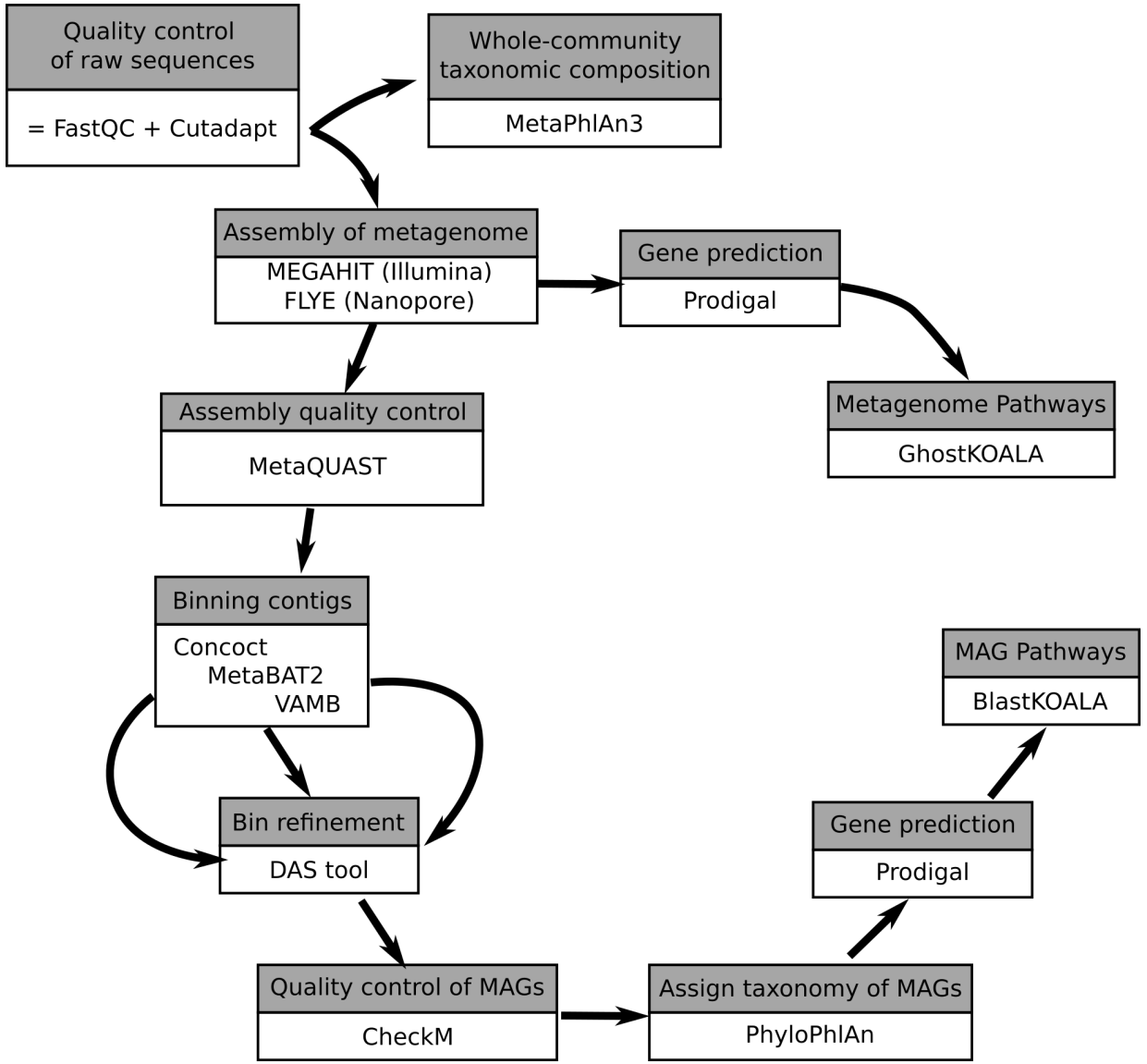
Functional microbiome research – bioinformatics section

