

# ARIELLA GLADSTEIN, PhD

algladstein@gmail.com

<https://github.com/agladstein> ♦ [www.linkedin.com/in/ariella-gladstein](https://www.linkedin.com/in/ariella-gladstein)

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

## EDUCATION

<b>University of Arizona, Tucson, AZ</b> PhD in Ecology and Evolutionary Biology, Minor in Mathematics	August 2018
<b>Beloit College, Beloit, WI</b> B.S. in Mathematical Biology & Russian, Cum Laude Departmental Honors: Mathematical Biology	May 2011
Undergraduate Study Abroad: Lomonosov Moscow State University, Moscow, Russia <i>Biology Department</i>	Fall 2009
Russian State University for the Humanities, Moscow, Russia <i>Russian Studies</i>	Spring 2009

## EXPERIENCE

<b>Open Source Contributor</b> <i>Tskit community</i> Contributing to writing and reviewing documentation and tutorials for the population genomics tskit ecosystem.	2024 - present <i>Remote</i>
<b>Research Scientist, Computational Biology</b> <i>Ancestry Science</i> Developed dog breed ancestry and relative-matching products using AWS cloud infrastructure and state of the art computational genomics methods. <ul style="list-style-type: none"><li>Proposed and developed a machine learning approach for village dog classification, enhancing accuracy and scalability.</li><li>Conducted literature reviews, tested, and reported recommendations on published methods for population clustering and relative inference.</li><li>Collaborated with scientists and engineers to improve local ancestry inference, implementing algorithm improvements for uncertainty estimates.</li><li>Played a key role in improving AWS pipeline automation from 40% to over 95% and scaling processes from hundreds of thousands to millions of samples.</li><li>Collaborated with a cross-functional team on a science initiative projected to save \$7.5 million annually.</li><li>Served on the Diversity, Equity, Inclusion, and Belonging (DEIB) committee, co-authoring an Inclusive Norms document and contributing to various DEIB initiatives.</li></ul>	2021 - 2024 <i>Embark Veterinary, Remote</i>
<b>Postdoctoral Fellowship</b> <i>Schrider Lab</i> Developed deep learning models for population genetics inference. 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.	2018 - 2020 <i>Department of Genetics, University of North Carolina, Chapel Hill</i>
<b>Dissertation Research</b> <i>Hammer Lab</i> Dissertation: Inference of recent demographic history of population isolates using genome-wide high density SNP arrays and whole genome sequences 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.	2011 - 2018 <i>Ecology and Evolutionary Biology, University of Arizona</i>

**Functional Genomics Research Rotation**  
*Restifo Lab*

Spring 2012  
*Neuroscience, University of Arizona*

**Computational Genomics Research Rotation**  
*Kececioglu Lab*

Fall 2011  
*Computer Science, University of Arizona*

**Lab technician intern**  
*Laboratory of Population Genetics*

Sp, Fall 2009  
*Russian Academy of Medical Sciences, Moscow, Russia*

## TEACHING EXPERIENCE

---

**Introduction to using stdpopsim Workshop**

2020

Designed and presented interactive workshop on using stdpopsim, the standard library for reproducible, bug-free simulations of genetic datasets from published demographic histories.

*Virtual, PopSim Consortium*

**NSF Cyber Carpentry Workshop**

2019

Assisted with 2-week workshop, where participants learned best practices for data-intensive computing, cloud infrastructure, and deep learning through hands-on projects.

*University of North Carolina, Chapel Hill*

**Research Bazaar Workshop on R**

2018

Assisted with hands-on intro R workshop.

*University of Arizona*

**CyVerse Container Camp**

2018

Assisted with 3-day hands-on workshop on using containerized workflows.

*University of Arizona*

**Software Carpentry Workshop**

2017

Assisted with Software Carpentry hands-on workshop on basic Unix, Python, and Git

*University of Arizona*

**Tucson Womens Hackathon workshop on Git**

2017

Assisted with hands-on workshop on basic Git

*University of Arizona*

**ECOL 320 Genetics**

2014 - 2016

Graduate Teaching Assistant

*University of Arizona*

**ECOL 182L Intro to Ecology and Evolutionary Biology Lab**

2015

Graduate Teaching Assistant

*University of Arizona*

**Population Genetics Module**

2013

Developed and taught a 3-day module on population genetics to high schoolers.

