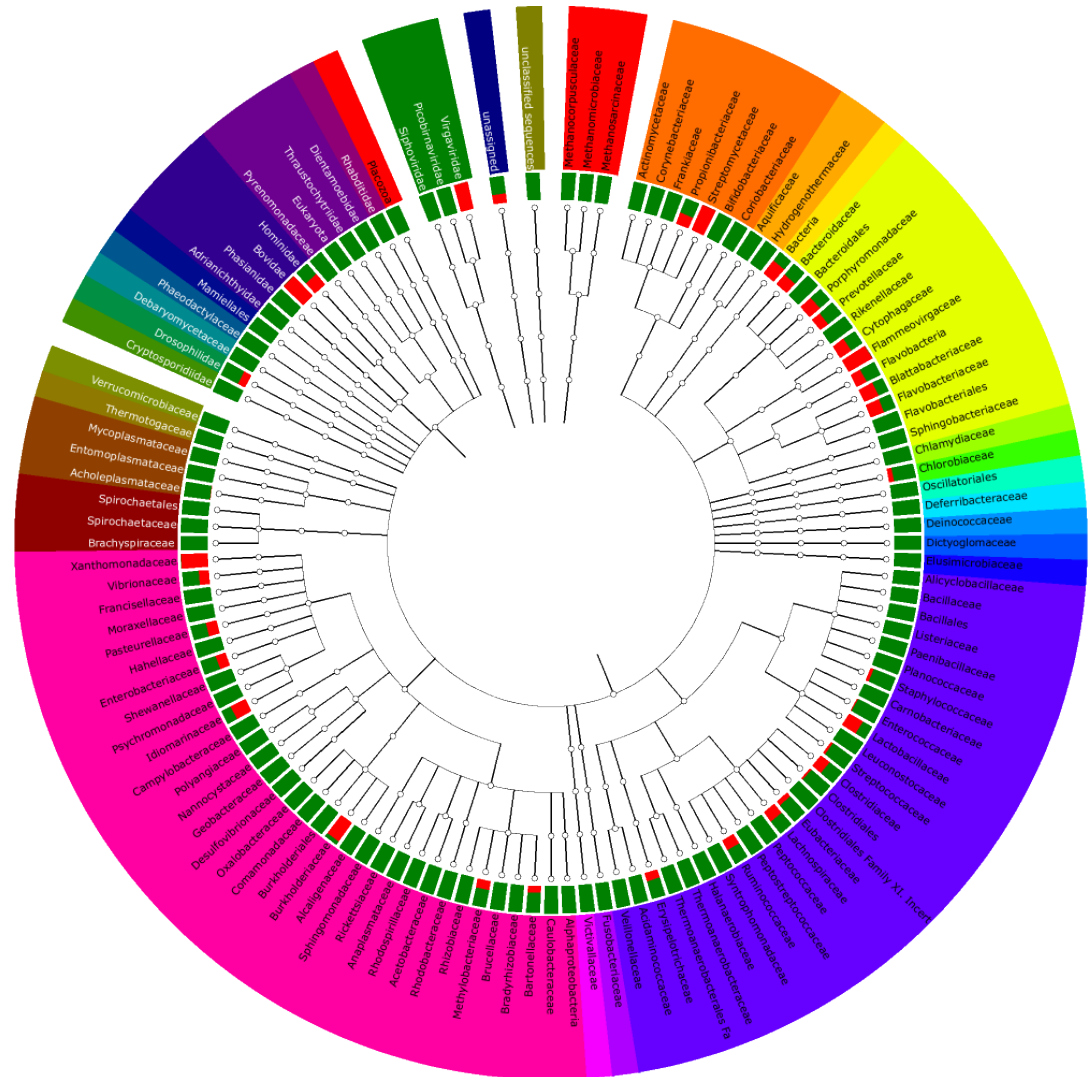


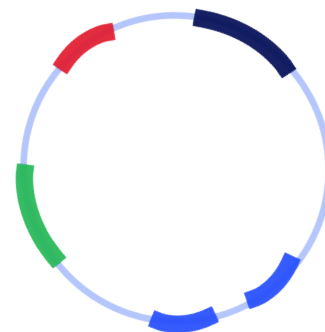
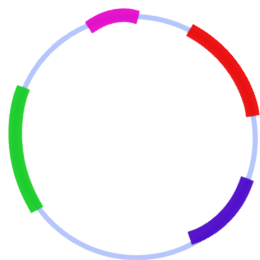
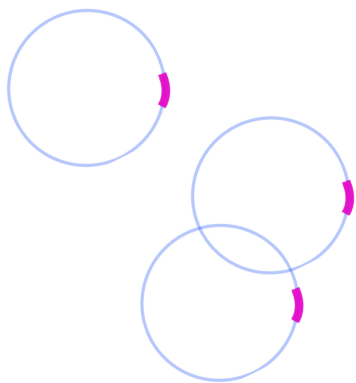
Metagenomic Analysis

Outline

- Gene targeted vs. random approaches
- Diagnostic versus ecological uses
- Case study
- mg-RAST
- MEGAN demonstration

