

JAMES DALGLEISH

MOLECULAR BIOSTATISTICS, MACHINE LEARNING, EPIGENOMICS, & CLINICAL BIOINFORMATICS

EDUCATION

Graduate Studies: Tulane University School of Medicine, Department of Cellular and Structural Biology

- **Anatomy Certification Program:** Completed Certificate Took medical student anatomy using dissection/microscopy
- Passed Gross Anatomy USMLE subject test

BACHELOR OF SCIENCE, MOLECULAR BIOLOGY Brigham Young University, Provo Utah, GPA 3.80, Major GPA 3.94

- **MCAT Score:** 38 overall (99th percentile)
 - 13 Physical Sciences (quantitative section) subscore (97th percentile),
 - 15 Biological Sciences (content based, **Perfect score**, percentile rank 100)

RESEARCH EXPERIENCE

CRTA Research Fellow, National Institutes of Health, National Cancer Institute, Bethesda, MD

2017-current

- Analyzed raw NGS chip and RNA (QC, alignment, peak calling, bedgraph comparisons), HiC data.
- Wrote an R package of many functions (HiClink) to simplify analysis and link Hi-C (3D DNA data) data types, tie them to known enhancer sites and RNAseq fold change data, and visualize them alongside ChIP tracks.
- Extensive knowledge of Hi-C data pipelines and many tools (including HiC QC), experience analyzing many such DNA loop datasets.
- Analyzed differential ChIP-seq, RNA with multifactor experimental designs, and differential loop data.
- Current work:
 - Integrative Analysis of Time-course ChIP-seq and RNA-seq to find patterns of TF binding over time in relation to their known target genes.
 - A highly-efficient method to pull down public data and combine with laboratory ChIP-seq experiments to swiftly find regions of close overlap that suggest presence of several transcription factors in a single region.
 - This method assess any number of chipseq runs in a combinatorial fashion
 - This method is faster and more robust than current tools to analyze multiple ChIP-seq runs.
 - Relevance: this method will allow the knowledge of which genes act in concert to drive expression of oncogenes, which can suggest new drug targets and possible clinical trials with existing drugs.

IRTA Research Fellow, National Institutes of Health, Bethesda, MD

2016-2017

- Primary author of a project analyzing the histopathology, proteomics, and Gene Expression of a cohort of patients with Nephrotic Syndrome (NS).
 - Used hierarchical clustering and a detailed analysis to elucidate a new therapeutic from existing drugs.
 - Developed superior linear and elastic net models that suggest new Kidney urine/serum biomarkers.
 - Also tested a replication RNAseq and Proteomic data that suggest utility in clinical diagnosis.
 - Promised first authorship during the internship
- Developed a novel statistical method of analyzing quantitative RNAseq data with continuous histopathological correlates.
- Prepared abstract for National Conference.
- Awarded co-authorship on a publication that involved TF-DNA bioinformatics prediction and validation.
 - This paper is currently a large manuscript under revision and will be submitted soon.
- Developed collaborative relationships within the KDS, in other NIH Institutes, and externally (UMich and University of Rochester).

Alder Laboratory (Telomerase), Department of Physiology and Developmental Biology, Brigham Young University **2014-2015**

Bioinformatics Analysis of Telomerase with traditional bench components in collaboration with Carol Grieder, **Nobel Laureate** and discoverer of Telomerase. I will be granted **co-authorship on the publication**.

- Used bioinformatics tools, coded in R and Ruby, literature, and current databases to find putative cancer driver molecules/mechanisms. Designed and orchestrated protocols to test mechanisms *in vitro*.

Prince Laboratory (Proteomics, Lipidomics, and Bioinformatics), BYU, Department of Chemistry & Biochemistry **2010-2014**

- **Designed experiments** and projects, built lab software, gained authorship on two publications, received two separate research awards.

JAMES DALGLEISH

MOLECULAR BIOSTATISTICS, MACHINE LEARNING, EPIGENOMICS, & CLINICAL BIOINFORMATICS

RESEARCH PUBLICATIONS AND AWARDS

- Mspire-Simulator: LC-MS Shotgun Proteomic Simulator for Creating Realistic Gold Standard Data 2013
Andrew B. Noyce, Rob Smith, James Dalgleish, Ryan M. Taylor, K. C. Erb, Nozomu Okuda, and John T. Prince.
Journal of Proteome Research 2013 12 (12), 5742-5749
- *Automated Structural Classification of Lipids by Machine Learning*
2014 Ryan Taylor; Ryan H. Miller; Ryan D. Miller; Michael Porter; James Dalgleish; John T. Prince
Bioinformatics 2014; doi: 10.1093/bioinformatics/btu723
- Research Awards: Undergraduate Research Award, given for solid research proposals (\$1500 each, \$3000 total in funding)
 - Fall 2010 (Universal Biochemical Extraction Method for Proteomic, Lipidomic, and Metabolomic LC/MS),
 - Spring 2014 (Peptide Behavior Factor Analysis by Machine Learning)

LABORATORY AND ANALYTICAL SKILLS

- Statistical software: SAS, R (3 years, with 1.5 years paid full-time experience), some exposure to MATLAB, JMP, and SPSS
- Programming Languages: exposure to Python/Snakemake, Ruby (4 years), C/C++, Java, others; able to quickly learn more.
- Machine learning packages: Weka, Rapid Miner
- Statistics coursework: Analysis of Variance (majors course for experimental design and analysis), SAS Programming (both basic and advanced SAS/SQL)
- Lab techniques: Sample prep for mass spectrometry, human cell culture, bacterial culture, lipid/protein extractions, MS sample prep (FASP), chemical synthesis, northern blots, DNA electrophoresis, PCR, other common nucleic acid techniques.
- Bioinformatics: familiarity with HPC (PBS, SLURM) systems (supercomputing job submission), alignment of RNA/DNA.
- Broad data experience: GWAS, RNAseq, Microarray, ChIP-seq, genomic, proteomic, and lipidomic data.
- Lab coursework: Molecular biology laboratory, general chemistry lab, organic chemistry lab, advanced molecular biology lab, RNAseq FAES course
- Extensive in pathway & gene regulation (IPA, WGCNA, igraph (including metrics), Pathway Studio, Metacore, Cytoscape)
- Big data familiarity:
 - Wrote & executed many parallelized HPC jobs on the NIH Helix/Biowulf System and the Mary Lou Fulton Supercomputing Cluster.
 - Some of these have been job arrays (rather than discrete jobs), involved large disk space (Human Proteome project), RAM, or cores (up to 4096 cores and 1T RAM).
- GWAS software, including PLINK
- Linux environment experience, using SSH and X11.

CLINICAL EXPERIENCE

- Shadowing with Dr. William Strong, Anesthesiologist and Dr. Tyler Crawford, Radiologist 2012-2014
- Vitals and Spanish Translation for patients in ER intake, Orem Community Hospital 2008-2010
- Genetics Observation, three workdays, Nicola Longo, MD, PhD, John Carey, MD, David Viskochil, MD, PhD 2016

SERVICE AND LEADERSHIP

- Club treasurer, Volunteer and Events Coordinator, Vice President, BYU. 2014-2015
Led many efforts to increase student wellness
Organized benefit concert "Chords of Hope"
- Student Advisory Council Representative, BYU 2014-2015
Led a group that worked with many administrators and was one of few chosen to present several prominent student issues and remarkable progress to University President and Vice President
- Teaching Assistant, Freshman Biology course, BYU. 2013
- Volunteer Translator, Orlando Food Bank 2007
- Missionary and Orlando Homeless Connect Volunteer, Orlando, FL. 2006