IMPROVED NEXT GENERATION SEQUENCING BASED CLASS I HLA TYPING THROUGH EXOME ENHANCEMENT


Background Precision immuno-oncology is increasingly relevant to cancer therapy given the ascendance of immunotherapy. While next-generation sequencing (NGS) based algorithms may elucidate immunotherapeutic response, many such algorithms require highly accurate Class I HLA typing. One major challenge of HLA type derivation resides in highly polymorphic HLA allelic diversity, which conventional exome sequencing technologies poorly capture. Further, accurate HLA typing requires definitive distinction between thousands of potential HLA alleles. These challenges may cause widely used NGS HLA typing tools, such as Polysolver and Optitype, to perform inaccurate HLA typing. Poor HLA coverage poses the risk of silently mistyping HLA alleles, yielding inaccurate downstream HLA loss of heterozygosity (LOH) detection and neoepitope predictions.

Methods We designed the ImmunoID NeXT Platform® to more comprehensively profile the HLA region. To evaluate the accuracy of conventional NGS-based Class I HLA typing, a widely used dbGaP project (phs000452, n=160) of melanoma NGS data was evaluated alongside a set of over 500 solid tumor cancer patient samples sequenced on the ImmunoID NeXT Platform. Read coverage was derived from both GRCh38 and HLA allele database alignments. To test whether Polysolver over represents specific HLA alleles under reduced read conditions, a Monte Carlo bootstrap approach predicted GRCh38 and HLA allele database alignments. To test whether Polysolver over represents specific HLA alleles under reduced read conditions, a Monte Carlo bootstrap approach predicted theoretical allele frequency ranges.

Results Below 20x read coverage, nearly 50% of Polysolver HLA calls (phs000452) are homozygous, representing a divergence from typical HLA homozygous rates of between 10–20%, with p<10^{-15} (Fisher’s Exact) compared to reference 1000 Genomes homozygous rates. Polysolver’s homozygous, heterozygous, and no-calls demonstrated a statistically significant difference in coverage (p<10^{-6}, Kruskal-Wallis) across all Class I HLA genes per Polysolver and public exome data (phs000452). The Personalis ImmunoID NeXT™ cohort did not demonstrate such a trend despite a similar exome-wide sequencing depth. Further, sixteen rare HLA alleles were identified with sample frequencies greater than expected from the dbGaP data set, with no such alleles identified from the Personalis ImmunoID NeXT data set.

Conclusions HLA typing may silently fail in the context of reduced read coverage without HLA-specific platform augmentation. This silent failure can have large implications for accurate neoantigen prediction and HLA LOH detection, both of which are becoming increasingly important for immuno-oncology treatment modalities such as personalized cancer vaccines, adoptive cell therapies, and blockade therapy response biomarkers. Studies utilizing neoepitope and HLA LOH prediction require careful validation for HLA calls, including assessments of coverage and homozygous rates, and may benefit from increased HLA locus coverage.

http://dx.doi.org/10.1136/jitc-2020-SITC2020.0056