Analysis and visualization of DNA methylation whole genome bisulfite sequencing data in diffuse large B-cell lymphoma

Chiara Facciotto[1], Amjad Alkodsi[1], Rainer Lehtonen[1], Sirpa Leppä[1,2], Sampsa Hautaniemi[1]

[1] *University of Helsinki*
[2] *Department of Oncology, Helsinki University Central Hospital*

Abstract
Diffuse large B-cell lymphoma (DLBCL) is an aggressive lymphoma. The current first-line therapy combines CD20 monoclonal immunotherapy to four chemotherapy drugs (R-CHOP). While R-CHOP cures 70-80% of the DLBCL patients those who relapse have abysmal prognosis.

DNA methylation is one of the key epigenetic mechanisms regulating gene expression. Alterations of DNA methylation patterns have been linked to several diseases including cancer.

In this project we analyzed the DNA methylomes of six DLBCL patients who, after an initial positive response to treatment, underwent relapse. DNA methylation samples were obtained before and after therapy using Whole Genome Bisulfite Sequencing (WGBS) technology. In addition, the DNA methylation profile of a healthy B-cell donor was obtained from the International Human Epigenome Consortium (IHEC) to be used as a healthy control.

In order to analyze WGBS data, a bioinformatic pipeline combining existing tools and in-house developed interactive visualizations was used to investigate the extent of differential methylation across all patients and between healthy and tumorigenic B-cells. An open source component-based workflow framework called Anduril was used to build our pipeline, allowing to integrate different softwares and perform WGBS data analysis in a modular, flexible and rapid fashion, through to the automatic parallelization of the workflow and to the possibility to fine-tune all parameters at every step of the analysis. The output of such analysis included differentially methylated regions either (i) between primary and relapse sample or (ii) between healthy and cancerous ones. In the first case, we were able to highlight methylation patterns caused by the drug resistance acquired by the patients during the administration of the therapy. In the second case, we focused on finding a DNA methylation profile that suggests an intrinsic drug resistant phenotype already present in the primary tumor and conserved in the relapse.

Interactive visualization tools were then used to compare these differentially methylated patterns across different patients, to investigate how much of the resistant phenotype is conserved across patients and how much is due to interindividual variability.

These bioinformatics tools enable advanced analysis and interpretation of WGBS DNA methylation data in cancer, helping us understanding the role played by this epigenetic mark in the onset of drug resistance.