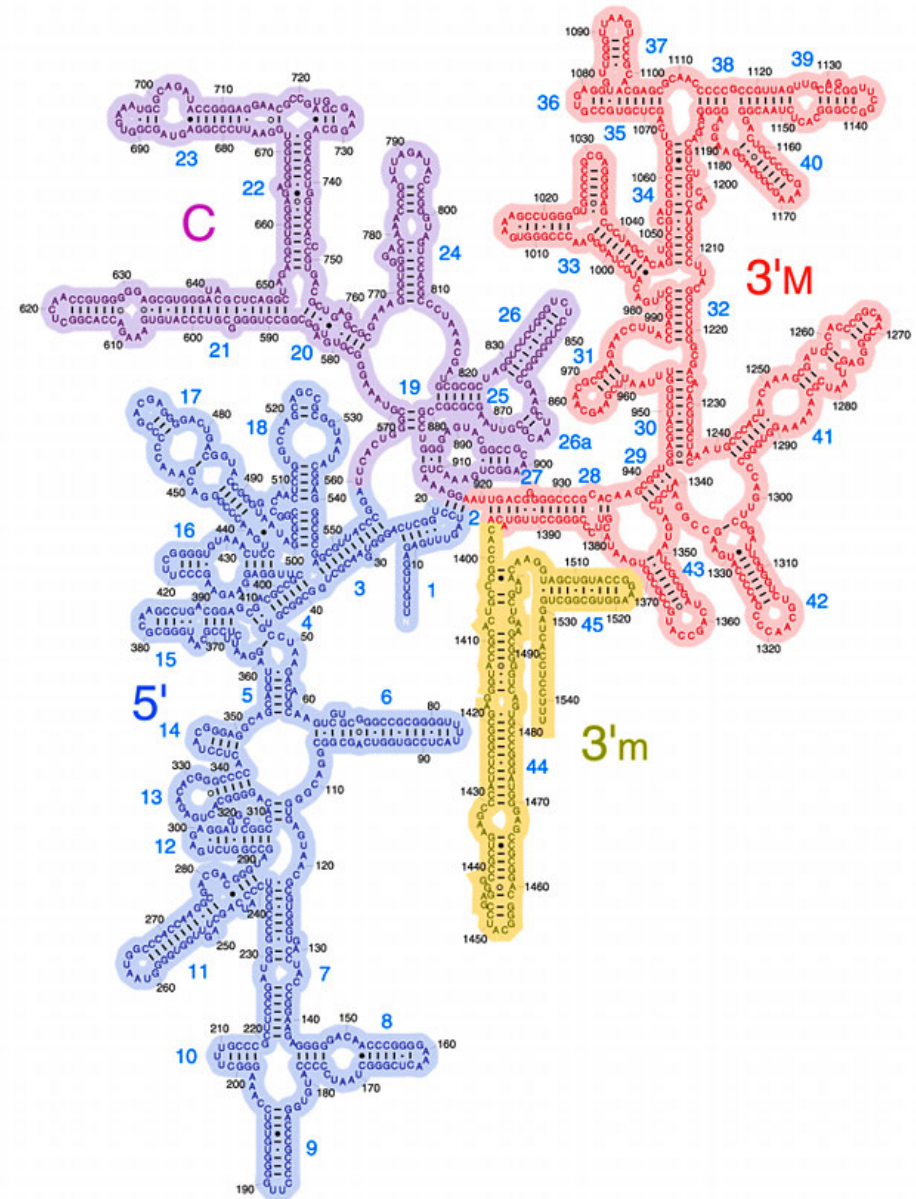


Gene Targeted and Random Shotgun Approaches

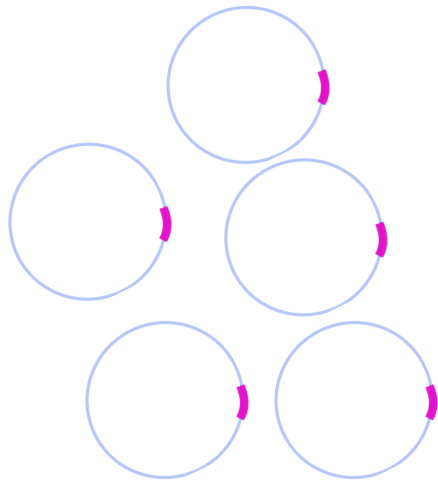


16s rRNA

- a section of prokaryotic DNA found in all bacteria and archaea
- encodes the small subunit of the ribosome
- 1.5 kb
- conserved and hypervariable regions
- used for decades in clinical microbiology and studies on bacterial evolution
- used extensively in metagenomic studies



targeted - 16s rRNA

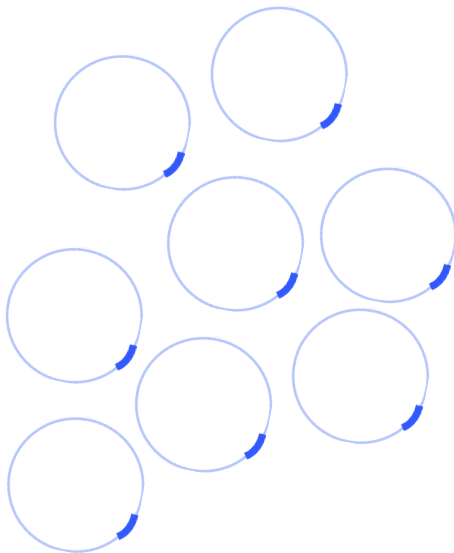


CTACAATGTGTTCTGTACGTACGGCACTTTTCNGCATTCTGACAGTACG
CTACAATCTGTACGTACGTTCGGCACTATCNGCATACTGACAGTACG
GTACAATGTGTACGTACCTACGGCACTATCNGTATACTTCCAGCACG
GTACAATGTGTACGTACGTACGGCACTTTTCNCCATACTGACAGTACC
CTACATTGTGTACGTACGTACGGCACTATCNGCATACTGACACTACG

Operational Taxonomic Unit

Database search

Escherichia coli



AACTGTCCTCTGGTATGCACGGCACTTTTTGCATACTGACAGTACGA
ATCTCTCCTCTGGTATGCACGGCACTTTTTGCATACTGACAGTACGA
AACTGTCCTCTGGTATGCACGGCACTTTAAGCATACTGACAGTACGA
AACTGTCCTCTGCTATGCACGGCACTTTTTGCATACTCACAGTACGA
ATCTGTCTGTGGTATGCACGGCACTTTTTGCATACTGACAGTACGA
AACTGTCGTCTGGTATGCACGGCACTTTTTGCATACTCACAGTACGA
AACTGTCTGTGGTAAGCACGGCACTTTTTGCATACTGACAGTACGA
AACTGACCTCTGGTATGCACGGCACTTTTTGCATACTGACAGTACGA
AACTGTCCTCTGGTATGCACGGCACTTTATGCATACTGACAGTACGA

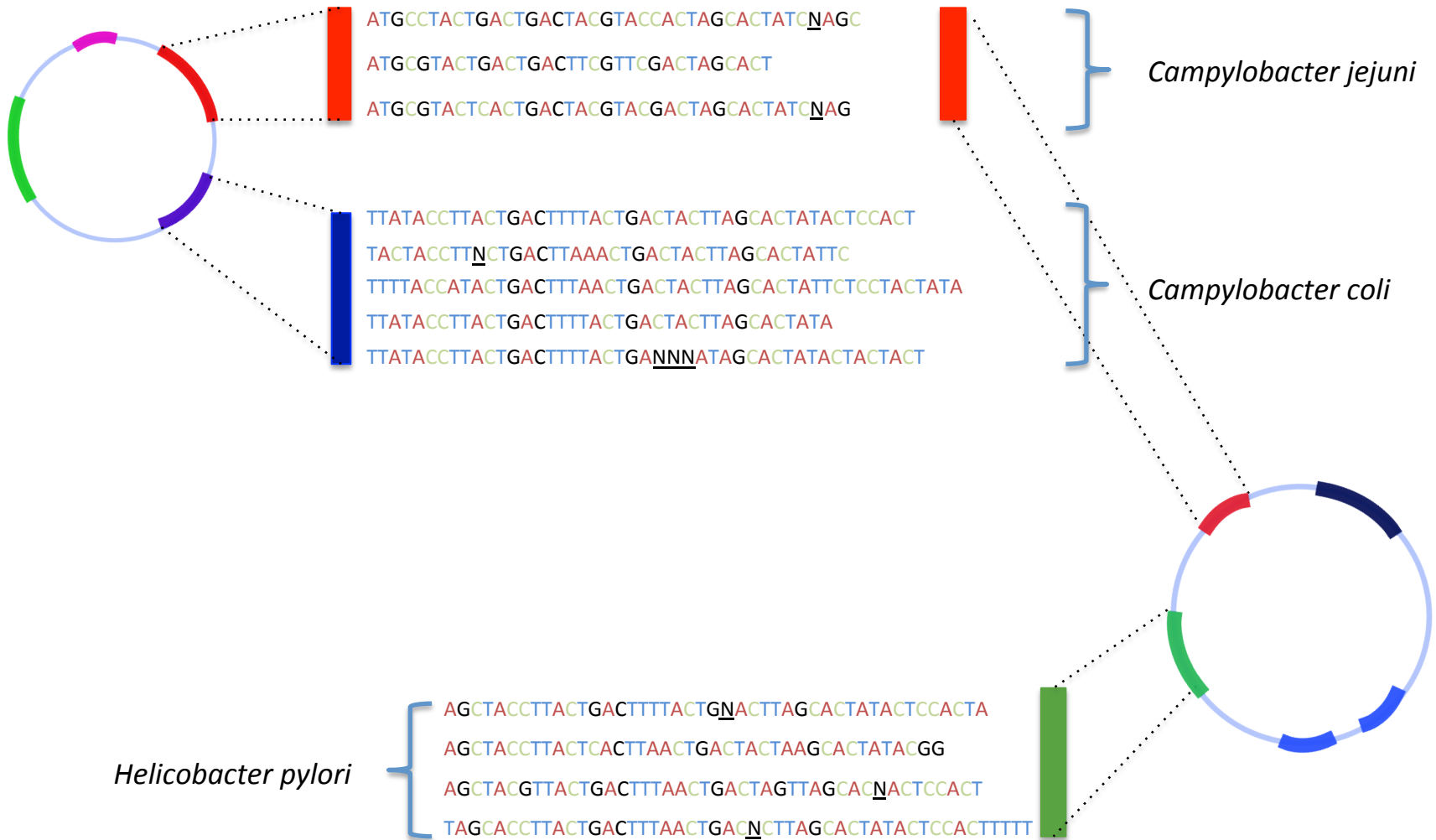
Operational Taxonomic Unit

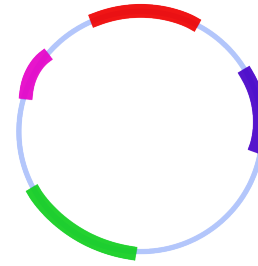
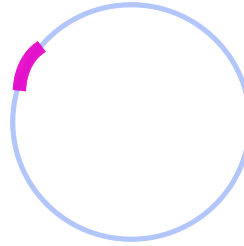
Database search

