Ayush Raman, Ph.D.   Curriculum vitae, July 2024

Contact Information

National Cancer Institute, NIH
41 Medlars Dr, NIH Building 41
Room B620, Bethesda, MD 20894
ORCID: 0000-0003-2872-2307
Linkedin:/in/aayushraman; Twitter: aayushraman
Website: araman.rbind.io

(412) 614-9490
ayush.raman@nih.gov
aayushraman09@gmail.com

Bio

I am a Staff Scientist at the National Cancer Institute (NCI), National Institutes of Health (NIH), Bethesda, MD. I am interested in Bioinformatics and computational analysis of bulk and single-cell (epi)genomics datasets using machine learning in development and diseases, including cancer and autism.

Education and Training

Broad Institute of MIT and Harvard, Cambridge, MA

Postdoctoral Associate, Epigenomics Program   Dec 2018 - April 2024

- Advisors: Martin Aryee, Ph.D. and Alexander Meissner, Ph.D.
- Research Summary: Understanding the dynamics of gene regulation using single-cell genomics and long read sequencing

Baylor College of Medicine, Houston, TX

Ph.D., Quantitative and Computational Biosciences   Sept 2013 - Sept 2018

- Dissertation: Decoding Big Genomic Datasets: Biases, hypotheses & regulation
- Advisors: Zhandong Liu, Ph.D. and Kunal Rai, Ph.D.
- Advisory Committee: Zhandong Liu, Ph.D. (Chair), Huda Y. Zoghbi, M.D., Aleksandar Milosavljevic, Ph.D., Olivier Lichtarge, M.D., Ph.D., Mirjana Maletić-Savatić, M.D., Ph.D., Kunal Rai, Ph.D.

Carnegie Mellon University, Pittsburgh, PA

M.S., Computational Biology   Aug 2007 - Dec 2008

- Academic Achievement Fellowship (2007 - 2009)

Vellore Institute of Technology, Vellore, India

B. Tech, Bioinformatics   Aug 2002 - May 2006

- Dissertation: Host protein mimicry of *M. tuberculosis* proteins in *H. sapiens*
- Distinction with highest honors (equivalent to *summa cum laude*)

SELECTED PUBLICATIONS (* CO-FIRST AUTHORS)

1. E Orouji*, **AT Raman***, AK Singh*, ..., S Kopetz, K Rai. Chromatin state dynamics confers specific therapeutic strategies in enhancer subtypes of colorectal cancer. ***Gut*** (2022) (PMID: 34059508; DOI)

2. H Gu*, **AT Raman***, X Wang, AW Mohammad, A Arczewska, ZD Smith, DA Landau, MJ Aryee, A Meissner, A Gnirke. Smart-RRBS for single-cell methylome and transcriptome analysis. ***Nature Protocols*** (2021) (PMID: 34244697; DOI)

3. V Kochat*, **AT Raman***, ..., K Rai, KE Torres. Enhancer reprogramming in PRC2-deficient malignant peripheral nerve sheath tumors induces a targetable de-differentiated state. ***Acta Neuropathologica*** (2021) (PMID: 34283254; DOI)

4. **AT Raman***, AE Pohodich*, YW Wan, HK Yalamanchili, HY Zoghbi, Z Liu. Apparent bias towards long gene misregulation in MeCP2 syndromes disappears after controlling for baseline variations. ***Nature Communications*** (2018) (PMID: 30104565; DOI)

5. H Yi*, **AT Raman***, H Zhang, G Allen and Z Liu. Detecting hidden batch factors through data-adaptive adjustment for biological effects. ***Bioinformatics*** (2018) (PMID: 29617963; DOI)

HONORS AND DISTINCTIONS

- Hechter Memorial Travel Award, Broad Institute (2020)
- Junior Research Parasite Award for Rigorous Secondary Data Analysis (2020)
- MCBIOS Young Scientist Excellence Postdoc Award (Third place; 2020)
- Gigascience Travel Award for Pacific Symposium on Biocomputing (2020)
- GSBS Scholarship, Baylor College of Medicine (2013-2014)
- Academic Achievement Fellowship, Carnegie Mellon University (2007-2009)
- Undergraduate thesis awarded highest grade, Vellore Institute of Technology (2006)
- Distinction with highest honors, Vellore Institute of Technology (2002-2006)
- Awarded Distinction in Mathematics Delhi Olympiad (2001)
- Awarded Distinction in New South Wales Mathematics Olympiad (1995-1998)

