Luca Pinello, PhD

The focus of **the Pinello laboratory** is to use innovative computational approaches and cutting-edge experimental assays, such as CRISPR genome editing and single cell sequencing, to systematically analyze sources of genetic and epigenetic variation and gene expression variability that underlie human traits and diseases. The lab uses AI, machine learning and high-performance computing technologies to solve computationally challenging and Big Data problems associated with functional genomics and sequencing data analysis. Our mission is to use computational strategies to further our understanding of disease etiology and to provide a foundation for the development of new drugs and novel targeted treatments.

Pinello Laboratory

Basheer Becerra*
Logan Blaine*
Lucas Ferreira Da Silva, PhD
Justin Delano*
Zhijian Li, PhD
Jiachen Li†
Zain Patel, PhD
Luca Pinello, PhD
Jayoung Ryu*
Cameron Smith, MD, PhD
Michael Vinyard*
shared with Gad Getz Lab
Lingfei Wang, PhD

* PhD student
† Visiting PhD student

Epigenetic variability in cellular identity and gene regulation

We are studying the relationship between epigenetic regulators, chromatin structure and DNA sequence and how these factors influence gene expression patterns.

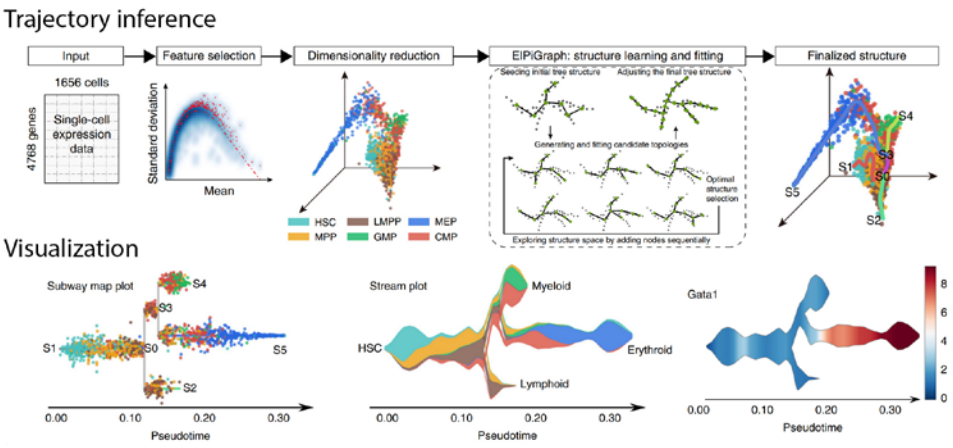
We recently developed an integrative computational pipeline called HAYSTACK HAYSTACK is a software tool (*https://github.com/lucapinello/Haystack*) to study epigenetic variability, cross-cell-type plasticity of chromatin states and transcription factor motifs and provides mechanistic insights into chromatin structure, cellular identity and gene regulation. By integrating sequence information, histone modification and gene expression data measured across multiple cell-lines, it is possible to identify the most epigenetically variable regions of the genome, to find cell-type specific regulators, and to predict cell-type specific chromatin patterns that are important in normal development and differentiation or potentially involved in diseases such as cancer.

Computational methods for genome editing

We embraced the revolution in functional genomics made possible by the novel genome editing approaches such as CRISPR/Cas9, base editing and prime editing by developing computational tools for the design, quantification of CRISPR edits and for the analysis of coding and non-coding tiling screens for functional genomics.

We have developed CRISPREsso2 (*http://crispresso2.pinellolab.org*), a software for the quantification of genome editing events that is now the standard de facto for the genome editing community. In collaboration with the groups of Daniel Bauer and Stuart Orkin, we applied our computational strategies to aid the development of several CRISPR screens for dissecting enhancer functionality in the blood system.

