

Dr. Olabiyi Obayomi

1400 W Swan Street, Apt 212, Stephenville, Texas, USA.

Email: obadbotanist@yahoo.com | **LinkedIn:** <https://www.linkedin.com/in/olabiyi-obayomi-94740b34/> | **GitHub:** <https://github.com/olabiyi/> | **Website:** <https://olabiyi.github.io/cv/> | **Phone:** +1 254-459-0603

Summary: I am a passionate bioinformatician skilled in omics data analysis particularly metagenomics, transcriptomics, microbial genomics, programming languages, and microbiology with the desire to develop and apply computational techniques to advance and accelerate biological and environmental research. I have over 5 years of experience working in bioinformatics and on microbial genomics projects. I have developed complex computational pipelines (<https://github.com/olabiyi/>) and provided insights on various next-generation sequencing datasets including but not limited to metagenomics, transcriptomics, and microbial genome assembly.

Education

Ben Gurion University of the Negev

Ph.D. Environmental Microbiology

Israel

2019

Cumulative GPA: 96.5

Core Modules: Microbial Pathogens, microbial communities, next-generation sequencing, microbial ecology

Skills and Learning:

- Characterized and tracked the source of microbiomes (bacteria and protists) including pathogens in soil, water and crop samples using both culture- dependent and independent approaches including amplicon (16S and 18S) sequencing
- Analyzed and processed data with various bioinformatics tools, including QIIME, SourceTracker, R Bioconductor packages
- Developed a tool for detection, statistical analysis, and visualization of pathogens in environmental samples using amplicon sequences https://github.com/olabiyi/16s_pathogen_analysis
- Wrote scripts in Bash, R and Python that used high performance clusters to decrease execution time on tasks
- Wrote and contributed to four manuscripts for publication

Technical Skills

- **Pipeline development:** Snakemake, Nextflow and Make
- **Biostatistics and machine learning:** multivariate statistics, hypothesis testing, regression, normalization, correlation, constrained analysis, correction for multiple hypothesis testing, ggplot2 and machine learning techniques
- **Bioinformatics analysis:** metagenomics (amplicon and shotgun), transcriptomics (model and non-model organisms), metatranscriptomics, genomics, variant calling, phylogenetics, gene functions and pathways enrichment analysis
- **Bioinformatics tools:** QIIME (1 and 2), PICRUST 2, Bioconductor, MEGA 7, MetaPhlan2, trinity, trinotate, cdhit, bwa, Stacks, Diamond, GATK4, bcftools, mauve, Humann2, megahit, samtools etc.
- **Computer programming:** R, Python, and Bash
- **Network analysis:** R
- **High performance computing:** SGE and SLURM
- **Report generation:** R markdown, R shiny, Jupyter, Jupyter notebook
- **Biological data repositories:** NCBI, EBI, JGI/IMG, KEGG, MetaCyc etc
- **Version control:** Git and GITHUB
- **Containerization:** Docker, Singularity and Conda
- **Database management:** SQL and MYSQL
- **Operating systems:** Linux, Windows
- **Cloud computing:** AWS EC2, S3, Cloud9, Batch with Nextflow

- **Microbiology:** Culture dependent and independent techniques for microbiomes

Personal Skills

- Poised communication, presentation, analytical, organizational, and problem-solving skills
- Excellent interpersonal skills, a team player
- Ability to handle multiple tasks concurrently and in a timely fashion
- Hardworking, result oriented, able to work under pressure with minimum or no supervision
- Proactive, dynamic, curious, and a quick learner

