











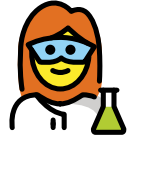


Efficient parallel taxonomic profiling for metagenomics with taxprofiler

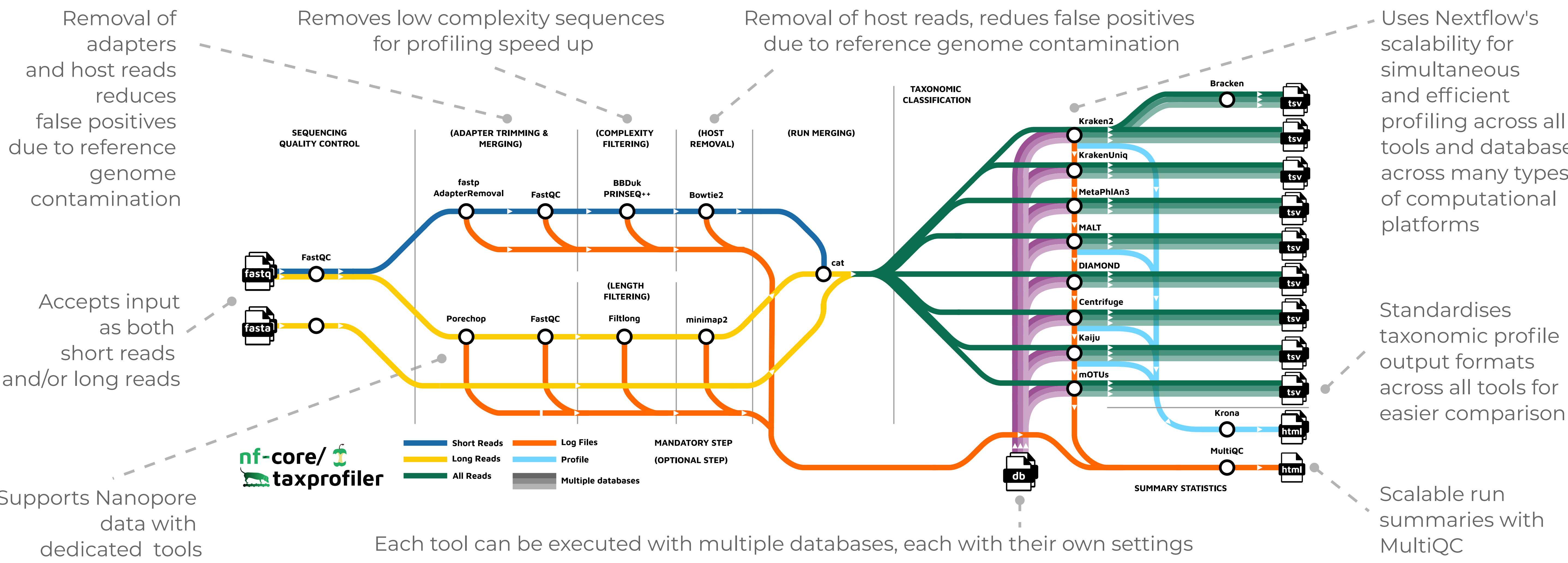
James A. Fellows Yates ^{1,2}, Moritz E. Beber ³, Sofia Stamouli ^{4,5,6}, Lauri Mesilaakso ⁷,
Tanja Normark ^{4,5,6}, Jianhong Ou ⁸, Thomas A. Christensen II ⁹, Mahwash Jamy ^{4,5,6},
Maxime Borry ², Rafał Stępień ¹⁰, & the nf-core community

¹. Dept. of Paleobiotechnology, Leibniz Institute for Natural Product Research and Infection Biology Hans Knöll Institute, Jena,  ². Dept. of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, Leipzig,  ³. Unseen Bio ApS, Copenhagen,  ⁴. Dept. of Microbiology, Tumor and Cell Biology, Karolinska Institute, Solna,  ⁵. Clinical Genomics, SciLifeLab, Solna,  ⁶. Dept. of Clinical Microbiology Karolinska University Hospital, Stockholm,  ⁷. Dept. of Clinical Microbiology, Region Östergötland, Linköping,  ⁸. Regeneration Center Duke University, Durham (NC),  ⁹. Veterinary Diagnostic Laboratory Kansas State University College of Veterinary Medicine, Manhattan (KS),  ¹⁰. Center for Public Health Genomics, University of Virginia, Charlottesville (VA), 



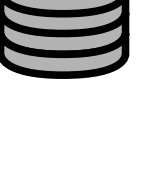
Motivation

-  **Shotgun metagenomics** is a powerful method for identifying all taxa present in a biological sample. However, each taxonomic profiling tool and database have their own **biases that may result in false positives or false negatives**. Furthermore, metagenomics is infamously renowned for being very computationally resource hungry.
-  To validate results and reduce computational requirements it has been often recommended to run samples against multiple tools and databases and compare output. However, this is challenging due to **different output file formats**.
-  Software allowing for **simultaneous profiling against multiple tools, databases, and settings**, with standardised output, would benefit many researchers - such as those working on **microbiomes and clinical microbiology**.


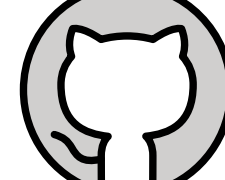
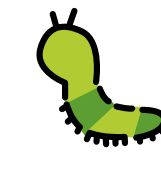
nf-core/taxprofiler workflow



Future

-  nf-core/taxprofiler is written in **Nextflow DSL2** with **nf-core modules**, making it easily extendable to **new taxonomic profilers required by the community**.
-  Investigating the ability to generate a **'consensus' profile** across all profilers.
-  Considering a separate workflow that **automates databases** generation for each tool **from the same set of input sequences** (where possible).

Get involved!

-  [nf-core/taxprofiler](#)
-  [nf-co.re/join](#)
-  [#taxprofiler](#)

Acknowledgements

Zandra Fagnäs, Robert Petit III, John Franklin Rawls, Christina Warinner, Microbiome Sciences Group (MPI-EVA). Icons from openmoji.org (CC-BY-SA 4.0)