1), 253–288. https://doi.org/10.1146/annurev.bi.59.070190.001345

Blasczyk, R. (2003). HLA Diagnostic Sequencing -Conception, Application and Automation. *Laboratoriums Medizin*, *27*(9–10), 359–368. https://doi.org/10.1046/j.1439-0477.2003.03069.x

Bleidorn, C. (2017a). Alignment and Mapping. In *Phylogenomics* (pp. 105–125). Springer, Cham. https://doi.org/10.1007/978-3-319-54064-1_6

Bleidorn, C. (2017b). Assembly and Data Quality. In *Phylogenomics* (pp. 81–103). Springer, Cham. https://doi.org/10.1007/978-3-319-54064-1_5

Boegel, S., Löwer, M., Bukur, T., Sahin, U., & Castle, J. C. (2014). A catalog of HLA type, HLA expression, and neoepitope candidates in human cancer cell lines. *OncoImmunology*, *3*(8). https://doi.org/10.4161/21624011.2014.954893/SUPPL_FILE/KONI_A_954893_SM1014.ZIP

Boegel, S., Löwer, M., Schäfer, M., Bukur, T., de Graaf, J., Boisguérin, V., Türeci, Ö., Diken, M., Castle, J. C., & Sahin, U. (2012). HLA typing from RNA-Seq sequence reads. *Genome Medicine*, *4*(12), 1–12. https://doi.org/10.1186/GM403

Bontadini, A. (2012). HLA techniques: Typing and antibody detection in the laboratory of immunogenetics. *Methods*, *56*(4), 471–476. https://doi.org/10.1016/j.ymeth.2012.03.025

Borst, L., van der Burg, S. H., & van Hall, T. (2020). The NKG2A-HLA-E Axis as a Novel Checkpoint in the Tumor Microenvironment. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research*, *26*(21), 5549–5556. https://doi.org/10.1158/1078-0432.CCR-19-2095

Braud, V. M., Allan, D. S. J., O'Callaghan, C. A., Söderström, K., D'Andrea, A., Ogg, G. S., Lazetic, S., Young, N. T., Bell, J. I., Phillips, J. H., Lanier, L. L., & McMichael, A. J. (1998). HLA-E binds to natural killer cell receptors CD94/NKG2A, B and C. *Nature*, *391*(6669), 795–799. https://doi.org/10.1038/35869

Bray, N. L., Pimentel, H., Melsted, P., & Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology 2016 34:5*, *34*(5), 525–527. https://doi.org/10.1038/nbt.3519

Buchkovich, M. L., Brown, C. C., Robasky, K., Chai, S., Westfall, S., Vincent, B. G., Weimer, E. T., & Powers, J. G. (2017). HLAProfiler utilizes k-mer profiles to improve HLA calling accuracy for rare and common alleles in RNA-seq data. *Genome Medicine*, *9*(1), 1–15. https://doi.org/10.1186/S13073-017-0473-6

Bunce, M., & Passey, B. (2013). HLA Typing by Sequence-Specific Primers. In A. A. Zachary & M. S. Leffell (Eds.), *Transplantation Immunology: Methods and Protocols* (pp. 147–159). Humana Press. https://doi.org/10.1007/978-1-62703-493-7_8

Cao, H., Wu, J., Wang, Y., Jiang, H., Zhang, T., Liu, X., Xu, Y., Liang, D., Gao, P., Sun, Y., Gifford, B., D'Ascenzo, M., Liu, X., Tellier, L. C. A. M., Yang, F., Tong, X., Chen, D., Zheng, J., Li, W., … Li, Y. (2013). An Integrated Tool to Study MHC Region: Accurate SNV Detection and HLA Genes Typing in Human MHC Region Using Targeted High-Throughput Sequencing. *PLoS ONE*, *8*(7). https://doi.org/10.1371/JOURNAL.PONE.0069388

Chen, J., Madireddi, S., Nagarkar, D., Migdal, M., vander Heiden, J., Chang, D., Mukhyala, K., Selvaraj, S., Kadel, E. E., Brauer, M. J., Mariathasan, S., Hunkapiller, J., Jhunjhunwala, S., Albert, M. L., & Hammer, C. (2021). In silico tools for accurate HLA and KIR inference from clinical sequencing data empower immunogenetics on individual-patient and population scales. *Briefings in Bioinformatics*, *22*(3), 1–11. https://doi.org/10.1093/BIB/BBAA223

Chowell, D., Morris, L. G. T., Grigg, C. M., Weber, J. K., Samstein, R. M., Makarov, V., Kuo, F., Kendall, S. M., Requena, D., Riaz, N., Greenbaum, B., Carroll, J., Garon, E., Hyman, D. M., Zehir, A., Solit, D., Berger, M., Zhou, R., Rizvi, N. A., & Chan, T. A. (2018). Patient HLA class I genotype influences cancer response to checkpoint blockade immunotherapy. *Science*, *359*(6375), 582–587. https://doi.org/10.1126/SCIENCE.AAO4572/SUPPL_FILE/AAO4572_CHOWELL_SM-TABLE-S9.XLSX

Claeys, A., Luijts, T., Marchal, K., & van den Eynden, J. (2021). Low immunogenicity of common cancer hot spot mutations resulting in false immunogenic selection signals. *PLOS Genetics*, *17*(2), e1009368. https://doi.org/10.1371/journal.pgen.1009368

Claeys, A., Staut, J., Merseburger, P., Marchal, K., & van den Eynden, J. (2022). Benchmark of tools for in silico prediction of MHC class I and class II genotypes from NGS data. *BioRxiv*, 2022.04.28.489842. https://doi.org/10.1101/2022.04.28.489842

Compagnone, M., Cifaldi, L., & Fruci, D. (2019). Regulation of ERAP1 and ERAP2 genes and their disfunction in human cancer. *Human Immunology*, *80*(5), 318–324. https://doi.org/10.1016/J.HUMIMM.2019.02.014

Contardi, E., Palmisano, G. L., Tazzari, P. L., Martelli, A. M., Falà, F., Fabbi, M., Kato, T., Lucarelli, E., Donati, D., Polito, L., Bolognesi, A., Ricci, F., Salvi, S., Gargaglione, V., Mantero, S., Alberghini, M., Ferrara, G. B., & Pistillo, M. P. (2005). CTLA-4 is constitutively expressed on tumor cells and can trigger apoptosis upon ligand interaction. *International Journal of Cancer*, *117*(4), 538–550. https://doi.org/10.1002/IJC.21155

Corchete, L. A., Rojas, E. A., Alonso-López, D., de Las Rivas, J., Gutiérrez, N. C., & Burguillo, F. J. (2020). Systematic comparison and assessment of RNA-seq procedures for gene expression quantitative analysis. *Scientific Reports 2020 10:1*, *10*(1), 1–15. https://doi.org/10.1038/s41598-020-76881-x

Corthay, A. (2009). How do Regulatory T Cells Work? *Scandinavian Journal of Immunology*, *70*(4), 326. https://doi.org/10.1111/J.1365-3083.2009.02308.X

