Stanford Center for Cancer Systems Biology - Seminar Series - 2024

WITH

Julia Salzman, PhD

Associate Professor, Stanford University, Dept. of Biomedical Data Science, Biochemistry, and Statistics (by Courtesy)

“Unified, ultra efficient genomic discovery without a reference with SPLASH”

Myriad mechanisms diversify the sequence content of DNA and of RNA transcripts. Currently, these events are detected using tools that first require alignment to a necessarily incomplete reference genome alignment in the first step; this incompleteness is especially prominent in diseases such as cancer. Second, today the next step in analysis requires as a custom choice of bioinformatic procedure to follow it: for example, to detect splicing, RNA editing or V(D)J recombination among many others. I will present a new statistics-first analytic approach to myriad problems in genomics —SPLASH (Statistically Primary aLignment Agnostic Sequence Homing). SPLASH performs unified, reference-free inference directly on raw sequencing reads without a reference genome or cell metadata and can be used for DNA or RNA-seq data. SPLASH is highly efficient and simple to run. As a snapshot of its discoveries, applying to 10,326 primary human single cells in 19 tissues profiled with SmartSeq2, we discover a set of splicing and histone regulators with highly conserved intronic regions that are themselves targets of complex splicing regulation, unreported transcript diversity in the heat shock protein HSP90AA1, and diversification in centromeric RNA expression, V(D)J recombination, RNA editing, and repeat expansions missed by existing methods. I will discuss these examples and their unpublished extensions to 10x genomics data, cancer transcriptomics, and other applications, time permitting.

Friday, January 19, 2024 10:00-11:00am

Clark Center, S360, 3rd floor

Refreshments will be provided

Click Here for Zoom Link