



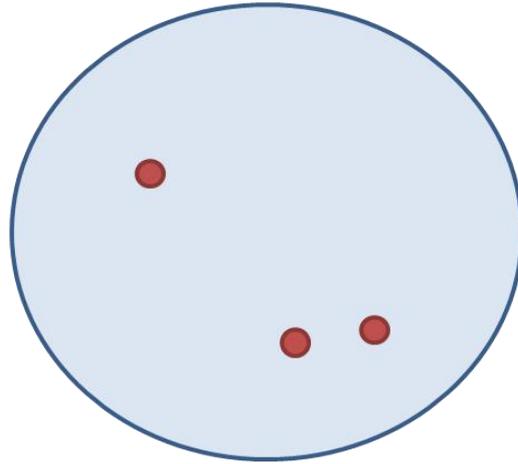
Metagenomics

Does exist the easy way how to study the microbiome?

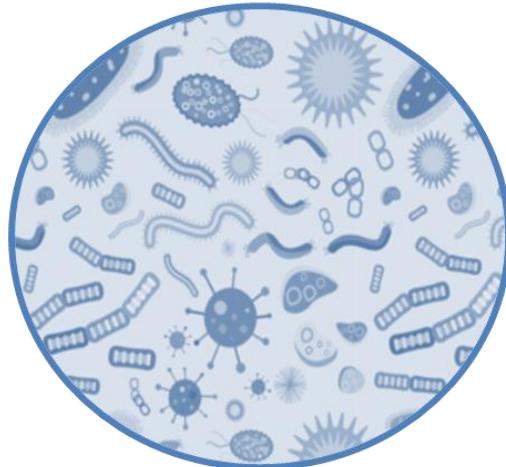
Petra Videnska

RECETOX, Masaryk University

Microbiome



Culture



Microscopy



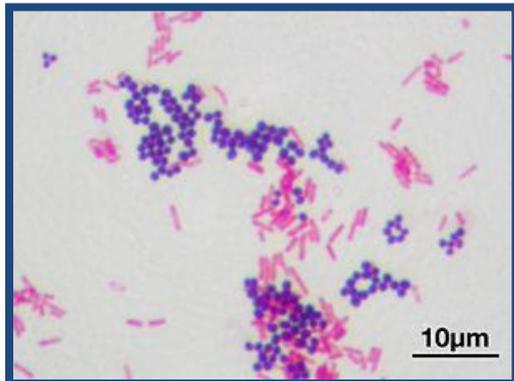
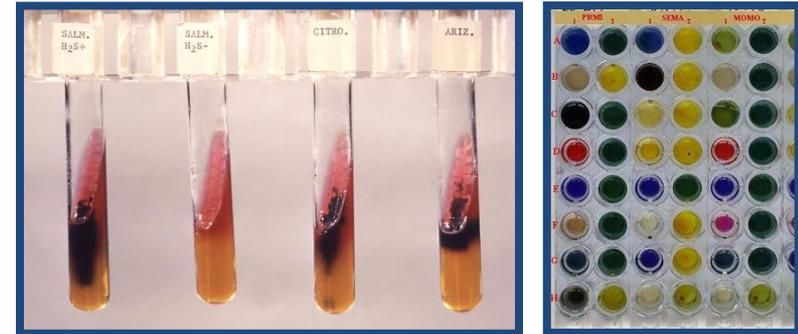
Culturable bacteria

Bacterial culture

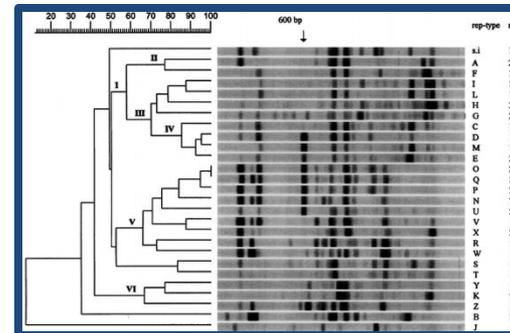


Isolation of pure cultures

Biochemical test



Microscopy



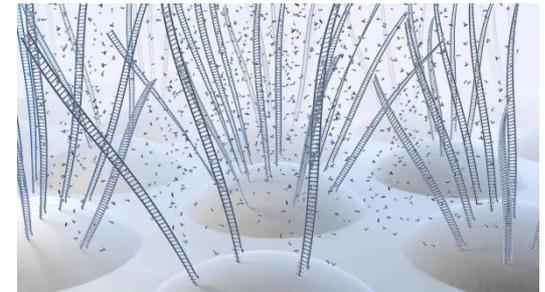
Molecular fingerprint methods

Unculturable bacteria

Study of DNA – uncultured bacteria (>99,9%)



Millions to billions
sequences per sample



Metagenomics



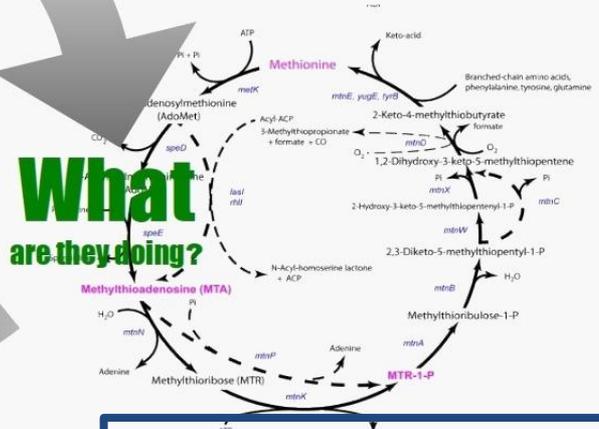
Targeted sequencing - 16S rDNA
Whole metagenomic sequencing

