Metagenomics 101

Session X: Metatranscriptomics

Anna Heintz-Buschart

April 2022

Metagenomics (+ other omics) pipeline


contaminant removal

taxonomic profile

quality control

reads

mOTU marker genes

assembly

gene predictions

functional annotations

functional categories

alignment

functional profile

binning

genome reconstructions

growth inference

variant calling

UNIVERSITY OF AMSTERDAM
Life Sciences

IMP3
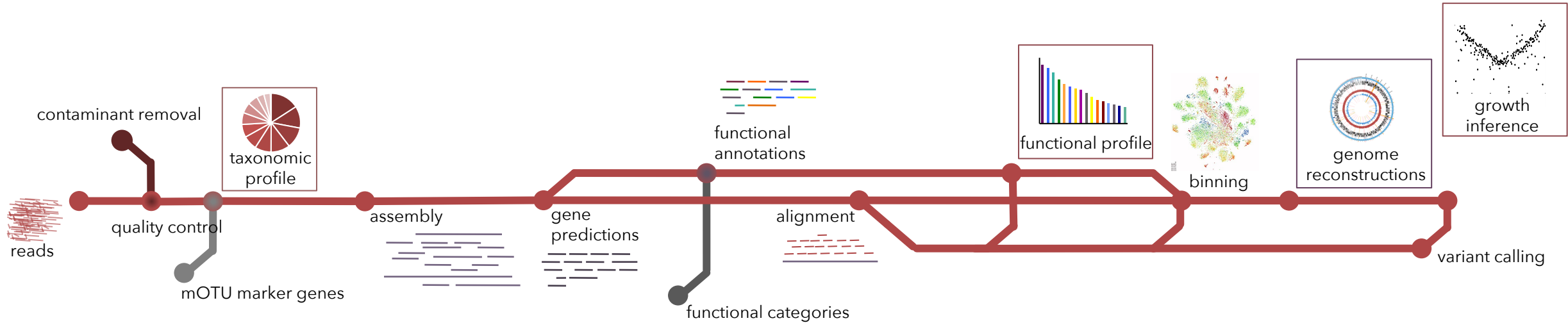
Today

- Why metatranscriptomics?
- Challenges in metatranscriptomics
- Workflows for metatranscriptomics

Metaomics

rRNA amplicon studies:
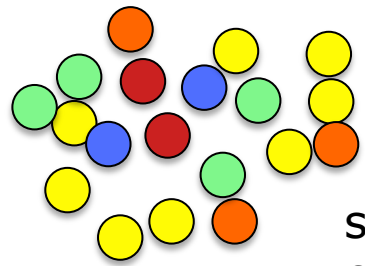
structure of community A
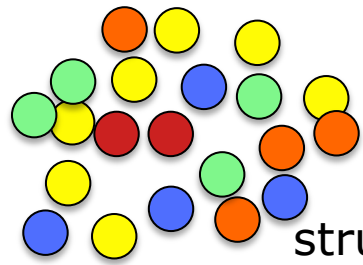
structure of community B

metagenomics:

potential of community A

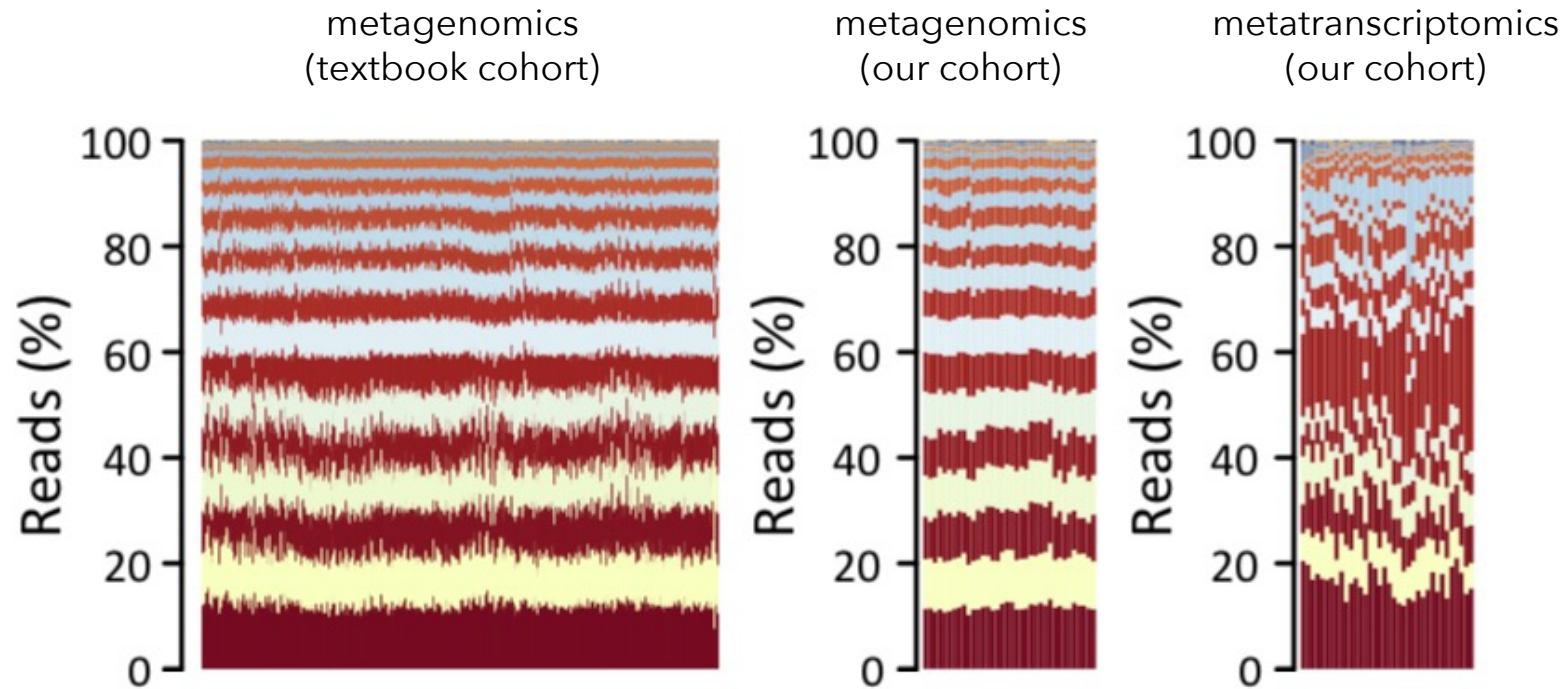potential of community B

functional omics:

activity of community A

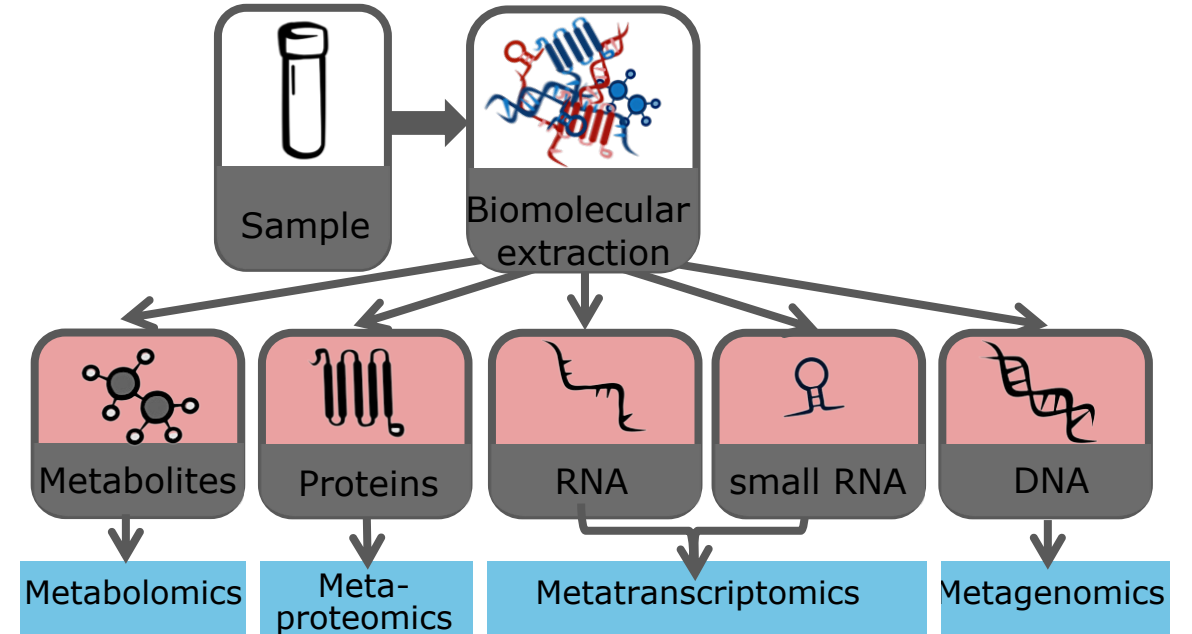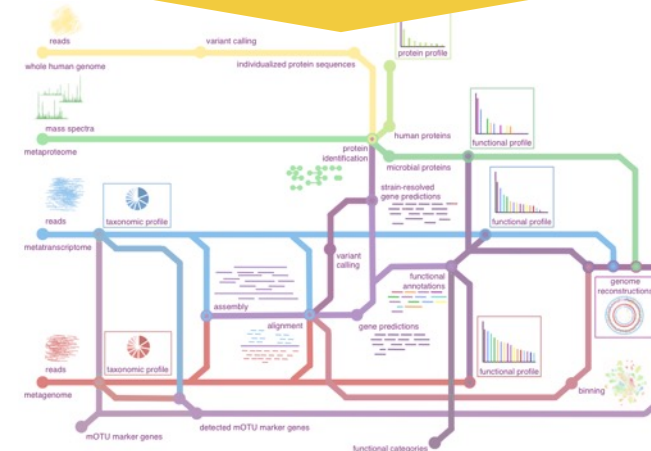activity of populations within communities

activity of community B

Metaomics

metagenomics
(textbook cohort)

metagenomics
(our cohort)

metatranscriptomics
(our cohort)

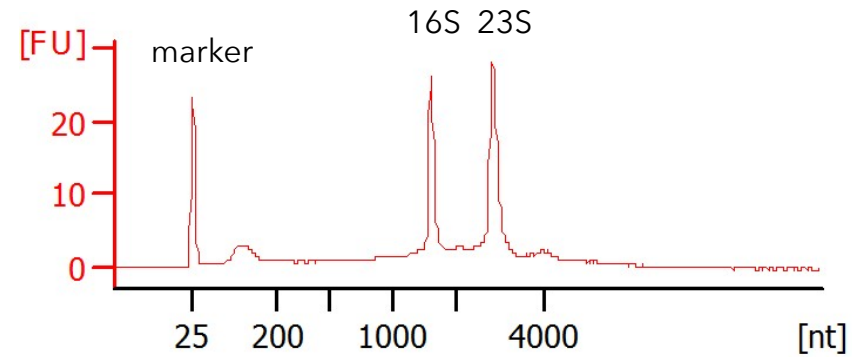- metatranscriptomic profiles are much more variable

UNIVERSITY OF AMSTERDAM
Life Sciences

5

Sample to data

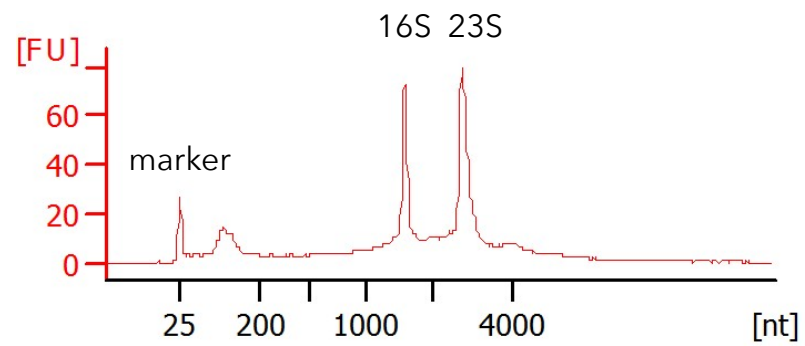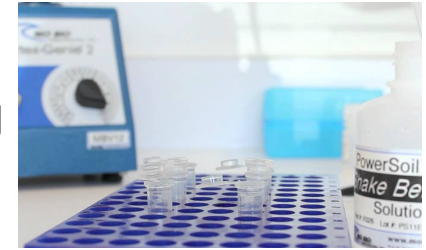# Sample to data

