Quadram Institute Science Health Food Innovation

Exploring the use of metagenomics in FMT

Microbes, Microbiomes and Bioinformatics MRC DTP MRC
CLIMB-BIG-DATA
Masterclass July 2023

Lizbeth Sayavedra

Postdoctoral Research Scientist (GHFS)

What will you learn?

- **Think about how you can use metagenomics for your own science**
- **Think about how to process your raw sequencing data to create MAGs**
- **Explore a metagenomics data set from FMT using Anvio**

Why use metagenomes as a tool?

- **Who is there?**
- **What are they doing?**

Can you think of a project where you would want to compare two or more metagenomes?

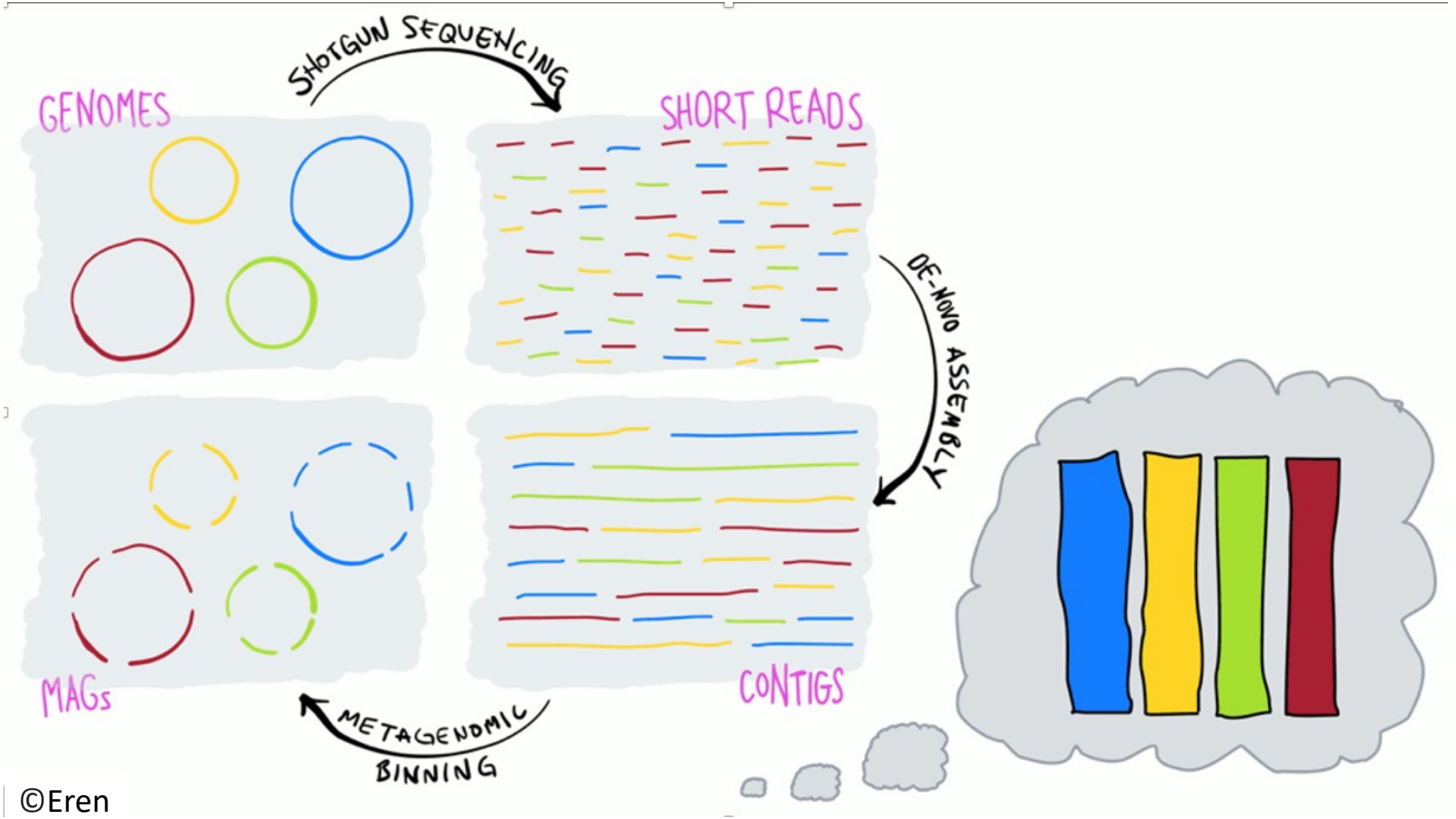
Which steps would you follow to go from reads to MAGs?

