Patrick J. Lawrence

(+1) 614-735-8139 | patrick.skillman-lawrence@osumc.edu | kingston487 | PatrickJLawrence487 | Patrick J. Lawrence

Summary

I am currently pursuing a PhD in Computational Biology and an MS in Computer Science at The Ohio State University. The research I conduct, under Dr. Xia Ning in the Biomedical Informatics Department, is focused on the development/application of deep learning methodologies for/to biological problems. I have specific interest in applying representation learning techniques on multi-omics data to aid in drug recommendation and repurposing tasks.

Education

The Ohio State University
Columbus, OH, USA

PhD, Computational Biology and Bioinformatics | GPA: 3.848 *2020 - present*

- *Advisor*: Xia Ning
- *Relevant Coursework*: Applied Regression Analysis, Bioinformatics, Concepts of Biomedical Sciences, Data Science, Linear Modeling, Machine Learning and Statistical Pattern Recognition

The Ohio State University
Columbus, OH, USA

MS, Computer Science *2022 - present*

- *Advisor*: Xia Ning
- *Relevant Coursework*: Advanced Operating Systems, Advanced Survey in Artificial Intelligence, Algorithms, Programming Languages, Image Processing, Neural Networks

Washington and Lee University
Lexington, VA, USA

BA, Chemistry | Minor, Philosophy

- *Relevant Coursework*: Applied Bayesian Regression, Data Science, Statistics, Endocrinology, Genetics, Immunology, Biochemistry, Inorganic/Bioinorganic Chemistry, Physical Chemistry: Quantum and Computational Chemistry

Research Experience

The Ohio State University
Columbus, OH, USA

Graduate Student — Ning Lab *Oct. 2020 - present*

1. **Enhancing drug and cell line representations via contrastive learning for anti-cancer drug prioritization**
   - Leveraged contrastive learning to learn both small molecule drug and cell line embeddings that can be used to predict individualized cancer drug response.
   - A random forest model applied to our novel embeddings achieved state-of-the-art performance with improvements of 37.74%, 53.27%, and 58.97% with respect to average cell line-wise precision@k=[1,2,3], respectively.
   - Pathway enrichment analysis of 40 genes whose expression was significantly correlated with our model's prioritization of docetaxel showed enrichment of three *Akt* pathways documented to induce docetaxel-resistance.
2. **Predicting MHC class I antigen processing via representation learning and cleavage site-specific kernels**
   - Facilitated binary antigen processing predictions by applying convolutional and multi-layer neural networks on learned representations of amino acids. Specific attention was given to the termini of input amino acid sequences as motifs in these regions indicate cleavage sites.
   - Improved performance over baseline methods by 5-10% with respect to NDCG@*k* and Precision@*k*.
3. **Stratifying patient outcomes following cochlear implantation through unsupervised learning**
   - Utilized unsupervised imputation and clustering techniques to identify subgroups of patients most and least likely to exhibit auditory improvements following cochlear implantation.
   - Identified preoperative measures capable of discerning patients who show significantly reduced post operative improvement, in terms of the reliable change index of target measures.
4. **Developing and optimizing applications for auditory studies**
   - Built and optimized applications which used multi-media output and input to evaluate cognitive and auditory capacity in patients with cochlear damage and hearing loss.
   - Specifically, applications tested test talker discrimination and sentence-final word identification and recall (SWIR) and conducted lexical decision tasks.

Nationwide Children's Hospital, Inst. for Genomic Medicine *Columbus, OH, USA*

Rotation Student — Computational Genomics Group *Feb. 2021 - Apr. 2021*

1. **Measuring immune cell composition of tumors from complementary transcriptomic data**
   - Implemented a method to perform batch correction and eliminate instrument biases through the mapping of scRNA-seq TPM counts and Nanostring immunoprofiler gene expression counts to a shared feature space.
   - Reduced the Euclidean distance on a tSNE plot between a sample's RNA-seq profile and the same sample's Nanostring immunoprofiler profile from greater than 20 to less than 1.
   - Correctly identified immune cell populations and their proportions in tumor tissue using both bulk RNA-seq and Nanostring gene expression counts after mapping.

The Ohio State University *Columbus, OH, USA*

Rotation Student — Bioinformatics and Mathematical Biosciences Lab *Jun. 2020 - Aug. 2020*

1. **Analyzing single cell RNA-seq data**
   - Employed machine learning and other bioinformatics techniques to analyze single cell RNA-seq data.

