SORS/WomenInBSC: Studying cell identity using single-cell epigenomic data

Abstract

Recent breakthroughs in high-throughput sequencing of single cells are revolutionizing the biological and biomedical sector. Among the different -omics layers that can be measured at the single-cell level, single-cell epigenomic measurements present a rich layer of regulatory information that stands between the genome and the transcriptome. These measurements can be obtained for large heterogeneous samples of single cells to profile tissues, organs and whole organisms, and to study dynamic processes like cellular differentiation, reprogramming or cancer evolution. These data types provide an unprecedented level of measurement resolution.

In this talk, I will discuss how single-cell ATAC-seq data, which measures chromatin openness at the single-cell level, can be used to study cell types. I will present how differential openness at the non-coding part of the genome, measured by scATAC-seq data, can be exploited to study cell identity. I will then introduce a new method, based on geometry regularized autoencoders, to embed the single-cell data into a low dimensional space. I will explain how we exploit this low dimensional representation to classify cells and learn new drivers of variation in the population.

Another level of genomic information that can be extracted from single-cell data are single-cell copy number variations (CNVs). I will present an algorithm that we have developed, epiAneufinder, which exploits the read count information from scATAC-seq data to extract genome-wide CNVs for individual single-cells, and I will show how the obtained CNVs are comparable to the ones obtained from single-cell whole genome sequencing data. Thanks to epiAneufinder it is possible to add a relevant extra layer of genomic information, namely single-cell copy number variation, to every scATAC-seq dataset without the need of additional experiments.

This picture belongs to the Helmholtz Center Munich

Short Bio

Maria Colomé Tatché is a theoretical physicist by training. She studied at the Autonomous University of Barcelona and at the École Normale Supérieure in Paris. She performed her PhD at the Laboratoire de Physique Théorique et Modèles Statistiques at Paris-Saclay and was a postdoctoral researcher at the Institute of Theoretical Physics at the University of Hannover. Her research focused on the properties of one-dimensional integrable systems.

Maria transitioned from theoretical physics to biology in 2010, attracted by the rapid and exciting developments in the field of epigenomics. Relying on her background in mathematical analysis, she develop ed computational and theoretical approaches to quantify the dynamical properties of the epigenome. During her postdoctoral position at the Groningen Bioinformatics Center, she worked on the study of DNA methylation inheritance patterns in populations of plants. After she became a junior group leader at the European Research Institute for the Biology of Ageing (ERIBA) (University Medical Centre Groningen) she went beyond applications in plants, and extended her computational methods to the analysis of epigenetic data in mammalian genomes, also at the single-cell level.

In 2016, Maria obtained a position as Helmholtz Young Investigator Group Leader at the Institute of Computational Biology (Helmholtz Center Munich). Since May 2021, she is also a professor at the Ludwig Maximilian University of Munich (Biomedical Center Munich BMC, Faculty of Medicine, LMU). She is strongly involved in the development of new algorithms for the analysis and interpretation of single cell -omic data.

Speakers

Speaker: Maria Colomé-Tatché. Professor at the Faculty of Medicine (LMU) and Group Leader at Institute of Computational Biology (Helmholtz Zentrum München)
Host: Alfonso Valencia. Life Sciences Department Director, BSC
Barcelona Supercomputing Center - Centro Nacional de Supercomputación

---