

Naveen Duhan, Ph.D.

 naveen.duhan@outlook.com  naveenduhan.com  github.com/navduhan
 [linkedin.com/in/navduhan](https://www.linkedin.com/in/navduhan)  [ResearchGate](https://www.researchgate.net/profile/Naveen-Duhan)  [Google Scholar](https://scholar.google.com/citations?user=Naveen-Duhan)

Research Profile

I am a computational scientist specializing in the development of scalable algorithms, machine learning models, and high-performance analysis frameworks for complex biological data. My research centers on designing reproducible computational methods for genomics, transcriptomics (encompassing single-cell and spatial resolutions), and multi-omics integration. I place particular emphasis on alignment-free modeling, deep learning architectures (utilizing PyTorch and TensorFlow), and the optimization of workflows for high-performance and cloud computing environments (HPC, AWS, Google Cloud).

My methodological contributions have yielded widely adopted open-source software and databases facilitating next-generation sequencing analysis, host–pathogen interaction modeling, and biological sequence classification. These tools, engineered using Python, R, and C++ and orchestrated via production-grade workflows (Snakemake, Nextflow), are extensively used in animal, plant, and human health research. I have directed the development of multiple standalone packages and interactive web platforms (R Shiny, Dash) that bridge the gap between advanced computational methods and accessible user interfaces.

I lead an externally funded research program at the nexus of computing and biology, with active grants (as PI and Co-PI) supporting the development of methods for genomic surveillance, disease diagnostics, and systems-level modeling. My publication portfolio comprises both methodological advancements in computational biology and applied studies that underscore the translational impact of these approaches on public health and precision medicine.

My long-term research agenda aims to advance core computational paradigms for biological data analysis—leveraging distributed computing and neural architectures—while fostering interdisciplinary collaborations through the creation of robust software, algorithms, and data-driven models that address global health challenges.

Academic Appointments and Professional Experience

Research Associate III, Bioinformatics

Nov 2023 – Present

Animal Disease Research & Diagnostic Laboratory South Dakota State University, Brookings, SD, USA

- Lead computational genomics research on avian viral pathogens; integrate multi-omics data for disease surveillance and diagnostics
- Co-PI on FFAR-funded project (\$150K) for diagnostic assay development
- Mentor graduate students in bioinformatics methods and project design
- Develop algorithms for viral genome analysis and phylogenetic reconstruction

Graduate Research Assistant

Jan 2019 – Nov 2023

Department of Plants, Soils and Climate Utah State University, Logan, UT, USA

- Developed machine learning pipelines for RNA-Seq, genome annotation, and variant analysis
- Created open-source tools (miPyRNA, SNVguru, pySeqRNA) adopted by global research groups
- Published 12 first-author papers; presented at national and international conferences
- Led collaborative projects with plant and animal genomics labs

- Led and contributed to projects analyzing high-throughput omics datasets from plant, animal, and human disease models, including host-pathogen interactions, immune response, and systems biology of disease.

Teaching Assistant, Bioinformatics

Aug 2011 – Dec 2018

School of Agricultural Biotechnology

Punjab Agricultural University, Ludhiana, India

- Assisted in teaching graduate and undergraduate courses in bioinformatics, genomics, and computational biology
- Designed and delivered hands-on workshops in NGS data analysis and Linux for biologists
- Supervised student research projects and laboratory sessions

Lecturer, Bioinformatics

Jul 2010 – Aug 2011

Department of Biotechnology

Guru Nanak Girls College, Ludhiana, India

- Taught undergraduate courses in bioinformatics and biotechnology
- Supervised student research projects and internships

Grants and Funding

Current Funding

- **Emerging avian metapneumovirus subgroup A and B in US poultry: development of diagnostic assays and control strategies**
Foundation for Food & Agriculture Research, Co-PI
Total: \$150,000
Status: Active
- **Broad-spectrum live recombinant vaccine and improved diagnostics for emerging avian metapneumovirus subgroups A and B causing severe outbreaks in US poultry**
Foundation for Food & Agriculture Research ROAR, Co-I
Total: \$150,000
Status: Active
- **Decoding the Poultry-HPAI Interactome: An Integrative Pipeline for Targeted Therapeutics Against Highly Pathogenic Avian Influenza**
USDA-APHIS, Call No. 10025-VSSP0000-25-0015, Co-PI
Total: \$1,969,435; Subaward: \$799,443
Status: Awarded

Submitted / Under Review

- **AI-Driven Genomic Surveillance and Control Strategies for Emerging Turkey Reoviruses (TRV)**
USDA-AFRI, PI
Total: \$300,000
Status: Submitted
- **BRIDGE: A host-pathogen interactome and machine learning platform to predict zoonotic risk of livestock and poultry viruses**
USDA-AFRI, PI
Total: \$300,000
Status: Submitted
- **Innovative approaches for early detection and control of zoonotic influenza viruses in poultry**
DHHS/National Institutes of Health (NIH), R21, Co-I

Total: \$400,500

Status: Submitted

- **Novel mass administration vaccine for humoral and cellular immunity using mRNA plus virus like particle**

USDA-AFRI, Co-PI

Total: \$300,000

Status: Submitted

- **Systems-Level Host Interactome Framework for Avian Metapneumovirus Mitigation: Advancing Therapeutics to Strengthen US Agricultural Biosecurity**

USDA-AFRI, Co-PI

Total: \$800,000

Status: Submitted

Education

Ph.D. in Plant Sciences (Bioinformatics and Computational Biology)

