ChIP-seq data analysis

Instructor: Ariel Madrigal Aguirre

Registration link: https://forms.gle/ifVFTLyXja91E4ZZA

Hours of instruction: 4

Prerequisites:

- Access to a terminal in your computer (MacOS and Linux users will have it already)

- Proficiency with command line interface (e.g., know how to navigate between directories, execute programs)

- Knowledge of basic programming

- Install:

    - R(>=4.0.0) and Rstudio

    - Python >=3.6

    - Bowtie2 (https://bowtie-bio.sourceforge.net/bowtie2/index.shtml)

    - Samtools (http://www.htslib.org/)

    - MACS2 (https://pypi.org/project/MACS2/)

    - Deeptools (https://deeptools.readthedocs.io/en/develop/#)

    - Bedtools (https://bedtools.readthedocs.io/en/latest/)

*Students that register will receive a guide to help set up requirements (through Docker or manual installation) for the workshop beforehand.

*There will be a pre-workshop session online to solve software installation issues on Tuesday, November 22nd 2:00 −3:00 PM EST.

Summary:

ChIP-seq is a method to identify genome-wide DNA binding sites for a protein of interest. This worskhop will guide you through the basics of ChIP-seq analysis with hands-on exercises. Worskhop participants will learn how to process ChIP-seq data: perform read alignment, peak calling, visualization through the genome browser, motif finding and gene set enrichment analysis.

Contents:

I. **Module 1: Introduction (20 min)**
   a. Outline (3 min)
   b. Intro to ChIP-seq (17 min)

II. **Module 2: Alignment and identification of binding sites (1h)**
   a. Alignment (10 min)
   b. Peak calling (15 min)
   c. Peak overlap (5 min)
   d. <mark>Hands-on (30 min)</mark>

10 min BREAK

III. **Module 3: Quality control (35 min)**

   a. Common metrics for QC (20 min)
   b. <mark>Hands-on (15 min)</mark>

IV. **Module 4: Visualization (45 min)**

   a. Generation of bigwig tracks (10 min)
   b. UCSC genome browser (10 min)
   c. <mark>Hands-on (25 min)</mark>

10 min BREAK

V. **Module 5: Motif finding and gene set enrichment analysis (45 min)**

   a. Motif enrichment (15 min)
   b. Gene set enrichment analysis (10 min)
   c. <mark>Hands-on (20 min)</mark>

VI. **Concluding remarks (10 min)**