Charlotte A. Darby

cdarby.github.io | charlotte.ay.darby@gmail.com | LinkedIn | @ChaDarby

Experience

New York Genome Center — New York, NY — 2020 — present
Senior Computational Biologist, Satjia Lab

10x Genomics — Pleasanton, CA — 2019
Computational Biology Intern
Projects: Allele-specific and cell-type specific expression of HLA
Sample index design for multiplexing

IBM Research Almaden — San Jose, CA — 2018
Graduate Research Assistant in Applied & Industrial Genomics
Project: Machine learning for classifying bacterial genomes

Cold Spring Harbor Laboratory — Cold Spring Harbor, NY — 2015
Undergraduate Research Program (URP) Researcher in Doreen Ware's lab
Project: Transcription factor networks in maize and *Arabidopsis*

Education

Ph.D. Computer Science — Johns Hopkins University — 2016 — 2020
Thesis: Computational methods addressing genetic variation in next-generation sequencing data (May 2020)
Major Projects: Vargas: fast heuristic-free read alignment
Samovar: somatic variant calling with linked reads
Other Projects: Algorithms for polyploid haplotype assembly
Species phylogeny and comparative genomics of *Trichomonadida*
LRSim: simulation software for linked reads
Comparing haplotype assemblies from linked and long reads
Co-Advisors: Michael Schatz, Ben Langmead

M.S. Computational Biology — Carnegie Mellon University — 2015 — 2016
GPA 3.55; Research Honors
Thesis: New nomenclature for horizontal gene transfer
Advisor: Dannie Durand

B.S. Computational Biology — Carnegie Mellon University — 2012 — 2015
GPA 3.70; University Honors

Teaching

Johns Hopkins University — Baltimore, MD — 2017 — present

- Teaching Academy Certificate of Completion
- Instructor, "Algorithms for hard problems in computer science" (seminar for first-year engineering students, class size 10, 2 semesters)
- Teaching Assistant, "Applied Comparative Genomics" (graduate elective course, class size 20, 1 semester)

Center for Talented Youth — Baltimore, MD — 2016 — 2017

- Teaching Assistant, Genetics & Genomics (residential summer academic program for high schoolers, class size 6-18, three 3-week sessions)

Carnegie Mellon University — Pittsburgh, PA

- Teaching Assistant, Introduction to Computational Biology (undergraduate core course, taught recitation sections and held office hours, class size 20/section, 2 semesters) — 2015 — 2016

Conference Posters and Presentations

Vargas: heuristic-free alignment for assessing graph and linear aligners (poster)	Genome Informatics 2019
Samovar: single-sample mosaic SNV calling with linked reads (talk) Best Paper Award	RECOMB-SEQ 2019
A heuristic algorithm for the k-ploid haplotype assembly problem (poster with undergraduate mentee George Botev)	Biological Data Science 2018
Leveraging linked reads for single-sample somatic variant calling (poster)	Genome Informatics 2017
Identification and analysis of somatic variants using linked read sequencing (poster)	Biology of Genomes 2017
Classifying Xenologs (talk)	GL-Bio/CCBC 2016
Classifying Xenologs (poster)	RECOMB-CG 2014

Outreach

Maryland New Directions — 2018 — 2020
Assist job-seeking clients in selecting professional clothing sourced from community donations

Johns Hopkins SABES — 2016 — 2017
Volunteer bi-weekly as a mentor for after-school elementary STEM program

Women@SCS (Carnegie Mellon University) — 2014 — 2016
Facilitate TechNights workshops for middle school students

Networking

CRA-W (Computing Research Association) — April 2018
Grad Cohort for Women Workshop attendee

Honors

Academic Achievement Fellowship 2015 — 2016
Carnegie Mellon University (M.S. Computational Biology program)

National Merit Scholarship 2012

Technical Skills

Advanced: Python, R, Unix. *Intermediate:* Rust, C++. *Proficient:* Matlab, Java, C, SML.
small variant calling, short read alignment, linked reads, haplotype assembly (phasing), high-performance computing environments, single-cell genomics, genome assembly, functional genomics, long read and third-generation sequencing technologies, structural variant detection, graph genomes

Publications

- **Darby, C. A.**, Gaddipati, R., Schatz, M. C., & Langmead, B. (2020). Vargas: heuristic-free alignment for assessing linear and graph read aligners. *Bioinformatics*, btaa265. https://doi.org/10.1093/bioinformatics/btaa265
- **Darby, C. A.**, Stubbington, M. J. T., Marks, P. J., Barrio, Á. M., & Fiddes, I. T. (2020). scHLAcount: Allele-specific HLA expression from single-cell gene expression data. *Bioinformatics*, btaa264. https://doi.org/10.1093/bioinformatics/btaa264
- **Darby, C. A.**, Fitch, J. R., Brennan, P. J., Kelly, B. J., Bir, N., Magrini, V., … Langmead, B., Schatz, M. C. (2019). Samovar: Single-Sample Mosaic Single-Nucleotide Variant Calling with Linked Reads. *IScience*, *18*, 1–10. https://doi.org/10.1016/J.ISCI.2019.05.037
- Sedlazeck, F. J., Lee, H., **Darby, C. A.**, & Schatz, M. C. (2018). Piercing the dark matter: Bioinformatics of long-range sequencing and mapping. *Nature Reviews Genetics*, *19*(6), 329–346. https://doi.org/10.1038/s41576-018-0003-4
- Luo, R., Sedlazeck, F. J., **Darby, C. A.**, Kelly, S. M., & Schatz, M. C. (2017). LRSim: A Linked-Reads Simulator Generating Insights for Better Genome Partitioning. *Computational and Structural Biotechnology Journal*, *15*, 478–484. https://doi.org/10.1016/j.csbj.2017.10.002
- **Darby, C. A.***, Stolzer, M.*, Ropp, P. J., Barker, D., & Durand, D. (2017). Xenolog classification. *Bioinformatics*, *27*(5), btw686. https://doi.org/10.1093/bioinformatics/btw686