

ABOUT

Fluent in the use of bioinformatics and statistical algorithms and software to transform and analyze genomics data in a rapidly evolving scientific discipline, providing an integrated analysis of multi-modal RNA, DNA, and protein information.

EDUCATION

Northeastern University

San Francisco, CA

M.S. Candidate in Bioinformatics

2020- Current

Thesis: "Clustering and dimensionality reduction to uncover trends in literature about heat shock proteins"

University of California, Santa Cruz

Santa Cruz, CA

Bachelor of Arts in Environmental Biology

2015- 2019

Thesis: "Genetic diversity of marine cyanobacteria and their sensitivity to added nutrients"

RESEARCH EXPERIENCE

- **Natural language processing and clustering - 2021**

Consolidated a literature corpus of 5,101 pdfs from PubMed and ScienceDirect, removed duplicate files, used Dockerized software called GeneRation Of Bibliographic Data to convert the documents into raw text, lxml and JSON. Parsed the plain text citations into a metadata table as well as to tokenize words. Used tf-idf on tokens to find relevant terms and then clustered terms with k-means and tSNE. Compare clustering methods and explored literature.

- **Assistant Scientist, Merck Research Laboratories, South San Francisco, CA - February 2021 - 2022**

Tracked DNA through the lab with excel spread sheets and email. Managed DNA inventory for transfections on a weekly basis. Trained colleagues to access company database and CESDS. Created Sharepoint List to manage DNA and protein inventory. Performed bacterial transformation, inoculation, transfection with mammalian cell lines, cell culture, sterile technique, protein characterization: UHPLC, CESDS.

- **Bacterial Genome Assembly, California State University, Monterey Bay, CA - 2020**

Assembled and polished a 6,868,303 bp bacterial genome with five different assemblers, benchmarked assembly methods with BUSCO and compared quality using quast reports, annotated the genome with hmmsearch and created phylogenetic trees for the phosphotriesterase protein family from all of the genomes. Worked with teammates to assemble five other genomes.

- **Hopkins Marine Station, Department of Biology at Stanford University, CA - 2019 - 2020**

Utilized flow cytometry to investigate coral bleaching on a cellular scale. Studied phenotypic variation, heat shock proteins and genome wide associations related to the disease state of coral.

COMPUTATIONAL SKILLS

- **Algorithms - 2022**

Calculated the weight of a protein by multiplying the length by the approximate weight of each amino acid, calculated the weight of a protein sequence when given a nucleotide sequence, gathered kmers of user defined length in a sliding window, calculated buffer volumes from stock concentration in a protocol with changing user input, calculated minor allele frequency, Leetcode problem 235. Lowest Common Ancestor of a Binary Search Tree.

- **Software Development - 2021/2022**

Gathered requirements for a fake app that's a combination of Tik Tok and Tinder, identified JavaScript classes, tasks and an underlying model. Made a JavaScript class for saving metadata about publications, implemented abstraction, encapsulation, polymorphism and inheritance to create classes for books, papers, webpages and calculate the percent composition of each letter in the title of each subclass of publication. Created python module with try and various exceptions, version control with GitHub, cleaned code with flake8 and unit tests with pytest.

- **Statistics - 2020/2022**

Performed linear regression on gene expression values from two probes. Summarized two genes' bivariate normal distribution with R. Identified chi-squared point estimator and one sided confidence interval given degree of freedom. Clustered microarray data for acute lymphatic leukaemia and acute myeloid leukaemia with Hierarchical, PCA and k-means. Identified confidence intervals (bootstrap, parametric) for a gene's expression in two groups.

PROFESSIONAL DEVELOPMENT

Articles on Medium <https://medium.com/@gmreed>

Grant writing with Coral Vita: assessed differential gene expression in corals with a SNP array - **October 2020**

Coauthor *In review*: Genome Assembly of bacteria with remediation potential for the organophosphate pesticide diazinon (**2020**) Heyer, N., Reed, G., Lee, L., Garcia, V., Boris, W., Ryder, R., Jue, N.

Volunteer: managed supplies and taught strawberry DNA extractions at a community festival in Santa Cruz, CA - **July 2018**

Awarded funding: for senior thesis on microbial ecology in the Zehr Lab - **June 2018- April 2019**