Metametabolomics
Metaproteomics



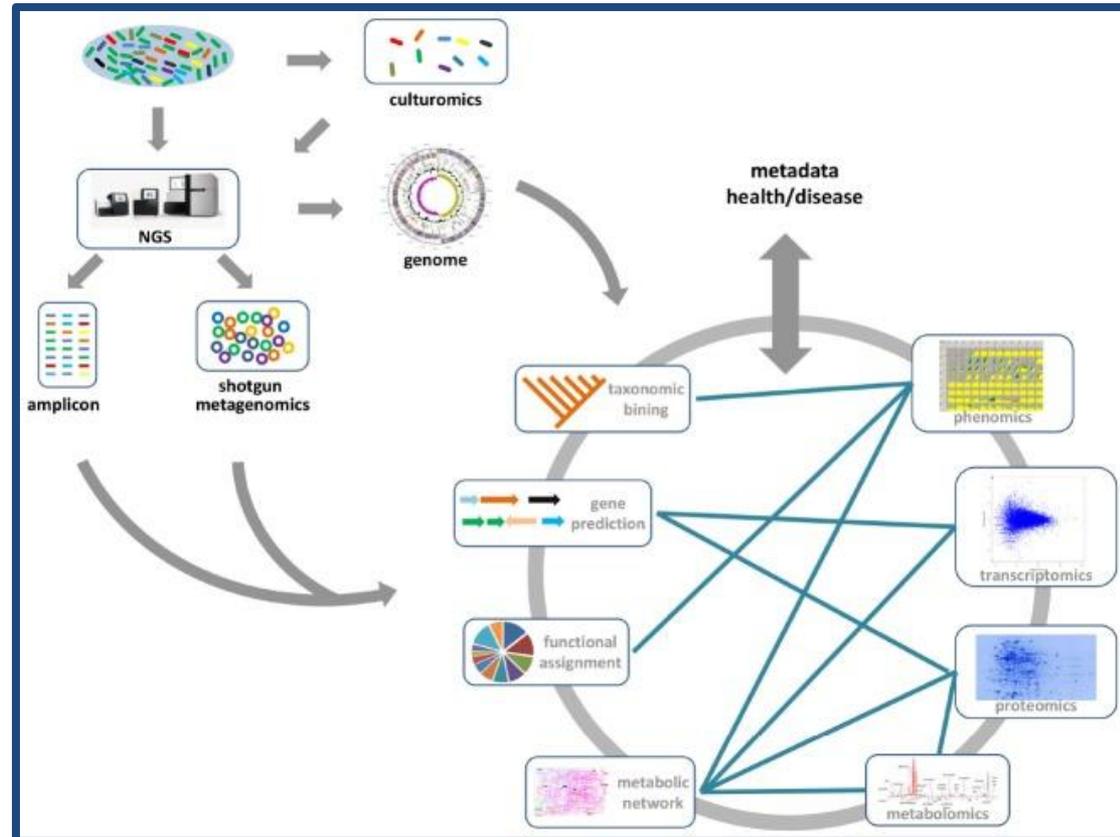
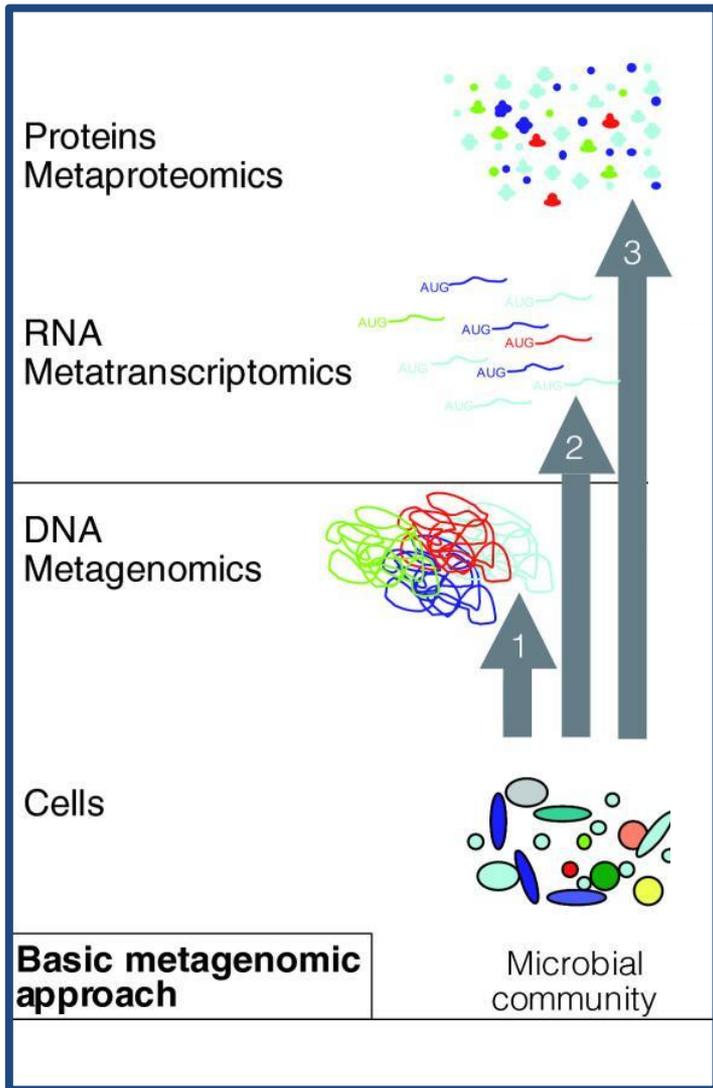
How
are they doing it?

What
are they doing?



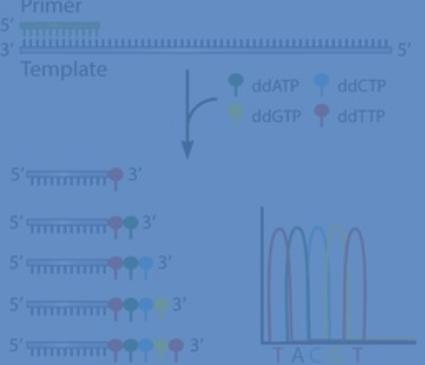
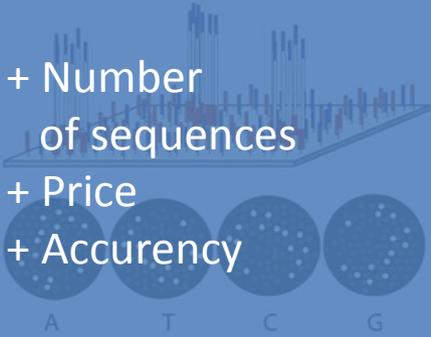
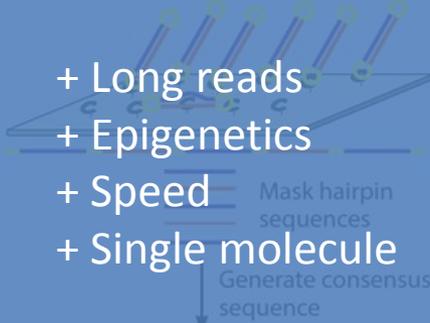
Whole metagenomic sequencing
Metatranscriptomics
Metaproteomics

