Pushkala Jayaraman
Pushkala.jayaraman@icahn.mssm.edu

APPOINTMENTS/EMPLOYMENT

Bioinformatics Scientist III Sept 2019 – Dec 2020

Clinical Genomic Diagnostics Laboratory, Pathology, Children's Hospital of Philadelphia, Philadelphia, PA 19104

Bioinformatics Scientist II Jan 2014 - Aug 2019

Clinical Genomic Diagnostics Laboratory, Pathology, Children's Hospital of Philadelphia, Philadelphia, PA 19104

Programmer/Analyst II Jan 2010 - Dec 2013

Rat Genome Database, Human and Molecular Genetics Center, Medical College of Wisconsin, Wauwatosa, WI, 53226

Academic Employment

Ph.D. Candidate June 2022 – present

Augmented Intelligence in Medicine and Science, Division of Data Driven and Digital Medicine (D3M), Icahn School of Medicine, Mt. Sinai, New York, NY, 10128

Ph.D. Student trainee (Research Assistant) Aug 2020 – June 2022

Nadkarni Lab, Department of Genetics and Genomic Sciences, Division of Data Driven and Digital Medicine (D3M), Icahn School of Medicine, Mt. Sinai, New York, NY, 10128

Graduate Research Assistant Jan 2009 - Dec 2009

Computational Genomics Group, School of Biological Sciences, Georgia Institute of Technology, Atlanta, GA, 30332

EDUCATION

Icahn School of Medicine, Mount Sinai, New York, NY Aug 2020 - present

Ph.D. in Biomedical Sciences, Artificial Intelligence and Emerging Technologies

Graduate Research Assistant (mentor: Prof. Girish Nadkarni, MD, MPH)

Georgia Institute of Technology, Atlanta, GA 2008 – 2009

Masters in Bioinformatics

Graduate Research Assistant (mentor: Prof. I. King Jordan)

Emory University, Atlanta, GA May 2009 - Aug 2009

Department of Human Genetics

Summer Internship (mentor: Prof. Madhuri Hegde, Assistant Professor, Director Emory Genetics Laboratory)

Rashtriya Vidyalaya College of Engineering, Bangalore, India 2004 - 2008

Department of Biotechnology

B.E. Biotechnology (Semester Average 82.8%)

Indian Institute of Sciences, Bangalore, India Jan 2008 - Jul 2008

Bioinformatics Center

Summer Internship (mentor: Prof. S. Ramakumar, Associate Professor and Chairman, Bioinformatics Center, IISC, Bangalore.)

Bioinformatic Analyses of thermophiles

HONORS/AWARDS

Computational Biology Faculty Research Award '09 Spring 2009

School of Biological Sciences, Georgia Institute of Technology

Best Outgoing Student (Valedictorian) '08 July 2008

Engineering Biotechnology, Rashtriya Vidyalaya College of Engineering

PUBLICATIONS

1. Kislyuk, A. O., L. S. Katz, S. Agrawal, M. S. Hagen, A. B. Conley, **P. Jayaraman**, V. Nelakuditi, J. C. Humphrey, S. A. Sammons, D. Govil, R. D. Mair, K. M. Tatti, M. L. Tondella, B. H. Harcourt, L. W. Mayer and I. K. Jordan (2010). "A computational genomics pipeline for prokaryotic sequencing projects." *Bioinformatics (Oxford, England)* 26(15): 1819-1826. IF: 7.33
2. Katz, L. S., J. C. Humphrey, A. B. Conley, V. Nelakuditi, A. O. Kislyuk, S. Agrawal, **P. Jayaraman**, B. H. Harcourt, M. A. Olsen-Rasmussen, M. Frace, N. V. Sharma, L. W. Mayer and I. K. Jordan (2011). " Neisseria Base: a comparative genomics database for Neisseria meningitidis." *Database: the journal of biological databases and curation 2011*: bar035-bar035. IF: 3.978
3. Dwinell, M., M. Shimoyama, R. Nigam, W. Liu, M. Tutaj, J. De Pons, S.J. Wang, J. Smith, T. Lowry, G. T. Hayman, S. Laulederkind, V. Petri**, P. Jayaraman**, E. Worthey, D. Munzenmaier and H. Jacob (2012). " PhenoMiner: an interactive tool for physiologists integrating phenotype data using multiple ontologies." *The FASEB Journal* 26(1 supplement): 717.711-717.711. IF: 5.498