Professional Experience

Bayer AG, Biologics Data Science and Engineering, Chesterfield, Missouri

Oct 2023 - present

Data Analyst

- Actively leading and providing technical contributions on multiple microbial genomics and metagenomics projects by analyzing data provided by third party stakeholders
- Leading the implementation of new and innovative methods for predicting microbial performance
- Designing and analyzing omics studies to enable discovery and improve product performance
- Working collaboratively with partners and stakeholders by scheduling and attending meetings to give insights on the analyses performed for decision making
- Hunted potential microbial contaminants in microbial isolate stocks using 16S rRNA amplicon sequencing and strain specific barcodes
- Refactoring microbial prevalence analysis pipeline codebase, updating its database and retrieving the prevalence data for microbes of interest to be used in product testing, regulatory assessment and decision making
- Performing secondary metabolites analysis using an in-house pipeline to determine the potential presence and transferability of antibiotic resistance genes and toxic metabolites in microbial strains of interest for product development
- Generating and identifying opportunities to extend and improve existing pipelines for microbial genome analysis by introducing new tools and analyses
- Predicting microbial performance to enable discovery and improve product performance.

Texas A&M AgriLife Research, Texas A&M University, Stephenville, Texas

June 2021 – Oct 2023

Postdoctoral Researcher in Bioinformatics

- Analyzed and derived insights on several metagenomics, transcriptomics, and population genetics projects by assisting PI-s across Texas A&M University system for bioinformatics analysis
- Characterized the microbiome of several environmental samples including water, soil, feces and skin samples of salamander, biochar, Texas beaches, cows, and crops
- Performed RNA-seq and pathway analysis for model plants (*Arachis ipaensis*) and humans
- Performed RNA-Seq and pathway analysis for non-model plants (*Arachis dardanii*, little blue stem and Indian grass)
- Compared the transcriptome profiles between plants. Between *Arachis ipaensis* and *Arachis dardanii* and between little bluestem and Indian grass.
- Developed a pipeline for the analysis of amplicon (16S, 18S, ITS) sequencing dataset using QIIME2 and PICRUSt - <https://github.com/olabiyi/snakemake-workflow-qiime2> and usearch <https://github.com/olabiyi/snakemake-usearch>
- Assembled and annotated the genomes of bacteria (*E. coli*) and virus (fowl pox virus)
- Analyzed RAD-Seq dataset of little blue stem grasses across Texas to understand the genetic basis for their diversity
- Assembled and assigned taxonomy to near full-length 16S rRNA sequences of peanut endophytes
- Identified antibiotic resistance genes in soil samples using next-generation sequencing

- Mined NCBI database for sequences to be used for primer design in a Texas sea turtles study
- Participated in project conceptualization and grant writing
- Taught graduate students how to perform statistics using R
- Submitted raw sequences to NCBI
- Wrote manuscripts for publication.

Ben Gurion University of the Negev, Bioinformatics Core facility, Beersheba

Nov 2019 – June 2021

Postdoctoral Researcher in Bioinformatics

- Applied omics tools for the development of pipelines to analyze Illumina and Oxford nanopore sequencing data to gain biological insights
- Developed a pipeline for RNA-Seq analysis of non-model organisms using NeatSeq_Flow to automate database creation, QC, mapping, assembly, annotation, statistical, clustering and enrichment analysis - https://neatseq-flow.modules.readthedocs.io/en/latest/Workflow_docs/RNA_seq_non_model.html
- Characterized the microbiome of Wolffia globosa and identified potential cobalamin synthesizing bacteria using 16S amplicon sequences and shotgun metagenomics
- Characterized the functional and taxonomic diversity of the microbiome of biofilters receiving sand and granular activated carbon treatments using shotgun metagenomics.
- Performed RNA-Seq and pathway analysis for non-model crustaceans (Hippolyte inermis and Cherax quadricarinatus)
- Performed comparative genomics between Z and W chromosomes of Macrobrachium rosenbergii in to identify sex linked genes
- Developed a pipeline for host associated shotgun metagenomics using snakemake - <https://github.com/olabiyi/shotgun-metagenomics-pipeline>
- Developed a pipeline for metatranscriptomics using both read based and assembly approaches - <https://github.com/olabiyi/snakemake-workflow-metatranscriptomics>
- Drafted custom scripts in R, Bash, Python and Perl for omics data analysis
- Presented and delivered my analyses results to customers on our bioinformatics core facility computer servers. This involved database management using SQL, MySQL and PHP, scripting with HTML and CSS and HTML report generation using RMarkdown
- Contributed to manuscripts for publication.