Metagenomics



NGS

Next generation sequencing

First Generation Shotgun Sequencing	Second Generation Massively Parallel Sequencing	Third Generation Single-molecule Sequencing
 <ul style="list-style-type: none"> • Sequencing by synthesis • High accuracy • Long read lengths • Relatively small amount of data generated <p>e.g., ABI capillary sequencer (ABI)</p>	 <ul style="list-style-type: none"> + Number of sequences + Price + Accuracy - Template amplification - Short reads <ul style="list-style-type: none"> • Sequencing by synthesis • Amplification generated during sequencing, reducing the need for starting material • High accuracy • Short read lengths <p>e.g., MiSeq (Illumina), Ion Torrent (Thermo Fisher Scientific)</p>	 <ul style="list-style-type: none"> + Long reads + Epigenetics + Speed + Single molecule - Price - Accuracy <ul style="list-style-type: none"> • Single-molecule templates • Low accuracy • Long read lengths <p>e.g., Single-Molecule Real-Time (SMRT) — Sequencing (Pacific Biosciences), MinION (Oxford Nanopore Technologies)</p>



©2014 Illumina, Inc. All rights reserved.



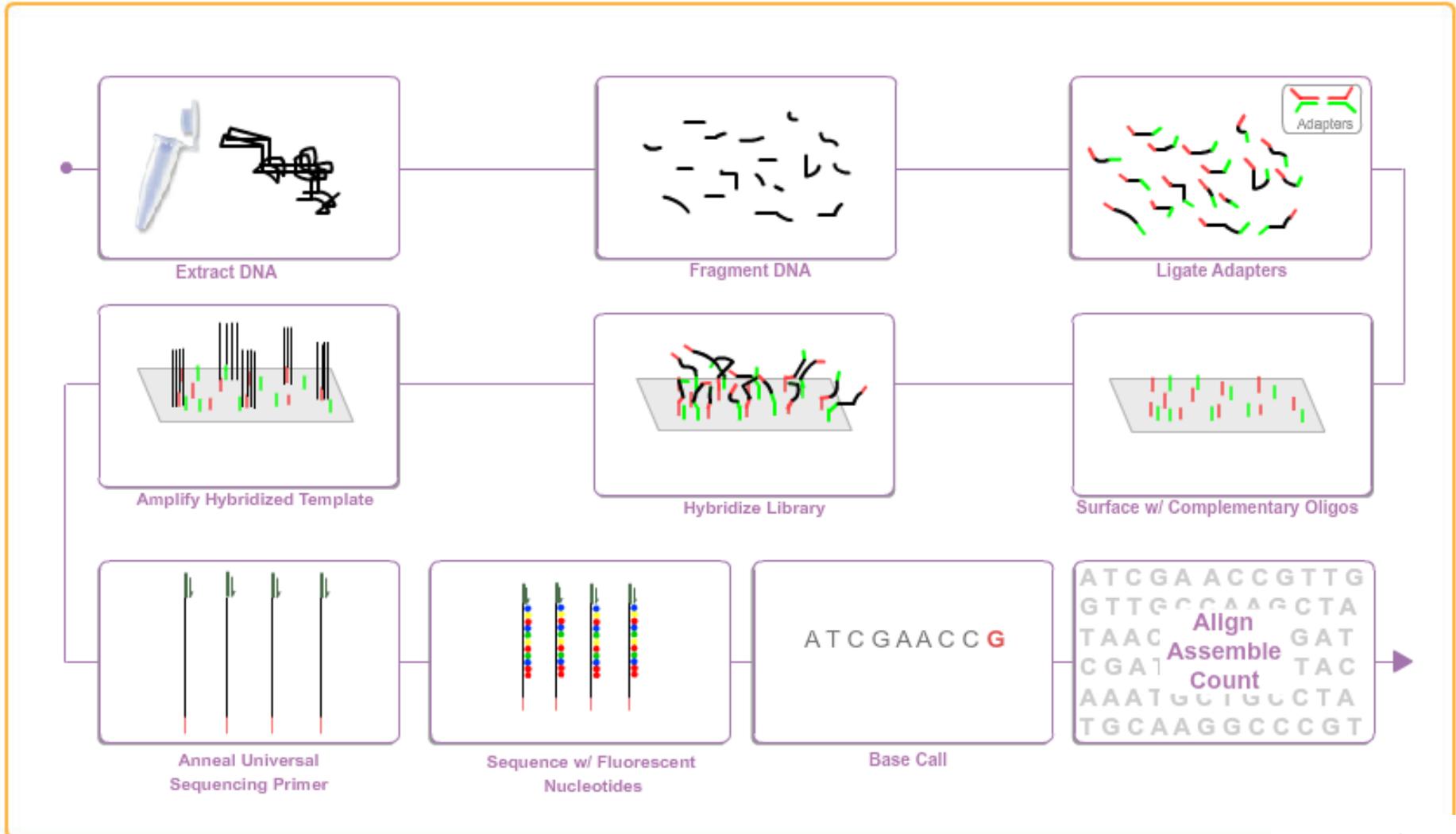
Available platforms

- 454 (Roche)
- SOLiD (Life Technologies)
- Illumina (Illumina)
- Ion Torrent (Life Technologies)
- **PACBIO, Sequel System (Pacific BioSciences)**
- **MiniON (Oxford Nanopore Technologies)**
- **BGISEQ-500 (BGI)**