Cristescu, R., Mogg, R., Ayers, M., Albright, A., Murphy, E., Yearley, J., Sher, X., Liu, X. Q., Lu, H., Nebozhyn, M., Zhang, C., Lunceford, J. K., Joe, A., Cheng, J., Webber, A. L., Ibrahim, N., Plimack, E. R., Ott, P. A., Seiwert, T. Y., … Kaufman, D. (2018). Pan-tumor genomic biomarkers for PD-1 checkpoint blockade–based immunotherapy. *Science (New York, N.Y.)*, *362*(6411). https://doi.org/10.1126/SCIENCE.AAR3593

de Kruijf, E. M., Sajet, A., van Nes, J. G. H., Natanov, R., Putter, H., Smit, V. T. H. B. M., Liefers, G. J., van den Elsen, P. J., van de Velde, C. J. H., & Kuppen, P. J. K. (2010). HLA-E and HLA-G Expression

in Classical HLA Class I-Negative Tumors Is of Prognostic Value for Clinical Outcome of Early Breast Cancer Patients. *The Journal of Immunology*, *185*(12), 7452–7459. https://doi.org/10.4049/jimmunol.1002629

Denzin, L. K., Sant'Angelo, D. B., Hammond, C., Surman, M. J., & Cresswell, P. (1997). Negative regulation by HLA-DO of MHC class II-restricted antigen processing. *Science (New York, N.Y.)*, *278*(5335), 106–109. https://doi.org/10.1126/SCIENCE.278.5335.106

Dilthey, A. T., Mentzer, A. J., Carapito, R., Cutland, C., Cereb, N., Madhi, S. A., Rhie, A., Koren, S., Bahram, S., McVean, G., & Phillippy, A. M. (2019). HLA*LA—HLA typing from linearly projected graph alignments. *Bioinformatics*, *35*(21), 4394–4396. https://doi.org/10.1093/BIOINFORMATICS/BTZ235

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., & Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, *29*(1), 15. https://doi.org/10.1093/BIOINFORMATICS/BTS635

Dunn, G. P., Old, L. J., & Schreiber, R. D. (2004). The Three Es of Cancer Immunoediting. *Annual Review of Immunology*, *22*(1), 329–360. https://doi.org/10.1146/annurev.immunol.22.012703.104803

Dupage, M., Mazumdar, C., Schmidt, L. M., Cheung, A. F., & Jacks, T. (2012). Expression of tumour-specific antigens underlies cancer immunoediting. *Nature*, *482*(7385), 405–409. https://doi.org/10.1038/nature10803

Edgerly, C. H., & Weimer, E. T. (2018). The past, present, and future of HLA typing in transplantation. *Methods in Molecular Biology*, *1802*, 1–10. https://doi.org/10.1007/978-1-4939-8546-3_1/COVER

Erlich, H., & Henry Erlich, C. (2012). HLA DNA typing: past, present, and future. *Tissue Antigens*, *80*(1), 1–11. https://doi.org/10.1111/J.1399-0039.2012.01881.X

Fabreti-Oliveira, R. A., Lasmar, M. F., Oliveira, C. K. F., Vale, E. M. G., & Nascimento, E. (2018). Genetic Mechanisms Involved in the Generation of HLA Alleles in Brazilians: Description and Comparison of HLA Alleles. *Transplantation Proceedings*, *50*(3), 835–840. https://doi.org/10.1016/J.TRANSPROCEED.2018.02.011

Frankish, A., Diekhans, M., Jungreis, I., Lagarde, J., Loveland, J. E., Mudge, J. M., Sisu, C., Wright, J. C., Armstrong, J., Barnes, I., Berry, A., Bignell, A., Boix, C., Sala, S. C., Cunningham, F., Domenico, T. di, Donaldson, S., Fiddes, I. T., Girón, C. G., … Flicek, P. (2021). GENCODE 2021. *Nucleic Acids Research*, *49*(D1), D916–D923. https://doi.org/10.1093/NAR/GKAA1087

Gabriel, C., Danzer, M., Hackl, C., Kopal, G., Hufnagl, P., Hofer, K., Polin, H., Stabentheiner, S., & Pröll, J. (2009). Rapid high-throughput human leukocyte antigen typing by massively parallel pyrosequencing for high-resolution allele identification. *Human Immunology*, *70*(11), 960–964. https://doi.org/10.1016/J.HUMIMM.2009.08.009

Gangaev, A., Rozeman, E. A., Rohaan, M. W., Isaeva, O. I., Patiwael, S., van den Berg, J. H., Ribas, A., Schadendorf, D., Schilling, B., Philips, D., Schumacher, T. N., Blank, C. U., Haanen, J. B. A. G., & Kvistborg, P. (2021). Differential effects of PD-1 and CTLA-4 blockade on the melanoma-reactive CD8 T cell response. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(43), e2102849118. https://doi.org/10.1073/PNAS.2102849118/SUPPL_FILE/PNAS.2102849118.SAPP.PDF

Garibyan, L., & Avashia, N. (2013). Polymerase Chain Reaction. *Journal of Investigative Dermatology*, *133*(3), 1–4. https://doi.org/10.1038/jid.2013.1

Garrido, G., Schrand, B., Rabasa, A., Levay, A., D'Eramo, F., Berezhnoy, A., Modi, S., Gefen, T., Marijt, K., Doorduijn, E., Dudeja, V., van Hall, T., & Gilboa, E. (2019). Tumor-targeted silencing of the peptide transporter TAP induces potent antitumor immunity. *Nature Communications*, *10*(1). https://doi.org/10.1038/S41467-019-11728-2

Gautreaux, M. D. (2017). Chapter 17 - Histocompatibility Testing in the Transplant Setting. In G. Orlando, G. Remuzzi, & D. F. Williams (Eds.), *Kidney Transplantation, Bioengineering and Regeneration* (pp. 223–234). Academic Press. https://www.sciencedirect.com/science/article/pii/B9780128017340000175

Gómez-Valenzuela, F., Escobar, E., Pérez-Tomás, R., & Montecinos, V. P. (2021). The Inflammatory Profile of the Tumor Microenvironment, Orchestrated by Cyclooxygenase-2, Promotes Epithelial-Mesenchymal Transition. *Frontiers in Oncology*, *11*. https://doi.org/10.3389/FONC.2021.686792

Gonzalez-Galarza, F. F., McCabe, A., Santos, E. J. M. dos, Jones, J., Takeshita, L., Ortega-Rivera, N. D., Cid-Pavon, G. M. D., Ramsbottom, K., Ghattaoraya, G., Alfirevic, A., Middleton, D., & Jones, A. R. (2020). Allele frequency net database (AFND) 2020 update: gold-standard data classification, open access genotype data and new query tools. *Nucleic Acids Research*, *48*(D1), D783–D788. https://doi.org/10.1093/NAR/GKZ1029

Gorer, P. A. (1937). The genetic and antigenic basis of tumour transplantation. *The Journal of Pathology and Bacteriology*, *44*(3), 691–697. https://doi.org/10.1002/PATH.1700440313