SKILL-SET SUMMARY

- **Next Gen-Sequencing Data Analysis**:
  - Proficiency in RNA-seq, ChIP-seq, ATAC-seq, Microarray, Detection and correction of Batch Effects, multi-omic analysis and data integration
  - Intermediate in scRNA-seq & scATAC-seq, Long-read seq, NanoString
- **Programming Languages**:
  - Proficiency in R, Perl
  - Intermediate in Python, MATLAB, Shell Scripting, WDL
  - Familiarity in Octave, C/C++, Java
- **High Performance Computing**: Load Sharing Facility (LSF), Condor
- **Cloud Computing**: Google Cloud Computing, Terra
- **Operating Systems**: Linux, Windows, Mac OS X, Unix
- **Machine Learning Algorithms**: Clustering (k-means, hierarchical clustering, NMF, semi-NMF), Dimensionality reduction methods (PCA, ICA, t-SNE, UMAP, MDS), Classification Algorithms (Naive Bayes, Linear Discriminant Analysis, Decision Trees, Random Forest, SVM, Logistic Regression), Regression (Linear, GLM), Regularization (LASSO, Ridge, Elastic-net)

ALL PUBLICATIONS (* CO-FIRST AUTHORS)

1. H Guo*, J Vuille,*, B Wittner, E Lachtara, Y Hou, M Lin, T Zhao, **AT Raman**, ..., DA Haber. DNA hypomethylation silences antitumor immune genes in early prostate cancer and CTCs. ***Cell*** (2023) (PMID: 37327786; DOI)

2. CB Nava Lauson, S Tiberti, PA Corsetto, F Conte, P Tyagi, M Machwirth, S Ebert, A Loffreda, L Scheller, D Sheta, Z Mokhtari, T Peters, **AT Raman**, ..., Luigi Nezi, Teresa Manzo. Linoleic acid potentiates CD8+ T cell metabolic fitness and antitumor immunity. ***Cell Metabolism*** (2023) (PMID: 36898381; DOI)

3. E Orouji, **AT Raman**. Computational methods to explore chromatin state dynamics. ***Briefings in Bioinformatics*** (2022) (PMID: 36208178; DOI)

4. SC Callahan, V Kochat, Z Liu, **AT Raman**, J Schulz, C Terranova, M Divenko, A Ghosh, M Tang, C Pickering, J Myers, K Rai. High enhancer activity is an epigenetic feature of HPV negative atypical head and neck squamous cell carcinoma. ***Frontiers in Cell and Developmental Biology*** (2022) (PMID: 35927986; DOI)

5. E Orouji*, **AT Raman***, AK Singh*, ..., S Kopetz, K Rai. Chromatin state dynamics confers specific therapeutic strategies in enhancer subtypes of colorectal cancer. ***Gut*** (2022) (PMID: 34059508; DOI)

   - New Epigenetic Target May Lead to Personalized Treatments

6. CJ Terranova*, KM Stemler*, P Barrodia, SL Jeter-Jones, Z Ge, M de la Cruz Bonilla, **A Raman**, KL Allton, E Arslan, OH Yilmaz, MC Barton, K Rai, H Piwnica-Worms. Reprogramming of H3K9bhb at regulatory elements is a key feature of fasting in the small intestine. ***Cell Reports*** (2021) (PMID: 34818540; DOI)

7. V Kochat*, **AT Raman***, ..., K Rai, KE Torres. Enhancer reprogramming in PRC2-deficient malignant peripheral nerve sheath tumors induces a targetable de-differentiated state. ***Acta Neuropathologica*** (2021) (PMID: 34283254; DOI)

8. S Shareef, S Bevill, **AT Raman**, M Aryee, P Galen, V Hovestadt, B Bernstein. Extended-representation bisulfite sequencing of gene regulatory elements in multiplexed samples and single cells. ***Nature Biotech*** (2021) (PMID: 33958785; DOI)

9. H Gu*, **AT Raman***, X Wang, AW Mohammad, A Arczewska, ZD Smith, DA Landau, MJ Aryee, A Meissner, A Gnirke. Smart-RRBS for single-cell methylome and transcriptome analysis. ***Nature Protocols*** (2021) (PMID: 34244697; DOI)

