

Stanford



Sujay Banerjee

Masters Student in Bioengineering, admitted Autumn 2025

Bio

BIO

Sujay Banerjee is a Master's student in Bioengineering at Stanford University, where he works in Dr. Russ Altman's lab on deep learning models for therapeutic binding affinity prediction.

Before Stanford, Sujay graduated summa cum laude from Middlebury College with a double major in Molecular Biology & Biochemistry and Computer Science. His undergraduate honors thesis introduced BASILISC, a transformer-based framework for structural variant analysis in genome sequencing, combining self-supervised learning with image-based genomic representations. He has conducted research at Harvard Medical School and Mass General Brigham on predictive modeling in endocrinology.

Sujay is passionate about bridging biology, computation, and medicine to advance personalized healthcare, and is specifically interested in AI-driven biotechnology.

EDUCATION AND CERTIFICATIONS

- Master of Science, Stanford University , Bioengineering (2026)
- Bachelor of Arts, Middlebury College , Computer Science and Molecular Biology/Biochemistry (2025)

Research & Scholarship

RESEARCH INTERESTS

- Data Sciences
- Research Methods

CURRENT RESEARCH AND SCHOLARLY INTERESTS

My research integrates bioengineering, computational biology, and artificial intelligence to advance precision medicine and therapeutic design. I am currently a Master's student in Bioengineering at Stanford University, working in Dr. Russ Altman's lab on deep learning models that predict therapeutic binding affinities for G protein–coupled receptors. This framework aims to improve our understanding of molecular recognition and inform drug discovery for GPCR-related diseases.

Previously, I conducted research at Harvard Medical School and Mass General Brigham under Dr. Alexander Turchin, developing AI models for longitudinal analysis of electronic health records to identify temporal patterns predictive of adverse outcomes in diabetes patients. This work emphasized the translational potential of machine learning in improving clinical decision-making and early risk detection.

As an undergraduate at Middlebury College, I completed an honors thesis titled “Self-Supervised Learning with Masked Images for Structural Variant Analysis in Short-Read Genome Sequencing.” In this project, I created BASILISC (BEiT-based Analysis with Self-supervised Image Learning In Structural variant Classification), a transformer-based framework that applied masked image modeling and discrete variational autoencoders to classify structural variants from genomic pileup images. BASILISC demonstrated how large-scale unlabeled sequencing data can train self-supervised models that generalize across variant types, improving the detection of clinically relevant genomic alterations.

Across these projects, my overarching goal is to bridge computational modeling with biological and clinical data to create scalable tools for understanding and treating disease. My broader interests include AI-driven drug discovery, computational genomics, systems pharmacology, and bioengineering approaches to personalized medicine. I am particularly motivated by how machine learning can enhance our ability to model complex biological systems—from DNA and protein structure to patient-level health outcomes—and ultimately translate these insights into therapies that improve human health.

LAB AFFILIATIONS

- Russ Altman, The Helix Group (10/5/2025)