4. Hayman, G. T., **P. Jayaraman**, V. Petri, M. Tutaj, W. Liu, J. De Pons, M. R. Dwinell and M. Shimoyama (2013). " The updated RGD Pathway Portal utilizes increased curation efficiency and provides expanded pathway information." *Human genomics* 7(1): 4-4. IF: 3.500
5. Sullivan, J., K. Karra, S. A. T. Moxon, A. Vallejos, H. Motenko, J. D. Wong, J. Aleksic, R. Balakrishnan, G. Binkley, T. Harris, B. Hitz, **P. Jayaraman**, R. Lyne, S. Neuhauser, C. Pich, R. N. Smith, Q. Trinh, J. M. Cherry, J. Richardson, L. Stein, S. Twigger, M. Westerfield, E. Worthey and G. Micklem (2013). " InterMOD: integrated data and tools for the unification of model organism research." *Scientific reports* 3: 1802-1802. IF: 4.122
6. Nigam, R., S. J. F. Laulederkind, G. T. Hayman, J. R. Smith, S.-J. Wang, T. F. Lowry, V. Petri, J. De Pons, M. Tutaj, W. Liu, **P. Jayaraman**, D. H. Munzenmaier, E. A. Worthey, M. R. Dwinell, M. Shimoyama and H. J. Jacob (2013). " Rat Genome Database: a unique resource for rat, human, and mouse quantitative trait locus data." *Physiological genomics* 45(18): 809-816. IF: 2.782
7. Petri, V., **P. Jayaraman**, M. Tutaj, G. T. Hayman, J. R. Smith, J. De Pons, S. J. F. Laulederkind, T. F. Lowry, R. Nigam, S.-J. Wang, M. Shimoyama, M. R. Dwinell, D. H. Munzenmaier, E. A. Worthey and H. J. Jacob (2014). " The pathway ontology – updates and applications." *Journal of Biomedical Semantics* 5(1): 7. IF: 1.600
8. Li, L., E. Chen, C. Yang, J. Zhu, **P. Jayaraman**, J. De Pons, C. C. Kaczorowski, H. J. Jacob, A. S. Greene, M. R. Hodges, A. W. Cowley, Jr., M. Liang, H. Xu, P. Liu and Y. Lu (2015). " Improved rat genome gene prediction by integration of ESTs with RNA-Seq information." *Bioinformatics* (Oxford, England) 31(1): 25-32. IF: 7.33
9. Guan, Q., J. Balciuniene, K. Cao, Z. Fan, S. Biswas, A. Wilkens, D. J. Gallo, E. Bedoukian, J. Tarpinian, **P. Jayaraman**, M. Sarmady, M. Dulik, A. Santani, N. Spinner, A. N. Abou Tayoun, I. D. Krantz, L. K. Conlin and M. Luo (2018). " AUDIOME: a tiered exome sequencing–based comprehensive gene panel for the diagnosis of heterogeneous nonsyndromic sensorineural hearing loss." *Genetics In Medicine*. IF:9.937
10. Wu, C., B. Devkota, P. Evans, X. Zhao, S. W. Baker, R. Niazi, K. Cao, M. A. Gonzalez, **P. Jayaraman**, L. K. Conlin, B. L. Krock, M. A. Deardorff, N. B. Spinner, I. D. Krantz, A. B. Santani, A. N. A. Tayoun and M. Sarmady (2019). " Rapid and accurate interpretation of clinical exomes using Phenoxome: a computational phenotype-driven approach." *European Journal of Human Genetics* 27(4): 612-620. IF: 4.349
11. **Pushkala Jayaraman**, Timothy Mosbruger, Taishan Hu, Nikolaos G Tairis, Chao Wu, Peter M Clark, Monica D'Arcy, Deborah Ferriola, Katarzyna Mackiewicz, Xiaowu Gai, Dimitrios Monos, Mahdi Sarmady, AnthOligo: automating the design of oligonucleotides for capture/enrichment technologies, *Bioinformatics*, Volume 36, Issue 15, 1 August 2020, Pages 4353–4356. IF: 6.937
12. Oh W, **Jayaraman P**, Sawant AS, Chan L, Levin MA, Charney AW, Kovatch P, Glicksberg BS, Nadkarni GN: Using sequence clustering to identify clinically relevant subphenotypes in patients with COVID-19 admitted to the intensive care unit. *Journal of the American Medical Informatics Association* 2021. IF:5.178
13. Paranjpe, I.*, **P. Jayaraman***, C.-Y. Su, S. Zhou, S. Chen, R. Thompson, D. M. Del Valle, E. Kenigsberg, S. Zhao, S. Jaladanki, K. Chaudhary, S. Ascolillo, A. Vaid, A. Kumar, M. Paranjpe, R. O'Hagan, S. Kamat, F. F. Gulamali, H. Xie, J. Harris, M. Patel, K. Argueta, C. Batchelor, K. Nie, S. Dellepiane, L. Scott, M. A. Levin, J. C. He, S. G. Coca, L. Chan, E. U. Azeloglu, E. Schadt, N. Beckmann, S. Gnjatic, M. Merad, S. Kim-Schulze, B. Richards, B. S. Glicksberg, A. W. Charney and G. N. Nadkarni. "Proteomic Characterization of Acute Kidney Injury in Patients Hospitalized with SARS-CoV2 Infection." *Nature Communications* (In Review)