No hits

Not *Escherichia coli*

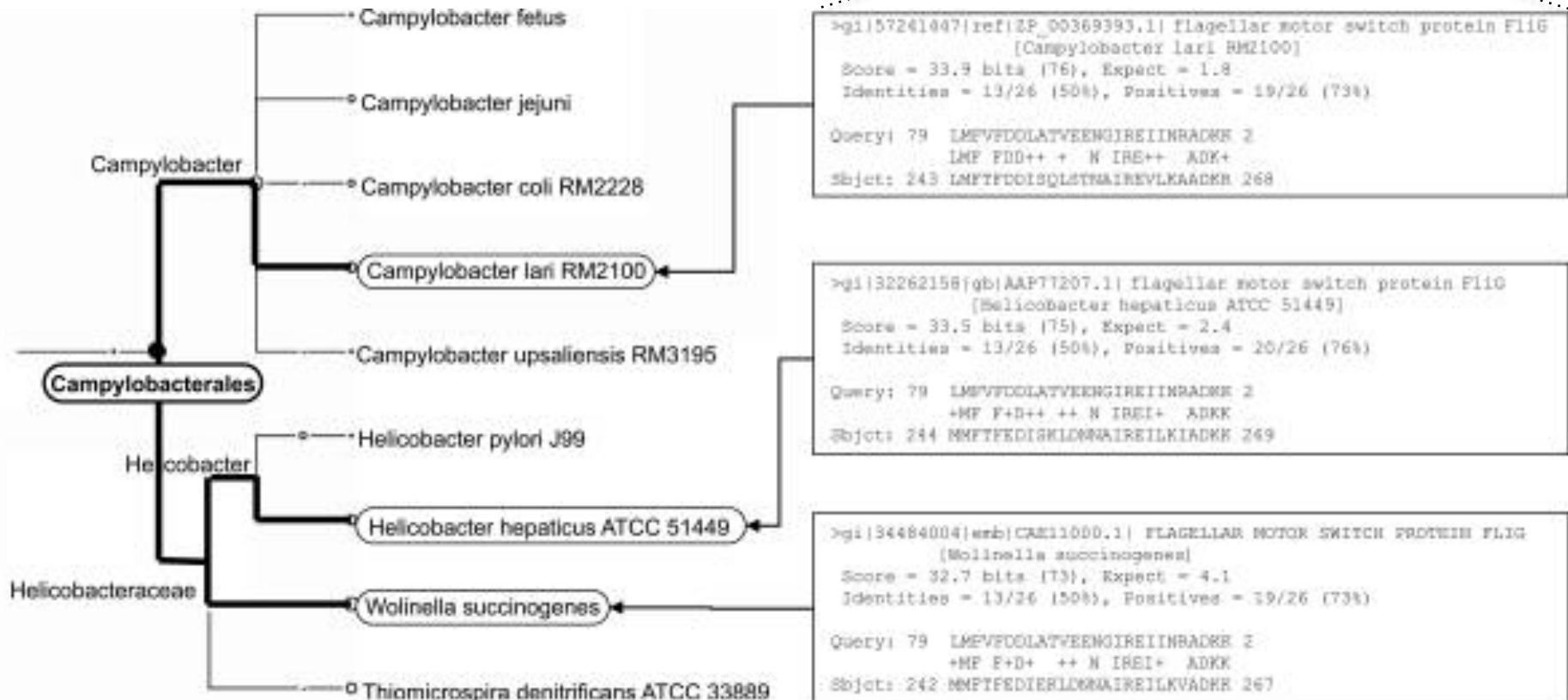
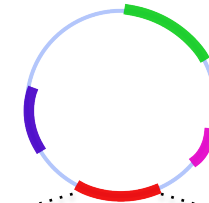
Random sequencing



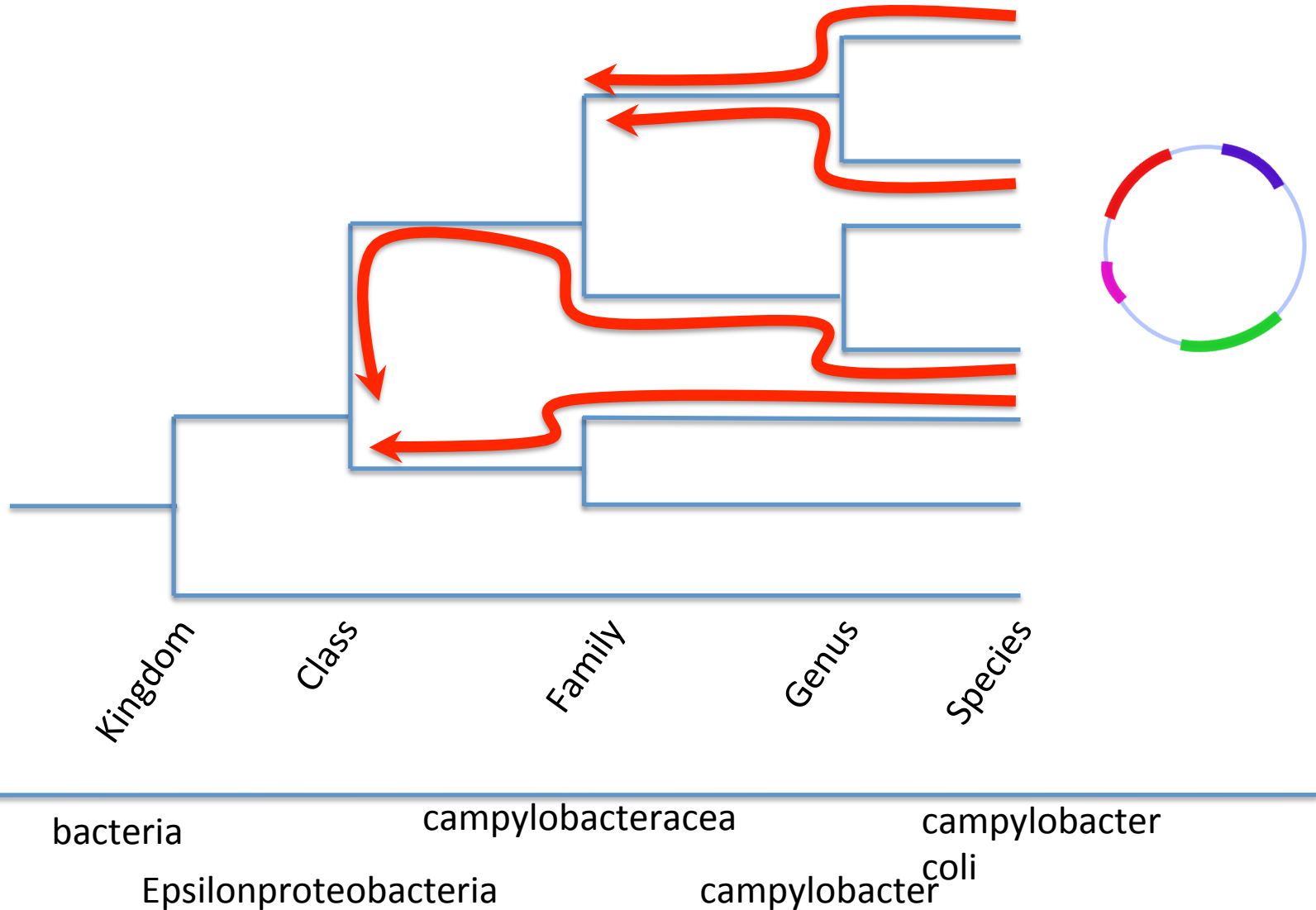


	16s rRNA	Random
Read quality	High	Variable
Reference Database	Good	Variable / Potentially more biased
Taxonomic Assignment	Partially database independent	Database dependent
Non-bacterial assignment	Impossible	Database dependent
Phenotype assessment	Assumed	Unknown

Lowest-common ancestor (MEGAN)