sampling + preservation

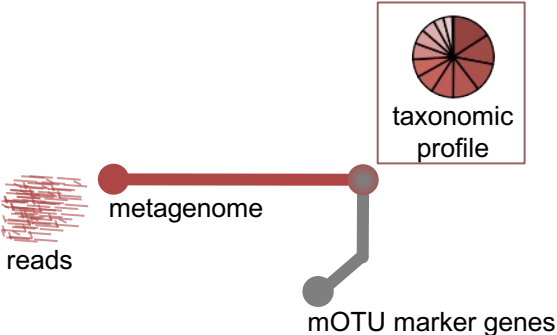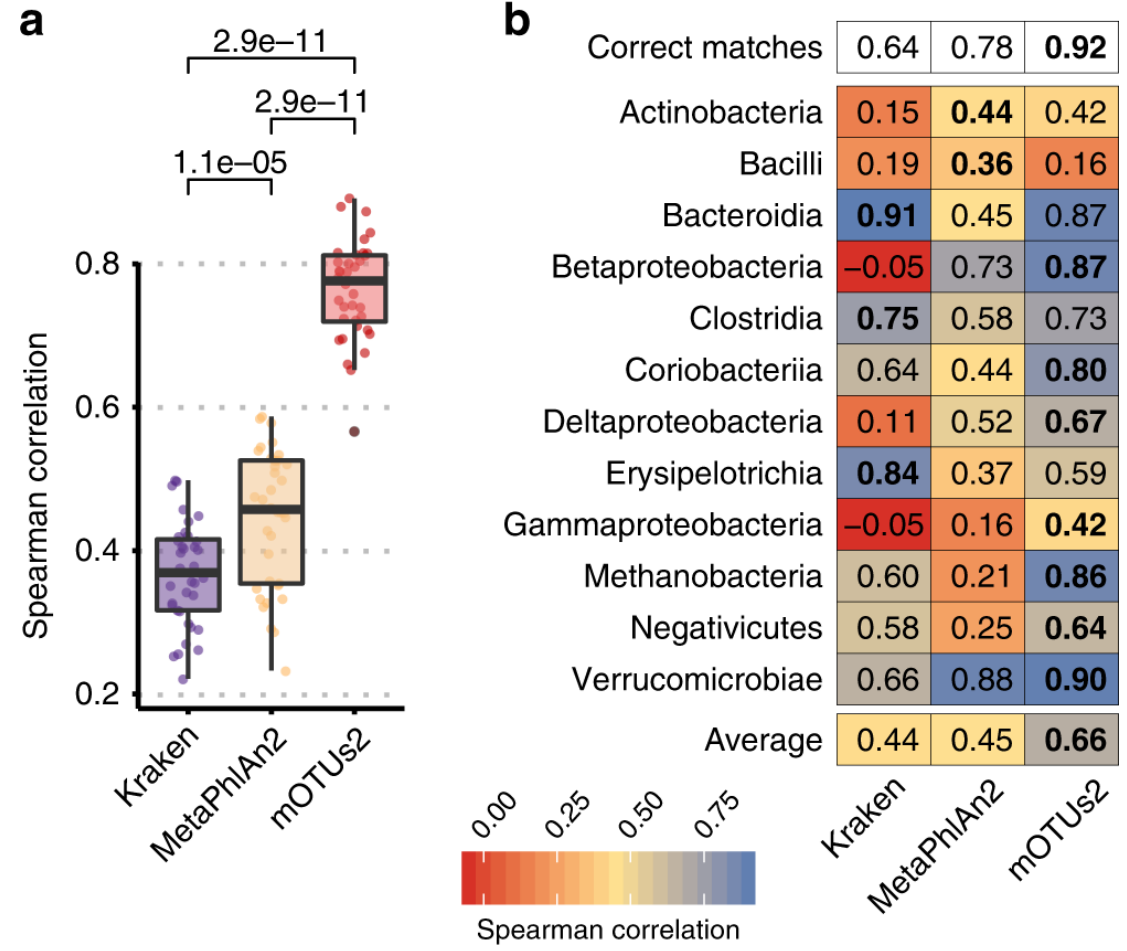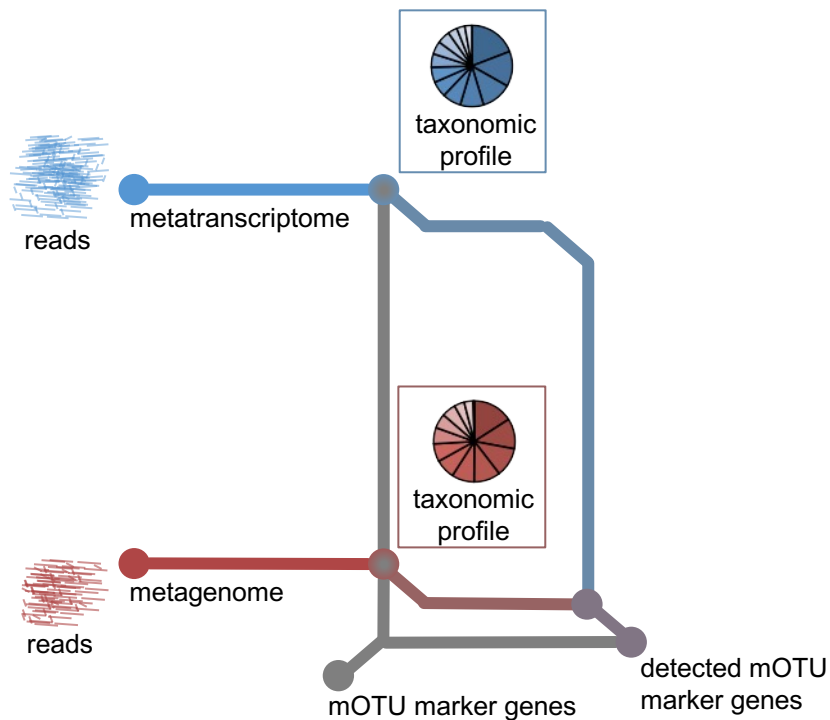sample processing

16S 23S

[FU]

60
40
20
0

marker

25   200   1000   4000   [nt]

16S 23S

[FU]

marker

20

10

0

25   200   1000   4000   [nt]

UNIVERSITY OF AMSTERDAM
Life Sciences

Multiomics integration
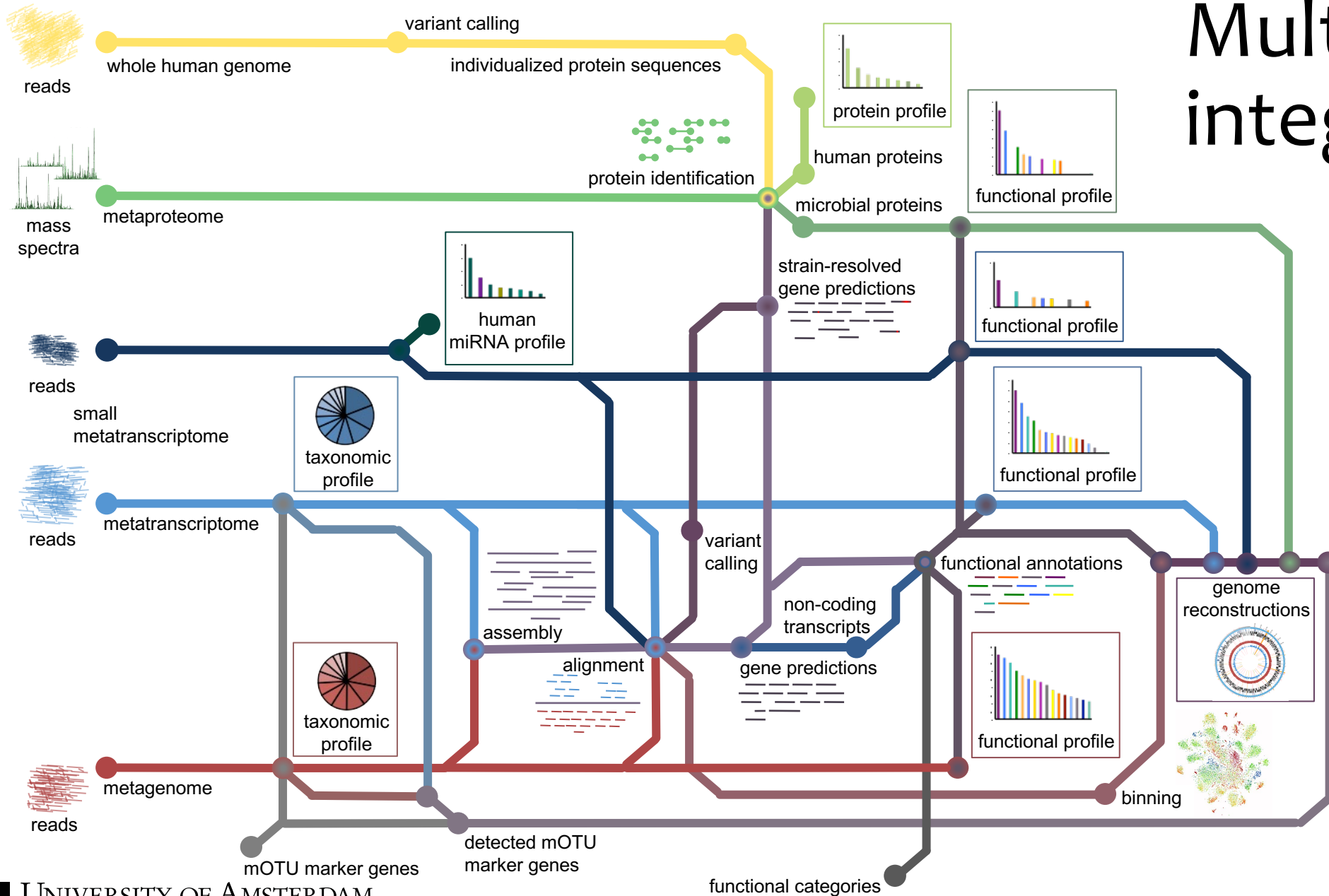
Workflow

taxonomic
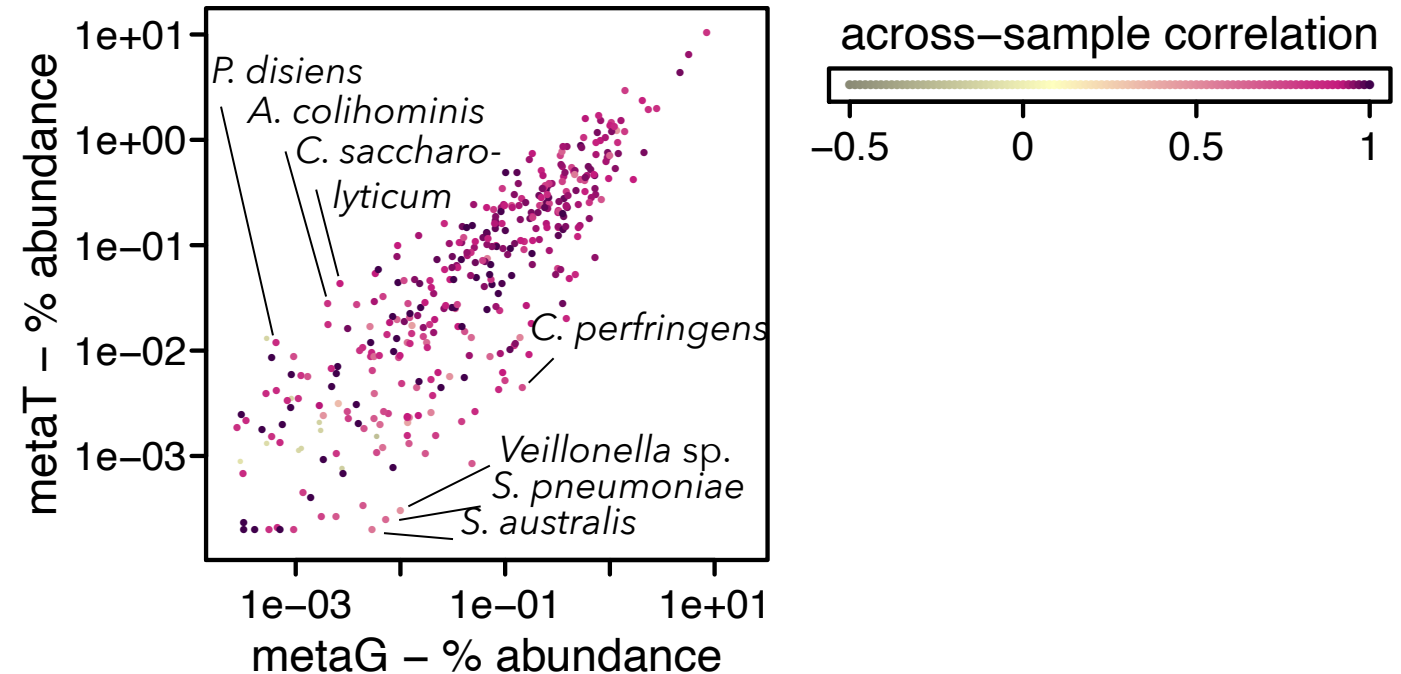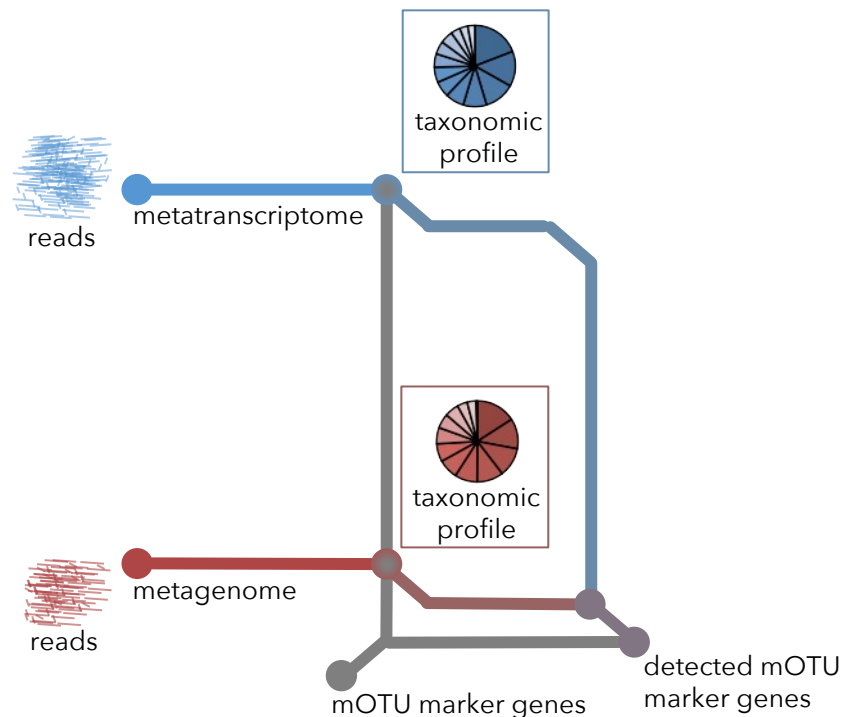profile
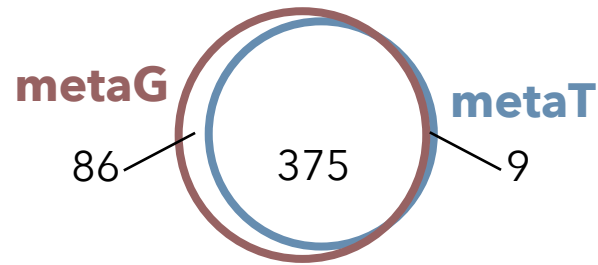
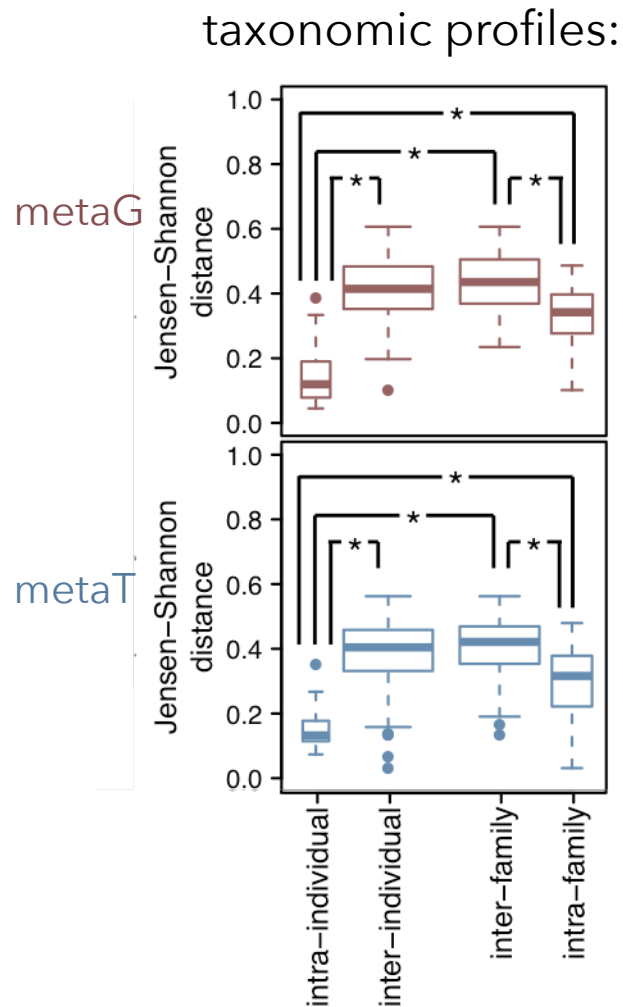metagenome

