

Chen Institute Retreat 2023

Presenter: Delaney Sullivan

Poster title: Flexible parsing and preprocessing of technical sequences with splitcode

Abstract: Next-generation sequencing libraries are constructed with numerous synthetic constructs such as sequencing adapters, barcodes, and unique molecular identifiers. Such sequences can be essential for interpreting results of sequencing assays, and when they contain information pertinent to an experiment, they must be processed and analyzed. We present a tool called splitcode, that enables flexible and efficient preprocessing, parsing, and manipulation of sequencing reads. The splitcode program is free, open source, and available for download at <http://github.com/pachterlab/splitcode>. This versatile tool will facilitate simple, reproducible preprocessing of reads from libraries constructed for a large array of single-cell and bulk sequencing assays.