* co-first authors

Pre-publication and Conference Proceedings

1. Hu, T., Tairis, N.G., Mosbruger, T., **Jayaraman, P**. and Monos, D.S., 2019. P053 Optimizing the targeting and sequencing of MHC class II region. *Human Immunology*, *80*, p.94.
2. Fan, J., S. Mulchandani, M. Dulik, J. Chen, A. Gleason, P. Jayaraman, M. Sarmady, E. Zackai, M. Luo, N. Spinner and L. Conlin (2016). "Assessing Copy Number Variants Involving ACMGG Secondary Finding Genes Identified by Routine Chromosomal SNP Array in a Clinical Pediatric Population." Cancer Genetics 209(5): 230. IF:2.351

VOLUNTARY PRESENTATIONS

- *Poster*: Whole genome annotation of bacterial genomes - a genome sequencing and computational genomics pipeline.Georgia Tech - Oak Ridge National Laboratory International Conference, Atlanta, GA, Nov 2009.
- *Poster*: The Rat Genome Database – Genome BrowserAnnual International conference on Intelligent Systems for Molecular Biology (ISMB), Boston, MA, 2010.

- *Poster*: The molecular pathway curation tool: the latest addition to the Rat Genome Database curation tool suite, enabling efficient assembly, editing and display of pathway descriptions, diagrams and associated biological data. Annual Georgia Tech, Emory and the Oak Ridge National Laboratory International Conference, Atlanta, GA, Nov 2011.
- *Poster*: Automated design of hybridization oligonucleotides for region- specific extraction of large contiguous DNA fragmentsAnnual meeting of the American Society of Human Genetics (ASHG), Baltimore, MA, Oct. 2015.
- *Poster*: Informatics and Infrastructure Support for Genomic Diagnostics in a Clinical Setting.Annual meeting Mid Atlantic Bioinformatics Conference, Philadelphia, PA, Oct. 2016.
- *Poster*: A Bioinformatics Pipeline for Clinical Mitogenome DiagnosticsAnnual meeting Mid Atlantic Bioinformatics Conference, Philadelphia, PA, Oct. 2018.

