Jared Lumpe

720-234-5870 | jared@jaredlumpe.com | jlumpe | jaredlumpe | Jared Lumpe

Experienced computational biologist with strong programming and analytical skills and desire to work in a highly collaborative, multi-disciplinary environment. Especially interested in single-cell technologies, computational systems biology and mathematical modeling.

Education

University of California, San Francisco

San Francisco, CA

MSc in Biomedical Informatics

Mar 2020

- Developed software and methods to improve development and calibration of computational systems biology models.
- Incorporated elements of ensemble modeling, global parameter sensitivity analysis, and automated model reduction to attempt to better explore the space of possible models for a given system.
- Worked on developing a more complete and accurate dynamic model of the yeast PKA signaling network.
- Used automated optimal experimental design to derive dynamic perturbations to a biological system that yield the most informative response for the purposes of model selection and fitting.

Santa Clara University

Santa Clara, CA

BS in Cellular and Molecular Biology

Mar 2012

Minor in mathematics

Professional Experience

Independent Researcher

Boulder, CO

Genomics Software Developer

Aug 2020 - Sep 2022

- Developed GAMBIT, a software tool for taxonomic identification of bacterial pathogens from WGS data.
- Research was published in PloS One in February 2023.
- Tool has been in regular use at multiple state- and county-level public health labs across the US for several years, where it has been used to diagnose actual patients.
- Created an automated workflow to reproduce all analysis and figures in the paper.

Independent Contractor

San Jose, CA

Genomics Software Developer

Jan 2017 - Sep 2017

Worked with the Alameda County Public Health Lab to refine my initial version of the GAMBIT software tool.

Primity Bio, Inc

Fremont, CA

Bioinformatics analyst and software developer

Jan 2015 - Aug 2017

- Created methods and tools for analyzing high-throughput, high-dimensional flow cytometry data sets.
- Developed a database and web application to annotate, validate, and store experimental data and later correlate across experiments.
- Automated the process of de-barcoding flow data by clustering cell populations in multiple dimensions.
- Work resulted in a many-fold decrease in analysis time for relevant projects.

Research associate

Jul 2012 - Dec 2014

Headed a project to develop a high-throughput assay panel for the discovery of new tyrosine kinase receptor inhibitors.

Publications

- Lumpe, J., Gumbleton, L., Gorzalski, A., Libuit, K., Varghese, V., Lloyd, T., Tadros, F., Arsimendi, T., Wagner, E., Stephens, C., Sevinsky, J., Hess, D., & Pandori, M. (2023). Gambit (genomic approximation method for bacterial identification and tracking): A methodology to rapidly leverage whole genome sequencing of bacterial isolates for clinical identification. *PLOS ONE*, *18*(2), 1–22. https://doi.org/10.1371/journal.pone.0277575
- Newberry, R. W., Arhar, T., Costello, J., Hartoularos, G. C., Maxwell, A. M., Naing, Z. Z. C., Pittman, M., Reddy, N. R., Schwarz, D. M. C., Wassarman, D. R., Wu, T. S., Barrero, D., Caggiano, C., Catching, A., Cavazos, T. B., Estes, L. S., Faust, B., Fink, E. A., Goldman, M. A., … Kampmann, M. (2020). Robust sequence determinants of α-Synuclein toxicity in yeast implicate membrane binding. *ACS chemical biology*, *15*(8), 2137–2153. https://doi.org/10.1021/acschembio.0c00339
- Sheikh, A. I., Caswell, D., Dick, C., Gang, S., Jarrell, J., Kohli, A., Lieu, A., Lumpe, J., Garrett, M., Parker, J., & Stephens, C. (2013). Regulation of d-galacturonate metabolism in caulobacter crescentus by HumR, a LacI-family transcriptional repressor. *Advances in bioscience and biotechnology*, *4*(10), 63–74. https://doi.org/10.4236/abb.2013.410A3008

Skills

Data Science and ML	Substantial knowledge of standard supervised learning models for regression and classification. Significant experience with clustering of high-dimensional data and dimensionality reduction, hierarchical Bayesian models and inference using MCMC methods.
Computational biology and bioinformatics	Mechanistic modeling of biological networks, parameter identification, ensemble modeling. Extensive experience in analysis of high-dimensional single-cell data sets (flow cytometry, scRNA seq). Analysis and QC of NGS data, *de novo* prokaryotic genome assembly. Quantitative microscopy image analysis (cell segmentation and tracking, feature extraction).
Mathematics	Significant knowledge of linear algebra, calculus/differential equations, and probability theory.
Programming languages	Python (advanced), Julia, C/C++, R, MATLAB, Bash scripting, SQL, Java, JavaScript/TypeScript, Mathematica, Lisp, HTML, PHP, LaTeX.
Scientific computing	Highly experienced with numeric and scientific computing libraries for Python (Numpy, Scipy, Pandas, Scikit-Learn, Sympy, BioPython). Significant experience with the Julia software ecosystem for scientific computing and mathematical modeling. Numerous contributions to open-source scientific computing software, such as BioPython. Chaining multiple tools into pipelines using Snakemake/WDL and creating reproducible analyses with the Jupyter ecosystem.
Software development	Version control (Git), automated testing, continuous integration, documentation, writing maintainable code using standard best practices.
Other computational skills	Significant experience with Linux system configuration and administration. Cloud computing with Google Cloud. Containerization using Docker.

Other Projects

pycyt

Python package for the analysis of flow cytometry data. Inspired by my work at Primity Bio and dissatisfaction with existing alternatives. Reads/writes FCS files, performs compensation, transformations, and gating, and creates specialized plots. Focused on high-throughput data sets, interactivity, and extensibility.

[Home page] [Usage examples]

Symbolic modeling toolkit

Python library for defining and manipulating symbolic representations of mathematical models, developed during my time at UCSF. Capable of translating model equations into optimized computational graphs, performing automatic symbolic differentiation and translating into high-performance code in C, Numba, or Theano. Intended to be used as the basis for computer-guided model reduction.

References available upon request.