

SARA E. POUR

github.com/spour ♦ sara.eslampour@mail.utoronto.ca

EDUCATION

University of Toronto / Temerty Faculty of Medicine

2020 - Present

Ph.D. Computational Biology (GPA: 4.0/4.0)

University of Toronto

2016 - 2020

H.B.Sc Molecular Genetics and Chemistry (GPA: 3.77/4.0)

EXPERIENCE

University of Toronto, Toronto, Canada

Dec 2020 - Present

PhD Researcher

- Developed and benchmarked sequence-to-activity machine-learning models to predict cleavage and polyadenylation (CPA) transcriptional cues from sequence data on a whole-genome scale.
 - Achieved a +5% improvement over previous methods and provided the first global census of CPA sequence elements in the human genome.
- Analyzed ChIP-seq datasets as part of an international consortium dedicated to mapping the binding specificities of all human transcription factors, including over 100 novel AT-hook proteins.
- Technologies: tensorflow, keras, numpy, matplotlib, scipy, scikit-learn, bash, R, L^AT_EX, git, slurm, SQL

University of Toronto, Toronto, Canada

Apr 2020 - Dec 2020

Research Fellow

- Investigated the cis-regulatory contribution of a DNA transposon to NF- κ B mediated inflammation over an evolutionary time period of 100 million years.
- Developed and implemented a convolutional neural network to accurately predict neural-specific splicing patterns in *C. elegans*.
- Conducted detailed analysis of scRNA-seq data from *C. elegans* to investigate the spatiotemporal dynamics of RNA-binding protein gene expression during nervous system development.
- Performed RNAcompete, molecular cloning, and SELEX experiments to study in vitro binding targets of transcription factors and RNA-binding proteins
- Technologies: pytorch, seaborn, jupyter notebook, bash, slurm, pandas, dask, Adobe Illustrator

PUBLICATIONS

1. **Pour, S.E.*** et al. Known sequence features can explain half of all human gene ends. *Nucleic Acids Research* (In press, 2023)
2. Ray, D., Lavery, K., Jolma, A., Nie, K., Samson, R., **Pour, S.E.*** et al., RNA-binding proteins that lack canonical RNA binding domains are rarely sequence-specific. *Scientific Reports* (In press, 2023).
3. Wang, L., Alizada, A., Rathnakumar, K., Khyzha, N., Tiegh, T., **Pour, S.E.*** et al., Multi-species analysis of inflammatory response elements reveals ancient and lineage-specific contributions of transposable elements to NF- κ B binding. *Genome Biology*. (In revision, 2023)
4. Lavery, K., Jolma, A., **Pour, S.E.*** et. al. PRIESSTESS: interpretable, high-performing models of the sequence and structure preferences of RNA binding proteins. *Nucleic Acids Research*.(2022)

INVITED PRESENTATIONS

- *ProteinQure Company Seminar* Feb 2023
Title: Analysis of 3'end Transcriptional Regulation Cues Using Machine Learning
- *Machine Learning in Computational Biology (MLCB) (Virtual)* Nov 2022
Title: Machine-Learning Models Capture Over 100,000 Cryptic CPA Sites Genome-Wide
- *RNA 3' end Formation and the Regulation of Eukaryotic Genomes (Oxford University)* Sep 2022
Title: Interpreting regulation of gene expression with Deep Learning
- *Departmental Seminar* Apr 2022
Title: Systematic Computational Interrogation of Gene 3' Ends

- *Molecular Genetics Retreat*
Title: AT-Hooks as an Under-Represented Class of DNA-binding Proteins

Sep 2019

TEACHING EXPERIENCE

- MMG1001: Foundational Genetics Approaches I (2022F)

MENTORSHIP

- Cameron Isaacs-Trepanier. *Currently: Ph.D. Student*
- Salman Basrai. *Currently: Ph.D. Student*
- Ali Fathi. *Currently: Ph.D. Student*
- Aiden Hiller. *Currently: undergraduate*

PROJECTS

- Developed a comprehensive Python suite of tools for the creation, evaluation, and interpretation of convolutional neural networks in genomic sequence classification tasks.
 - https://github.com/spour/tf_tools
- Implemented and trained a Variational Autoencoder (VAE) to uncover patterns in genomic sequences and engineer synthetic DNA, resulting in a novel data augmentation method for imbalanced classification tasks.
 - https://github.com/spour/VAEs_DNAseqs
- Designed and implemented a Generative Adversarial Network (GAN) to learn and generate polyphonic baroque music
 - https://github.com/spour/music_gan
- Built and trained a CNN to classify cellular components; visualized weight matrices and layer activations, and analyzed learning through saliency maps.
 - https://github.com/spour/computer_vision_cellular_components

SELECTED HONORS AND AWARDS

Jennifer Dorrington Graduate Research Award 2023 (\$2,000); School of Graduate Studies Conference Grant 2022 (\$1,200); Cecil Yip Doctoral Research Award 2021 (\$3,000); Faculty of Medicine Merit Entrance Fellowship 2020 (\$2,000); St. Michael's College Silver Medal for Exceptional Academic Standing 2020 (\$1,000); Undergraduate Research Opportunity Program (UROP) 2019 (\$5,000); ChemClub UROP 2019 (\$5,000); Hosinec Family Scholarship 2018 (\$7,000). Total: \$36,400.