10. C Terranova*, M Tang*, M Maitituoheti*, **A Raman**, ..., K Rai. Reprogramming of bivalent chromatin states in NRAS mutant melanoma suggests PRC2 inhibition as a therapeutic strategy. ***Cell Reports*** (2021) (PMID: 34289358; DOI)

11. **AT Raman**. A research parasite's perspective on establishing a baseline to avoid errors in secondary analyses. ***GigaScience*** (2021) (PMID: 33710326; DOI)

12. M Maitituoheti, EZ Keung, M Tang, L Yan, H Alam, G Han, **AT Raman**, ..., K Rai. Enhancer reprogramming confers dependence on glycolysis and IGF signaling in KMT2D mutant melanoma. ***Cell Reports*** (2020) (PMID: 33086062; DOI)

13. T Manzo, B Prentice, K Anderson, **A Raman**, ..., Giulio Draetta, Luigi Nezi. Accumulation of long-chain fatty acids in the tumor microenvironment drives dysfunction in intrapancreatic CD8+ T cells. ***Journal of Exp. Medicine*** (2020) (PMID: 32491160; DOI)

14. CA Chen, W Wang, SE Pedersen, **A Raman**, ..., HY Zoghbi, CP Schaaf. Nr2f1 heterozygous knockout mice recapitulate neurological phenotypes of Bosch-Boonstra-Schaaf optic atrophy syndrome and show impaired hippocampal synaptic plasticity. ***Hum Mol Genet.*** (2019) (PMID: 31600777; DOI)

15. S Adhikary*, D Chakravarti*, C Terranova*, I Sengupta, M Maitituoheti, A Dasgupta, D Srivastava, J Ma, **A Raman**, ..., K Rai, C Das. Atypical plant homeodomain of UBR7 functions as an H2BK120Ub ligase and breast tumor suppressor. ***Nature Communications*** (2019) (PMID: 30923315; DOI)

16. **AT Raman***, AE Pohodich*, YW Wan, HK Yalamanchili, HY Zoghbi, Z Liu. Apparent bias towards long gene misregulation in MeCP2 syndromes disappears after controlling for baseline variations. ***Nature Communications*** (2018) (PMID: 30104565; DOI)

    - Recommended by the Faculty of 1000
    - Synopsis in Texas Children's Hospital website: Long genes are not preferentially altered in Rett and MeCP2 duplication syndromes

17. C Terranova, M Tang, E Orouji, M Maitituoheti, **A Raman**, S Amin, Z Liu, K Rai. An integrated platform for genome-wide mapping of chromatin states using high-throughput ChIP-sequencing in tumor tissues. ***J Vis Exp.*** (2018) (PMID: 29683440; DOI)

18. H Yi*, **AT Raman***, H Zhang, G Allen and Z Liu. Detecting hidden batch factors through data-adaptive adjustment for biological effects. ***Bioinformatics*** (2018) (PMID: 29617963; DOI)

19. A Pohodich, HK Yalamanchili, **AT Raman**, YW Wan, M Gundry, S Hao, H Jin, J Tang, Z Liu, H Zoghbi. Forniceal deep brain stimulation induces gene expression and splicing changes that promote neurogenesis and plasticity. ***elife*** (2018) (PMID: 29570050; DOI). *Computational and Integrative analysis lead (equal contributor)*

    - Featured in TheScientist. Press release - Deep Brain Stimulation Affects the Activity of Hundreds of Genes

20. **AT Raman**, K Rai. Loss of histone acetylation and H3K4 methylation promotes melanocytic malignant transformation. ***Molecular & Cellular Oncology*** (2018) (PMID: 30250885; DOI)

21. U Oleksiewicz*, M Gladych*, **AT Raman***, ..., Kunal Rai, Maciej Wiznerowicz. TRIM28 and interacting KRAB-ZNFs control self-renewal of human pluripotent stem cells through epigenetic repression of pro-differentiation genes. ***Stem Cell Reports*** (2017) (PMID: 29198826; DOI)

22. The Cancer Genome Atlas Network. Genomic classification of cutaneous melanoma. ***Cell*** (2015) (PMID: 26091043; DOI)