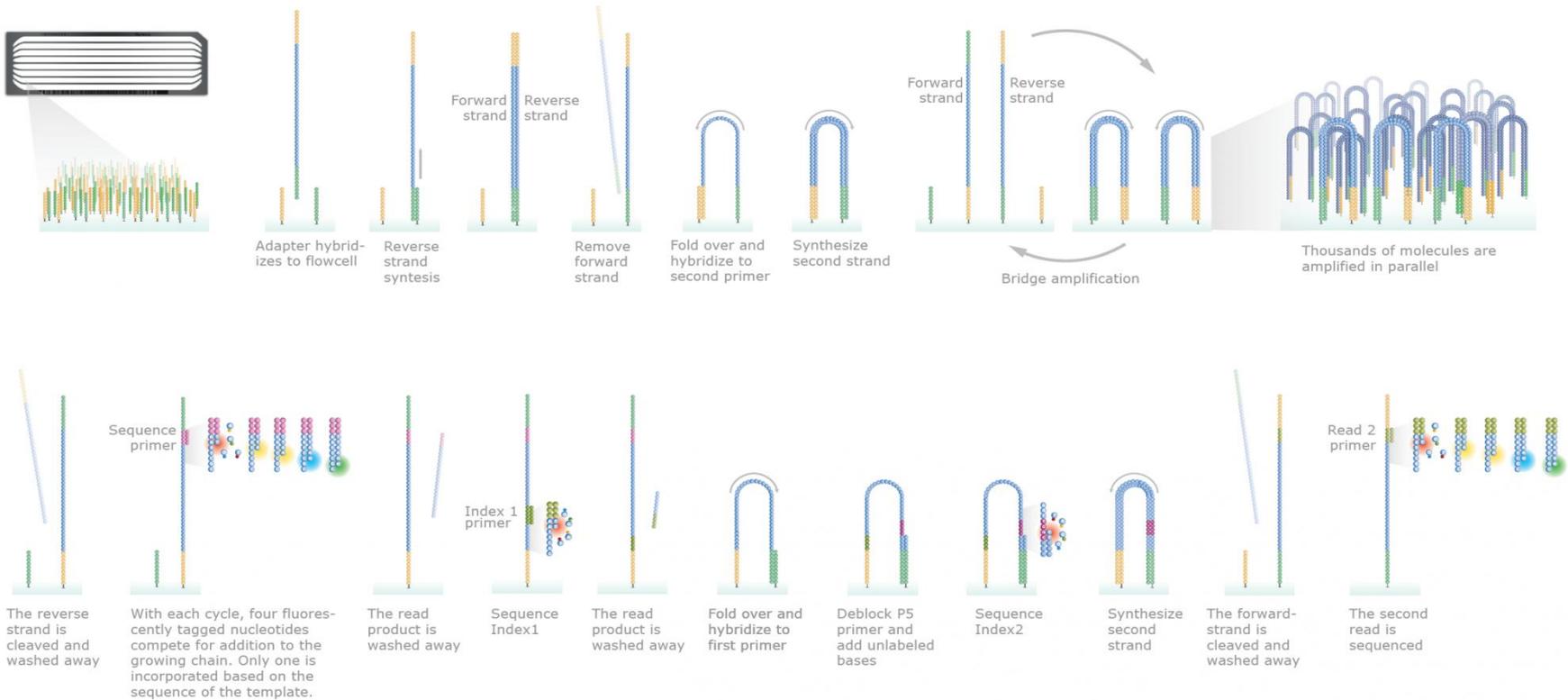
Illumina sequencing overview

Sequencing by Synthesis (SBS) Overview



Illumina sequencing overview

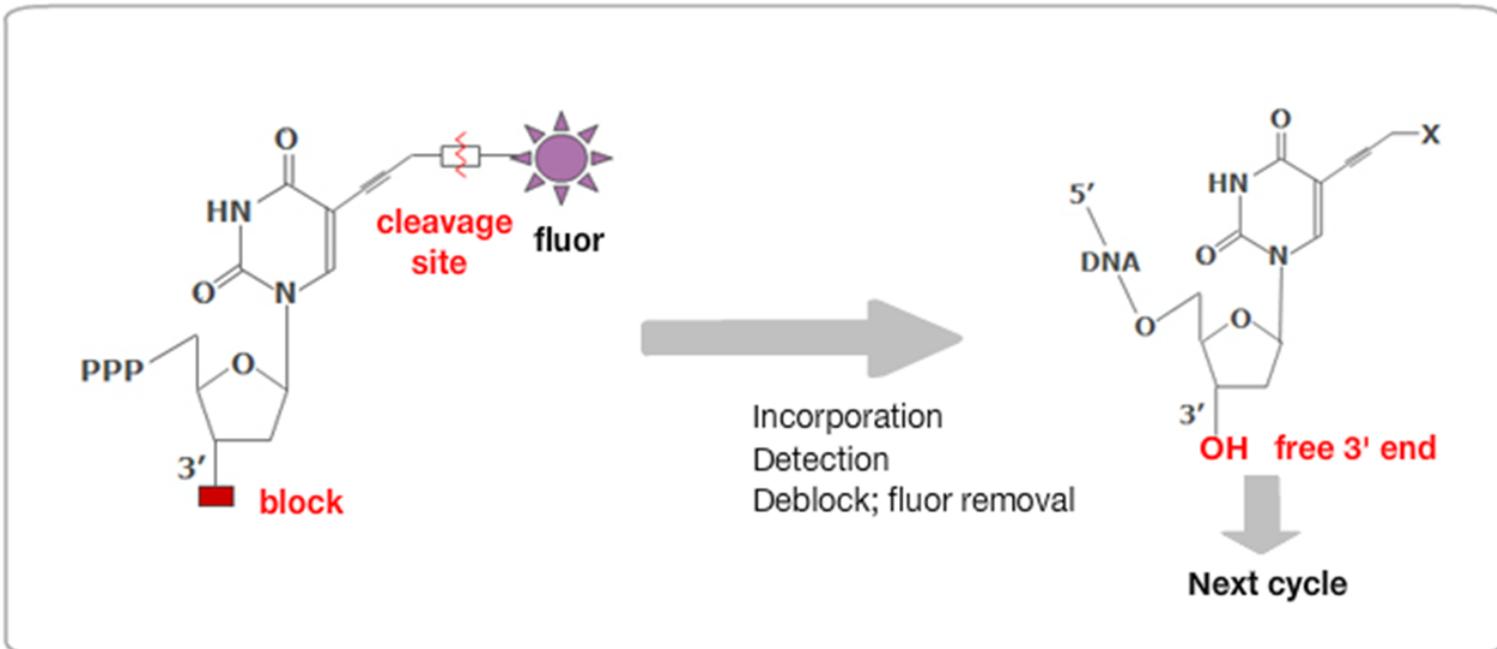
SBS sequencing technology



Illumina sequencing overview

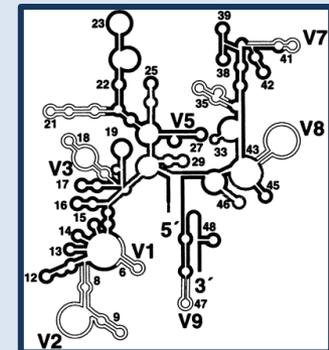
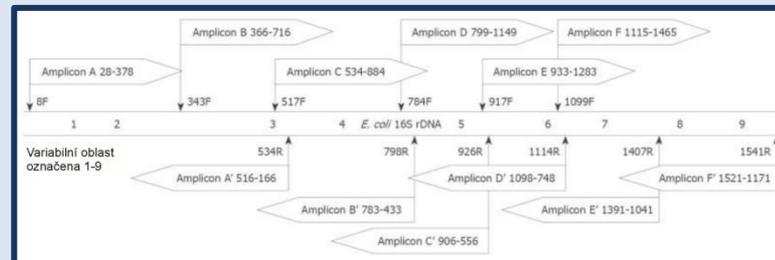
The MiSeq sequences the DNA clusters using Illumina's Sequencing By Synthesis (SBS) Chemistry which relies on Reversible Terminator Chemistry (RTC).

