KRISHNA REDDY GUJJULA

Contact Information

Email : gujjulakrishnareddy@tamu.edu
Webpage : https://kianfar.engr.tamu.edu/krishna-reddy-gujjula/
Linkedin : www.linkedin.com/in/gujjulakrishnareddy
Phone : (979)739-1664
Address : Department of Industrial and Systems Engineering,
Texas A&M University, College Station, TX 77843-3131

Education

- **Ph.D. in Systems Engineering (Bioinformatics)** (GPA: 3.94/4.00) May 2018
  Texas A&M University, College Station, TX, USA
  Dissertation: "Map2Peak: From Unmapped Reads to ChIP-Seq Peaks in Half the Time".
  Advisor: Dr. Kiavash Kianfar

- **Bachelor of Technology in Mechanical Engineering** (GPA: 3.93/4.00) June 2010
  National Institute of Technology, Kurukshetra, India

Career Goal

Design and implementation of bioinformatics algorithms and statistical analysis of biomedical data aimed at discovery of (personalized) therapies for complex diseases such as cancer.

Highlights of Bioinformatics Expertise

4+ years of research experience in the field of bioinformatics, especially working with Next Generation Sequencing (NGS) data; Knowledge and skills to analyze and build statistical models, mathematical models, design algorithms for different bioinformatics problems using state of the art computer programs such as C++, R and Python; Hands-on internship experience of working in a clinical sequencing laboratory and collaboration experience with molecular biologists.

- **NGS analysis pipelines:** ChIP-Seq, RNA-Seq, Identifying Genomic variants.
- **NGS database:** UCSC, ENCODE project, ArrayExpress, GEO, SRA, ENSEMBL.
- **NGS tools:** Working knowledge of well known bioinformatics tools e.g. read aligners (bowtie, BWA), peak callers (MACS2, PeakSeq), variant callers (GATK) and ability to develop bioinformatics analysis pipelines for RNA-Seq, ChIP-Seq and variant calling studies.
- **NGS utilities and format:** Experience with NGS utility tools such as SAMTools, BEDTools, BamTools etc and NGS formats such as SAM, BAM, VCF on Linux/Unix platform.

Other Computer Skills

- Statistical and optimization tools: R, Base SAS, CPLEX Concert technology, AMPL
- Scripting: Shell scripting, MATLAB, Python, awk, LaTeX
- Programming Language : C++

Relevant Coursework

- **Bioinformatics & Statistics :** Large Scale Biological Dataset Analysis, Statistical Bioinformatics, The Methods of Statistics I, Theory of Statistics-Distribution Theory, Theory of Statistics-Inference, Statistical Computations and analysis, Probability for Engineering Decisions, Design and Analysis of Experiments, Machine Learning with Networks.
- **Algorithm & Optimization :** Analysis of Algorithms, Linear and Non-Linear Optimization, Heuristic Optimization, Integer Optimization.

Publications

- **Gujjula, K** and Kianfar, K., *"Map2Peak: From Unmapped Reads to ChIP-Seq Peaks in Half the Time,"* working paper, To be submitted to **Bioinformatics**, in April 2018. Draft is available upon request. Map2Peak is available at https://kianfar.engr.tamu.edu/krishna-reddy-gujjula/.
- **Gujjula, K**, Seshadrinathan, K and Meisami, A., *"A Hybrid Metaheuristic for the Maximum k-Plex problem",* **Examining Robustness and Vulnerability of Networked Systems**, IOS Press, 2014, pp 83-92.

Software

Map2Peak: Map2Peak is an ultrafast ChiP-Seq peak calling bioinformatics tool. It combines read alignment and peak calling, and considers multi-mappable reads in its algorithm. Input to Map2Peak is a FASTQ file and output is a peaks BED file. Map2Peak is available at https://kianfar.engr.tamu.edu/krishna-reddy-gujjula/.

Professional Positions

- **Research Assitant** Fall 2013-Present
  Texas A&M University, College Station, TX
- **Course Instructor (ISEN302)** Fall 2017
  Texas A&M University, College Station, TX
- **Senior Intern** Summer 2017
  R&D Clinical Sequencing Lab, ThermoFisher Scientific, Austin, TX

- **Course Instructor (ISEN302)** Spring 2017
  Texas A&M University, College Station, TX
- **Guest Lecturer (ISEN668)** Spring 2017, Fall 2017
  Texas A&M University, College Station, TX
- **Lab Instructor/Teaching Assistant (ISEN314)** Fall 2016
  Texas A&M University, College Station, TX
- **Course Instructor (ISEN315)** Summer 2015
  Texas A&M University, College Station, TX
- **Teaching Assistant (ISEN302)** August 2013 – May 2016
  Texas A&M University, College Station, TX

Key Responsibilities and Achievements

- **Internship-Bioinformatics**
  - Worked closely with molecular biologists on liquid biopsy (cfDNA) project to develop panels for early identification of genomic variants.
  - Worked on development of Tagseq and Ampliseq panels to identify low frequency mutations.
  - Performed statistical data analysis and developed a real time dashboard to help visualize NGS sequencing lab metrics and evaluate lab performance.
  - Dashboard was developed in R-shiny and was deployed on ThermoFisher server.
- **Research**
  - Developed ultrafast ChIP-Seq algorithm to identify transcription factor binding sites with capabilities to identify binding sites in repeat elements.
  - Developed a bioinformatics tool in C++ for ChIP-Seq peak calling.
  - Performed statistical analysis on microarray gene expression data in collaboration with molecular biologists.
  - Developed a novel heuristic for maximum k-plex problem (NP-hard). The algorithm was applied on 15 well known graphs and near-optimal solutions were achieved.
- **Teaching**
  - Taught Economic analysis of engineering projects (2 times), Total Quality Engineering (1 time), Production Systems Planning (1 time).
  - Undertook full responsibility of teaching & evaluating students.
  - Developed course handouts & exams, held help sessions.
  - Received highly positive course evaluation metrics.

Professional Projects

- **Intern Project**
  Performed statistical analysis on NGS datasets and developed a real time scalable dashboard to visualize NGS sequencing lab metrics and evaluate lab performance. Dashboard was developed in R-Shiny and was deployed on ThermoFisher server.

- **Ph.D. Dissertation**
  Developed fast and accurate ChIP-Seq peak calling algorithm to detect **transcription factor binding sites**. Algorithm has the ability to resolve location of multi-mappable reads and identify binding sites in **repeat elements** of DNA.

- **Gene expression analysis**
  Developed statistical model for time based differential gene expression analysis to find miR-150 role in negative regulation of B-cell growth.

Honors, Awards & Certifications

- Graduate Teaching Fellowship — Spring & Fall 2017
  Texas A&M University, College Station, TX

- Academy of Future Faculty Certification — Spring 2017
  Texas A&M University, College Station, TX