Plenty of software around, which has it's pros and cons

Time to think about how to process raw sequencing data from metagenomes

- **Clean the reads**
- **Get rid of the host**
- **Taxonomic profiling**
- **Merge the sequencing data**
- **Generate a consensus assembly**
- **Make a stacked abundance plot**
- **Create metagenome-assembled genomes (MAGs)**

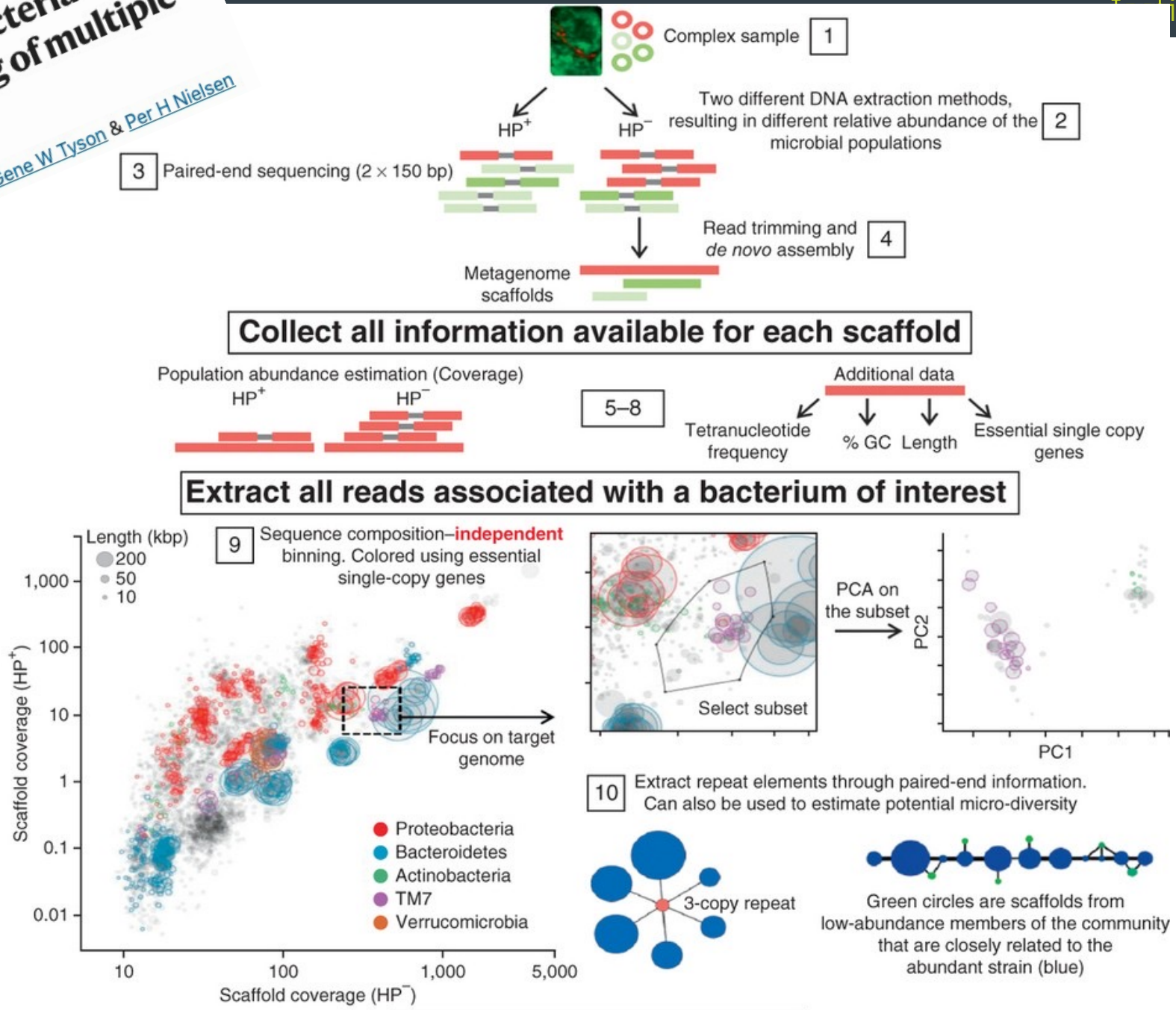
MAG creation



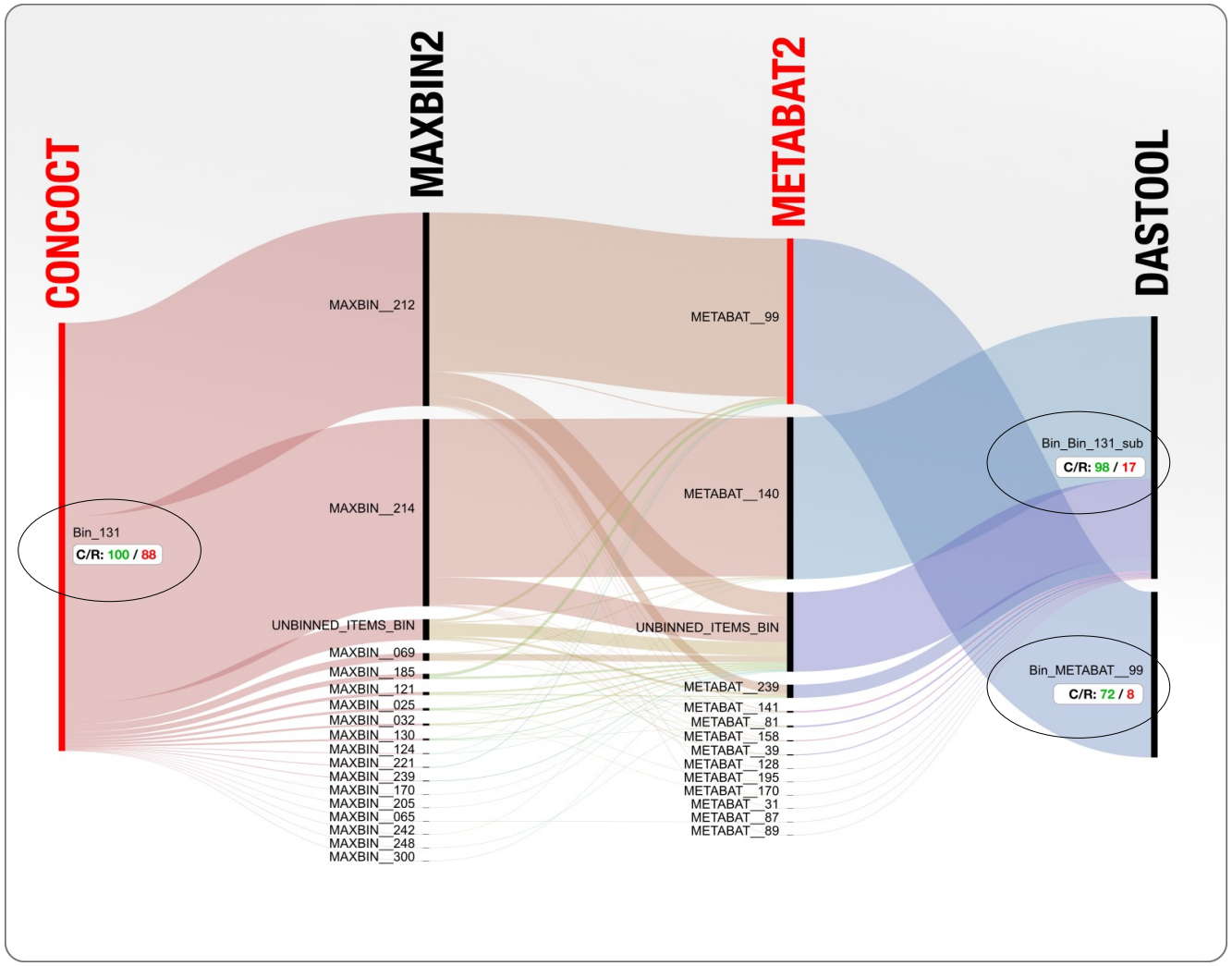
©Eren

Genome sequences of rare, uncultured bacteria obtained by differential coverage binning of multiple metagenomes