- All 4 labeled nucleotides in 1 reaction
- Higher accuracy

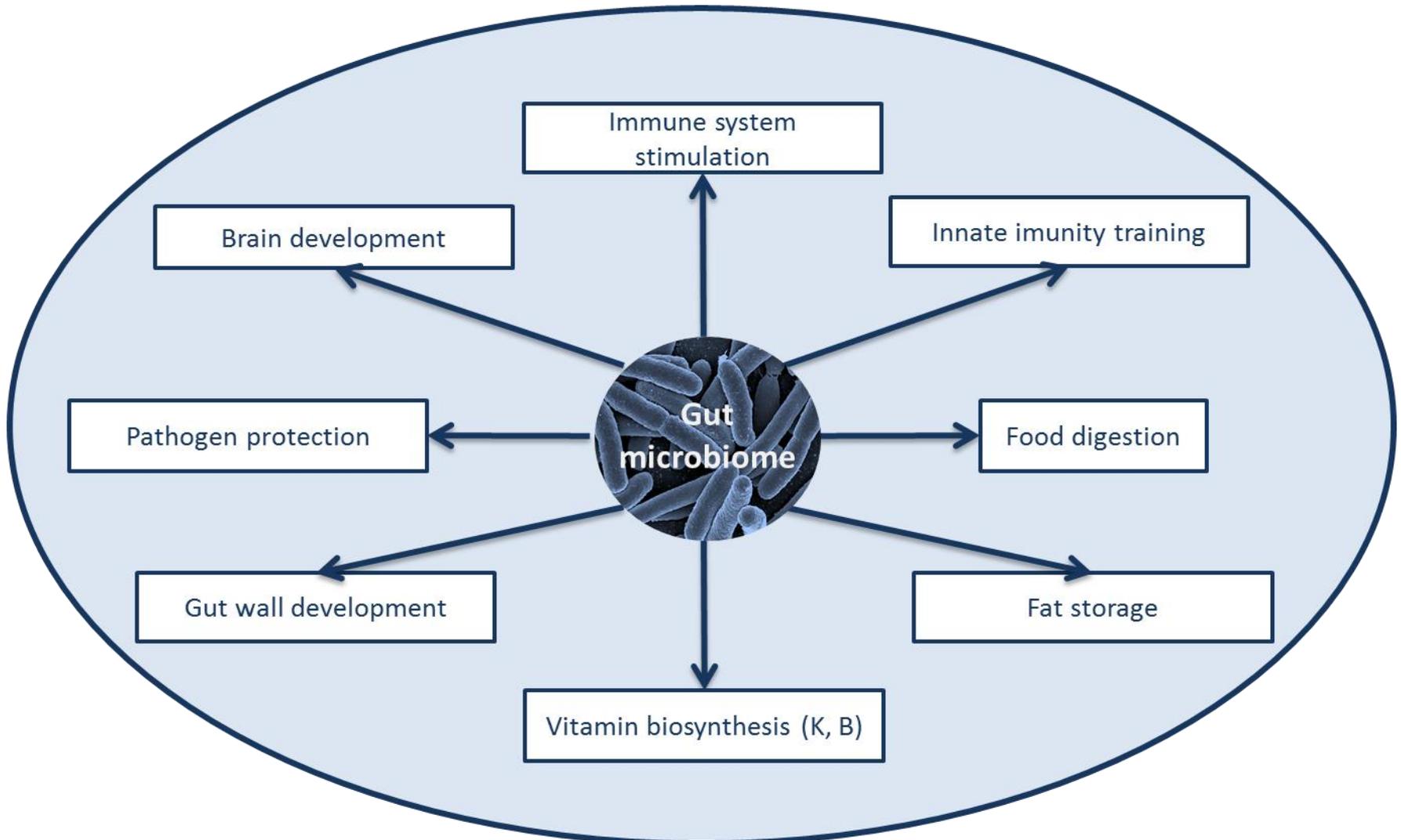


Microbiome studies – methodological problems

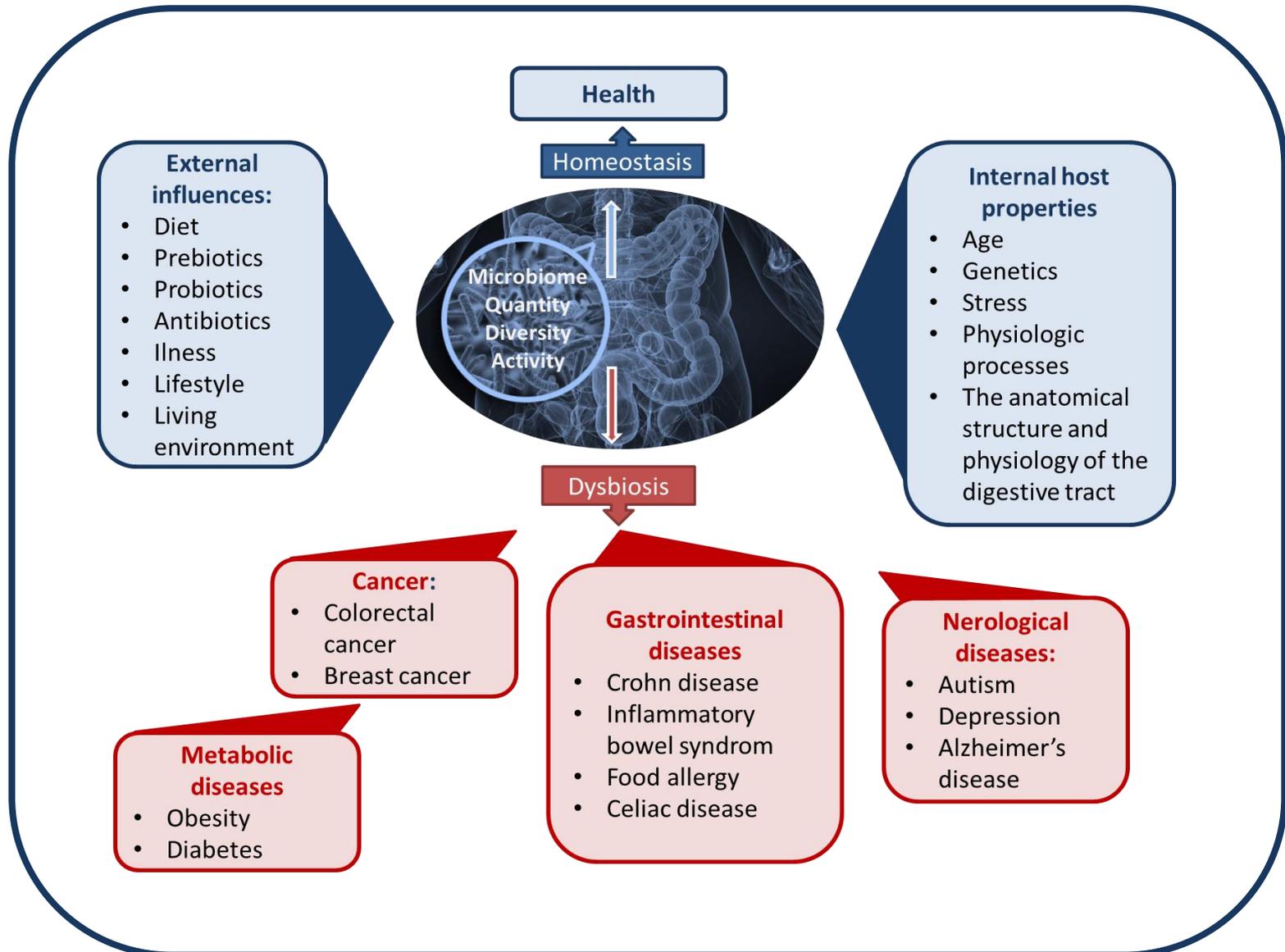
- Inconsistent approaches microbiome studies
- The composition of bacteria is influenced
 - Sampling method
 - Sample storage conditions
 - **DNA extraction**
 - Primers targeting different parts of 16S rRNA gene
 - Bioinformatics analyses