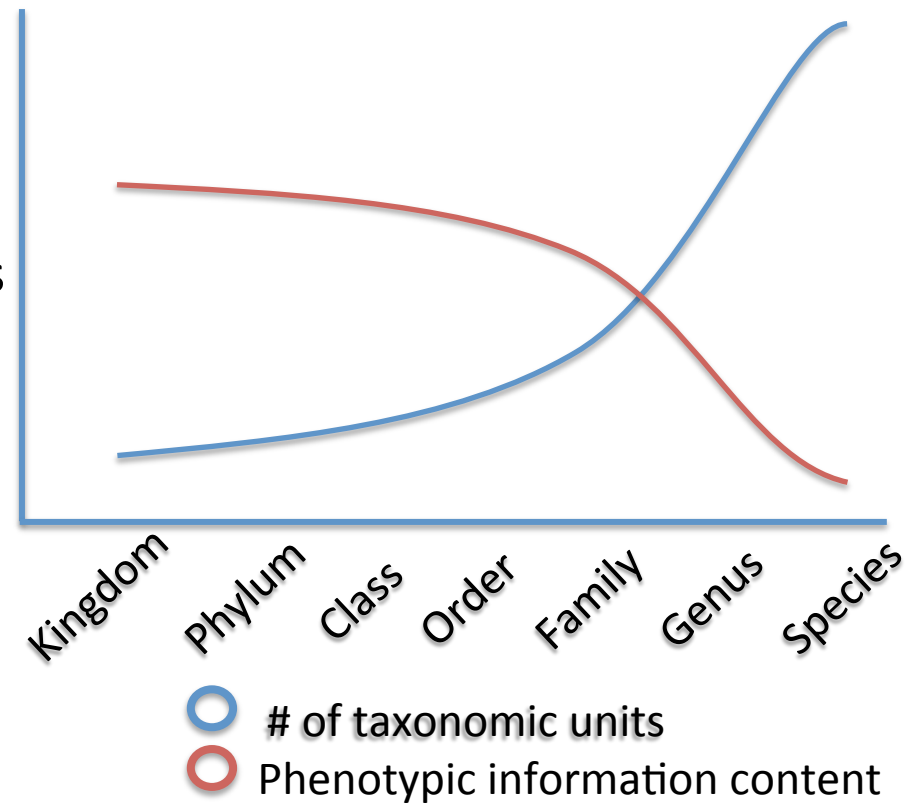
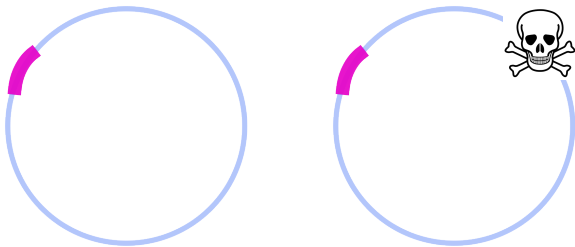


Assignment tends to move up the tree

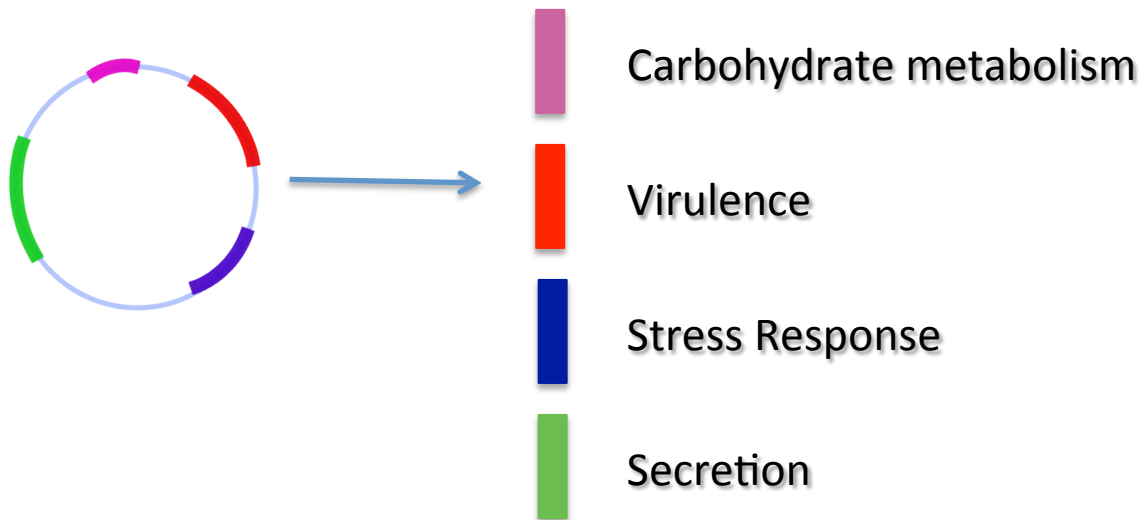
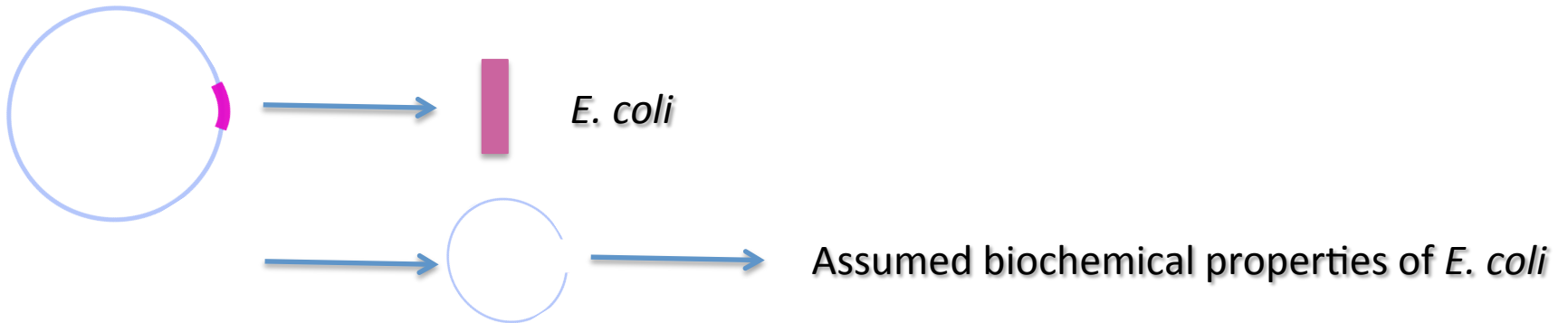


What level of taxonomic resolution do we need?

- Depends on what you want
- Diagnostics vs. ecology
- The functional ecology promise of metagenomics might not benefit from species level analysis

