

EDUCATION

University of California, Berkeley <i>PhD Student, Department of Electrical Engineering and Computer Science</i>	Berkeley, USA <i>Aug 2021 - Present</i>
University of Toronto <i>Masters in Computer Science</i> – Thesis: <i>Contrastive Learning of Protein Representations by Mutual Information Maximization</i>	Toronto, Canada <i>Jan 2019 – May 2020</i>
University of Waterloo <i>Bachelors of Science, Honours Science, Bioinformatics Option</i> – Thesis: <i>Interpreting Convolutional Neural Networks for Discovering Regulatory Motifs of Femur Growth</i>	Waterloo, Canada <i>Sept 2014 – May 2018</i>

RESEARCH EXPERIENCE

Prescient Design (Genentech) <i>Student Intern</i> – Generative methods for antibody design with biological priors.	New York, USA <i>June 2023 – Present</i>
UC Berkeley/Berkeley AI Research <i>PhD Student — Advisor: Pieter Abbeel</i> – Developing methods for ML-guided protein design.	Berkeley, USA <i>July 2021 – Present</i>
Google Brain <i>Student Researcher — Host: Andreea Gane</i> – Understanding protein functions using large language models.	Mountain View, USA <i>May 2022 – Dec 2022</i>
Insitro <i>Machine Learning Engineer III</i> – Engineering representations of microscopy image phenotypes for drug discovery. – Supporting ML-guided small molecule design through chemoinformatic analyses and developing novel representation learning strategies.	South San Francisco, USA <i>July 2020 – July 2021</i>
University of Toronto/Vector Institute <i>Masters Student — Advisors: Alan Moses, Marzyeh Ghassemi</i> – Developed self-supervised representations for proteins using contrastive mutual information maximization (<i>MLCB 2020, PLOS Computational Biology</i>). – Benchmarked self-supervised computer vision methods for microscopy images to address generalization challenges to natural covariate shifts (<i>NeurIPS 2019</i>). – Quantitative and qualitative evaluation of bias in contextual word embeddings of clinical notes (<i>Spotlight, ACM CHIL 2020</i>).	Toronto, Canada <i>Jan 2019 – May 2020</i>
Stanford University <i>Visiting Student Researcher</i> – Using domain adaptation methods to enable transcription factor binding prediction when evaluating on holdout cell lines.	Stanford, USA <i>Sept 2019 – Jan 2020</i>
Harvard Medical School/Boston Children’s Hospital <i>Intern, Research Computing — Advisor: Piotr Sliz</i> – Understanding genotype-phenotype relationships in childhood epilepsy. Interpreted important model features to seek novel disease-associated variants from whole exome (WES) data.	Boston, USA <i>Jun 2018 – Jan 2019</i>
University of Waterloo <i>Undergraduate Thesis Student — Advisor: Andrew Doxey</i> – Prediction of accessible chromatin regions for femur growth regulation. Reconstructed convolutional filters as a position-weighted matrix (PWM) with statistical matches in JASPAR, a database of known motifs.	Waterloo, Canada <i>Sept 2017 – May 2018</i>

- Used molecular dynamics (MD) and GROMACS to simulate enzyme-membrane interaction mechanisms of NDM-1, an enzyme which confers antibiotic resistance.

AWARDS

D. E. Shaw Graduate & Postdoc Women's Fellowship	2022
– Fellowship program for graduate & postdoc women in computational drug discovery	
Paula Hawthorn Fellowship	2021
– UC Berkeley Computing, Data Science and Society departmental fellowship.	
NSERC Postgraduate Scholarships – Doctoral program (PGS D) Award	2020
– Federal doctoral scholarship tenurable abroad, selected in the Committee for Computing Sciences.	
NSERC Canada Graduate Scholarships – Doctoral (CGS D) Award	2020
– Federal doctoral scholarship tenurable only at a Canadian institution [DECLINED].	
NSERC Michael Smith Foreign Supplement	2019
– Supports high-calibre Canadian graduate students in pursuing research abroad.	
Alexander Graham Bell Canada Graduate Scholarships – Master's (CGS M) Award	2018
– Federal research scholarship for high-calibre Master's research students.	
EPFL Scholarship of Excellence in Research	2017
– Sponsors students for research internship at EPFL.	
President's Scholarship of Distinction, Arebi Family Science Scholarship	2014
– Entrance scholarships, University of Waterloo.	
Royal Conservatory of Music (RCM)	2013
– ARCT Performer's Diploma in Piano.	

