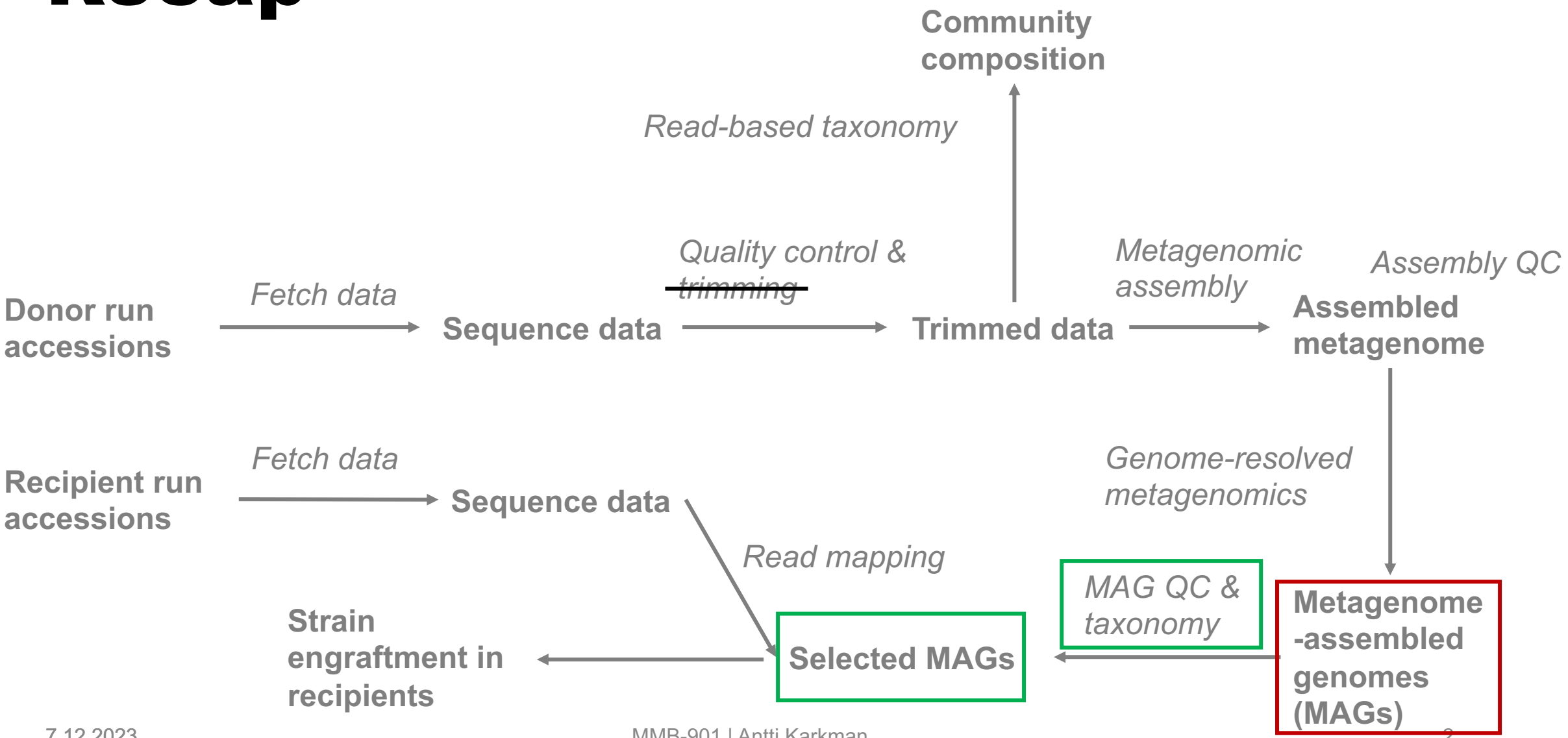
Microbial metagenomics

MMB-901

Recap


Community composition
Read-based taxonomy
Quality control & ~~trimming~~
Metagenomic assembly
Assembly QC
Donor run accessions
Fetch data
Sequence data
Trimmed data
Assembled metagenome
Recipient run accessions
Fetch data
Sequence data
Genome-resolved metagenomics
Read mapping
MAG QC & taxonomy
Metagenome -assembled genomes (MAGs)
Strain engraftment in recipients
Selected MAGs