Jan 2019 – Dec 2024

Department of Plants, Soils and Climate

Utah State University, Logan, UT, USA

- Thesis: "Machine Learning and Data Mining in Complex Genomics Big Data: Developing Efficient Tools to Advance Computational Systems Biology"
- Advisor: Dr. Rakesh Kaundal

M.Sc. in Bioinformatics

Aug 2008 – Jul 2010

Punjabi University Patiala, India

- Thesis: "In Silico Prediction of MicroRNA in *Catharanthus roseus* and Their Role in Metabolomics"

B.Sc. in Biology

Jul 2005 – Jun 2008

University College

Kurukshetra University, Kurukshetra, India

- Subjects: Zoology, Botany, Chemistry

Add-on Certificate Diploma in Bioinformatics

2007 – 2008

Department of Bioinformatics

Kurukshetra University, Kurukshetra, India

Awards and Honors

- **USU Robins Awards Finalist for Doctoral Student Researcher of the Year (2022-2023)**, Utah State University, Logan, UT, USA
- **Doctoral Student Researcher of the Year (2022-2023)**, College of Agriculture and Applied Sciences, Utah State University, Logan, UT, USA
- **Doctoral Student Researcher of the Year (2022-2023)**, Department of Plants, Soils and Climate, Utah State University, Logan, UT, USA
- **Young Scientist Award (2017)**, National Conference on Technological Challenges in Social, Environmental, and Agricultural Reforms (TECHSEAR-2017), IIRR Hyderabad, India

Bioinformatics Tools and Resources Developed

Software and Computational Tools (5)

- **miPyRNA** (<https://github.com/navduhan/mipyRNA>): An automated python package for small RNA sequencing data analysis. Used by 100+ researchers worldwide.
- **SNVguru** (<https://github.com/usubioinfo/snvguru>): An automated python package for single nucleotide variation (SNV) analysis. Adopted by 50+ research groups.

- **pySeqRNA** (<https://github.com/navduhan/pyseqrna>): An automated python package for next-generation sequencing data analysis. Complete documentation website with 112+ users.
- **MINpred** (<https://bioinfo.usu.edu/minpred/download/minpred.tar.gz>): A standalone alignment-free deep learning based tool for nitrogen mineralization enzyme classification.
- **deepNEC** (<https://navduhan@bitbucket.org/navduhan/deepnec.git>): A standalone alignment-free deep learning based tool for nitrogen metabolism enzyme classification. Web server with 700+ users.

Databases and Web Resources (13)

- **HPInet** (<http://bioinfo.usu.edu/hpinet/>): A Cereal crops and pathogen protein-protein interaction database (20+ users)
- **HuPoxNET** (<http://bioinfo.usu.edu/hupoxnet/>): A database for Human-Monkeypox virus interactome (20+ users)
- **myDockDB** (<http://bioinfo.usu.edu/myDockDB/>): A Multiple Myeloma Proteins and Natural Compounds Docking WEB-resource (20+ users)
- **ranchSATdb** (<http://bioinfo.usu.edu/ranchSATdb/>): A database of ranch animals microsatellite markers (30+ users)
- **TritiKBdb** (<http://bioinfo.usu.edu/tritikbdb/>): A Host-Pathogen interaction database for understanding Karnal Bunt disease infection mechanisms in Wheat (540+ users)
- **HuCoPIA** (<http://bioinfo.usu.edu/hucopia/>): An Atlas of Human vs. SARS-CoV-2 interactome and the comparative analysis with other Coronaviridae family viruses (60+ users)
- **deepHPI** (<http://bioinfo.usu.edu/deepHPI/>): A Deep Learning framework for the prediction of host-pathogen protein interactions and visualization (1.03K+ users)
- **GreeningDB** (<http://bioinfo.usu.edu/GreeningDB/>): A database of host-pathogen interactions and studying comparatomics of citrus and citrus greening disease (HLB) (828+ users)
- **covidTracker** (<http://bioinfo.usu.edu/covidTracker/>): A COVID 19 disease infection statistics tracker (8,200+ users)
- **legumeSSRdb** (<http://bioinfo.usu.edu/legumeSSRdb/>): A Legume species microsatellite markers database (712+ users)
- **citSATdb** (<http://bioinfo.usu.edu/citSATdb/>): A Citrus species microsatellite markers database (530+ users)

Publications

I have **32 publications** in international peer-reviewed journals, including **12 first-author** articles. My **h-index = 9**, **i10-index = 9**, and **total citations = 333+** (as per [Google Scholar Citations](#)).

^S denotes equal contribution; underlined names indicate undergraduate mentees; underlined italic names indicate graduate mentees.

Journal Articles

- [1] Debnath, S., Tyagi, W., Rai, M., Singh, K., Majumder, S., **Duhan, N.**, and Meetei, N. T. “Discovery of potential haplotypes associated with varying levels of vicine content due to the InDel1.4 and a coding-SNP in the VC1 gene in faba bean (*Vicia faba* L.)” In: *Plant Gene* 41 (2025), p. 100481. ISSN: 2352-4073. URL: <https://doi.org/10.1016/j.plgene.2024.100481>.