Gourraud, P. A., Khankhanian, P., Cereb, N., Yang, S. Y., Feolo, M., Maiers, M., Rioux, J. D., Hauser, S., & Oksenberg, J. (2014). HLA Diversity in the 1000 Genomes Dataset. *PLOS ONE*, *9*(7), e97282. https://doi.org/10.1371/JOURNAL.PONE.0097282

Greten, F. R., & Grivennikov, S. I. (2019). Inflammation and Cancer: Triggers, Mechanisms, and Consequences. *Immunity*, *51*(1), 27–41. https://doi.org/10.1016/J.IMMUNI.2019.06.025

Halenius, A., Gerke, C., & Hengel, H. (2014). Classical and non-classical MHC I molecule manipulation by human cytomegalovirus: so many targets—but how many arrows in the quiver? *Cellular & Molecular Immunology 2015 12:2*, *12*(2), 139–153. https://doi.org/10.1038/cmi.2014.105

Hanahan, D. (2022). Hallmarks of Cancer: New Dimensions. *Cancer Discovery*, *12*(1), 31–46. https://doi.org/10.1158/2159-8290.CD-21-1059

Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of cancer: The next generation. *Cell*, *144*(5), 646–674. https://doi.org/10.1016/j.cell.2011.02.013

Havel, J. J., Chowell, D., & Chan, T. A. (2019). The evolving landscape of biomarkers for checkpoint inhibitor immunotherapy. *Nature Reviews Cancer 2019 19:3*, *19*(3), 133–150. https://doi.org/10.1038/s41568-019-0116-x

Hayashi, S., Moriyama, T., Yamaguchi, R., Mizuno, S., Komura, M., Miyano, S., Nakagawa, H., & Imoto, S. (2019). ALPHLARD-NT: Bayesian Method for Human Leukocyte Antigen Genotyping and Mutation Calling through Simultaneous Analysis of Normal and Tumor Whole-Genome Sequence Data. *Journal of Computational Biology*, *26*(9), 923–937. https://doi.org/10.1089/cmb.2018.0224

Hewitt, E. W. (2003). The MHC class I antigen presentation pathway: strategies for viral immune evasion. *Immunology*, *110*(2), 163. https://doi.org/10.1046/J.1365-2567.2003.01738.X

Horton, R., Wilming, L., Rand, V., Lovering, R. C., Bruford, E. A., Khodiyar, V. K., Lush, M. J., Povey, S., Talbot, C. C., Wright, M. W., Wain, H. M., Trowsdale, J., Ziegler, A., & Beck, S. (2004). Gene map of the extended human MHC. *Nature Reviews Genetics 2004 5:12*, *5*(12), 889–899. https://doi.org/10.1038/nrg1489

Hosomichi, K., Shiina, T., Tajima, A., & Inoue, I. (2015). The impact of next-generation sequencing technologies on HLA research. *Journal of Human Genetics 2015 60:11*, *60*(11), 665–673. https://doi.org/10.1038/jhg.2015.102

Howell, W. M., Carter, V., & Clark, B. (2010). The HLA system: immunobiology, HLA typing, antibody screening and crossmatching techniques. *Journal of Clinical Pathology*, *63*(5), 387–390. https://doi.org/10.1136/JCP.2009.072371

Huang, Y., Yang, J., Ying, D., Zhang, Y., Shotelersuk, V., Hirankarn, N., Sham, P. C., Lau, Y. L., & Yang, W. (2015). HLAreporter: A tool for HLA typing from next generation sequencing data. *Genome Medicine*, *7*(1), 1–12. https://doi.org/10.1186/S13073-015-0145-3

Hurley, C. K. (2021). Naming HLA diversity: A review of HLA nomenclature. *Human Immunology*, *82*(7), 457–465. https://doi.org/10.1016/J.HUMIMM.2020.03.005

Jackman, S. D., Vandervalk, B. P., Mohamadi, H., Chu, J., Yeo, S., Hammond, S. A., Jahesh, G., Khan, H., Coombe, L., Warren, R. L., & Birol, I. (2017). ABySS 2.0: Resource-efficient assembly of large genomes using a Bloom filter. *Genome Research*, *27*(5), 768–777. https://doi.org/10.1101/GR.214346.116/-/DC1

Jhunjhunwala, S., Hammer, C., & Delamarre, L. (2021). Antigen presentation in cancer: insights into tumour immunogenicity and immune evasion. *Nature Reviews Cancer 2021 21:5*, *21*(5), 298–312. https://doi.org/10.1038/s41568-021-00339-z

Jia, X., Han, B., Onengut-Gumuscu, S., Chen, W. M., Concannon, P. J., Rich, S. S., Raychaudhuri, S., & de Bakker, P. I. W. (2013). Imputing Amino Acid Polymorphisms in Human Leukocyte Antigens. *PLOS ONE*, *8*(6), e64683. https://doi.org/10.1371/JOURNAL.PONE.0064683

Jiang, Y., Jiang, Y., Wang, S., Zhang, Q., & Ding, X. (2019). Optimal sequencing depth design for whole genome re-sequencing in pigs. *BMC Bioinformatics*, *20*(1), 556. https://doi.org/10.1186/s12859-019-3164-z

Johnson, D. K., Magoffin, W., Myers, S. J., Finnell, J. G., Hancock, J. C., Orton, T. S., Persaud, S. P., Christensen, K. A., & Weber, K. S. (2021). CD4 Inhibits Helper T Cell Activation at Lower Affinity Threshold for Full-Length T Cell Receptors Than Single Chain Signaling Constructs. *Frontiers in Immunology*, *11*, 3473. https://doi.org/10.3389/FIMMU.2020.561889/BIBTEX

Jurgens, S. J., Choi, S. H., Morrill, V. N., Chaffin, M., Pirruccello, J. P., Halford, J. L., Weng, L. C., Nauffal, V., Roselli, C., Hall, A. W., Oetjens, M. T., Lagerman, B., vanMaanen, D. P., Abecasis, G., Bai, X., Balasubramanian, S., Baras, A., Beechert, C., Boutkov, B., … Ellinor, P. T. (2022). Analysis of rare genetic variation underlying cardiometabolic diseases and traits among 200,000 individuals in the UK Biobank. *Nature Genetics 2022 54:3*, *54*(3), 240–250. https://doi.org/10.1038/s41588-021-01011-w

Ka, S., Lee, S., Hong, J., Cho, Y., Sung, J., Kim, H. N., Kim, H. L., & Jung, J. (2017). HLAscan: Genotyping of the HLA region using next-generation sequencing data. *BMC Bioinformatics*, *18*(1), 1–11. https://doi.org/10.1186/S12859-017-1671-3

Kaslow, R. A., Carrington, M., Apple, R., Park, L., Muñoz, A., Saah, A. J., Goedert, J. J., Winkler, C., O'Brien, S. J., Rinaldo, C., Detels, R., Blattner, W., Phair, J., Erlich, H., & Mann, D. L. (1996). Influence of combinations of human major histocompatibility complex genes on the course of HIV–1 infection. *Nature Medicine 1996 2:4*, *2*(4), 405–411. https://doi.org/10.1038/nm0496-405

