



MONTY: an open-source, modular pipeline for virus metagenomics



Charles S.P. Foster^{1,2,3}, Preston Leung^{2,3}, Beatriz Delgado Corrales^{2,3}, Legana Fingerhut^{2,3}, Maria E. Craig^{1,3,4}, William D. Rawlinson^{1,2,3} and Ki Wook Kim^{1,2,5}

1. School of Biomedical Sciences, Faculty of Medicine and Health, UNSW Sydney, New South Wales, Australia; 2. School of Biotechnology and Biomolecular Sciences (BABS), UNSW Sydney, New South Wales, Australia; 3. Virology Research Laboratory, Serology and Virology Division (SAVID), NSW Health Pathology, Prince of Wales Hospital, Sydney, NSW 2031, Australia; 4. Children's Hospital Westmead, Faculty of Medicine and Health, University of Sydney, New South Wales, Australia; 5. School of Clinical Medicine, Faculty of Medicine and Health, UNSW Sydney, New South Wales, Australia.



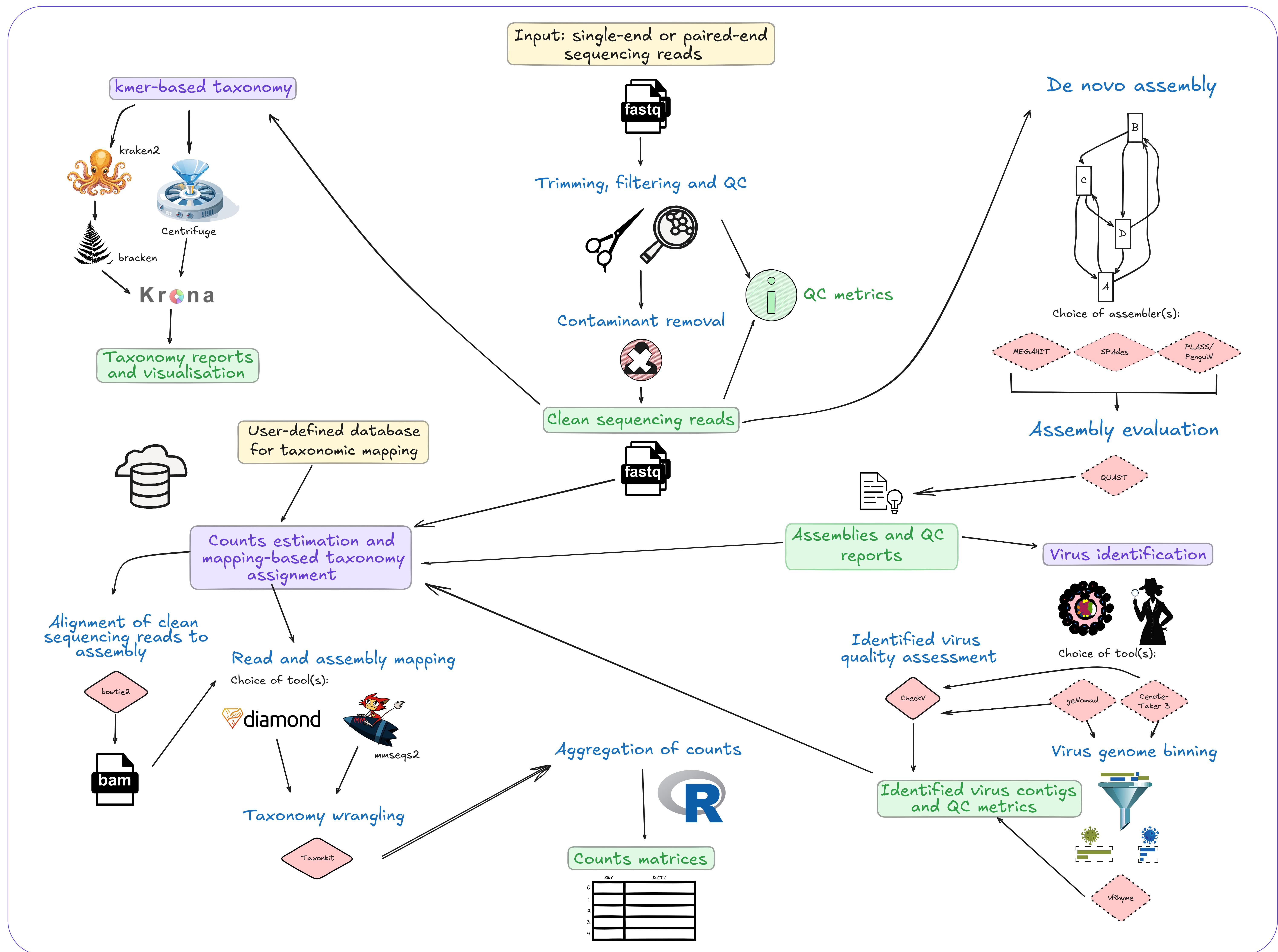
charles.foster@unsw.edu.au

The Problem

- The microbiome can cause both acute and chronic disease, and act as an environmental trigger of autoimmune diseases.
- Quantifying the presence and abundance of taxa in the microbiome is important, but most existing methods either do not account for viruses or only have viruses as a secondary consideration.
- In many cases, output count data are not appropriately normalised.

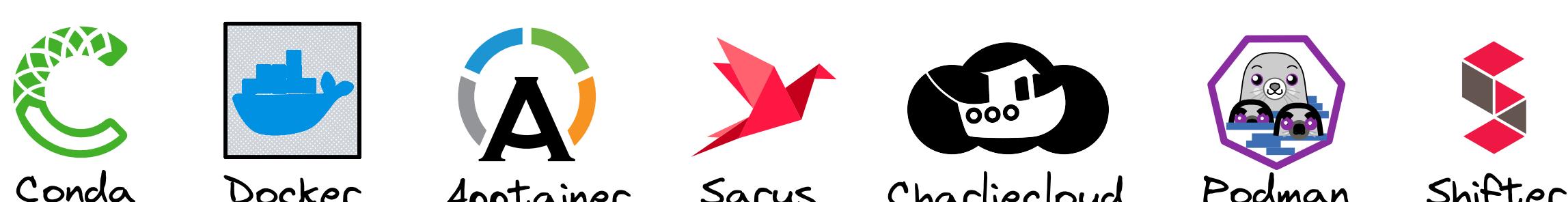
Our Solution

- Develop a novel pipeline specifically targeted towards **virus metagenomics**.
- Goals: easy to run, reproducible, open-source, modular, highly scalable.
- Result: the “Metagenomic Analysis of Existing and Novel Threats in Virology” workflow, a.k.a. **MONTY**.
- Why “MONTY”? To honour our lab’s unofficial mascot (my dog), and hey, bioinformaticians love a forced “backronym”.



MONTY's Strengths

- Virus-focused, including de novo virus identification
- Developed in **nextflow** according to the best practices of **nf-core**: modular, reproducible, freely available, open source.
- All dependencies handled by:



- Easily deployed on, and integrated with:



What's next?

- Continued benchmarking against existing workflows (e.g. CZ-ID)
- Continued development and optimisation of the workflow, including the addition of new modules on request.
- Improved normalisation of count matrices.
- Provision of MONTY as an online service: workflow and associated databases hosted on AWS to allow secure and easy deployment using your own data.