23. V Litvak, A Ratushny, A Lampano, F Schmitz, A Huang, **A Raman**, A. Rust, A. Bergthaler, J. Aitchison, A. Aderem. A FOXO3-IRF7 gene regulatory circuit limits inflammatory sequelae of antiviral responses. ***Nature*** (2012) (PMID: 22982991; DOI)

- Recommended by the Faculty of 1000

24. S Ramachandran, A Katiyar, A Sinha, A Bharadwaj, A Dutta, **A Raman** *et al*. Mycobacterium tuberculosis systems biology data in R. ***Biobytes, Indian Journal*** (2009) (URL)

PEER-REVIEW & PREPRINTS (* CO-FIRST AUTHORS)

1. CJ Terranova, M Maitituoheti, **AT Raman**, M Tang, ..., K Rai. H3K4me1-marked Enhancer Activation in Resistant Prostate Cancers Implicates SOX4 and MENIN Inhibition as Therapeutic Strategies. ***bioRxiv*** (2021) (DOI)

2. M Maitituoheti*, E Arslan*, **AT Raman***, ..., K Rai. Pan-cancer Chromatin Landscape Reveals Principles of Epigenetic Heterogeneity and Dependencies in Cancer Cells. ***In revision, Nature Cell Biology (NCB-R44336)*** (2023)

RESEARCH EXPERIENCE

Staff Scientist April 2024 - Present
Oberdoerffer and Larson Labs, National Cancer Institute, NIH Bethesda, MD

- Elucidate mRNA splicing patterns and chemical modifications of nucleotides using high throughput datasets

Postdoctoral Associate Dec 2018 - April 2024
Aryee and Meissner Labs, Broad Institute of MIT and Harvard Cambridge, MA

- Understanding the dynamics of gene regulation using single-cell genomics (Published in *Nature Protocol* and *Nature Biotechnology*) and long read sequencing
- Understanding the role of DNA methylation and its writers in diseases in development and aging
- Elucidating regulatory mechanisms of tumor suppressor genes in senescence (Calico)

Graduate Student / Research Assistant Sept 2013 - Sept 2018
Liu's Lab, NRI, Baylor College of Medicine Houston, TX
Rai Lab, Genome Medicine, MD Anderson Cancer Center

- Amplification-based expression data show bias towards long gene misregulation in synaptic disorders (Published in *Nature Comm.*)
- Developed batch detection algorithm in gene expression datasets (Published in *Bioinformatics*)
- Mechanism behind the effects of deep brain stimulation in Rett syndrome, meta-analysis of activity dependent genes and intellectual disorders (Published in *elife* and *Hum Mol Genet.*)
- Epigenomic landscape in colorectal tumors (Published in *Gut*)
- Oncogenic Role of EZH2 in the pathogenesis of MPNST tumors (Published in *Acta Neuropathologica*)
- Role for the TRIM28/KRAB repressors during early development, with a particular emphasis in nuclear reprogramming (Published in *Stem Cell Reports*)
- Lymphocytic Infiltration as a basis of response in melanoma tumors (Published in *Cell*)
- Multi-omics integrative analysis of ChIP-seq and RNA-seq datasets
- Genome wide analysis of non-coding regulatory mutations in cancer

- Lab rotation between Sept 2013 - May 2014 before accepting thesis lab –
  - Han Liang's Lab: Analyzed RNA-seq Data for various TCGA Datasets for the potential unique transcription factors per subtype
  - Lynda Chin's Lab: Analyzed known mutations in leukemia dataset for Apollo Project and the reported somatic mutations in COSMIC
  - Zhandong Liu's Lab: Correlation network for 90,000 samples from Affymetrix Human Genome U133 Plus 2.0 Array

Bioinformatics Scientist   July 2010 - July 2013
Shmulevich Lab, Institute for Systems Biology   Seattle, WA

- Rhinovirus vs Influenza Virus Project: Analyzing the similarities and differences between the two viruses using microarray datasets (Proctor & Gamble Project)
- Tol-II Receptors Project: Analysis of Pol-II, Transcription Factor, HDAc ChIP-seq datasets, Integrative analysis with gene Expression data for the discovery of gene regulatory network that gets activated in the presence of Tol-II receptors (Published in *Nature*)
- Poised Genes Project: Predicting poised genes in the macrophages using machine learning algorithms
- In-House ChIP-seq Pipeline: Development and integration of various functionalities within the pipeline