Kawaguchi, S., Higasa, K., Shimizu, M., Yamada, R., & Matsuda, F. (2017). HLA-HD: An accurate HLA typing algorithm for next-generation sequencing data. *Human Mutation*, *38*(7), 788–797. https://doi.org/10.1002/HUMU.23230

Kern, R., & Panis, C. (2021). CTLA-4 Expression and Its Clinical Significance in Breast Cancer. *Archivum Immunologiae et Therapiae Experimentalis*, *69*(1). https://doi.org/10.1007/S00005-021-00618-5

Kim, H. J., & Pourmand, N. (2013). HLA Haplotyping from RNA-seq Data Using Hierarchical Read Weighting. *PLOS ONE*, *8*(6), e67885. https://doi.org/10.1371/JOURNAL.PONE.0067885

Kiyotani, K., Mai, T. H., & Nakamura, Y. (2016). Comparison of exome-based HLA class I genotyping tools: identification of platform-specific genotyping errors. *Journal of Human Genetics 2017 62:3*, *62*(3), 397–405. https://doi.org/10.1038/jhg.2016.141

Klasberg, S., Surendranath, V., Lange, V., & Schöfl, G. (2019). Bioinformatics Strategies, Challenges, and Opportunities for Next Generation Sequencing-Based HLA Genotyping. *Transfusion Medicine and Hemotherapy*, *46*(5), 312–325. https://doi.org/10.1159/000502487

Klein, J. (2001). George Snell's First Foray Into the Unexplored Territory of the Major Histocompatibility Complex. *Genetics*, *159*(2), 435–439. https://doi.org/10.1093/GENETICS/159.2.435

Klein, J., & Sato, A. (2000). The HLA System. *New England Journal of Medicine*, *343*(10), 702–709. https://doi.org/10.1056/NEJM200009073431006

Krijgsman, D., Roelands, J., Hendrickx, W., Bedognetti, D., & Kuppen, P. J. K. (2020). HLA-G: A New Immune Checkpoint in Cancer? *International Journal of Molecular Sciences 2020, Vol. 21, Page 4528*, *21*(12), 4528. https://doi.org/10.3390/IJMS21124528

Lange, V., Böhme, I., Hofmann, J., Lang, K., Sauter, J., Schöne, B., Paul, P., Albrecht, V., Andreas, J. M., Baier, D. M., Nething, J., Ehninger, U., Schwarzelt, C., Pingel, J., Ehninger, G., & Schmidt, A. H. (2014). Cost-efficient high-throughput HLA typing by MiSeq amplicon sequencing. *BMC Genomics*, *15*(1), 1–11. https://doi.org/10.1186/1471-2164-15-63/FIGURES/8

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357. https://doi.org/10.1038/NMETH.1923

Lappalainen, T., Sammeth, M., Friedländer, M. R., 'T Hoen, P. A. C., Monlong, J., Rivas, M. A., Gonzàlez-Porta, M., Kurbatova, N., Griebel, T., Ferreira, P. G., Barann, M., Wieland, T., Greger, L., van Iterson, M., Almlöf, J., Ribeca, P., Pulyakhina, I., Esser, D., Giger, T., … Dermitzakis, E. T. (2013). Transcriptome and genome sequencing uncovers functional variation in humans. *Nature 2013 501:7468*, *501*(7468), 506–511. https://doi.org/10.1038/nature12531

Larkin, J., Chiarion-Sileni, V., Gonzalez, R., Grob, J.-J., Rutkowski, P., Lao, C. D., Cowey, C. L., Schadendorf, D., Wagstaff, J., Dummer, R., Ferrucci, P. F., Smylie, M., Hogg, D., Hill, A., Márquez-Rodas, I., Haanen, J., Guidoboni, M., Maio, M., Schöffski, P., … Wolchok, J. D. (2019). Five-Year Survival with Combined Nivolumab and Ipilimumab in Advanced Melanoma. *New England Journal of Medicine*, *381*(16), 1535–1546. https://doi.org/10.1056/NEJMOA1910836/SUPPL_FILE/NEJMOA1910836_DATA-SHARING.PDF

Lee, H., & Kingsford, C. (2018a). Accurate assembly and typing of HLA using a graph-guided assembler kourami. *Methods in Molecular Biology*, *1802*, 235–247. https://doi.org/10.1007/978-1-4939-8546-3_17/TABLES/2

Lee, H., & Kingsford, C. (2018b). Kourami: Graph-guided assembly for novel human leukocyte antigen allele discovery. *Genome Biology*, *19*(1), 1–16. https://doi.org/10.1186/S13059-018-1388-2

Lee, M., Seo, J. H., Song, S., Song, I. H., Kim, S. Y., Kim, Y. A., Gong, G., Kim, J. E., & Lee, H. J. (2021). A New Human Leukocyte Antigen Typing Algorithm Combined With Currently Available Genotyping Tools Based on Next-Generation Sequencing Data and Guidelines to Select the Most Likely Human Leukocyte Antigen Genotype. *Frontiers in Immunology*, *12*, 4080. https://doi.org/10.3389/FIMMU.2021.688183

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, *25*(14), 1754–1760. https://doi.org/10.1093/BIOINFORMATICS/BTP324

Li, X., Zhou, C., Chen, K., Huang, B., Liu, Q., & Ye, H. (2021). Benchmarking HLA genotyping and clarifying HLA impact on survival in tumor immunotherapy. *Molecular Oncology*, *15*(7), 1764–1782. https://doi.org/10.1002/1878-0261.12895

Lind, C., Ferriola, D., Mackiewicz, K., Heron, S., Rogers, M., Slavich, L., Walker, R., Hsiao, T., McLaughlin, L., D'Arcy, M., Gai, X., Goodridge, D., Sayer, D., & Monos, D. (2010). Next-generation sequencing: the solution for high-resolution, unambiguous human leukocyte antigen typing. *Human Immunology*, *71*(10), 1033–1042. https://doi.org/10.1016/J.HUMIMM.2010.06.016

Liu, C., Yang, X., Duffy, B., Mohanakumar, T., Mitra, R. D., Zody, M. C., & Pfeifer, J. D. (2013). ATHLATES: accurate typing of human leukocyte antigen through exome sequencing. *Nucleic Acids Research*, *41*(14), e142–e142. https://doi.org/10.1093/NAR/GKT481

Liu, D., Schilling, B., Liu, D., Sucker, A., Livingstone, E., Jerby-Amon, L., Zimmer, L., Gutzmer, R., Satzger, I., Loquai, C., Grabbe, S., Vokes, N., Margolis, C. A., Conway, J., He, M. X., Elmarakeby, H., Dietlein, F., Miao, D., Tracy, A., … Schadendorf, D. (2019). Integrative molecular and clinical modeling of clinical outcomes to PD1 blockade in patients with metastatic melanoma. *Nature Medicine 2019 25:12*, *25*(12), 1916–1927. https://doi.org/10.1038/s41591-019-0654-5