Nationwide Children's Hospital, Inst. for Genomic Medicine *Columbus, OH, USA*

Bioinformatics Analyst — Computational Genomics Group *Jun. 2019 - Jun. 2020*

1. **Discovering pathogenic genes using genomic sequencing of multiplex families with left-sided cardiac defects**
   - Aimed to identify novel, pathogenic variant by analyzing and interpreting the whole genome and whole exome sequence data from over 1000 patients with congenital heart defects.
   - Discovered 3 novel, heritable, genetic variants in GATA6, CASZ1, and NRP2, which result in tricuspid atresia, bicuspid aortic valve, and ventral septal defect.
   - Expanded tertiary analysis annotations of whole genome sequence data to include heart enhancer regions. These identified additional candidates for functional analysis, specifically in subjects whose defects are currently idiopathic.
2. **Performing *in silico*, quantitative, functional analysis on variants of unknown significance**
   - Modeled protein folding of both wild type and mutant sequences to assess the impact variants may have on conformation, stability, and its microenvironment. Changes to these properties were in turn used to elucidate disease etiology.
   - Validated candidate pathogenic variants in GATA6, CASZ1, and LMNNA using these techniques.
3. **Developing WHAAP, a comprehensive wound healing automated analysis pipeline**
   - Built software capable of tracking and statistically analyzing the changes in cell density and migration in images taken across multiple time points and conditions.
   - Used this software to show that, *in vitro*, an identified NRP2 variant significantly reduces cellular migration during cardiac valve development, which explained the the molecular etiology of an instance of familial bicuspid aortic valve.

Research Aide – Computational Genomics Group *May 2017 - Jun. 2019*

1. **Creating and implementing QC metrics for genetic analysis of heritable, congenital heart defects**
   - Created a quality control pipeline to confirm that the whole genome and/or exome sequence data of more than 1000 subjects, participating in an on-going, multi-site genetic study on congenital heart defects, agreed with reported pedigree and demographic data.
   - Implementing the pipeline identified multiple sample swaps, non-paternity issues, and unknown consanguinities.

Publications

1. **Lawrence P**, Burns B, Ning X. S2P: Predicting drug perturbation-induced transcriptomic profiles for repurposing via latent space alignment. ['In Progress']
2. Patro A*, **Lawrence P***, Tamati T, Ning X, Moberly A. Use of unsupervised machine learning to stratify patient outcomes following cochlear implantation. ['In Progress']
3. **Lawrence P**, Ning X. Enhancing drug and cell line representations via contrastive learning for improved anti-cancer drug prioritization. *npj Precision Oncology,* ['In Review']
4. Xiang S*, **Lawrence P***, Peng B, ChienWei C, Kim D, Shen L, Ning X. (2023) Modeling Path Importance for Effective Alzheimer's Disease Drug Repurposing. *Pacific Symposium on Biocomputing,* https://arxiv.org/abs/2310.15211

** indicates authors contributed equally to the work*

5. Yasuhara J, Manivannan S, Majumdar U, Gordon D, **Lawrence P**, Zender G, Myers K, Stiver C, Bigelow A, Galantowicz M, Yamagishi H, McBride K, White P, Garg V. (2023) A novel pathogenic *GATA6* variant identified in a family with congenital heart disease, childhood-onset diabetes mellitus and necrotizing enterocolitis. *Pediatric Research,* https://doi.org/10.1038/s41390-023-02811-y
6. **Lawrence P**, Ning X. (2022) Improving MHC Class I antigen processing predictions using learned embeddings and cleavage site-specific kernels. *Cell Reports Methods, 2*(9), https://doi.org/10.1016/j.crmeth.2022.100293
7. Gordon D, Cunningham D, Zender G, **Lawrence P**, Penzaloza J, Lin H, Fitzgerald-Butt S, Myers K, Corsmeier D, Gaither J, Kuck H, Witjeratne S, Moreland B, Kelly B, Askwith C, Garg V, White P, McBride K. (2022) Exome Sequencing in Multiplex Families with Left-Sided Cardiac Defects has High Yield for Disease Gene Discovery. *PLOS Genetics, 18*(6), https://doi.org/10.1371/journal.pgen.1010236
8. Bennett J, Gordon D, Majumdar U, **Lawrence P**, Matos Nieves A, Myers K, Kamp A, Leonard J, McBride K, White P, Garg V. (2022) Use of Machine Learning to Classify High Risk Variants of Uncertain Significance in Lamin A/C Cardiac Disease. *Heart Rhythm, 19*(4), https://doi.org/10.1016/j.hrthm.2021.12.019
9. Yu B, Chen C, Qi R, **Lawrence P**, Zheng R, Wang X, Ma A, Gu H, Ma Q. (2021) scGMAI: a Gaussian Mixture Model for Clustering Single-Cell RNA-Seq Data Based on Deep Autoencoder. *Briefings in Bioinformatics, 22*(4), https://doi.org/10.1016/j.compbiomed.2020.103899
10. Wu Z*, **Lawrence P***, Ma A, Zhu J, Xu D, Ma Q. (2020) Single-cell techniques and deep learning in predicting drug response. *Trends in Pharmacological Sciences, 41*(12), http://doi.org/10.1016/j.tips.2020.10.004
11. Chen C, Zhang Q, Yu B, Yu Z, **Lawrence P**, Ma Q, Zhang Y. (2020) Improving protein-protein interaction prediction accuracy using XGBoost feature selection and stacked ensemble classifiers. *Computers in Biology and Medicine. 123*(103899), https://doi.org/10.1016/j.compbiomed.2020.103899