reads

mOTU marker genes

Taxonomic profiles?

Milanese, A., Mende, D.R., Paoli, L. et al. Microbial abundance, activity and population genomic profiling with mOTUs2. Nat Commun 10, 1014 (2019). https://doi.org/10.1038/s41467-019-08844-4
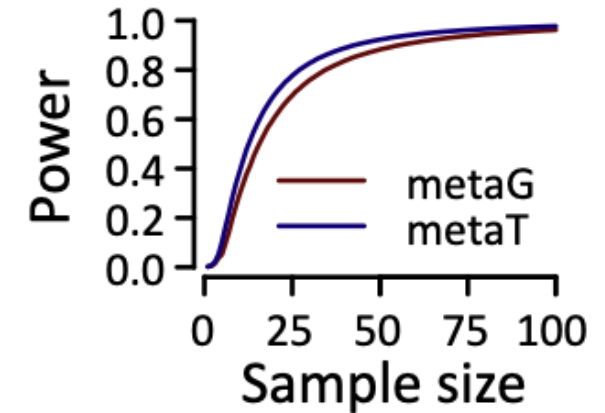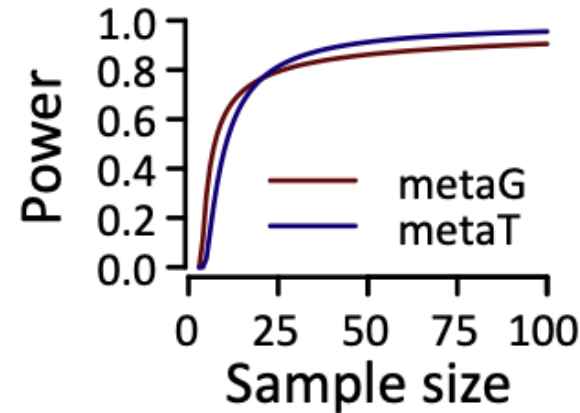
MetaT vs metaG – taxonomic profiles

• **taxonomic profiles** at metagenomic **and** metatranscriptomic levels are **very similar**

Heintz-Buschart *et al.*, Nature Microbiology, 2017
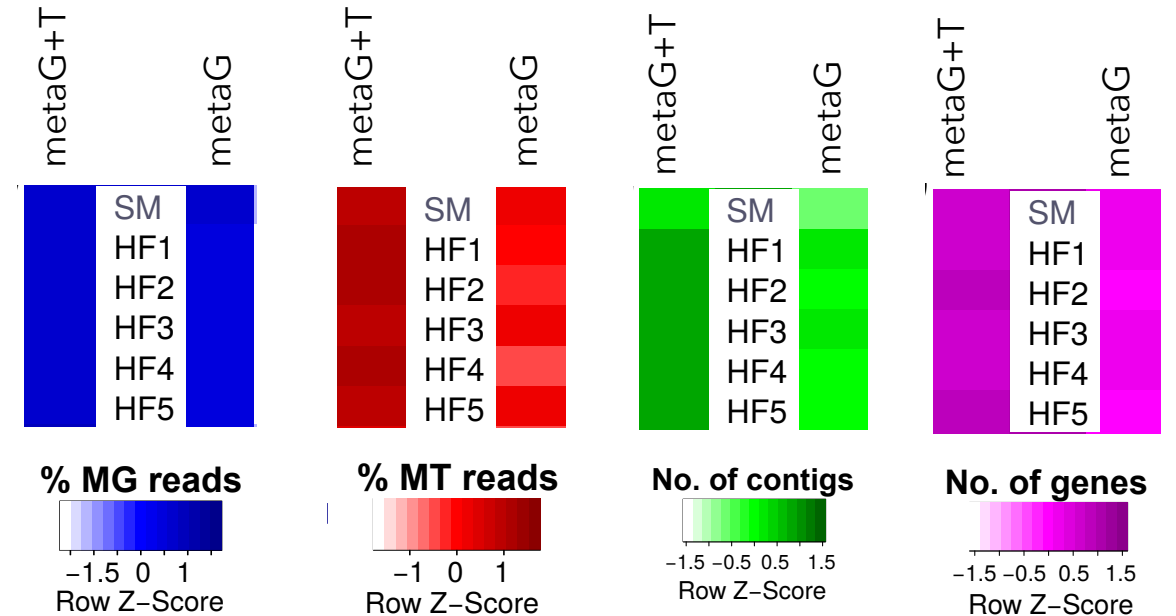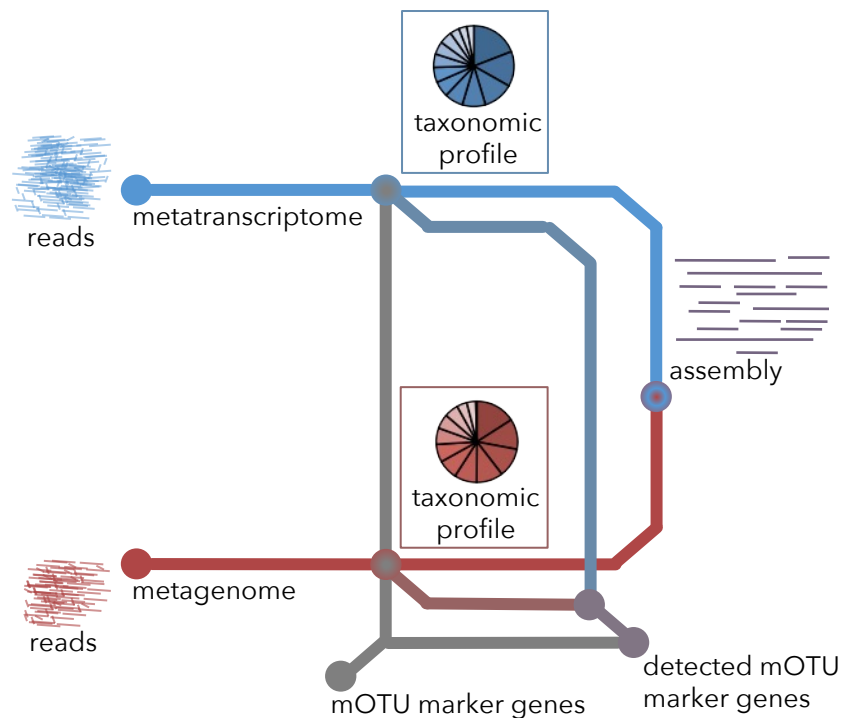
Patterns

taxonomic profiles:

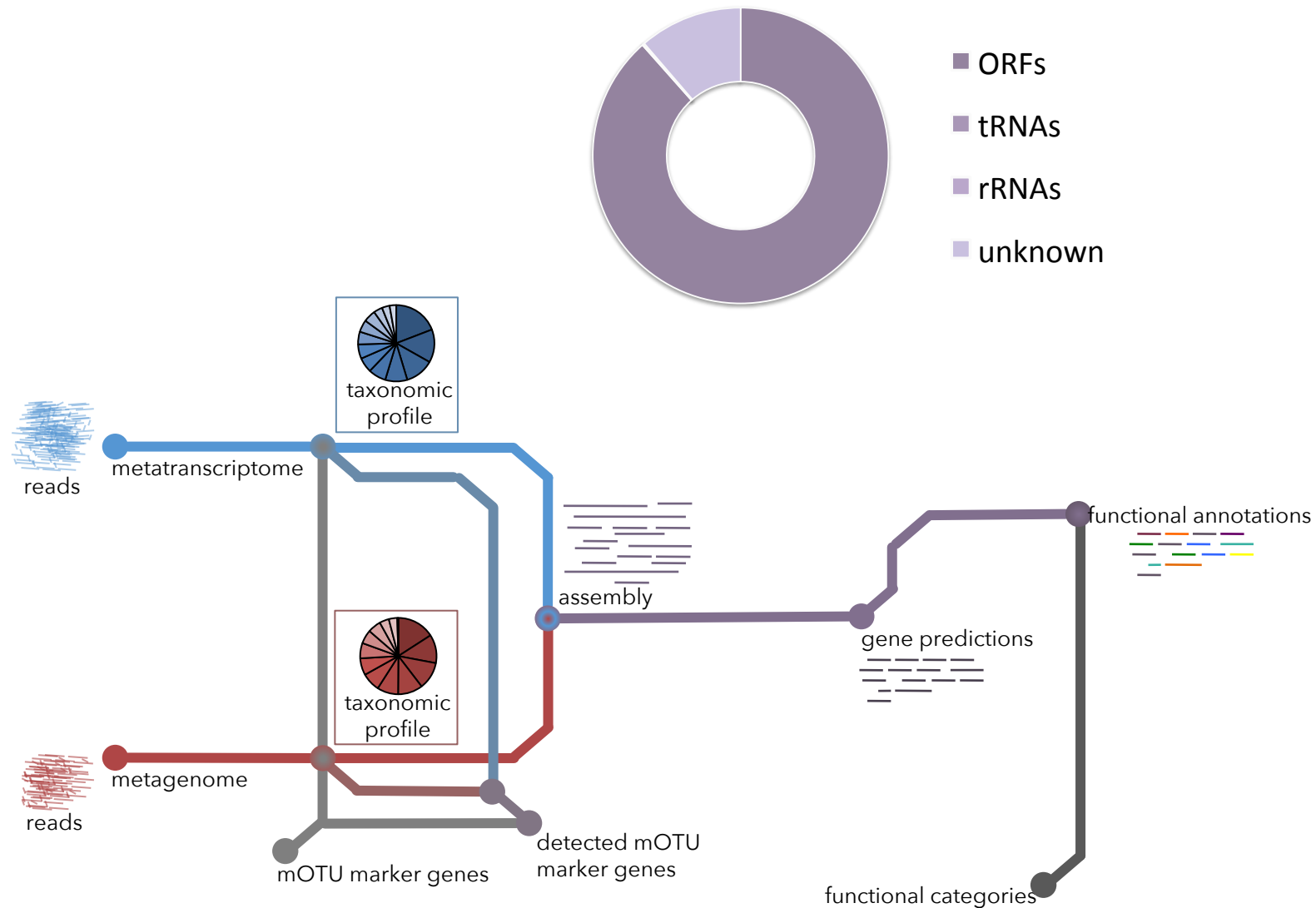- metatranscriptomics-based taxonomic profiles show similar results to metagenomics