PREPRINTS AND PUBLICATIONS

R Boger*, **AX Lu***, S Chithrananda*, K Yang, P Skopintsev, B Adler, E Wallace, P Yoon, P Abbeel, J Doudna. TOPH: Adapting A Contrastive Question-Answering Framework for Protein Search. *ICML Workshop on Computational Biology*, 2023.

AJ Reddy, MH Herschl, S Kolli, **AX Lu**, X Geng, A Kumar, PD Hsu, S Levine, NM Ioannidis. Pretraining strategies for effective promoter-driven gene expression prediction. *bioRxiv*, 2023.

AX Lu, **AX Lu**, I Pritišanac, T Zarin, JD Forman-Kay, AM Moses. Discovering molecular features of intrinsically disordered regions by using evolution for contrastive learning. *PLOS Computational Biology*, 2022

S Kolli, **AX Lu**, X Geng, A Kumar, S Levine. Data-Driven Optimization for Protein Design: Workflows, Algorithms and Metrics. *ICLR Workshop on Machine Learning for Drug Discovery (MLDD)*, 2022.

C Dallago, K Schütze, M Heinzinger, T Olenyi, M Littmann, **AX Lu**, KK Yang, S Min, S Yoon, JT Morton, B Rost. Using protein sequence representations from deep learning to visualize and predict protein sets. *Current Protocols*, 2021.

AX Lu, H Zhang, M Ghassemi, AM Moses. Self-supervised contrastive learning of protein representations by mutual information maximization. *Machine Learning for Computational Biology (MLCB)*, 2020.

AX Lu, AX Lu, AM Moses. Evolution Is All You Need: Phylogenetic Augmentation for Contrastive Learning. *Machine Learning for Computational Biology (MLCB)*, 2020.

H Zhang*, **AX Lu***, M Abdalla, M McDermott, M Ghassemi. Hurtful Words: Quantifying Biases in Clinical Contextual Word Embeddings. *Spotlight, ACM Conference on Health, Inference, and Learning (CHIL)*, 2020.

AX Lu, **AX Lu**, W Schormann, M Ghassemi, DW Andrews, AM Moses. The Cells Out of Sample (COOS) dataset and benchmarks for measuring out-of-sample generalization of image classifiers. *Neural Information Processing Systems (NeurIPS)*, 2019.

M Abdalla, H Zhang, **AX Lu**, I Chen, M Ghassemi. Quantifying Fairness in a Multi-Group Setting and its Impact in the Clinical Setting. *NeurIPS Workshop on Fair ML for Health*, 2019.

AM Moses, AX Lu, **AX Lu**, M Ghassemi. Transfer Learning vs. Batch Effects: what can we expect from neural networks in computational biology? *Machine Learning for Computational Biology (MLCB)*, 2019.

AX Lu, **AX Lu**, AM Moses. Paired Cell Inpainting: A Multiple-Instance Extension of Self-Supervised Learning for Bioimage Analysis. *ICML Workshop on Self-Supervised Learning*, 2019.

J Ban, M Tadrous, **AX Lu**, EA Cicinelli, SM Cadarette. Diffusion of indirect comparison meta-analytic methods to study drugs: a systematic review and co-authorship network analysis. *BMJ Open*, 2018.

SERVICE

Co-organizer, ML Protein Engineering Seminar Series	2022 – Present
– Biweekly research seminars for the broader ML for protein engineering community.	
Research to the People, Core Team	2018 – 2020
– Non-profit connecting patients of rare genomic diseases to the academic community to collaborative research initiatives.	
Waterloo Residence Don	2016 – 2018
– Leader for the Velocity Residence (spin-off program of the Velocity start-up incubator) and first-year residences.	
Volunteer, Tosamaganga Hospital	2016
– Supported operations and shadowed physicians at a rural Tanzanian hospital	

REVIEWING

Machine Learning for Health (ML4H)	2020, 2021, 2022, 2023
Machine Learning for Computational Biology (MLCB)	2021
ICLR Workshop on AI4Science	2022
ICML Workshop on AI4Science	2022
NeurIPS Workshop on AI for Science	2021, 2022, 2023
NeurIPS Workshop on ML for Structural Biology	2022, 2023
NeurIPS Workshop on Distribution Shifts	2021, 2022
NeurIPS Workshop on Robustness in Sequence Modelling	2022

TEACHING

Teaching Assistant, Genetics (BIOL 239)	2016
---	------

MENTORSHIP

Sathvik Kolli	Undergrad, 5th year MS
Seyone Chithrananda	Undergrad