Ethan Alexander García Baker

ethanagbaker@pitt.edu
ethanagbaker.com

SUMMARY

Undergraduate student at the University of Pittsburgh pursuing a Bachelor of Philosophy in Neuroscience and a Bachelor of Arts in History and Philosophy of Science with minors in Computer Science and Chemistry. Interests include application of novel computational and data science methods to NGS datasets to study human neuropsychiatric disorders and cancers, development of genomics-based precision medicine technologies, and issues of ethical use and equitable access in genomics.

EDUCATION

Doctor of Philosophy, Biology, Massachusetts Institute of Technology, Cambridge, MA
Beginning August 2017

April 2017 - Bachelor of Philosophy, Neuroscience, University of Pittsburgh, Pittsburgh, PA
GPA: 3.73
Thesis: *Characterizing* anticipatory control of autonomic regulatory processes in conscious animal models: a neuroinformatic approach.
Supervisor: Billy J. Yates, Ph.D. (Departments of Neuroscience and Otolaryngology)

April 2017 - Bachelor of Arts, History and Philosophy of Science, University of Pittsburgh, Pittsburgh, PA (2017)

PUBLICATIONS

Baker, Ethan Alexander García, Goodwin, S., McCombie, W.R., and Ramos, O.M. "SiLiCO: A Simulator of Long Read Sequencing in PacBio and Oxford Nanopore". BioRxiv. doi: 10.1101/07690

Baker, Ethan Alexander García, Sezen, U., Falk, T., Maloney P., Vogler, D. ,Jensen, C., Mitton, J., Wright, J., Knaus B., Cronn, R., Rai, H., Gonzalez-Ibeas,D., Vasquez-Gross, H., Famula, R., Liu, J., Kueppers, L., Neale D., Wegrzyn, J.L. "Comparative transcriptomics among four white pine species." *G3: Genes, Genomes, and Genetics*. (Under review)

Baker, Ethan Alexander García, Patel, N., Miller, D., and Yates, Billy J**.** "Characterizing anticipatory autonomic responses in conscious animal models.", *J. Neurophysiology.* (Manuscript in preparation)

Baker, Ethan Alexander García, Ramos, O.M., Eskipehlivan, S.M., Goodwin, S., Antoniou, E., and McCombie, W.R. "Comparative analysis of PacBio libraries reveals non-stochastic biases in sites of DNA nicking". (Manuscript in preparation)

ACHIEVEMENTS AND AWARDS

2016 - Barry M. Goldwater Scholarship Honorable Mention
2016 - Chancellor's Undergraduate Research Fellowship (Spring)
2016 - University Honors College Travel Grant
2016 - Office of Undergraduate Research Travel Grant
2015 - Brackenridge Undergraduate Research Fellowship (Fall)
2015 - William Shakespeare Fellowship, Cold Spring Harbor Laboratory
2013 - Present - Dean's List
2013 - Honors College Full Tuition Scholarship, University of Pittsburgh

RELEVANT EXPERIENCE

May 2016-August 2016 - Computational Genomics Intern, Gerstein Lab, Departments of Computer Science and Molecular Biology & Biochemistry, Yale University, New Haven, CT
Developed pipelines for genome-wide enhancer discovery and annotation via massively-parallel peak calling in STARR-seq datasets. Applied pipelines to very large (4-5x existing datasets) datasets from the emerging Enhancer-seq assay to build enhanced gene regulatory networks and map regulatory regions.

January 2014 - Present — B. Phil. Candidate, Yates Lab, Department of Otolaryngology, University of Pittsburgh School of Medicine, Pittsburgh, PA

Developing novel computational strategies to process mass quantities of neural recording data to build a computational model of neural pathways implicated in anticipatory autonomic regulation. Conducting neurophysiological study on the rostral ventrolateral medulla to determine mechanisms of anticipatory regulation of blood flow in response to vestibular stimuli. Examining neuronal inputs to vestibular systems to deduce autonomic control pathways implicated in balance disorders and maintenance of homeostasis.

August 2015-September 2016 - Research Intern
June-Aug 2015—William Shakespeare Fellow, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
Recipient of highly competitive fellowship to conduct independent research in computational genomics of cancer and psychiatric disorders under Dr. W. Richard McCombie. Worked to develop bioinformatic and bench approaches to better characterize the mechanics of next-generation sequencing library failure to build better genomic diagnostic tools.

September 2015-Present — Bioinformatics Analyst, Department of Developmental Biology, Children's Hospital of Pittsburgh of UPMC, Pittsburgh, PA
Process and manage bioinformatic data pipelines, perform variant discovery, functional annotations, and expression analyses. Developing analysis pipeline for variant calling for single-cell RNA-seq data. Design figures for inclusion in publications.

May-December 2014 — Student Research Specialist, Computational Genomics Laboratory, Department of Ecology & Evolutionary Biology, University of Connecticut, Storrs, CT
Assembled and analyzed transcriptomes of various white pine species. Performed the first comparative transcriptomic analysis in the white pines to identify conserved gene families. Identified genes corresponding to resistance to disease and climate change under positive selection as targets for breeding programs.

September 2012 – March 2013 — Student Researcher, Gogarten Laboratory, Department of Molecular and Cell Biology, University of Connecticut, Storrs, CT
Studied molecular evolution with a special focus on the role on inteins and homing endonucleases using computational methods.

SKILLS

Bioinformatics: genome and transcriptome assembly, functional annotation, clustering, phylogenetic analysis, sequence alignment, variant discovery (GATK), quality control, single-cell transcriptomics, in-silico read simulation, cluster management
Statistics: R and Minitab
Programming Languages: Python, MATLAB, R, UNIX shell, Java, Perl (basic)
Applications: LaTeX, Jekyll, Git, Jupyter, Octave, ggplot2, matplotlib,
Operating Systems: UNIX, Linux, Mac OS X, Windows
Laboratory Techniques: electrophysiology experimental design in large animal models, NGS library preparation, TUNEL assays, gel electrophoresis, SDS-PAGE
Language: English (native), Spanish (fluent)

ORAL PRESENTATIONS

February 2016 — Poster in Advances in Genome Biology and Technology (AGBT), Orlando, FL
"Comparative analysis of PacBio libraries reveals non-stochastic biases in sites of DNA nicking"

February 2016 — Poster in Advances in Genome Biology and Technology (AGBT), Orlando, FL
"New approaches to rescue poorly performing PacBio libraries by quantifying and repairing single stranded DNA nicks"

January 2016 — Poster in Plant and Animal Genome (PAG) XXIV, San Diego, CA
"Comparative Transcriptomics Among Four White Pine Species"

CONTINUING EDUCATION

The Data Scientist's Toolbox, Johns Hopkins University via Coursera (2015)

R Programming, Johns Hopkins University via Coursera (2016)
Machine Learning, Stanford University via Coursera (In Progress)

COMMUNITY ENGAGEMENT
November 2015 — Guest Speaker, Upper Saint Clair High School, Upper Saint Clair, PA
Taught 9th grade biology course about next-generation sequencing, genomics, and bioinformatics.
September 2016 - Introduction to Genomic Data Science
Begun development of supplemental activities for high school biology students introducing elementary concepts of data science, particularly in relation to genomics, data visualization, and bioinformatics.

VOLUNTEER
Student Health Advisory Board, President (2014-Present)

Birmingham Free Clinic, Community Health Fellow (January 2015-Present)

Salud Para Niños, Volunteer (October 2015-Present)

VA Pittsburgh Healthcare, Veteran Experience Volunteer (September 2014-Present)