Immunoprofiling- Learn how we characterize antibodies and T-Cell receptors in blood and other sources

**Keith Loritz**

[LoritzK@UCI.edu](mailto:LoritzK@UCI.edu) , [KRLoritz@gmail.com](mailto:KRLoritz@gmail.com) , Interdisciplinary Science & Engineering Building, Samueli School of Engineering, University of California, Irvine, CA 92697

**Oscar Camargo**

[ocamargo@ufl.edu](mailto:ocamargo@ufl.edu) , [oscardcamargo@gmail.com](mailto:oscardcamargo@gmail.com) , Herbert Wertheim College of Engineering, University of Florida, 300 Weil Hall, 1949 Stadium Road, PO Box 116550, Gainesville, FL 32611

**Ashly Corona Rodriguez**

[ashlycorogz@hotmail.com](mailto:ashlycorogz@hotmail.com), [a01275775@exatec.tec.edu](mailto:a01275775@exatec.tec.edu), Tecnológico de Monterrey, Prol. Canal de Miramontes, Coapa, San Bartolo el Chico, Tlalpan, 14380 Ciudad de México, CDMX

**Rashmi Naidu**

**Miguel Angel Rizzo Ignacio**

[leugim\_9430@hotmail.com](mailto:leugim_9430@hotmail.com),Spain-Madrid, Lucano street, 55, 28022

**Todd Smith** [todd@digitalworldbiology.com](mailto:todd@digitalworldbiology.com), [digitaltodd@me.com](mailto:digitaltodd@me.com), Digital World Biology, LLC. 2442 NW Market St. PMB 160, Seattle, WA 98107, USA

# Abstract

Immunoprofiling (immune profiling), the quantitative measurement of antigen receptors (ARs; antibodies or T-cell receptors) in a sample, is a hot topic in biotechnology. In basic research, immunoprofiling is used to assess the diversity of antigen receptors (ARs: antibodies and T-Cell receptors) and how this diversity changes in response to allergens, infections, or vaccines. Understanding AR diversity helps scientists develop new vaccines, improve existing vaccines, and treat allergies and diseases.

Immunoprofiling assays often use massively parallel, next-generation, DNA sequencing to simultaneously determine the sequences of either the DNA or RNA that encodes the variable regions of AR molecules from millions of AR cells. The resulting datasets containing millions of sequence reads (reads) are processed and aligned to AR reference data to annotate specific AR regions. These regions are then compared within and between samples to understand AR biology and how the diversity of ARs change in response to vaccines, allergy, and disease.

In terms of course-based undergraduate research (CURE), AR datasets from single samples and collections of samples ranging from raw reads to very large tables of annotated data make ideal projects for teaching programming, linking bioinformatics and biology, and finding and analyzing data.

Research goal: Using bioinformatics to understand antigen receptor biology

Database: IEDB, iReceptor, NCBI, EBI

**Keywords: bioinformatics, antibody, T-Cell receptor, antigen receptor, immunome, Immunoprofiling, immune profiling**

# Be sure to include in the sections below these points:

1. What is our research goal / question? How does antibody diversity and structure influence immune responses? By analyzing antibody sequence data, we seek to unravel the intricate relationships between antibody variations and their functional implications. This profiling of antibodies promises insights into the adaptive immune system's ability to recognize diverse antigens. Through a comprehensive understanding of antibody diversity and structure, we aspire to advance our comprehension of immune responses, potentially leading to innovations in personalized medicine and therapeutic antibody development.
2. Data - what data will we investigate? What criteria do we use for choosing it? In our antibody profiling study, our focus rests on investigating sequences derived from raw data. Our criteria for data selection encompass diversity in antibody sequences, ensuring representation across various sources and contexts. By delving into these sequences, we aim to uncover intricate patterns, elucidating the underlying mechanisms that govern antibody diversity and function. This data-driven approach allows us to discern valuable insights into immune responses, potentially revolutionizing our understanding of disease mechanisms and therapeutic interventions.
3. Tools - reagents, software, websites:iReceptor, EBI, NCBI, SabDab, IEDB (antibody - epitopes), IgBlast, fastp, unix command lines, Python, Python libraries (PANDAS, MatPlotLib, Seaborn, Biopython), Jetstream2, Slack, Google Docs/Drive (writing tools), ChatGpt, Google searches, GitHub, Dockers/Containers. . .
4. What instructions do we need? Learning programs, setting up computing environments, getting computing resources, Jupyter notebook basics, making accounts on resources. Install and utilize freely available Python resources to process and present data in useful manners. Basics of human immune system functionality.
5. What needs to be tracked and recorded during the project? programs used, data sources, and queries/filters used to access data. Good documentation of code listing author contributions, intent, concerns and suggestions of potential future paths to pursue.
6. What questions need to be answered? How do we find interesting data?

For a given starting point - DNA sequences, tab (csv) delimited dataset, what tools are needed for characterizing the data?

How do datasets from different data sources differ?

How can we summarize data in a dataset? Bioinformatics, charts, graphs, statistical estimates such as average, mean, mode(s), spread, outliers, . . .

How can we assess data quality? “Look at the data.” Are there any outliers? Is the data set complete and consistent? Are quality metrics included in the dataset? Does the metadata include notes that need to be interpreted/utilized to properly understand scope and intent?  
How will the work be assessed? How well the presentation of summary information encapsulates/distills the pertinent information obtained from the data source. How quickly (as measured in both man/cpu hours) can the raw data be pushed through the entire discovery, modification, presentation pipeline. Any one of combination of, Were team members able to find data that interested them,

# Introduction

# Methods

# Results/Discussion

# Conclusion

# References

# Figures and Tables