Matthew H. Bailey

Cell: (801) 592-8253 | Email: m.hawkins.bailey@gmail.com | GitHub: https://github.com/MHBailey

PROFILE

- Led team projects to completions with The Cancer Genome Atlas (TCGA) and International Cancer Genetics Consortia (ICGC).
- Experienced in two aging phenotypes: Alzheimer's disease and Pan-Cancer gene discovery analysis.
- Regularly analyze genomic datasets (whole genome, whole exome, GWAS, and pedigrees) with 10,000+ samples.
- Expert in big-data figure generation: created 17 figures and 3 graphical abstracts for 3 TCGA PanCancer Atlas manuscripts (2018) and 6 figures for ICGC PanCancer analysis working groups (2020).
- 8 years of experience assisting and leading international genomic collaboration.
- Realized personal goal of implementing personalized precision oncology strategies into a clinical setting

EDUCATION

Ph.D. Washington University in St. Louis: *Mentor: Li Ding.* Human and Statistical Genetics. 2013-2018
B.S. Brigham Young University: *Mentor: John S. K. Kauwe.* Bioinformatics. 2009-2013

COLLABORATIONS & PROJECT MANAGEMENT

- Patient derived xenograft network (PDXnet): Designed a blinded drug in-vivo validation experiment from predictions made using expression biomarkers. We longitudinally tested specific drugs using PDX models of inter-institutional collaborators. Assisted in quality control teams and data-analysis teams. 2019-Present
- International cancer genetics consortia (ICGC): Assisted in validation sequencing for the 64 initial whole genome samples for the mutation-calling group and explored the intersection of somatic variant calls for 750 samples with WGS and WES technology. 2015-2020
- The Cancer Genome Atlas (TCGA): PanCanAtlas: Led two projects that resulted in two Cell publications. Assisted in consensus mutation calling for 10,000+ samples, contributed to the MC3 (mutation calling), SARC-AWG (sarcoma analysis working group), and CESC-AWG (cervical cancer analysis working group). 2014-2018
- Alzheimer's Disease Neuroimagine Initiative & AD genetics consortia (ADNI & ADGC): Curated CSF-biomarker phenotypes, performed GWAS analysis using PLINK and supplied Manhattan plots for main text. 2011-2013

RESEACH EXPERIENCE

- *Postdoctoral Research Associate* University of Utah and Huntsman Cancer Institute. Under the direction of Gabor Marth, Bryan Welm and Alana Welm, integrated multi-omics platforms with patient derived organoid drug responses into clinically actionable findings. Fall 2018 - Present
- *Graduate student* Department of Medicine and Oncology Washington University in Saint Louis. Led multi-institutional consortium to produce Pan-Cancer genetic research under the direction and mentorship of Li Ding. Thesis, "A tail of two distributions: The landscape of positive and negative selection in cancer." Rotations with Mike Province, Don Conrad, Zac Pincus, Alison Goate, and Li Ding Fall 2013-Fall 2018
- *Research assistant* College of Life Sciences BYU. Established a strong foundation in data cleaning and computational research. Performed genome-wide-association studies in Alzheimer's disease endophenotypes. Fall 2010-Spring 2013
- *Research assistant* University of Utah. Learned the challenges and advantages of linkage studies using the Utah Population Database with Lisa Cannon Albright, a co-author of the BRCA2 breast cancer discovery. Summer 2012
- *Research fellow* Simmons Center for Cancer Research BYU. Re-applied my skills in running genome-wide-association studies using relevant cancer proteins as endophenotypes. Summer 2011

SERVICE & OUTREACH

- American Association for Cancer Research — Winter 2017
- American Society of Human Genetics — Fall 2012
- Americorp Member: Volunteered and worked at the Boys and Girls club. Enjoyed developing relationships with foster kids and children from less fortunate upbringings. — Summer 2009-2010
- Church of Jesus Christ of Latter-day Saints Mission to Paris, France: Learned French, the importance of hard work, dedication, and the impact of meaningful service. — 2006-2008

PUBLICATIONS (Sorted by year and category, *equal contribution, Google Scholar: Matthew H. Bailey)