RESEARCH PROFILE (expanded)

Graduate Ph.D. Candidate 2021 – present

Augmented Intelligence in Medicine and Science, Division of Data Driven and Digital Medicine (D3M), Icahn School of Medicine, Mt. Sinai, New York, NY

Mentor: Prof. Girish Nadkarni, MD, MPH
Fishberg Professor of Medicine, the inaugural chief of the Division of Data Driven and Digital Medicine (D3M) and the co-director of the Mount Sinai Clinical Intelligence Center

Omics Research:
- Proteomic Characterization of AKI in COVID-19 patients in the ICU
- Transcriptomic analysis of AKI in COVID-19 patients in critical care

Machine Learning:
- COVID-19 sub-phenotypes from longitudinal information in Electronic Health Records (EHR) for ICU patients
- Learning effective treatment policies for COVID-19 patients in a critical care seting

Bioinformatics Scientist III 2014 – 2021

Department of Pathology and Laboratory Medicine, Children's Hospital of Philadelphia, Philadelphia, PA

Bioinformatics Scientist II Jan 2014 - Aug 2019

Department of Pathology and Laboratory Medicine, Children's Hospital of Philadelphia, Philadelphia, PA

Director: Dr. Mahdi Sarmady, Ph.D.

Director of Bioinformatics, Division of Genomic Diagnostics, Children's Hospital of Philadelphia Assistant Professor of Clinical Pathology And Laboratory Medicine, Perelman School of Medicine

Clinical Genomic Diagnostics Test Development:

– R&D and clinical implementation of a high-throughput targeted Mitochondrial genome NGS diagnostic test. Clinical pipeline detects both single nucleotide variants and large copy number variants and runs on Cromwell workflow engine designed in WDL.

– Implementation, upgrade, and execution of the Clinical Whole Exome Sequencing (CWES), Constitutional and Cancer pipelines for more than 300 tests.

– Implemented data-integrity check for sample gender match, sample identity concordance and coverage thresholds throughout the NGS pipelines for QC (quality control).

– Devised a secure protocol to scale up raw data release to share of patient raw BAM (Binary Sequence Alignment File) and VCF (Variant Call Format) files to internal and external requesters when an official request is made.

Bioinformatics Tool Development:

– Developed a critical in-house web application to fingerprint patient DNA samples via Single Nucleotide Polymorphism (SNP) genotyping concordance and ensure sample integrity for clinical diagnostics.

– Developed and maintained a complex oligo-design application in collaboration with the Clinical Immunogenetics Group at Division of Genomic Diagnostics (DGD) for use with a patented HLA typing method used in RSE (Region Specific Extraction). Publicly available at http://antholigo.chop.edu.

– Upgrade of existing HLA (Human Leukocyte Antigen) disease associations and analysis software SKDM (https://github.com/chopdgd/CHOP SKDM) including automation of the HLA sequence alignments download from IMGT database.

– Developed and maintained an in-house clinical web-based application (available within CHOP at http://dgdgenecoverage.chop.edu) for the analysis of coverage information for WES (Whole Exome Sequencing) data from patient samples.

– Developed and maintained a custom primer design application for Sanger Confirmation of clinical NGS analysis variants for Constitutional and Cancer NGS tests.

– Developed and maintained clinical CNV calling pipeline for SNPArray analysis. A python-based pipeline with a *PennCNV* core to call CNVs on SNP Array data, annotate and create a filtered set of clinically significant CNVs.

Programmer/Analyst II Jan 2010 - Dec 2013

Rat Genome Database (RGD), Human Molecular Genetics Center(HMGC), Medical College of Wisconsin, Milwaukee, WI

Manager Bioinformatics: Jeff De Pons

Director: Dr. Mary Shimoyama, Ph.D.

Associate Professor, Department of Biomedical Engineering, Marquette University and Medical College of Wisconsin

Rat Genomic Data Warehouse for Cross-Platform Sharing

– Developer for an open-source application for mining data from Rattus Norvegicus MOD (Model Organism Database) – Ratmine built on the Intermine platform. Responsible for storing and maintenance of data, building templates for complex queries and data analysis widgets thus allowing automated cross species data sharing)

Applications for Rat-Mouse-Human Cross-functional research

– Developed a simple Primer Design Application for RGD to confirm known and novel variants in the clinical variant analysis program.

– Expanded the in-house curation platform to aid curators in creating and editing pathway information for the Rat Genome Database Pathway Suite. Automated creation of new pathway reports and link associated diseases and references to the report.

