Short talk at the sponsor time slot
EpiBesançon 2022 10 & 11 MAY
6th edition of international epigenetics congress

Solving the multi-alignment challenge for interrogating non-coding RNA-seq

Andrea Hita Ardiaca
Diagenode, Liège, Belgium

Presenter: Andrea Hita Ardiaca (ESR14)

Total-RNA sequencing (total-RNA-seq) allows the simultaneous study of both the coding and the non-coding transcriptome. However, transcripts from distinct RNA biotypes vary in length, biogenesis, and function, can overlap in a genomic region, and may be present in the genome with a high copy number. Consequently, reads from total-RNA-seq libraries may cause ambiguous genomic alignments, demanding flexible quantification approaches capable of analyzing these reads. To overcome this challenge, we present Multi-Graph count (MGcount), a flexible quantification framework that firstly, it hierarchical assigns reads to transcripts to account for loci length disparity between small-RNA and long-RNA and secondly, it collapses locus where reads systematically multimap into communities. Ultimately, this strategy maximizes the information extracted from totalRNA-seq and improves the interrogation of non-coding RNA.

This project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 813282, H2020-MSCA-ITN "PEP-NET: Predictive Epigenetics".