ARIELLA GLADSTEIN, PhD

algladstein@gmail.com

`https://ariella-gladstein.netlify.app` ◇ `https://github.com/agladstein` ◇ `www.linkedin.com/in/ariella-gladstein`

Expertise: *population genetics, genomics, computational biology, data science, machine learning, research computing*

EDUCATION

University of Arizona, Tucson, AZ *August 2018*
PhD in Ecology and Evolutionary Biology, Minor in Mathematics

Beloit College, Beloit, WI *May 2011*
B.S. in Mathematical Biology & Russian, Cum Laude

EXPERIENCE

Research Scientist, Computational Biology 2021 - 2024
Ancestry team *Embark Veterinary, Remote*
Developed dog breed ancestry and relative-matching products, enhancing computational efficiency while maintaining rigorous quality standards. Played a key role in improving pipeline automation from 40% to over 95% within one year and scaling processes from hundreds of thousands to millions of samples. Collaborated with a cross-functional team on a science initiative projected to save $7.5 million annually. Served on the Diversity, Equity, Inclusion, and Belonging (DEIB) committee, co-authoring an Inclusive Norms document, designing and analyzing company-wide DEIB surveys, and contributing to various DEIB initiatives.

Postdoctoral Fellowship 2018 - 2020
Schrider Lab *Department of Genetics, University of North Carolina, Chapel Hill*
Developed deep learning models for population genetics inference, utilizing high-throughput computing for simulations and high-performance GPUs, as well as a Kubernetes cluster, for model training. Contributed to writing documentation and tutorials for the population genomics tskit software ecosystem. Actively participated in the PopSim Consortium, contributing to the development of the stdpopsim library.

Dissertation Research 2011 - 2018
Hammer Lab *Ecology and Evolutionary Biology, University of Arizona*
Developed and applied computational methods for population genetics analysis, addressing ascertainment bias in identifying runs of homozygosity and demographic inference. Created and released software for high-throughput genomic simulations. Applied Approximate Bayesian Computation (ABC) to model the demographic history of Ashkenazi Jews, utilizing over 11 million CPU hours in under six months to run large-scale simulations on the Open Science Grid and multiple HPC systems with a high-throughput workflow.

SKILLS

Programming	Python, R, Bash, unit testing, profiling, multiprocessing, Git
HPC/HTC & Cloud	Linux/Unix, SLURM, PBS, AWS, JupyterHub, Kubernetes
Reproducibility	Workflow development (Pegasus, Makeflow, Snakemake), containers (Singularity, Docker), virtual environments, Jupyter, Knitr
Genomics	PLINK, Bedtools, VCFtools, phasing, imputation, haplotype detection, population structure, genome simulation
Data Science	Numpy, Pandas, Scikit-learn, SciPy, TensorFlow, Keras, Matplotlib, Seaborn, Plotly
Documentation	LaTeX, Markdown, reStructuredText, JupyterBook, Confluence

TEACHING EXPERIENCE

Introduction to using stdpopsim Workshop 2020, *Virtual, PopSim Consortium*
Designed and presented interactive workshop on using stdpopsim, the standard library for reproducible, bug-free simulations of genetic datasets from published demographic histories.

NSF Cyber Carpentry Workshop 2019, *University of North Carolina, Chapel Hill*
Assisted with 2-week workshop, where participants learned best practices for data-intensive computing, cloud infrastructure, and deep learning through hands-on projects.

Research Bazaar Workshop on R 2018, *University of Arizona*
Assisted with hands-on intro R workshop

CyVerse Container Camp   2018, *University of Arizona*
Assisted with 3-day hands-on workshop on using containerized workflows for more reproducible science.

Software Carpentry Workshop   2017, *University of Arizona*
Assisted with Software Carpentry hands-on workshop on basic Unix/Bash, Python, and Git

Tucson Womens Hackathon workshop on Git   2017, *University of Arizona*
Assisted with hands-on workshop on basic Git

ECOL 320 Genetics   2014 - 2016, *University of Arizona*
Graduate Teaching Assistant

ECOL 182L Intro to Ecology and Evolutionary Biology Lab   2015, *University of Arizona*
Graduate Teaching Assistant

BIOL 247 Biometrics   2011, *Beloit College*
Teaching Assistant

AWARDS

XSEDE Research allocation (250,000 CPU hrs)	*2020*
XSEDE Supplemental allocation (250,000 CPU hrs)	*2018, 2020*
XSEDE Startup allocation (100,000 CPU hrs, 6,500 GPU hrs, 5.5 Tb storage)	*2019*
XSEDE Startup allocation (150,000 CPU hrs)	*2017*
Open Science Grid User School (travel, lodging, meals, cost of program)	*2017*
NIH Computational and Mathematical Modeling of Biological Systems Traineeship ($71,064)	*2013-14*
NSF Integrative Graduate Education and Research Traineeship in Genomics ($97,083)	*2011-13*

WORKSHOPS AND HACKATHONS

SMBE satellite meeting on Speciation Genomics, Tjarno, Sweden (3 days)	*06/2019*
NCBI RNA-Seq in the Cloud hackathon, Chapel Hill, NC (3 days)	*03/2019*
Cyber Carpentry, Chapel Hill, NC (2 weeks)	*06/2018*
XSEDE HPC Workshop: Big Data, Tucson, AZ (2 days)	*02/2018*
Open Science Grid User School, Madison, WI (1 week)	*07/2017*

SELECT PUBLICATIONS, POSTERS, & TALKS

Gladstein A.L.*, et al. A Novel Machine Learning Method for Classification of Village Dogs. Probabilistic Modeling in Genomics. 2023. Cold Spring Harbor, NY. (Poster)

Baumdicker, F.*, Bisschop, G.*, Goldstein, D.*, Gower, G.*, Ragsdale, A. P.*, Tsambos, G.*, Zhu, S.*, …, **Gladstein, A. L.**, et al. 2022. Efficient ancestry and mutation simulation with msprime 1.0. *Genetics.* Volume 220, Issue 3. https://doi.org/10.1093/genetics/iyab229

Adrion, J. R.*, Cole, C. B.*, Dukler, N.*, Galloway, J. G.*, **Gladstein, A. L.**[†], Gower, G.*, Kyriazis, C.C.*, Ragsdale, A.P.*, Tsambos, G.*, et al. 2020. A community-maintained standard library of population genetic models. *eLife.* 9:e54967. doi: https://doi.org/10.1101/2019.12.20.885129

Gladstein A.L.[†], Schrider R.D. Demographic model selection with deep learning. Probabilistic Modeling in Genomics. 2019. Aussois, France. (Poster)

Gladstein A.L.* and Hammer M.F. 2019. Substructured population growth in the Ashkenazi Jews inferred with Approximate Bayesian Computation. *Molecular Biology and Evolution.* 36(6): 1162-1171. doi: https://dx.doi.org/10.1093/molbev/msz047

Inference of evolutionary history with Approximate Bayesian Computation. Open Science Grid All-Hands Meeting. 2018. Salt Lake City, Utah. (Talk)

Code optimization for research scientists. Research Bazaar. 2018. Tucson, AZ. (Talk)

Gladstein A.L.* et al. 2018. SimPrily: A Python framework to simplify high-throughput genomic simulations. *SoftwareX*, 7, 335-340. https://doi.org/10.1016/j.softx.2018.09.003

---

*Presenter
[†]First author