Liu, P., Yao, M., Gong, Y., Song, Y., Chen, Y., Ye, Y., Liu, X., Li, F., Dong, H., Meng, R., Chen, H., & Zheng, A. (2021). Benchmarking the Human Leukocyte Antigen Typing Performance of Three Assays and Seven Next-Generation Sequencing-Based Algorithms. *Frontiers in Immunology*, *12*, 840. https://doi.org/10.3389/FIMMU.2021.652258/BIBTEX

Liu, Z., Huang, C. J., Huang, Y. H., Pan, M. H., Lee, M. H., Yu, K. J., Pfeiffer, R. M., Viard, M., Yuki, Y., Gao, X., Carrington, M., Chen, C. J., Hildesheim, A., & Yang, H. I. (2021). HLA Zygosity Increases Risk of Hepatitis B Virus-Associated Hepatocellular Carcinoma. *The Journal of Infectious Diseases*, *224*(10), 1796. https://doi.org/10.1093/INFDIS/JIAB207

Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., He, G., Chen, Y., Pan, Q., Liu, Y., Tang, J., Wu, G., Zhang, H., Shi, Y., Liu, Y., Yu, C., Wang, B., Lu, Y., Han, C., … Wang, J. (2012). SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *GigaScience*, *1*(1), 18. https://doi.org/10.1186/2047-217X-1-18

Mardis, E. R. (2008). The impact of next-generation sequencing technology on genetics. *Trends in Genetics : TIG*, *24*(3), 133–141. https://doi.org/10.1016/J.TIG.2007.12.007

Markov, P. v., & Pybus, O. G. (2015). Evolution and Diversity of the Human Leukocyte Antigen(HLA). *Evolution, Medicine, and Public Health*, *2015*(1), 1. https://doi.org/10.1093/EMPH/EOU033

Marsh, S. G. E. (2022). Nomenclature for factors of the HLA system, update April, May and June 2022. *International Journal of Immunogenetics*, *49*(4), 279–315. https://doi.org/10.1111/IJI.12591

Marsh, S. G. E., Albert, E. D., Bodmer, W. F., Bontrop, R. E., Dupont, B., Erlich, H. A., Fernández-Viña, M., Geraghty, D. E., Holdsworth, R., Hurley, C. K., Lau, M., Lee, K. W., MacH, B., Maiers, M., Mayr, W. R., Müller, C. R., Parham, P., Petersdorf, E. W., Sasazuki, T., … Trowsdale, J. (2010).

Nomenclature for factors of the HLA system, 2010. *Tissue Antigens*, *75*(4), 291. https://doi.org/10.1111/J.1399-0039.2010.01466.X

Marty, R., Kaabinejadian, S., Rossell, D., Slifker, M. J., van de Haar, J., Engin, H. B., de Prisco, N., Ideker, T., Hildebrand, W. H., Font-Burgada, J., & Carter, H. (2017). MHC-I Genotype Restricts the Oncogenic Mutational Landscape. *Cell*, *171*(6), 1272-1283.e15. https://doi.org/10.1016/J.CELL.2017.09.050

Marty, R., Thompson, W. K., Salem, R. M., Zanetti, M., & Carter, H. (2018). Evolutionary Pressure against MHC Class II Binding Cancer Mutations. *Cell*, *175*(2), 416-428.e13. https://doi.org/10.1016/j.cell.2018.08.048

Matern, B. M., Olieslagers, T. I., Voorter, C. E. M., Groeneweg, M., & Tilanus, M. G. J. (2020). Insights into the polymorphism in HLA-DRA and its evolutionary relationship with HLA haplotypes. *HLA*, *95*(2), 117–127. https://doi.org/10.1111/TAN.13730

Matey-Hernandez, M. L., Brunak, S., & Izarzugaza, J. M. G. (2018). Benchmarking the HLA typing performance of Polysolver and Optitype in 50 Danish parental trios. *BMC Bioinformatics*, *19*(1), 1–12. https://doi.org/10.1186/S12859-018-2239-6

McCombie, W. R., McPherson, J. D., & Mardis, E. R. (2019). Next-Generation Sequencing Technologies. *Cold Spring Harbor Perspectives in Medicine*, *9*(11), a036798. https://doi.org/10.1101/cshperspect.a036798

Mellins, E. D., & Stern, L. J. (2014). HLA-DM and HLA-DO, key regulators of MHC-II processing and presentation. *Current Opinion in Immunology*, *26*(1), 115–122. https://doi.org/10.1016/J.COI.2013.11.005

Metzker, M. L. (2009). Sequencing technologies — the next generation. *Nature Reviews Genetics 2010 11:1*, *11*(1), 31–46. https://doi.org/10.1038/nrg2626

Milner, C. M. (2001). Genetic organization of the human MHC class III region. *Frontiers in Bioscience*, *6*(1), d914. https://doi.org/10.2741/Milner

Motzer, R. J., Penkov, K., Haanen, J., Rini, B., Albiges, L., Campbell, M. T., Venugopal, B., Kollmannsberger, C., Negrier, S., Uemura, M., Lee, J. L., Vasiliev, A., Miller, W. H., Gurney, H., Schmidinger, M., Larkin, J., Atkins, M. B., Bedke, J., Alekseev, B., … Choueiri, T. K. (2019). Avelumab plus Axitinib versus Sunitinib for Advanced Renal-Cell Carcinoma. *New England Journal of Medicine*, *380*(12), 1103–1115. https://doi.org/10.1056/NEJMOA1816047/SUPPL_FILE/NEJMOA1816047_DATA-SHARING.PDF

Naranbhai, V., Viard, M., Dean, M., Groha, S., Braun, D. A., Labaki, C., Shukla, S. A., Yuki, Y., Shah, P., Chin, K., Wind-Rotolo, M., Mu, X. J., Robbins, P. B., Gusev, A., Choueiri, T. K., Gulley, J. L., & Carrington, M. (2022). HLA-A*03 and response to immune checkpoint blockade in cancer: an epidemiological biomarker study. *The Lancet. Oncology*, *23*(1), 172–184. https://doi.org/10.1016/S1470-2045(21)00582-9

Nariai, N., Kojima, K., Saito, S., Mimori, T., Sato, Y., Kawai, Y., Yamaguchi-Kabata, Y., Yasuda, J., & Nagasaki, M. (2015). HLA-VBSeq: Accurate HLA typing at full resolution from whole-genome sequencing data. *BMC Genomics*, *16*(2), 1–6. https://doi.org/10.1186/1471-2164-16-S2-S7

Neefjes, J., Jongsma, M. L. M., Paul, P., & Bakke, O. (2011). Towards a systems understanding of MHC class i and MHC class II antigen presentation. *Nature Reviews Immunology*, *11*(12), 823–836. https://doi.org/10.1038/nri3084

Ng, P. C., & Kirkness, E. F. (2010). Whole genome sequencing. In *Methods in Molecular Biology* (Vol. 628, pp. 215–226). Humana Press Inc. https://doi.org/10.1007/978-1-60327-367-1_12/FIGURES/2

