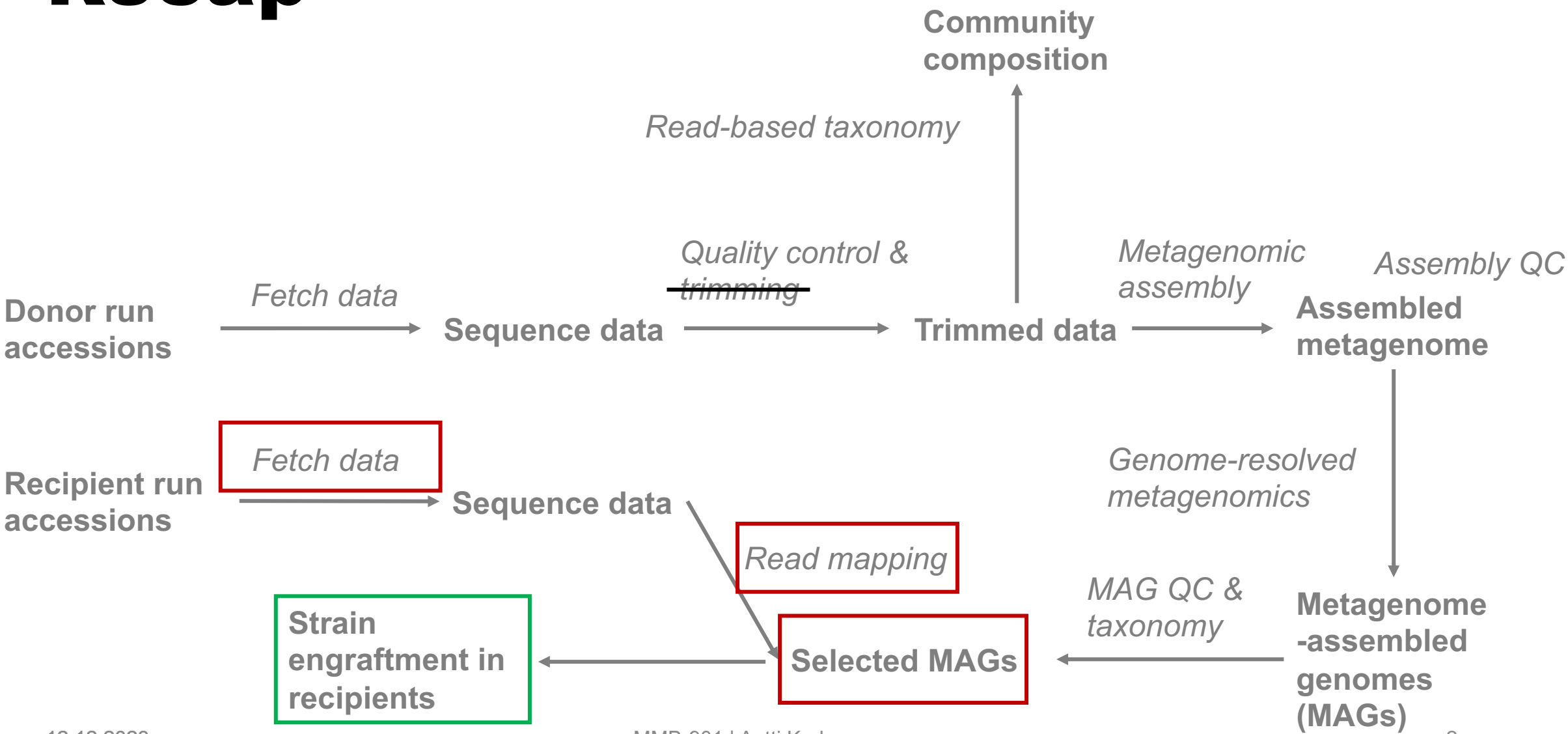
Microbial metagenomics

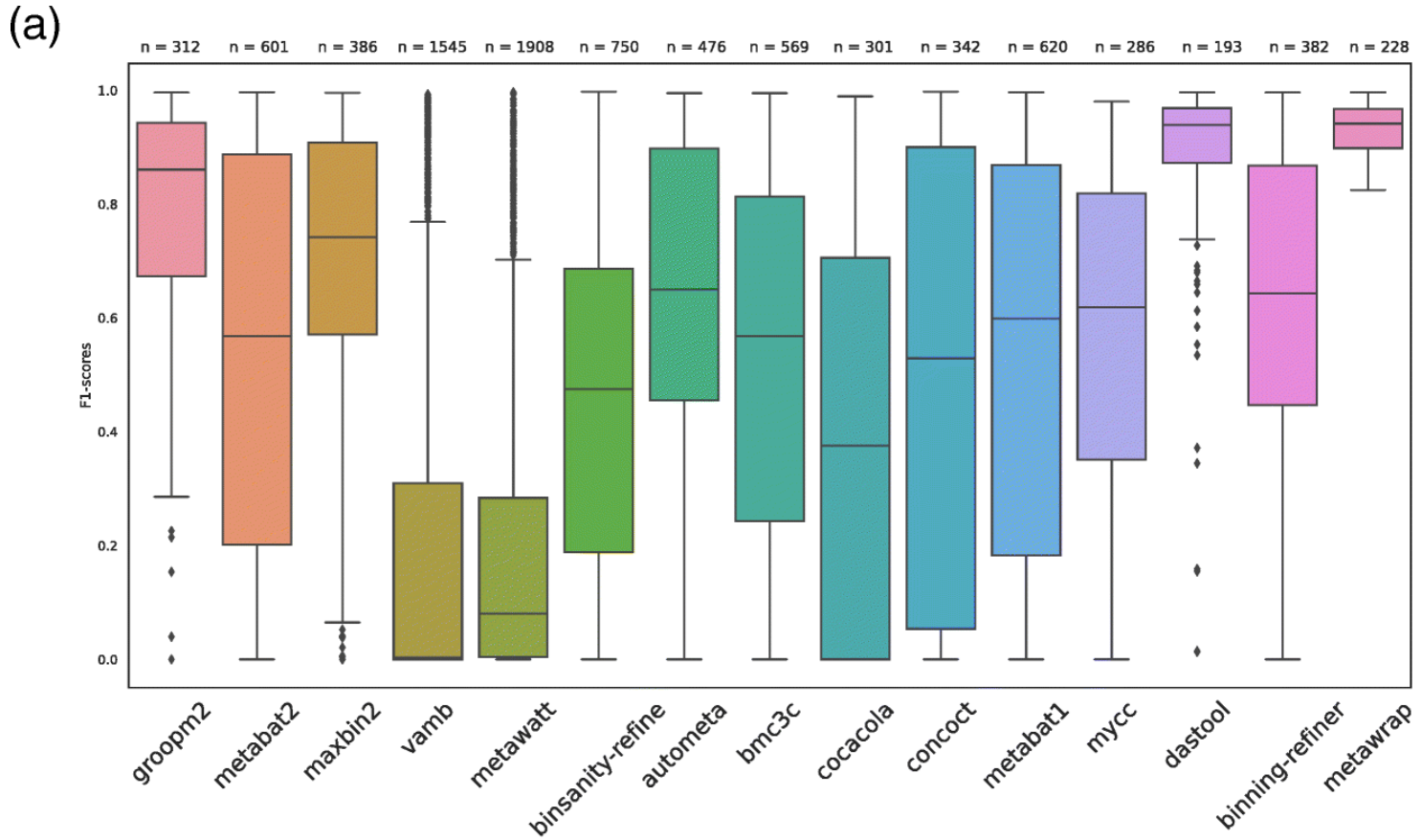
MMB-901

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
Strain engraftment in recipients
Selected MAGs
Metagenome -assembled genomes (MAGs)

Automatic binning

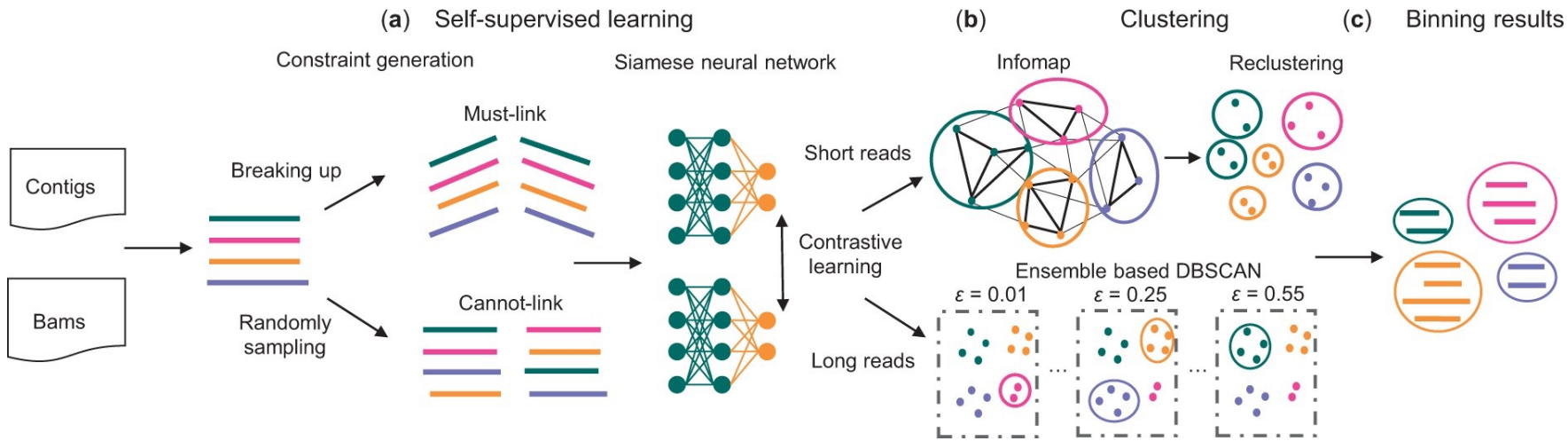
- Several tools available, with varying performance



https://doi.org/10.1186/s12859-020-03667-3

Semibin2

- Based on self-supervised learning
  - Unsupervised methods use kmer distribution and abundance
- Pre-trained models for single sample binning – fast
- Multi-sample binning improves binning results – slower



https://doi.org/10.1093/bioinformatics/btad209

Our workflow in short

- Run Semibin2 with a pre-trained model with single sample
- Do quality control and annotation of resulting bins
- Import Semibin2 collection to anvi'o and inspect the results
- Optional:
  - Run Semibin2 in multi-sample mode with all 9 samples
  - Generate plots on the fate of contigs with different binning approaches

Let’s get to work

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