Anupama Jha

Curriculum Vitae

Personal Information

Primary Email	anupamaj@uw.edu
Secondary Email	anupamajha48@gmail.com
Website	www.anupamajha.com
Primary Address	Department of Genome Sciences, University of Washington, Seattle

---

Education

2014-20	**Ph.D.** in Computer and Information Science, *University of Pennsylvania*, Philadelphia, Pennsylvania, GPA: 3.87/4.0
2012-14	**M.S.** in Informatics, *Technical University of Munich*, Munich, Bavaria, Germany, Grade: 1.5, Passed with Distinction
2006-10	**B.Tech.** in Information Technology, *Guru Gobind Singh Indraprastha University*, Delhi, India, CPI: 85.25%, First Class with Distinction

---

Research Experience

Mar 2021-Present	Postdoctoral Scholar in Genome Sciences, University of Washington, Seattle, Noble Lab (Advisor: Dr. William Stafford Noble)
Jan 2021-Feb 2021	Postdoctoral Scholar in Genetics, University of Pennsylvania, Philadelphia, BioCiphers Lab, (Advisor: Dr. Yoseph Barash)
Aug 2014-Dec 2020	Graduate Student in Computer and Information Science, University of Pennsylvania, Philadelphia, BioCiphers Lab, (Advisor: Dr. Yoseph Barash)
Nov 2013-Aug 2014	Master Thesis Student in Informatics, Technical University of Munich, Munich, Germany, Model-Based Systems and Qualitative Reasoning Group, (Advisor: Dr. Peter Stuss)
Nov 2013-Mar 2014	Research Assistant in Informatics, Technical University of Munich, Munich, Germany, Human Brain Project, (Advisor: Dr. Florian Röhrbein)
Mar 2013-Sep 2013	Research Assistant in Informatics, Technical University of Munich, Munich, Germany, Rost Lab, (Advisor: Dr. Burkhard Rost)

Professional Experience

Jun 2010-Sep 2012	Systems Engineer, Infosys Ltd., India and TomTom Int. BV., Netherlands

---

Professional Organizations

2021-Present	Member of the Encyclopedia of DNA Elements consortium (ENCODE)
2021-Present	Member of the DNA Zoo consortium
2021-Present	Member of the 4D Nucleome consortium (4DN)
2014-Present	Member of the International Society for Computational Biology (ISCB)

Research Funding

Aug 2024-July 2029	K99/R00 Pathway to Independence Award from the National Human Genome Research Institute.

Awards & Honors

2021-Present	UW Data Science Fellow at the eScience Institute, University of Washington
Jul 2023	Travel Fellowship, ISMB/ECCB 2023
Sep 2022	Parker Travel Award, Genome Sciences Annual Retreat, University of Washington
Jun 2019	Selected student abstract for oral presentation, SAGES 2019, Philadelphia
Sep 2018	GHC 18 Scholarship, Grace Hopper Celebration 2018
Jul 2017	Travel Fellowship, ISMB/ECCB 2017
Jun 2017	Best Poster Award, SAGES 2017, Philadelphia.
May 2017	Best Poster Award, RNA Biology & Cancer 2017 Symposium, University of Pennsylvania.

Bibliography (* → Equal contribution, Google Scholar)

First Author

1. **Jha**, A., Hristov, B., Wang, X., Wang, S., Greenleaf, W. J., Kundaje, A., Aiden, E. L., Bertero, A. & Noble, W. S. Prediction and functional interpretation of inter-chromosomal genome architecture from DNA sequence with TwinC. *bioRxiv* (2024).
2. **Jha***, A., Bohaczuk*, S. C., Mao, Y., Ranchalis, J., Mallory, B. J., Min, A. T., Hamm, M. O., Swanson, E., Dubocanin, D., Finkbeiner, C., Li, T., Whittington, D., Noble, S, W., Stergachis, A. B. & Vollger*, M. R. DNA-m6A calling and integrated long-read epigenetic and genetic analysis with fibertools. *Genome Research*, gr–279095 (2024).
3. **Jha***, A., Quesnel-Vallières*, M., Wang, D., Thomas-Tikhonenko, A., Lynch, K. W. & Barash, Y. Identifying common transcriptome signatures of cancer by interpreting deep learning models. *Genome Biology* **23,** 117 (2022).