Part 2 – Metagenomics

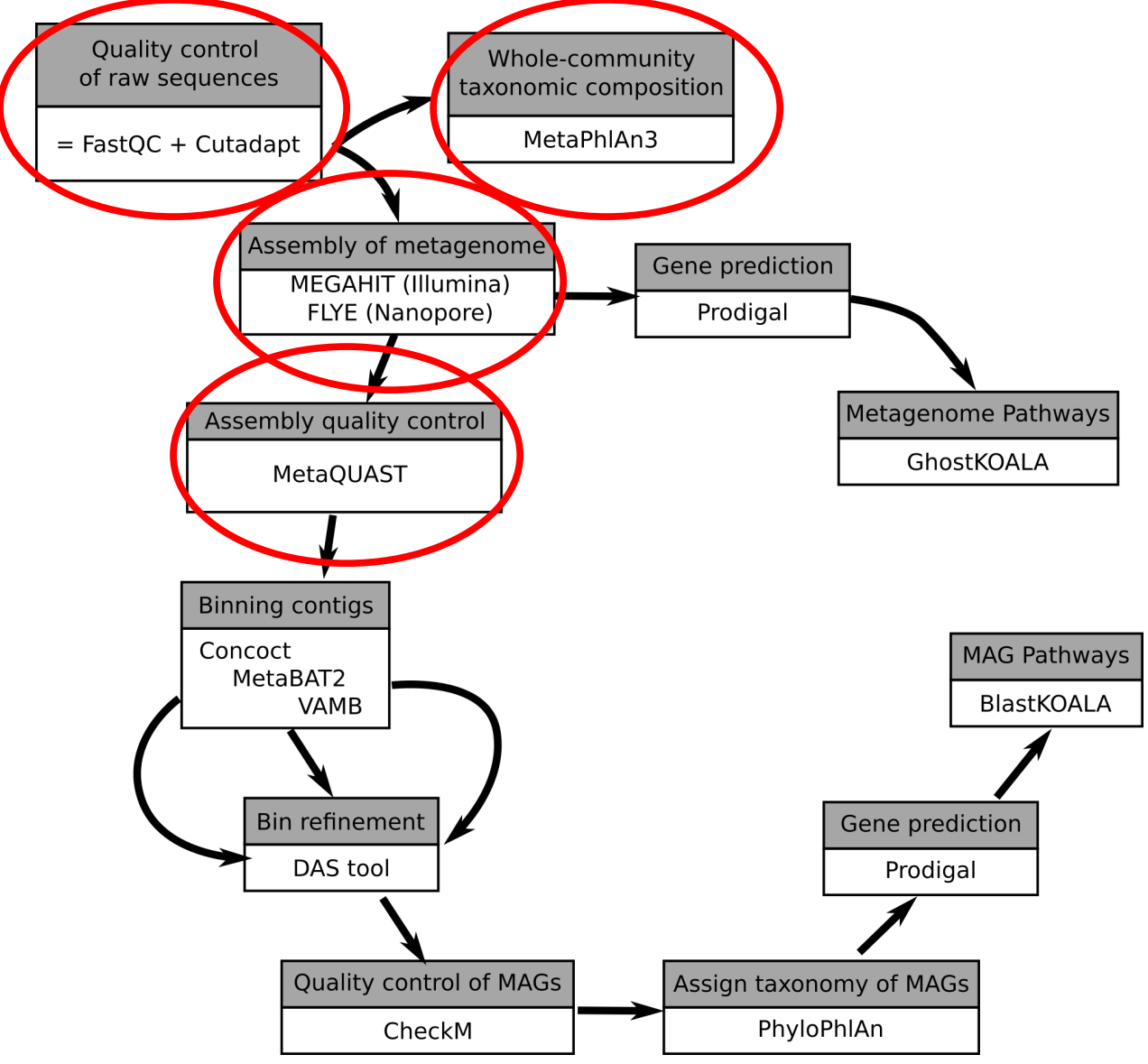
You can grab today's scripts here:

```
wget https://raw.githubusercontent.com/danchurch/FunctionalMicrobiomePractical2022/main/funmic2023/funMetagenomicScript.txt
```



