Tutorial 5 – ChIP-seq Analysis

MICB405 – BIOINFORMATICS – 2021W-T1

15 OCTOBER 2021

AXEL HAUDUC

What is MACS2

"Peak-calling" software

Calls areas of the genome enriched for your epigenomic mark of interest based on presence of FASTQ reads

Outputs BED (".narrowPeak" in MACS2 output) file which describe the genomic region where your mark is present
◦ This is distinct from sequence alignment!

Tutorial background

In this tutorial we will be analyzing a dataset generated as part of a project designed by Dr. Perona-Wright and Nicolette Fonsesca.

These data were generated from Naive T-cells purified from mice and subjected to ChIP-seq

Specifically, we will using ChIP-sq data for H3K27ac
◦ An active mark found at gene promoters and enhancer elements
◦ We will also be using an input control that represents a sequencing library generated from the sheared DNA fragments prior to immunoprecipitation
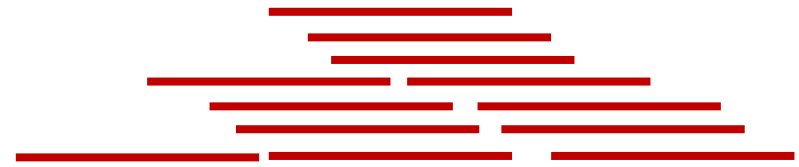
Variant Calling

Variant Calling

Single-nucleotide variant

Insertion or deletion (indel)

Single-nucleotide variant
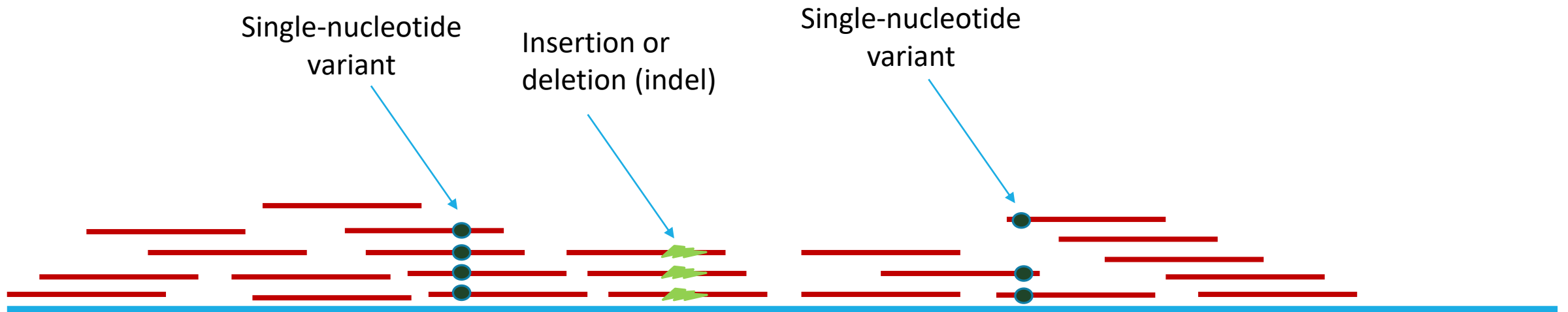
Let's find them!

Peak Calling
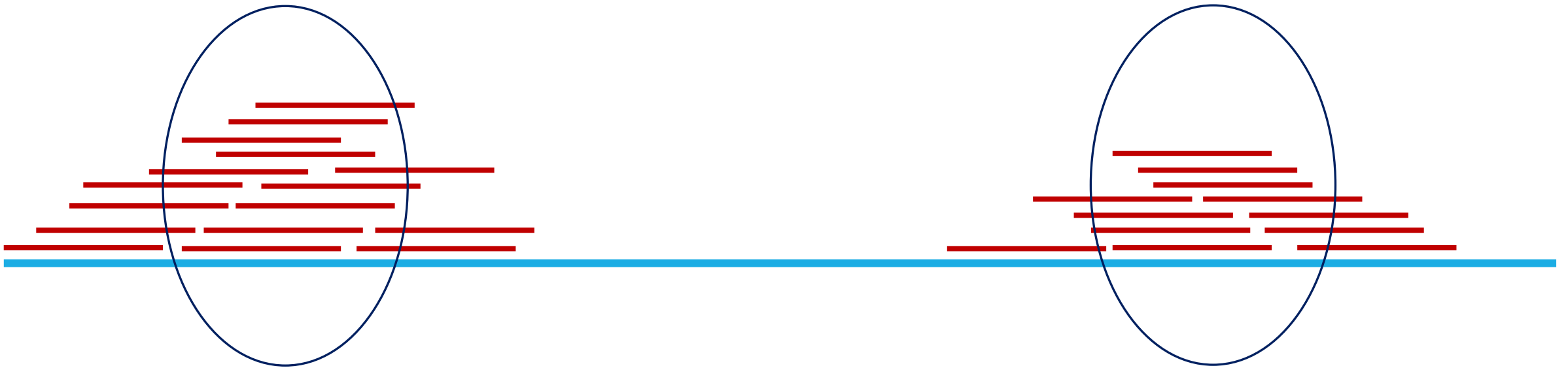
Peak Calling

These are peaks – their locations are linked to the epigenomic mark that is being studied