Matthew T. Ralston

E-mail: [professional.bio.coder@gmail.com](mailto:professional.bio.coder@gmail.com)

Phone: 610-470-8028

Website: [https://matthewralston.github.io](https://matthewralston.github.io/)

Github and Social Media: [https://matthewralston.github.io](https://matthewralston.github.io/)

SUMMARY

Technical expert, engineer, scientist, DevOps, cloud computing, 5 years of industry experience, expert Python programmer, Linux/OSX, numpy/scipy, pandas, matplotlib, R, bioinformatics, custom database technologies, full-stack web application development, microservices, kubernetes/k8s, custom pipeline architectures |

EDUCATION

University of Delaware 2021

Masters (M.S.) in Data Science (ongoing)

University of Delaware 2012-2015

Masters (M.S.) in Bioinformatics and Computational Biology

* Thesis: *Assembling Improved Gene Annotations in Clostridium acetobutylicum with RNA Sequencing.*
* 3.97 GPA

University of Delaware 2008-2012

Bachelors (B.S.) in Biochemistry

* Minor in Biochemical Engineering
* Merit Scholarships
* Undergraduate Research Fellow

University of Delaware 2007

High-school Summer College Program

WORK EXPERIENCE

Terra Informatics LLC June-December 2022

Principal Consultant

* Django, HTML5, Javascript, CSS, Stripe Checkout, PostgreSQL, Kubernetes(k8s), Docker, bioinformatics tools (bowtie2, fastqc, limma, R, ANOVA) and much more.
* Building IP backbone for a consultancy offering. Pipeline-as-a-Service (microSaaS)
* 48CPU server conformant kubenetes, 38-component pipeline for multi-cloud and cloud-agnostic deployment for < $5k per project.

Bayer Crop Sciences June-December 2020

Genomics Pipeline Developer – Microbial Genomics and Data Science (GDS)

* Used AWS CLI, developer console, and custom in-house tools (AWFL) to manage resources related to pipeline scale-up, fault tolerance, proper interfaces and co-development with the RDBMS team.
* Upgraded a pipeline from 5 modules to 24 modules in 6 months, adding substantial new bioinformatics capabilities, including biosynthetic gene discovery, phylogenetic annotations, genome completeness metrics, and other functional genomics annotations.
* Developed custom parsers, json-schema specifications for data structures, and initialization scripts in Python.

Bristol Myers Squibb 2015-2019

Research Scientist II – Genomics and Computer Assisted Drug Design (CADD)

* Galaxy bioinformatics server administrator, sequencer automation, Amazon Web Services (AWS) integration, tool builder, NGS RNAseq QC pipeline and large sample (20+) summary reports, and user support (Comp Bio)
* Developed expertise in expression technologies (Affymetrix, Illumina RNAseq) from dataset QC, processing, to statistical modeling.
* Built a large deduplicated dataset version control system, provided git expertise (Comp. Bio)
* Built a mirrored chemical database, structural and UID deduplication. (CADD)
* Expanded a multi-platform theoretical chemistry framework (Schrodinger, Openeye, ChemAxon) to enumerate all protomers, tautomers, and rotamers/conformers (2D->3D enumeration) (CADD)
* Parameter optimization and scalability investigation of the framework across multiple chemotypes. Built SMARTS filters for eliminating chemically implausible substructures from the enumerations.
* Increased skills with AWS, Flask/Ruby MVC, Linux/Unix, git, bash, NodeJS, Python, R, PostgreSQL and Oracle, software engineering

Papoutsakis Laboratory 2012-2015

Graduate Research Assistant

* Utilized batch fermentation in bioreactors to produce bacterial cultures for sampling and nucleic acid extraction
* Developed a laboratory protocol to produce strand-specific Illumina RNA-seq libraries from high quality RNA.
* Ensured RNA quality by designing and implementing quality control analytical checkpoints.
* Constructed a bioinformatics pipeline to process, align, and assess 1.5 billion paired-end Illumina reads (RNAseq)
* Produced and corrected a transcriptome assembly from a range of culture conditions.
* Supported a team of scientists and engineers with computational analyses.

Center for Translational Cancer Research 2010-2012

Undergraduate Research Fellow

* Characterized the apoptotic response of colon cancer cell lines to statin treatment in human cell culture with Annexin V assay.
* Investigated differential BMP-pathway signaling in fresh and FFPE colon tissue with immunohistochemistry, fluorescence microscopy, Western blot.
* Assisted researchers with fluorescence assisted cell sorting (FACS) and flow cytometry.
* Presented BMP-pathway IHC micrographs and culture kinetics in response to statins in fellowship poster session.

PROJECTS

Consultancy Django application 2022

Principal Consultant – Terra Informatics LLC

Vanilla Django + HTML5 templating, 4 apps with PostgreSQL, marketing pages, login functionality, Stripe checkout, HTML, ticket access to microSaaS pipeline offerings (RNASeq, metagenomics). Dashboards and run logging. 38 component, conformant kubernetes SaaS offering with Stripe checkout.

K-mer Database (.kdb) 2019-2023

Software Engineer and Researcher

<https://github.com/MatthewRalston/kmerdb>

<https://matthewralston.github.io/kmerdb>

Developed a custom database format under .bgzf compression to store the 8 nearest neighbors of each k-mer in a pseudo graph format, with an index function. Characterized the sensitivity of the analyses to errors from simulated data, examining the relationship between k-mer profile distances and phylogenetic distance.

