Richard J. Abdill

Bioinformatics and bibliometrics

📱 +1-856-266-0764 • ✉ richard.abdill@gmail.com • 🌐 richabdill.com

Education

University of Minnesota **Minneapolis, Minn.**
Ph.D. candidate, Bioinformatics and Computational Biology *2018–2022 (anticipated)*
Supervisor: Dr. Ran Blekhman

University of Maryland **College Park, Md.**
B.A., Print Journalism *2008–2011*

Research

Thesis work (in progress)

"Large-scale curation of human microbiome datasets for machine learning"
A chronic challenge in evaluating the microbial communities that colonize the human gut is the noisy, intercorrelated, and highly variable nature of metagenomic assays—issues that are difficult to mitigate even in the largest studies. To better characterize broad ecological patterns in enteric microbiota, I built an automated pipeline for collecting, processing and integrating publicly available data into **the largest human microbiome dataset in the world**. Rather than evaluating each study separately, this "big data" approach involves processing hundreds of thousands of samples using a uniform pipeline, the output of which can be used not only to characterize variation across the human gut microbiome, but also to **train dimensionality reduction algorithms** that can be applied to smaller datasets. This **transfer learning tool**, informed by patterns identified in the compendium, can be used with much smaller datasets to find meaningful signals that would otherwise be lost in the noise.

Publications

- **2021 (preprint; under review): Abdill RJ**, Adamowicz EM, Blekhman R. Public human microbiome data dominated by highly developed countries. BioRxiv. DOI: 10.1101/2021.09.02.458641.

- **2021 (preprint; under review):** Fink I, **Abdill RJ**, Blekhman R, Grieneisen L. BiomeHorizon: visualizing microbiome time series data in R. BioRxiv. DOI: 10.1101/2021.08.29.458140.

- **2020:** Carneiro CFD, Queiroz VGS, Moulin TC, Carvalho CAM, Haas CB, Rayêe D, Henshall DE, De-Souza EA, Amorim FE, Boos FZ, Guercio GD, Costa IR, Hajdu KL, van Egmond L, Modrák M, Tan PB, **Abdill RJ**, Burgess SJ, Guerra SFS, Bortoluzzi VT, Amaral OB. Comparing quality of reporting between preprints and peer-reviewed articles in the biomedical literature. *Research Integrity and Peer Review*. DOI: 10.1186/s41073-020-00101-3.

- **2020: Abdill RJ**, Adamowicz EM, Blekhman R. International authorship and collaboration across bioRxiv preprints. *eLife* 9:e58496. DOI: 10.7554/eLife.58496.

- **2019:** Mangul S, Mosqueiro T, **Abdill RJ**, Duong D, Mitchell K, Sarwal V, Hill B, Brito J, Littman RJ, Statz B, Lam AKM, Dayama G, Grieneisen L, Martin LS, Flint J, Eskin E, Blekhman R. Challenges and recommendations to improve installability and archival stability of omics computational tools. *PLOS Biology* 17(6): e3000333. DOI: 10.1371/journal.pbio.3000333.

- **2019: Abdill RJ** & Blekhman R. Rxivist.org: Sorting biology preprints using social media and readership metrics. *PLOS Biology* 17(5): e3000269. DOI: 10.1371/journal.pbio.3000269.

- **2019: Abdill RJ** & Blekhman R. Tracking the popularity and outcomes of all bioRxiv preprints. *eLife* 8:e45133. DOI: 10.7554/eLife.45133.

Conference proceedings

- **2021:** Microbiome Data Congress. Oral presentation: "Resolving pipeline installation conflicts with containerization." Virtual conference, 9–10 June 2021.
- **2021:** Bioinformatics and Computational Biology Research Symposium. Poster presentation: "Measuring large-scale variation in the human microbiome using machine learning." Virtual conference (University of Minnesota), 14 Jan 2021.
- **2019:** Computational Genomics Summer Institute. Oral presentation: "The perilous geometry of compositional data." Los Angeles, Calif., 15–19 July 2019.
- **2019:** Great Lakes Bioinformatics Conference. Oral presentation: "Tracking the popularity and outcomes of all bioRxiv preprints." Madison, Wisc., 19–22 May 2019.

Software and tools

GitHub profile: `https://github.com/rabdill`

- **Rxivist** – A web application (Python) and data repository (PostgreSQL) that indexes all bioRxiv preprints and enables sorting based on Twitter and download activity. `https://rxivist.org`
- **PanLingua** – A web application built on Google machine-translation services that enables search and reading of bioRxiv preprints in 104 world languages. `https://panlingua.rxivist.org`
- **UMN OB-GYN Research Database** – A database (MS Access) and reporting tool (Python and Javascript) for the University of Minnesota Medical School to track publications and grants.
- **DADA2 Docker image** – A publicly available Docker image to help users process amplicon data with the R tool "DADA2." `https://hub.docker.com/r/blekhmanlab/dada2`

Teaching experience

- **GCD 8920, Genomic Analysis:** Worked as a teaching assistant for this graduate-level course and taught class sessions covering quality control of genetic sequencing data, genome-wide association studies and RNA-seq analysis. *University of Minnesota, spring 2020.*

Service and outreach

- **Ad-hoc peer review**: *eLife*, *Journal of Open Source Software*, *Nature Biotechnology* (supervised), *PLOS Computational Biology*, *PLOS ONE*, *PLOS Biology* (supervised).
- **Skype A Scientist**, 2019–present. Scientist participant in this program that matches researchers with primary and secondary school classrooms to facilitate discussions about scientific research.
- **Technovation Girls**, 2020. Senior-level judge in this international competition for young girls who learn to develop mobile applications to solve problems in their communities.
- **Microbiome Digest**, 2020–2021. Volunteer author at `microbiomedigest.com`, a daily aggregator of new microbiome research.
- **Tutoring**, 2020. Volunteer tutor at the International Institute of Minnesota's College Readiness Academy, which helps prepare recent immigrants for college-level courses.

Honors and awards

2019:
- Award, best talk – Great Lakes Bioinformatics Conference, Madison, Wisc.
  - Travel fellowship – Great Lakes Bioinformatics Conference
- Travel award – Triangle Scholarly Communication Conference, Durham, NC
- Travel award – Chan Zuckerberg Initiative Open Science Workshop, San Francisco, Calif.

Employment

Full employment history available upon request.

- **Dimensions** **https://dimensions.ai**
  Senior software engineer *2021–present*
- **Minnesota Supercomputing Institute** **Minneapolis, Minn.**
  System programmer *2018*
- **Gannett/*USA Today*** **McLean, Va.; Minneapolis, Minn.**
  Software engineer *2014–2015; 2016–2018*
- **Target** **Minneapolis, Minn.**
  Lead engineer *2015–2016*
- **Amazon Web Services** **Herndon, Va.**
  DevOps consultant *2015*
- **Village Voice Media** **Fort Lauderdale, Fla.**
  Staff writer *2012*