Translational medicine

1. Katrin P. Guillen, Andrew J. Butterfield, Sandra D. Scherer, **Matthew H. Bailey**, Zhengtao Chu, Yoko S. DeRose, Yi Qiao, Jeffery Vahrenkamp, Emilio Cortes-Sanchez, Chieh-Hsiang Yang, Satya S. Pathi, Fadi Haroun, Diane Hernandez, Gabor T. Marth, Jeff Chaung, Michael T. Lewis, Jason Gertz, Katherine E. Varley, Alana L. Welm, Bryan E. Welm, Predicting breast cancer therapy response using a patient-derived xenograft organoid screening platform — Cell 2020 *In preparation*
2. Katrin P. Guillen, Sandra D. Scherer, Andrew J. Butterfield, Ling Zhao, Yoko S. DeRose, **Matthew H. Bailey**, Alana L. Welm, Bryan E. Welm. Generating Long-Term Breast Cancer Organoid Cultures from Patient-Derived Xenografts for improved throughput in determining therapeutic response — STAR Protocols 2020 *In preparation*

Whole genome and multi-'omic' projects

3. **Matthew H Bailey**, William U Meyerson, Lewis Jonathan Dursi, Liang-Bo Wang, Guanlan Dong, Wen-Wei Liang, Amila Weerasinghe, Shantao Li, Sean Kelso, MC3 Working Group, Novel somatic mutation calling methods, Gordon Saksena, Kyle Ellrott, Michael C Wendl, David Wheeler, Gad Getz, Jared T Simpson, Mark B Gerstein, Li Ding, PCAWG Consortium. Retrospective evaluation of whole exome and genome mutation calls in 746 cancer samples. — Nature Communication 2020 *Accepted in principle*
4. The ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Consortium. Pan-cancer analysis of whole genomes. Contributions: Deep sequencing validation for WGS caller performance. — Nature 2020
5. Antonio Colaprico, Catharina Olsen, **Matthew H Bailey**, ... Xi Steven Chen, Elena Papaleo. Interpreting pathways to discover cancer driver genes with Moonlight. — Nature Communication 2019
6. Li Ding*, **Matthew H. Bailey***, Eduard Porta-Pardo*, Vesteinn Thorsson, … Vincent, Carolyn M Hutter, Jean Claude Zenklusen, Katherine A Hoadley, Michael C. Wendl, llya Shmulevich, Alexander J. Lazar, David Wheeler, Gad Getz, The Cancer Genome Atlas Research Network Perspective on Oncogenic Processes At the End of the Beginning of Cancer Genomics. — Cell 2018

Exome sequencing projects

7. Galen F Gao, Joel S Parker, Sheila M Reynolds, Tiago C Silva, Liang-Bo Wang, Wanding Zhou, Rehan Akbani, **Matthew Bailey**, ... The Genomic Data Analysis Network, Han Liang, Michael S Noble. Before and after: Comparison of legacy and harmonized TCGA genomic data commons' data. — Cell Systems 2019
8. Xing Yi Woo, Jessica Giordano, Anuj Srivastava, Zi-Ming Zhao, Michael W Lloyd, Roebi de Bruijn, Yun-Suhk Suh, Rajesh Patidar, Li Chen, Sandra Scherer, **Matthew Bailey**, ... Jeffrey H Chuang, PDXNet Consortium, EurOPDX Consortium Conservation of copy number profiles during engraftment and passaging of patient-derived cancer xenografts — BioRxiv 2019
9. **Matthew H. Bailey***, Collin Tokheim*, Eduard Porta-Pardo*, … The Cancer Genome Atlas Research Network, Michael Lawrence, Adam Godzik, Nuria Lopez-Bigas, Josh Stuart, David Wheeler, Gad Getz, Ken Chen, Alexander J. Lazar, Gordon B Mills, Rachel Karchin, Li Ding. Comprehensive Characterization of Driver Genes and Mutations in Cancer. — Cell 2018
10. Kyle Ellrott, **Matthew H. Bailey,** Gordon Saksena, … Gad Getz, David Wheeler, Li Ding, The Cancer Genome Atlas Research Network. Scalable Open Science Approach for Mutation Calling of Tumor Exomes Using Multiple Genomic Pipelines. — Cell Systems 2018
11. Francisco Sanchez-Vega, Marco Mina, Joshua Armenia, Walid K. Chatila, Augustin Luna, Konnor La, Sofia Dimitriadoy, David L. Liu, Havish S. Kantheti, Sadegh Saghafinia, Debyani Chakravarty, Foysal Daian, Qingsong Gao, **Matthew H. Bailey**, … Andrew D. Cherniack, Giovanni Ciriello, Chris Sander, Nikolaus Schultz. Oncogenic Signaling Pathways in The Cancer Genome Atlas. — Cell 2018
12. Yanan Cao, Weiwei Zhou, Lin Li, Jiaqian Wang, Zhibo Gao, Yiran Jiang, Xiuli Jiang, Aijing Shan, **Matthew H Bailey**, Kuan-lin Huang, Sam Q Sun, Michael D McLellan, Beifang Niu, Weiqing Wang, Li Ding, Guang Ning. Pan-cancer analysis of somatic mutations across 21 neuroendocrine tumor. — Cell Research 2018

