From genomic dark matter to functional traits: an innovative bioinformatic strategy to investigate their genomic bases within non-model organisms transcriptomes

Lucie Bittner*[1], Arnaud Meng , Camille Marchet , Erwan Corre , Pierre Peterlongo , Adriana Alberti , Corinne Da Silva , Patrick Wincker , Eric Pelletier , Stéphane Le Crom , and Fabrice Not

[1]Sorbonne Université, Evolution Paris Seine (SU, UPMC, EPS) – Université Pierre et Marie Curie [UPMC] - Paris VI – Université Pierre et Marie Curie Institut de Biologie Paris Seine 7-9 Quai Saint Bernard 75252 Paris cedex 05, France

Abstract

The advent of high-throughput sequencing approaches has unveiled the extent of Earth biodiversity and revealed our ignorance with respect to the role of this diversity in ecosystems' functioning. The fundamental molecular mechanisms associated to the functional traits of organisms are poorly known, and often restricted to model organisms. For instance, symbiotic relationships are widespread and are critically important for the functioning of ecosystems but the genomic bases of the establishment and the maintenance of these associations remain largely unknown. The study of holobiont transcriptomes involving non-model lineages represents bioinformatic challenges in order to circumvent the production of chimera or to distinguish the taxonomic origins of the assembled sequences. In addition, the vast majority of the sequences obtained remain functionally unknown, limiting the analyses to a subpart of the genomic data newly produced. Here we present an innovative bioinformatic approach and applied it to marine holobionts, in order to investigate the genomic basis of some of their traits. Using a k-mer based similarity method, holobiont reads were sorted out and independent assemblies were performed for each partners, leading to a significant diminution of de novo assembled chimeras compared to classical assembly methods. Thereafter, sequence similarity network analyses were used to perform comparative studies including the functionally unknown sequences. Following this strategy, candidate protein domains associated to symbiosis were identified. These genomic markers characterizing functional traits constitute working hypotheses to be further confirmed by targeted molecular studies. To conclude, our efficient and innovative analysis strategy allows to study the genomic of non-model organisms and their dark matter on a massive scale, and represents one of the very few studies available to date to expand our genomic knowledge about functional traits.

*Speaker