

NOAH DUKLER

Boulder, CO 80304

(+1)845-417-8166 ♦ ndukler@gmail.com ♦ noahdukler.netlify.app

PROFESSIONAL SUMMARY

Domain knowledge in regulatory genomics, transcriptomics, phylogenetics, and gene regulation from PhD and subsequent post-doctoral research position. Developed novel computational analysis and pipelines for deployment and insuring quality using continuous integration and testing. Experience primarily in R and Python with significant development time in C++ for performance optimization. Industry experience at a biotech startup where I implemented analysis at and data visualization scale using AWS as part of an integrative bioinformatics web portal.

WORK EXPERIENCE

Post-doctoral researcher

January 2019 - August 2020

- Led project in comparative inter-species epigenomics which resulted in publication and release of R package *phyloGLM*
- Participated as part of large software development team focused on standardizing population genetic simulation resulting in co-first author paper
- Led project developing method for isoform level quantification of nascent RNA with R package (DENR)

Computational Scientist

September 2020 - Present

- Custom bioinformatics analysis for large industry clients and internal projects
- Full stack development for bioinformatics web portal built using Python, MySQL, JavaScript, and AWS
- API development for programatic use of bioinformatics platform
- AWS cost optimization

SELECTED PROJECTS

Inferring the determinants of cis-regulatory evolution

- Developed novel method fitting parameters of phylogenetic processes as a function of site-wise covariates, enabling users to directly and jointly test multiple hypothesis about the evolution of discrete traits
- Applied method to apply these methods to previously published epigenetic data that serves as a proxy for gene regulatory element in liver tissue from nine mammals and found evidence for a variety of genomic feature correlating with differential evolutionary rates accross multiple classes of cis-regulatory elements
- Implemented in R and C++ as phyloGLM package on GitHub (<https://github.com/ndukler/phyloGLM>).

Estimation of nascent transcript abundance

- Developed novel method for quantifying nascent RNA sequencing data at the gene and transcript levels

- Addresses the lack of exonic structure and increased noise particular to nascent RNA data with ML based filtering using TensorFlow that constrains the subsequent deconvolution step
- Implemented in R as the tuSelector2 package on GitHub (<https://github.com/ndukler/tuSelector2>)

Analysis of rapid transcriptional dynamics after drug treatment

- Analysis of nascent RNA sequencing data to detect changes in transcription levels within in minutes to distinguish primary from secondary responses to regulatory signals in response to celastrol, a compound with reported anti-inflammatory, tumor-inhibitory, and obesity-controlling effects
- Applied an auto-regressive clustering model to identify distinct temporal profiles of gene responses and identify associated pathways.
- Regressed estimated transcription factor affinities in promoters against time-course expression profiles to temporally resolve transcription factors driving transcriptional responses

EDUCATION

SUNY Geneseo

September 2009 - May 2013

Bachelors of Biochemistry (B.S.) & Mathematics (B.A.)
Department of Biology, Department of Mathematics

Weill Cornell Graduate School of Medical Sciences

July 2013 - December 2018

PhD. in Computational Biology

SKILLS

Machine Learning	Graphical Models, Deep Learning, Generalized Linear Models
Programming Languages	R, Python, C++, Perl, Bash
Pipeline Development	Snakemake
Software Development	GitHub, Travis CI, CircleCI
Data Visualization	ggplot2, Adobe Illustrator
Databases	MySQL
Web development	JavaScript
AWS	S3, BATCH, ECS, EC2

SELECTED TALKS

Leveraging comparative epigenomics to understand cis-regulatory evolution Oct 2019

Presented for Probabilistic Modeling in Genomics; Aussois, France

Characterizing the genomic determinants of cis-regulatory element evolution with phylogenetic models Jan 2019

Presented for New York Population Genetics Meeting at Mt. Sinai; NYC, NY

SELECTED PUBLICATIONS

- (1) Yixin Zhao et al. “Deconvolution of Expression for Nascent RNA sequencing data (DENR) highlights pre-RNA isoform diversity in human cells”. In: *Bioinformatics* (Aug. 2021). ISSN: 1367-4803. DOI: 10.1093/bioinformatics/btab582. eprint: <https://academic.oup.com/bioinformatics/advance-article-pdf/doi/10.1093/bioinformatics/btab582/39678541/btab582.pdf>. URL: <https://doi.org/10.1093/bioinformatics/btab582>.
- (2) Jeffrey R Adrion et al. “A community-maintained standard library of population genetic models”. In: *eLife* 9 (June 2020). Ed. by Graham Coop et al., e54967. ISSN: 2050-084X. DOI: 10.7554/eLife.54967. URL: <https://doi.org/10.7554/eLife.54967>.
- (3) Noah Dukler, Yi-Fei Huang, and Adam Siepel. “Phylogenetic Modeling of Regulatory Element Turnover Based on Epigenomic Data”. In: *Molecular Biology and Evolution* (Mar. 2020). DOI: 10.1093/molbev/msaa073.
- (4) Hussein A Hejase, Noah Dukler, and Adam Siepel. “From Summary Statistics to Gene Trees: Methods for Inferring Positive Selection”. In: *Trends in Genetics* (2020).
- (5) Noah Dukler et al. “Is a super-enhancer greater than the sum of its parts?” In: *Nature genetics* 49.1 (2017), p. 2.
- (6) Noah Dukler et al. “Nascent RNA sequencing reveals a dynamic global transcriptional response at genes and enhancers to the natural medicinal compound celastrol”. In: *Genome research* 27.11 (2017), pp. 1816–1829.

TEACHING

Cold Spring Harbor Laboratory

2019

ML course tutor

Tutored machine learning course for graduate students and post-docs. Lectured, assisted with problem sets, and provided support for the application of course topics to individual course attendees projects.

Cold Spring Harbor Laboratory

2015-2016

CSHL PFF tutor

Tutor for CSHL Partners for the Future Program. Tutored high school student in research project focused on gene regulation and transcription. Focused on developing scripting skills in R and applications of linear models.