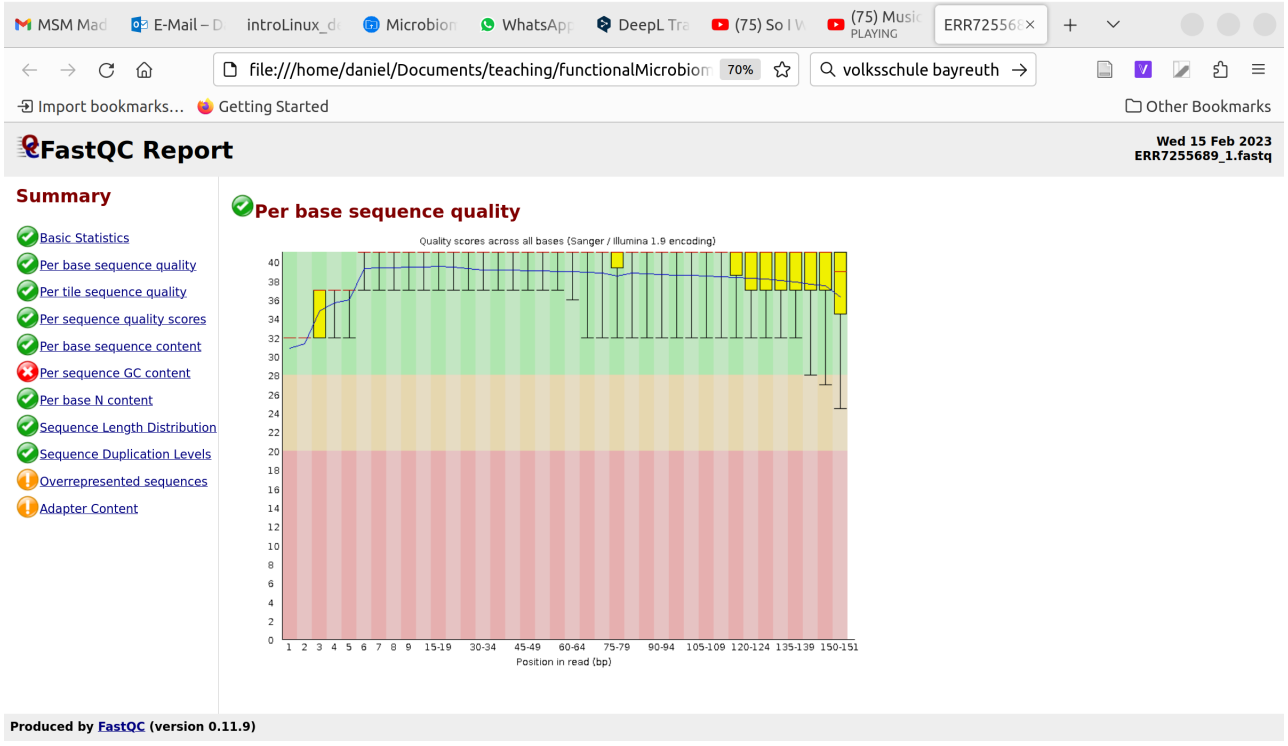
This morning we will attempt to:

- look at the quality of our illumina and nanopore reads
- profile the members of the microbial community
- assemble our illumina-based metagenome
- check the quality of our nanopore and illumina assemblies

