



Sequencing Techniques & the Discovery of Transposable Elements

June 15, 2021

Background: With the advances of sequencing techniques, the availability and quality of genomes is ever increasing, leading to a more refined understanding of the processes underlying evolution. One of the major limitation of the previous sequencing technique (short-read sequencing) is its reduced ability to recover sequences with a high number of repeats. Those sequences are of high interest as recent evidence points to the role of Transposable Elements (TEs), composed of repeats, in the emergence of phenotypic innovations.

To investigate the role of TEs during evolutionary processes, the evolution of TEs needs to be compared among species. However, since many genomes have been sequenced with various techniques, how it can impact future analyses remains unknown.

Objectives: The present project aims to understand to which extent the use of several sequencing techniques will impact the detection and categorization of TEs by comparing genomes of the same species obtained with different sequencing techniques (i.e. short- vs. long-read sequencing). One main point is to understand if different techniques have only quantitative differences in TE detection or also qualitative ones. This project, therefore, aims to develop strategies and guidelines when analysing the evolution of TEs among species with different sequencing techniques to avoid any bias.

Requirements:

- Good knowledge of literature and databases searches
- Interest in genome evolution and sequencing techniques
- Basic knowledge of command line, scripting and data handling
- Ability to work independently

Methods: After literature search and gathering 12 (best case scenario) species sequenced with both short- and long-read techniques, several tools (repeatmodeler, LTRharvest, ...) will be used to retrieve TEs in each genome. From newly obtained TE data, comparison between techniques will be undertaken through general linear mixed models in R to account for all factors which could influence TE detection and categorization.

Supervision: Dr. Bertrand Fouks, b.fouks@uni-muenster.de, Molecular Evolution and Bioinformatics Group (<http://bornberglab.org/>).