Single molecule detection and quantification of RNA modifications in human cancer samples from Nanopore data

Lead Groups: Leonardi and Nicassio (Center for Genomic Science, IIT)
Partner Group: Birney (EMBL-EBI)

Rationale & Hypothesis

The recent advent of Nanopore sequencing allowed for the first time to directly sequence native, full-length RNA molecules without retro-transcription or amplification, combining in a single technique quantification and sequence-specific detection of RNA modifications [1]. We have recently developed Nanocompore, a software for RNA modification detection that implements a model-free algorithm based on 2-dimensional Gaussian mixture models comparing an experimental RNA sample against a reference lacking modifications [2]. This work led to the generation of proof-of-concept results showing the feasibility of identifying RNA modifications at the level of individual RNA molecules, opening up the possibility to investigate the existence of an RNA *modification code* governing modifications on single molecules [3].

Aims

Building upon this foundation, we aim 1) to develop a robust analytical framework for the identification of RNA modifications on single molecules; 2) to explore the repertoire of RNA modifications at the single-molecule level in mammalian non-coding RNAs in the context of cancer.

Significance & Impact

This project will involve designing, developing and applying methods for single-molecule detection and quantification of RNA modifications from Nanopore direct RNA Sequencing data. This line of work will require to devise efficient algorithms for signal processing of the raw current data generated by the Nanopore sequencer and to implement robust approaches based on statistical modelling and/or machine learning for the detection and quantification of RNA modifications. This will constitute a significant technical advancement for the field of RNA modifications research. The fellow will also have the opportunity to apply these methods to a unique dataset of targeted direct RNA sequencing of long non-coding RNAs in the context of breast cancer, drawing novel insight into the regulation of these non-coding elements in tumour progression and chemoresistance.

Integration of Expertise of Partners

The Center for Genomic Science of IIT (based in Milan) has broad expertise in modern genomic technologies applied to the study of complex biological processes and diseases, with a particular emphasis on cancer. EMBL-EBI (Hinxton) is a world leading institute in the fields of computational biology and genome analysis. The project will be conducted in the context of a close collaboration between the the two institutes, and the fellow will be based at the Center for Genomic Science with the opportunity of frequent exchanges with EMBL-EBI.

References

1. Workman et al., ‘Nanopore Native RNA Sequencing of a Human Poly(A) Transcriptome’. Nature Methods 2019
2. Leger et al., ‘RNA Modifications Detection by Comparative Nanopore Direct RNA Sequencing’. bioRxiv 2019
3. Ries et al., ‘m6A Enhances the Phase Separation Potential of mRNA’. Nature 2019