Taylor Joren

taylorjoren | taylorjoren | taylor [dot] m [dot] joren [at] gmail

Experience

Prescient Design / Genentech
Machine Learning Scientist June 2023 – Present

- Co-lead of Universal Molecular Encoder (UME), a 740M-parameter multi-modal encoder trained on multiple biological "languages", including over 1B protein, DNA, and small molecule sequences and structures. Lead evaluation, training data collection, post-training and manuscript writing. Implemented 25 evaluation tasks across AA/DNA/SM; improved affinity prediction by up to 17% via domain-adaptive post-training.
- Inventor of Conditional Walk–Jump Sampling (cWJS), a discrete, conditional generative model for protein sequences, trained on 85M+ antibody sequences from 3,000 unique clonotype modes. The method achieved 95%+ accuracy in generating from target modes while maintaining high novelty and diversity.
- Portfolio contributions: Generated and triaged 415 antibody designs across 7 targets, testing 2 generative methods and 1 repertoire mining method in active antibody portfolio projects. Improved pKD up to 15x over a starting hit.
- Core contributor and advisor to Next Generation Sequencing-guided machine learning (NGS-ML) task force. Lead research and engineering for 3 design rounds to improve antibody hit expansion.
- Co-authored 3 publications (incl. first-author ICLR LMRL 2025 spotlight); inventor on 1 U.S. patent.
- Engineering: Deployed three internal methods and tools across three APIs, impacting 5+ portfolio projects. Contributed 50+ peer-reviewed PRs over 8 internal packages for biological discovery.

Sanofi AI and Deep Analytics (AIDA)
Machine Learning Scientist Feb 2021 – May 2023

- Built end-to-end in-silico generation pipeline, accelerating discovery of 4 therapeutic antibodies.
- Mentored 4 interns/PhD students; delivered ML talks to 500+ colleagues.
- Contributed to 2 internal papers and 2 MIT JClinic collaborations applying ML to drug discovery.

Broad Institute of MIT & Harvard
Associate Computational Biologist (Machine Learning) Sept 2019 – Feb 2021

Built a computer vision pipeline integrating spatial Raman spectra processing, cell segmentation/classification, and inference of single-cell gene expression.

Massachusetts General Hospital
Research Intern (Machine Learning) May 2017 – Sept 2017

Combined statistical modeling with ML to improve prediction of cancer cell responses to molecular perturbations.

University College London Cancer Institute
Research Intern (Bioinformatics) Jun 2018 – Sept 2018

Identified correlation between lung adenocarcinoma and chromothripsis-like chromosome shattering using a 100-patient genomic dataset.

Education

2021 **Masters in Computer Science and Electrical Engineering**, Massachusetts Institute of Technology
Thesis: Interpreting Raman spectra using machine learning: towards a non-invasive method of characterizing single cells.
Advisors: Aviv Regev, Tommaso Biancalani

2019 **B.Sc. Computer Science and Molecular Biology**, Massachusetts Institute of Technology

Publications

Ismail, A. A., T. Oikarinen, A. Wang, J. Adebayo, S. Stanton, **Joren, Taylor**, J. Kleinhenz, A. Goodman, H. C. Bravo, K. Cho, and N. C. Frey (2025). "Concept Bottleneck Language Models for Protein Design". In: *Proceedings of the International Conference on Learning Representations (ICLR)*.

Joren, Taylor, S. A. Robinson, H. MohammadiPeyhani, S. P. Mahajan, E. Lee, S. Lillington, Q. Bei, S. Saremi, J. H. Lee, R. Bonneau, I. Hotzel, V. Gligorijevic, S. Kelow, and N. C. Frey (2025). "Guided Generation of B-cell Receptors with Conditional Walk–Jump Sampling". In: *ICLR 2025 Workshops: Learning and Modeling for Representation Learning (LMRL) & GenBio*. Spotlight.

Singh, R., C. Im, **Joren, Taylor**, P. Ponraj, Y. Qiu, H. Minoux, M. Wendt, B. Bryson, and B. Berger (2025). "Learning the Language of Antibody Hypervariability". In: *Proceedings of the National Academy of Sciences*.

Kobayashi-Kirschvink, K. J., C. S. Comiter, S. Gaddam, **Joren, Taylor**, E. I. Grody, J. R. Ounadjela, K. Zhang, B. Ge, J. W. Kang, R. J. Xavier, P. T. C. So, T. Biancalani, J. Shu, and A. Regev (2024). "Prediction of Single-cell RNA Expression Profiles in Live Cells by Raman Microscopy with Raman2RNA". In: *Nature Biotechnology*.

Frey, N. C., **Joren, Taylor**, A. A. Ismail, A. Goodman, R. Bonneau, K. Cho, and V. Gligorijevic (2023). "Cramming Protein Language Model Training in 24 GPU Hours". In: *NeurIPS 2023 Workshop on Machine Learning for Structural Biology (MLSB)*.

Riahi, S., J. H. Lee, **Joren, Taylor**, S. Wei, S. Jager, R. Olfati-Saber, A. Park, M. Wendt, H. Minoux, and Y. Qiu (2023). "Surface ID: A Geometry-aware System for Protein Molecular Surface Comparison". In: *Bioinformatics*.

Deist, T. M., A. Patti, Z. Wang, D. Krane, **Sorenson, Taylor**, and D. Craft (2019). "Simulation-assisted Machine Learning". In: *Bioinformatics*.

Patents

GUIDED MACHINE LEARNING ENABLED GENERATION OF ANTIGEN RECEPTORS (U.S. Patent App. No. 63/756594)

Awards

- Spotlight paper at ICLR LMRL Workshop 2025, *Guided Generation of B-cell Receptors with Conditional Walk-Jump Sampling (cWJS)*
- *cWJS* Paper selected for fast-track submission to Nature Biotechnology at ICLR GemBio Workshop 2025
- Spotlight paper at ICML 2024 Workshop on Accessible and EFficient Foundation Models for Biological Discovery, *Cramming Protein Language Model Training in 24 GPU Hours*
- Mt. Auburn Hospital Externship Fellow, Radiation Oncology (2017)
- Energy Solutions Foundation Scholarship (2013)
- Sterling Scholar of Utah: Douglas F. Bates Community Service, State Winner (2013)

Service

- Reviewer for ICLR 2025 LMRL workshop
- MITvote (2020): Led a group of graduate students to increase student voter turnout for the 2020 elections
- AddisCoder (2019): Created curriculum, gave lectures, and worked individually with students in summer computer science course for 180 top Ethiopian high school students
- MIT PKG Center (2018): Led a group of 20 student through an exploration of the Criminal Justice System in MA. Built community compost bins, volunteered at soup kitchens, tutored students, and led activities for after-school programs in Camden, New Jersey

Personal

Running and Triathlons

10K	San Diego Gaslamp 10K (1st place overall)	2025
Marathon	Boston Marathon	2025
Marathon	St. George Marathon (BQ)	2023
Marathon	Mt. Nebo Marathon	2021
Triathlon	Lake Perris, CA Olympic Triathlon	2023
Triathlon	San Diego International Triathlon	2022