Ben Gurion University of the Negev, Zuckerberg Inst. For Water Research, Sde Boker

May 2019 – October 2019

Post-doctoral Researcher in Bioinformatics

- Characterized the microbiomes of soil, water and crop samples that were sequenced with Illumina and Oxford Nanopore platforms
- Analyzed and processed data with various bioinformatics tools, including QIIME, SourceTracker, R, and others
- Assisted and mentored fellow students (Masters and PhD) with analyzing their bioinformatics-related data
- Developed a bioinformatics pipeline for the analysis of amplicon sequences generated by Oxford Nanopore
- Developed scripts for detecting pathogens using amplicon sequences
- Wrote and contributed to manuscripts for publication.

Publications

<https://scholar.google.com/citations?user=zppxGbKAAAAJ>

Publications acknowledged for my bioinformatics contributions

- Shai A. Shaked, Shai Abehsera, Tome Levey, Vered Chalifa-Capsi, Amir Sagi. (2020). From sporadic single genes to a broader transcriptomic approach: Insights into the formation of the biomineralized exoskeleton in

decapod crustaceans. *Journal of Structural Biology*. 212:107612. <https://doi.org/10.1016/j.jsb.2020.107612>. (*Transcriptomics*)

- Yang Yang, Anne Bogler, Zeev Ronen, Gideon Oron, Moshe Herzberg, and Roy Bernstein (2020). Initial Deposition and Pioneering Colonization on Polymeric Membranes of Anaerobes Isolated from an Anaerobic Membrane Bioreactor (AnMBR). 55:5832-5842. <https://doi.org/10.1021/acs.est.9b06763>. (*16S Amplicon sequencing*).
- Vered Chalifa-Caspi (2021). RNA-Seq in Nonmodel Organisms. In: Noam Shomron (ed.), *Deep Sequencing Data Analysis, Methods in Molecular Biology*, vol. 2243: 143-167, Springer Science Business Media, LLC, part of Springer Nature. https://doi.org/10.1007/978-1-0716-1103-6_8. (*Non-model RNA-Seq pipeline*)
- Talia Gabay, Guy Rotem, Osnat Gillor, Yaron Ziv (2022). Understanding changes in biocrust communities following phosphate mining in the Negev Desert. *Environmental Research*. 207: 112200. <https://doi.org/10.1016/j.envres.2021.112200>. (*16S Amplicon sequencing*).

Professional Development

- **Next generation sequencing analysis training for clinical pathogen genomics by the American society for microbiology (ASM)** 2023
3-month long workshop and seminars on next generation sequencing infrastructure, techniques, and analysis for pathogen genomics.
- **Biological data science workshop, online via zoom by Drexel University** 2022/2023
Five-day workshop on bash, cloud computing, containerization, NCBI, Biopython, machine learning and deep learning.
- **Texas A&M University workshops** online via zoom 2021-
Annual short train workshops on high performance computing, containerization, python, Fortran, etc
- **Linux bootcamp: Command-line, AWK and SED; Vi editor and Grep** on udemy.com 2019
10 hours of learning to program with Linux and Linux tools
- **Advanced R** on udemy.com 2019
4.5 hours of training on advanced topics in R such as environments, creating packages, parallelization etc.
- **Data Analysis with Python and Pandas** on stoneriverelearning.com 2019
- **4th Haifa Winter Workshop in Bioinformatics:** statistics and data visualization for omics studies 2018
University of Haifa, Israel
Five-day workshop on statistical analysis and visualization of various omics studies
- **Next Generation Sequencing Workshop** 2018
Tel Hai college, Israel
Five-day workshop on analysis of NGS dataset: phylogenetics analysis, variant calling, RNA-seq etc.
- **Data Analysis of Metagenomic Samples using QIIME** 2017
Ben Gurion University of the Negev, Beersheba, Israel
One-day workshop on the analysis of metagenomics samples using QIIME