CheckM2

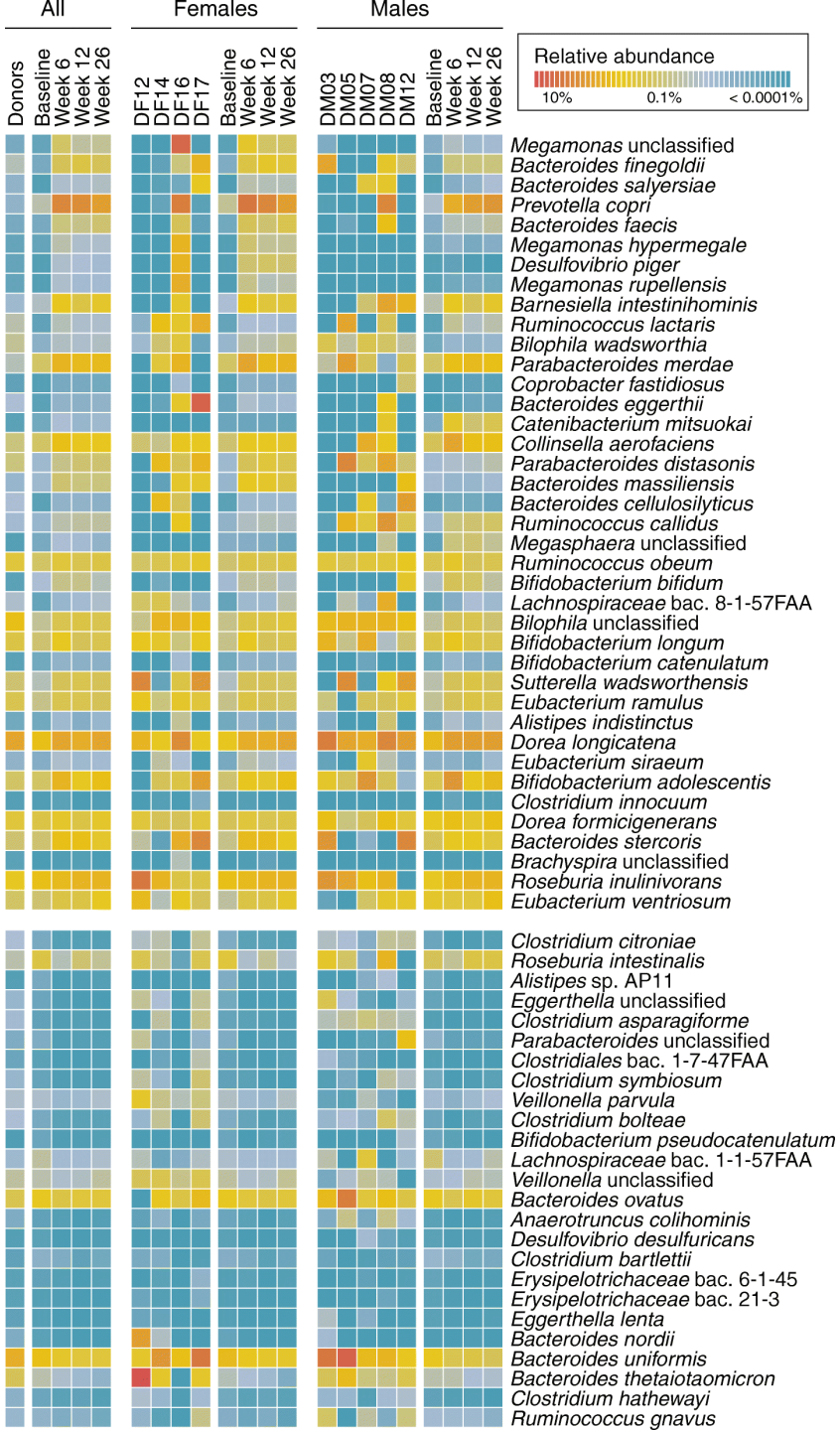
- Predicts the completeness and contamination of MAGs
- Uses machine learning (not SCGs like e.g. CheckM1)
- Trained on simulated genomes with known completion and contamination – “random-protein-sampling”
- Should work better for lineages with reduced genome sizes compared to SCG-approach

GTDB-Tk

- Genome Taxonomy Database toolkit
- GTDB
  - Standardised microbial taxonomy based on phylogeny
  - Based on SCGs: 120 for bacteria and 122 for archaea
- Taxonomc annotation based on concatenated protein reference trees

Selected MAGs

- One either *P. copri* or *Megamonas funiformis*
- Select one additional donor MAG that seems to colonize the recipients
- Original article Figure 2



Next steps in strain engraftment

- Annotation of selected MAGs
- Construct contigs DB from annotated genomes
- Run HMM and add SCG taxonomy annotations to contigs DB
- Fetch all sequence data for 3 recipients
- Map recipient data to each of the selected MAGs
- Construct profile DBs and merge them
- Visualize the results

Next steps in strain engraftment

- Annotation of selected MAGs
- Construct contigs DB from annotated genomes
- Run HMM and add SCG taxonomy annotations to contigs DB
- Fetch all sequence data for 3 recipients
- Map recipient data to each of the selected MAGs
- Construct profile DBs and merge them
- Visualize the results

Let’s get to work

https://github.com/karkman/MMB-901_Metagenomics