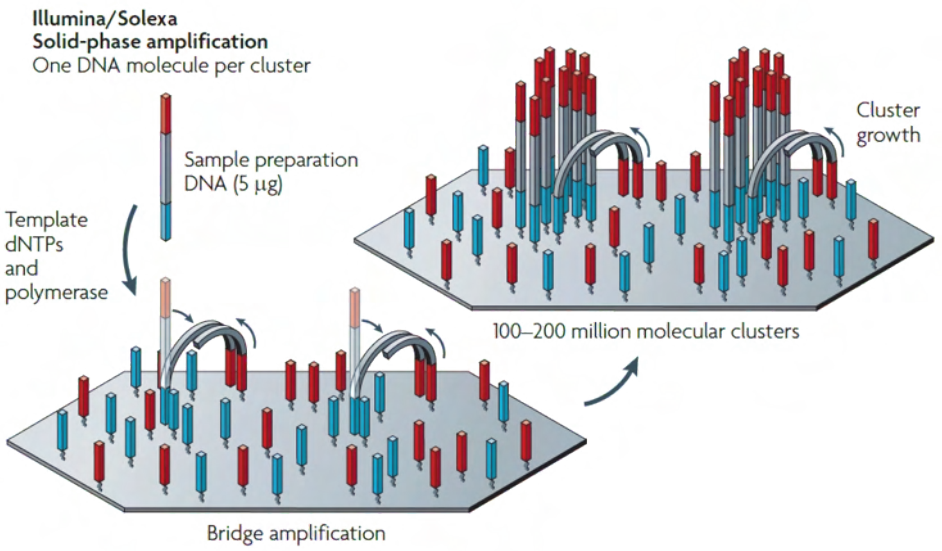


This morning we will attempt to:

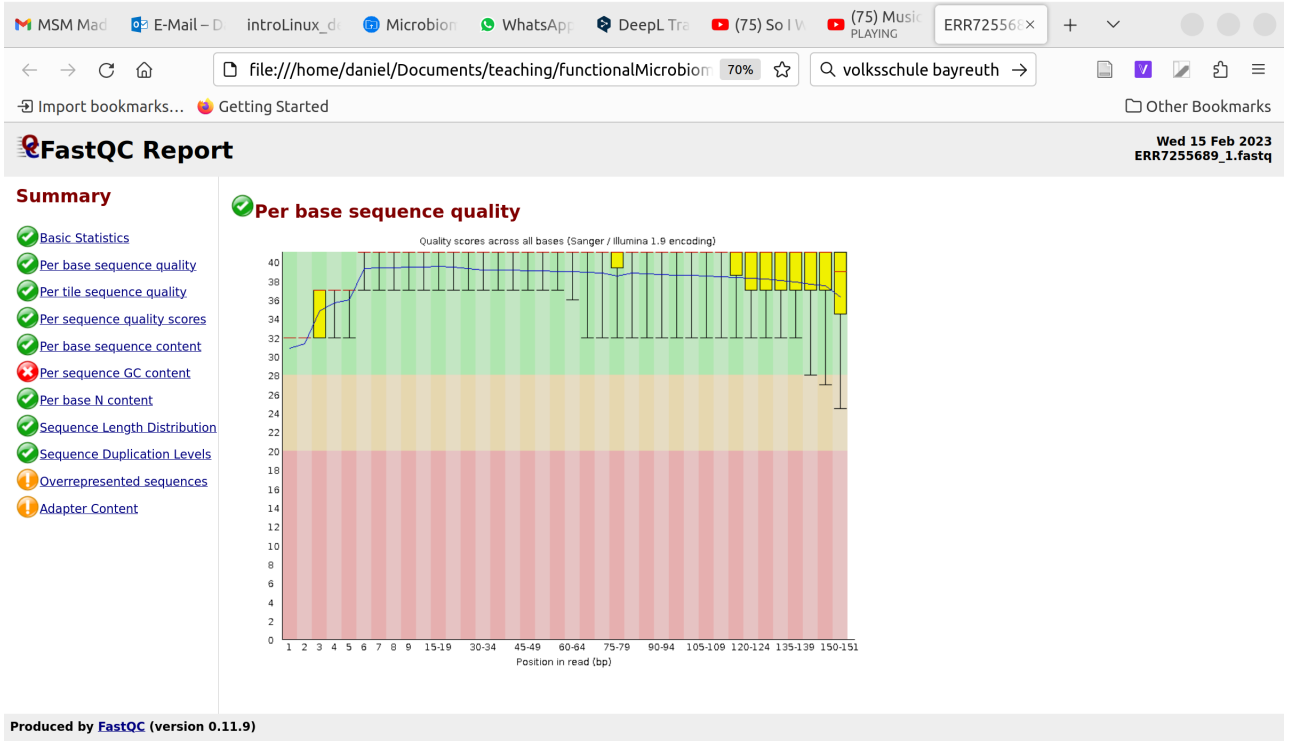
- look at the quality of our illumina and nanopore reads
- profile the members of the microbial community
- assemble our illumina-based metagenome
- check the quality of our nanopore and illumina assemblies



