A CASE STUDY FOR HLA TYPING IN CLINICS: REPRODUCIBILITY AND FLOWSPACE ALIGNMENT OF NGS READS

Kristina Rigó 1, György Horváth 1, Szilveszter Juhos 1, Sabine Canivet 2, Philippe Pellet 2, Abdelali Boudifa 1, Veronique Bertrand 2, Ioannis Theodorou 2

1 Omixon Biocomputing Ltd, Budapest, Hungary
2 Groupe Hospitalier Pitié Salpêtrière, Uf d’Histocompatibilité et Immunogénétique, Paris, France

Introduction

Accuracy, precision and reproducibility are the utmost goals of clinical HLA typing. Next-generation sequencing (NGS) typing algorithms have been available for a while, and for clinical application it is expected that subsequent versions of the typing software including improvements give the same results. With the evolution of the typing software changes in the core algorithm, improved handling of sequencing errors and fixing software bugs can change the outcome of the typing results – i.e. typing based only on the coding part (CDS - exons only) of the reference database allows maximum 6 digits resolution. Introducing introns and UTRs will allow 8 digits typing, and a new software feature allowing novel allele detection gives yet another layer of complexity.

In this study we had analysed 56 samples using three different typing approaches. Since many of the samples were pooled, altogether 165 typings were performed for testing each algorithm. The concordance through the different software versions were perfect, in this poster we are presenting those differences that are due to changes in reporting ambiguities or the resolution of alleles.

Methods

Short reads from 454 Roche sequencing targeting HLA-A, -B, -C and HLA-DRB1 for some samples were analysed. The samples contained reads from the whole genomic region of these genes; besides the exons intrinsic and UTR regions were also sequenced. Three different typing approaches are presented:

- Align reads to the CDS only (excluding introns and UTRs altogether) and perform typing based on coverage statistics.
- Align reads to exons, introns and UTRs (if they are available) as separate segments, perform typing based on coverage statistics.
- Align reads to continuous genomic references if they are available, and build a consensus sequence for typing.

Typing results are compared for concordance, ambiguities and possible novel alleles.

Conclusions

- NGS reads spanning through whole HLA genes can achieve high accuracy and high resolution HLA typing in clinical practice.
- Typing is reproducible; with the evolution of the typing software ambiguities are better reported or even resolved.
- Special care is needed for rare/null alleles: in many cases these allele calls are due to sequencing artifacts (i.e. homopolymer errors).
- Consensus sequences built through whole genomic references can help in finding novel alleles.