Mads Albertsen, Philip Hugenholtz, Adam Skarshewski, Kåre L Nielsen, Gene W Tyson & Per H Nielsen

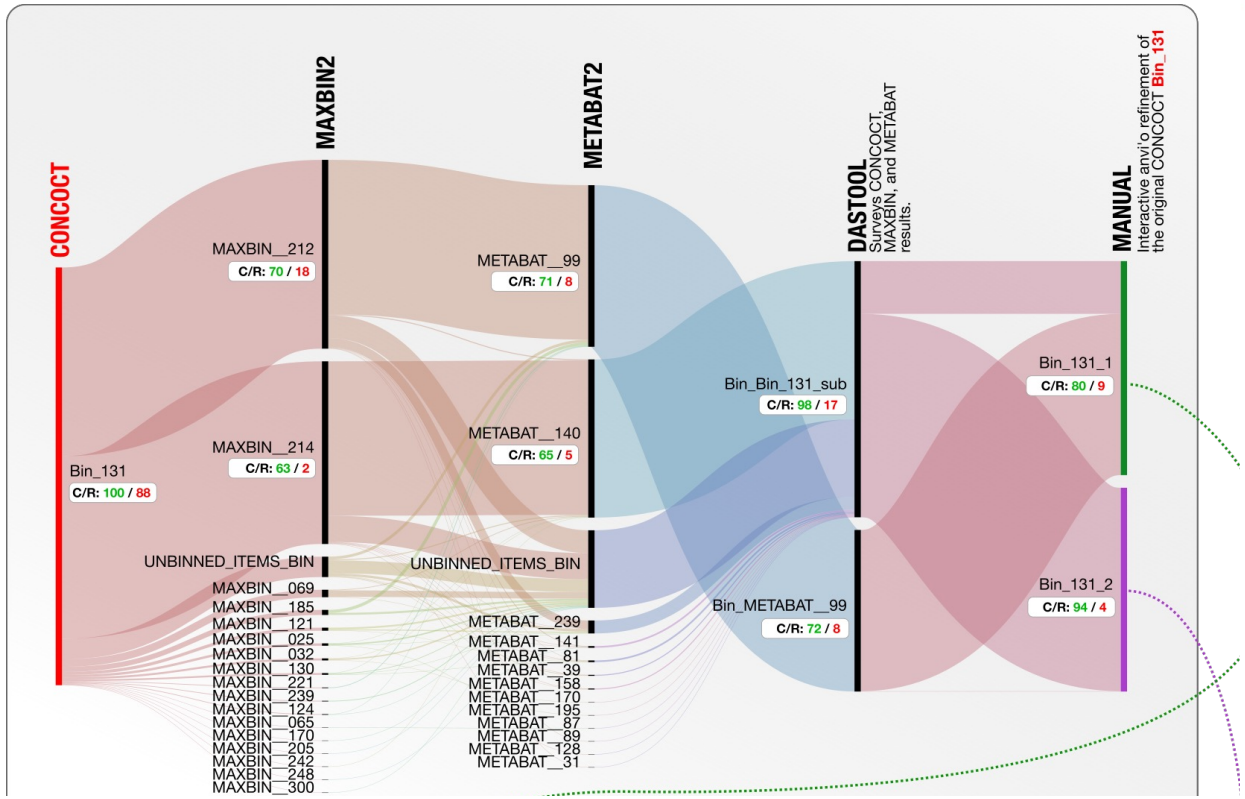


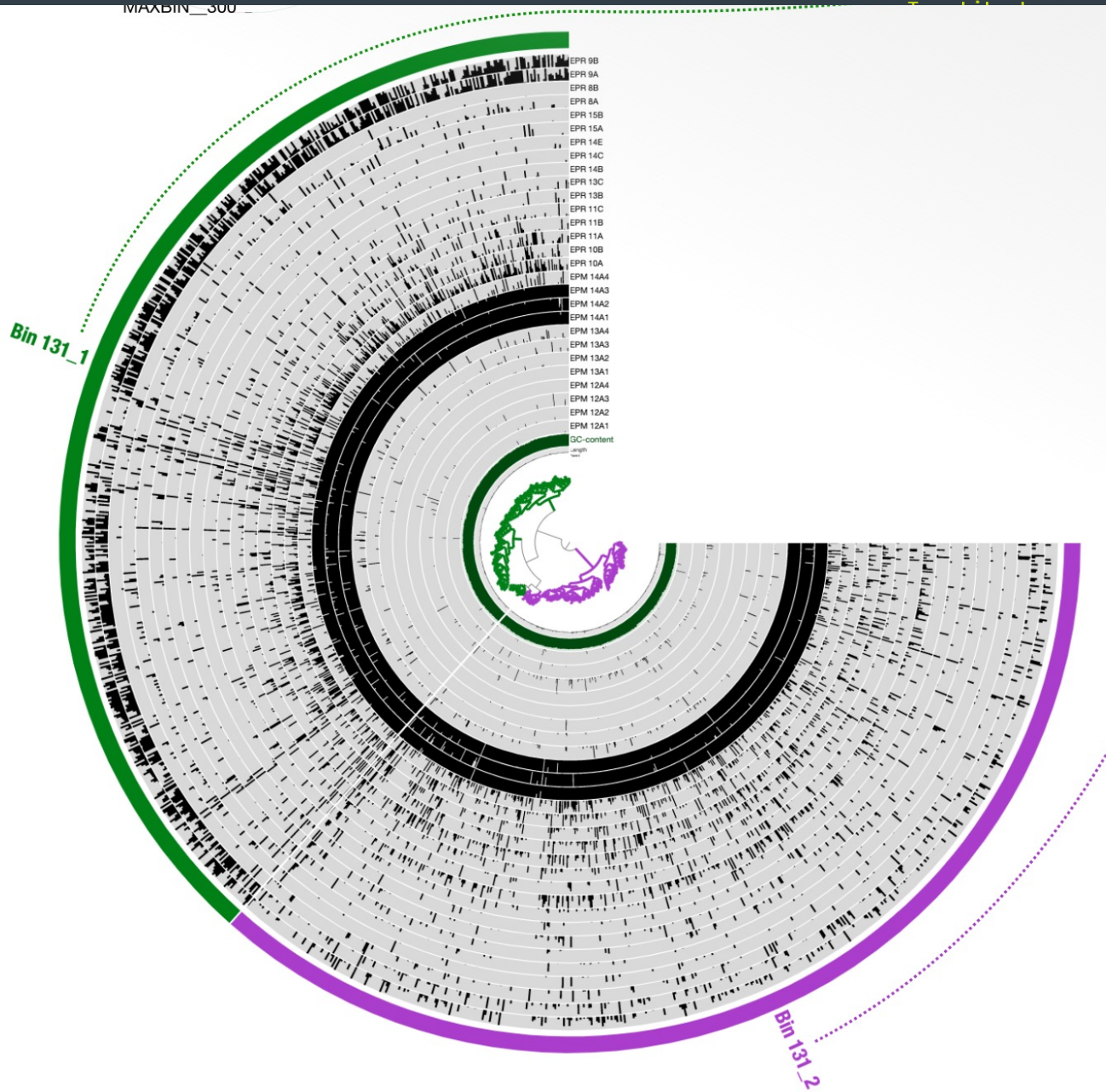
Be aware that MAG creation is not perfect



Different automatic binning tools will decide on different bin creation when the data sets are not easy to tease apart

Manual binning can help refining MAGs





©Eren

Why do we care about quality of MAGs?

e.g.

