Gur Yaari

Title: AIRR-seq data reveal mosaic deletion patterns of the human antibody heavy and light chain gene loci

Abstract: Adaptive Immune Receptor Repertoire sequencing (AIRR-seq) analysis is a potent approach to study adaptive immune responses. Many aspects of the analysis crucially depend on a reliable knowledge of the genomic loci encoding antibody genes. These loci are extremely variable between individuals, and potentially play an important role in determining genetic predisposition to a wide range of diseases. Despite its importance, we do not know a lot about variations in these loci, mostly due to technical difficulties in aligning short reads to these highly repetitive regions. One approach to study these variations is to utilize AIRR-seq data of naive cells to infer individual genotype and haplotype. For this, we generated the largest multi individual data set of naive B-cell repertoires to date, and studied these variations by applying a novel method for Bayesian genotyping and haplotyping. Our method extends haplotype inference to IGHD and IGHV-based analysis, thereby enabling inference of complex genetic events like deletions and copy number variations in the entire population. We present evidence for allele usage bias, as well as a mosaic, tiled pattern of IGHD and IGHV genes. Our findings greatly expand the knowledge that can be extracted from antibody repertoire sequencing data.

Affiliation: Faculty of Engineering, Bar Ilan University, Ramat Gan, Israel