Belmont, CA
☎ +1 513 560 4978
✉ shruti.marwaha@gmail.com
www.linkedin.com/in/shrutimarwaha
www.github.com/ShrutiMarwaha

Shruti Marwaha

Skills

Programming Languages	R (& Bioconductor), Python, SQL, Perl.
High Throughput Data Analysis	Variant Calling, Variant Prioritization, bulk RNA-seq, Mass Cytometry.
Machine Learning	Caret (R), Scikit-learn (Python).
Tools	Amazon Web Services (AWS), Google Cloud Platform (GCP), High-Performance Computing (HPC) cluster, Git.
Operating Systems	Linux, Mac, Windows.

Education

Aug 2010 - Aug 2015 **PhD in Systems Biology & Physiology**, *University of Cincinnati*, OH.

Thesis title *Genomics and Mathematical Modeling approach to study Helicobacter pylori associated Gastritis and Gastric Cancer.*

Description Key research projects:

Drug Repurposing
- *Analyzed high-throughput gene expression data* to identify key genes and pathways dysregulated in gastric cancer and gastric atrophy.
- Developed computational pipeline on AWS for RNAseq analysis and gene-signature based drug repurposing.

Mathematical Modeling
- Developed an Ordinary Differential Equation (ODE) model of the host's immune response to *H. pylori*, with focus on crosstalk between cytokines and sonic hedgehog.
- The model helped to identify a novel feedback loop and potential oscillatory behavior of the system.

2005 - 2007 **Masters in Bioinformatics (Post Graduate Diploma)**, *Institute of Bioinformatics & Applied Biotechnology*, Bangalore, India.

Internship Project Developed an interaction network of Angiopoietin pathway in cancer by literature curation and mathematically modeled it using Ordinary Differential Equations.

2002 - 2005 **B.Sc (Hons) in Zoology**, *University of Delhi*, Delhi, India.

Experience

Dec 2017 - Present **Research Engineer**, *Stanford University*, CA.

Project I Title **Undiagnosed Disease Network** (UDN) consortium, whose primary goal is to improve the diagnostic outcomes of patients who have struggled to obtain a clear diagnosis for years through traditional clinical means. `https://undiagnosed.hms.harvard.edu/`

Dec 2017 - Present; Description

My role involves applying and developing computational pipelines and using the latest biological assays that are often not yet available in the clinic, to help identify diagnostic leads for the ultra-rare-disease patients.

- Prioritization of small (SNV and INDEL) and large (Structural Variants) variants. Benchmarked Exomiser/Genomiser tools using 386 diagnosed cases from UDN and identified optimal parameters for the variant prioritization pipeline.
- Analyze high throughput Immunological data (Luminex, Mass Cytometry) to identify immune cells and proteins that are dys-regulated in a patient as compared to healthy controls.
- Wrote a seed grant on Inborn Errors of Immunology (IEI), which was awarded $80K to fund future experiments.
- Mentor students involved in the research projects, and write manuscripts and grants.

Project II Title

Genomics Research to Elucidate the Genetics of Rare diseases (GREGoR) consortium, whose primary goal is to improve the diagnostic yield for rare disease patients using multi-omics approach. `https://gregor.stanford.edu/`

Aug 2021 - Present; Description

My role involves applying and developing computational pipelines and also building a bioinformatics team.

- Responsible for generating and processing metadata and genomic files for data sharing across the GREGoR consortium.
- Bioinformatics Hiring: Hired four bioinformaticians. Streamlined the Bioinformatician hiring process and was actively involved in the first four rounds of screening and evaluations. Designed a coding exercise and the calibrated a criterion required for candidate evaluation on the coding exercise.

Project III Title

Molecular Transducers of Physical Activity Consortium (MoTrPAC), a consortium to study the molecular changes that occur during and after exercise. `https://www.motrpac.org/`

Dec 2017 - July 2022; Description

- Study the impact of exercise on Mitochondrial Heteroplasmy.
- Performed quality control for the Luminex assay.

.

Aug 2016 - Nov 2017

Sr. Computational Biologist, *Stanford University*, CA.

Project Title

Undiagnosed Disease Network (UDN).

Project Description

- Developed computational pipelines for analysis of DNA sequencing data and Variant Prioritization.
- Evaluated different variant calling tools against gold standard Genome In A Bottle dataset and calculated their precision and recall.
- Using API calls, extracted and parsed data from UDN portal.