Ng, S. B., Turner, E. H., Robertson, P. D., Flygare, S. D., Bigham, A. W., Lee, C., Shaffer, T., Wong, M., Bhattacharjee, A., Eichler, E. E., Bamshad, M., Nickerson, D. A., & Shendure, J. (2009). Targeted capture and massively parallel sequencing of 12 human exomes. *Nature*, *461*(7261), 272–276. https://doi.org/10.1038/NATURE08250

Nielsen, M., Lundegaard, C., Blicher, T., Lamberth, K., Harndahl, M., Justesen, S., Røder, G., Peters, B., Sette, A., Lund, O., & Buus, S. (2007). NetMHCpan, a Method for Quantitative Predictions of Peptide Binding to Any HLA-A and -B Locus Protein of Known Sequence. *PLOS ONE*, *2*(8), e796. https://doi.org/10.1371/journal.pone.0000796

Orenbuch, R., Filip, I., Comito, D., Shaman, J., Pe'Er, I., & Rabadan, R. (2020). arcasHLA: high-resolution HLA typing from RNAseq. *Bioinformatics*, *36*(1), 33–40. https://doi.org/10.1093/BIOINFORMATICS/BTZ474

Orenbuch, R., Filip, I., & Rabadan, R. (2020). HLA typing from RNA sequencing and applications to cancer. *Methods in Molecular Biology*, *2120*, 71–92. https://doi.org/10.1007/978-1-0716-0327-7_5/FIGURES/6

Parla, J. S., Iossifov, I., Grabill, I., Spector, M. S., Kramer, M., & McCombie, W. R. (2011). A comparative analysis of exome capture. *Genome Biology*, *12*(9), 1–17. https://doi.org/10.1186/GB-2011-12-9-R97/FIGURES/8

Pasinetti, N., Pirtoli, L., Buglione, M., Triggiani, L., Borghetti, P., Tini, P., Maria, S., Pasinetti, M. N., Buglione, • M, Triggiani, • L, Borghetti, P., Pirtoli, L., Tini, P., & Magrini, S. M. (2016). From Molecular to Clinical Radiation Biology of Glioblastoma. In *Current Clinical Pathology* (pp. 275–292). Humana Press, Cham. https://doi.org/10.1007/978-3-319-28305-0_17

Patro, R., Duggal, G., Love, M. I., Irizarry, R. A., & Kingsford, C. (2017). Salmon: fast and bias-aware quantification of transcript expression using dual-phase inference. *Nature Methods*, *14*(4), 417. https://doi.org/10.1038/NMETH.4197

Patro, R., Mount, S. M., & Kingsford, C. (2014). Sailfish enables alignment-free isoform quantification from RNA-seq reads using lightweight algorithms. *Nature Biotechnology*, *32*(5), 462–464. https://doi.org/10.1038/NBT.2862

Paz-Ares, L., Ciuleanu, T. E., Cobo, M., Schenker, M., Zurawski, B., Menezes, J., Richardet, E., Bennouna, J., Felip, E., Juan-Vidal, O., Alexandru, A., Sakai, H., Lingua, A., Salman, P., Souquet, P. J., de Marchi, P., Martin, C., Pérol, M., Scherpereel, A., … Reck, M. (2021). First-line nivolumab plus ipilimumab combined with two cycles of chemotherapy in patients with non-small-cell lung cancer (CheckMate 9LA): an international, randomised, open-label, phase 3 trial. *The Lancet Oncology*, *22*(2), 198–211. https://doi.org/10.1016/S1470-2045(20)30641-0

Pedersen, B. S., & Quinlan, A. R. (2018). Mosdepth: quick coverage calculation for genomes and exomes. *Bioinformatics*, *34*(5), 867–868. https://doi.org/10.1093/BIOINFORMATICS/BTX699

Philipps, C., McMillan, M., Flood, P. M., Murphy, D. B., Forman, J., Lancki, D., Womack, J. E., Goodenow, R. S., & Schreiber, H. (1985). Identification of a unique tumor-specific antigen as a novel class I major histocompatibility molecule. *Proceedings of the National Academy of Sciences of the United States of America*, *82*(15), 5140–5144. https://doi.org/10.1073/PNAS.82.15.5140

Pruis, M. A., Groenendijk, F. H., Badloe, K. S., van Puffelen, A., Robbrecht, D., Dinjens, W. N. M., Sleijfer, S., Dingemans, A. M. C., von der Thüsen, J. H., Roepman, P., & Lolkema, M. P. (2022). Personalised selection of experimental treatment in patients with advanced solid cancer is feasible using whole-genome sequencing. *British Journal of Cancer 2022 127:4*, *127*(4), 776–783. https://doi.org/10.1038/s41416-022-01841-3

Raghavan, M., & Geng, J. (2015). HLA-B polymorphisms and intracellular assembly modes. *Molecular Immunology*, *68*(2), 89–93. https://doi.org/10.1016/J.MOLIMM.2015.07.007

Reck, M., Rodríguez-Abreu, D., Robinson, A. G., Hui, R., Csőszi, T., Fülöp, A., Gottfried, M., Peled, N., Tafreshi, A., Cuffe, S., O'Brien, M., Rao, S., Hotta, K., Leiby, M. A., Lubiniecki, G. M., Shentu, Y., Rangwala, R., & Brahmer, J. R. (2016). Pembrolizumab versus Chemotherapy for PD-L1–Positive Non–Small-Cell Lung Cancer. *New England Journal of Medicine*, *375*(19), 1823–1833. https://doi.org/10.1056/NEJMOA1606774/SUPPL_FILE/NEJMOA1606774_DISCLOSURES.PDF

Reinhold, W. C., Varma, S., Sunshine, M., Elloumi, F., Ofori-Atta, K., Lee, S., Trepel, J. B., Meltzer, P. S., Doroshow, J. H., & Pommier, Y. (2019). RNA Sequencing of the NCI-60: Integration into CellMiner and CellMiner CDB. *Cancer Research*, *79*(13), 3514–3524. https://doi.org/10.1158/0008-5472.CAN-18-2047

Riaz, N., Havel, J. J., Makarov, V., Desrichard, A., Urba, W. J., Sims, J. S., Hodi, F. S., Martín-Algarra, S., Mandal, R., Sharfman, W. H., Bhatia, S., Hwu, W. J., Gajewski, T. F., Slingluff, C. L., Chowell, D., Kendall, S. M., Chang, H., Shah, R., Kuo, F., … Chan, T. A. (2017). Tumor and Microenvironment Evolution during Immunotherapy with Nivolumab. *Cell*, *171*(4), 934-949.e16. https://doi.org/10.1016/J.CELL.2017.09.028

Richmond, C. (2009). Jean Dausset. *The Lancet*, *374*(9698), 1324. https://doi.org/10.1016/s0140-6736(09)61813-4