- [2] **Duhan, N.** and Kaundal, R. “AtSubP-2.0: An integrated web server for the annotation of Arabidopsis proteome subcellular localization using deep learning”. In: *The Plant Genome* 18.1 (Mar. 2025), e20536. URL: <https://doi.org/10.1002/tpg2.20536>.
- [3] **Duhan, N.** and Kaundal, R. “PRGminer: harnessing deep learning for the prediction of resistance genes involved in plant defense mechanisms”. In: *Frontiers in Plant Science* 16 (June 2025), p. 1411525. URL: <https://doi.org/10.3389/fpls.2025.1411525>.
- [4] **Duhan, N.** and Kaundal, R. “RSLpred2: An Integrated Web Server for the Annotation of Rice Proteome Subcellular Localization Using Deep Learning”. In: *Rice* 18 (2025), p. 58. URL: <https://doi.org/10.1186/s12284-025-00767-7>.
- [5] Selim, M., Jangra, S., Luqman, M., **Duhan, N.**, Temeeyasen, G., Serrao, T., Kaushik, R. S., Sharafeldin, T., and Mor, S. “Isolation and Molecular Characterization of Avian Metapneumovirus Subgroup B Virus Isolated from a Chicken Farm in the United States”. In: *Avian Dis* 69.S1 (Dec. 2025), pp. 562–573. URL: <https://doi.org/10.1637/aviandiseases-D-25-00071>.
- [6] Selim, Mohamed, Jangra, Sumit, Luqman, Muhammad, Clement, Travis, Duhan, Naveen, Temeeyasen, Gun, Sharafeldin, Tamer, and Mor, Sunil. “Tracking the Virus Spread and Evolution of Emerging Avian Metapneumovirus Subtypes A and B in U.S. Poultry”. In: *Avian Diseases* 69.S1 (Dec. 2025), pp. 548–561. URL: <https://doi.org/10.1637/aviandiseases-D-25-00072>.
- [7] Selim, Mohamed, Jangra, Sumit, Luqman, Muhammad, Duhan, Naveen, Temeeyasen, Gun, Serrao, Tricia, Kaushik, Radhey S., Sharafeldin, Tamer, and Mor, Sunil. “Isolation and Molecular Characterization of Avian Metapneumovirus Subgroup B Virus Isolated from a Chicken Farm in the United States”. In: *Avian Diseases* 69.S1 (Dec. 2025), pp. 562–573. URL: <https://doi.org/10.1637/aviandiseases-D-25-00071>.
- [8] Kataria, Raghav, **Duhan, N.** and Kaundal, Rakesh. “Navigating the human-monkeypox virus interactome: HuPoxNET atlas reveals functional insights”. In: *Frontiers in Microbiology* 15 (2024). ISSN: 1664-302X. URL: <https://doi.org/10.3389/fmicb.2024.1399555>.
- [9] Kaur, S., Seem, K., **Duhan, N.**, Kumar, S, Kaundal, R, and Mohapatra, T. “Comparative miRNome and transcriptome analyses reveal the expression of novel miRNAs in the panicle of rice implicated in sustained agronomic performance under terminal drought stress”. In: *Planta* 259.6 (2024), p. 128. ISSN: 1432-2048. URL: <https://doi.org/10.1007/s00425-024-04399-x>.
- [10] Luqman, M., **Duhan, N.**, Temeeyasen, G., Selim, M., *Jangra, S.*, and Mor, S.K. “Geographical Expansion of Avian Metapneumovirus Subtype B: First Detection and Molecular Characterization of Avian Metapneumovirus Subtype B in US Poultry”. In: *Viruses* 16 (4 2024), p. 508. URL: <https://doi.org/10.3390/v16040508>.
- [11] **Duhan, N.** and Kaundal, R. “HuCoPIA: An Atlas of Human vs. SARS-CoV-2 Interactome and the Comparative Analysis with Other Coronaviridae Family Viruses”. In: *Viruses* 15.2 (2023), p. 492. URL: <https://doi.org/10.3390/v15020492>.
- [12] **Duhan, N.**, Kaur, S., and Kaundal, R. “ranchSATdb: a Genome-wide Simple Sequence Repeat (SSR) Markers Database of Livestock Species for Mutant Germplasm Characterization and Improving Farm Animal Health.” In: *Genes* 14.2 (2023), p. 1481. URL: <https://doi.org/10.3390/genes14071481>.
- [13] Kaur, S., Seem, K., **Duhan, N.**, Kumar, S., Kaundal, R., and Mohapatra, T. “Transcriptome and Physio-Biochemical Profiling Reveals Differential Responses of Rice Cultivars at Reproductive-Stage Drought Stress”. In: *International Journal of Molecular Sciences* 24.2 (2023), p. 1002. URL: <https://doi.org/10.3390/ijms24021002>.
- [14] **Duhan, N.**, Kataria, R., and Kaundal, R. “TritiKBdb: A Functional Annotation Resource for Deciphering the Complete Interaction Networks in Wheat-Karnal Bunt Pathosystem”. In: *International Journal of Molecular Sciences* 23.13 (2022), p. 7455. URL: <https://doi.org/10.3390/ijms23137455>.
- [15] **Duhan, N.**, Norton, J. M., and Kaundal, R. “deepNEC: a novel alignment-free tool for the identification and classification of nitrogen biochemical network-related enzymes using deep learning”.