Microbiome function

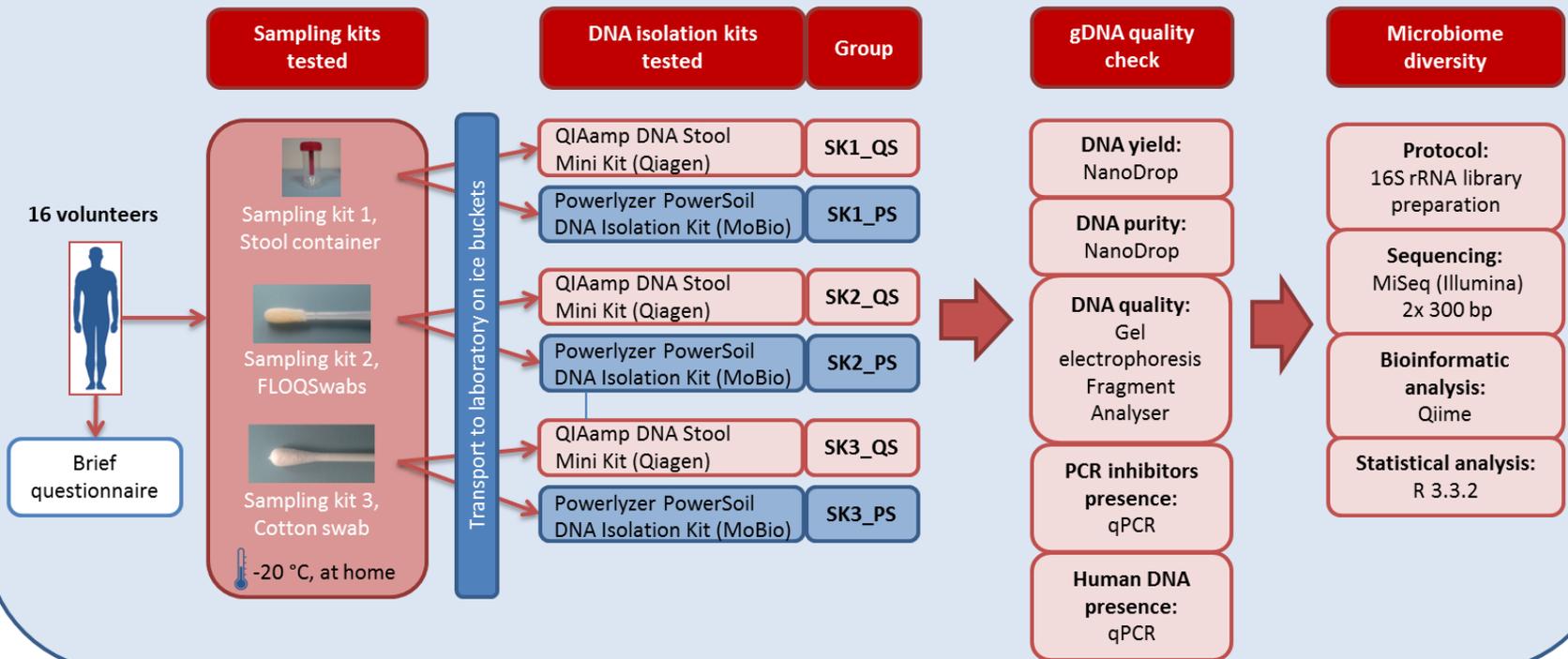


Microbiome and health

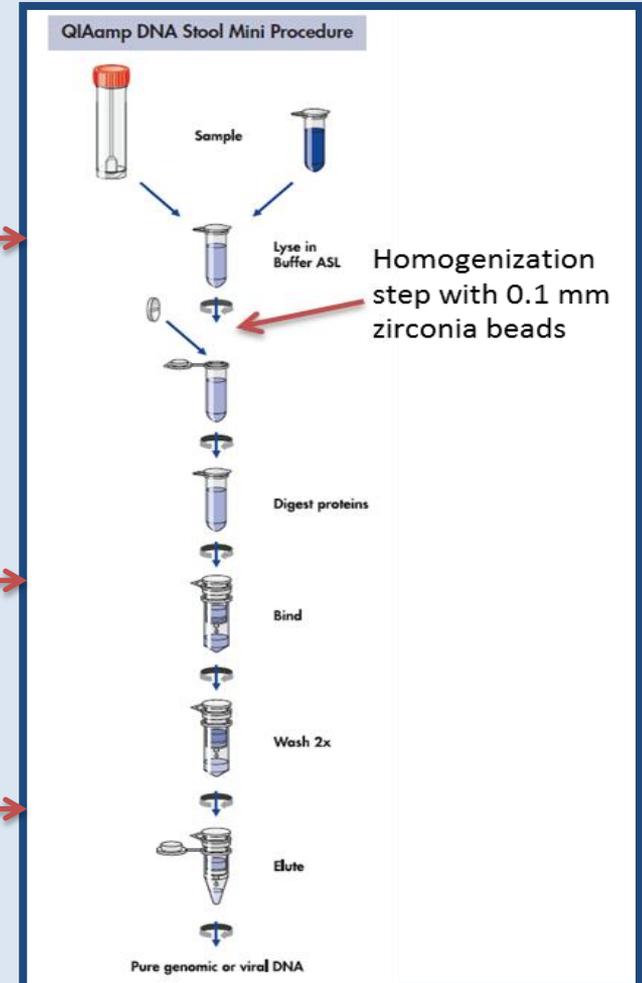
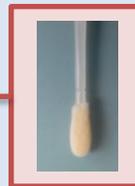
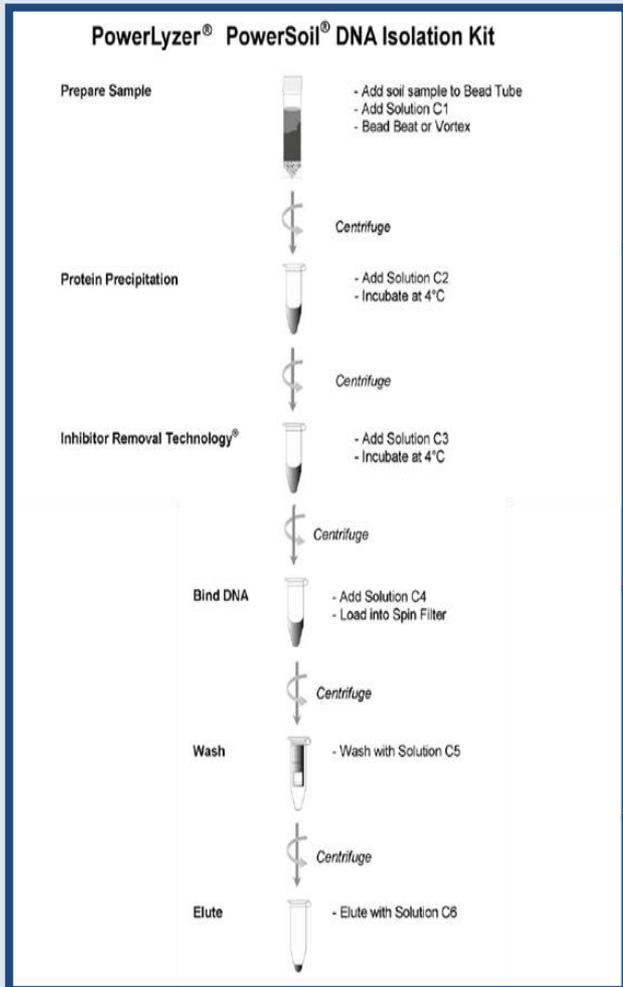