Oct 2015 - Dec 2015

Research Associate, *Cincinnati Children's Hospital Medical Center*, OH.

Project Title

Drug Repurposing for Head and Neck Squamous Cell Carcinoma (HNSCC).

Project Description

- Identified the genes and pathways that are differentially expressed between HPV (Human Papilloma Virus) positive and HPV negative Head and Neck Cancer cells.
- Identified potential compounds that can be used for treatment of HNSCC.

Nov 2007 - Mar 2010

Associate Team Lead Scientist, *Cellworks Research India Ltd*, India.

Project Description

- Developed Ordinary Differential Equation models of molecular pathways that are perturbed in cancer to *identify potential drug targets.*
- Case-study example: The model predicted higher efficacy of mTORC1 inhibitor in PTEN and KRAS wild type system as compared to that in PTEN and KRAS mutant model, highlighting the importance of mTORC1-IRS1-AKT-ERK feedback loop.
- Managed and mentored a team of 7 members.

May 2007 - Oct 2007 **Associate Biomodeling Scientist**, *Cellworks Research India Ltd*, India.

Project Description
- Mathematically modeled mTOR-HIF, glycolysis and TCA pathways to simulate tumor metabolism.

Data Science Projects

April 2016 - May 2016 **Kaggle Challenge**, *San Francisco Crime Classification*.

Project Description
- Developed a machine learning model to predict the category of a crime that occurred in past 12 years in San Francisco.
- Built a multiclass classifier from ~880,000 samples to classify the crime into one of the 39 possible categories.
- `https://github.com/ShrutiMarwaha/KaggleSFCrime/`

Feb 2016 - Mar 2016 **Chronic Kidney Disease Prediction**.

Project Description
- Developed a predictive model from clinical features to identify patients at the risk of Chronic Kidney Disease using Support Vector Machine.
- Implemented the model as a web application `https://shruti-marwaha.shinyapps.io/KidneyDiseasePredictor/`

Key Publications

Until 09/11/2022 Citations: 1664, h-index: 11

`https://scholar.google.com/citations?hl=en&user=58b8fRwAAAAJ`

Peer Review Publications

Isabelle B. Cooperstein , Marwaha S, , Alistair Ward, Shilpa N. Kobren , Jennefer N. Carter, Undiagnosed Diseases Network, Matthew T. Wheeler and Gabor T. Marth. **An optimized variant prioritization process for rare disease diagnostics: recommendations for Exomiser and Genomiser**, medRxiv, 2025 April.

John E. Gorzynski*, Marwaha S*, Chloe Reuter*, Tanner D. Jensen, Alexis Ferrasse, Archana Natarajan Raja, Liliana Fernandez, Elijah Kravets, Jennefer Carter, Devon Bonner, Shirley Sutton, Undiagnosed Diseases Network, Maura Ruzhnikov,Louanne Hudgins, Paul G Fisher, Jonathan A. Bernstein, Matthew T. Wheeler, Euan A.Ashley. **Clinical application of Complete Long Read genome sequencing identifies a 16kb intragenic duplication in EHMT1 in a patient with suspected Kleefstra syndrome**, medRxiv, 2024 June. *joint first authors

Sarah Fazal, Matt C Danzi, Isaac Xu, Shilpa Nadimpalli Kobren, Shamil Sunyaev, Chloe Reuter, Marwaha S, Matthew Wheeler, Egor Dolzhenko, Francesca Lucas, Stefan Wuchty, Mustafa Tekin, Stephan Züchner, Vanessa Aguiar-Pulido. **RExPRT: a machine learning tool to predict pathogenicity of tandem repeat loci**, Genome biology, 2024 Jan.

Monica H Wojcik, Chloe M Reuter, Marwaha S, Medhat Mahmoud, Michael H Duyzend, Hayk Barseghyan, Bo Yuan, Philip M Boone, Emily E Groopman, Emmanuèle C Délot, Deepti Jain, Alba Sanchis-Juan, Lea M Starita, Michael Talkowski, Stephen B Montgomery, Michael J Bamshad, Jessica X Chong, Matthew T Wheeler, Seth I Berger, Anne O'Donnell-Luria, Fritz J Sedlazeck, Danny E Miller. **Beyond the exome: what's next in diagnostic testing for Mendelian conditions**, American Journal of Human Genetics, 2023 Aug.