4. **Jha**, A., K Aicher, J., R Gazzara, M., Singh, D. & Barash, Y. Enhanced integrated gradients: improving interpretability of deep learning models using splicing codes as a case study. *Genome biology* **21,** 1–22 (2020).
5. **Jha**, A., Gazzara, M. R. & Barash, Y. Integrative deep models for alternative splicing. *Bioinformatics* **33,** i274–i282 (2017).

Collaborative Author

6. Fang, T., Liu, Y., Woicik, A., Lu, M., **Jha**, A., Wang, X., Li, G., Hristov, B., Liu, Z., Xu, H., Noble, W. S. & Wang. Enhancing Hi-C contact matrices for loop detection with Capricorn: a multiview diffusion model. *Bioinformatics* **40,** i471–i480 (2024).
7. Yang, K., Islas, N., Jewell, S., **Jha**, A., Radens, C. M., Pleiss, J. A., Lynch, K. W., Barash, Y. & Choi, P. S. Machine learning-optimized targeted detection of alternative splicing. *bioRxiv* (2024).
8. Vaquero-Garcia, J., Aicher, J. K., Jewell, S., Gazzara, M. R., Radens, C. M., **Jha**, A., Norton, S. S., Lahens, N. F., Grant, G. R. & Barash, Y. RNA splicing analysis using heterogeneous and large RNA-seq datasets. *Nature Communications* **14,** 1230 (2023).
9. Bone, W. P., Siewert, K. M., **Jha**, A., Klarin, D., Damrauer, S. M., Chang, K.-M., Tsao, P. S., Assimes, T. L., Ritchie, M. D., *et al.* Multi-trait association studies discover pleiotropic loci between Alzheimer's disease and cardiometabolic traits. *Alzheimer's research & therapy* **13,** 1–14 (2021).
10. Ji, X., **Jha**, A., Humenik, J., Ghanem, L. R., Kromer, A., Duncan-Lewis, C., Traxler, E., Weiss, M. J., Barash, Y. & Liebhaber, S. A. RNA-binding proteins PCBP1 and PCBP2 are critical determinants of murine erythropoiesis. *Molecular and Cellular Biology* **41,** e00668–20 (2021).
11. Slaff, B., Radens, C. M., Jewell, P., **Jha**, A., Lahens, N. F., Grant, G. R., Thomas-Tikhonenko, A., Lynch, K. W. & Barash, Y. MOCCASIN: a method for correcting for known and unknown confounders in RNA splicing analysis. *Nature communications* **12,** 3353 (2021).
12. Gazzara, M. R., Mallory, M. J., Roytenberg, R., Lindberg, J. P., **Jha**, A., Lynch, K. W. & Barash, Y. Ancient antagonism between CELF and RBFOX families tunes mRNA splicing outcomes. *Genome research* **27,** 1360–1370 (2017).
13. Chakraborty, P., Taneja, S., Anand, N., **Jha**, A., Malik, D. & Nayar, A. An Optimizing Compiler for Turing Machine Description Language. *IUP Journal of Computer Sciences* **5** (2011).

Submitted

14. Wu, D., **Jha**, A., Jewell, S., Maus, N., Gardner, J. R. & Barash, Y. Generative modeling for RNA splicing code predictions and design. *Submitted* (2024).

Invited Talks ($^+$ → Travel Fellowship)

Aug 2024	**Speaker**, *Department of Genome Sciences Seminar Series*, University of Washington, Seattle. Integrative models of nuclear DNA organization.

