

# FAYAZ SEIFUDDIN



## BACKGROUND

- Experienced bioinformatics leader with 15+ years in computational genomics, data science, and machine learning, driving innovation in multi-omics research
- Proven track record in developing and overseeing bioinformatics strategies, optimizing data processing pipelines for genomics, transcriptomics, and epigenomics
- Expertise in managing large-scale NGS datasets, ensuring data integrity, reproducibility, and actionable insights for translational and clinical applications
- Strong background in leading cross-functional teams, collaborating with researchers, clinicians, and data scientists to advance bioinformatics initiatives
- Published scientist with a history of driving impactful research, translating complex genomic data into meaningful discoveries



## EMPLOYMENT

Current  
|  
2021  
  
2021  
|  
2019  
  
2016  
|  
2012  
  
2012  
|  
2008  
  
2008  
|  
2007

- **Senior Scientist in Computational Genomics**  
Bristol Myers Squibb 📍 Lawrenceville, NJ
- **Bioinformatics Staff Scientist**  
National Institutes of Health (NIH), National Heart, Lung, and Blood Institute (NHLBI) 📍 Bethesda, MD
- **Senior Bioinformatics Research Data Analyst**  
Mood Disorders Center, Department of Psychiatry and Behavioral Sciences, The Johns Hopkins University, School of Medicine 📍 Baltimore, MD
- **Bioinformatics Research Data Analyst**  
Mood Disorders Center, Department of Psychiatry and Behavioral Sciences, The Johns Hopkins University, School of Medicine 📍 Baltimore, MD
- **Bioinformatics Intern**  
J. Craig Venter Institute (JCVI) 📍 Rockville, MD

## CONTACT

✉ [seifudd@gmail.com](mailto:seifudd@gmail.com)  
📞 (469) 583-0651  
📍 12000 Amber Ridge Circle,  
Germantown, MD 20876  
🌐 [github.com/seifudd](https://github.com/seifudd)

## PUBLICATIONS

🔍 [scholar.google.com/citations?user=x8yJW3MAAAAJ&hl=en](https://scholar.google.com/citations?user=x8yJW3MAAAAJ&hl=en)

## COMPUTATIONAL SKILLS



Made with the R package [pagedown](#).

The source code is available on [github.com/seifudd/cv](https://github.com/seifudd/cv).

Last updated on 2025-02-18.



## EDUCATION

2019  
|  
2016

### PhD., Bioinformatics and Computational Biology

George Mason University

📍 Fairfax, VA

- **Dissertation:** lncRNAKB- A Knowledgebase of Tissue-specific Functional Annotation and Trait Association of Long Noncoding RNA

2008  
|  
2006

### M.S., Bioinformatics and Computational Biology

George Mason University

📍 Fairfax, VA

- **Thesis:** Identification of Lung Adenocarcinoma Progression Biomarkers

2005  
|  
2001

### B.S., Computer Science (minor-Mathematics)

Linfield University

📍 McMinnville, OR



## INDUSTRY & ACADEMIC EXPERIENCE

Current  
|  
2021

### Senior Scientist in Computational Genomics

Bristol Myers Squibb

📍 Lawrenceville, NJ

- **Leading NGS data analysis** for clinical trials, exploratory, and research studies, deriving actionable insights for precision medicine
- **Developing and optimizing bioinformatics pipelines**, integrating AWS, Docker, CWL, in Seven Bridges and Arvados platforms for scalable, reproducible workflows
- **Directing innovative research projects** on Microsatellite Instability (MSI), alternative splicing analysis, neoantigen prediction, Tumor Mutational Burden (TMB), genetic ancestry analysis, cfDNA nucleosome profiling and RNA-Seq lite in clinical applications
- **Designing novel tools and workflows** to support long-read sequencing (PacBio) data and advanced transcriptomics analyses

2021  
|  
2016

### Bioinformatics Staff Scientist & PhD. Research Fellow (in Bioinformatics and Computational Biology)

National Institutes of Health (NIH), National Heart, Lung, and Blood Institute (NHLBI)

📍 Bethesda, MD

- **RNA-Seq (Transcriptomics):** Led large-scale transcriptomic analyses (GTEx, 11,000+ samples), developing HPCC pipelines for expression quantification, differential analysis, pathway enrichment, eQTL and sQTL mapping
- **DNA-Seq (Genomics):** Designed and implemented HPCC workflows for WGS, WES, targeted DNA-seq, and CRISPR studies, including germline/somatic mutation calling, CNV detection, and mutational burden analysis
- **Methyl-Seq (Epigenomics):** Built WGBS pipelines for genomic and cfDNA analysis, performing CpG methylation analysis, deconvolution for tissue-of-origin identification, and processing ChIP-Seq, CUT&Tag, and CUT&RUN data
- **Proteomics:** Applied machine learning models (random forest and mixed-effects modeling) to predict cardiovascular risk using high-density lipoprotein proteome data
- **lncRNA Analysis:** Developed lncRNAKB (<https://osf.io/ru4d2/>), a knowledgebase integrating lncRNA tissue-specificity and disease associations, incorporating cis-eQTL, GWAS associations (UK Biobank), and co-expression modules from GTEx RNA-Seq data