** indicates authors contributed equally to the work*

Presentations

1. Xiang S*, **Lawrence P***, ChienWei C, Kim D, Shen L, Ning X. (2024, January). *Modeling Path Importance for Effective Alzheimer's Disease Drug Repurposing* [Oral Presentation]. Pacific Symposium on Biocomputing, Waimea, HI, USA.
2. Moberly A*, **Lawrence P***, Tamati T, Ning X. (2023, July). *Use of Machine Learning to Predict Adult Cochlear Implant Benefit Using Reliable Change Index* [Poster Presentation]. Conference on Implantable Auditory Prothesis, Tahoe City, CA, USA.
3. Moberly A*, **Lawrence P***, Tamati T, Ning X. (2023, March). *Use of Machine Learning to Predict Adult Cochlear Implant Outcomes* [Podium Presentation]. American Auditory Society, Scottsdale, AZ, USA.
4. **Lawrence P**, Ning X. (2021, November). *Improving MHC Class I Antigen Processing Prediction via Representation Learning and Cleavage Site-Specific Kernel* [Podium Presentation]. American Medical Informatics Association Annual Symposium, San Diego, CA, USA.
5. **Lawrence P**. (2020, June). *WHAAP: Wound Healing Automated Analysis Pipeline* [Investor Pitch]. Nationwide Children's Hospital Technology Develop Fund Committee Meeting Columbus, OH, USA.
6. **Lawrence P**, Zender G, White P, McBride K. (2019, December). *Reducing the Burden of Genetic Functional Analysis: The Use of Image Recognition in Automating the Analysis of Wound Healing Assays* [Poster Presentation]. The Ohio State University's Center for Clinical and Translational Science's 9th Annual Meeting Columbus, OH, USA.
7. Bennett J, Gordon D, **Lawrence P**, McBride K, White P, Garg V. (2019, December). *Use of Machine Learning to Identify High Risk Variants of Uncertain Significance in Lamin A/C Cardiomyopathy* [Poster Presentation]. The Ohio State University's Center for Clinical and Translational Science's 9th Annual Meeting Columbus, OH, USA.
8. **Lawrence P**, Gordon D, Kelly B, White P. (2018, December). *Implementing Quality Control Metrics in Congenital Heart Defect Variant Analysis with Big Data* [Poster Presentation]. Nationwide Children's Hospital Summer Student Poster Competition Columbus, OH, USA.

** indicates authors contributed equally to the work*

Intellectual Property Disclosures

1. **Lawrence P**, Gordon D, White P, McBride K. (2020, March). *WHAAP: A Comprehensive Wound Healing Automated Analysis Pipeline to Reduce the Burden of Genetic Variant Functional Analysis* (Invention Disclosure). Office of Technology Commercialization, The Research Institute at Nationwide Children's Hospital.

Services

1. **Reviewer for conference paper submissions**
   - ACM The Web Conference. (2024)
   - IEEE International Conference on Bioinformatics and Biomedicine. (2024)
   - SIAM International Conference on Data Mining. (2023)
   - ACM SIGKDD Conference on Knowledge Discovery and Data Mining. (2022-present)
   - ACM International WSDM Conference. (2022-present)
   - Association for the Advancement of Artificial Intelligence. (2022-present)
   - The Medical Image Computing and Computer Assisted Intervention Society. (2022-present)
   - IEEE International Conference on Data Mining. (2021-present)
   - Special Interest Group on Information Retrieval. (2021-present)
2. **Reviewer for journal submissions**
   - Briefings in Bioinformatics. (2023)

Honors and Awards

1. **ISCB Fellowship for the 2024 Pacific Symposium on Biocomputing**. (2023). International Society for Computational Biology.
2. **LEAD Trainee Meeting Scholarship**. (2021). American Medical Informatics Association.
3. **Dean's List**. (2015-2019). Washington and Lee University.
4. **Honor Roll**. (2015-2019). Washington and Lee University.

Skills

Programming Languages	**Python**, **Bash**, **R**, LaTeX, MATLAB, Lua, Java
Development Tools	**Jupyter**, **Git**, Colab, Conda, Neovim, RStudio, PySpark
Machine Learning Tools	**Keras**, **TensorFlow**, Scikit-learn, SciPy Scikit-image, PyTorch, DeepChem
Visualization Tools	Bokeh, Plotly, R Shiny
Genomics Tools	UCSC Genome Browser, Ensembl, BLAST, PRIMUS
Protein Modeling Tools	RDKit, I-Tasser, PyMOL, PyRosetta, Chimera