- To predict the functional potential of a bacterium of interest, we need high quality MAGs
- If getting rid of low-quality genomes, there is a risk of loosing taxa that play an important role

nature biotechnology

RESOURCE

https://doi.org/10.1038/s41587-020-0603-3

Check for updates

OPEN

A unified catalog of 204,938 reference genomes from the human gut microbiome

Alexandre Almeida[1,2]✉, Stephen Nayfach[3,4], Miguel Boland[1], Francesco Strozzi[5], Martin Beracochea[1], Zhou Jason Shi[6,7], Katherine S. Pollard[6,7,8,9,10,11], Ekaterina Sakharova[1], Donovan H. Parks[12], Philip Hugenholtz[12], Nicola Segata[13], Nikos C. Kyrpides[3,4] and Robert D. Finn[1]✉

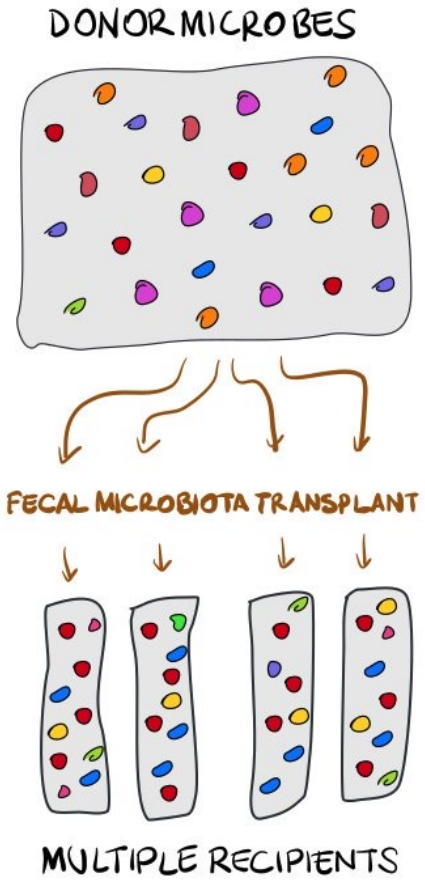
Comprehensive, high-quality reference genomes are required for functional characterization and taxonomic assignment of the human gut microbiota. We present the Unified Human Gastrointestinal Genome (UHGG) collection, comprising 204,938 non-redundant genomes from 4,644 gut prokaryotes. These genomes encode >170 million protein sequences, which we collated in the Unified Human Gastrointestinal Protein (UHGP) catalog. The UHGP more than doubles the number of gut proteins in comparison to those present in the Integrated Gene Catalog. More than 70% of the UHGG species lack cultured representatives, and 40% of the UHGP lack functional annotations. Intraspecies genomic variation analyses revealed a large reservoir of accessory genes and single-nucleotide variants, many of which are specific to individual human populations. The UHGG and UHGP collections will enable studies linking genotypes to phenotypes in the human gut microbiome.

Can you think of strategies that avoid MAG reconstruction?

Quadram Institute

Metabolic independence drives gut microbial colonization and resilience in health and disease

Andrea R. Watson, Jessika Füssel, Iva Veseli, Johanna Zaal DeLongchamp, Marisela Silva, Florian Trigodet, Karen Lolans, Alon Shaiber, Emily Fogarty, Joseph M. Runde, Christopher Quince, Michael K. Yu, Arda Söylev, Hilary G. Morrison, Sonny T. M. Lee, Dina Kao, David T. Rubin, Bana Jabri, Thomas Louie & A. Murat Eren



Key points:

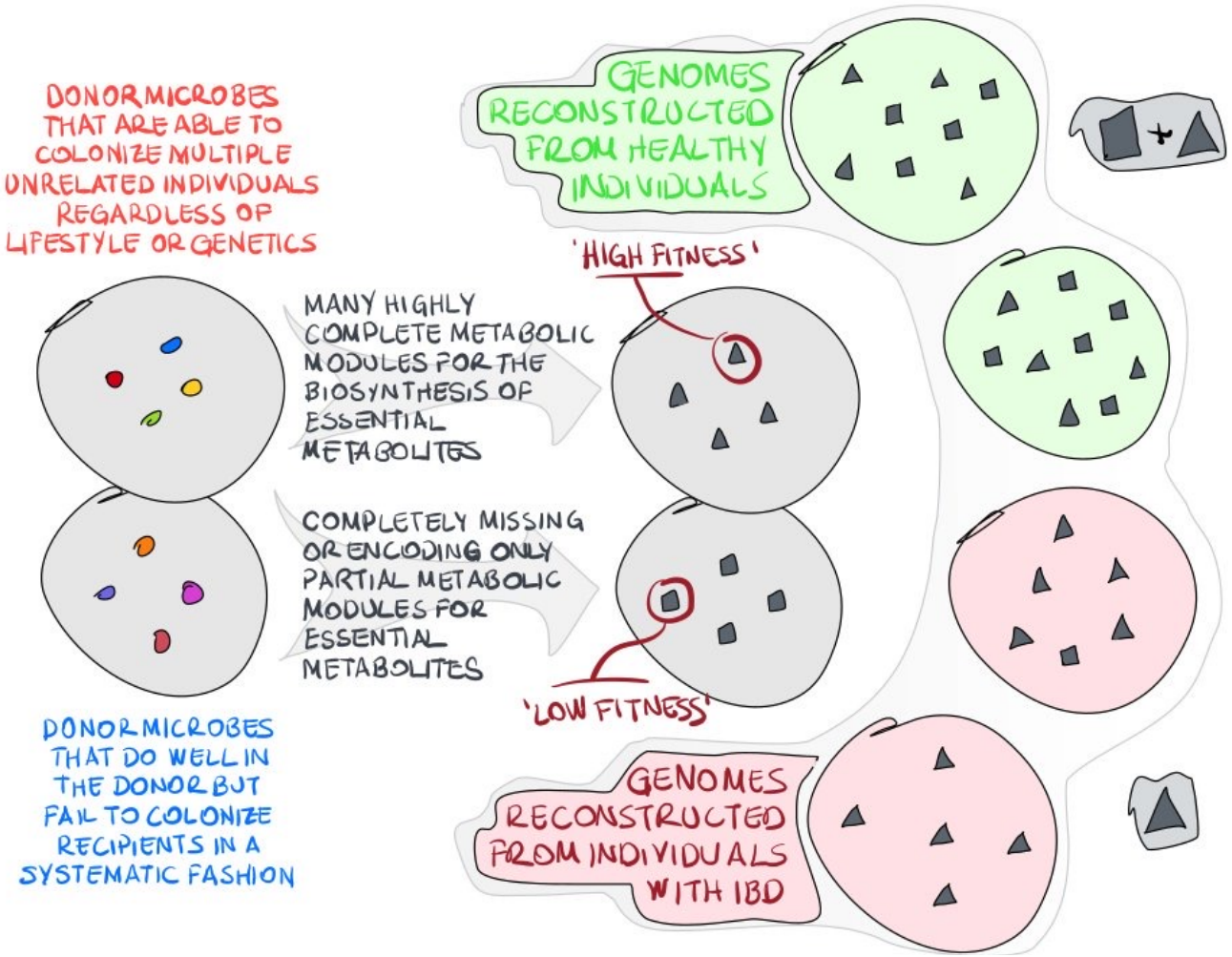
- **Healthy** individuals often have a **more diverse microbiota** compared to unhealthy individuals
- Metabolically independent microbes associated with human gut “dysbiosis” may not in fact be causal of disease, but **instead selected for** under stressful conditions
- Bacteria with **greater functional potential** to synthesize essential metabolites are overrepresented in the human gut both **during inflammation and after FMT**

Proposed model

- In a **healthy gut,** microbes that are **metabolically independent co-occur** in harmony
- In a **stressed environment**, microbes that can self-sustain are selected for

Watson et al., 2023

In other words



©Eren

https://merenlab.org/data/fmt-gut-colonization

Practical

Let's look at the FMT data of the donors and the recipients in detail with Anvio. We will only take a look at their abundance distribution, not the functional potential

For detailed instructions go to:

https://github.com/lsayaved/Masterclass_QIB2023

Questions:
a) From a randomly selected single genome:

1. Can you identify the taxonomic classification, the completeness and the redundancy (contamination)?

b) From the whole collection:
1. Can you change the taxonomic level displayed?
2. What are the dominant taxa at the phylum level for the FMT donor?
3. What are the dominant taxa for the recipients before and after FMT (phylum level)?