Harihara Subrahmaniam Muralidharan

📞 +1(240)481-0542 ✉ hsmurali@umd.edu in hsmurali-umd ○ hsmurali ⬤ 0000-0002-2482-2266
🐦 Harihara_subrah 🌐 https://cs.umd.edu/~hsmurali

EDUCATION

University of Maryland, College Park   Expected May 2024
Doctor of Philosophy in Computer Science   *College Park, MD, USA*
- **Specialization:** Bioinformatics and Computational Biology.
- **Thesis:** Data-driven algorithms to characterize structural variation in metagenomic data.

University of Maryland, College Park   Aug 2018 – May 2023
Master of Science in Computer Science   *College Park, MD, USA*
- **Course Work:** Machine Learning, Advanced Numerical Optimization, Computational Geometry, Computational Linguistics, Computational Genomics, Unsupervised Learning, Biological Networks Across Scales, Data Structures and Algorithms for High-Throughput Genomics, Computational Phylogenetics, Randomized Algorithms

SASTRA University   Jul 2011 – Jul 2015
Bachelor of Technology in Computer Science and Engineering   *Thanjavur, India*
- **Selected Course Work:** Data structures, Algorithms, Theory of computation, Operating systems, Artificial Intelligence, Linux Programming, Compiler Design, Linear Algebra, Calculus, Cloud Computing, Databases, Parallel and Distributed Systems
- <u>Semester abroad</u> at Universitat Politècnica de Catalunya (UPC), Barcelona and worked on my undergraduate thesis titled "Enhancing Learning Analytics Platform for Secondary Schools: Design and Development of Indicators".

TECHNICAL SKILLS

Languages - Python, C++, Java, R, shell, C
Machine Learning - Scikit sklearn, Keras, Scipy, Numpy, PyTorch, TensorFlow
Bioinformatics Tools - SAMtools, BEDtool for sequence alignment arithmetic. BowTie 2, BWAMEM, BLAST, Minimap2 for sequence alignments. Has worked extensively with NGS data preprocessing (quality control, read trimming and adapter sequence removal), genome assembly (long and paired end short reads), genome polishing, metagenomic binning and OTU analysis with amplicon sequencing data. Gene prediction and annotation using function and taxonomy databases. Computational phylogenomic analyses.
Bioinformatic Databases - UniProt, NCBI RefSeq, NCBI NR/NT, GTDB, RDP, SILVA, GO, KEGG
Platforms - Linux, MATLAB, EnergyPlus, SLURM for HPC Environment, Google cloud platform, git/github, SQL, Jupyter notebooks, RMarkdown
Professional Skills - Strong oral and written communication, Cross-functional collaboration, Project leadership, Criticial thinking

EXPERIENCE

University of Maryland   Jan 2018 – Present
Graduate Research Assistant   *College Park, MD, USA*
Research Areas: Computational Genomics, Bioinformatics, Genomic Data Science
- Developed tools and data analysis pipelines using multiomic NGS data such as illumina short read sequencing, long read sequencing(ONT and pacbio), Hi-C, amplicon sequencing and, single-cell amplified metagenomes to,
  - characterize structural variants from metagenomic assembly graphs, discover novel phages through PIRATE and improve metagenome bins using an ensemble of coverage and assembly graph features with Binnacle
  - cluster large 16S rRNA gene sequence datasets using SCRAPT
  - study the impact of transitive annotations on the sensitivity of taxonomic classifiers
  - characterize variants in the *Synechococcus spp* from the Yellow stone national park microbial mats.
- Mentored 3 undergraduate students, 4 summer interns, and, 2 high school students
- Teaching assistant for data science in Python (fall 2018), data science in R (spring 2019) and algorithms (fall 2019).

Tata Consultancy Services, Innovation Labs   Aug 2015 – Aug 2018
Researcher   *Chennai, India*
Research Areas: Machine Learning, Cyberphysical Systems, Time Series Analysis, Optimization

- Identified suitable models using machine learning and domain physics to analyze real time sensory signals from large scale engineering systems such as VCRS(Vapor Compression Refrigeration Systems), buildings, and, wind turbines.
- Developed software modules to benchmark buildings along 2 dimensions (HVAC & Non-HVAC) using energy signals.
- Developed a scalable decision-making framework for demand response with a district cooling plant. This work resulted in a paper and a patent.

PUBLICATIONS

*-Denotes equal Contribution, #- Authors listed alphabetically

1. <u>Harihara Subrahmaniam Muralidharan</u>, Noam. Y Fox, Mihai Pop, The impact of transitive annotation on the training of taxonomic classifiers, Frontiers in Microbiology, 2024.
2. <u>Harihara Subrahmaniam Muralidharan</u>*, Tu Luan*, Marwan Alshehri, Ipsa Mittra, Mihai Pop, SCRAPT: an iterative algorithm for clustering large 16S rRNA gene data sets, Nucleic Acids Research, 2023, gkad158.
3. <u>Harihara Subrahmaniam Muralidharan</u>*, Nidhi Shah*, Jacquelyn S Meisel, Mihai Pop. Binnacle: Using Scaffolds to Improve the Contiguity and Quality of Metagenomic Bins. Frontiers in Microbiology, 2021.
4. Gabriel Birzu, <u>Harihara Subrahmaniam Muralidharan</u>, Danielle Goudeau, Rex Malmstrom, Daniel S Fisher, Devaki Bhaya, Hybridization breaks species barriers in long-term coevolution of a cyanobacterial population. eLife, 2023.
5. Seth Commichaux, Tu Luan, <u>Harihara Subrahmaniam Muralidharan</u>#, Mihai Pop. Database size positively correlates with the loss of species-level taxonomic resolution for the 16S rRNA and other prokaryotic marker genes, bioRxiv, 2023.
6. Srinarayana Nagarathinam, <u>Harihara Subrahmaniam Muralidharan</u>, Arunchandar Vasan, Venkatesh Sarangan, Sermisha Narayana, Anand Sivasubramaniam. One for all, All for one: a scalable decision-making framework for demand response with a district cooling plant. ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation(Buildsys'19), 2019.

Manuscripts in preparation

1. Assembly graph-based variant discovery reveals novel dynamics in the human microbiome.
2. Comparative Metagenomic Genome Analysis of *Synechococcus spp.* in Microbial Mats Across a Temperature Gradient in Hot Springs From Yellowstone National Park.

TALKS & POSTER PRESENTATIONS

$-Poster presentation, #- Talk

1. One for all, All for one: a scalable decision-making framework for demand response with a district cooling plant. [Buildsys '19]#
2. Binnacle: Using Scaffolds to Improve the Contiguity and Quality of Metagenomic Bins. [ISMB'20, CSHL Microbiome Workshops'20, ASM Microbe'22]$
3. PIRATE: Phage Identification fRom Assembly-graph varianT Elements. [ISMB'20, CSHL Microbiome Workshops'20] #$
4. Comparative Metagenomic Genome Analysis of *Synechococcus spp.* in Microbial Mats Across a Temperature Gradient in Hot Springs From Yellowstone National Park. [CSHL Genome Informatics'21, ASM Microbe'22]$
5. SCRAPT: An Iterative Algorithm to Cluster Large 16S Gene Datasets. [ISMB'22]#$

VOLUNTARY SERVICE

External reviewer for IEEE CDC (2018), WABI (2020), ISMB (2021), ISMB(2022), WABI(2022), RECOMB Seq (2023), ESA (2023), Subreviewed an article for Nature methods.