13. Yanan Cao, Weiwei Zhou, Lin Li, Jiaqian Wang, Zhibo Gao, Yiran Jiang, Xiuli Jiang, Aijing Shan, **Matthew H Bailey**, Kuan-lin Huang, Sam Q Sun, Michael D McLellan, Beifang Niu, Weiqing Wang, Li Ding, Guang Ning. Pan-cancer analysis of somatic mutations across 21 neuroendocrine tumor. — Cell Research 2018
14. The Cancer Genome Atlas Research Network Comprehensive and Integrated Genomic Characterization of Adult Soft Tissue Sarcomas. Contributions: **M.H.B** (Mutation and SMG analysis) — Cell 2017
15. The Cancer Genome Atlas Research Network, Contributions: **M.H.Bailey** (mutation analysis) Integrated genomic and molecular characterization of the cervical cancer. — Nature 2017
16. Beifang Niu, Adam D Scott, Sohini Sengupta, **Matthew H Bailey**, … Li Ding, Protein-structure-guided discovery of functional mutations across 19 cancer types. — Nature Genetics 2016

GWAS projects

17. Lyndsay A. Staley, Mark T. W. Ebbert, Daniel Bunker, **Matthew Bailey**, … Perry G. Ridge, Alison M. Goate and John S. K. Kauwe. Variants in ACPP are associated with cerebrospinal fluid Prostatic Acid Phosphatase levels. — BMC Genomics 2016
18. Lyndsay A. Staley, Mark T. W. Ebbert, Sheradyn Parker, **Matthew Bailey**, … Perry G. Ridge, Alison M. Goate and John S. K. Kauwe. Genome-wide association study of prolactin levels in blood plasma and cerebrospinal fluid. — BMC Genomics 2016
19. Yuetiva Deming, Jian Xia, Yefei Cai, Jenny Lord, Jorge L Del-Aguila, Maria Victoria Fernandez, David Carrell, Kathleen Black, John Budde, ShengMei Ma, Benjamin Saef, Bill Howells, Sarah Bertelsen, **Matthew Bailey**, … Carlos Cruchaga, Alzheimer's Disease Neuroimaging Initiative (ADNI). Genetic studies of plasma analytes identify novel potential biomarkers for several complex traits. — Scientific reports 2016
20. Kauwe, John SK, **Bailey, Matthew H**, … Celeste M, Harari, Oscar, Cruchaga, Carlos. Genome-wide association study of CSF levels of 59 Alzheimer's disease candidate proteins: significant associations with proteins involved in amyloid processing and inflammation. — PLoS Genetics 2014
21. Sharp, Aaron R, Ridge, Perry G, **Bailey, Matthew H**, … Kauwe, John SK. Population substructure in Cache County, Utah: the Cache County study. — BMC bioinformatics 2014
22. Ebbert, Mark TW, Ridge, Perry G, Wilson, Andrew R, Sharp, Aaron R, **Bailey, Matthew**, …, Kauwe, John SK. Population-based analysis of Alzheimer's disease risk alleles implicates genetic interactions. — Biological psychiatry 2014
23. Cao, Ying, Wei, Peng, **Bailey, Matthew,** Kauwe, John SK, Maxwell, Taylor J. A versatile omnibus test for detecting mean and variance heterogeneity. — Genetic epidemiology 2014
24. Carlos Cruchaga, John SK Kauwe, Oscar Harari, Sheng Chih Jin, Yefei Cai, Celeste M Karch, Bruno A Benitez, Amanda T Jeng, Tara Skorupa, David Carrell, Sarah Bertelsen, **Matthew Bailey**, … Anne M Fagan, David M Holtzman, John C Morris, Alison M Goate, GERAD Consortium, Alzheimer's Disease Neuroimaging Initiative (ADNI, Alzheimer Disease Genetic Consortium). GWAS of cerebrospinal fluid tau levels identifies risk variants for Alzheimer's disease. — Neuron 2013

