Xavi Loinaz

email xloinaz@broadinstitute.org
phone (650) 380-3875
website xaviloinaz.com

Education

Brown University, Honors Sc.B. in Computational Biology, A.B. in Engineering (Sept. 2017 - May 2021):

❖ **Computational Biology major GPA:** 3.81

❖ **Relevant coursework (29 STEM courses taken in total):** Deep Learning, Deep Learning in Genomics, Machine Learning, Data Science, Computational Molecular Biology, Accelerated Introduction to Computer Science (data structures/algorithms course sequence accelerated into one semester), Honors Statistical Inference, Recent Applications of Probability and Statistics, Statistical Inference in Genomics, Signal Processing, Digital Electronics Systems Design, Neuroengineering, Electrical Circuits and Signals, Electricity & Magnetism, Genetics, Organic Chemistry, Biochemistry, Intro to Neuroscience, Linear Algebra, Multivariable Calculus, Ordinary Differential Equations, Partial Differential Equations

❖ **Activities & Communities:** Brown Healthcare Investment Group, Brown Ground Breakin' (breakdancing club), Brown-RISD Catholic Community, Intramural Softball

Coursera: AI for Medicine Track (May 2020 - June 2020):

❖ Completed AI for Medical Diagnosis course

Treehouse: iOS Development in Swift (July 2015 - June 2017):

❖ Took a bunch of courses in Treehouse's iOS Development track

Engineering Summer Academy at University of Pennsylvania in Biotechnology (July 2015):

❖ Took three-week compressed college credit course in biotechnology with a laboratory component.

Research Experience

Associate Computational Biologist II, I - Prof. Gad Getz Lab (officially co-advised by Prof. Esther Rheinbay starting Dec. 2023), Broad Institute of MIT and Harvard (July 2021 - Present):

❖ Lab's **lead analyst** for the NCI's Clinical Trial Sequencing Project on diffuse large B-cell lymphoma (DLBCL) to characterize DLBCL and understand its treatment response. Key contributions for this project include: **finding a prominent mutational signature not previously linked specifically to DLBCL**, generating mutation/copy number/structural variant calls, **identifying and patching 3 bugs** in the Getz Lab's mutation/copy number/structural variant pipelines, **devising and applying a methodology to filter NovaSeq sequencing artifact for Broad Institute-sequenced cohorts** (also applied to other cohorts), identifying/validating low-purity tumor samples unsuitable for downstream analysis, and **finding novel structural variant driver gene candidates for DLBCL**.

❖ Lab's **lead analyst** for the NCI's Human Cancer Models Initiative. Performed analysis to validate model and tumor concordance via mutational signatures, generated structural variant calls for 1,000+ cases, and performed quality control analyses of 1,000+ genomes and exomes. Manuscript to be submitted to **Nature** by the end of November 2024.

❖ **Conceptualized and developed from scratch** (initially as an unfunded side project) SSVGAR, a model to detect **significant structural variant (SV) drivers in cancer cohorts, outperforming existing methods** (published in Nature Communications in 2022) in the field. Initially **recapitulated ~10 known driver candidates** in CTSP DLBCL cohort. **Sole first-author manuscript** in preparation for **Nature Methods**. Interviewed, hired, and continuing to mentor paid **undergraduate intern** on project since June 2023.

❖ Applied SSVGAR to 1,000+-patient multiple myeloma cohort and **found 14 candidate SV driver hits, recapitulating known multiple myeloma drivers. Such candidate drivers were integrated into an "MM-like score"** predicting the likelihood of progression to full-blown multiple myeloma from precursor conditions. Found enrichment of either AID or APOBEC mutational signatures for specific SV driver candidates. Manuscript under revision for **Nature Genetics**.

❖ Helped to finalize mutation/copy number/structural variant/quality control/loss-of-Y-chromosome-calling pipelines for **processing of 8,529 whole genomes** from The Cancer Genome Atlas Program (TCGA), and applied such pipelines to 71 uveal melanoma cases. **Developed wrapper** to run pipelines together. **Devised a new transcript override list (over 50 times smaller)** for the widely-used annotation tool Funcotator to fix numerous annotation issues (**with my list soon-to-be-used by the Broad Institute's Genomics Platform** to regularly annotate their clinical cases). **Mentored undergraduate intern** on improving SV pipeline and **integrating NovaSeq sequencing artifact filter I created**.