- metatranscriptomics have similar power to metagenomics
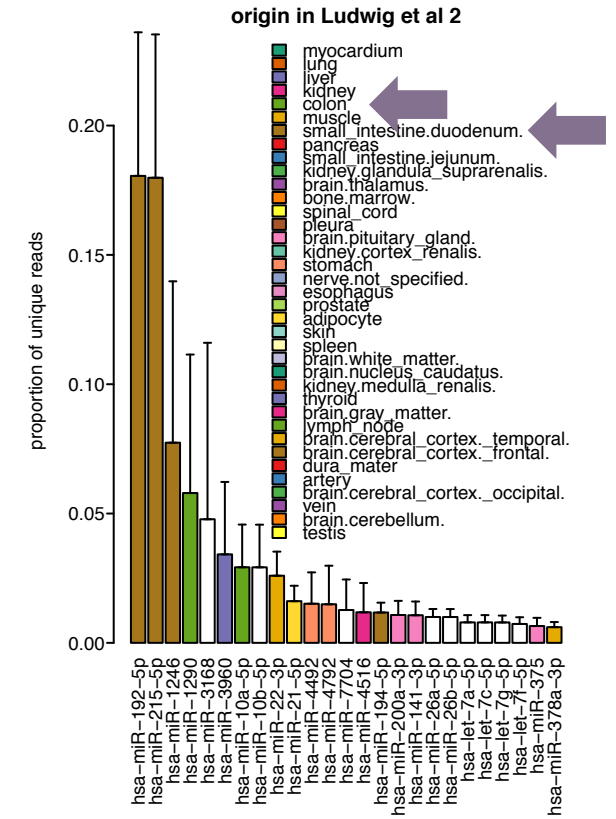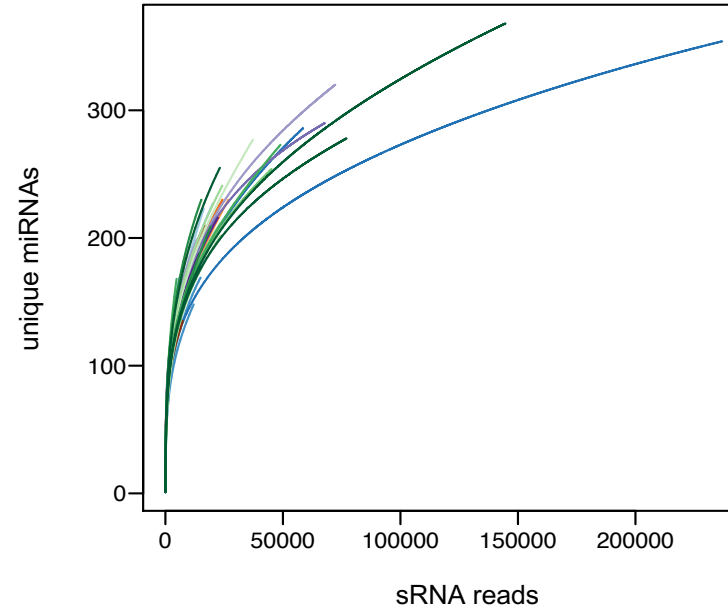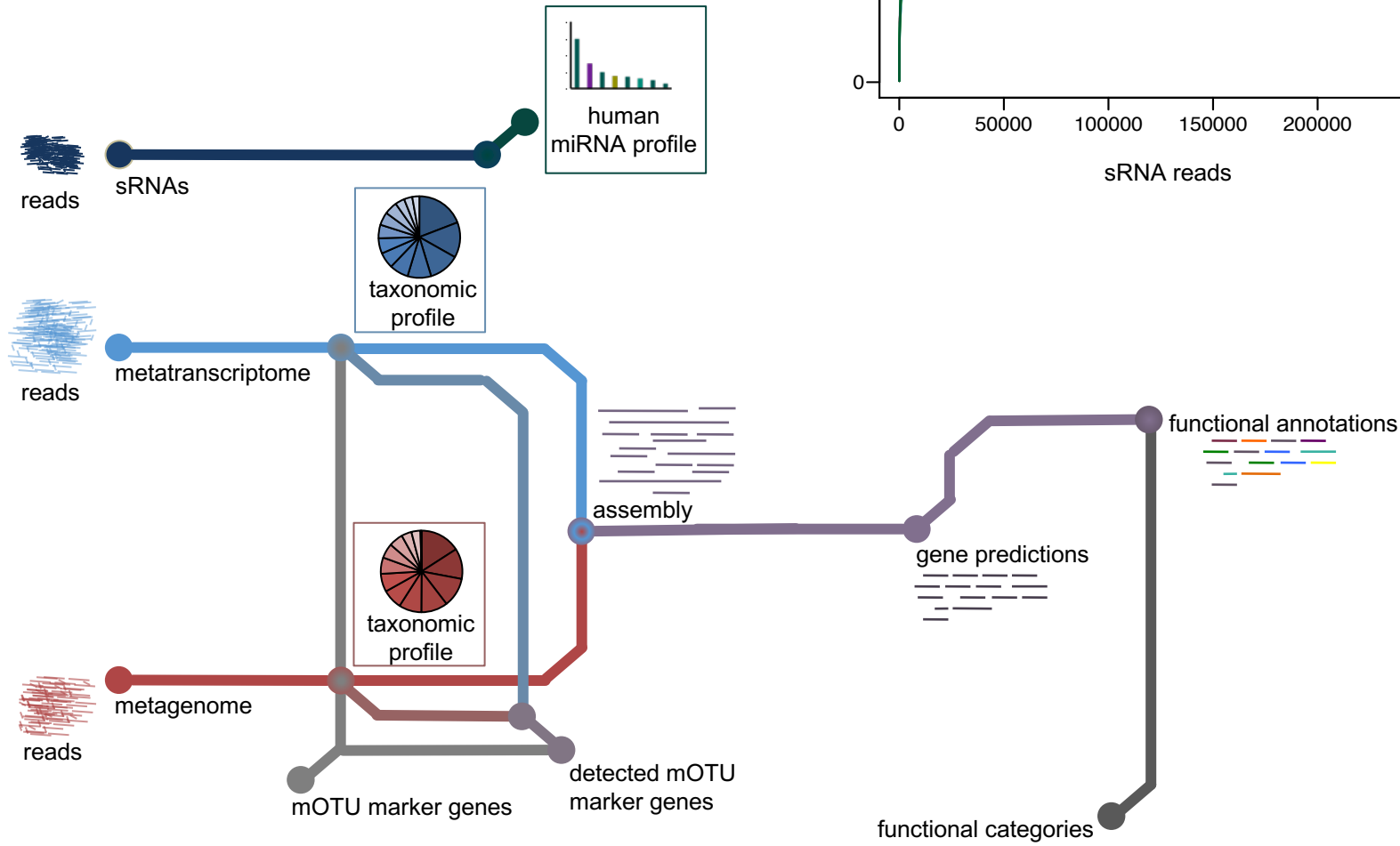
Workflow

reads

metatranscriptome

taxonomic profile

assembly

reads

metagenome

mO...

metaG+T

metaG

SM
HF1
HF2
HF3
HF4
HF5

of genes

0.5 0.5 1.5

Z-Score

Workflow

- ORFs
- tRNAs
- rRNAs
- unknown

taxonomic profile

reads

metatranscriptome

taxonomic profile

functional annotations

assembly

gene predictions

metagenome

reads

detected mOTU marker genes

mOTU marker genes
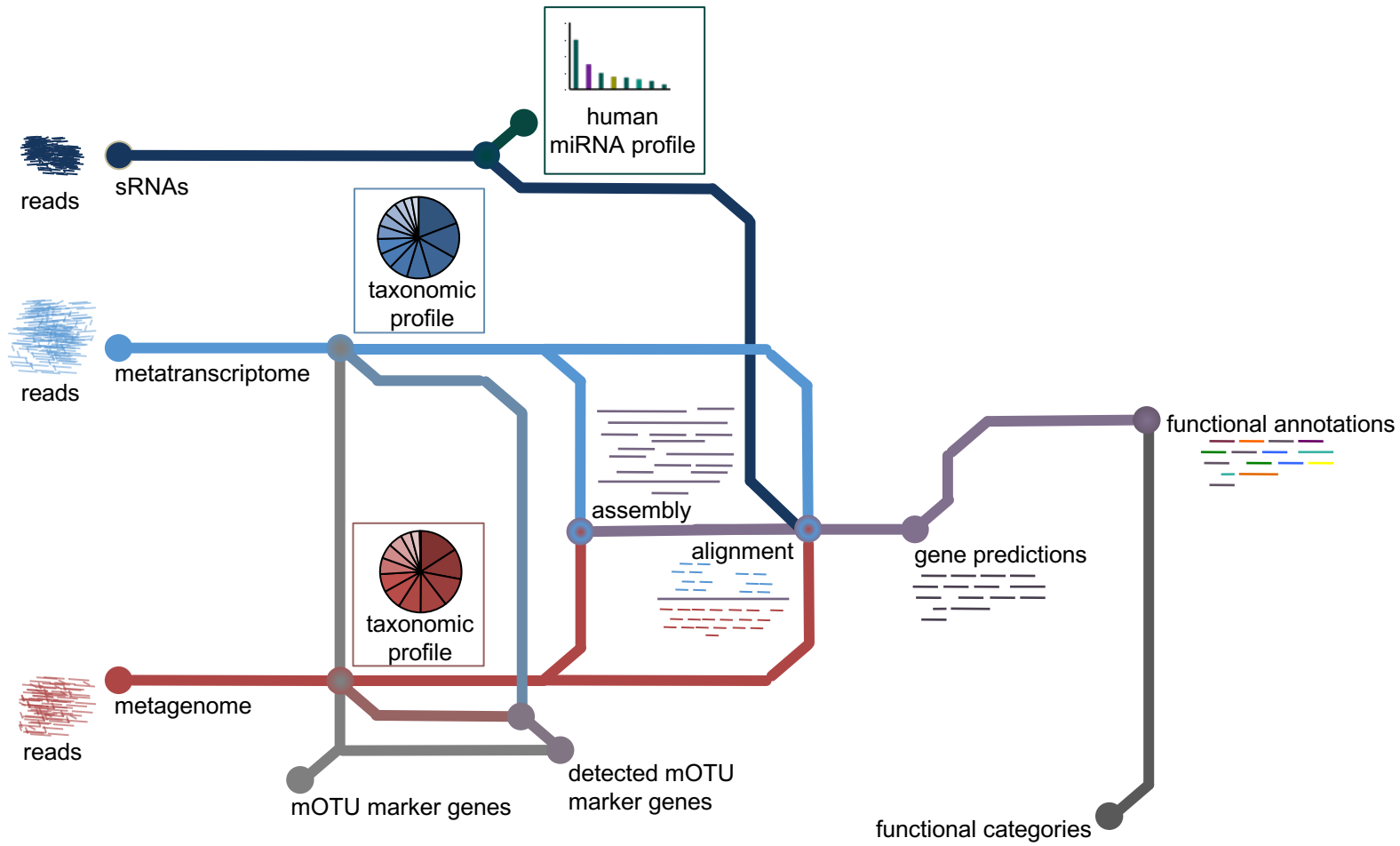
functional categories
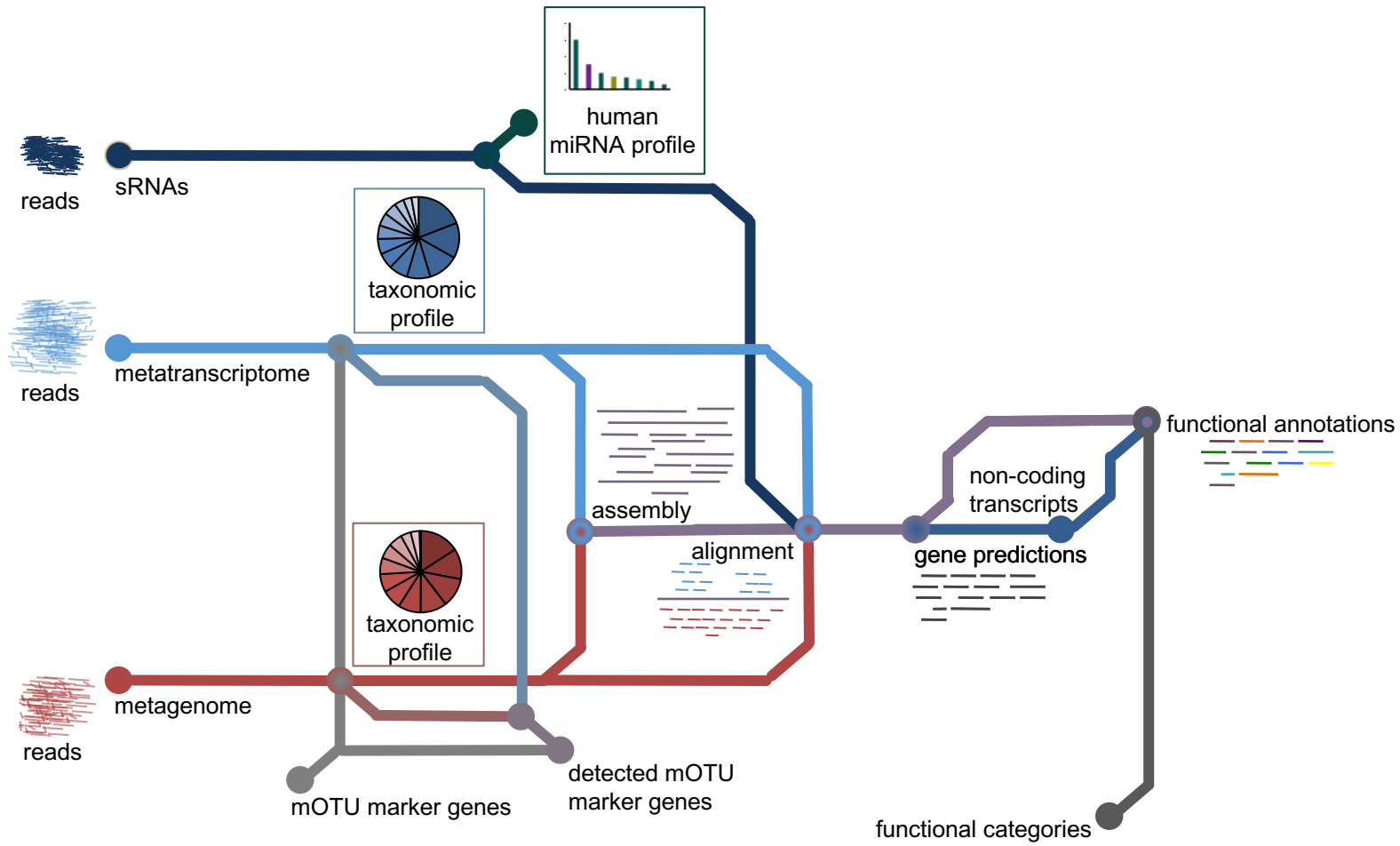
Workflow