Statistician   Nov 2009 - July 2010
Computational Genetics Lab, WPIC-UPMC   Pittsburgh, PA

- Palau Project: Analyzed GWAS datasets for the detection of causal genes in psychotic disorders and schizophrenia and
- Implemented genomic distance based multivariate regression model for the estimation of Identity by Descent (IBD) in the haplotype datasets
- Worked on the detection of IBD for the use of the population based linkage analysis among the case and control pairs

Research Assistant Programmer   Apr 2009 - Oct 2009
Department of Biomedical Informatics, University of Pittsburgh   Pittsburgh, PA

- Feature/covariates selection from comparative and high-throughput data for the prediction of protein-protein interaction pairs
- Computational prediction and accuracy of different statistical based algorithms such as discriminative and generative algorithms, and clustering based algorithms

Research Assistant   Sept 2007 to Dec 2007
Murphy Lab, Carnegie Mellon University   Pittsburgh, PA

- Development of database using PostgreSQL and JSP as back-end and front-end respectively

Research Trainee   Dec 2005 to Nov 2006
Institute of Genomics and Integrative Biology   New Delhi, India
Advisers: Dr. S. Ramachandran and Dr. Vani Brahmachari

- B.Tech Thesis: Study of Host Protein Mimicry of M. tuberculosis proteins in *H.*

sapiens

- Comparative genomic analysis of *M. tuberculosis* H37Rv with *H. sapiens* using sequence and structural based approaches

<table>
<tr><td>

INVITED PRESENTATIONS AND SEMINARS

</td><td>

1. Amplification-based expression data show bias towards long gene misregulation in synaptic disorders in MeCP2 syndrome. Virtual/oral presentation for the Young Scientist Award session, *Annual Conference Of The Midsouth Computational Biology & Bioinformatics Society (MCBIOS) and MAQC Annual Meeting* (2020)

2. Decoding Big Genomic Datasets: Biases, hypotheses & regulation at *IGIB, New Delhi.* Host: Debasis Dash, Ph.D. and Anurag Agrawal, M.D., Ph.D. (2019)

3. Decoding Big Genomic Datasets: Biases, hypotheses & regulation at *Harvard Medical School, Boston.* Host: Steven McCarroll, Ph.D. (2018)

4. Deciphering epigenomic landscape in colorectal cancer at *Genome Medicine Science Day, MD Anderson Cancer Center*, Houston, TX (2018)

5. Decoding Big Genomic Datasets: Biases, hypotheses & regulation at *Broad Institute of MIT and Harvard, Cambridge.* Host: Evan Macosko, M.D., Ph.D. and Aviv Regev, Ph.D. (2018)

6. Decoding Big Genomic Datasets: Biases, hypotheses & regulation at *Broad Institute of MIT and Harvard, Cambridge.* Host: Martin Aryee, Ph.D., Alexander Meissner, Ph.D. and Andreas Gnirke, Ph.D. (2018)

7. Decoding Big Genomic Datasets: Biases, hypotheses & regulation at *University of Toronto, Toronto.* Host: Mathieu Lupien, Ph.D. (2018)

8. Decoding Big Genomic Datasets: Biases, hypotheses & regulation at *University of California, San Diego.* Host: Bing Ren, Ph.D. (2018)

9. Decoding Big Genomic Datasets: Biases, hypotheses & regulation at *University of California, Los Angeles.* Host: Daniel Geschwind, M.D., Ph.D. (2018)

10. ChIP-seq Analysis: Pipelines and Data Interpretations at *ISB Nucleic Acid Course*, Seattle, WA (2013)

11. Identification of FOXO3/IRF7 circuits that limits inflammatory sequel of antiviral response at *ISB Retreat*, Seattle, WA (2012)

</td></tr>
<tr><td>

TEACHING EXPERIENCE

</td><td>

- Teaching Assistant for Bioinformatics Workshop, MD Anderson Cancer Center (Dec. 2017)
    - Supervised by Kunal Rai, Ph.D.
- Teaching Assistant for Computational Math for Quantitative Biomedicine, BCM
    - Supervised by Zhandong Liu, Ph.D. and Aleksandar Milosavljevic, Ph.D.
    - Evaluation from the students: **6/7** (Fall 2017)
- Taught a course on ChIP-seq analysis at ISB Nucleic Acid Course (2013)