Research Experience (continued) 🔬

Associate Computational Biologist II, I - Prof. Gad Getz Lab (officially co-advised by Prof. Esther Rheinbay starting Dec. 2023), Broad Institute of MIT and Harvard (July 2021 - Present), continued:

❖ **Took it upon myself** to develop dRanger2, an **improved** version of the SV-calling algorithm dRanger that aims to increase sensitivity of detection via integration of split-read data. **Implemented and successfully tested dRanger's main algorithm** in Python, and currently **mentor undergraduate intern** on the project.

❖ **Rewrote and maintain** the **50+-person lab**'s in-use structural variant-calling pipeline. **Fixed 4 bugs** in the pipeline and implemented a scatter/gather parallelization of the SV-calling tool SvABA, **reducing SvABA's runtime by a factor of 10** as well as reducing compute costs.

❖ Became **maintainer** of lab's quality control pipeline. **Fixed 4 bugs** and further developed pipeline for ease-of-use (mostly pertaining to file localization).

❖ **Contributed to various other manuscripts**, including the NCI's Multicentric Italian Lung Detection project (produced structural variant calls), the NCI's Clinical Trial Sequencing Project on kidney renal clear cell carcinoma (ran quality control on tumor and normal exomes, as well as sorted out data file location discrepancies for the **entire consortium**), the NCI's project on cancer subtype classification using compact feature sets (produced tables summarizing model results), and a collaboration on DLBCL treatment relapse (**mentored undergraduate student** for running pipelines).

Undergraduate Researcher in Machine Learning - Prof. Ritambhara Singh Lab, Brown University Center for Computational Molecular Biology (June 2020 - Oct. 2021):

❖ Worked with PhD student Jeremy Bigness to develop a novel graph convolutional neural network model (GC-MERGE) to predict gene expression from epigenetic marks integrating long-range regulatory interactions, resulting in **second-author journal publication**. Conducted hyperparameter tuning, implemented and ran baselines, interpreted results. Also **caught bug in data processing** for neural network features, and **caught and addressed** initial lack of standardization for test set.

❖ **Conceptualized and implemented** an **improved** version of the above model (featured in my **honors senior thesis**) which **outperformed it for all tested cell lines** in terms of AUROC, as well as **tied or outperformed all baselines** featured in the GC-MERGE paper for all tested cell lines in terms of AUROC.

❖ Worked on clinical project to predict gene expression in glioblastoma based off various -seq data. **Led two more junior students** in development and testing of various machine learning models, and presented bi-weekly to clinical collaborator Dr. Nikos Tapinos.

Undergraduate Researcher - Prof. David Borton Lab, Brown University Department of Engineering (Sept. 2019 - Oct. 2020):

❖ Had **my own project** where I developed a proof-of-concept for data synchronization between video and various electrophysiological sensors for studies in the lab. Read in and decoded LTC signals from Timecode Systems devices on an Arduino and wrote software to test synchronization.

❖ **Resolved a problem** in drift between GoPro and behavioral sensor data for one of the lab's projects — an ongoing issue for **around a year which the graduate researcher had not resolved.**

❖ Wrote up 14-page (single-spaced) report on my project and a tutorial for future doctoral students intending to utilize my system.

Undergraduate Researcher - Prof. Vicki Colvin Lab, Brown University Department of Chemistry (June 2019 - Aug. 2019):

❖ Worked on two different research projects: one observing protein crystallization in the presence of nanoparticles, and one creating gadolinium-oxide MRI contrast agents for characterizing glioma.

❖ **Determined a way to synthesize and demonstrate** PAMPS-LA polymer of desired size, **an issue that had been unresolved in the lab for approximately a year**.

Science, Technology, and Reconstructive Surgery (STaRS) Program Research Intern - Prof. Jill Helms Lab, Stanford University Department of Plastic and Reconstructive Surgery (June 2018 - Sept. 2018):

❖ Performed histological stainings, prepared samples, assisted in tooth implant/extraction of mice, and various other wet lab work.

❖ **Solely developed a software side project** in Python that automatically quantifies histological staining in microscope images using OpenCV.

Research Intern - Neurosurgery Department, Kaiser Permanente, Redwood City, CA (April 2016 - May 2017):

❖ **Designed visual recognition software** using OpenCV to analyze video of deep brain stimulation neurosurgery. Produced a 17-page report on work.

