# VIVEK RAI

+1-734-846-5791 • mail@raivivek.in • raivivek.in • **github**/raivivek • **orcid**/0000-0002-2058-7238

## EDUCATION

| | |
|---|---|
| 04-2022 (exp) | **UNIVERSITY OF MICHIGAN**<br>*Doctor of Philosophy,* Bioinformatics<br>*Master of Arts,* Statistics |
| 2017 | **INDIAN INSTITUTE OF TECHNOLOGY, KHARAGPUR**<br>*Bachelor & Master of Technology (Dual Degree),* Biotechnology & Biochemical Engineering |

## RESEARCH EXPERIENCE

| | |
|---|---|
| 2018 (–now) | **UNIVERSITY OF MICHIGAN** \| RESEARCH ASSISTANT<br>Mentor: Stephen CJ Parker, Ph.D.<br>• Collaborated with multi-institution and inter-disciplinary research teams on multiple research projects to investigate epigenetic mechanisms underlying regulatory genetic variation in type 2 diabetes (T2D).<br>• Co-led projects integrating multi-omics datasets (RNA, ATAC, Imaging) at bulk and single-cell resolution using statistical and machine learning approaches to identify disease defining features of T2D.<br>• Contributed to open-source bioinformatics tool design and development; and created sequencing data processing pipelines using workflow-programming languages. |
| 2018 (Jan-Apr) | **UNIVERSITY OF MICHIGAN,** RESEARCH ASSISTANT<br>Mentor: Cristen Willer, Ph.D.<br>• Programmatically mined the GWAS catalog data to identify genomic loci associated with cardiovascular disease markers to prioritize candidate drug targets.<br>• Built a polygenic risk score (PRS) predictive model to identify individuals at the risk of cardiovascular diseases using LDL-cholesterol data in UK Biobank (UKBB). |
| 2015–2017 | **IIT KHARAGPUR,** MASTER'S THESIS<br>Mentor: Amit K Das, Ph.D.<br>• Systematically profiled protein structures from Protein Data Bank (PDB) to using knot finding algorithms in Python and identified structural and functional patterns that characterize a knotted-protein structure. |
| 2016 (May-Jun) | **INSTITUTE FOR SYSTEMS BIOLOGY,** RESEARCH INTERN<br>Mentor: Gustavo Glusman, Ph.D.<br>• Analyzed >2000 individual whole-genome coverage profiles to characterize technical biases in the downstream analyses due to errors in sequencing mapping pipeline.<br>• Showed that sex-specific biases in whole-genome analysis resulted from mapping biases due to segmental duplication between chromosomes. |
| 2016 (May-Aug) | **GOOGLE SUMMER OF CODE,** DEVELOPER<br>Mentors: Bastian Greshake, Ph.D. and Philip Bayer, Ph.D.<br>• Wrote and conceived the proposal to enhance the user-experience of OpenSNP portal by allowing the users to see phenotypes linked to their variants of interest.<br>• Designed the database schema and Ruby-on-Rails backend API to link SNPs ("Genotypes") to Phenotypes based on trait-variant association data mined from literature. |
| 2015–2016 | **DEVELOPER** — WIKIMEDIA (INDIVIDUAL ENGAGEMENT GRANT)<br>Team Leader: Maximilien Klein |

- Contributed to development of "Wikidata Human Gender Indicators" (WHGI), an open source, longitudinal dataset that helped quantify gender disparities in the representation of women across multiple language Wikipedias.
- Developed interactive visualizations and co-wrote manuscript that showcased WHGI as a suitable index for monitoring gender disparities across time, space, culture, occupation, and language.

## TECHNICAL SKILLS

| | |
|---|---|
| *domains* | **Multi-Omics Bioinformatics • Functional Genomics • Results-Oriented Research** |
| *languages* | Python • R • Shell • Scripting (Bash) • Ruby • SQL |
| *web* | HTML/CSS • ReactJS • JavaScript • Static-site Development |
| *tools* | Git • Snakemake • Nextflow • Unix CLI • Singularity/Docker |
| *computing* | HPC (Slurm) • Google Cloud Platform • Amazon Web Services (AWS) • Azure DevOps |
| *open-source* | Bioconda • conda-forge • scikit-bio • Afra • SequenceServer • makebio • awesome-biology |

## PUBLICATIONS* (5 OF 10)

| | |
|---|---|
| 2021 | Walker JT†, Saunders CD†, **Rai V**†, et al. *RFX6-mediated dysregulation mediates human β cell dysfunction in early type 2 diabetes.* BiorXiv (In review at **Nature**). |
| 2020 | **Rai V**†, Quang DX†, Erdos MR, Cusanovich DA, Daza RM, Narisu N, et al. *Single-cell ATAC-Seq in human pancreatic islets and deep learning upscaling of rare cells reveals cell-specific type 2 diabetes regulatory signatures.* Molecular Metabolism 32: 109–121. ([featured on **Cover Page**](#)) |
| 2019 | Priyam A, Woodcroft BJ, **Rai V**, Moghul I, Munagala A, Ter F, et al. *Sequenceserver: A Modern Graphical User Interface for Custom BLAST Databases.* Mol Biol Evol. 36(12), 2922-2924. |
| 2018 | Grüning B, Dale R, Sjödin A, Chapman BA, Rowe J, Tomkins-Tinch CH, Köster J, & **The Bioconda Team**. *Bioconda: sustainable and comprehensive software distribution for the life sciences.* Nature Methods. 15(7), 475-476. |
| 2016 | Klein M, Gupta H, **Rai V**, Konieczny P, Zhu H. *Monitoring the Gender Gap with Wikidata Human Gender Indicators.* Proceedings of the 12th International Symposium on Open Collaboration - OpenSym '16. Berlin, Germany: ACM Press; 2016 p. 1–9. |

## CONFERENCE PRESENTATIONS (2 OF 4)

| | |
|---|---|
| 2021 | (**Oral**, Midwest Islet Club 2021, Virtual) — *Integrative Multi-Omic Approaches Define a RFX6 Network Module as Critical for Human β cell Dysfunction in Type 2 Diabetes.* |
| 2019 | (**Oral**, Midwest Islet Club 2019, Ann Arbor) — *Single-Nuclei ATAC-Seq in Human Pancreatic Islets Reveals Cell-Type Specific Type 2 Diabetes Regulatory Signatures.* |

## PROFESSIONAL DEVELOPMENT

| | |
|---|---|
| 2021 | Advancing Career Targets (ACT) Program — Univ of Michigan |
| 2020–now | **Reviewer** - Journal of Open-Source Software (JOSS), ROpenSci |
| 2020 | **ComSciCon-MI** (Communicating Science Workshop), Michigan — Online |
| 2018 | **Course Reviewer**, Python Tips, Tricks and Techniques Course — Packt Publishing |
| 2018–now | President, **Michigan Argentine Tango Club** (MATC) — Univ. of Michigan |
| 2018–now | Office Hours Assistant, Girls Who Code – DCM&B, Univ. of Michigan |
| 2016 | Intelligent Systems for Molecular Biology (ISMB) — Orlando, Florida, US |
| 2012–2018 | Contributor (**40+ articles** and 1800+ edits) — English Wikipedia |

† Co-first authors