Ethan Alexander García Baker

eagb@mit.edu
ethanagbaker.com

SUMMARY

Graduate (PhD) student in the Department of Biology at the Massachusetts Institute of Technology and a graduate of the University of Pittsburgh. Interested in gene regulatory networks, biomedical data science, design of computational methods for NGS data analysis for precision medicine, and issues of equitable access to emerging medical technologies.

EDUCATION

Doctor of Philosophy, Biology   September 2017 - Present
Massachusetts Institute of Technology, Cambridge, MA, USA

Bachelor of Science, Neuroscience   September 2013 - April 2017
Magna Cum Laude
University of Pittsburgh, Pittsburgh, PA, USA
Thesis: Characterizing anticipatory control of autonomic regulatory processes in conscious animal models: a neuroinformatic approach
Supervisor: Billy J. Yates, Ph.D.

Bachelor of Arts, Philosophy of Science   September 2013 - April 2017
Magna Cum Laude
University of Pittsburgh, Pittsburgh, PA, USA

PUBLICATIONS

1. **Ethan A.G. Baker**, J.L. Wegrzyn, U. Sezen, T. Falk, P. Maloney, D.Vogler, A. Delfino-Mix, C. Jensen, J. Mitton, J. Wright, B.Knaus, H. Rai, R. Cronn, D. Gonzalez Ibeas, H.Vasquez-Gross, R. Famula, J. Liu, L.Keuppers, and D. Neale. Comparative transcriptomics among four white pines. *G3*. 2018. *(Accepted, Awaiting Publication)*

2. N. M. Patel, **E.A.G. Baker**, S.R. Wittman, I.C. Engstrom, G.H. Bourdages, A.A. McCall, D.M. Miller, Bill. J. Yates. Cardiovascular adjustments during anticipated postural changes. Physiol Rep, 6 (1), 2018.

3. **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.* (2017) doi: 10.1101/07690

FELLOWSHIPS AND AWARDS

2017 - National Science Foundation Graduate Research Fellowship (NSF GRFP)
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 - Honors College Full Tuition Scholarship, University of Pittsburgh

RELEVANT EXPERIENCE

Computational Genomics Intern, Gerstein Lab   May 2016 - August 2016
Departments of Computer Science and Molecular Biology & Biochemistry
Yale University, New Haven, CT, USA

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.

Undergraduate Researcher, Yates Lab   January 2014 - April 2017
Department of Otolaryngology
University of Pittsburgh School of Medicine, Pittsburgh, PA, USA
Developed computational strategies to process mass quantities of neural recording data to build a computational model of neural pathways implicated in anticipatory autonomic regulation. Conducted neurophysiological study on the rostral ventrolateral medulla to determine mechanisms of anticipatory regulation of blood flow in response to vestibular stimuli.

Bioinformatics Analyst   September 2015 - February 2017
Department of Developmental Biology
Childrens Hospital of Pittsburgh of UPMC, Pittsburgh, PA
Processed and managed bioinformatic data pipelines, perform variant discovery, functional annotations, and expression analyses. Developed analysis pipeline for variant calling for single-cell RNA-seq data. Designed figures for inclusion in publications.

Research Intern   August 2015-September 2016
William Shakespeare Fellow   June 2015 - August 2015
Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA
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.

Student Research Specialist   May 2014 - December 2014
Department of Ecology & Evolutionary Biology
University of Connecticut, Storrs, CT, USA
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.

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
Language: English (native), Spanish (intermediate fluency)

ORAL PRESENTATIONS

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

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

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

COMMUNITY ENGAGEMENT

Guest Speaker, Upper Saint Clair High School, Upper Saint Clair, PA
November 2015
Taught 9th grade biology course about next-generation sequencing, genomics, and bioinformatics.

Introduction to Genomic Data Science
September 2016
Began development of supplemental activities for high school biology students introducing elementary concepts of data science, particularly in relation to genomics, data visualization, and bioinformatics.