Jul 2024	**Section talk**, *ISMB 2024*, Montreal, Canada, *International*. Predicting trans Hi-C contacts from DNA sequence with TwinC.
Jan 2024	**Invited talk**, *Department of Biosciences and Bioengineering*, IIT Bombay. Integrative models of nuclear DNA organization.
Dec 2023	**Speaker**, *4D Nucleome Genomic Organization of Mammalian Embryogensis Center*, University of Washington, Seattle, *local*. Predicting trans Hi-C contacts from DNA sequence with TwinC.
Jul 2023^{+}	**Section talk**, *ISMB/ECCB 2023*, Lyon, France, *International*. Bidirectional translation between Hi-C and ATAC-seq data across species using Chrome-Zoo.
May 2023	**Speaker**, *Combi Seminar Series*, University of Washington, Seattle. Bidirectional translation between Hi-C and ATAC-seq data across species using Chrome-Zoo.
Sep 2022	**Lightening talk**, *Department of Genome Sciences Annual Retreat*, University of Washington, Seattle. Bidirectional translation between Hi-C and ATAC-seq data across species using Chrome-Zoo.
Mar 2022	**Invited talk**, *DNA Zoo Consortium Monthly Meeting*, Virtual. Bidirectional translation between Hi-C and ATAC-seq data across species using Chrome-Zoo.
Jan 2022	**Invited talk**, *ENCODE Consortium Monthly Meeting*, Virtual. Bidirectional translation between Hi-C and ATAC-seq data across species using Chrome-Zoo.
Jul 2021	**Section talk**, *ISMB/ECCB 2021*, Virtual, *International*. Identifying common transcriptome signatures of cancer by interpreting deep learning models. Presented by co-author Mathieu Q.V.
Jul 2020	**Section talk**, *ISMB 2020*, Virtual, *International*. RBP-Pokedex: Prediction of RBP knockdown effect via DNN experiment modeling.
Jul 2019	**Section talk**, *ISMB/ECCB 2019*, Basel, Switzerland, *International*. Improving interpretation of deep learning models: splicing codes as a case study.
Jun 2019	**Invited talk**, *SAGES 2019*, University of Pennsylvania, Philadelphia, *National*. Improving interpretation of deep learning models: splicing codes as a case study.
Oct 2017	**Speaker**, *Penn Research in Machine Learning Scholar's Seminar*, University of Pennsylvania, Philadelphia, *local*. Integrative Deep Models for Alternative Splicing.
Jul 2017^{+}	**Section talk**, *ISMB/ECCB 2017*, Prague, Czech Republic, *International*. Integrative Deep Models for Alternative Splicing.

Posters ($^{+} \rightarrow$ best poster award)

Sep 2024	**Predicting trans Hi-C contacts from DNA sequence with TwinC**. *MLCB 2024*, Seattle.
Dec 2023	**Predicting trans Hi-C contacts from DNA sequence with TwinC**. *4D Nucleome Annual Meeting*, Boston.

Jul 2023	**Chrome-Zoo: cross-species chromatin profile prediction using DNA Zoo data**. *ISMB/ECCB 2023*, Lyon, France.
Sep 2022^{+}	**Chrome-Zoo: cross-species chromatin profile prediction using DNA Zoo data**. *Genome Sciences Annual Retreat*, University of Washington, Seattle.
Jul 2022	**Identifying universal cancer transcriptome patterns by interpreting deep learning models**. *ISMB/ECCB 2022*, Virtual. Presented by co-author Mathieu Q.V.
Jul 2020	**RBP-Pokedex: Prediction of RBP knockdown effect via DNN experiment modeling**. *ISMB 2020*, Virtual.
Jul 2019	**Interpretation of deep learning models in genomics: splicing codes as a case study**. *ISMB/ECCB 2019*, Basel, Switzerland.
Jun 2019	**Interpretation of deep learning models in genomics: splicing codes as a case study**. *SAGES 2019*, Philadelphia, USA.
Jul 2018	**Assembling the Building Blocks for a Unified Splicing Code**. *ISMB 2018*, Chicago, USA.
Jul 2017	**Integrative Deep Models for Alternative Splicing**. *ISMB/ECCB 2017*, Prague, Czech Republic.
Jun 2017^{+}	**Integrative Deep Models for Alternative Splicing**. *SAGES 2017*, Philadelphia, USA.
May 2017^{+}	**Integrative Deep Models for Alternative Splicing**. *RNA Bio & Cancer Symposium 2017*, University of Pennsylvania, Philadelphia, USA.

Media Coverage

Jun 2024	Brotman Baty Institute, UW Genome Sciences Hackathon Leads to New Long-Read Sequencing Technology.

