

Benchmarking metagenomic software with the new CAMI web portal

Fernando Meyer^{1,2,3,4}, Gary Robertson^{1,2,4} & Alice C. McHardy^{1,2,3,4}

¹Computational Biology of Infection Research, Helmholtz Centre for Infection Research (HZI), Braunschweig, Germany; ²Braunschweig Integrated Centre of Systems Biology (BRICS), Technische Universität Braunschweig, Braunschweig, Germany; ³Cluster of Excellence RESIST (EXC 2155), Hannover Medical School, Hannover, Germany; ⁴NFDI4Microbiota, German National Research Data Infrastructure, Germany.

Introduction

To enable objective and comprehensive benchmarking of metagenomic software, the community-led initiative for the Critical Assessment of Metagenome Interpretation (CAMI) [1] promotes standards and good practices agreed upon by the community, by providing comprehensive datasets for developers to assess their software on, together with software facilitating their assessment, benchmarking guidelines, and challenges. However, assessing methods in between challenges and comparing their performances to other techniques still requires substantial time and technical expertise.

Results

We introduce the new CAMI benchmarking web portal, which allows users to assess software for common metagenome analyses using the CAMI benchmarking metrics and visualizations available in the AMBER [2], MetaQUAST [3], and OPAL [4] evaluation software. As input, the portal accepts assemblies, genome or taxonomic binning, and taxonomic profiles of a CAMI benchmark dataset. As output, it displays the performance metrics and visualizations employed in the CAMI challenges [1] (Figs.1-4).

The portal improves the FAIRness of computationally created benchmarking data by making it more:

- Findable through relevant, searchable metadata and Digital Object Identifiers (DOIs)
- Accessible through the portal and public repositories
- Interoperable by providing results in common formats that are machine readable
- Reusable and reproducible for more sustainable research efforts

It is available for testing at <https://cami.bifo.helmholtz-hzi.de>.



Figs. 1-4: CAMI portal example visualizations and metrics for results of the CAMI II Challenge on the marine (1-3) and strain madness (4) datasets. 1a-c. Assembly results. 2a-c. Genome binning results. 3. Taxonomic binning results at the genus rank. 4a-c. Taxonomic profiling results.

Conclusions: The CAMI benchmarking portal frees users from the need to install and execute other methods and evaluation software for a comprehensive assessment of individual tools. It further facilitates an interactive exploration of method results, to identify the most suitable software meeting a specific performance profile and provides a continuously adapting evaluation frame for metagenome analyses methods.

References

1. Meyer, F., Fritz, A. et al. Nat. Methods 19, 429–440 (2022)
2. Meyer, F. et al. Gigascience 7, giy069 (2018)
3. Mikheenko, A. et al. Bioinf 32, 1088–1090 (2016)
4. Meyer, F. et al. Genome Biology 20 (2019)