Project 3.9 Comprehensive, multimodal profiling of chromatin states and gene expression in single cells

Supervisor: prof. dr hab. Maciej Wojtkowski / dr Marcin Tabaka
Institute: Institute of Physics Chemistry PAS
Unit: Research group No. 3. Physical Optics and Biophotonics Group - prof. dr hab. Maciej Wojtkowski
www: http://groups.ichf.edu.pl/wojtkowski

Background:

Single-cell profiling examines DNA sequences of the genome from individual cells using Next Generation Sequencing. Nowadays, hundreds or even thousands of millions of such sequences can be simultaneously measured by state-of-the-art sequencers. To study cellular functions, many of the features along the genome can be processed in experimental protocols to be encoded by DNA sequences. Using bioinformatic tools DNA sequences can be mapped to the genome to annotate positions and levels of the features across genome. Depending on the number of features measured simultaneously from the same cell, we call such methods unimodal (one feature) or multimodal (multiple features). Most of the single-cell profiling protocols are unimodal and allow us to measure such modalities as RNA molecules, DNA-bound proteins, chromatin accessible regions such as active promoters or enhancers, or chromatin proteins and their modifications like histone modifications. In the last three years, we have seen advent of multimodal profiling protocols mostly for profiling of two modalities such as RNA molecules and chromatin accessibility or histone modifications.

Aim:

The main aim of this project is to develop an experimental protocol to simultaneously measure dozens of histone modifications and DNA-bound proteins, accessibility of the chromatin, and RNA molecules that are polyadenylated (protein coding genes) and non-polyadenylated (regulatory RNAs). This comprehensive approach will allow us to study single-cell genome-wide features at unprecedented resolution. To analyze big data generated by the new profiling protocol, we will develop a novel computational tool that would allow us to handle the multimodal data in a highly parallel manner.

Requirements:

- M.Sc. degree in Math, Statistics, Physics, Computer Science, Bioinformatics, Computational Biology, or related quantitative discipline,
- proficiency in Python,
- familiarity with Unix command-line tools,
- background in statistics and machine learning,
- knowledge of bioinformatics tools