- In: *Briefings in Bioinformatics* 23.3 (2022), bbac071. URL: <https://doi.org/10.1093/bib/bbac071>.
- [16] Kataria, R., **Duhan, N.**, and Kaundal, R. “Computational systems biology of Alfalfa–bacterial blight host–pathogen interactions: uncovering the complex molecular networks for developing durable disease resistant crop”. In: *Frontiers in Plant Science* 12 (2022), p. 807354. URL: <https://doi.org/10.3389/fpls.2021.807354>.
- [17] Kaundal, R., Loaiza, C. D., **Duhan, N.**, and Flann, N. “deepHPI: a comprehensive deep learning platform for accurate prediction and visualization of host–pathogen protein–protein interactions”. In: *Briefings in Bioinformatics* 23.3 (2022), bbac125. URL: <https://doi.org/10.1093/bib/bbac125>.
- [18] Mahalingam, R., **Duhan, N.**, Kaundal, R., Smertenko, A., Nazarov, T., and Bregitzer, P. “Heat and drought induced transcriptomic changes in barley varieties with contrasting stress response phenotypes”. In: *Frontiers in Plant Science* 13 (2022), p. 4652. URL: <https://doi.org/10.3389/fpls.2022.1066421>.
- [19] Rutigliano, H. M., Thomas, A. J., Umbaugh, J. J., Wilhelm, A., Sessions, B. R., Kaundal, R., **Duhan, N.**, Hicks, B. A., Schlafer, D. H., White, K. L., and J., Davis C. “Increased expression of pro-inflammatory cytokines at the fetal–maternal interface in bovine pregnancies produced by cloning”. In: *American Journal of Reproductive Immunology* 87.3 (2022), e13520. URL: <https://doi.org/10.1111/aji.13520>.
- [20] **Duhan, N.** and Kaundal, R. “LegumeSSRdb: A Comprehensive Microsatellite Marker Database of Legumes for Germplasm Characterization and Crop Improvement”. In: *International Journal of Molecular Sciences* 22.21 (2021), p. 11350. URL: <https://doi.org/10.3390/ijms222111350>.
- [21] Kaundal, R., **Duhan, N.**, Acharya, B. R., Pudussery, M. V., Ferreira, J. FS., Suarez, D. L., and Sandhu, D. “Transcriptional profiling of two contrasting genotypes uncovers molecular mechanisms underlying salt tolerance in alfalfa”. In: *Scientific Reports* 11.1 (2021), p. 5210. URL: <https://doi.org/10.1038/s41598-021-84461-w>.
- [22] Loaiza, C. D., **Duhan, N.**^S, Lister, M., and Kaundal, R. “In silico prediction of host–pathogen protein interactions in melioidosis pathogen *Burkholderia pseudomallei* and human reveals novel virulence factors and their targets”. In: *Briefings in Bioinformatics* 22.3 (2021), bbz162. URL: <https://doi.org/10.1093/bib/bbz162>.
- [23] Loaiza, C. D., **Duhan, N.**, and Kaundal, R. “GreeningDB: A Database of Host–Pathogen Protein–Protein Interactions and Annotation Features of the Bacteria Causing Huanglongbing HLB Disease”. In: *International Journal of Molecular Sciences* 22.19 (2021), p. 10897. URL: <https://doi.org/10.3390/ijms221910897>.
- [24] **Duhan, N.** and Kaundal, R. “pySeqRNA: an automated Python package for RNA sequencing data analysis”. In: *f1000research* (2020). URL: <https://doi.org/10.7490/f1000research.1118314.1>.
- [25] **Duhan, N.**, Meshram, M., Loaiza, C. D., and Kaundal, R. “citSATdb: genome-wide simple sequence repeat (SSR) marker database of Citrus species for germplasm characterization and crop improvement”. In: *Genes* 11.12 (2020), p. 1486. URL: <https://doi.org/10.3390/genes11121486>.
- [26] **Kaur, K.**, **Duhan, N.**, Singh, J., Kaur, G., and Vikal, Y. “Computational identification of maize miRNA and their gene targets involved in biotic and abiotic stresses”. In: *Journal of biosciences* 45 (2020), pp. 1–17. URL: <https://doi.org/10.1007/s12038-020-00106-6>.
- [27] Malhotra, P. K., Verma, G., Sidhu, G. S., and **Duhan, N.** “EPIGENOMICS: ROLE, APPROACHES AND APPLICATIONS IN PLANTS”. In: *JAPS, Journal of Animal and Plant Sciences* 30.5 (2020), pp. 1071–1082. URL: <https://doi.org/10.36899/JAPS.2020.5.0122>.
- [28] Sharma, B., Batz, T. A., Kaundal, R., Kramer, E. M., Sanders, U. R., Mellano, V. J., **Duhan, N.**, and Larson, R. B. “Developmental and molecular changes underlying the vernalization-induced transition to flowering in *Aquilegia coerulea* (James)”. In: *Genes* 10.10 (2019), p. 734. URL: <https://doi.org/10.3390/genes10100734>.

