Nur M Shahir, PhD

Ellicott City, MD, USA
nmshahir@proton.me • +1 (443) 745-4467 • https://nmshahir.com

EDUCATION

University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

- Ph.D. in Bioinformatics and Computational Biology — Aug 2014 – May 2020
  - Dissertation: Inflammatory Bowel Disease Differentially Affects Region Specific Composition and Aerotolerance Profiles of Mucosally-Adherent Bacteria
  - Adviser: Dr. Terrence S. Furey, Dr. Shezad Z. Sheikh
  - Committe: Dr. Michael I. Love, Dr. Ian Carroll, Dr. Yufeng Liu
  - Focus: Inflammatory Bowel Disease, gut microbiota, bioinformatics, 16S amplicon sequencing.

University of Maryland, Baltimore County, Baltimore, Maryland, USA

- M.S. in Statistics — Aug 2011 – Dec 2013
  - Track: Biostatistics
  - Thesis: Longitudinal Analysis of Urea Cycle Disorder Patients
  - Adviser: Dr. DoHwan Park

Massachusetts Institute of Technology, Cambridge, MA, USA

- B.S. in Mathematics — 2006 – 2010

SKILLS

PROGRAMMING
R, Python, SQL, MATLAB, HTML, bash/shell scripting, LaTeX

TOOLS & PLATFORMS
QIIME, DADA2, AWS, GCP, Git, CI/CD, Containerization (Docker, Singularity), HPC clusters, Adobe Illustrator, Adobe Premiere, ArcGIS, Excel, Rstudio, Positron

WORKFLOW MANAGERS
extflow, Snakemake

RESEARCH
16S rRNA-seq, bulk RNA-seq, scRNA-seq, WGS, GWAS, Rare Variant Association Analysis, Public Health Genomics, Statistical Modeling, Machine Learning

RELEVANT INDUSTRY EXPERIENCE

Booz Allen Hamilton Remote

- Lead Scientist — Aug 2022 – Oct 2024
  - Served as a federal contractor bioinformatician, contributing to public health genomic surveillance, pathogen genomics, and large-scale human genomics projects across CDC and NIH.
  - Redesigned and optimized public health genomics workflows to align with Nextflow nf-core standards, implementing rigorous pipeline validation and automated unit testing. Streamlined collaboration through Git, ensured reproducibility via Docker containerization, and enhanced project delivery efficiency by integrating Agile tracking in JIRA.
  - Engineered reproducible bioinformatics pipelines with Snakemake on high-performance computing (HPC) clusters, enabling robust rare variant detection and uncovering genetic associations with ulcer development in sickle cell disease.
  - Led the design and implementation bioinformatics pipelines in R and Python on Google Cloud Platform (GCP) via the NIH All of Us Researcher Workbench, spanning genomic data ingestion, QC, and advanced downstream analysis. Delivered cloud-native, production-ready workflows that accelerated large-scale genomic insights for precision medicine applications.

RESEARCH EXPERIENCE

Davenport Lab, Pennsylvania State University

- Postdoctoral Fellow — Jun 2020 – Jul 2022
  - Piloted a benchmarking study on computational approaches to identify viral transcripts from bulk and single-cell RNA sequencing data, assessing the accuracy and precision of transcript identification.
  - Designed bioinformatics workflows in snakemake for efficient data processing on HPC computing environments, employing tools including samtools, bwa, bowtie2, Kraken2, and STAR, and utilizing R for downstream analysis and visualization.
  - Mentored undergraduate students, enhancing their research skills and academic performance.

Furey Lab, University of North Carolina at Chapel Hill

- Graduate Research Assistant — Nov 2014 – May 2020
  - Identified key microbial associations using R for data analysis and visualization of 16S rRNA amplicon data from IBD patients and controls.

- Developed a consensus analysis method in R with DEseq2 and Lefse, significantly enhancing the understanding of microbial dysbiosis in IBD.
- Presented research findings at various conferences, including the American Society of Human Genetics, to engage the scientific community.
- Authored a peer-reviewed journal article, contributing to the advancement of knowledge in the field of microbial dysbiosis.

PUBLICATIONS   Shahir, NM, et.al, "Crohn's Disease Differentially Affects Region-Specific Composition and Aerotolerance Profiles of Mucosally Adherent Bacteria." *Inflammatory bowel diseases*, vol. 26, no. 12, pp. 1843–1855, 2020.

