ENASearch: A Python library for interacting with ENA's API

Bérénice Batut and Björn Grüning

1 Bioinformatics group, Freiburg, Germany

Summary

The European Nucleotide Archive (ENA) (Leinonen et al. 2010) is a database with a comprehensive record of nucleotide sequencing information (raw sequencing data, sequence assembly information and functional annotation). The data contained in ENA can be accessed manually or programmatically via REST URLs. However, building HTTP-based REST requests is not always straightforward - a user friendly, high-level access is needed to make it easier to interact with ENA programmatically.

We developed ENASearch, a Python library to search and retrieve data from ENA database. It also allows for rich querying support by accessing different fields, filters or functions offered by ENA. ENASearch can be used as a Python package, through a command-line interface or inside Galaxy (Afgan et al. 2016).

References