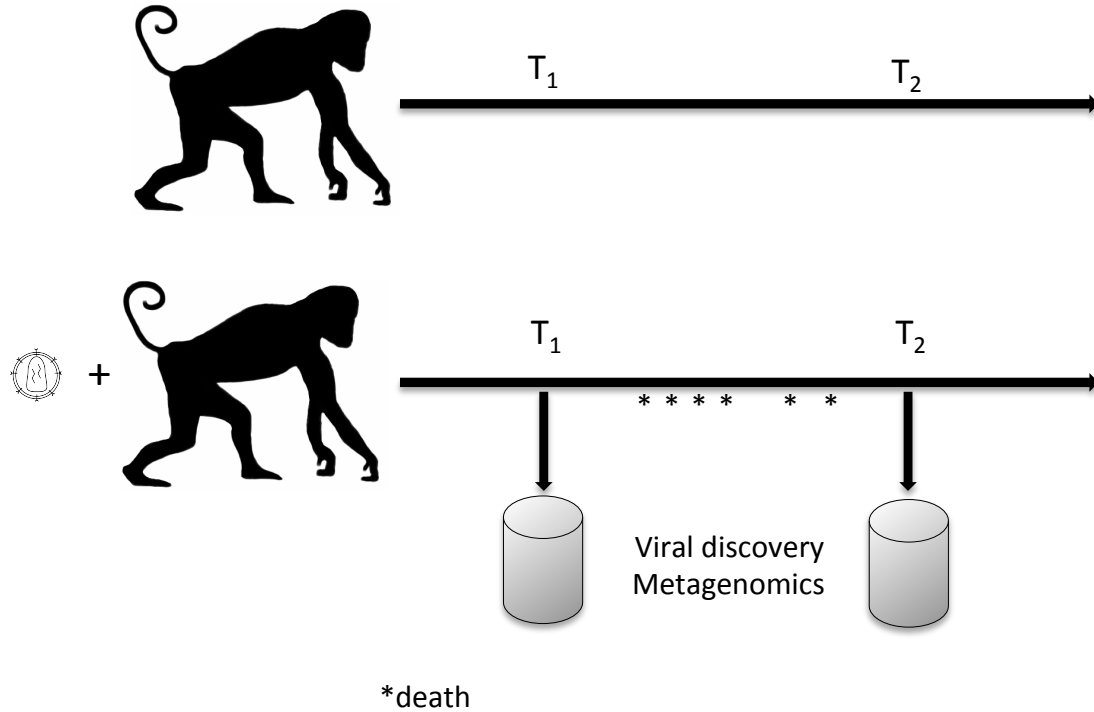


16s vs. random

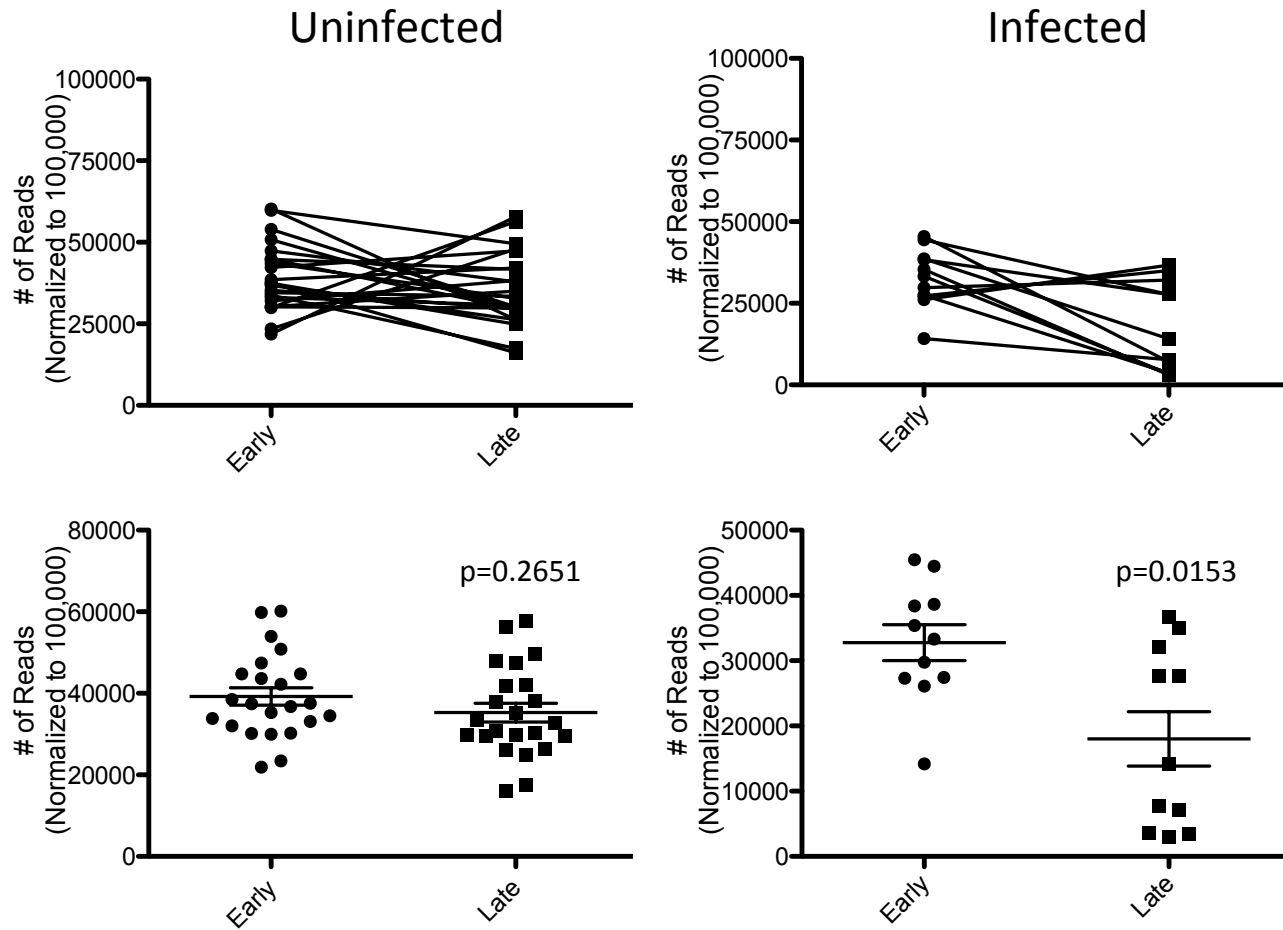


SIV Macaque Gut Metagenomics

SIV Study Cohort

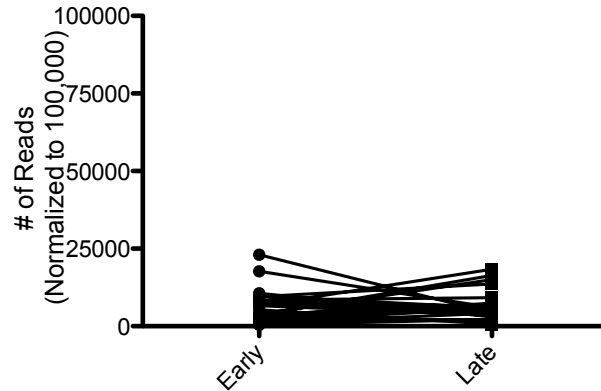


Bacteria

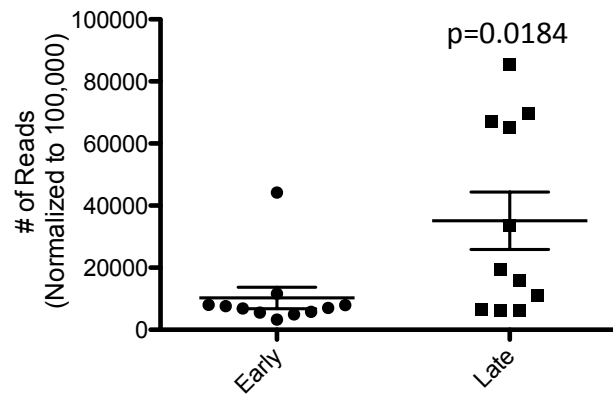
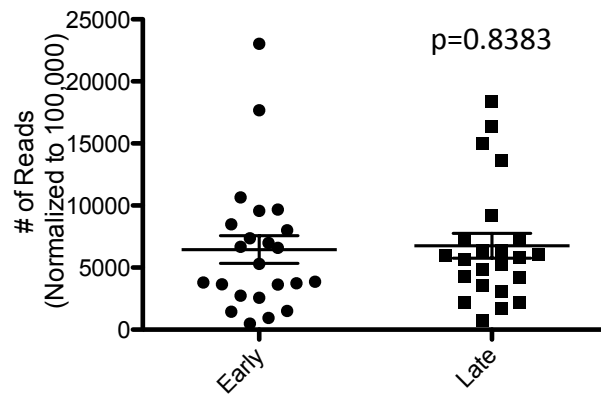
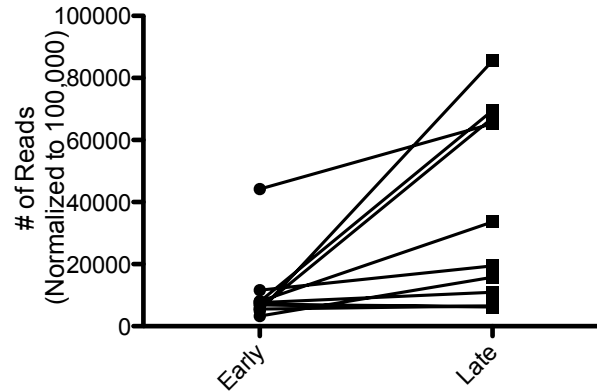


