PULSUS
WWW.PULSUSCONFERENCE.COM

J Genet Disor Genet Rep 2018, Volume 7
DOI: 10.4172/2327-5790-C1-003

Large scale HLA typing and repository of the Asian population

Anjali Kumari
Medgenome Labs Ltd., India

The Human Leukocyte Antigen (HLA) region encompasses a crucial set of genes that regulate the immune function. The HLA gene complex resides on a 3 Mbp stretch within chromosome 6p21 and is the most densely polymorphic region of the genome. This high sequence variability is important to fine tune the adaptive immunity. HLA testing is often required in support of patients with cancer; for example, HLA testing is used in both solid organ and Hematopoietic Stem Cell Transplantation (HSCT) and for selected pharmacogenomics testing for a personalized medicine approach (immunotherapy). Numerous studies have been conducted on HLA allele distribution across different populations and such HLA registries act as a rich resource in several aspects. In this study we have benchmarked some of the top ranked HLA typing methods available for WGS and RNA-Seq datasets. The dataset included 61 samples sequenced using in-solution targeted capturing of the classical class I (HLA-A, HLA-B, HLA-C) and class II HLA genes (HLA-DRB1, HLA-DQA1, HLA-DQB1) to a mean >80-100X coverage on Illumina sequencing platform. Five methods were tested using the WGS method which includes PHLAT (Precise HLA Typing), HLA-Vbseq and HLA miner, Polysolver (only HLA Class 1) and Optitype. The same dataset was used to simulate RNA-Seq data by restricting the number of reads to those mapped to exon regions and RNA-Seq based methods that include PHLAT, Seq2HLA, HLA-Vbseq and HLA miner were benchmarked. Based on the accuracy percentage, PHLAT, seq2HLA and HLA-Vbseq outperform with an average accuracy of 94, 80 and 75%, respectively for up to 4-digit HLA typing for both class-I and II HLA alleles. We further used an automated pipeline to do HLA typing on ~4500 samples from the Asian population. The most frequently found Class-I alleles were A*11:01:01 (29%), A*24:02:01 (25%), C*06:02:01 (23%), C*04:01:01 (20%), C*15:02:01 (19%), B*40:06:01 (18%), A*33:03:01 (18%) and Class-II alleles DQA1*01:03:01 (45%), DQB1*06:01:01 (34%), DQA1*02:01 (31%), DRB1*07:01:01 (31%), DQB1*05:03:01 (21%), DRB1*15:01:01 (21%) and DQB1*03:01:01 (21%) to be the most frequently found HLA alleles in the Asian population. We are building an Indian HLA Database (IHLAdb), by storing all the samples HLA typing results. This database will be useful for studying many of the autoimmune diseases, where HLA alleles play a crucial role such as Rheumatoid arthritis and others. HLA sequencing allows a better assessment of the HLA diversity in human populations. Analyzing HLA alleles at the population level may improve our comprehension of population genetic relationships by facilitating the identification of demographic events that marked human evolution. Imputation-Based HLA Typing with SNPs (Single nucleotide polymorphisms) in GWAS Studies will be very useful in detecting functionally significant SNPs in the MHC (Major histocompatibility complex) region.

anjalij939@gmail.com