Project 8
Project title: Detection and quantification of RNA modifications from nanopore direct RNA sequencing data
Lead group: Nicassio/Leonardi, Center for Genomic Science, IIT
Partner group: Birney group, EMBL-EBI

Rationale
The capacity of a cell to finely regulate its processes has been recognized as a fundamental property of every biological system since the early days of molecular biology [1]. One of the central regulatory mechanisms common to all kingdoms of life is the covalent modification of biopolymers after their synthesis. In fact, RNA has been known for decades to undergo several chemical modifications that modulate its properties, such as stability, structure or localization [2]. For technical reasons, the identification of RNA modifications has been historically mostly limited to abundant transfer and ribosomal RNAs, but recent technical advances have greatly expanded this landscape, leading to the identification of over 150 individual types of modifications in virtually every class of RNA molecule. Despite these recent advances, mapping modifications still mostly relies on indirect techniques based on reverse transcription signatures, chemical treatments or immunodetection with specific antibodies [3]. However, the recent advent of Nanopore sequencing allowed for the first time to directly sequence native, full-length RNA molecules without retrotranscription or amplification, combining in a single technique quantification and sequence-specific detection of RNA modifications [4].

Aims
To exploit the full potential of Nanopore direct RNA Sequencing, the scientific community needs robust analytical strategies and computational tools to extract information from the sequencing data and address biologically relevant questions. Despite fast progress in the detection of RNA modifications from Nanopore data, we are still lacking tools capable of accurately calling multiple RNA modifications from the raw sequencing signal at the single molecule level. The overall aim of this fellowship is to design, implement and apply algorithms and computational methods for the *de novo* identification of RNA modifications from raw sequencing data.

Objectives
The primary objective of this fellowship will consist in developing a method based on machine learning and/or statistical modelling for the identification of RNA modifications from direct RNA sequencing data. Such method will provide an extremely valuable tool to study the largely unexplored landscape of RNA modifications in a broad spectrum of contexts, with a high potential for discovering novel biologically relevant mechanisms. Depending on the fellows' scientific interests, background and expertise, we propose as a secondary objective the application of Nanopore-based RNA modifications detection to one or multiple of the following projects:

- Characterization of the interplay between transcriptional dynamics and RNA modifications through the parallel identification of incorporated nucleoside analogues (RNA metabolic labelling) and endogenous RNA modifications.
- Characterization of the expression profile and RNA modification profile of complex transcriptional units, such as long non-coding RNAs, in breast cancer.
- Characterization of RNA transcription and RNA modification profiles upon DNA damage.
- Identification of RNA modification QTLs in a large scale DNA/RNA sequencing project of Medaka fish.

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. The Birney group (EMBL-EBI, Hinxton) has very extensive experience in computational biology, and in particular on the development of algorithms for sequence analysis. The partnership between the two groups will provide both the ideal environment for designing and developing methods (EMBL-EBI) as well as for applying them to biologically relevant contexts (CGS).

References
[1] Britten and Davidson, Science (1969).
[2] Cohn and Volkin, Nature (1951).
[3] Helm and Motorin, Nat Rev Gen (2017).
[4] Workman et al., Biorxiv (2018).