

Introduction to Metagenomics for Clinical Virology

Sarah Buddle

UCL Great Ormond Street Institute of Child Health

Session developed by Dr Cristina Venturini

Session structure

11:00-12:00: Introduction to metagenomics

12:00-13:00: Metagenomics bioinformatics practical

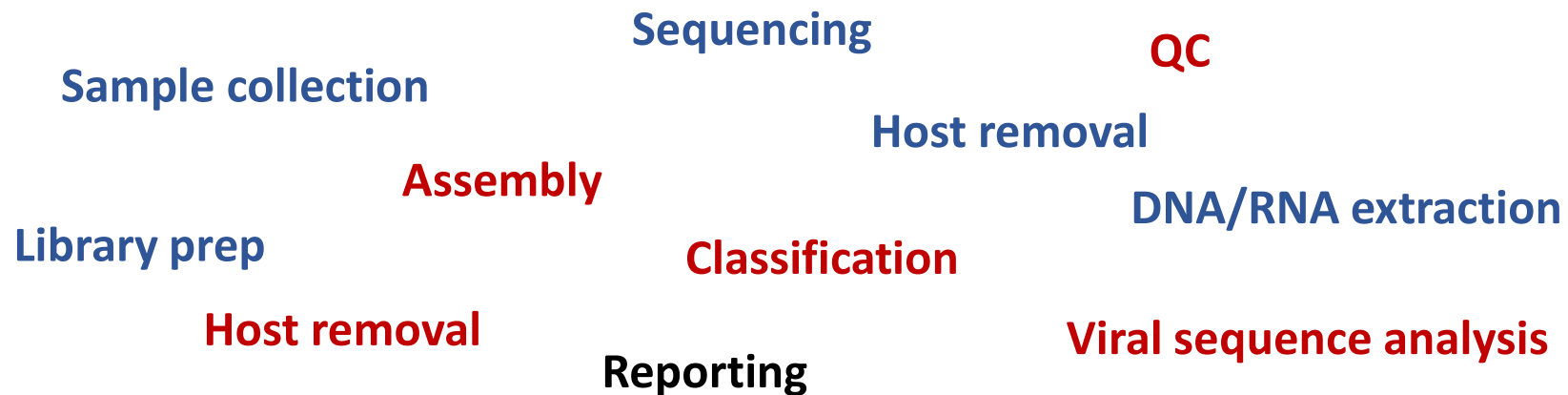
1. What is metagenomics?
2. What clinical questions can we answer with metagenomics?
3. What are the advantages and disadvantages of metagenomics over other techniques you might use to answer those questions?
4. *(Optional)* What might you need to consider before implementing metagenomics in a clinical or public health setting? If you have used metagenomics before, what difficulties did you encounter?

Protocol

What are the key steps in a metagenomics protocol?

What is the purpose of each step?

What methods might you use?



Optional: What sequencing platforms could you use for metagenomics and what are the advantages/disadvantages of each?

Classification

What factors should we consider when choosing:

1: a classifier

2: sequences to include in your database

Contamination

1. Where might contamination come from?
2. How can we reduce/deal with contamination?

Practical

Part 1: Metagenomics analysis with Kraken2/Bracken (command line)

Try to work out the commands yourself rather than looking at the answers!

Part 2: Metagenomics analysis with CZID (online)

Use the login details on the board.

Choosing bioinformatics protocols for metagenomics

The protocol shown in the practical may not be the best one for your research or clinical question!

Some other tools: a non-exhaustive list

nf-core/taxprofiler

nf-core is a set of community-curated best practice bioinformatics pipelines built in Nextflow.

Taxprofiler Includes Kraken2/Bracken, DIAMOND, Centrifuge etc



Online, cloud-based, user-friendly tool



Illumina Dragen Metagenomics / Nanopore EPI2ME labs wf-metagenomics

Illumina and Nanopore's tools. Simple to run and can be automated.



Check benchmarking papers for lots of other options!