Abstract

An unprecedented wealth of biological data has been generated since the human and other genome projects were started. The huge request for raw data analysis and interpretation is being controlled by the evolving science of bioinformatics. Bioinformatics is defined as the application of computational tools and analysis approaches to capture and help to interpret results. It is an interdisciplinary field which harnesses computer science, mathematics, statistics, physics. It is essential for management of data in comprehensive medical research.

This Ph.D. thesis deals with the topic of bioinformatic pipelines implemented for end-to-end analysis of high-throughput sequencing data (DNA/RNA) in different research areas including cancer and assisted reproduction trough case studies.

First, example studies from primary central nervous system lymphoma (**PCNSL**) and chronic lymphocytic leukemia (**CLL**) research fields will be presented with the aim of identifying disease related somatic variants (SNPs, short INDELs) from custom targeted gene panels to create mutation profiles. Next, the reduced-representation bisulfite sequencing method and its related bioinformatic analysis will be demonstrated for DNA CpG methylation profiling to better understand one of the most aggressive cancers, known as glioblastoma multiforme (**GBM**). Additional to DNA related changes, microRNAs have emerged as promising biomarkers that can contribute more effective to early detection of lung cancer. The analysis of global miRNAome in early-stage non-small cell lung cancer (**NSCLC**) patients will be showed for the development of accurate predictive biomarkers of relapse, following surgery.

Infertility impacts the reproductive period of millions in the World and it influences their families and communities. *In vitro* fertilization helps in getting pregnant, embryo development and implantation, but also there is a high need for genetic examination before embryo transfer. Therefore, as the last application field, a comprehensive bioinformatic workflow for non-invasive pre-implantation genetic testing for aneuploidy (**NIPGT-A**) in assisted reproduction treatments will be introduced.

The described bioinformatic workflows in my thesis are essential in better understanding the molecular background of the previously mentioned disorders.

Key words: cancer, IVF, CLL, PCNSL, GBM, NCSCL, NIPGT-A, NGS, bioinformatics, workflow building, mutation profiling, methylome, CNVs, miRNAs