Family based and molecular projects

25. Manda, KR, Tripathi, P, Hsi, AC, Ning, J, Ruzinova, MB, Liapis, H, **Bailey, M**, … Feng, C. NFATc1 promotes prostate tumorigenesis and overcomes PTEN loss-induced senescence. — Oncogene 2015
26. Cannon-Albright, Lisa A, Farnham, James M, **Bailey, Matthew,** … A, Thomas, Alun. Identification of specific Y chromosomes associated with increased prostate cancer risk. — The Prostate 2014
27. Ridge, Perry G, Maxwell, Taylor J, Foutz, Spencer J, **Bailey, Matthew H**, … Cawthon, Richard M, Kauwe, John SK. Mitochondrial genomic variation associated with higher mitochondrial copy number: the Cache County Study on Memory Health and Aging. — BMC bioinformatics 2014
28. Ridge, Perry G, Koop, Andre, Maxwell, Taylor J, **Bailey, Matthew H**, … Kauwe, John SK, Honea, Robyn A. Mitochondrial haplotypes associated with biomarkers for Alzheimer's disease. — PloS one 2013

FUNDING & AWARDS

- T32 Training Grant: Two-year NIH Ruth L. Kirschstein National Research Service Award from the National Human Genome Research Institute Training in Genomic Medicine. $100,008 — Summer 2019-Present
- Precision Medicine Pathway: Invited fellowship where I learned from physicians on how they implement genomics in their current research. $2,000 — Fall 2014-Fall 2018
- Myriad Genetics Scholarship: Was awarded a financial scholarship for my work in the field of Alzheimer's disease in the department of biology. $5,000 — Spring 2012

- Travel Award - Alzheimer's Disease International Conference: Awarded sponsored attendance to conference. Registration fee — Summer 2011 & 2012
- AmeriCorps Education Award: Tuition scholarship awarded for my work as an AmeriCorps member at the Boys and Girls Club of Provo. $2,500 — Spring 2010

ORAL PRESENTATIONS

- Retreat Talk, St. Louis, Missouri: Tri-department retreat student talks. Presented on "The forefront of driver gene and mutation discovery using TCGA data." — Fall 2017
- ICGC Face to Face, Boston, Massachusetts: International cancer genetics consortia talk at the BROAD Institute at Harvard. Contrasted whole exome and whole genome mutations from ~750 matched samples. — Fall 2016
- ICGC Face to Face, Aptos, California: International cancer genetics consortia talk hosted by UCSC. Presented ultra-deep sequencing validation for 64 whole genome sequenced samples. — Summer 2015
- Graduate Student Qualification Exam, St. Louis, Missouri: Required oral exam for Ph.D. candidacy. Presented on recent finding and the genetic application of synthetic lethality in cancer treatment. — Fall 2015
- Cancer Research Center BYU, Provo, Utah: End of summer oral presentation. Spoke on the relevance of genome-wide-association studies to the progression of finding cures to cancer. — Summer 2011
- President's Leadership Council, Provo, Utah: Invited to give a presentation to 200 university donors, and the university president about how the Kauwe Lab was contributing to curing Alzheimer's disease. — Spring 2011

