Albert Tian Chen

Interdisciplinary researcher with 4+ years of statistics, data science and data visualization experience, as well as hands-on experience with bench biochemistry, sequencing, and mass-spec proteomics.

Phone: 339-927-0113
Email: chena@broadinstitute.org
Portfolio: https://atchen.me
GitHub: https://github.com/atc3
ORCID ID: 0000-0002-5387-0208
LinkedIn: https://linkedin.com/in/achen339

EDUCATION

Northeastern University, *College of Engineering*, Boston, MA   *9/2015 – 05/2020*
Candidate for Bachelor of Science in *Bioengineering*

- 3.7 / 4.0 GPA

PUBLICATIONS

Chen AT, Altschuler K, Zhan SH, Chan AY, Deverman BE. COVID-19 CG: Tracking SARS-CoV-2 mutations by locations and dates of interest. *bioRxiv* (2020) https://doi.org/10.1101/2020.09.23.310565
Website: https://covidcg.org

Chen AT, Franks A, Slavov N. DART-ID increases single-cell proteome coverage. *PLOS Computational Biology* (2019) https://doi.org/10.1371/journal.pcbi.1007082
Website: https://dart-id.slavovlab.net/

Huffman RG*, **Chen AT***, Specht H*, Slavov N. DO-MS: Data-Driven Optimization of Mass Spectrometry Methods. *Journal of Proteome Research* (2019) https://doi.org/10.1021/acs.jproteome.9b00039.
*Co-first authors
Website: https://do-ms.slavovlab.net/

RESEARCH EXPERIENCE

Vector Engineering Team, Broad Institute of MIT and Harvard (Link)
Associate Computational Biologist, Cambridge MA   *5/2019 – Present*

- With a Broad Shark Tank-winning team, accelerate discovery of promising recombinant adeno-associated viruses (AAV) for gene therapy in the central nervous system, by developing a novel computational method of identifying functional variants from large screens of randomized gene libraries.
- Leading development of COVID-19 CG, a web application for filtering and visualizing mutations in over 250,000+ SARS-CoV-2 genomes in a location- and date-specific manner. Assisted in securing outside funding for the project from a leading COVID-19 vaccine developer.
- Investigate AAV production dynamics in cell lines, with 10X Genomics single cell transcriptome sequencing. Led a group of 6 Northeastern University undergraduate students in analyzing this data as part of their Capstone project.
- Running weekly programming tutoring sessions for research assistants in the lab.
- Characterize cell lines and primary cells used in AAV selection assays with bulk RNA-seq.
- Assist lab members with library prep, sequencing, and data analysis.
- Generate alternative encodings of AAV variants, for use in models, with RaptorX and UniRep.
- Write efficient and multithreaded tools for extracting AAV variants from sequencing data, filtering on quality control metrics, and correcting for sequencing error.

- Built a user-friendly and code-free data processing pipeline for next-generation sequencing data, and store/query results with PostgreSQL, using the Google Cloud Platform.
- Mined complex atomic interactions in AAV capsid proteins with PyMOL scripts, and visualize them in an interactive application built with Python and Bokeh.

Slavov Laboratory, Northeastern University (Link)
Research Assistant, Boston MA   *7/2016 – 05/2020*

- First author on DART-ID project, a Bayesian framework for increasing peptide identification rates in mass-spec proteomics experiments, by incorporating peptide retention time as additional evidence. Wrote software, drafted initial manuscript and figures, and responded to reviewers
- Co-first author on DO-MS, a modular, extensible, and automated data analysis and visualization platform for rationally optimizing mass spectrometry methods.
- Led P-SCoPE (Single cell phosphoproteomics by mass spectrometry), a modification of the Single Cell Proteomics by Mass Spectrometry (SCoPE-MS) method (http://scope2.slavovlab.net/). Designed and ran LC/MS proteomics experiments, from cell culture to sample prep to instrument operation. Analyzed data and presented findings within lab meetings and at university events.
- Won two university awards ($3000 and $5000) to support research on the P-SCoPE project.
- Ran weekly programming office hours for research assistants in the lab

Waters Corporation, **Separations R&D**
Chemistry Co-op, Milford MA   *7/2017 – 12/2017*

- Prototyped high-throughput LC/MS analysis workflow that doubled throughput to 4800 samples/day, through hardware and firmware modifications to existing instrumentation
- Implemented efficient low-level data processing software in C++ and Armadillo
- Presented prototypes to upper management and solicited feedback from customers
- Performed maintenance, repairs, and installed custom modifications on Acquity LC instruments and Waters mass spectrometry instruments

AWARDS / FELLOWSHIPS

NSF Graduate Research Fellowships Program (GRFP), Honorable Mention	*2020*
Summer Scholars Independent Research Fellowship, Northeastern University ($5000)	*2019*
Advanced Research/Creative Endeavor Award, Northeastern University ($3000)	*2018*
ASBMB Active Site Poster / Travel Award Winner	*2018*
Bioengineering Research Achievement Award, Northeastern University	*2018*
Dean's Merit Scholarship	*2015 – 2020*
Dean's List	*2015 – 2020*

TECHNICAL SKILLS

Programming*:*
Advanced: Python, R, JavaScript, HTML/CSS
Proficient: MATLAB, LaTEX, SQL, PL/pgSQL, bash

Software / Libraries:

Python	multiprocessing, numpy, scipy, pandas, scikit-learn, matplotlib, snakemake, networkx
R	dplyr/tidyverse ecosystem, shiny, ggplot
Sequencing	bcl2fastq, fastqc, bowtie2, STAR, kallisto, samtools, pysam, IGV, cellranger
JavaScript	node.js, React, MobX, Vega, Webpack
HPC	UGER, Slurm, Google Cloud Platform (Compute Engine, Storage, SQL, Functions)

Other   git, Linux, STAN, PostgreSQL, PyMOL, SnapGene, MaxQuant, Adobe Photoshop, Illustrator

Laboratory:
Mammalian cell culture (suspension and adherent), Vector assembly + cloning, Transfection, PCR, qPCR, ddPCR, FACS (Sony SH800z, MA900), Illumina NGS sample prep, Bulk RNAseq, Single cell RNAseq (10X), Proteomics sample prep, LC/MS operation + maintenance

Hands-On:
Shop and hand tools, Basic bench electronic equipment, Computer building and repair

OTHER EXPERIENCE

Journal of Medical Insight (JOMI), *Lead Web Developer*, Boston MA   *6/2014 – 9/2016*
- Led development of http://old.jomi.com with no prior web development experience, in 3 months
- Rebuilt site in Node.js, MongoDB, and React as https://jomi.com a year later, again in 3 months
- Assisted in filming live surgeries with video team

Bikeway Source, *Sales Associate/Bike Mechanic*, Bedford MA   *8/2012 – 6/2014*
- Performed bike repair, facility construction, sales, and customer service

PRESENTATIONS

Poster: **<u>Chen AT</u>**, Chan YA, Deverman BE, Elzahraa-Eid F. Clustif: Motif-based clustering of short peptides for protein engineering applications. Presented at 15th annual Broad Retreat (Broad Institute of MIT and Harvard), December 2019, Boston MA
PDF: https://atchen.me/images/clustif_poster_v10.pdf

Poster: **<u>Chen AT</u>**, Perlman DH, Emmott E, Huffman RG & Slavov N. P-SCOPE: Carrier-enabled low-input phosphoproteomics by mass spectrometry. Presented at American Society of Biochemistry and Molecular Biology (ASBMB) Active Site Northeast Regional Meeting, November 2019, Boston MA
PDF: https://atchen.me/images/PSCOPE_poster_v2.pdf

Poster: **<u>Chen AT</u>**, Perlman DH, Emmott E, Huffman RG & Slavov N. P-SCOPE: Carrier-enabled low-input phosphoproteomics by mass spectrometry. Presented at Northeastern University Creative Endeavor and Research Experience Symposium (CERES), October 2019, Boston MA
PDF: https://atchen.me/images/NUSPURF_Vertical_2019.pdf

Poster: **<u>Chen AT</u>**, Perlman DH, Emmott E, Huffman RG & Slavov N. P-SCOPE: Carrier-enabled low-input phosphoproteomics by mass spectrometry. Presented at Northeastern Bioengineering Research Symposium, June 2019, Boston MA
PDF: https://atchen.me/images/PSCOPE_poster_v1.pdf

Presentation: **<u>Chen AT</u>**, Franks A & Slavov N. DART-ID Increases Single-Cell Proteome Coverage. Presented at Single Cell Proteomics (SCP) Conference, June 2019, Boston MA
Video: https://www.youtube.com/watch?v=IWeasIXs5F0
Slides: https://atchen.me/images/DART_SCP_v3.pptx

Poster: **<u>Chen AT</u>**, Franks A & Slavov N. DART-ID Increases Single-Cell Proteome Coverage. Presented at Experimental Biology (EB) as part of the American Society of Biochemistry and Molecular Biology (ASBMB), April 2019, Orlando FL
PDF: https://atchen.me/images/DART_poster_v6.pdf

Poster: **<u>Chen AT</u>**, Franks A & Slavov N. DART-ID Increases Single-Cell Proteome Coverage. Presented at Northeastern Research, Innovation, and Scholarship Expo (RISE), April 2019, Boston MA
PDF: https://atchen.me/images/DART_RISE_Poster.pdf

Poster: **<u>Chen AT</u>**, Franks A & Slavov N. DART-ID Increases Proteome Coverage with Data-Driven Alignment of Retention Times. Presented at American Society of Biochemistry and Molecular Biology (ASBMB) Active Site Northeast Regional Meeting, November 2018, Boston MA
PDF: https://atchen.me/images/dart_poster_v2_3.pdf

Poster: **<u>Chen AT</u>**, Franks A & Slavov N. DART-ID Data-Driven Alignment of Retention Times for Peptide Identification increases peptide coverage by 75%. Presented at Single Cell Proteomics (SCP) Conference, June 2018, Boston MA
PDF: https://atchen.me/images/scp_poster_v6.pdf

RELEVANT COURSEWORK

Biology:
Molecular Engineering, Cellular Engineering, Tissue Engineering, Genetics and Molecular Biology

Chemistry:
Organic Chemistry 1 & 2, Biochemistry

Math / Statistics:
Supervised Machine Learning (Audited), Unsupervised Machine Learning (Audited), Dynamic Modeling in Systems Biology, Differential Equations, Statistics (MIT OCW 18.05), Linear Algebra (MIT OCW 18.06)

Physics:
Physics 1 & 2, Statics and Dynamics, Transport and Fluids, Circuits and Signal Processing