We have also recently developed a powerful tool called CRISPRme, which considers both SNPs and indel genetic variants to identify and prioritize off-target sites, offering a more comprehensive and accurate assessment of off-target risks. By utilizing CRISPRme, we discovered and validated a previously overlooked off-target site for a guide RNA (gRNA) targeting the BCL11A enhancer, currently being used in clinical trials for sickle cell disease and β-thalassemia.



STREAM on transcriptomic data from the mouse hematopoietic system. Top, STREAM workflow to recover hierarchical structure composed of curves approximating the inferred trajectories. Single cells are represented as circles and colored according to the FACS sorting labels. Bottom, from left to right, Subway map plot representation at single cell resolution; branches are represented as straight lines. The length of the branches and the distances between cells and assigned branches are proportional to the original representation in the 3D space. Rainbow plot: intuitive visualization to show cell type distribution and density along different branches. Relative expression of GATA1 in each branch using the reconstructed structure.

Exploring single cell gene expression variation in development and cancer

Cancer often starts from mutations occurring in a single cell that results in a heterogeneous cell population. Although traditional gene expression assays have provided important insights into the transcriptional programs of cancer cells, they often measure a combined signal from a mixed population of cells and hence do not provide adequate information regarding subpopulations of malignant cells. Emerging single cell assays now offer exciting opportunities to isolate and study individual cells in heterogeneous cancer tissues, allowing us to investigate how genes transform one subpopulation into another. Characterizing stochastic variation at the single cell level is crucial to understand how healthy cells use variation to modulate their gene expression programs, and how these patterns of variation are disrupted in cancer cells. We are developing tools to characterize cellular types and states at single cell resolution by using data from single cell transcriptomic or epigenomics data. For example, we developed STREAM (Single-cell Trajectories Reconstruction, Exploration And Mapping), an interactive computational pipeline for reconstructing complex cellular developmental trajectories from sc-qPCR, scRNA-seq or scATAC-seq data available at *http://stream.pinellolab.org*. This method can be used for disentangling complex cellular types and states in development, cancer, differentiation or in perturbation studies.

Selected Publications:

Chen H, Ryu J, Vinyard ME, Lerer A, **Pinello L**. SIMBA: single-cell embedding along with features. *Nat Methods*. 2023 May 29.

Cancellieri S, Zeng J, Lin LY, Tognon M, Nguyen MA, Lin J, Bombieri N, Maitland SA, Ciuculescu MF, Katta V, Tsai SQ, Armant M, Wolfe SA, Giugno R†, Bauer DE†, **Pinello L**†. Human genetic diversity alters off-target outcomes of therapeutic gene editing. *Nat Genet*. 2023 Jan;55(1):34-43.

Hsu JY, Grünewald J, Szalay R, Shih J, Anzalone AV, Lam KC, Shen MW, Petri K, Liu DR, Joung JK†, **Pinello L**†. PrimeDesign software for rapid and simplified design of prime editing guide RNAs. *Nat Commun*. 2021 Feb 15;12(1):1034.

Chen H, Albergante L, Hsu JY, Lareau CA, Lo Bosco G, Guan J, Zhou S, Gorban AN, Bauer DE, Aryee MJ, Langenau DM, Zinovyev A, Buenrostro JD, Yuan GC†, **Pinello L**†. Single-cell trajectories reconstruction, exploration and mapping of omics data with STREAM. *Nat Commun*. 2019. Apr 23;10(1):1903.

Clement K, Rees H, Canver MC, Gehrke JM, Farouni R, Hsu JY, Cole MA, Liu DR, Joung JK, Bauer DE†, **Pinello L**†. CRISPResso2 provides accurate and rapid genome editing sequence analysis. *Nat Biotechnol*. 2019 Mar;37(3):224-226.

Pinello L*†, Farouni R*, Yuan GC†. Haystack: systematic analysis of the variation of epigenetic states and cell-type specific regulatory elements. *Bioinformatics*. 2018; 34(11):1930-1933.

*Co-first authors
†Co-corresponding authors