POSTER PESENTATIONS

Translational medicine

- Center for Genomic Medicine: Utah, United States: Precision drug selection predicts response for advanced breast cancer patients (better poster format) — Winter 2020
- Biology of Genomes: Cold Spring Harbor, United States: Organoids maintain molecular properties of xenograft tumor and drug screening of organoids support biomarker discovery. — Summer 2019

Whole genome and whole exome sequencing projects

- American Association of Cancer Research, Chicago, United States: Characterized 750 samples that were sequenced by both whole genome and exome technologies. — Spring 2018
- Human and Statistical Genetics Retreat, St. Louis, United States: Presented on National Science Foundation grant proposal and thesis ideas. — Fall 2014 & 2015

GWAS studies

- Biotechnology and Bioinformatics Symposium, Provo, United States: Presented on preferential disequilibrium in the *APOE* region. This work led to foundation work in trying to identify individuals with *APOE ε4* allele that don't get Alzheimer's disease. — Fall 2012
- International Conference on Alzheimer's Disease, Vancouver, Canada: Genome-wide screen for variants that modify the known relationship between cerebrospinal fluid amyloid-42 (beta) and ptau levels – interaction modifiers. — Summer 2012
- Annual Mentored Learning Conference & Student Poster Sessions, Provo, United States. F2 -isoprostane (oxidative stress) cerebrospinal fluid levels as an interaction modifier in Alzheimer's genomics. — Spring 2012
- International Conference on Alzheimer's Disease, Paris, France: Presented on correlations between top Alzheimer's genetic variants and cerebral spinal fluid levels. — Summer 2011
- Annual Mentored Learning Conference & Student Poster Sessions, Provo, United States: Presented on association between top Alzheimer's disease genetic variants and cerebral spinal fluid levels. — Spring 2011

Family-base genomics

- American Society Human Genetics, San Diego, United States: Presented on data collected by the Long-Life Family study describing allele persistence in Y-Chromosome markers. Fall 2014
- American Society Human Genetics, San Francisco, United States: Presented research on melanoma genetic modifiers in p16 gene mutations pedigrees Fall 2012

TEACHING

- *Guest speaker* Taught fledgling medical students in their Foundations in Personalized Health Care course about the role genomics and real-time analytics will play in the future. Fall 2019
- *Teacher's assistant* Helped implement a Python based auto-grading software called nbgrader which sped up grading process. Assisted with in-class demos and filming. Course name: Introduction to computational biostatistics. Fall 2017 & Fall 2016
- *Teacher's assistant* Taught 2 lectures and assisted with updating lectures and grading student assignments. Course name: Human linkage and association Spring 2014
- *Teacher's assistant* Worked behind the scenes to ensure the class ran smoothly, grading assignments and resolving student concerns. Course name: Introduction to communications. Fall 2010-Spring 2011
- *Youth development professional.* Responsible for young teenagers' safety, helped develop strong social skills, and reported any inappropriate incident. AmeriCorps/Boys and Girls Club Fall 2009-Spring 2010

LANGUAGES

Advanced

Scripting and coding: Python, Perl, snakemake, and SQL
Statistical: R and SAS
Technology specific tools:

- *SNParray*: PLINK, Gemma, EMMAX, PedGenie, PennCNV, and GADA
- *Exome sequencing*: MuSiC, HotSpot3D, Bedtools, Samtools, VEP, and vcf2maf, SevenBridges
- *Immunogenomics:* ATHLATES, PvacSeq, and HLAminer
- *Drug screening:* GRmetrics

Graphics: R, Circos, Adobe Illustrator, and Adobe InDesign
Operating systems: OSX, Linux, and Windows
Version control: GitHub
Spoken: English (native), French

Experience

Pipeline: GATK
Cloud: Google Cloud

Wet lab techniques: Q5 mutagenesis, PCR, lipofectamine transfection, DNA mini-prep, enzyme digestion, western-blot, microscopy, and cell culture.

REFERENCES

Name	Contact information	Relationship
Li Ding Ph.D.	lding@wustl.edu	Research mentor
Alex Lazar MD	alazar@mdanderson.org	Collaborator
Bryan Welm Ph.D.	Bryan.Welm@hci.utah.edu	Postdoc mentor