

EDUCATION

Columbia University *Sep. 2019 - May. 2021 (Expected)*
Bachelor of Engineering, Computer Science, GPA: 3.79/4.00

Colgate University *Sep. 2016 - May. 2019*
Bachelor of Arts, Physics, GPA: 3.88/4.00

Aquincum Institute of Technology *Spring 2019*
CS Study Abroad, affiliated with Budapest University of Technology and Economics (BME)

RESEARCH EXPERIENCE

Cell Image Segmentation - single-cell level spatial profiling *May 2020 - Present*
Computational Cancer Biology Lab, Columbia University
Supervisor: Dr. Elham Azizi

- Implemented and trained an end-to-end Feature Pyramid Networks for cell nuclei image segmentation
- Refined segmentation results for adjacent, cluttered cells with incorporation of geometrical features, multi-level watershed and morphological post-processing steps
- Transferred the segmentation knowledge learnt from *in vitro* images to *in situ* tissue images

Predicting intercellular Ligand-Receptor interactions *May 2020 - Aug. 2020*
Computational Cancer Biology Lab, Columbia University
Supervisor: Dr. Elham Azizi

- Performed preprocessing, clustering & cell-type inference on multiple liver tissue scRNA-seq datasets
- Calculated pairwise cluster distance within and across datasets for cell-type imputation, predicted cluster-wise ligand-receptor interactions via NicheNet

PICASSO: Copy Number Variation Inference *Sep. 2019 - Present*
Computational Biology Lab, Columbia University
Supervisors: Dr. Itsik Pe'er, Dr. Sitara Persad

- Assisted developing PICASSO, an iterative HMM model to infer copy number variation regions (CNVs) and intra tumor heterogeneity from single-cell RNA-seq data
- Inferred cellular clonal structures and phylogenetic lineage using EM-clustering, determined the stopping criterion of tree iterations guided by BIC and hierarchical clustering
- Implemented large-scale, full genome length randomized simulations synthesized from real scRNA-seq samples for benchmark
- Optimized the running time over 60% by parallelizing the core computation steps of the framework

NanoFusion: Oxford Nanopore Sequencing Analysis *May 2019 - Aug. 2019*
Computational Biomedicine Workgroup, BME
Supervisor: Dr. Péter Sárközy

- Designed pipelines to compare performances and accuracy metrics of multiple Oxford Nanopore Sequencing basecallers on NA12878 Chr1 data
- Performed PHRED-score biased Multiple Sequence Alignments to basecalling reads, reduced soft-clipped regions and improved the mapping accuracy

Seismic Data Analysis *May 2017 - Sep. 2017*
Geophysics Lab, Colgate University
Supervisor: Dr. Aubrey Adams

- Retrieved and analyzed Earthquake events in New York state from 2016 to 2017 with Python and AWK, refined data on earthquake arrival time, epicentral locations & depth
- Characterized and distinguished shallow earthquakes from mining events in the Adirondacks region

PROJECTS	Ancestral Sequence Reconstruction (ASR) Course: Computational Genomics <i>Supervisor: Dr. Itsik Pe'er</i>	Jan. 2020 - May. 2020
	<ul style="list-style-type: none"> · Implemented algorithms of among-site rate variations (ASRV) and site-specific constraints for marginal ASR, extending features to <i>TreeTime</i>, a python library for phylogenetic analysis · Evaluated reconstruction results on public mitochondrial datasets, improved the overall root state likelihood scores comparing against the baseline model 	
	Gene-level Analysis of scRNA-seq data Course: Machine Learning for Functional Genomics <i>Supervisor: Dr. David Knowles</i>	Oct. 2019 - Dec. 2019
	<ul style="list-style-type: none"> · Predicted cell types of human SC-islets dataset with XGBoost model, extracted cell-type specific differential expression features with SHAP · Constructed gene co-expression networks from the trained model, evaluated the reliability of the networks with WGCNA 	
	Phylogenetics tree with 18S rRNA sequences Course: Computational Biology and Medicine <i>Supervisor: Dr. Péter Sárközy</i>	Mar. 2019 - May. 2019
	<ul style="list-style-type: none"> · Refined a phylogenetic tree with progressive alignments & hierarchical clustering on selected Vertebrata 18S rRNA sequences · Evaluated the results with NCBI taxonomy database & MUSCLE package 	
MANUSCRIPTS	Yinuo Jin* , Alexandre Toberoff*, and Elham Azizi. “Transfer learning framework for cell segmentation with incorporation of geometric features” Accepted to <i>NeurIPS 2020 Workshop on Learning Meaningful Representations of Life (LMRL)</i>	
	Aubreya Adams, Yinuo Jin , Monica Dimas, and Isabel Dove. “Leveraging USArray and Regional Networks to Characterize Natural and Mining Seismicity at the Intersection of the Grenville and Taconic Provinces” Submitted to <i>Seismological Research Letters: Eastern Section</i>	
RELEVANT COURSEWORKS	Computer Science: <ul style="list-style-type: none"> · Computational Genomics · Computational Biology and Medicine · Artificial Intelligence · Analysis of Algorithms Statistics & Biology: <ul style="list-style-type: none"> · Probability Theory · Biostatistics for Engineers · Genetics 	<ul style="list-style-type: none"> · Machine Learning for Functional Genomics · Data Mining · Machine Learning · Advanced Programming
HONORS & AWARDS	Sigma Pi Sigma (SPS) Physics Honor Society Charles A. Dana Scholars (top 5% seniors, juniors and sophomores) Colgate University Douglas Rankin '53 Fellowships Colgate University Hackett-Rathmell 1968 Memorial Fund	2020 2018 2017 2017
SKILLS	Languages: Python, C++, Java, Bash, R Framework/Tools: Pytorch, GATK, Samtools, Minimap2	