*Kino School*

**BIOL 247 Biometrics**

2011

Teaching Assistant

*Beloit College*

**RUSS 110, 115, 210, 215 Russian**

2009 - 2011

Tutor

*Beloit College*

## SKILLS

---

**Programming  
HPC/HTC & Cloud  
Genomics**

Python, R, Bash, unit testing, profiling, multiprocessing, Git  
Linux/Unix, SLURM, PBS, AWS, SageMaker, JupyterHub, Kubernetes  
PLINK, Bedtools, VCFtools, phasing, imputation, haplotype detection, population structure, genome simulation

**Data Science  
Reproducibility**

Numpy, Pandas, Scikit-learn, SciPy, Keras, Matplotlib, Seaborn, Plotly, ggplot  
Workflow development (Pegasus, Makeflow, Snakemake), containers  
(Singularity, Docker), virtual environments, Jupyter, Knitr

**Documentation  
Mathematical Skills  
Language Skills**

L<sup>A</sup>T<sub>E</sub>X, Markdown, reStructuredText, JupyterBook, Confluence  
Probability theory, statistical inference, linear algebra  
English (native speaker), Russian (fluent)

## PUBLICATIONS

---

Lauterbur, M. E.\* et al. 2023. Expanding the stdpopsim species catalog, and lessons learned for realistic genome simulations *eLife* 12:RP84874. <https://doi.org/10.7554/eLife.84874.3>

Baumdicker, F.\* 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>

Bernstein M.N.\* et al. Jupyter notebook-based tools for building structured datasets from the Sequence Read Archive [version 2; peer review: 2 approved]. F1000Research 2020. 9:376. doi: <https://doi.org/10.12688/f1000research.23180.2>

**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>

**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>

**Gladstein A.\*** and Hammer M.F. 2016. Population Genetics of the Ashkenazim. In: *eLS*. John Wiley & Sons, Ltd: Chichester. pp. 1-8. <https://doi.org/10.1002/9780470015902.a0020818.pub2>

Behar D.\* et al. 2013. No evidence from genome-wide data of a Khazar origin for the Ashkenazi Jews. *Human Biology*. 85.6:859-900. <https://doi.org/10.3378/027.085.0604>

**Gladstein A.L.\*** 2011. Split decomposition analysis groups Jewish populations together between European and Middle Eastern populations. *The Beloit Biologist*. 30:29-36.

## POSTERS AND PRESENTATIONS

---

### Invited Talks

Deep learning for demographic model choice. Human Evolutionary and Population Genomics Seminar. 2020. LANGE BIO, UGA-Cinvestav. Irapuato-Guanajuato, Mexico (virtual).

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

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

### Posters

**Gladstein A.L.<sup>†</sup>**, et al. A Novel Machine Learning Method for Classification of Village Dogs. Probabilistic Modeling in Genomics. 2023. Cold Spring Harbor, NY.

**Gladstein A.L.<sup>†</sup>**, Schrider R.D. Demographic model selection with deep learning. Probabilistic Modeling in Genomics. 2019. Aussois, France.

**Gladstein A.L.<sup>†</sup>**, Hammer M.F. Substructured population growth in the Ashkenazi Jews inferred with Approximate Bayesian Computation. UNC Women in Computing Research Symposium. 2019. Chapel Hill, NC.

**Gladstein A.L.<sup>†</sup>**, Hammer M.F. Substructured population growth in the Ashkenazi Jews inferred with Approximate Bayesian Computation. Probabilistic Modeling in Genomics. 2018. Cold Spring Harbor, NY.

**Gladstein A.L.<sup>†</sup>**, et al. Efficient pipeline for whole genome simulation and summary statistic calculation with flexible demographic models. Meeting of the American Society for Human Genetics. 2017. Orlando, FL.

**Gladstein A.L.<sup>†</sup>**, et al. The effect of SNP array ascertainment bias on the distribution of runs of homozygosity lengths. Annual Meeting of the American Society for Human Genetics. 2015. Baltimore, MD.

---

\*First author

<sup>†</sup>Presenter

## 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

## AWARDS

---

XSEDE Research allocation (250,000 CPU hrs)	2020
XSEDE Supplemental allocation (250,000 CPU hrs)	2018, 2020
Probabilistic Modeling in Genomics Grant (registration, meals, lodging)	2018, 2019
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
GPSC Travel Grant (\$761)	2015, 2016, 2017
University of Arizona Galileo Circle Scholarship (\$1,000)	2015
NIH Computational and Mathematical Modeling of Biological Systems Traineeship (\$71,064)	2013 - 2014
NSF Integrative Graduate Education and Research Traineeship in Genomics (\$97,083)	2011 - 2013
Society for Learning Unlimited Grant (\$2,000)	2009
Study Abroad Enhancement Grant, Beloit College (\$250)	2009
Beloit College Presidential Scholar (\$60,000)	2007 - 2011

## COMMUNITY SERVICE

---

Judged Graduate & Professional Student Council Travel Grants, Tucson, AZ	2012, 2016, 2017, 2018
Mentored undergraduate interns in computer science, Tucson, AZ	2017
Judged Tucson Magnet High School Science Fair, Tucson, AZ	2012, 2015
Judged EEB Undergraduate Research Poster Session, Tucson, AZ	2012
Arizona Assurance Mentor	2012

## OTHER ACTIVITIES

---

<b>Circus Arts</b>	2012 - present
Trained, performed, and taught aerial silks, aerial rope, lyra, flying trapeze, tightwire, handbalancing, contortion.	
<b>Figure Skating</b>	1998 - 2013, 2021 - present
Trained, performed, competed, and taught singles freestyle.	