Assessing the Consistency of Computational Tools for Quantifying Key Antibody Repertoire Metrics

JIANCHE LIU¹, Helen Huang²,³, Haripriya Vaidehi Narayanan²,³, Alexander Hoffmann²,³

¹BIG Summer Program, Institute for Quantitative and Computational Biosciences, UCLA
²Signaling Systems Laboratory, Institute for Quantitative and Computational Biosciences, UCLA
³Department of Microbiology, Immunology, and Molecular Genetics (MIMI), UCLA
Bioinformatics Interdepartmental PhD Program, UCLA

Abstract

B-cell undergoes somatic recombination and hypermutation to construct diverse B-cell receptors (BCRs) to bind different antigens. The diversity in these BCR sequences poses challenges in drawing biologically meaningful conclusions, calling for effective computational software. In the past decade, several tools have been developed to assign germline genes, determine complementarity-determining region 3 (CDR3), and characterize sequence mutation frequency and selection pressure. However, no work has been done to benchmark the performance of these tools and guide the selection of software. Here, we implemented a few of the highly cited software (Change-O, MixCR, and Parts) to compare the BCR repertoire between NFkB mutant and wild-type mice. We found that these software packages showed consistency in summarizing clonal diversity and CDR3 length distribution but diverged in estimating mutation frequency and selection pressure. Our results demonstrate the value of comparing software using real data, and provide insights into software selection in BCR repertoire analysis.

Methods & Results

3. Many computational tools were developed to analyze BCR repertoire

- **Clonal AIRR-tsv**: performs clonal AIRR-tsv analysis
- **Change-O**: analyzes BCR repertoire
- **MixCR**: analyzes BCR repertoire
- **Parts**: analyzes BCR repertoire
- **SHazaM**: analyzes BCR repertoire

3.1. Consistent estimates of CDR length and the rank abundance curve indicate greater clonal diversity in iBc/-mice

- CDR3 length distribution
- All clones show similar to a Gaussian-like distribution
- More even the curve is, the more diverse the sample is

3.2. Increased mutation rate is consistently observed in iBc/- over WT, but softwares diverge in their numerical quantification

- Change-O and MixCR have a completed BCR profiling pipeline and is more user-friendly.
- Typical repertoire metrics include CDR3 characteristic, clonal diversity, mutation rate and selection pressure.

Conclusion

- Antibody profiling is a dynamic research field – multiple software have been developed.
- Different software showed consistency in concluding clonal diversity and CDR3 length distribution in the antibody repertoire.
- Different softwares diverged in the exact numerical quantification of mutation rate and selection pressure in antibody repertoire, with consitency in drawing qualitative conclusion between iBc/− sample and WT.

References