Our Study Design

Study design

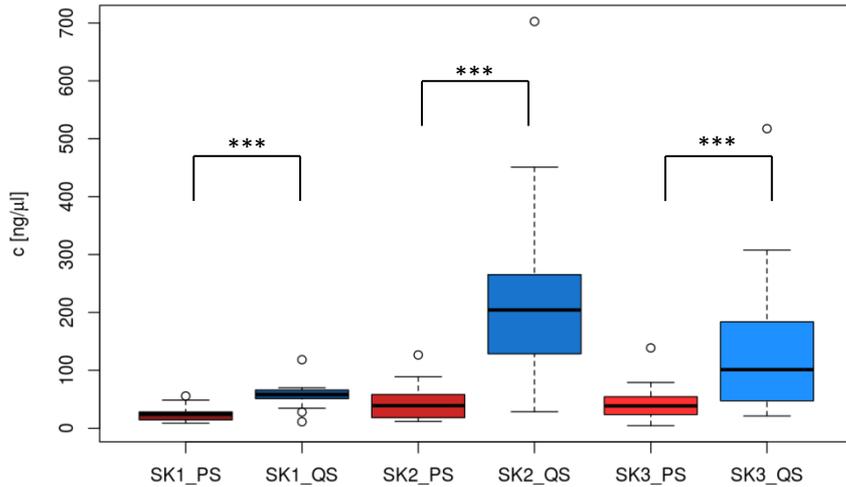


DNA Extraction



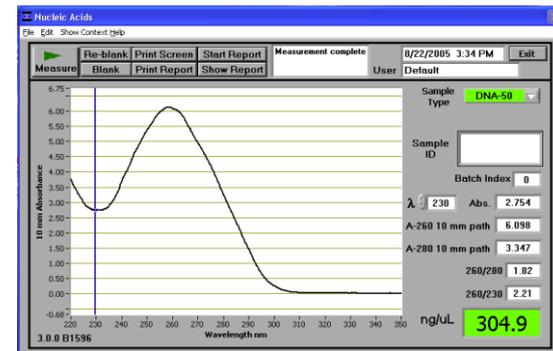
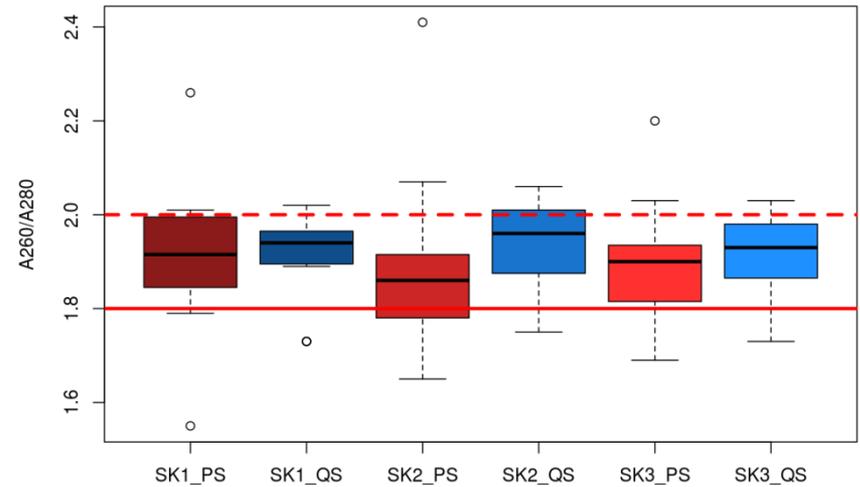
gDNA quality check

