YIHENG (SAM) SU

Sam.Su@utexas.edu
201 E 21st Street #W0523 • Austin, TX 78705 • (214) 609-8100
GitLab Account: https://gitlab.com/Sam.Su1
GitHub Account: https://github.com/SamSoup

EDUCATION

The University of Texas at Austin Bachelor of Science in Computer Science May 2021
- Applied Statistical Modeling Certificate (18 hours of coursework) GPA 4.0

Honors: University Honors (Fall 2018, Spring 2019, Fall 2019, Spring 2020, Fall 2020)

RELEVANT COURSEWORK

Data Structures, Discrete Math, Multivariable Calculus, Computer Organization and Architecture, Intro to Probability and Statistics, Big Data in Biology (FRI stream), Principles of Computer System, Matrices/Matrix Calculations, HDF Research Practicum, Cyber-physical Systems, Multicore OS Implementations, Objected Oriented Programming (OOP), Competitive Programming, Biostatistics, Artificial Intelligence, Algorithms and Complexity, Statistical Learning/Inference, Computational Biology & Informatics, Computer Programming: C++ and Python.

ACADEMIC PROJECTS

Analyzing Differentially Expressed Genes of COPD Fall 2018 – Spring 2019
- Identified differentially expressed genes and enriched pathways of Chronic Obstructive Pulmonary Diseases using RNA-Seq data from three different brain regions
- Specialized in data visualization and interpretation using Matplotlib, Seaborn, and KNN in Python

Life History Transitions of Salmonids Spring 2019 – Fall 2019
- Investigated whether genes highly responsive to life history transitions are also plastic across other contexts (ie. age, sex, size, etc.) and thus exhibit fewer regulatory interactions and/or higher rates of molecular evolution.
- Authored BLAST scripts to map microarray data to salmonid genome using LoneStar5 under TACC
- Conducted and interpreted results from Principal Component Analysis and Differentially Expression Analysis using R packages such as tidyverse, LIMMA, pheatmap, and UpSetR.

Expression Variance and Evolution (EVE) Model and Wrasse cleaners Spring 2020 – Current
- Mentored project under Dr. Rebecca Young from the Hofmann Lab at UT Austin (youngrl@utexas.edu)
- Examined how we can integrate EVE model (phylogenetic ANOVA) with other traditional comparative analysis approaches to explore the effects of relatedness on associations between expression and phenotypes using TAG-Seq data on six Wrasse species.
- Administered Weighted Gene Co-Expression Network Analysis (WGCNA) as well as differential expression analysis using LIMMA underlying an overall data analysis pipeline including EVE.
- Generated data visualizations such as volcano plots, heatmaps, dendrograms using ggplot2 in R.

Predicting Sexually Antagonistic Selection in DNA Sequences Spring 2021 – Current
- Mentored project under Dr. Mark Kirkpatrick (kirkp@austin.utexas.edu) and Dr. Andrius Jonas Dagilis (adagilis@gmail.com) from the Kirkpatrick Lab at UT Austin.
- Applied machine learning techniques such as Random Forest to capture signals of Sexually Antagonist Selection from X, Y chromosomes of Japan Sea Sticklebacks.
- Performed dimensionality reduction, feature engineering, and cross validation to optimize said models using Python packages such as Scikit -Learn, Pandas, Numpy, and Matplotlib.

RESEARCH EXPERIENCE

Office Research Assistant for Project SEED Summer 2019 – Spring 2020

- Schedules and coordinates home visits with Spanish-speaking families using REDCAP (UT database)

Research Assistant for Hofmann Lab Spring 2020 – Current

- Utilizing the EVE model (C based package) that seeks to explore inter and intra species variation in a phylogenetic approach using TAG-Seq data from the Hofmann Lab

Research Assistant for Huk Lab Spring 2021 – Current

- Working with DeepLabCut (DLC), an open-source package for animal pose estimation tracking locations of key points via deep neural networks trained by hand-annotated images.

Research Assistant for Kirkpatrick Lab Spring 2021 – Current

- Applying Random Forest and other machine learning techniques to predict if genetic sequences demonstrate signals of Sexually Antagonistic Selection (SAS) and determine significant statistical metrics that sheds light on the underlying genetic basis of SAS

LEADERSHIP EXPERIENCE AND ACTIVITIES

Unmanned Aerial Vehicle Austin (UAVA) – *Active Member* Fall 2018 – Spring 2019

- Engaged in weekly coding sessions that explores how to use MakeFiles, Node JavaScript, and Docker
- Visualized data from InfluxDB using Grafana API interface

HONORS

- UT Retired Faculty/Staff Scholarship Fall 2020
- Tracor/Frank McBee, Jr. Scholarship Fall 2020
- UT Distinguished College Scholar Honoree Spring 2020
- UT College of Natural Sciences Second Year Excellence Award Spring 2020
- Angus G. and Erna H. Pearson Undergraduate Scholarship in CS Fall 2019
- Coolest Strategy Award at SpaceCraft SXSW Spring 2019
- Best Use of HERE API at HackTX 2018 Fall 2018

Skills

Computer Skills: Proficient in R, Python, Java, C++, and C; Learning x86 Assembly.
Languages: Fluent in Mandarin, Basic Spanish