- [29] Mohanpuria, P., **Duhan, N.**, Sarao, N. K., Kaur, M., and Kaur, M.. “In silico identification and validation of potential microRNAs in Kinnow Mandarin (*Citrus reticulata* Blanco)”. In: *Interdisciplinary Sciences: Computational Life Sciences* 10 (2018), pp. 762–770. URL: <https://doi.org/10.1007/s12539-017-0235-6>.
- [30] *Jhinjer, R. K.*, **Duhan, N.**, and Chhuneja, P. “Comparative Structure Modeling of Glyoxalase II Protein of *Oryza sativa*”. In: *Int. J. Curr. Microbiol. App. Sci* 6.8 (2017), pp. 1380–1387. URL: <https://doi.org/10.20546/ijcmas.2017.608.168>.
- [31] **Duhan, N.**, Sarao, N.K., Mohanpuria, P., and Sidhu, M.K. “In silico identification and characterization of potential miRNAs from *Capsicum annum*.” In: *HortFlora Research Spectrum* 5.4 (2016), pp. 306–309. URL: <https://www.cabdirect.org/cabdirect/abstract/20173074226>.
- [32] Mohanpuria, P., Sarao, N. K., **Duhan, N.**, Brar, G., Kaur, A., and Singh, K. “Computational identification and characterization of potential miRNAs from Asiatic cotton (*Gossypium arboreum* L.)” In: *Crop Improv* 40.2 (2013), pp. 121–128.
- [33] Omer, A., Singh, S., and **Duhan, N.** “Identification of miRNAs in *C. roseus* and their potential targets”. In: *Bioinformatics* 8.2 (2012), p. 75. URL: <https://doi.org/10.6026/97320630008075>.

Under Review

- [1] **Duhan, N.** and Kaundal, R. “pySeqRNA: an all-in-one NGS to Spreadsheet Python package for RNA-Seq Data, including single, multi-gene quantification and error correction”. *Bioinformatics*. 2025.
- [2] Larson, R.B., **Duhan, N.**, and Kaundal, R. “In silico identification of potential drug-like natural compounds for the treatment of Multiple Myeloma: A comprehensive molecular docking and simulation study”. *Frontiers in Chemistry*. 2025.
- [3] Perisse, I., Kuczynski, C., Liu, Y., Alipour, E., LaPradd, C., Regouski, M., Baker, S., Poudel, L., Babatunde, O., Kaundal, R., **Duhan, N.**, Fan, Z., Xu, H., Lau, C., Stott, R., Sweat, A., Beaty, M., Kim-Shapiro, D., Atala, A., Porada, C., and Polejaeva, I. “Introduction of the Human Sick Cell Mutation in Sheep Beta-globin by CRISPR/Cas9 Models Human SCD in Cloned Sheep”. Under review, *Lab Animals*. 2025.

Manuals

- [1] **Duhan, N.**, Chhuneja, P., and Singh, K. *Introduction to Bioinformatics: A Practical Manual*. School of Agricultural Biotechnology, Punjab agricultural University, Ludhiana. COA/2012/Manual-10, 2012, pp. 1–100.

Abstract / Oral / Conferences

- [1] **Duhan, N.**, Temeeyasen, G., Luqman, M., Selim, M., Khaled, A., T., Sharafeldin, Jangra, S., and Mor, S. “trevoDB: A Comprehensive Resource for Turkey Orthoreovirus Classification and Comparative Analysis”. In: *75th North Central Avian Disease Conference*. 2024.
- [2] **Duhan, N.** and Kaundal, R. “PySeqRNA: pySeqRNA: an Automated Python Package for Next-Generation Sequencing Data Analysis.” In: *Plant and Animal Genomes Conference 30 (PAG International 2023)*. 2023.
- [3] **Duhan, N.** and Kaundal, R. “PySeqRNA: pySeqRNA: an Automated Python Package for Next-Generation Sequencing Data Analysis.” In: *The 2022 Conference on Applied Statistics in Agriculture and Natural Resources, Utah State University, Logan, Utah, USA; May 16-19, 2022*. 2022.
- [4] Kaundal, R., **Duhan, N.**, Acharya, B. R., Pudussery, M. V., Ferreira, J. FS., Suarez, D. L., and Sandhu, D. “Transcriptional profiling of two contrasting genotypes uncovers molecular mechanisms underlying salt tolerance in Alfalfa.” In: *The 2022 Conference on Applied Statistics in Agriculture and Natural Resources, Utah State University, Logan, Utah, USA; May 16-19, 2022*. 2022.
- [5] **Duhan, N.** and Kaundal, R. “PySeqRNA: An Automated Python Package for Next-Generation Sequencing Data Analysis”. In: *ASA, CSSA, SSSA International Annual Meeting*. 2021.