– As RGD GBrowse (https://github.com/GMOD/GBrowse) administrator, was responsible for periodic update and maintenance of more than 30 annotation databases. Also customized plugins that communicated with NCBI BLAST tool and UCSC BLAT application.

– Developed a web application to help the end user load and edit SSLP (Simple Sequence Length Polymorphisms) markers into the database directly from a user interface.

RGD Data Import Pipelines

– Developed an automated pipeline to format and import pathway ontology annotations from the PID (Pathway Interaction Database) at NCI (National Cancer Institute) and from KEGG (Kyoto Encyclopedia of Genes and Genomes) into the RGD Pathway Ontology after manual curation.

– Created and maintained an automated reference data pipeline that imports abstract and DOI data from PubMed into RGD and then creates the necessary associations between other data types within the RGD schema.

– Created automated data generation pipelines for more than 30 visualization tracks on Rat, Mouse and Human GBrowse for data transfer into the Rat genome browser backend database. Also created data import and export pipelines for visualization of external collaborator data in GBrowse.

Graduate Research Assistant Jan 2009 - Dec 2009

Computational Genomics group, School of Biology, Georgia Institute of Technology, Atlanta, GA

Mentor: Dr. I. King Jordan, Associate Professor & Director of the Bioinformatics Graduate Program at Georgia Tech

Computational genomics pipeline for prokaryotic sequencing projects

– Automated and improved the process of computational functional annotation for N. meningitidis genome

– Added QC checks in the bulk data transfer between each segment of the pipeline - genome assembly, gene prediction, functional annotation, and comparative genomics.

Pattern Recognition and Machine Learning

– Improvise Netflix Recommendation Systems scores by updating the Netflix scoring algorithm

Bioinformatics Summer Intern May 2009 - Aug 2009

Department of Human Genetics, Emory University, Atlanta, GA

Mentor: Dr. Madhuri Hegde, Professor & Executive Director Emory Genetics Laboratory

Probe design

– Design of microarray chips and analysis of SNPs for Congenital Muscular Dystrophy (CMD) and Deuchene Muscular Dystrophy (DMD)

– Created a program to analyze genotyped and sequenced microarray SNPs in a custom designed Affymetrix chip

TECHNICAL SKILLS

- **Programming language Platforms/Software Technologies:** R, Python Flask, Django, Java (Spring and J2EE Web application design using Model-View-Controller architecture), Perl, Perl CGI, Python, HTML, JavaScript, Bootstrap, HTML5, Web Application development using JSP2.0, WDL
- **Database systems:** MYSQL, ORACLE SQL, Postgres SQL, Seqfeature::Store and GFF schema
- **Machine Learning:** Regression models: Linear and Logistic regression, SVM, Classification models: Clustering (Supervised and Unsupervised), Tree-based (Random Forest, Decision Trees, XGBoost), Deep learning (RNN, CNN), Reinforcement Learning, Time-series analysis
- **Software Packages suites:** Hadoop, Intermine Data mining tool, Gbrowse GMOD tools, EMBL EBI tools, NCBI tools, VCFTools, BCFTools, SAMTools, IGV, Matlab, R statistical package, Docker, SAMtools, BEDTools
- **File formats**: GFF (General Feature Format), GTF, GFF3, VCF (Variant Call format), UCSC BED, WIG (wiggle density), SAM (Sequence Alignment/Map), BAM, FASTA, FastQ, Illumina CASAVA output
- **Open-Source Applications:** Variant Calling tools (GATK, FreeBayes, Varscan, Mutect, Pindel, Delly), Genome alignment tools (Novoalign, BWA, GMAP-GSNAP), Picard, MACS, BLAST, NIH DAVID (functional analysis), IPA, GSEA, Cytoscape, PLINK, RNASeq tools (Bowtie, Tophat, Cufflinks), FastQC
- **Operating Systems**: Linux, Unix, MacOS, Windows
- **Version Control:** Git, Subversion
- **Pipeline/Job management systems:** HTC Condor, Pegasus Job scheduler, Cromwell workflow engine, Cromwell WDL workflow language

Template Updated June 2022