origin in Ludwig et al 2

Ludwig *et al.*, Distribution of miRNA expression across human tissues. Nucleic Acids Res. 2016 May 5;44(8):3865-77. doi: 10.1093/nar/gkw116

15

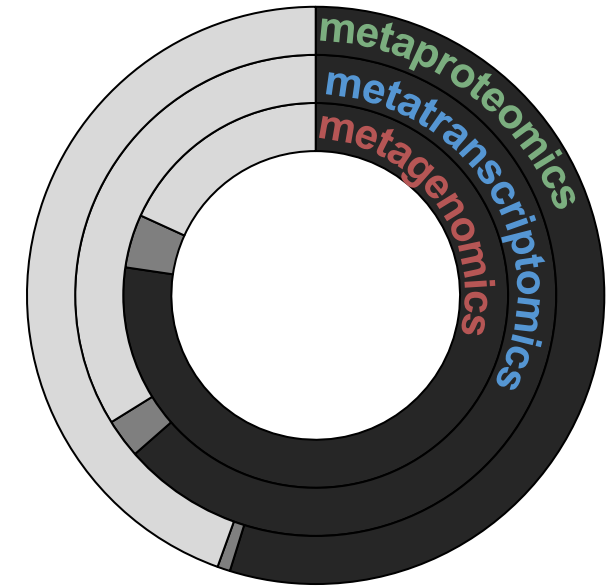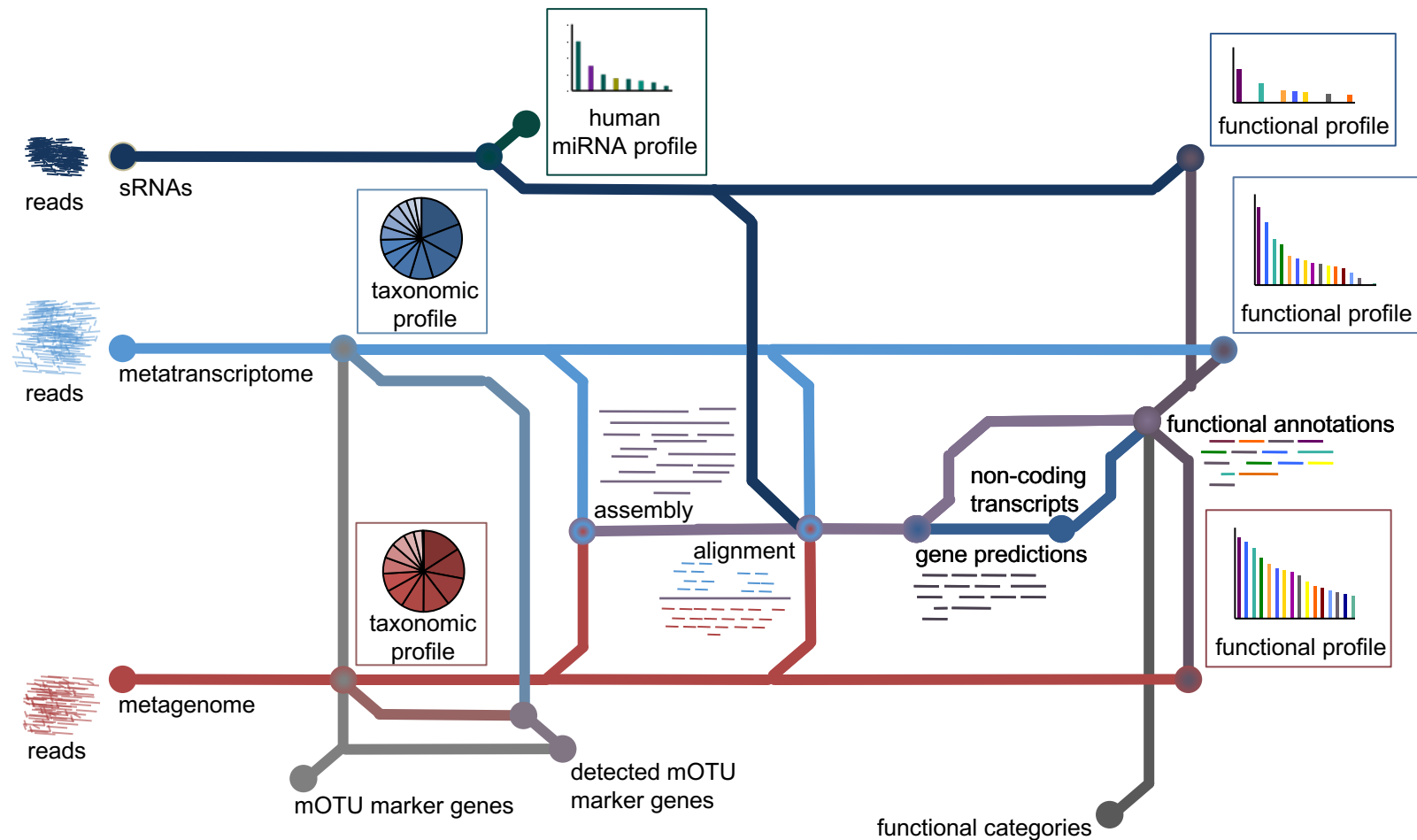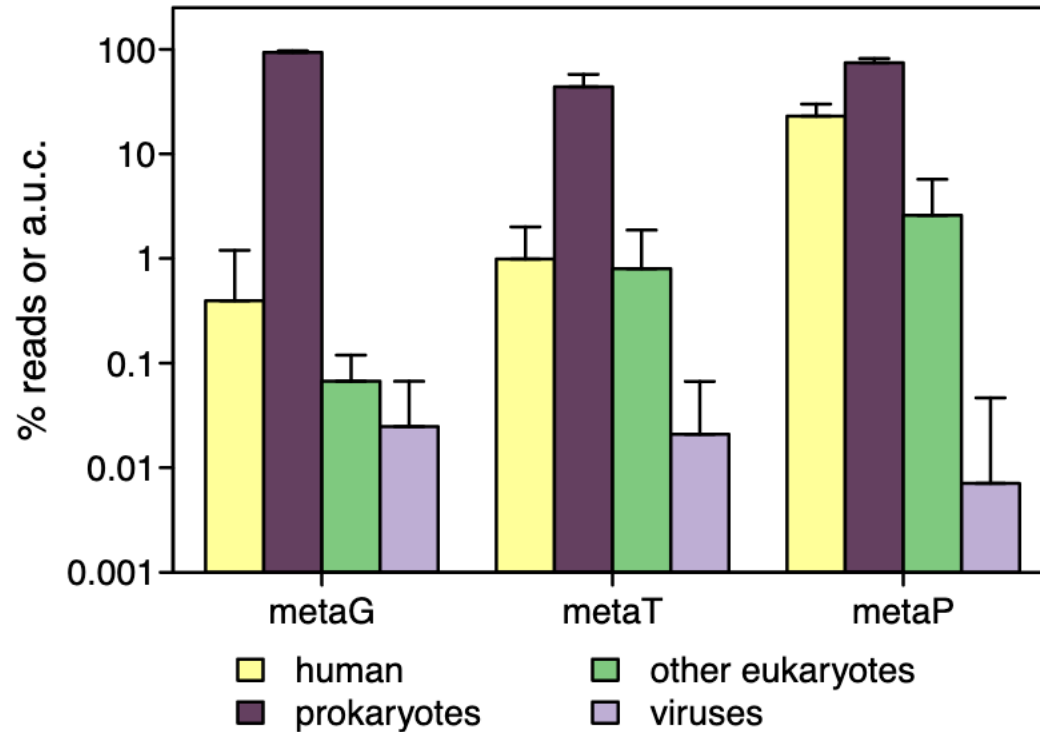# Workflow

Workflow

reads
sRNAs

human
miRNA profile

taxonomic
profile

reads
metatranscriptome

reads
metagenome

taxonomic
profile

assembly

alignment

mOTU marker genes

detected mOTU
marker genes

non-coding
transcripts

gene predictions

functional annotations
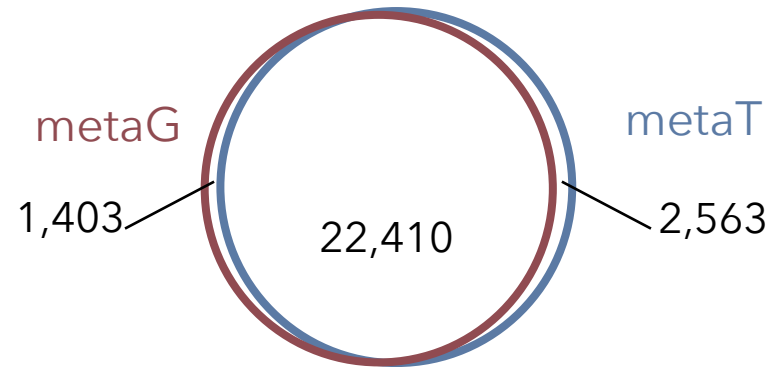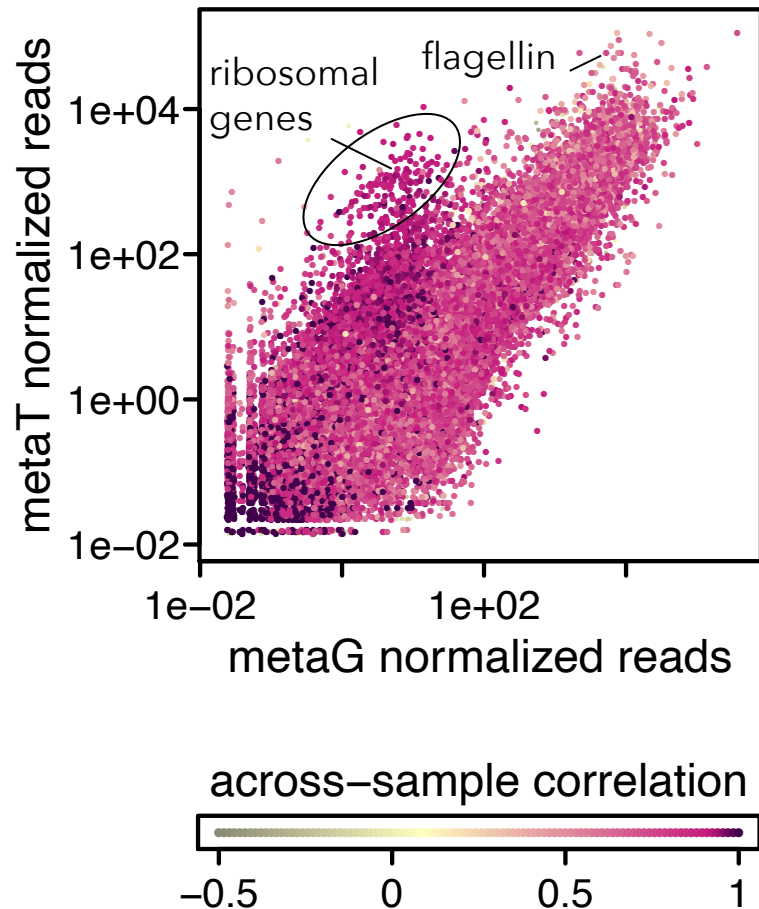
functional categories

Workflow

MetaT vs metaG – what do we detect?