</td></tr>
</table>

ACADEMIC SERVICES AND MEMBERSHIPS	• **Peer Review (*ad hoc*)**: Nature Communications, PLOS Computational Biology, IUBMB Life, Stem Cell Research & Therapy, Frontiers in Genetics, Frontiers in Immunology, Frontiers in Molecular Biosciences
	• **Guest Editor**: Frontiers in Pharmacology, Frontiers in Immunology

	• **Professional Societies and Memberships**:
	  – Core RSG India Member, ISCB (2016-2017)
	  – Contributors of SysBorgTB & Open Source Drug Discovery by CSIR, India

	• **Wikipedia articles**: Tumor Metabolism for Cancer Biology course (Spring 2014)

RELEVANT COURSES	• **Massive Open Online Courses**: Machine Learning (Coursera), RNA-seq data analysis (edX)
	• **Baylor College of Medicine** (Ph.D. course work): Computational Mathematics for Quantitative Biomedicine, Methods and Logic in Molecular Biology
	• **Rice University** (Ph.D. course work): Molecular Biophysics, Statistical Computing & Graphics, Statistical Machine Learning, Cancer Biology, Probability & Statistics
	• **Carnegie Mellon University** (M.S. course work): Data Mining, Applied Machine Learning, Computational Biology, Computational Methods for Biological Modeling & Simulation, Mathematical Models, Advanced Molecular & Cell Biology, Data Structure & Algorithms
	• **Vellore Institute of Technology** (B.Tech course work): Biochemistry, Cell Biology, Immunology, Probability and Biostatistics, Discrete Math, Genomics, Proteomics, Systems Biology, Programming Languages, Databases, Machine Learning, Structure and Molecular Modeling, Molecular Evolution and Phylogeny, Artificial Intelligence, Data Mining and Warehousing

COURSE PROJECTS	• Loan Default Prediction - Imperial College London (Kaggle dataset). A. Raman and M. Chen. (Comp 540, Spring 2014; Prof. Devika Subramanian)
	• Analysis and Forecast of Start-Up Companies. A. Raman, Y. Miao and G. Breternitz (Stat 405, Fall 2013; Dr. Roberto Bertolusso) - Our project was used as a reference for Spring/Fall 2014-2015 class
	• CMBMS Final Project: B-cell HIV Epitope Predictor (B-HIVE P). S Garimalla, A Raman, V Pillalamarri (03-712 Course Project, Fall 2008; Prof. Russell Schwartz, CMU)
	• Developed ORF and Gene Prediction Tool in prokaryotes in Perl and HTML (Genomics & Computational Molecular Biology Course Project, Spring 2005; Prof. Ghosh, VIT) – Awarded as the best class project

REFERENCES	Zhandong Liu, Ph.D. Associate Professor, Department of Pediatrics Chief of Computational Sciences, Texas Children's Hospital Jan and Dan Duncan Neurological Research Institute Baylor College of Medicine, Houston, TX 77030 Huda Zoghbi, M.D. Professor, Depts of Neuroscience, Molecular and Human Genetics, Depts of Pediatrics and Neurology

Investigator, Howard Hughes Medical Institute
Jan and Dan Duncan Neurological Research Institute
Baylor College of Medicine, Houston, TX 77030

Kunal Rai, Ph.D.
Associate Professor, Dept. of Genome Medicine
Scientific Director MDACC Epigenomics Therapy Initiative (METI)
MD Anderson Cancer Center, Houston, TX 77054

Martin Aryee, Ph.D.
Associate Professor, Dept of Data Science, Dana Farber Cancer Institute
Associate Professor, Dept of Biostatistics, Harvard TH Chan School of Public Health
Merkin Institute Fellow at the Broad Institute of MIT and Harvard
Institute Member, Broad Institute of Harvard and MIT
Boston, MA 02215

Alexander Meissner, Ph.D.
Director and Scientific member, Department of Genome Regulation,
Max Planck Institute for Molecular Genetics, Berlin, Germany
Visiting Scientist and Professor, Department of Stem Cell and Regenerative Biology,
Visiting Scientist and Professor, Harvard University and Broad Institute
Berlin, Germany 14195

Additional references available upon request.