Robinson, J., Halliwell, J. A., Hayhurst, J. D., Flicek, P., Parham, P., & Marsh, S. G. E. (2015). The IPD and IMGT/HLA database: Allele variant databases. *Nucleic Acids Research*, *43*(D1), D423–D431. https://doi.org/10.1093/nar/gku1161

Rooney, M. S., Shukla, S. A., Wu, C. J., Getz, G., & Hacohen, N. (2015). Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell*, *160*(1–2), 48–61. https://doi.org/10.1016/j.cell.2014.12.033

Sabbatino, F., Liguori, L., Polcaro, G., Salvato, I., Caramori, G., Salzano, F. A., Casolaro, V., Stellato, C., Col, J. D., & Pepe, S. (2020). Role of Human Leukocyte Antigen System as A Predictive Biomarker for Checkpoint-Based Immunotherapy in Cancer Patients. *International Journal of Molecular Sciences 2020, Vol. 21, Page 7295*, *21*(19), 7295. https://doi.org/10.3390/IJMS21197295

Safaeian, M., Johnson, L. G., Yu, K., Wang, S. S., Gravitt, P. E., Hansen, J. A., Carrington, M., Schwartz, S. M., Gao, X., Hildesheim, A., & Madeleine, M. M. (2014). Human leukocyte antigen class I and II alleles and cervical adenocarcinoma. *Frontiers in Oncology*, *4 JUN*. https://doi.org/10.3389/FONC.2014.00119/ABSTRACT

Schumacher, T. N., & Schreiber, R. D. (2015). Neoantigens in cancer immunotherapy. *Science*, *348*(6230), 69–74. https://doi.org/10.1126/SCIENCE.AAA4971/ASSET/9BC707C7-9B25-479A-8F2D-54AE27164ADA/ASSETS/GRAPHIC/348_69_F4.JPEG

Seal, R. L., Wright, M. W., Gray, K. A., & Bruford, E. A. (2013). Vive la différence: Naming structural variants in the human reference genome. *Human Genomics*, *7*(1), 1–3. https://doi.org/10.1186/1479-7364-7-12/FIGURES/1

Segawa, H., Kukita, Y., & Kato, K. (2017). HLA genotyping by next-generation sequencing of complementary DNA. *BMC Genomics*, *18*(1), 1–12. https://doi.org/10.1186/S12864-017-4300-7/FIGURES/5

Seidel, J. A., Otsuka, A., & Kabashima, K. (2018). Anti-PD-1 and anti-CTLA-4 therapies in cancer: Mechanisms of action, efficacy, and limitations. *Frontiers in Oncology*, *8*(MAR), 86. https://doi.org/10.3389/FONC.2018.00086/BIBTEX

Shiina, T., Hosomichi, K., Inoko, H., & Kulski, J. K. (2009). The HLA genomic loci map: expression, interaction, diversity and disease. *Journal of Human Genetics 2009 54:1*, *54*(1), 15–39. https://doi.org/10.1038/jhg.2008.5

Shiina, T., Suzuki, S., Kulski, J. K., & Inoko, H. (2018). Super high resolution for single molecule-sequence-based typing of classical HLA loci using ion torrent PGM. *Methods in Molecular Biology*, *1802*, 115–133. https://doi.org/10.1007/978-1-4939-8546-3_8/TABLES/6

Shukla, S. A., Rooney, M. S., Rajasagi, M., Tiao, G., Dixon, P. M., Lawrence, M. S., Stevens, J., Lane, W. J., Dellagatta, J. L., Steelman, S., Sougnez, C., Cibulskis, K., Kiezun, A., Brusic, V., Wu, C. J., & Getz, G. (2015). Comprehensive analysis of cancer-associated somatic mutations in class I HLA genes. *Nature Biotechnology*, *33*(11), 1152. https://doi.org/10.1038/NBT.3344

Shyamala, V., & Ames, G. F. L. (1989). Genome walking by single-specific-primer polymerase chain reaction: SSP-PCR. *Gene*, *84*(1), 1–8. https://doi.org/10.1016/0378-1119(89)90132-7

Sibinga, C. S., Klein, H. G., & Red Cross Blood Bank Noord-Nederland. (2000). *Molecular biology in blood transfusion : proceedings of the Twenty-Fourth International Symposium on Blood Transfusion, Groningen 1999*. Kluwer Academic Publishers. https://books.google.com/books/about/Molecular_Biology_in_Blood_Transfusion.html?hl=nl&id=wPi4_qqN98sC

Simmonds, M., & Gough, S. (2009). The HLA Region and Autoimmune Disease: Associations and Mechanisms of Action. *Current Genomics*, *8*(7), 453–465. https://doi.org/10.2174/138920207783591690

Sims, D., Sudbery, I., Ilott, N. E., Heger, A., & Ponting, C. P. (2014). Sequencing depth and coverage: key considerations in genomic analyses. *Nature Reviews Genetics*, *15*(2), 121–132. https://doi.org/10.1038/nrg3642

Snell, G. D. (1948). Methods for the study of histocompatibility genes. *Journal of Genetics*, *49*(2), 87–108. https://doi.org/10.1007/BF02986826

Sommer, S. (2005). The importance of immune gene variability (MHC) in evolutionary ecology and conservation. *Frontiers in Zoology*, *2*(1), 1–18. https://doi.org/10.1186/1742-9994-2-16/TABLES/1

Stark, R., Grzelak, M., & Hadfield, J. (2019). RNA sequencing: the teenage years. *Nature Reviews Genetics*, *20*(11), 631–656. https://doi.org/10.1038/s41576-019-0150-2

Stratikos, E., Stamogiannos, A., Zervoudi, E., & Fruci, D. (2014). A Role for Naturally Occurring Alleles of Endoplasmic Reticulum Aminopeptidases in Tumor Immunity and Cancer Pre-Disposition. *Frontiers in Oncology*, *4*(NOV). https://doi.org/10.3389/FONC.2014.00363

Su, X. Z., Wu, Y., Sifri, C. D., & Wellems, T. E. (1996). Reduced Extension Temperatures Required for PCR Amplification of Extremely A+T-rich DNA. *Nucleic Acids Research*, *24*(8), 1574–1575. https://doi.org/10.1093/NAR/24.8.1574

Sun, Z., Chen, F., Meng, F., Wei, J., & Liu, B. (2017). MHC class II restricted neoantigen: A promising target in tumor immunotherapy. *Cancer Letters*, *392*, 17–25. https://doi.org/10.1016/J.CANLET.2016.12.039

Sverchkova, A., Anzar, I., Stratford, R., & Clancy, T. (2019). Improved HLA typing of Class I and Class II alleles from next-generation sequencing data. *HLA*, *94*(6), 504–513. https://doi.org/10.1111/TAN.13685

Szolek, A., Schubert, B., Mohr, C., Sturm, M., Feldhahn, M., & Kohlbacher, O. (2014). OptiType: precision HLA typing from next-generation sequencing data. *Bioinformatics*, *30*(23), 3310–3316. https://doi.org/10.1093/bioinformatics/btu548