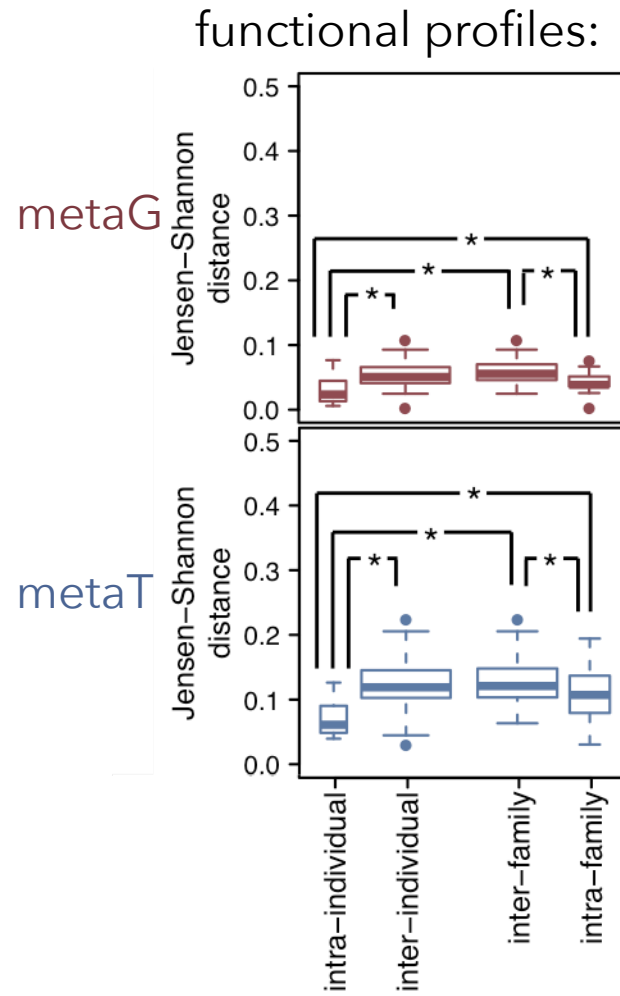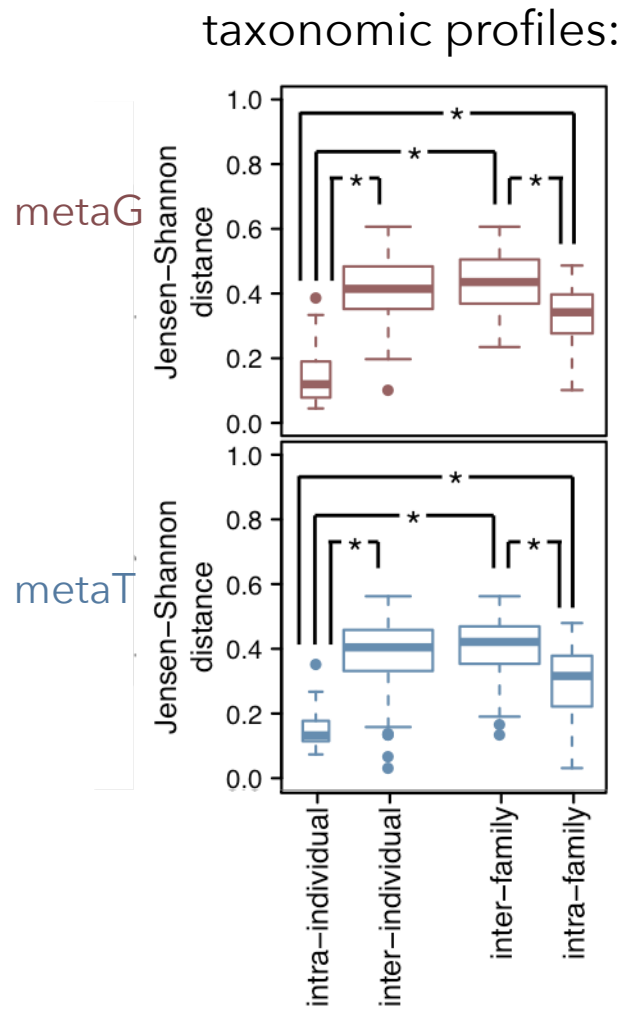- overall similar origins, but larger organisms are more represented

- often more host!

UNIVERSITY OF AMSTERDAM
Life Sciences

MetaT vs metaG – functional profiles

functional profiles:

ribosomal genes

flagellin

metaT normalized reads

1e+04
1e+02
1e+00
1e−02

1e−02   1e+02
metaG normalized reads

across−sample correlation

−0.5   0   0.5   1
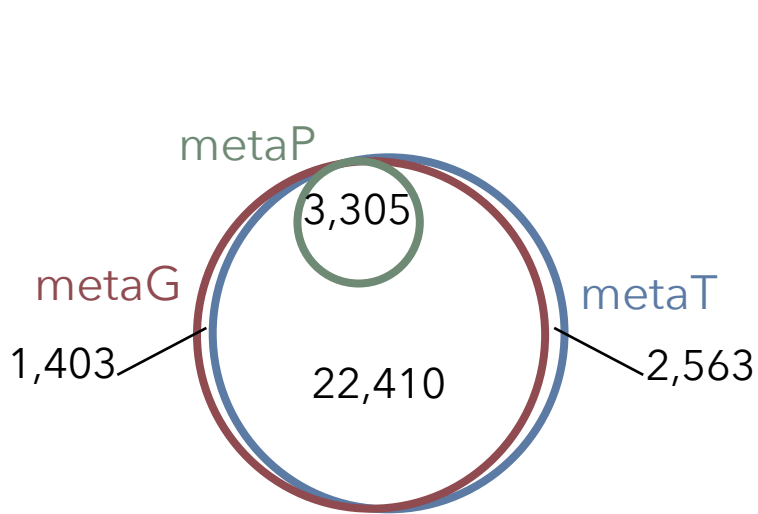
metaG   metaT

1,403   22,410   2,563

o **functional profiles** at metagenomic and metatranscriptomic levels are **less similar than taxonomic profiles**

UNIVERSITY OF AMSTERDAM
Life Sciences

Heintz-Buschart *et al.*, Nature Microbiology, 2017
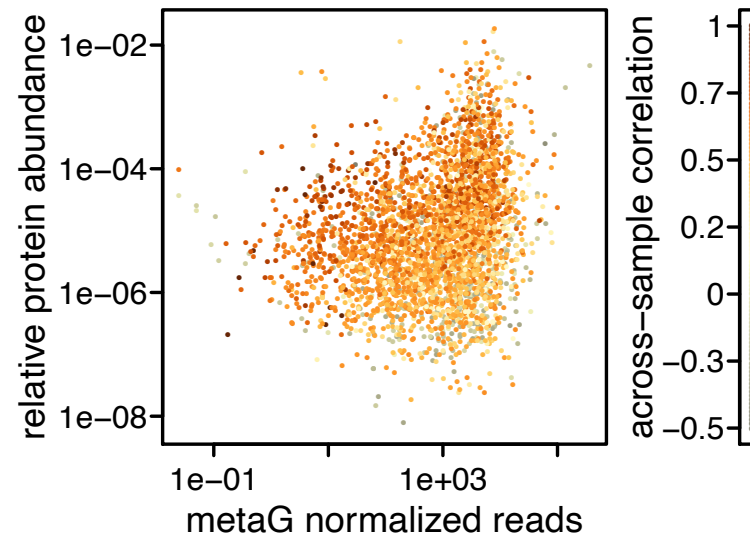
Patterns

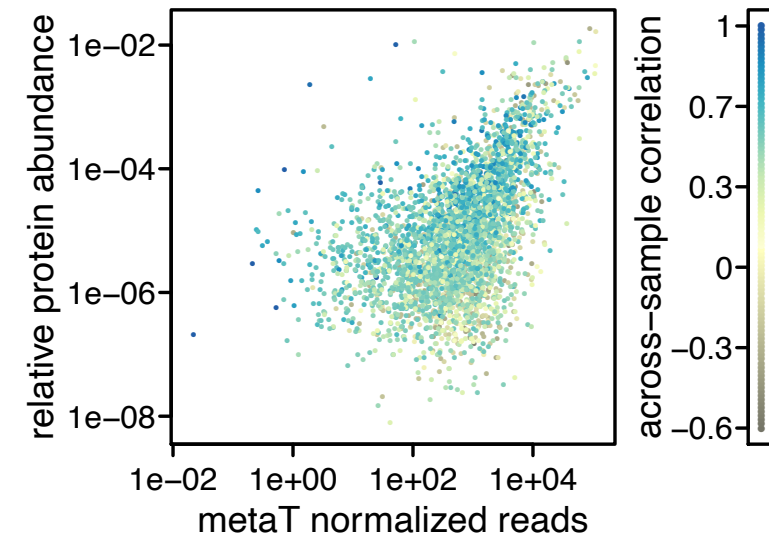taxonomic profiles:

functional profiles:

MetaT vs metaP – functional profiles
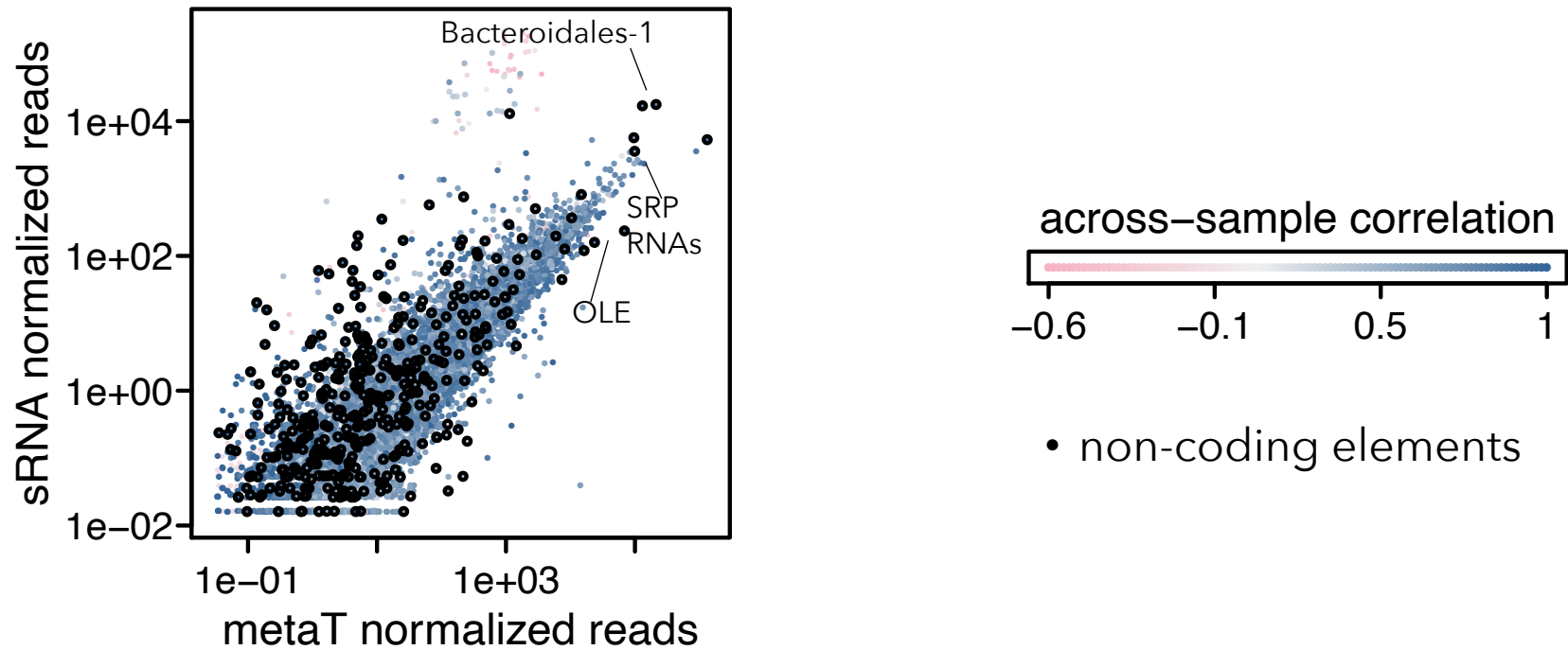

metaP vs metaG:

metaP vs metaT:
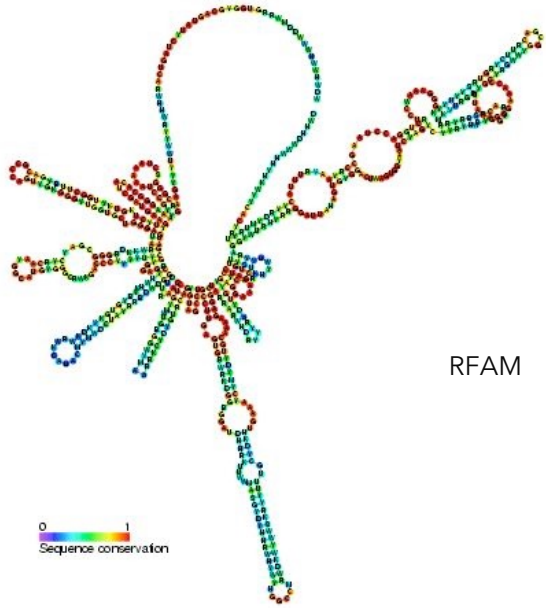
- metaproteome is more related to metatranscriptome
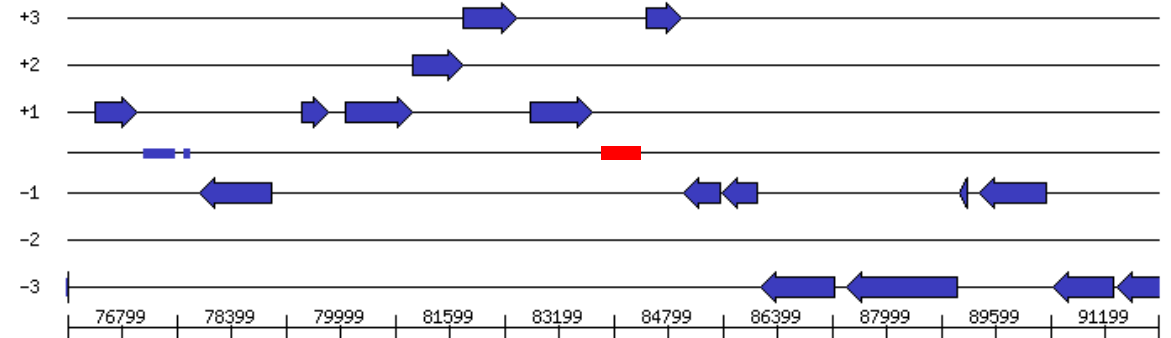
Coding and non-coding transcripts

- some highly expressed transcripts are non-coding

UNIVERSITY OF AMSTERDAM
Life Sciences

Example - OLE (ornate, large, extremophilic) RNA

RFAM

o found in Firmicutes
o ~ 600 nt length

Firmicutes mOTU:

Clostridia mOTU:
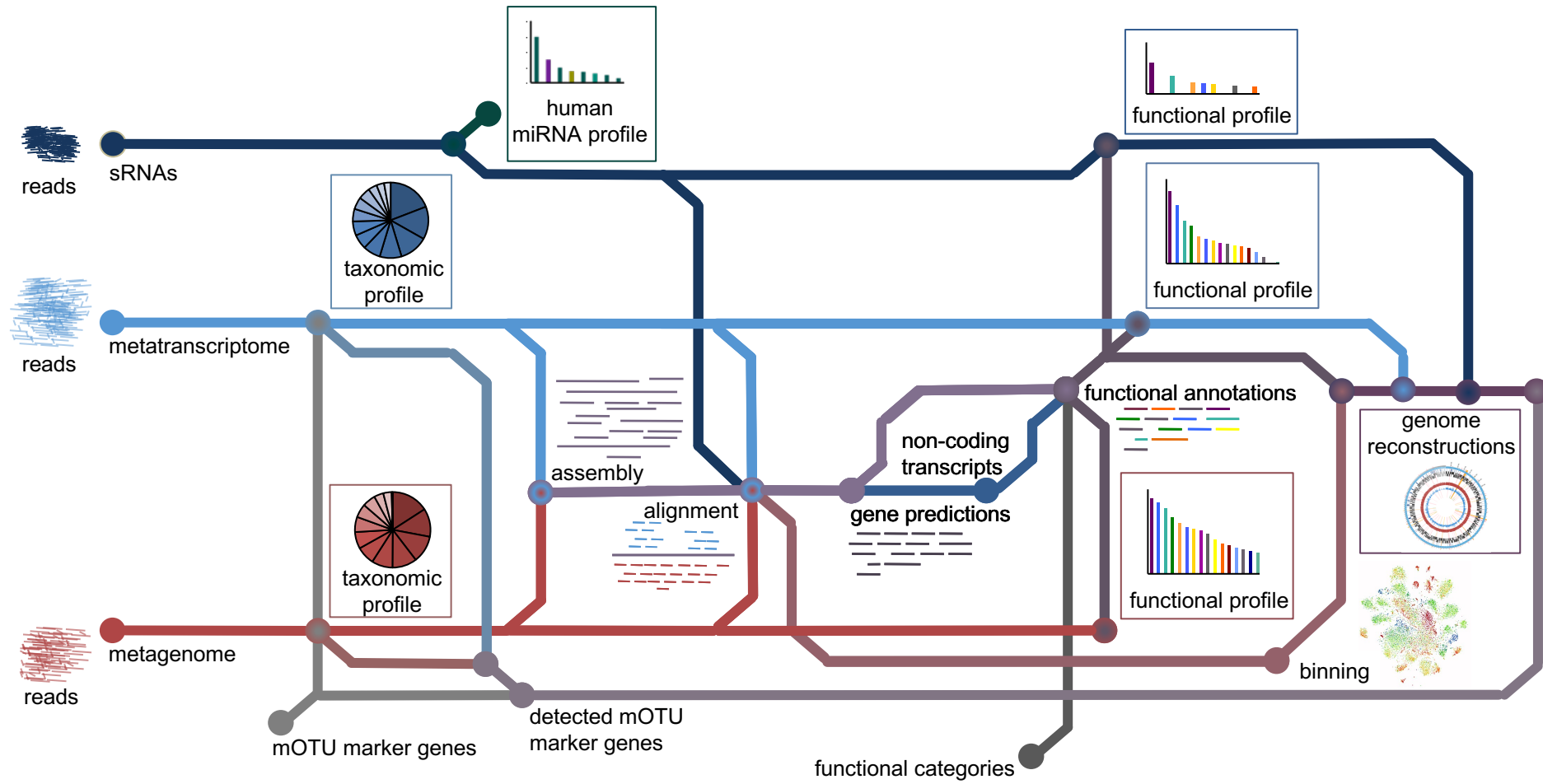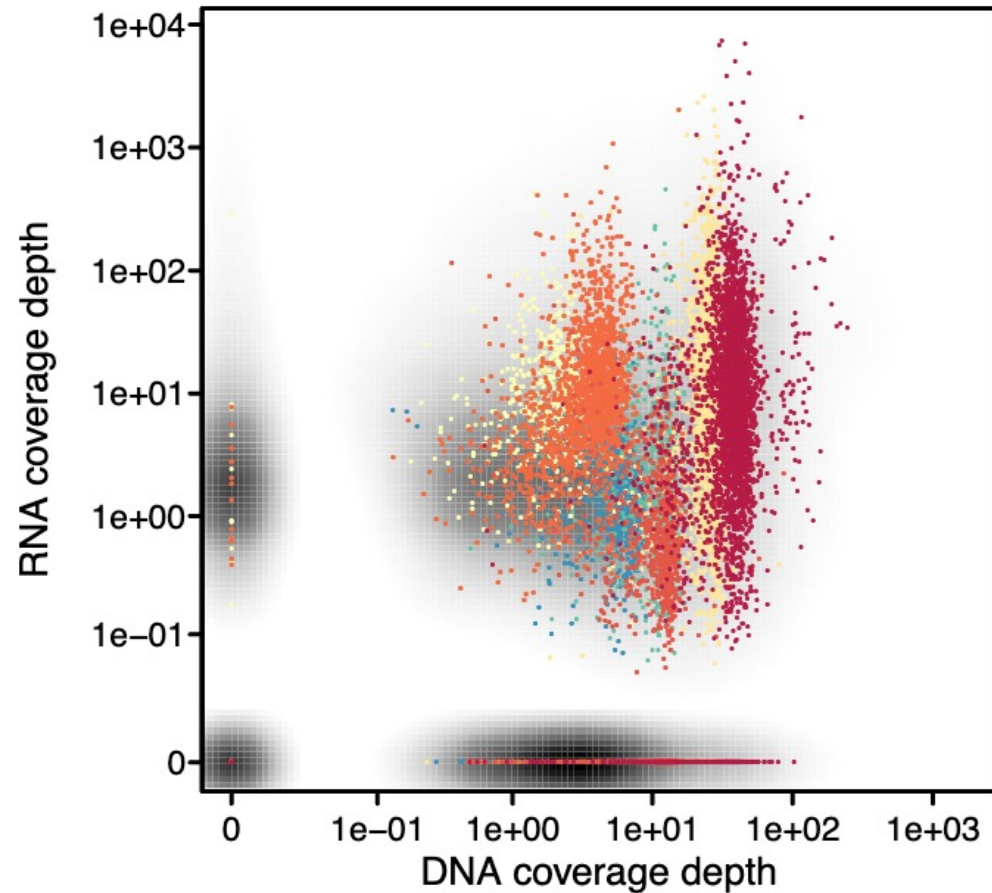
UNIVERSITY OF AMSTI
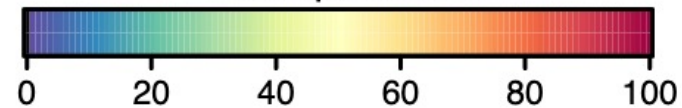Life Sciences

Workflow

Workflow

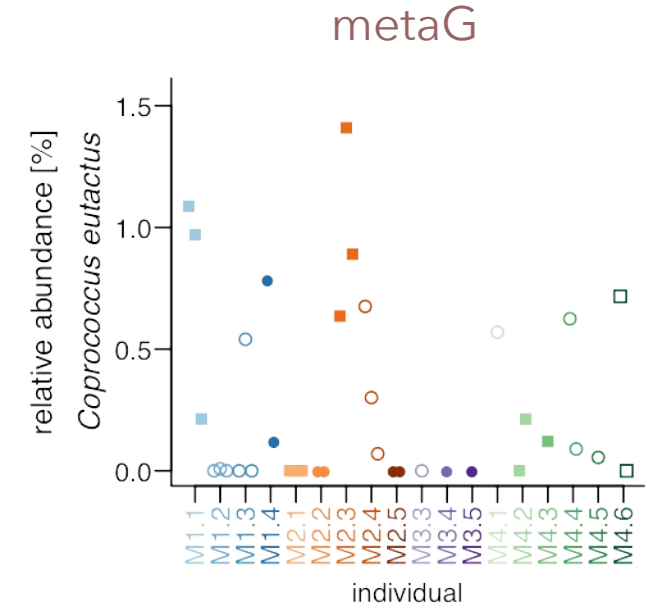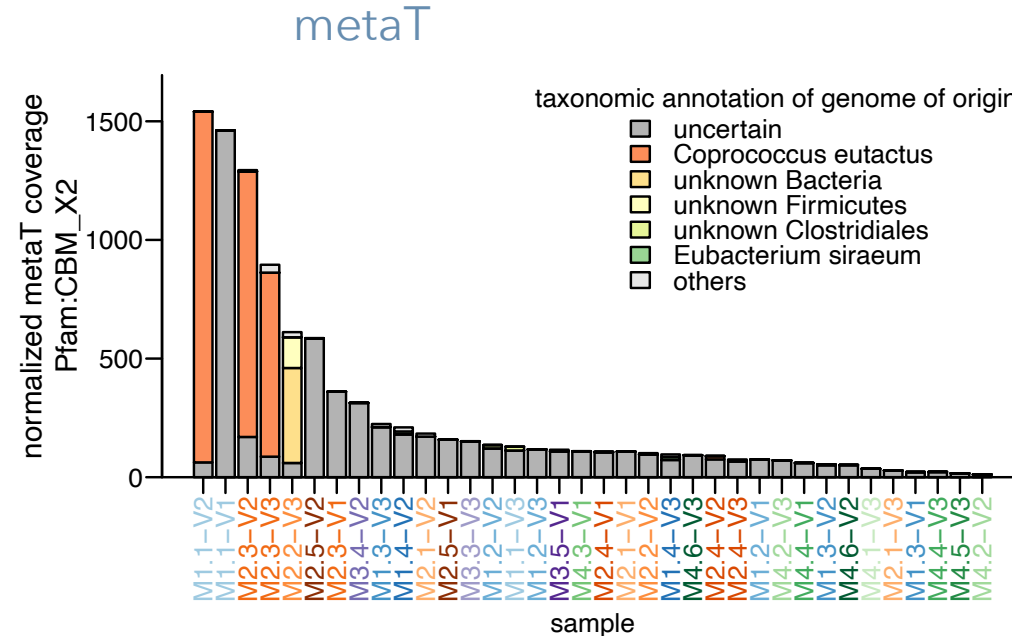# Expression per genome

binned population–level genomes – taxonomy
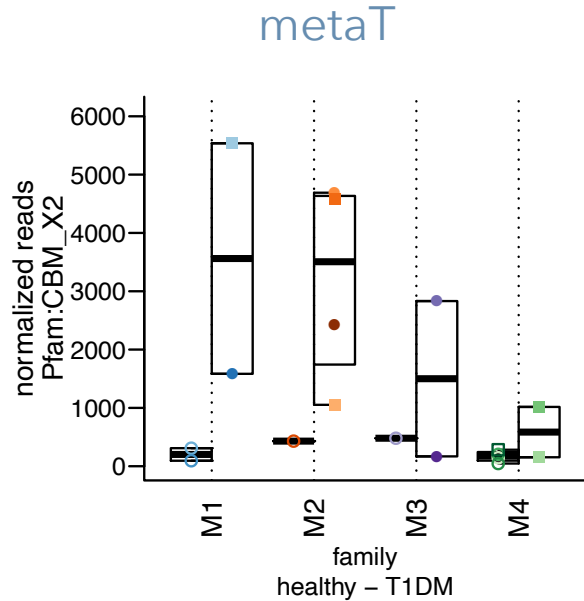
- G2.2.2 – Butyrivibrio crossotus
- G11 – Firmicutes
- G31 – unclassified Clostridiales
- O32 – unclassified Clostridiales
- O50 – Ruminococcus sp.
- C1.1.2 – Collinsella aerofaciens
- C24.1 – unclassified Clostridiales

