Luca Pinello, PhD

The focus of **the Pinello laboratory** is to use innovative computational approaches and cutting-edge experimental assays, such as 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 machine learning, data mining and high performance computing technologies, for instance parallel computing and cloud-oriented architectures, to solve computationally challenging and Big Data problems associated with next generation 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

Matthew Canver, PhD
Huidong Chen*
(shared with Guo-Cheng Yuan Lab)
Kendell Clement, PhD
Rick Farouni, PhD
Jonathan Hsu*
(shared with Keith Joung Lab)
Luca Pinello, PhD
Qiuming Yao, PhD

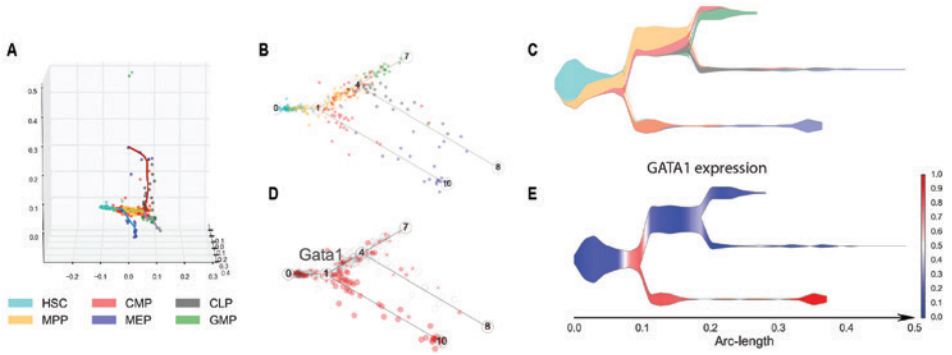
* PhD candidate

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 proposed an integrative computational pipeline called HAYSTACK (*https://github.com/lucapinello/Haystack*). HAYSTACK is a software tool 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

Recent genome editing technologies such as CRISPR/Cas9 are revolutionizing functional genomics. However computational methods to analyze and extract biological insights from data generated with these powerful assays are still in an early stage and without standards. We have embraced this revolution by developing cutting-edge computational tools to quantify and visualize the outcome of CRISPR/Cas9 experiments. We created a novel computational tool called CRISPResso (*http://github.com/lucapinello/CRISPResso*), an integrated software pipeline for the analysis and visualization of CRISPR-Cas9 outcomes from deep sequencing experiments, as well as a user-friendly web application that can be used by non-bioinformaticians (*http://crispresso.rocks*). In collaboration with Daniel Bauer's and Stuart Orkin's groups, we recently applied CRISPResso and other computational strategies to aid the development of an *in situ* saturation mutagenesis approach for



ARIADNE on transcriptomic data from the mouse hematopoietic system. A) Dimensionality reduction, reconstructed hierarchical structure composed of curves approximating the inferred trajectories. Single cells are represented as circles and colored according to the FACS sorting labels. B) Flat tree representation at single cell resolution; branches are represented as straight lines, (cells are represented as in A). The length of the branches and the distances between cells and assigned branches are proportional to the original representation in the 3D space. C) Rainbow plot: intuitive visualization to show cell type distribution and density along different branches. D) Single cell resolution expression pattern of GATA1, each circle is red filled proportionally to the relative expression of GATA1 in the whole population. E) Relative expression of GATA1 in each branch using the representation in C.

dissecting enhancer functionality in the blood system with the aim of developing potential therapeutic genome editing applications for hemoglobin disorders.

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 currently developing a method called ARIADNE to model the variability of gene expression at single cell resolution, and to reconstruct developmental trajectories (see illustrative image) using data from single cell assays such as scRNA-seq, multiplexed qPCR or sc-ATAC-seq. This method can be used for disentangling complex cellular types and states in development, cancer, differentiation or in perturbation studies.

Selected Publications:

Beyaz S*, Kim JH*, **Pinello L***, Xifaras ME, Hu Y, Huang J, Kerenyi MA, Das PP, Barnitz RA, Herault A, Dogum R, Haining WN, Yilmaz ÖH, Passegue E, Yuan GC, Orkin SH, Winau F. The histone demethylase UTX regulates the lineage-specific epigenetic program of invariant natural killer T cells. *Nat Immunol*. 2017 Feb;18(2):184-195.

Pinello L, Canver MC, Hoban MD, Orkin SH, Kohn DB, Bauer DE, Yuan GC. Analyzing CRISPR genome-editing experiments with CRISPResso. *Nat Biotechnol*. 2016 Jul 12;34(7):695-7.

Guo G*, **Pinello L***, Han X, Lai S, Shen L, Lin TW, Zou K, Yuan GC, Orkin SH. Serum-Based Culture Conditions Provoke Gene Expression Variability in Mouse Embryonic Stem Cells as Revealed by Single-Cell Analysis. *Cell Rep*. 2016 Feb 2;14(4):956-65.

Canver MC*, Smith EC*, Sher F*, **Pinello L***, Sanjana NE*, Shalem O, Chen DD, Schupp PG, Vinjamur DS, Garcia SP, Luc S, Kurita R, Nakamura Y, Fujiwara Y, Maeda T, Yuan G-C, Zhang F, Orkin SH & Bauer DE. BCL11A enhancer dissection by Cas9-mediated in situ saturating mutagenesis. *Nature*. 2015 Sep 16.

Wu JN*, **Pinello L***, Yissachar E, Wischhusen JW, Yuan GC, Roberts CW. Functionally distinct patterns of nucleosome remodeling at enhancers in glucocorticoid-treated acute lymphoblastic leukemia. *Epigenetics Chromatin*. 2015 Dec 2;8:53.

Pinello L*, Xu J*, Orkin SH, Yuan GC. Analysis of chromatin state plasticity identifies cell-type specific regulators of H3K27me3 patterns, *PNAS*. 2014 Jan 6; 10.1073/pnas.1322570111.

*Co-first authors