Publications 📄

SSVGAR: A computational algorithm to discover likely driver structural variant genes within cohorts of tumor genomes (under preparation for Nature Methods; preprint imminent for bioRxiv)
❖ **X. Loinaz**, J. Dagan, C. Stewart, J. Alberge, J. Hess, E. Rheinbay, G. Getz

Human Cancer Models Initiative (under preparation for Nature by end of 2024)
❖ Author order not finalized yet

Genomic landscape of multiple myeloma and of its precursor conditions, and its clinical implications (revisions submitted for Nature Genetics)
❖ J. Alberge*, A. Dutta*, A. Poletti*, T. Coorens, E. Lightbody, R. Toenges, **X. Loinaz**, (...), G. Getz**, I. Ghobrial**

Classification of non-TCGA Cancer Samples to TCGA Molecular Subtypes Using Compact Feature Sets (accepted to Cancer Cell)
❖ K. Ellrott*, C. Wong*, C. Yau*, M. Castro*, J. Lee*, B. Karlberg*, J. Grewal*, V. Lagani*, B. Tercan*, V. Friedl, T. Hinoue, V. Uzunangelov, L. Westlake, **X. Loinaz**, (...), A. Cherniack**, P. Laird**

Integrating Long-Range Regulatory Interactions to Predict Gene Expression Using Graph Convolutional Networks (Journal of Computational Biology, May 2022)
❖ J. Bigness, **X. Loinaz**, S. Patel, E. Larschan, R. Singh

Exploring Graph-Based Neural Networks for Modeling Long-Range Epigenetic Gene Regulation (Brown University Honors Senior Thesis, May 2021)
❖ **X. Loinaz** (advised by R. Singh)

A Brief Analysis of Predictive Pitching Metrics (FanGraphs Community Research Blog, Jan. 2018)
❖ **X. Loinaz**

Abstracts/Posters/Presentations 📊

A 1,000-genome map for multiple myeloma and its precursor stages, and its impact on clinical outcome (abstract for American Society of Hematology Annual Meeting & Exposition, Dec. 2024)
❖ J. Alberge* , A. Dutta*, A. Poletti*, (...), **X. Loinaz**, (...), G. Getz**, I. Ghobrial**

Leveraging compact feature sets for TCGA-based molecular subtype classification on new samples (oral presentation for American Association for Cancer Research Annual Meeting, April 2024)
❖ K. Ellrott*, C. Wong*, C. Yau*, M. Castro*, J. Lee*, B. Karlberg*, J. Grewal*, V. Lagani*, (...), **X. Loinaz**, (...), A. Cherniack**, P. Laird**

Genome Sequencing to Discover Drivers of Clonal Expansion in Smoldering Multiple Myeloma (abstract for American Society of Hematology Annual Meeting & Exposition, Dec. 2023)
❖ J. Alberge*, A. Dutta*, (...), **X. Loinaz**, (...), G. Getz**, I. Ghobrial**

SVGAR_sf: A novel computational algorithm to discover driver structural variant genes within cancer patient cohorts (poster for Annual Broad Institute Scientific Retreat, Dec. 2023)
❖ **X. Loinaz**, C. Stewart, J. Dagan, J. Alberge, A. Dunford, J. Hess, E. Rheinbay, G. Getz

Characterization of somatic events for diffuse large B-cell lymphoma in the Clinical Trial Sequencing Project (poster for Annual Broad Institute Scientific Retreat, Dec. 2021)
❖ **X. Loinaz**, Q. Zhang, J. Hess, J. Ma, D. Heiman, C. Stewart, G. Getz

Graph-Based Neural Networks for Modeling Long-Range Epigenetic Regulation (poster for Annual Brown CS Research Symposium, May 2021)
❖ **X. Loinaz**, J. Bigness, S. Patel, E. Larschan, R. Singh.

Optimizing PAMPS-LA Polymer Chain Size for Gadolinium-Oxide Nanoparticle MRI Contrast Agents for Detecting Glioma (poster for Brown Undergraduate Summer Research Symposium, Aug. 2019)
❖ **X. Loinaz**, X. Guo, J. Villanova, H. Kim, V. Colvin

The Design of Visual Recognition Software to Analyze Deep Brain Stimulation Video (poster for Advanced Authentic Research Symposium, May 2017)
❖ **X. Loinaz**, E. Sabelman

Awards/Academic Fellowships 🏆

National Science Foundation CSGrad4US Fellowship Winner (Aug. 2023)
❖ **One of 64 fellows chosen nationally** to receive **$159,000 of government-based funding** across 3 years for pursuing a computer science/informatics-based PhD.

Sigma Xi (April 2021)

Best Final Project Presentation for ENGN 0030: Introduction to Engineering (Dec. 2017)
❖ Recognized as part of a 4-person group in a class of 185 students.

National AP Scholar (July 2017)

Gunn High School Faculty Cup Nominee (May 2017)
❖ Nominated for award which 2 students from a ~500-student class receive for **exemplifying creative thinking, adaptability, and social/ethical responsibility.**

TSA TEAMS: Technology Student Association Tests of Engineering Aptitude, Mathematics, and Science 5th Place Team Nationally for Grades 11/12 (June 2016)
❖ Out of ~100 teams following regional qualification.