Teaching

University of Washington

2024	Instructor, Introduction To Computational Molecular Biology: Molecular Evolution (GENOME 541)
2021	Teaching Assistant, Software Carpentry Python Workshop, eScience Institute

University of Pennsylvania

2016-2018	Teaching Assistant, Advanced Computational Biology (GCB 537)
2016-2018	Guest Lecturer, Advanced Computational Biology (GCB 537)
2016, 2018	Teaching Assistant, Ph.D. Special Topics: Deep Learning (CIS 800)

Academic Advising

Graduate (Secondary mentor)

2024-Present	Annie Staker, University of Washington, Data Science Master's student, Noble Lab.
2023-Present	Shengqi Hang, University of Washington, CS Ph.D. student, Noble Lab.
2022-Present	Nathanial Islas, University of Pennsylvania, CIS Ph.D. student, Barash Lab.
2019-2021	Di Wu, University of Pennsylvania, CIS Ph.D. student, Barash Lab.
2017-2019	Deependra Singh, University of Pennsylvania, CIS Master's student, Barash Lab.

Undergraduate (Primary mentor)

2023-Present	Catherine L Rasgatis, University of Washington, Noble Lab.
2017-2018	Joseph Kupferberg, University of Pennsylvania, Barash Lab.
2017-2018	Liam Murphy, University of Pennsylvania, Barash Lab.
2017-2018	Nazih Bissat, University of Pennsylvania, Barash Lab.

Academic Service

2023-2024	**Postdoctoral representative, Department of Genome Sciences, University of Washington**. Invited seminar speakers, attended faculty meetings, organized postdoc townhall, bimonthly postdoc lunches and conveyed postdoc concerns to faculty and department chair.
2018-Present	**Journal Reviewer**, *Genome Biology, Nature Communications, PLOS Computational Biology, Briefings in Bioinformatics, Bioinformatics, BMC Bioinformatics, Nature Scientific Reports.*
2021-Present	**Journal co-reviewer with Prof. William Stafford Noble**, *Genome Research, Nature Communications, Nature Genetics.*
2015-2020	**Journal co-reviewer with Prof. Yoseph Barash**, *PLOS Computational Biology, Bioinformatics, Nucleic Acids Research.*
2018-Present	**Conference reviewer**, *MLCB 2024, RECOMB 2022-2024, ISMB 2022-2024, ACM-BCB 2022, TEAMC-2018*
2015-2020	**Conference co-reviewer with Prof. Yoseph Barash**, *NeurIPS 2015-2020, ISMB/ECCB 2016-2019, ICLR 2018-2020, ICML 2019-2020.*

Outreach

Mar 2024	**Organizer and Team Leader, 4D Nucleome Hackathon**, Organized a NHGRI-sponsored hackathon with 50 participants on behalf of the 4D Nucleome consortium. Led a team of 5 students to find tissue-specificity of existing sequence-to-Hi-C models.

Sep 2023	**Team Leader, Genome Sciences Hackathon**, Led a team of 10 students to improve Fibertools m6A caller for ONT long reads Fiber-seq.
Aug 2023	**Guest Lecture: Genome Hackers Python camp for high school students**, Taught Python to high school students who identify as a gender minority.
Sep 2022	**Team Leader, Genome Sciences Hackathon**, Led a team of 8 students to make an m6A caller for Fiber-seq using PacBio long reads. The project lead to Fibertools, a tool used for m6A calling by ENCODE, Human Pangenome Research Consortium (HPRC), All of Us Network, Undiagnosed Diseases Network (UDN), and the GREGoR Consortium.
Mar 2019	**Panelist: Is Graduate School for me?**, participated in a panel at CAPWIC 2019 to discuss entry requirements and graduate school experience for gender minorities.
2018-2019	**Guest Lecture: AI and Computational Biology** with Prof. Yoseph Barash at the WICS High School Day for Girls at the University of Pennsylvania.

Skills

Languages, Systems, and Tools	**Proficient:** Python, Tensorflow, PyTorch, High Performance Computing, UNIX, LaTeX. **Past work experience:** R, Perl, MATLAB, Java, Android, C#, .NET Framework, C++, Oracle, SQL.
Relevant Graduate Courses	Machine Learning, Bayesian Statistics, Mathematical Statistics, Deep Learning, Advanced Computational Biology, RNA World, High-throughput Datasets for Biologists, Interpretation of Deep Learning Models, Adversarial and Secure Machine Learning, Computational Linguistics.

References

William Stafford Noble	wnoble@uw.edu
Yoseph Barash	yosephb@seas.upenn.edu
Andrew Stergachis	absterga@uw.edu
Mathieu Quesnel-Vallières	Mathieu.Quesnel-Vallieres@usherbrooke.ca

Last updated: November 3, 2024