Virus

Uninfected

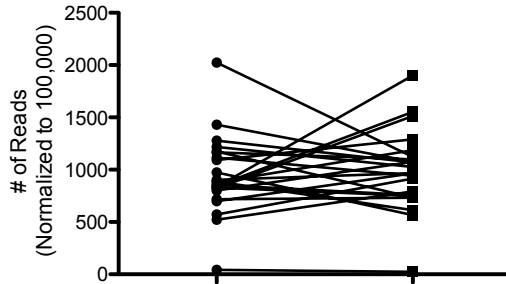


Infected

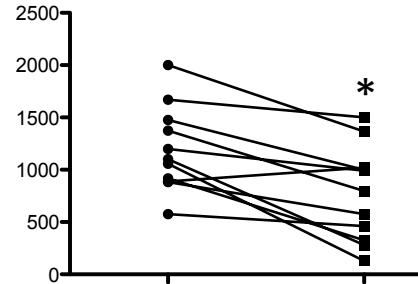


Other Kingdoms

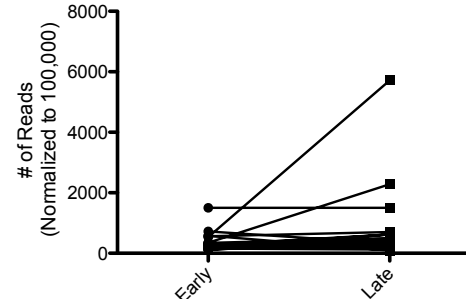
SIV (-) Fungi



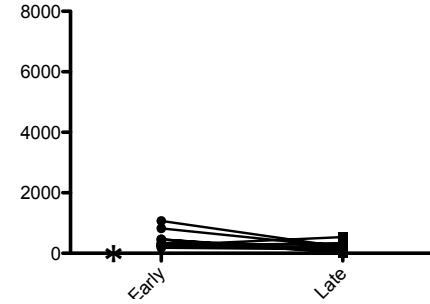
SIV (+) Fungi



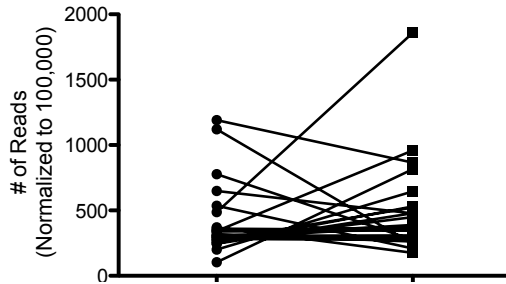
SIV (-) Amoebozoa



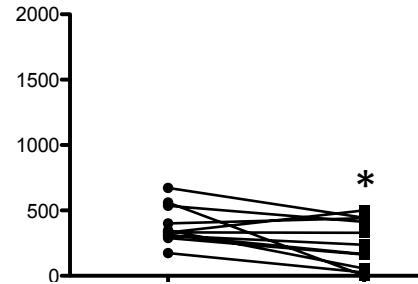
SIV (+) Amoebozoa



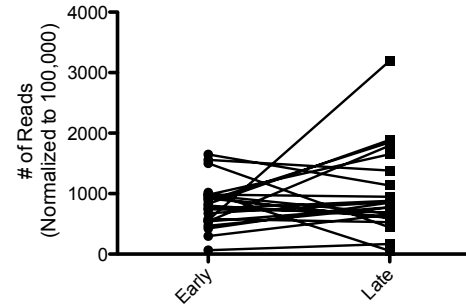
SIV (-) Archaea



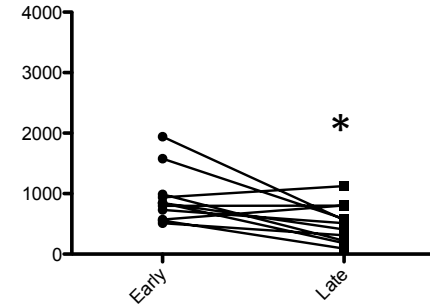
SIV (+) Archaea



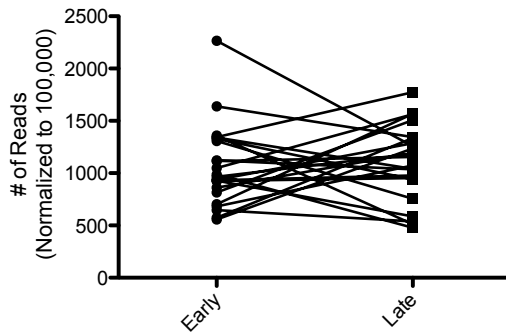
SIV (-) Viridiplantae



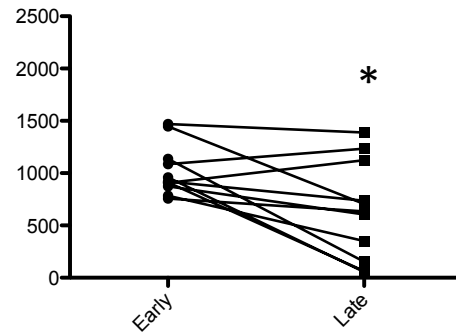
SIV (+) Viridiplantae



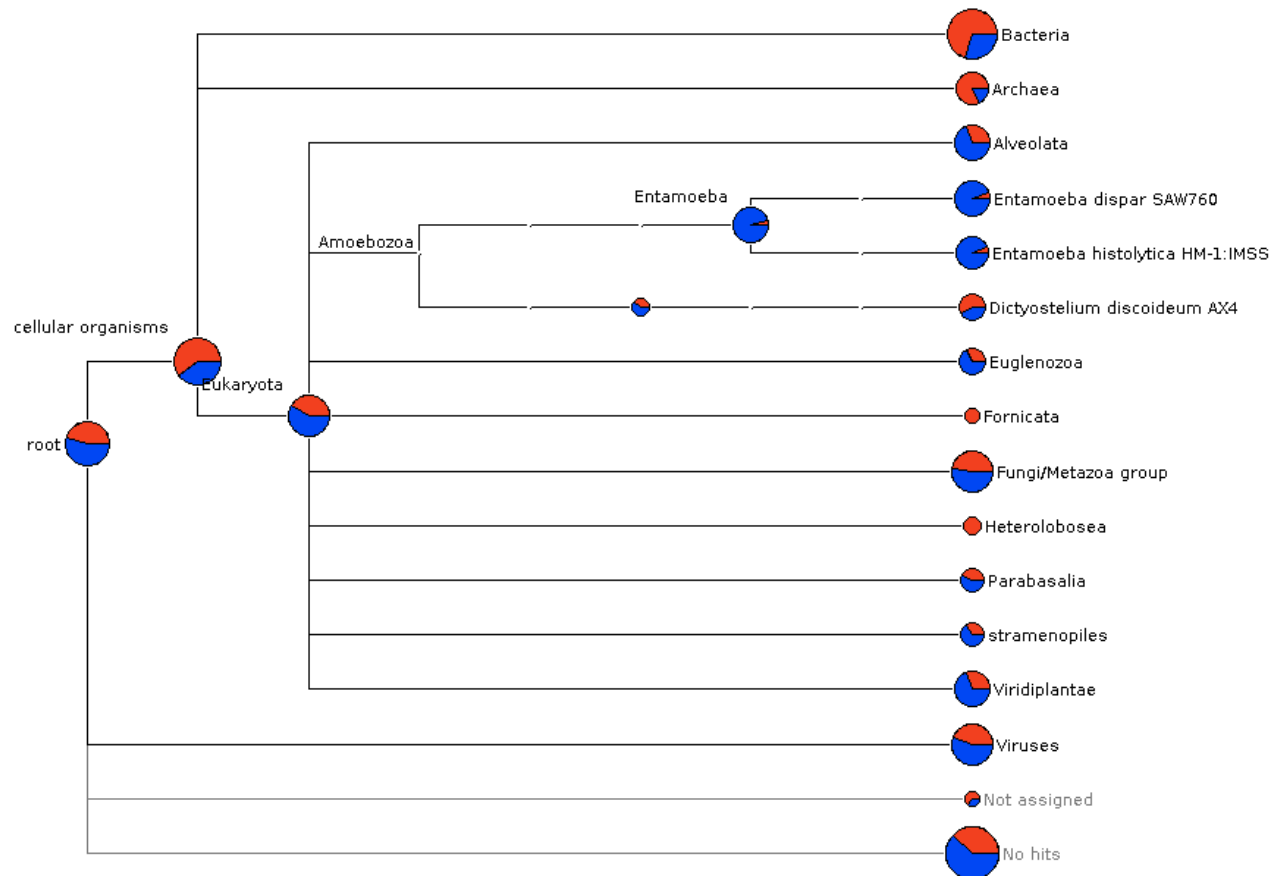
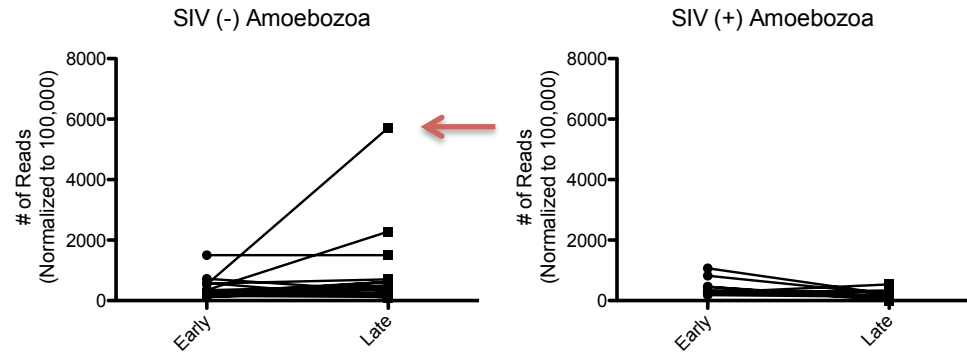
SIV (-) Alveolata