Rebecca C Spillmann, Queenie K-G Tan, Chloe Reuter, Kelly Schoch, Undiagnosed Diseases Network, Jennefer Kohler, Devon Bonner, Diane Zastrow, Anna Alkelai, Evan Baugh, Heidi Cope, Marwaha S, Matthew T Wheeler, Jonathan A Bernstein, Vandana Shash et. al. **A concurrent dual analysis of genomic data augments diagnoses: experiences of 2 clinical sites in the Undiagnosed Diseases Network**, Genetics in Medicine, 2023 April.

Marwaha S, Joshua W Knowles, Euan A Ashley. **A guide for the diagnosis of rare and undiagnosed disease: beyond the exome**, Genome Medicine, 2022 Feb.

Kohler JN, Kelley EG, Boyd BM, Sillari CH, Marwaha S, UDN, Wheeler M. **Genetic counselor roles in the undiagnosed diseases network research study: Clinical care, collaboration, and curation**, J Genet Couns. 2022 Apr.

Sardana D, Marwaha S, Bhatnagar R. **Supervised and Unsupervised Machine Learning Methodologies for Crime Pattern Analysis**, International Journal of Artificial Intelligence and Applications, 2021 Nov.

Chen YH, Zastrow DB, Metcalfe RD, Gartner L, Krause F, Morton CJ, Marwaha S, Fresard L, Huang Y, Zhao C, McCormack C, Bick D, Worthey EA, Yang Y, Eng CM, Gold J, Undiagnosed Diseases Network, Montgomery SB, Fisher PG, Ashley EA, Wheeler MT, Parker MW, Shanmugasundaram V, Putoczki TL, Schmidt-Arras D, Laurence A, Bernstein JA, Griffin MDW, Uhlig HH. **Functional and structural analysis of cytokine-selective IL6ST defects that cause recessive hyper-IgE syndrome**, J Allergy Clin Immunol. 2021 Aug.

Kobren SN, Baldridge D, Velinder M, Krier JB, LeBlanc K, Esteves C, Pusey BN, Züchner S, Blue E, Lee H, Huang A, Bastarache L, Bican A, Cogan J, Marwaha S, Alkelai A, Murdock DR, Liu P, Wegner DJ, Paul AJ, Undiagnosed Diseases Network, Sunyaev SR, Kohane IS. **Commonalities across computational workflows for uncovering explanatory variants in undiagnosed cases.** Genetics in Medicine, June 2021.

Burke EA, Sturgeon M, Zastrow DB, Fernandez L, Prybol C, Marwaha S, Frothingham EP, Ward PA, Eng CM, Fresard L, Montgomery SB, Enns GM, Fisher PG, Wolfe LA, Harding B, Carrington B, Bishop K, Sood R, Huang Y, Elkahloun A, Toro C, Bassuk AG, Wheeler MT, Markello TC, Gahl WA, Malicdan MCV. **Compound heterozygous KCTD7 variants in progressive myoclonus epilepsy.** J Neurogenet., June 2021.

Reuter CM, Kohler JN, Bonner D, Zastrow D, Fernandez L, Dries A, Marwaha S, Davidson J, Brokamp E, Herzog M, Hong J, Macnamara E, Rosenfeld JA, Schoch K, Spillmann R; Undiagnosed Diseases Network, Loscalzo J, Krier J, Stoler J, Sweetser D, Palmer CGS, Phillips JA, Shashi V, Adams DA, Yang Y, Ashley EA, Fisher PG, Mulvihill JJ, Bernstein JA, Wheeler MT. **Yield of whole exome sequencing in undiagnosed patients facing insurance coverage barriers to genetic testing.**, J Genet Couns. 2019 Sep.

Hom J, Marwaha S, Postolova A, Kittle J, Vasquez R, Davidson J, Kohler J, Dries A, Betancourt L, Majcherska M, Dearlove J, Vogel H, Bernstein J, Fisher P, Ashley E, Sampson J, Wheeler M. **A Patient with Sjogren's Syndrome and Subsequent Diagnosis of Inclusion Body Myositis and Light-Chain Amyloidosis**, Journal of General Internal Medicine, 2019 Jun;34(6):1058-1062.

