




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

B er nice Batut¹ and Bj orn Gr uning¹

¹ Bioinformatics group, Freiburg, Germany

DOI: [10.21105/joss.00418](https://doi.org/10.21105/joss.00418)

Software

- [Review](#) 
- [Repository](#) 
- [Archive](#) 

Licence

Authors of JOSS papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC-BY](#)).

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

- Afgan, Enis, Dannon Baker, Marius Van den Beek, Daniel Blankenberg, Dave Bouvier, Martin  ech, John Chilton, et al. 2016. "The Galaxy Platform for Accessible, Reproducible and Collaborative Biomedical Analyses: 2016 Update." *Nucleic Acids Research* 44 (W1). Oxford University Press: W3–W10.
- Leinonen, Rasko, Ruth Akhtar, Ewan Birney, Lawrence Bower, Ana Cerdeno-T arraga, Ying Cheng, Iain Cleland, et al. 2010. "The European Nucleotide Archive." *Nucleic Acids Research* 39 (suppl_1). Oxford University Press: D28–D31.