PRESENTATIONS   **EXTERNAL TALKS**

- *Crohn's Disease Differentially Affects Intestinal Region Composition and Aerotolerance Profiles of Mucosally-Adherent Bacteria*, Remote   May 2020
  Virtual Microbiome Summit
- *IBD differentially affects region specific composition and aerotolerance profiles of mucosal-adherent bacteria*, Remote   May 2020
  MIT and UNC Joint Virtual Microbiome Seminar Series

UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL

- *Crohn's Disease and the Intestinal Microbiota*,
  Chapel Hill, NC   Dec 2016
  Genetics Research Colloquium
- *Alterations in the Mucosal-Adherent Enteric Microbiota Between CD and nonIBD*,
  Chapel Hill, NC   Oct 2016
  Translations Medicine Closed Door Talks
- *A distinct microbiota signature characterizes patients with penetrating Crohn's disease*,
  Chapel Hill, NC   Oct 2015
  Center for Gastrointestinal Biology and Disease
- *Analysis of the Composition and Diversity of the Colonic Mucosa Microbiota in Crohn's Disease*,
  Chapel Hill, NC   May 2015
  Bioinformatics and Computational Biology Curriculum New Student Talks
- *Identification of SERPINA1 Splice Variants from Next-Gen Sequencing Data*,
  Chapel Hill, NC   Oct 2014
  Bioinformatics and Computational Biology Research in Progress Talks

POSTERS   **EXTERNAL**

- *Analysis of mucosal adherent 16S rRNA reveals altered microbial composition and decreased diversity in patients with Crohns disease*,
  Baltimore, MD   Oct 2015
  American Society for Human Genomics

UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL

- *Characterizing the Intestinal Mucosal Landscape in Inflammatory Bowel Disease*,
  Chapel Hill, NC   Aug 2017
  Genetics Department Retreat
- *Analysis of the Composition and Diversity of the Colonic Mucosa Microbiota in Crohn's Disease*,
  Chapel Hill, NC   Aug 2016
  Genetics Department Retreat
- *Analysis of the Composition and Diversity of the Colonic Mucosa Microbiota in Crohn's Disease*,
  Chapel Hill, NC   Jul 2015
  Center for Gastrointestinal Biology and Disease Poster Session
- *Analysis of the Composition and Diversity of the Colonic Mucosa Microbiota in Crohn's Disease*,
  Chapel Hill, NC   May 2015
  Information Technology Services Research Computing Symposium

HONORS & AWARDS
- **NIH T32 Training Fellow**   2015
  Bioinformatics and Computational Biology Predoctoral Training Grant

OTHER RESEARCH EXPERIENCE

National Human Genome Research Institute, Bethesda, MD

- Mentors: Dr. Julie Segre, Dr. Sean Conlan   Jun 2012– Aug 2012
  Summer Fellow
  - Piloted a study on the viral diversity of the human skin through the use of metagenomic datasets acquired from the human microbiome project.
  - Assisted with fungal speciation of Malassezia species through bioinformatic tools
  - Applied various bioinformatic tools including: Clustal, BioPython, BioPerl, and Bowtie
  - Extracted full viral genomes from metagenomic datasets
  - Presented research at NHGRI and NIH poster sessions.

National Human Genome Research Institute, Bethesda, MD

- Mentors: Dr. Ellen Sidransky, Dr. Nahid Tayebi   Jun 2009– Aug 2009
  Summer Fellow
  - Worked on defining the association between glucocerebrosidase mutations and Parkinsons disease.
  - Learned and applied biological methods: including sequencing, PCR, westerns, RNA and protein extractions.
  - Performed statistical analysis on gene expression data via Excel
  - Trained incoming fellows in lab protocols
  - Presented research at NHGRI and NIH poster sessions

National Cancer Institute, Bethesda, MD

- Mentors: Dr. David D. Roberts, Dr. Michael Pendrak   Jun 2007– Aug 2007
  Summer Fellow
  - Worked on the development of a modification of a tetracycline-regulating system for developmental regulation in Candida albicans.
  - Learned and applied methods in molecular biology and bioinformatics including: PCR, gel electrophoresis, BLAST, ClustalW, DNA purification.
  - Trained incoming high school intern in lab protocols.

LANGUAGES

- English: Native language.
- French: Intermediate (reading); basic (speaking, writing).