Frésard L, Smail C, Ferraro NM, Teran NA, Li X, Smith KS, Bonner D, Kernohan KD, Marwaha S, Zappala Z, Balliu B, Davis JR, Liu B, Prybol CJ, Kohler JN, Zastrow DB, Reuter CM, Fisk DG, Grove ME, Davidson JM, Hartley T, Joshi R, Strober BJ, Utiramerur S; Undiagnosed Diseases Network; Care4Rare Canada Consortium, Lind L, Ingelsson E, Battle A, Bejerano G, Bernstein JA, Ashley EA, Boycott KM, Merker JD, Wheeler MT, Montgomery SB. **Identification of rare-disease genes from whole blood transcriptome sequencing of diverse undiagnosed cases using large control cohorts**, Nat Med. 2019 Jun.

McGuire JL, Depasquale EA, Funk AJ, O'Donnovan SM, Hasselfeld K, Marwaha S, Hammond JH, Hartounian V, Meador-Woodruff JH, Meller J, McCullumsmith RE. **Abnormalities of signal transduction networks in chronic schizophrenia**, NPJ Schizophrenia. 2017 Sep.

Marwaha S, Schumacher MA, Zavros Y, Eghbalnia HR. **Crosstalks between Cytokines and Sonic Hedgehog in Helicobacter pylori Infection: A Mathematical Model**, PLoS One. 2014 Nov.

Key Poster Presentations

Marwaha S, Ashley E, Wheeler M. **Comparison of Indel Callers and metrics to evaluate performance of new tools.** *American Society for Human Genetics Conference*, Oct 19, 2018. San Diego, CA.

Marwaha S, Hom J, Postolova A, Sampson J, Dearlove J, Kittle J, Vasquez R, Kohler J, Betancourt ELF, Dries AM, Vogel OH, Members of the Undiagnosed Diseases Network, Fisher P, Bernstein J, Ashley E, Wheeler M. **Mass Cytometry Immunophenotyping to augment diagnosis of a Rare Disease.** *Stanford-EMBL Personalized Health Conference*, Nov 2, 2017. Stanford, CA.

Marwaha S, Eghbalnia H. **Drug Repurposing for Gastric Cancer using Genomics Data.** *Southern California Systems Biology Conference*, January 31, 2015. Irvine, CA.

Marwaha S, Schumacher M, Zavros Y, Eghbalnia H. Interactions between cytokines and sonic hedgehog in *H. pylori* mediated gastritis implicates a novel feedback circuit. *International Conference on Computational Cell Biology*, August 14-16, 2013. Blacksburg, Virginia.

Oral Presentation

Marwaha S, Schumacher M, Zavros Y, Eghbalnia H. Feedbacks in sonic hedgehog circuit with cytokines in *H. pylori* mediated gastritis. *Ohio Physiological Society*, October 6-7, 2011. Cincinnati, OH.

Awards

2021 **Bio-X Star Mentor Award**, *Stanford*, for mentoring an undergraduate student with research.

2018 **Reviewers' Choice Abstract**, *at American Society for Human Genetics Conference*, for "Comparison of Indel Callers and metrics to evaluate performance of new tools", abstract scored in the top 10% of poster abstracts.

2011 **Peter K Lauf Award**, *at Ohio Physiological Society*, for presentation on "Feedback in Sonic hedgehog circuit with cytokines in H. pylori mediated gastritis".

2010 - 2015 **University Graduate Scholarship**, *University of Cincinnati*.

Journal Review Activities

Sept 2022 - Sept 2024 **Junior Editorial Board Member**, *Journal of Translational Genetics and Genomics*.
`https://publons.com/researcher/4063963/shruti-marwaha/peer-review/`

Aug 2023 **NEJM-AI**.

July 2023 **Molecular Biology Reports**.

May 2023 **Nature Communications**.

Jan 2023 **BMC Genomic Dat**.

Nov 2022 **Frontiers in Genetics**.

Aug 2022 **Journal of Medical Genetics**.

May 2022 **Journal of Translational Genetics and Genomics**.

May 2022 **Journal of Medical Genetics**.

Oct 2021 **Orphanet Journal of Rare Diseases**.

June 2021 **Orphanet Journal of Rare Diseases**.

Mar 2021 **Orphanet Journal of Rare Diseases**.

Mar 2021 **Frontiers in Genetics**.

Dec 2020 **BMC Bioinformatics**.

Dec 2020 **Orphanet Journal of Rare Diseases**.

Professional Affiliations

2014 - 2015 Vice President for Physiology Student Organization, University of Cincinnati.

2012 - 2013 Secretary for Health Sciences Graduate Student Association, University of Cincinnati.

References available on request