

# Zach Collester

zrcollect@gmail.com | [linkedin.com/zachcollester/](https://www.linkedin.com/in/zachcollester/)

## EXPERIENCE

---

### STYLUS MEDICINE

Scientist I, Computational Biology

Cambridge, MA

Sep 2023 – Present

- Developed bioinformatics pipelines to identify engineered recombinase plasmid insertions in the human genome
- Conducted in-silico protein mutagenesis using advanced protein language models (ESM, AMPLIFY), generating a prioritized list of mutations for experimental validation
- Automated data workflows including sequencer-to-cloud integration and well-plate format processing, saving wet lab scientists approximately 5 hours per week
- Designed impactful data visualizations used in investor pitch decks to effectively communicate biological insights

### KSQ THERAPEUTICS

Scientist, Computational Biology

Cambridge, MA

Aug 2021 – Sep 2023

- Led computational analysis for a novel engineered TIL therapy leveraging scRNA-seq and TCR repertoire data to characterize tumor-infiltrating T cells, resulting in a co-first author on a [peer-reviewed publication](#)
- Designed and analyzed genome-wide CRISPR screening experiments to discover novel autoimmune drug targets, contributing to intellectual property development

### TREX BIO

Bioinformatics and Data Science Associate

San Francisco, CA (remote)

Jan 2021 – Aug 2021

- Conducted single-cell RNA-seq analysis on tissue Treg populations to identify key transcription factors and signaling pathways regulating Treg activity
- Collaborated effectively with cross-functional teams spanning multiple time zones in a fully remote environment

### SEQUENTIA BIOTECH

Bioinformatics Intern

Barcelona, Spain

Apr 2020 – Jul 2021

- Developed pan-cancer classification models across 20+ tissue types using RNA-seq data (TCGA, GTEx) and machine learning algorithms (XGBoost, neural networks)

## PROJECTS

---

### scRegNetWrapper

Jan 2021 – Aug 2021

- Created [open-source R package](#) to streamline workflows for three single cell regulatory network analysis tools (DoRothEA, PROGENy, pySCENIC) with added functionality for downstream analysis

### MACRO-COMPLEX MODELING OF BIOMOLECULES

Jan 2020 – Mar 2020

- Developed python-based application that utilizes pairwise interaction data from the Protein Data Bank (PDB) to generate models of biomolecules

## EDUCATION

---

### UNIVERSITAT POMPEU FABRA

Master of Science in Bioinformatics

Barcelona, Spain

Jul 2021

### BATES COLLEGE

Bachelor of Science in Neuroscience

Cumulative GPA: 3.7/4.0

Lewiston, ME

May 2019

## SKILLS

---

**Programming:** Advanced in R, Python, Bash; experience with workflow automation tools (Snakemake, Nextflow), and environment management (Anaconda)

**Data Types:** Extensive experience with scRNA-seq, bulk RNA-seq, TCR repertoire analysis, CRISPR screening, and amplicon sequencing

**Cloud Computing:** Proficient in AWS S3 and EC2