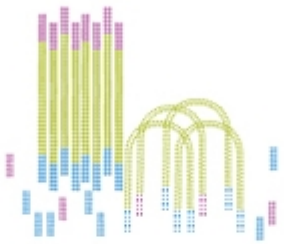
Check our reads with FASTQC.



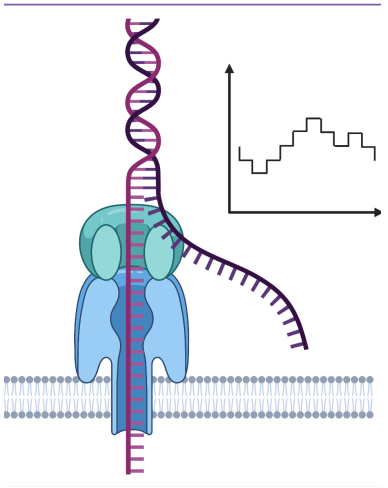
Illumina data often has two, paired-end read files. Why?

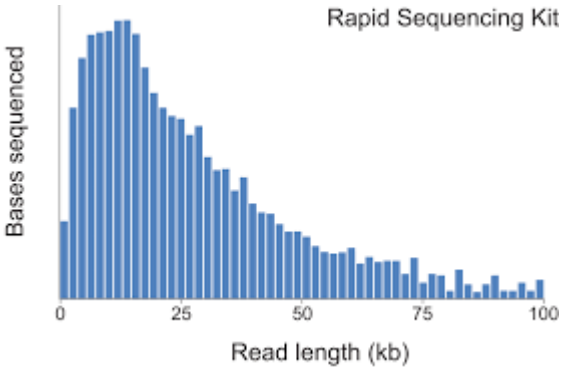


Let’s look at what FASTQC can tell us.