Not Very Humerus 2018-2020

Blog Author

<https://matthewralston.github.io/blog>

A simple blog on beginner programming topics for the most part.

Kmer.js 2018-2019

Software Developer

<https://github.com/MatthewRalston/kmer.js>

A very very simple in-memory k-mer array.

Liganator v2.0 2019

Software Developer – Bristol Myers Squibb

Developed and planned benchmarking for a utility to explore different vendors, parameters, and more regarding the process of preparing a 2D chemical structure for all its physicochemically identical isoforms, as well as rotational perturbations, in preparation for ligand docking studies. A product designed for protein biochemists.

Project Utilities for Reproducible Research 2016

Software Developer – Bristol Myers Squibb

A NodeJS REST-API and accompanying CLI client for a reproducibility database. It was similar in principle to git-lfs. It’s pretty easy to imagine what we did if you have a technical background, we used certain checksums to identify a chunk of data, and then stored metadata associated with it, to make a hard drive on a selected server a reproducibility hub.

SKILLS

RNA-seq library preparation, qRT-PCR, RT-PCR, PAGE: Western blot, Northern blot, HPLC

Nucleic acid extraction, enzymatic treatment, FACS, Light microscopy

Fluorescence microscopy, immunohistochemistry, immunofluorescence

Spectrometry: UV-Vis, Fluorometry, Nanodrop

GC-MS

Cell culture: Human cell culture, anaerobic bacterial cell culture, fermentation

Sequence alignment: Hisat2, Bowtie, Tophat, BWA, Blast, Blat

Assembly: Cufflinks, Trinity, Velvet

Sequence Processing: Fastqc, Samtools, Picard, Bedtools

Miscellaneous: Cuffdiff, DEseq, Circos, David, Amazon Web Services (AWS)

Programming: Python, R, Bash, Javascript/NodeJS, Ruby, Emacs, D3, Julia, Perl, awk, grep/sed, LaTeX, Matlab, HTML/CSS, MySQL, Oracle, PostgreSQL

Data Management: Oracle SQL, PostgreSQL, rdflib, snapshotting

Cloud Computing: S3, Elastic Compute (EC2), CodeBuild, Elastic Container Repository (ECR), Docker, IAM

Statistics/Data Science: Hypothesis testing, ANOVA/regression, distribution fitting, R/Rstudio/Rshiny, Bioconductor, discrete/continuous probability, multivariate statistical analysis, PCA, clustering, classification

Systems Administration: systemd, DE/WM, tmux, Nginx, Apache, conky, web servers, SGE/UGE, bash, parallel, time

LEADERSHIP

Terra Informatics LLC 2022

CEO and Principal Consultant

* Building an IP backbone structure for RNA-Seq, DNA-Seq, metagenomics, and more. Using conformant kubernetes(k8s) to deploy the pipeline on multi-cloud. A MicroSaaS offering primarily for academic researchers in the Ammon Pinizzotto Engineering and Bioinformatics departments. Targeting collaborations in the incubation community amongst the notorious Chemical Engineering dept of the Univ. of Delaware. Sequencing is needed more than ever.

Bioinformatics Student Association 2013-2014

President

* Organized and presented in a UNIX and high performance computing (HPC) workshop.
* Secured funding for this and other student development workshops.

IGERT Business, Ethics, and Communication in Life Sciences 2013

Project Manager

* Won a proposal for a revision to Fraunhoffer CMB’s GMP vaccine production process.

National Honor Society, A.I. chapter 2008

Vice President

* Coordinated volunteer activities

A.I. Marching Band 2008

Vice President

* Coordinated field show, trip to Rose Bowl

AWARDS

IGERT Business, Ethics, and Communication in Life Sciences 2013

IGERT SBE2 Faculty

* Team of students won tuition for a semester by providing a proposal for a revision to Fraunhoffer CMB’s GMP vaccine production process.

Undergraduate Research Fellowship 2010

Undergraduate Research Program

* Proposed a fractional-factorial investigation of statin treatment on colon cancer cell lines.
* Presented findings in a Research Symposium

2nd Place – Sound of Music event 2008

Delaware Science Olympiad

* Engineered a guitar and PVC chime set
* Wrote original melodies, harmonies

VOLUNTEER

Carisma Peru 2014

Clinical assistant

* Assisted physicians in Peru with clinical laboratory sampling, spirometry, blood pressure measurements in ObGyn and E.R. settings.

PUBLICATIONS

* Ralston, Matthew T., and Eleftherios T. Papoutsakis. "RNAseq‐based transcriptome assembly of Clostridium acetobutylicum for functional genome annotation and discovery." AIChE Journal 64.12 (2018): 4271-4280.
* "Complex and extensive post-transcriptional regulation revealed by integrative proteomic and transcriptomic analysis of metabolite stress response in Clostridium acetobutylicum," contributing author. Biotechnology for biofuels 8.1 (2015): 81.
* “Overexpression of the L. plantarum peptidoglycan biosynthesis murA2 gene increases the tolerance of E. coli to alcohols and enhances ethanol production,” contributing author. Applied Microbiology and Biotechnology. 98.19 (2014): 8399-8411.
* “The Clostridium small RNome that responds to stress: the paradigm and importance of toxic metabolite stress in C. acetobutylicum”, contributing author. BMC Genomics. 14.1 (2013): 849.