Gunn@YourService Grant for Buddies4Math (May 2016)
❖ **Applied for and received a $300 grant** to buy equipment for Buddies4Math, the volunteer tutoring program **I helped lead.**

Service/Volunteering ✋

RATalks Organizing Committee - BroadRATS Steering Committee (March 2022 - Present):
❖ Part of the organizing committee for arranging monthly talks for the Broad's early-career researcher community.
❖ **Personally conceptualized, recruited for, and helped organize and advertise** for an alumni talk with **~25 attendees** given by Grace Tiao, Broad's former Associate Director for Computational Genomics in the Medical and Population Genetics Program, on career advice and life lessons for early-career scientists.

Broad Institute Cancer Program Bootcamp Mentor (Feb. 2023):
❖ **Acted as assigned mentor** for a bootcamp program to onboard non-computational postdocs to computational methods in the Broad Institute's Cancer Program.

Brown Engineering Department Course Grader - Electrical Circuits and Signals (Jan. 2019 - May 2019):
❖ Graded homework assignments for Brown University's engineering course ENGN 0520: Electrical Circuits and Signals.

Classroom Lead - Buddies4Math (Sept. 2014 - April 2017):
❖ Volunteer and **classroom lead** for Buddies4Math tutoring program for underprivileged elementary school students. Worked to improve students' math skills and enthusiasm for math, coordinate volunteering logistics, and recruit new volunteers. Sometimes **tutored in Spanish** for Spanish-speaking students.

Mentorship 👥

Undergraduate research students mentored:
❖ **Johnathan Dagan, NYU '25 (June 2023 - Present):** Interviewed and hired Johnathan to help me work on SSVGAR.
❖ **Audrey Li, MIT '23 (May 2023 - August 2023):** Assigned to be Audrey's official mentor while her boss recovered from major surgery. Oversaw her running genome characterization pipelines for DLBCL relapse project.
❖ **Zachary Everton, BYU '25 (Sept. 2023 - Present):** Mentored Zach on pipeline development for the processing of 8,529 TCGA whole genomes, as well as development of dRanger2.

Selected Projects 📁

True Shooting Percentage Predictor (2021):
- ❖ Personal project. Worked to develop a model using multiple regression to best predict players' shooting efficiency year-to-year in the NBA.

Are.na Recommendation System (2020):
- ❖ Final project for Brown's upper-level CS course Data Science (CSCI 1951A). Helped develop a recommendation system for the art-sharing website Are.na.
- ❖ Used singular value decomposition to predict desirable content for website collaborators to contribute to. Generated visualizations of results.

TheGunnApp (2016-2018):
- ❖ **Led development of and maintained** mobile app (officially backed by my high school) with useful features for the Gunn High School community, such as the current schedule and staff directory. Originated as a project from my high school's iOS app development club which **I co-founded and was co-president of**.
- ❖ **Had over 5,000 total downloads** for iOS and Android, and was **regularly used by students, parents, and faculty.**

FanGraphs: A Brief Analysis of Predictive Pitching Metrics (2018):
- ❖ **Published article** on baseball analytics website FanGraphs on the predictive ability of certain pitching statistics in the MLB.

Unwynd (2017):
- ❖ Worked with a Brown University alumnus to develop a proof-of-concept platform as a **startup idea** where people can find and register for fitness classes. **Developed first iteration of prototype**, integrating with a Knack database.

Various Graphic Design (2016-2017):
- ❖ Developed the initial logo for Brown Speaks, a video series where Brown community members and alumni are interviewed by Brown students.
- ❖ Designed logos for sections of GunnTV, the morning announcements show at Gunn High School.

Skills/Interests ✏️

Programming
- ❖ Python (NumPy, SciPy, TensorFlow, PyTorch, scikit-learn, pandas), Swift (iOS development), Bash, R, SQL, Java, MATLAB, Lisp/Racket, Pyret, LaTeX, HTML, CSS, JavaScript, Arduino, Verilog

Wet Lab
- ❖ Mass spectrometry, NMR spectroscopy, immunohistochemistry

Graphic Design
- ❖ Adobe Illustrator, Adobe Photoshop

Interests
- ❖ Reading (read 120+ books since COVID pandemic; favorite authors: Min Jin Lee, Gary Shteyngart, Celeste Ng, Cal Newport), baseball (played 3 years of high school varsity), basketball, sports statistics, CrossFit, movies (favorite films: Past Lives, 500 Days of Summer, Toy Story 3), writing

Resume designed by Xavi Loinaz on Adobe Illustrator.