Illumina

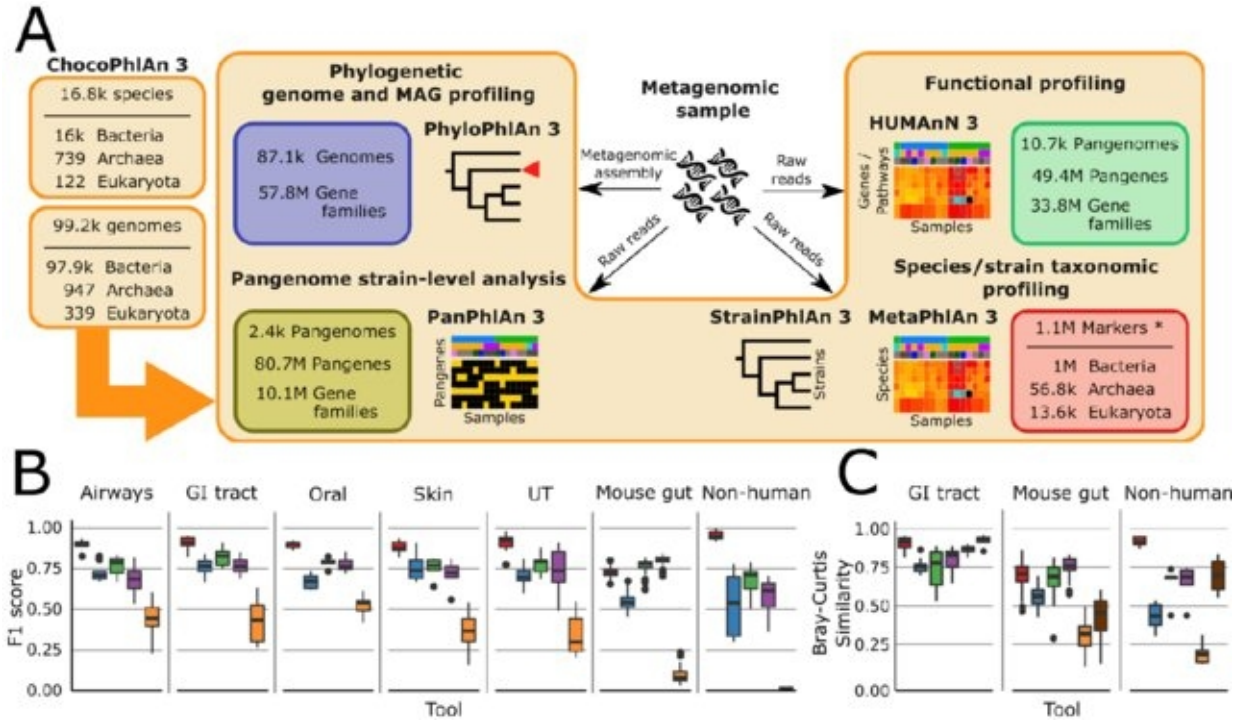




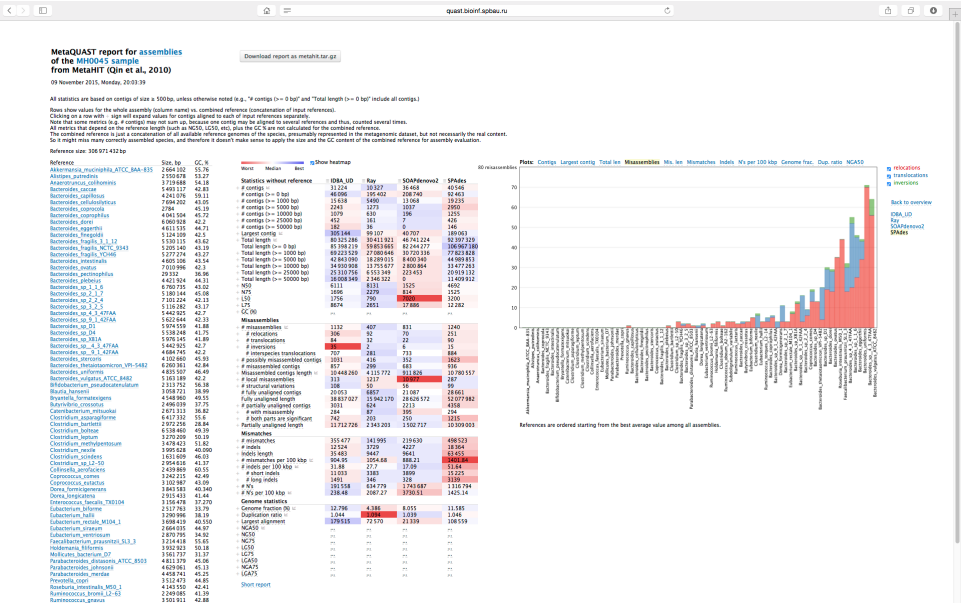
Nanopore

Metagenome assemby: we use two assemblers:

- MetaFLYE can handle the longer reads, shallower depth and higher error rates of Nanopore data.
- MEGAHIT is optimized for short reads with high accuracy (Illumina data).



MetaPhlan and biobakery – marker-based discovery of who is in our metagenome? Not just 16s!! MetaPhlan is still only available for illumina data.



Finally, we can also use MetaQUAST to preview our metagenome. MetaQuast extracts 16s barcodes from our metagenome assembly, looks for the closest match on public databases, and downloads these matches for comparison.