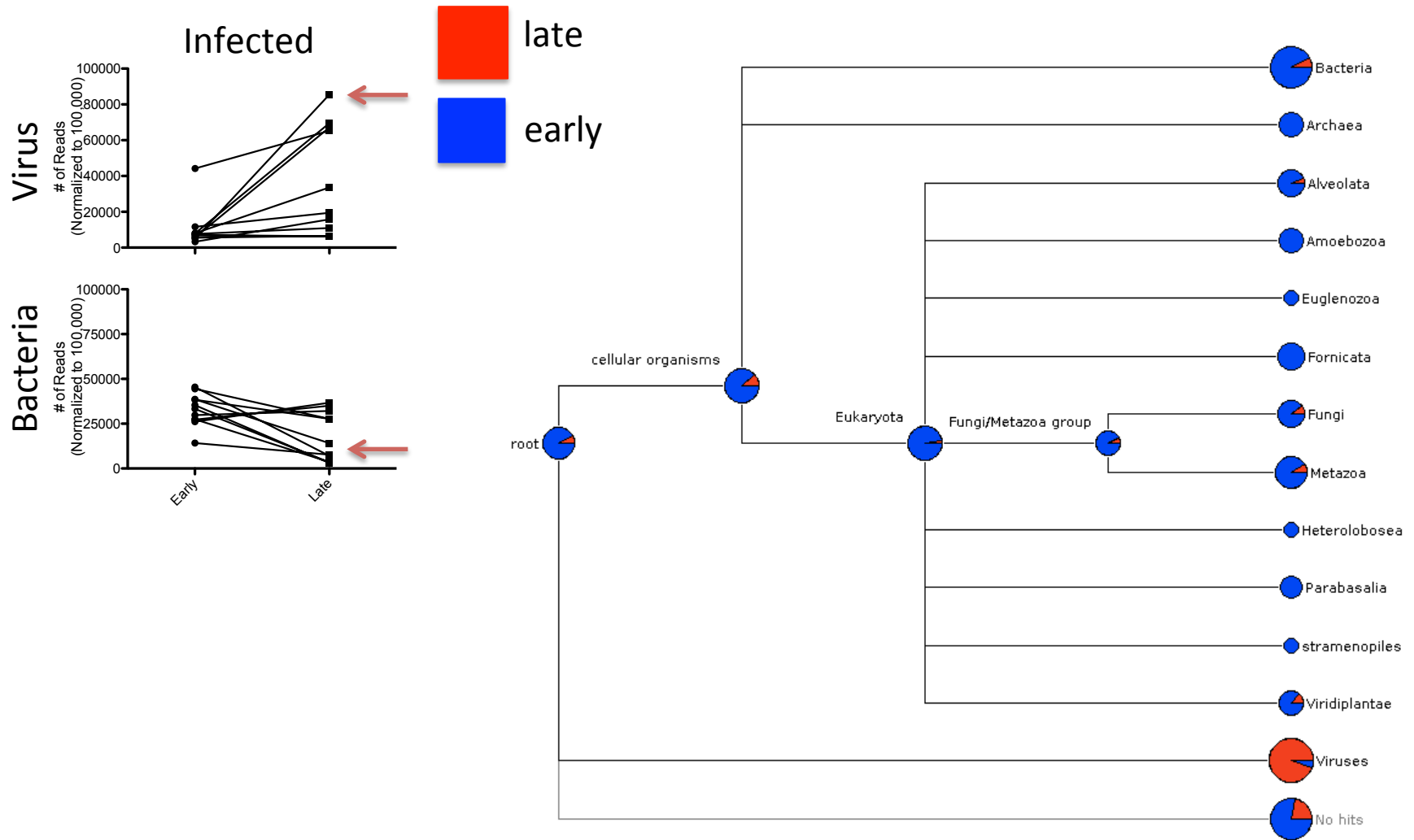
SIV (+) Alveolata



Diagnostic potential

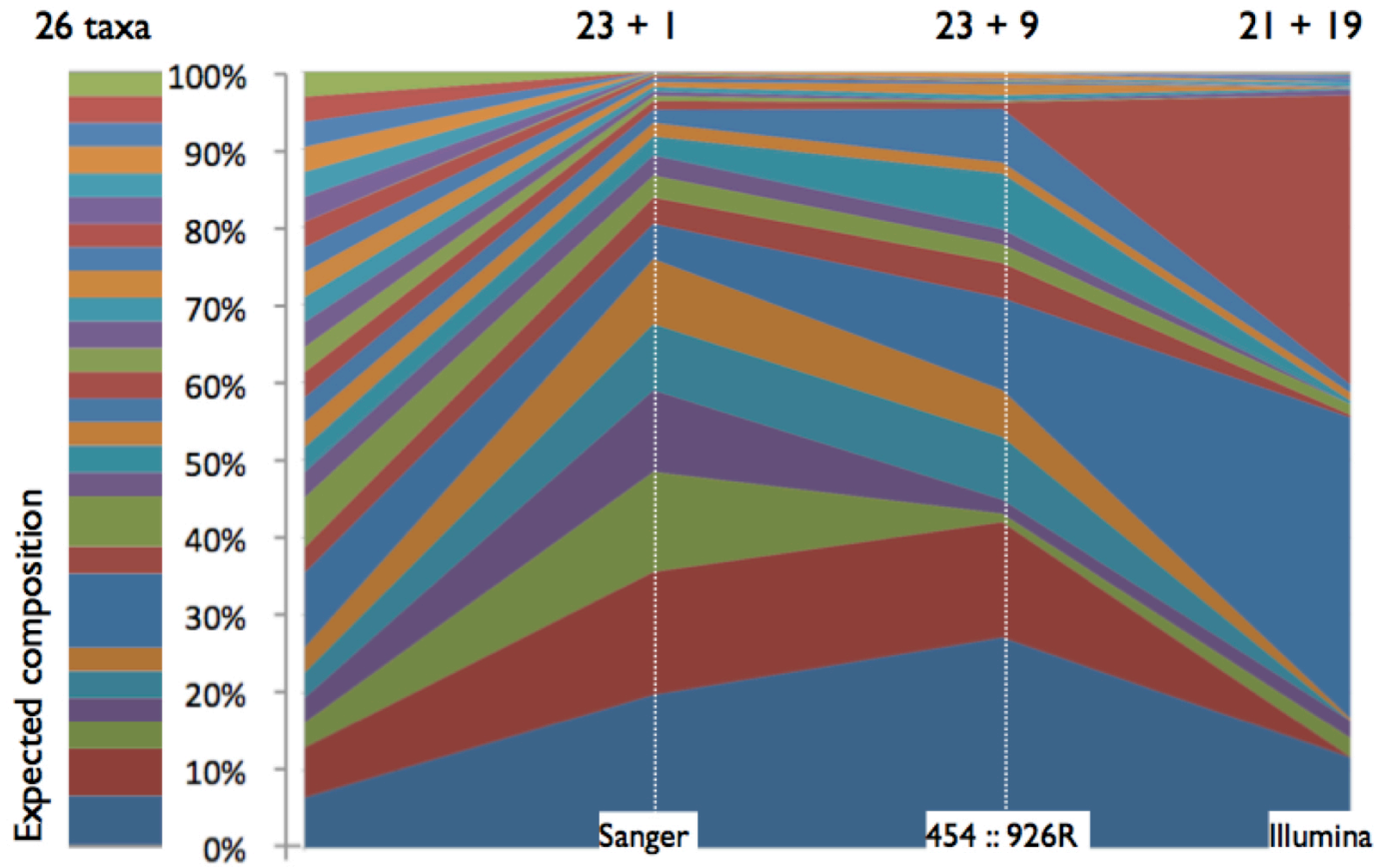


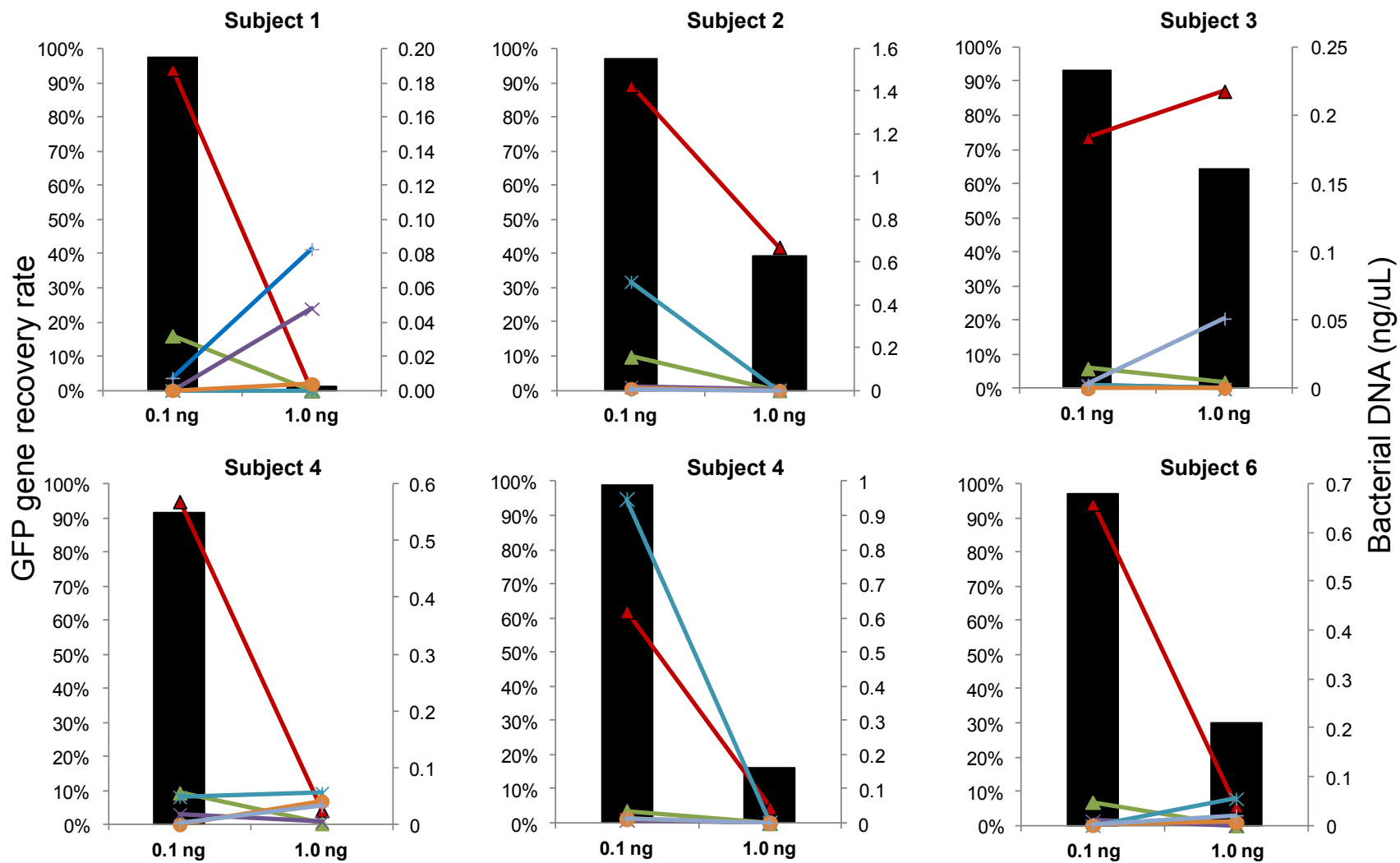
Extreme example





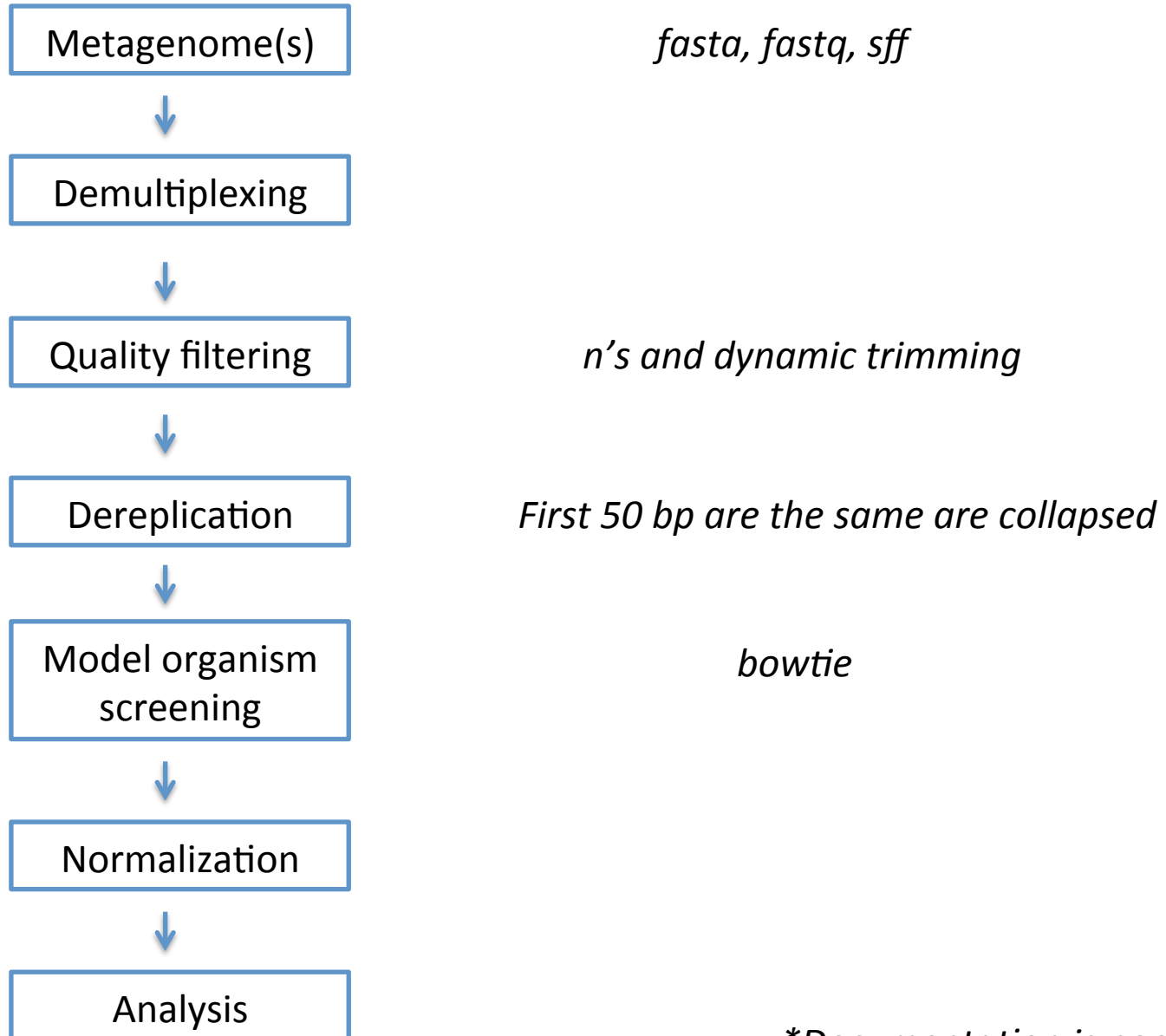
Challenges





mg-RAST

Automated Pipeline*



**Documentation is poor at best*

<http://metagenomics.anl.gov/>

[http://evomics.org/learning/
genomics/metagenomics/mg-rast/](http://evomics.org/learning/genomics/metagenomics/mg-rast/)