% completeness

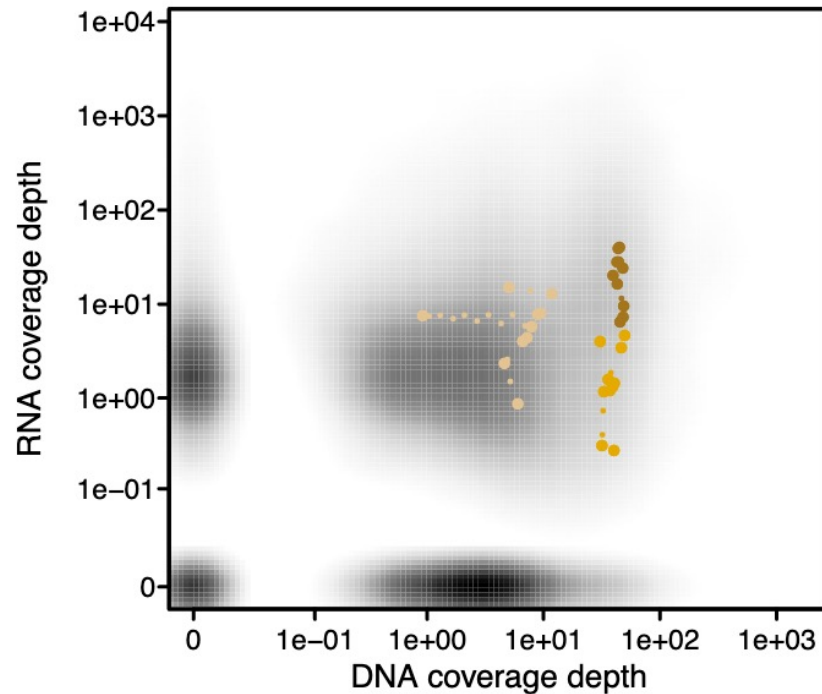# Example –
who expresses a gene of interest?
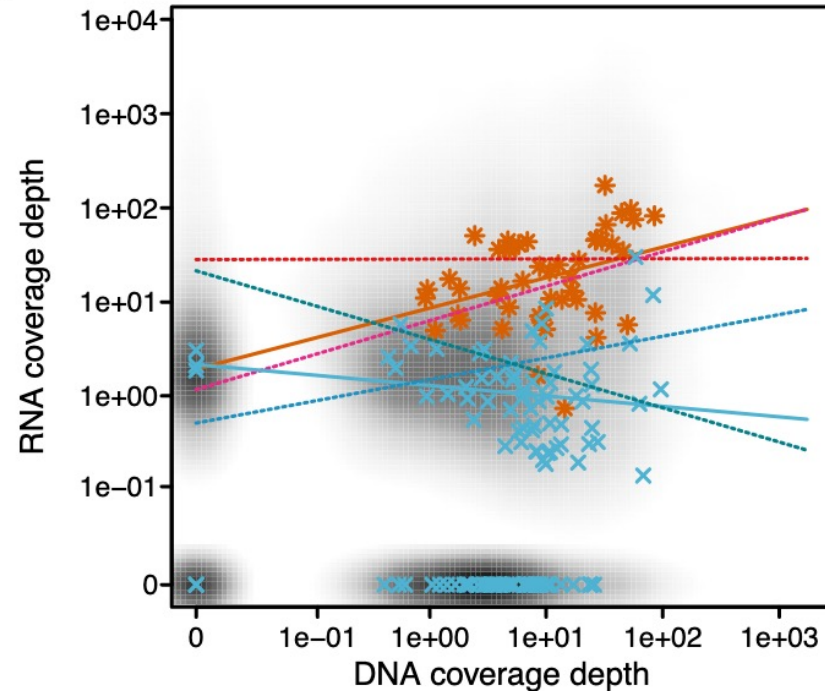
example: protein with cellulose-binding domain

Heintz-Buschart *et al.*, Nature Microbiology, 2017

Expression per and across genomes

mOTU genes:

non-house-keeping genes:

Example – non-coding elements in a *Parabacteroides distasonis* genome

- 4.7 Mbp
- ~4,600 ORFs
  - ~3,600 expressed
  - 16 proteins detected
- 2,949 UTRs
- 84 annotated non-coding regions (+ 4 5S rRNAs and 55 tRNAs)
- 46 CRISPR repeats
- 1 bacterial nc RNA (*Bacteroidales*-1)
- 12 riboswitches
- other potential regulatory elements

CRISPR region:
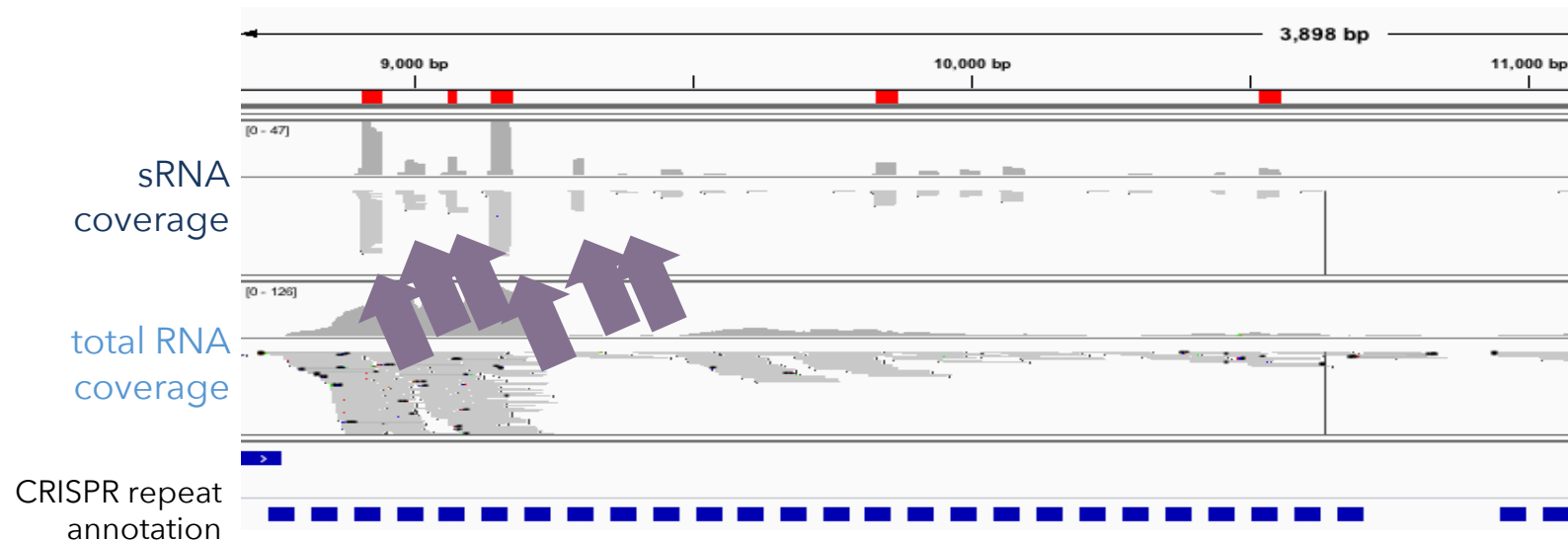
UNIVERSITY OF AMSTERDAM
Life Sciences

Heintz-Buschart *et al.*, unpublished

Example – non-coding elements in a *Parabacteroides distasonis* genome

- 4.7 Mbp
- ~4,600 ORFs
  - ➤ ~3,600 expressed
  - ➤ 16 proteins detected
- 2,949 UTRs
- 84 annotated non-coding regions (+ 4 5S rRNAs and 55 tRNAs)
- 46 CRISPR repeats
- 1 bacterial nc RNA (*Bacteroidales*-1)
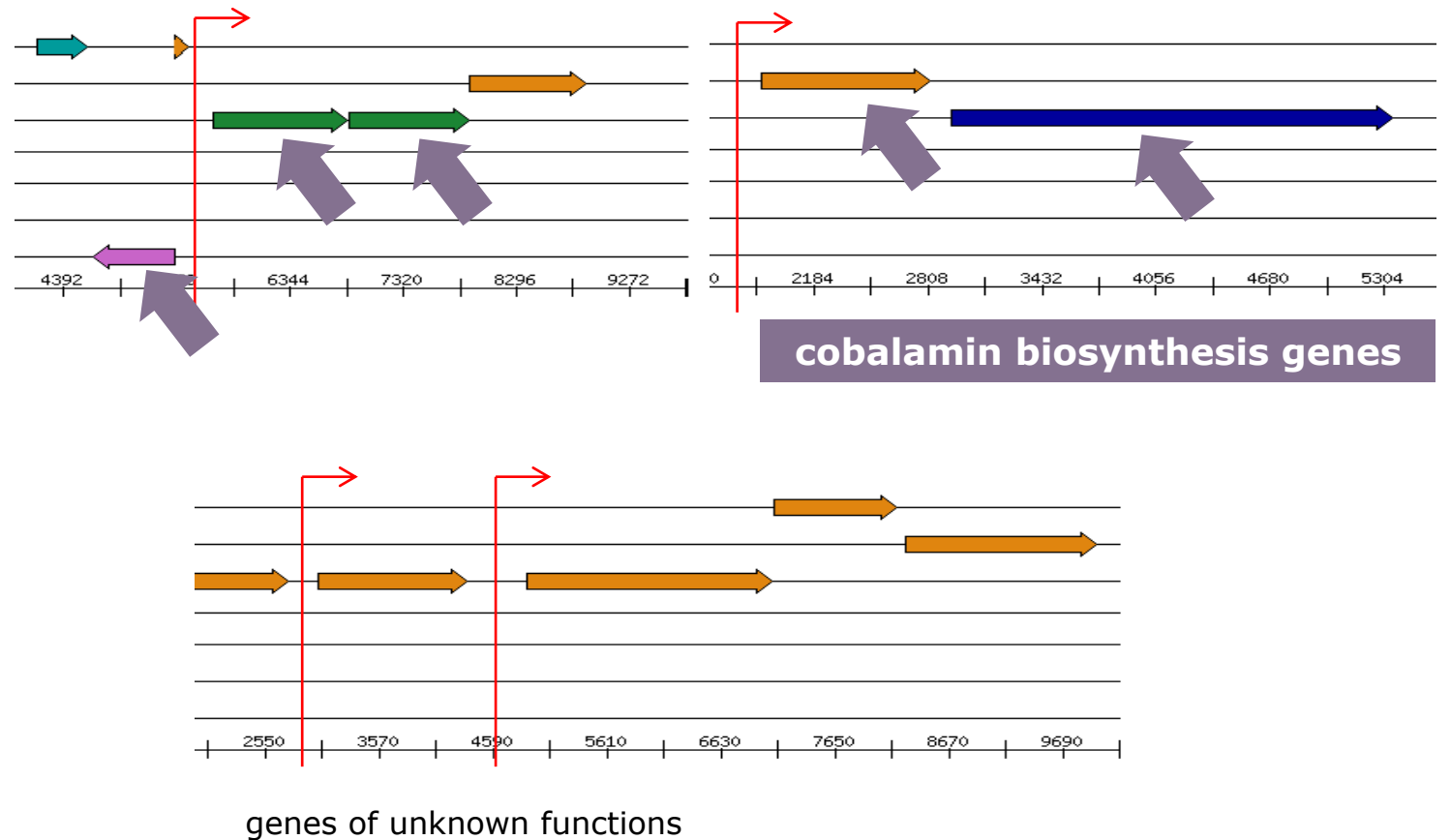- 12 riboswitches
- other potential regulatory elements

cobalamin riboswitches and gene function:

cobalamin biosynthesis genes

genes of unknown functions

Thanks for your attention!

a.u.s.heintzbuschart@uva.nl

SP C2.205

github.com/a-h-b

twitter.com/_a_h_b_