- [6] **Duhan, N.**, Norton, J. M., and Kaundal, R. “deepNEC: a novel alignment-free tool for the identification and classification of nitrogen biochemical network-related enzymes using deep learning”. In: *ASA, CSSA, SSSA International Annual Meeting*. 2021.
- [7] Kataria, R., **Duhan, N.**, and Kaundal, R. “Computational Systems Biology of Alfalfa-Pseudomonas Syringae Protein-Protein Interactions: Uncovering the Complex Molecular Networks of Host-Pathogen Interactions to Control Bacterial Blight in Alfalfa”. In: *ASA, CSSA, SSSA International Annual Meeting*. 2021.
- [8] Kataria, R., **Duhan, N.**, and Kaundal, R. “Translating-omics big data: Comprehensive understanding of host-pathogen interactions to control bacterial blight in alfalfa using computational approaches”. In: *MOLECULAR PLANT-MICROBE INTERACTIONS*. Vol. 34. 7. 2021.
- [9] **Duhan, N.** and Kaundal, R. “PySeqRNA: an Automated Python Package for Next-Generation Sequencing Data Analysis.” In: *The 28th Annual International Conference on Intelligent Systems for Molecular Biology*. 2020.
- [10] **Duhan, N.**, Meshram, M., Loaiza, C. D., and Kaundal, R. “citSATdb: genome-wide simple sequence repeat (SSR) marker database of Citrus species for germplasm characterization and crop improvement”. In: *Plant and Animal Genomes Conference XXVIII (PAG International 2020)*. 2020.
- [11] **Duhan, N.**, Norton, J. M., and Kaundal, R. “deepNEC: a novel alignment-free tool for the identification and classification of nitrogen biochemical network-related enzymes using deep learning”. In: *The 2020 Annual ASM Intermountain Branch Meeting, American Society of Microbiology*. 2020.
- [12] **Duhan, N.**, Norton, J. M., and Kaundal, R. “NECminer: a novel alignment-free tool for the identification and classification of Nitrification-related enzymes using Deep Learning.” In: *The 28th Annual International Conference on Intelligent Systems for Molecular Biology*. 2020.
- [13] Loaiza, C. D., **Duhan, N.**, and Kaundal, R. “GreeningDB: a Database of Protein Features and Protein- Protein Interactions of the Bacteria causing HLB.” In: *Plant and Animal Genomes Conference XXVIII (PAG International 2020)*. 2020.
- [14] Mahalingam, R., Bregitzer, P., **Duhan, N.**, and Kaundal, R. “Transcriptome Analysis of Barley Varieties with Contrasting Responses to Heat, Drought and Combined Heat and Drought Stress”. In: *ASA, CSSA, SSSA Annual International Agronomy Meeting*. 2020.
- [15] Kaundal, R. and **Duhan, N.** “Comparative transcriptomics study of alfalfa (*Medicago sativa* L.) reveals novel genotype-specific salt tolerance mechanisms”. In: *The Annual Plants, Soils, and Climate Showcase 2019, College of Agriculture and Applied Sciences; Utah State University, Logan, Utah, USA*. 2019.
- [16] Yadav, I. S., Malhotra, P. K., **Duhan, N.**, and Sharma, A. “Functional annotation of sugarcane ESTs using bioinformatics tools”. In: *International conference on recent trends in bioinformatics and biotechnology for sustainable development, Punjab Agricultural University, Ludhiana, India*. 2016.
- [17] **Duhan, N.**, Sarao, N. K., Mohanpuria, P., Kaur, R., and Sandhu, P. “In silico identification and characterization of Potential miRNAs from *Capsicum annuum*”. In: *National Symposium Crop Improvement for Inclusive Sustainable Development, Punjab Agricultural University, Ludhiana, India 7-9 November 2014*. 2014.
- [18] **Duhan, N.** and Chunneja, P. “In silico identification of *Catharanthus roseus* miRNAs and their role in human metabolomics”. In: *National Symposium on Bioinformatics Feb 02, 2012, Jammu University, Jammu, India*. 2012.
- [19] Jhinjer, R. K., **Duhan, N.**, Yadav, B., and Chunneja, P. “Comparative structure modelling of glyoxalase II protein of *Oryza sativa*”. In: *National Symposium on Bioinformatics Feb 02, 2012, Jammu University, Jammu, India*. 2012.
- [20] Kaur, P, **Duhan, N.**, Yadav, B., and Chunneja, P. “In silico detection of simple sequence repeats in expressed sequence tags of wheat”. In: *National Symposium on Bioinformatics Feb 02, 2012, Jammu University, Jammu, India*. 2012.

- [21] Manchanda, P., **Duhan, N.**, Yadav, B., and Chunneja, P. “In silico prediction of microRNAs from pea (*Pisum sativum*)”. In: *National Symposium on Bioinformatics Feb 02, 2012, Jammu University, Jammu, India*. 2012.
- [22] Verma, L., Yadav, B., **Duhan, N.**, and Chunneja, P. “In silico gene prediction in *Vitis vinifera* using fourth generation tool”. In: *National Symposium on Bioinformatics Feb 02, 2012, Jammu University, Jammu, India*. 2012.

Presentations and Lectures

Invited Talks

- **LaTeX: A Useful Tool for Scientific Writing and Presentations** (November 22, 2017), School of Agricultural Biotechnology, Punjab Agricultural University, Ludhiana, India
- **Role of Bioinformatics in Conservation and Use in Plant Breeding** (November 1, 2017), Training programme on “Enhancing germplasm use through pre-breeding, evaluation, and frontier”, Centre for Advanced Faculty training in Genetics and Plant Breeding, Punjab Agricultural University, Ludhiana, India
- **Open-source computing for young biologists: Overview of Linux** (February 3, 2017), Hands-on training on text mining at “Biotechnological platforms for utility production maize”, ICAR-Indian Institute of Maize Research, Ludhiana, India
- **Differential Gene Expression Studies** (February 6, 2017), Hands-on training of raw data assembly, identification, and annotation of differential genes, “Biotechnological platforms for utility production maize”, ICAR-Indian Institute of Maize Research, Ludhiana, India
- **Computational identification of disease responsive microRNAs** (August 30, 2016), “Innovative breeding approaches for the development of climate smart crops”, Centre for Advanced Faculty training in Genetics and Plant Breeding, Punjab Agricultural University, Ludhiana, India
- **Bioinformatics: A step towards innovation in agriculture** (August 6-26, 2015), Training programme on “Mendelian Genetics to Molecular Genetics in Relevance to Plant Breeding”, Centre for Advanced Faculty training in Genetics and Plant Breeding, Punjab Agricultural University, Ludhiana, India
- **Bioinformatics: A step towards innovation** (August 23, 2014), “Milk Genomics workshop”, DAV College Karnal, India
- **In silico gene prediction and annotation** (August 18, 2014), Department of Plant Breeding, Punjab Agricultural University, Ludhiana, India
- **Role of miRNAs in Clinical Research** (November 13, 2013), Workshop of Bioinformatics and Human Genomics, Christian Medical College, Ludhiana, India
- **Searching Biological Databases** (November 13, 2013), Workshop of Bioinformatics and Human Genomics, School of Agricultural Biotechnology, Punjab Agricultural University, Ludhiana, India
- **Role of Bioinformatics in Genomics and Primer Designing** (October 23, 2013), University College, Kurukshetra University, Kurukshetra, India
- **In silico gene prediction and annotation** (September 27, 2013), Department of Plant Breeding, Punjab Agricultural University, Ludhiana, India
- **Protein structure and function prediction** (March 15, 2013), Department of Biotechnology, Kumaun University Nainital, India
- **Open-source for budding Computational Biologist** (March 13-14, 2013), National Symposium and Workshop on Current Advances in Bioinformatics, Aligarh Muslim University, Aligarh, India