2016  
|  
2012

### Senior Bioinformatics Research Data Analyst

Mood Disorders Center, Department of Psychiatry and Behavioral Sciences, The Johns Hopkins University, School of Medicine

📍 Baltimore, MD

- **Led a systematic review and mega-analysis** of microarray gene expression studies in brain tissue for mood disorders
- **Developed Metamoodics**, a web-based bioinformatics platform integrating genome-wide linkage, expression, and association data for Major Depression and Bipolar Disorder
- **Conducted polygenic risk score (PRS) analysis**, linking genetic predisposition to stressful life events and depressive symptoms in older adults
- **Designed and implemented HPCC pipelines** for epigenetic data analysis, including Whole Genome Bisulfite Sequencing (WGBS), to study mood disorder associations

2012  
|  
2008

### Bioinformatics Research Data Analyst

Mood Disorders Center, Department of Psychiatry and Behavioral Sciences, The Johns Hopkins University, School of Medicine

📍 Baltimore, MD

- **Designed and executed scalable bioinformatics pipelines** on High-Performance Computing Clusters (HPCCs) for GWAS analysis of mood disorders, integrating Illumina & Affymetrix array data
- **Led imputation and statistical analysis** of GWAS data, enhancing the resolution of genetic association studies
- **Conducted a comprehensive meta-analysis** of candidate gene studies in Bipolar Disorder, identifying statistically significant polymorphisms
- **Applied machine learning and data mining techniques** to analyze high-dimensional genomic data from psychiatric disorders
- **Developed analytical workflows** for genome-wide DNA methylation (DNAm) analysis

2008  
|  
2007

### Bioinformatics Intern

J. Craig Venter Institute (JCVI)

📍 Rockville, MD

- **Developed SMURF** (<http://smurf.jcvi.org/index.php>), a web-based tool for genomic mapping of fungal secondary metabolite clusters, enabling systematic prediction of biosynthetic gene clusters
- **Designed algorithms** to analyze genomic sequences and identify gene clusters responsible for biosynthesis, export, and transcriptional regulation of secondary metabolites
- **Enhanced data-driven insights** into fungal metabolomics, providing researchers with a streamlined approach to studying secondary metabolism



## TEACHING EXPERIENCE

2019

### Trainer, Bioinformatics Analysis for NGS Data

The NHLBI Workshop Series

📍 Bethesda, MD

- Trained 20 staff scientists on NGS platforms and data analysis techniques
- Delivered sessions on data quality assessment (FastQC), preprocessing (Trimmomatic, Cutadapt), and alignment (BWA, Bowtie2)
- Taught variant calling (GATK, samtools) and annotation (Annovar, VEP)
- Guided gene expression analysis (Salmon, Kallisto, DESeq2, edgeR)
- Introduced data visualization tools (IGV, R/Bioconductor)
- Covered functional enrichment and pathway analysis (DAVID, GOSep)
- Emphasized best practices, troubleshooting, and validation in NGS analysis

- 2019  
|  
2017

**Guest Lecturer**

George Mason University, Department of Bioinformatics and Computational Biology

  - Lectured M.S. and Ph.D. students on current trends, methods, and analysis of biological sequence and NGS functional genomics data
  - Engaged and instructed a class of 15 students

📍 Fairfax, VA
  
- 2006  
|  
2005

**Tutor Coordinator & Mentor**

Dallas College North Lake Campus, TRiO Upward Bound

  - Hired, trained, and mentored a team of 10 tutors for the Upward Bound program, supporting first-generation and low-income high school students in college preparation

📍 Dallas, TX
  
- 2005  
|  
2004

**Teaching & Lab Assistant**

Linfield University, Department of Computer Science

  - Designed, assisted, and graded C++ and Java programming assignments for 20 first- and second-year students in Programming and Data Structures courses

📍 McMinnville, OR
  
- 2005  
|  
2001

**Tutor**

Linfield University, TRiO Upward Bound

  - Tutored 60 students in Math and Science, focusing on Algebra I & II, Geometry, Pre-Calculus, and Calculus

📍 McMinnville, OR



## HONORS AND AWARDS

- 2019  
|  
2016

**NIH Pre-Doctoral Intramural Research Training Award (IRTA)**

NHLBI

📍 Bethesda, MD
  
- 2019

**Graduate Student Research Award in Bioinformatics / Biostatistics / Epidemiology / Computational & Systems Biology**

NIH, Graduate Student Research Symposium

📍 Bethesda, MD
  
- 2019

**Outstanding Digital Presentation Award in Bioinformatics / Computational & Systems Biology**

George Mason University, Student Research Day

📍 Fairfax, VA
  
- 2019

**Reviewers Choice Poster Abstracts, Bioinformatics & Computational Biology**

The American Society of Human Genetics (ASHG)

📍 Houston, TX
  
- 2005  
|  
2001

**Sumie Yamamoto Endowed International Students Scholarship**

Linfield University

📍 McMinnville, OR