Teer, J. K., & Mullikin, J. C. (2010). Exome sequencing: the sweet spot before whole genomes. *Human Molecular Genetics*, *19*(R2), R145–R151. https://doi.org/10.1093/hmg/ddq333

Thorsby, E., & Thorsby, E. (2009). A short history of HLA. *Tissue Antigens*, *74*(2), 101–116. https://doi.org/10.1111/J.1399-0039.2009.01291.X

Tiwary, B. K. (2022). Next-Generation Sequencing. In *Bioinformatics and Computational Biology* (pp. 117–135). Springer, Singapore. https://doi.org/10.1007/978-981-16-4241-8_7

Trowsdale, J. (1993). Genomic structure and function in the MHC. *Trends in Genetics*, *9*(4), 117–122. https://doi.org/10.1016/0168-9525(93)90205-V

van den Eynden, J., Jiménez-Sánchez, A., Miller, M. L., & Larsson, E. (2019). Lack of detectable neoantigen depletion signals in the untreated cancer genome. *Nature Genetics*, *51*(12), 1741–1748. https://doi.org/10.1038/s41588-019-0532-6

Vogelstein, B., Papadopoulos, N., Velculescu, V. E., Zhou, S., Diaz, L. A., & Kinzler, K. W. (2013). Cancer genome landscapes. *Science*, *340*(6127), 1546–1558. https://doi.org/10.1126/science.1235122

Wang, Y. Y., Mimori, T., Khor, S. S., Gervais, O., Kawai, Y., Hitomi, Y., Tokunaga, K., & Nagasaki, M. (2019). HLA-VBSeq v2: improved HLA calling accuracy with full-length Japanese class-I panel. *Human Genome Variation 2019 6:1*, *6*(1), 1–5. https://doi.org/10.1038/s41439-019-0061-y

Warren, R. L., Choe, G., Freeman, D. J., Castellarin, M., Munro, S., Moore, R., & Holt, R. A. (2012). Derivation of HLA types from shotgun sequence datasets. *Genome Medicine*, *4*(12), 1–8. https://doi.org/10.1186/GM396

Wells, D. K., van Buuren, M. M., Dang, K. K., Hubbard-Lucey, V. M., Sheehan, K. C. F., Campbell, K. M., Lamb, A., Ward, J. P., Sidney, J., Blazquez, A. B., Rech, A. J., Zaretsky, J. M., Comin-Anduix, B., Ng, A. H. C., Chour, W., Yu, T. v., Rizvi, H., Chen, J. M., Manning, P., … Defranoux, N. A. (2020). Key Parameters of Tumor Epitope Immunogenicity Revealed Through a Consortium Approach Improve Neoantigen Prediction. *Cell*, *183*(3), 818-834.e13. https://doi.org/10.1016/J.CELL.2020.09.015

Welsh, R. A., & Sadegh-Nasseri, S. (2020). The love and hate relationship of HLA-DM/DO in the selection of immunodominant epitopes. *Current Opinion in Immunology*, *64*, 117–123. https://doi.org/10.1016/J.COI.2020.05.007

Williams, A. G., Thomas, S., Wyman, S. K., & Holloway, A. K. (2014). RNA-seq Data: Challenges in and Recommendations for Experimental Design and Analysis. *Current Protocols in Human Genetics*, *83*(1), 11.13.1-11.13.20. https://doi.org/10.1002/0471142905.HG1113S83

Wittig, M., Anmarkrud, J. A., Kässens, J. C., Koch, S., Forster, M., Ellinghaus, E., Hov, J. R., Sauer, S., Schimmler, M., Ziemann, M., Görg, S., Jacob, F., Karlsen, T. H., & Franke, A. (2015). Development of a high-resolution NGS-based HLA-typing and analysis pipeline. *Nucleic Acids Research*, *43*(11), e70. https://doi.org/10.1093/NAR/GKV184

Wittig, M., Juzenas, S., Vollstedt, M., & Franke, A. (2018). High-resolution HLA-typing by next-generation sequencing of randomly fragmented target DNA. *Methods in Molecular Biology*, *1802*, 63–88. https://doi.org/10.1007/978-1-4939-8546-3_5/TABLES/22

Xie, C., Yeo, Z. X., Wong, M., Piper, J., Long, T., Kirkness, E. F., Biggs, W. H., Bloom, K., Spellman, S., Vierra-Green, C., Brady, C., Scheuermann, R. H., Telenti, A., Howard, S., Brewerton, S., Turpaz, Y., & Venter, J. C. (2017). Fast and accurate HLA typing from short-read next-generation sequence data with xHLA. *Proceedings of the National Academy of Sciences*, *114*(30), 8059–8064. https://doi.org/10.1073/PNAS.1707945114

Xie, T., Rowen, L., Aguado, B., Ahearn, M. E., Madan, A., Qin, S., Campbell, R. D., & Hood, L. (2003). Analysis of the Gene-Dense Major Histocompatibility Complex Class III Region and Its Comparison to Mouse. *Genome Research*, *13*(12), 2621. https://doi.org/10.1101/GR.1736803

Yi, J., Chen, L., Xiao, Y., Zhao, Z., & Su, X. (2021). Investigations of sequencing data and sample type on HLA class Ia typing with different computational tools. *Briefings in Bioinformatics*, *22*(3), 1–6. https://doi.org/10.1093/BIB/BBAA143

Yu, Y., Wang, K., Fahira, A., Yang, Q., Sun, R., Li, Z., Wang, Z., & Shi, Y. (2021). Systematic comparative study of computational methods for HLA typing from next-generation sequencing. *HLA*, *97*(6), 481–492. https://doi.org/10.1111/TAN.14244

Zerbino, D. R., & Birney, E. (2008). Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Research*, *18*(5), 821–829. https://doi.org/10.1101/GR.074492.107

Zhao, H., Wu, L., Yan, G., Chen, Y., Zhou, M., Wu, Y., & Li, Y. (2021). Inflammation and tumor progression: signaling pathways and targeted intervention. *Signal Transduction and Targeted Therapy 2021 6:1*, *6*(1), 1–46. https://doi.org/10.1038/s41392-021-00658-5

Zheng-Bradley, X., Streeter, I., Fairley, S., Richardson, D., Clarke, L., Flicek, P., & Consortium, the 1000 G. P. (2017). Alignment of 1000 Genomes Project reads to reference assembly GRCh38. *GigaScience*, *6*(7), 1. https://doi.org/10.1093/GIGASCIENCE/GIX038

Zhong, G., Castellino, F., Romagnoli, P., & Germain, R. N. (1996). Evidence that binding site occupancy is necessary and sufficient for effective major histocompatibility complex (MHC) class II transport through the secretory pathway redefines the primary function of class II-associated invariant chain peptides (CLIP). *The Journal of Experimental Medicine*, *184*(5), 2061–2066. https://doi.org/10.1084/JEM.184.5.2061

Zhuang, B., Shang, J., & Yao, Y. (2021). HLA-G: An Important Mediator of Maternal-Fetal Immune-Tolerance. *Frontiers in Immunology*, *12*. https://doi.org/10.3389/FIMMU.2021.744324