- **In silico gene prediction and annotation** (March 12, 2013), Department of Plant Breeding, Punjab Agricultural University, Ludhiana, India

Research Interests

1. Multi-Omics and Next-Generation Sequencing Analysis

- Integration of genomics, transcriptomics, proteomics, and metabolomics data to improve animal, plant, and human health responses to environmental changes
- Comprehensive pipelines for genome assembly, RNA-seq, ChIP-seq, single-cell RNA-seq, variant sequencing, Digenome-seq, and metagenomics data analysis
- Genome-scale functional annotation and development of novel algorithms for biochemical pathway prediction and structure/function analyses

2. Artificial Intelligence and Machine Learning in Bioinformatics

- Computational modeling to analyze large-scale sequence data using deep learning approaches for biological discovery
- Development of alignment-free deep learning tools for enzyme classification and biological sequence analysis
- AI-driven prediction of protein structures, drug discovery, and biological interactions

3. Systems Biology and Host-Pathogen Interactions

- Integration of phenotypic, genotypic, and environmental data for host-pathogen interactions and protein-protein interactions (intra- and inter-species)
- Metabolic pathway analysis, gene regulatory networks, and protein interaction network modeling
- Computational analysis of disease mechanisms and host-pathogen co-evolution dynamics

4. Web Server and Database Development

- Creation of prediction tools, web servers, and databases to aid the research community in computational biology and bioinformatics applications
- Development of comprehensive host-pathogen interaction databases for understanding disease mechanisms across multiple species
- Specialized databases for microsatellite markers, protein interactions, and biological resources for genetic diversity and population studies

5. Translational and Disease-Focused Bioinformatics

- Application of computational biology and bioinformatics to understand mechanisms of inflammation, vascular disease, and fibrosis using human and animal model datasets, with a focus on translational and precision medicine.

Teaching Experience

Utah State University, Logan, UT, USA

- **PSC 4150/6150:** Bioinformatics and Big Data Mining (Graduate/Undergraduate), 2022

Punjab Agricultural University, Ludhiana, India

- **Biotech 509:** Bioinformatics Tools and Their Application in Agriculture (Graduate), 2013-2018
- **Biotech 607:** Advances in Bioinformatics (Graduate), 2013-2018

- **Biotech 307:** Introduction to Bioinformatics (Undergraduate), 2013-2018
- **Biotech 401:** Introduction to Genomics and Proteomics (Undergraduate), 2013-2018
- **Biotech 404:** Computational Biology (Undergraduate), 2013-2018
- **Biotech 499:** In-house Training (Undergraduate), 2013-2018

Workshops and Training

- **NGS Data Analysis Workshop:** Hands-on training on RNA-Seq, ChIP-Seq, and variant calling
- **Linux for Biologists:** Command-line tools and scripting for biological data analysis
- **Machine Learning in Bioinformatics:** Python-based workshops on ML applications

Leadership and Service

Academic Leadership

- **President, Indian Student Association** (2021-2022), Utah State University, Logan, UT, USA
- **Workshop,** Conducted small-group workshops on NGS, Linux, and Machine Learning for Biologists

Professional Service

- **Reviewer, Undergraduate Research and Creative Opportunities (URCO) Grant Program** (2022-2023), Utah State University, Logan, UT, USA
- **Moderator and Evaluator** (2023), Utah State University Research Symposium, Logan, UT, USA
- **Manuscript Reviewer:** Briefings in Bioinformatics, BMC Genomics, Journal of Plant Biochemistry and Biotechnology, Scientific Reports-Nature, Frontiers in Cell and Developmental Biology, Genes, and Forests

Curriculum Development

- **Designed 8 courses** for the Bioinformatics module in B. Tech Biotechnology Degree under Indian Council of Agricultural Research (ICAR) 5th Dean committee for course revision (2013-2018), Punjab Agricultural University, Ludhiana, India

Technical Skills

Proficient in a comprehensive suite of programming languages, bioinformatics tools, and high-performance computing environments.

