Zachary L. Maas

Summary

Computational scientist with expertise in high throughput sequencing data and using machine learning to solve problems in a biologically informed way. I have worked with hundreds of datasets, across a range of protocols. My work focuses on new multi-omic methods, pulling more information out of the data we do have, and understanding experiments when our assumptions fail or data quality is low.

Education

PhD, Computer Science

4TH YEAR, EXPECTED 2024, UNIVERSITY OF COLORADO BOULDER

> Statistical Tool and Algorithm Development in the Robin Dowell Lab
> Interdisciplinary Quantitative Biology Program

Undergraduate

MAY 2019, UNIVERSITY OF COLORADO BOULDER

> B.A. Chemistry, *Summa Cum Laude*
> B.A. Mathematics, Emphasis in Computational Mathematics
> Honors Thesis: *Method Optimization in the Analysis of Nascent Sequencing Data*

Research Experience

Graduate Student, Computer Science

2020-PRESENT, DOWELL LAB, CU BOULDER, COMPUTER SCIENCE

> Current work — combining nascent sequencing data (PRO-seq, GRO-seq) and single cell sequencing assays (scRNA-seq, scATAC-seq) for the identification of cell-type specific transcriptional regulatory programs, comparing classical and deep learning approaches.
> Developed novel algorithms and statistical methods for the analysis of high throughput sequencing data
> Created a technique for normalization of nascent sequencing assays in the absence of traditional external spike-ins, which performs with high sensitivity and accuracy.
> Extended deconvolution (*in-silico* microdissection) techniques to the analysis of nascent sequencing data in bulk, identifying issues with literature best-practices when applied to much larger sets of data.

Collaborative Graduate Projects

2020-PRESENT, CU BOULDER

> Prototyped methods to infer biologically relevant conclusions from large knowledge graphs, integrating ontology and sequencing data.   MEDICAL SCHOOL
> Developed a deep-learning based pipeline for segmentation of live-cell movies in cyanobacteria, improving segmentation speed by 5x with significant quality improvements.   BIOCHEMISTRY
> Improved code on a cross-institutional project to discover feasible metabolic networks on early earth. Developed API and front-end code to improve accessibility of project tools.   GEOLOGY

Professional Research Assistant

2019-2020, DOWELL LAB, CU BOULDER

> Solved statistical and computational issues in the development of a model to simultaneously infer the activity of hundreds of transcription factors from a single experiment.
> Initiated the analysis, curation, and quality control of thousands of sequencing samples in published experiments.
> Established best-practices in analysis pipelines, facilitating better scalability and accessibility.

Undergraduate Research Assistant

2017–2019, Taatjes Lab / Dowell Lab, CU Boulder

> Optimized existing computational methods used in the analysis of nascent sequencing data, including pipeline development.
> Led computational analysis of international collaborations, studying transcriptional regulation by general transcription factors.

PUBLICATIONS

[1]   **Z. Maas**, C. Huffine, C. Brininger, J. W. Tay, and J. Cameron. "CyPose: Segmentation and classification of cyanobacterial cells with deep learning". 2024. In preparation.

[2]   **Z. Maas**, R. Sigauke, and R. Dowell. "Deconvolution of Nascent Sequencing Data Using Transcriptional Regulatory Elements". In: *Proceedings of the 2024 Pacific Symposium on Biotechnology*. Oct. 15, 2023. DOI: 10.1101/2023.10.11.561942. URL: http://biorxiv.org/lookup/doi/10.1101/2023.10.11.561942.

[3]   **Z. L. Maas** and R. D. Dowell. "Internal and External Normalization of Nascent RNA Sequencing Run-On Experiments". In: *BMC Bioinformatics* (Dec. 2023). In press.

[4]   J. D. Rubin, J. T. Stanley, R. F. Sigauke, C. B. Levandowski, **Z. L. Maas**, J. Westfall, D. J. Taatjes, and R. D. Dowell. "Transcription Factor Enrichment Analysis (TFEA) Quantifies the Activity of Multiple Transcription Factors from a Single Experiment". In: *Communications Biology* 4.1 (1 June 2, 2021), pp. 1–15. ISSN: 2399-3642. DOI: 10.1038/s42003-021-02153-7. URL: https://www.nature.com/articles/s42003-021-02153-7.

[5]   C. B. Fant, C. B. Levandowski, K. Gupta, **Z. L. Maas**, J. Moir, J. D. Rubin, A. Sawyer, M. N. Esbin, J. K. Rimel, O. Luyties, M. T. Marr, I. Berger, R. D. Dowell, and D. J. Taatjes. "TFIID Enables RNA Polymerase II Promoter-Proximal Pausing". In: *Molecular Cell* 78.4 (May 21, 2020), 785–793.e8. ISSN: 1097-2765. DOI: 10.1016/j.molcel.2020.03.008. URL: http://www.sciencedirect.com/science/article/pii/S1097276520301568.

[6]   J. K. Rimel, Z. C. Poss, B. Erickson, **Z. L. Maas**, C. C. Ebmeier, J. L. Johnson, T.-M. Decker, T. M. Yaron, M. J. Bradley, K. B. Hamman, S. Hu, G. Malojcic, J. J. Marineau, P. W. White, M. Brault, L. Tao, P. DeRoy, C. Clavette, S. Nayak, L. J. Damon, I. H. Kaltheuner, H. Bunch, L. C. Cantley, M. Geyer, J. Iwasa, R. D. Dowell, D. L. Bentley, W. M. Old, and D. J. Taatjes. "Selective Inhibition of CDK7 Reveals High-Confidence Targets and New Models for TFIIH Function in Transcription". In: *Genes & Development* (Oct. 15, 2020). ISSN: 0890-9369, 1549-5477. DOI: 10.1101/gad.341545.120. PMID: 33060135. URL: http://genesdev.cshlp.org/content/early/2020/10/14/gad.341545.120.

HONORS AND AWARDS

> National Science Foundation GRFP Honorable Mention   2021
> CU Boulder Computer Science Professional Development Fellowship   2021
> CU Boulder BioFrontiers Slosky Research Award   2021
> American Chemical Society Award in Analytical Chemistry   2020
> American Institute of Chemists Student Award   2020
> Membership — Pi Mu Epsilon Mathematical Honor Society   2020
> CU Boulder Undergraduate Research Grant, Academic Year   2019-2020

TEACHING AND OUTREACH

Dowell And Allen Lab Short Read Sequencing Workshop

2019-Present

> Annual Two-Week Graduate-Level Workshop funded by INCLUDE grants to serve underrepresented early career researchers.
> Led instruction of a two-week graduate workshop covering best practices in handling and processing high throughput sequencing data.
> Managed curriculum design of new course material for emerging techniques of interest to attendees.
> Helped to manage $30+$ graduate students and postdoctoral researchers flown in to attend the course.

Dowell And Allen Lab Code Club

2019-Present

> Established regular code club covering topics of interest to researchers within our lab and the broader local research community.
> Developed content and taught topics ranging on statistical techniques and best practices as well emerging techniques in deep learning.

SCOPE (Scientific Outreach Program and Education

2022-Present

> Led classroom volunteers in education and outreach to local middle-schools.
> Implemented hands-on experiments and activities to improve student engagement with the sciences.

Skills

Software Development

> (Intermediate Skill) Python, R, Shell Scripting
> (Basic Skill) C++
> Bayesian modeling with pymc3
> Deep learning with pytorch

Scientific

> High performance computing on large datasets in a distributed environment
> Sequencing analysis and quality control of RNA-seq, ATAC-seq, Nascent sequencing, as well as scRNA-seq and scATAC-seq data.
> Automated pipeline development