Programming Languages & Development

- **Python:** Advanced bioinformatics tool development, machine learning pipelines, data analysis frameworks, web application development (Flask/Django), API development
- **R:** Statistical genomics, Bioconductor packages, differential expression analysis, pathway enrichment, visualization (ggplot2, plotly), R Shiny applications
- **Perl:** High-throughput sequence processing, text mining, pipeline automation, legacy bioinformatics tool maintenance
- **Bash/Shell:** Advanced scripting, pipeline development, system administration, cluster job management, automation workflows
- **SQL:** Database design, complex queries, PostgreSQL, MySQL, SQLite for biological data management

- **JavaScript/HTML/CSS:** Web-based bioinformatics tools, interactive visualizations, dashboard development
- **C++/C:** High-performance computing, algorithm optimization, memory-efficient data structures

Advanced Bioinformatics Analysis

- **Single-Cell RNA-Seq:** Seurat, Scanpy, Monocle, cell type identification, trajectory analysis, spatial transcriptomics (10x Genomics, Visium)
- **Multi-Omics Integration:** WGCNA, MOFA, mixOmics, multi-modal data fusion, cross-platform validation
- **ChIP-Seq & ATAC-Seq:** Peak calling, motif analysis, chromatin accessibility, transcription factor binding, epigenomic analysis
- **Metagenomics:** QIIME2, MetaPhlAn, HUMAnN, taxonomic profiling, functional annotation, microbiome analysis
- **Structural Bioinformatics:** AlphaFold, I-TASSER, PyMOL, protein structure prediction, molecular docking, drug discovery
- **Population Genomics:** PLINK, VCFtools, STRUCTURE, ADMIXTURE, GWAS, selection scans, demographic inference

Genome Analysis & Assembly

- **De Novo Assembly:** SPAdes, Velvet, Canu, Flye, assembly quality assessment, scaffolding, gap filling
- **Genome Annotation:** MAKER, BRAKER, AUGUSTUS, gene prediction, functional annotation, comparative genomics
- **Variant Analysis:** GATK, FreeBayes, variant calling, filtering, annotation, clinical interpretation
- **Comparative Genomics:** OrthoMCL, BLAST, synteny analysis, gene family evolution, whole-genome alignments

Machine Learning & AI

- **Deep Learning:** TensorFlow, PyTorch, Keras, neural networks for sequence analysis, protein structure prediction
- **Supervised Learning:** Random Forest, SVM, XGBoost, ensemble methods for classification and regression
- **Unsupervised Learning:** Clustering, dimensionality reduction, PCA, t-SNE, UMAP for data exploration
- **Natural Language Processing:** BERT, transformers for biological text mining, literature analysis

High-Performance Computing

- **Cluster Computing:** SLURM, PBS, job scheduling, parallel processing, resource optimization
- **Cloud Computing:** AWS, Google Cloud, Docker, Kubernetes, scalable bioinformatics workflows
- **Big Data:** Apache Spark, Hadoop, distributed computing for large-scale genomic analysis
- **GPU Computing:** CUDA, GPU-accelerated bioinformatics algorithms, deep learning optimization

Bioinformatics Tools & Pipelines

- **Sequence Analysis:** BLAST, BWA, Bowtie2, STAR, HISAT2, sequence alignment and mapping
- **Transcriptomics:** DESeq2, edgeR, limma, differential expression, pathway analysis, gene set enrichment
- **Phylogenetics:** RAxML, IQ-TREE, MrBayes, phylogenetic reconstruction, molecular evolution
- **Network Analysis:** Cytoscape, STRING, protein-protein interaction networks, pathway analysis
- **Workflow Management:** Snakemake, Nextflow, Galaxy, reproducible research, pipeline automation

Data Visualization & Communication

- **Scientific Visualization:** matplotlib, seaborn, plotly, D3.js, publication-quality figures
- **Interactive Dashboards:** R Shiny, Plotly Dash, web-based data exploration tools
- **3D Visualization:** PyMOL, Chimera, molecular structure visualization and analysis
- **Geographic Visualization:** Spatial data analysis, mapping tools for geographic genomics

Databases & Resources

- **Sequence Databases:** NCBI GenBank, RefSeq, Ensembl, UniProt, comprehensive sequence resources
- **Functional Databases:** KEGG, Reactome, GO, pathway and functional annotation databases
- **Interaction Databases:** STRING, IntAct, BioGRID, protein-protein interaction networks
- **Variant Databases:** dbSNP, ClinVar, COSMIC, human genetic variation resources
- **Specialized Resources:** Cancer Genome Atlas (TCGA), 1000 Genomes, specialized genomic projects

Professional & Soft Skills

- **Grant Writing:** NIH, NSF, USDA proposals, budget development, project planning
- **Scientific Communication:** Manuscript writing, conference presentations, public speaking
- **Project Management:** Agile methodologies, team leadership, timeline management, stakeholder communication
- **Teaching & Mentoring:** Curriculum development, workshop organization, student supervision
- **Collaboration:** Cross-disciplinary research, international partnerships, industry collaborations

Media and Outreach

Media Coverage

- **Data Scientists Uncovering Genes that Protect Alfalfa against Salinity Stress** (2021), [USU Today](#), Utah State University
- **USU Bioinformatics Expert Hopes Big Data Lab Will Revolutionize More Fields of Research** (2022), [USU Today](#), Utah State University
- **SDSU awarded research grant to help fight poultry virus** (2025), [SDSU News](#), South Dakota State University

Outreach Activities

- **Invited Speaker** at 10+ national and international conferences and workshops

- **Workshop Organizer** for NGS, Linux, and Machine Learning for Biologists (2019-2023)
- **Science Communication** through presentations and demonstrations at university open houses
- **Student Mentoring** through formal and informal programs