DNA yield



5x diluted DNA in
extraction DNA

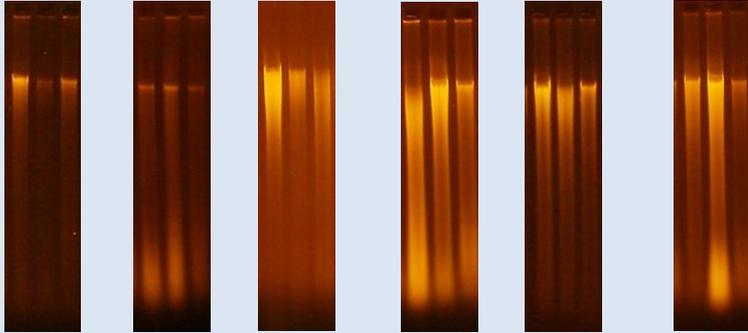
DNA purity



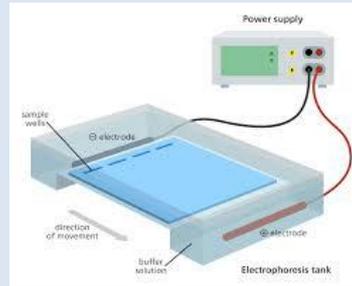
gDNA quality check

Gel electrophoresis

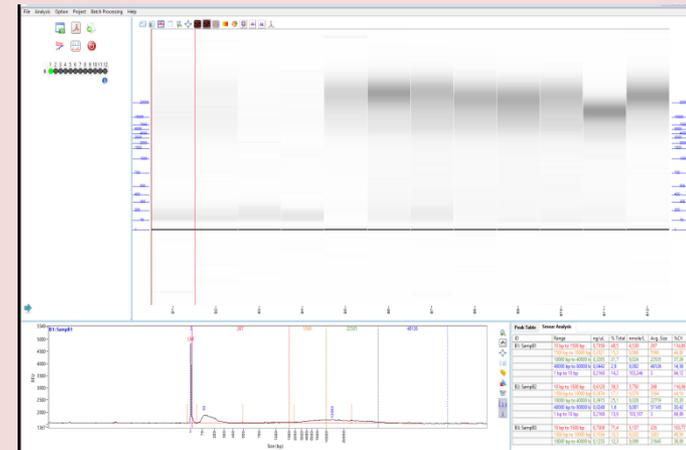
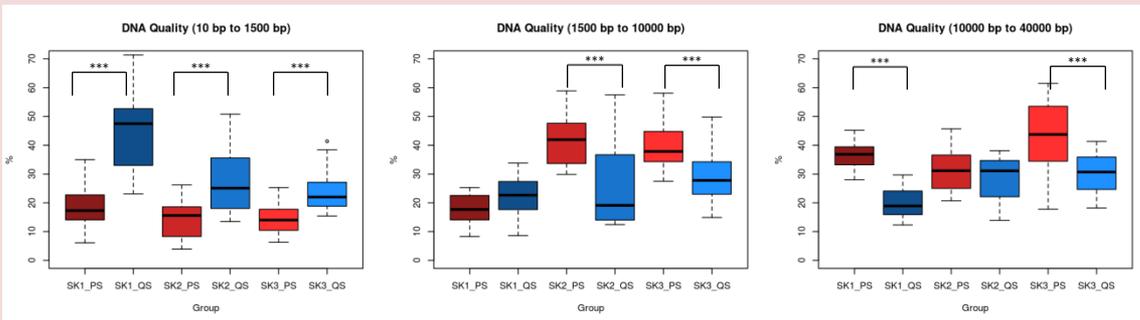
SK1_PS SK1_QS SK2_PS SK2_QS SK3_PS SK3_QS



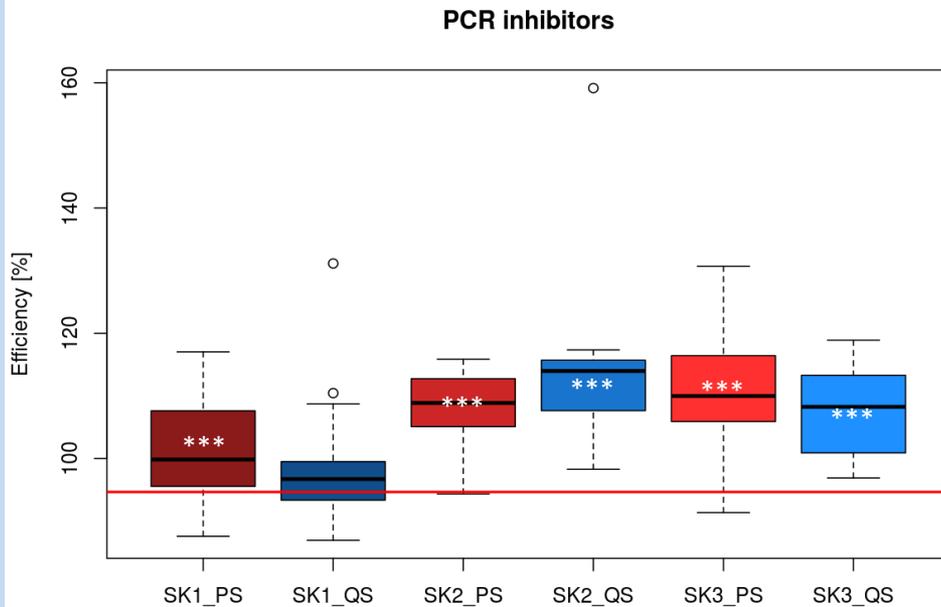
5x diluted DNA in extraction DNA



Fragment analyzer



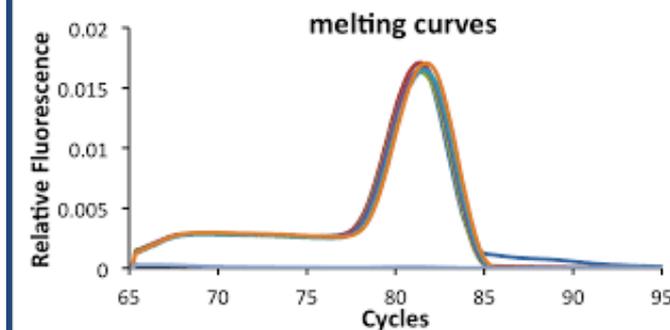
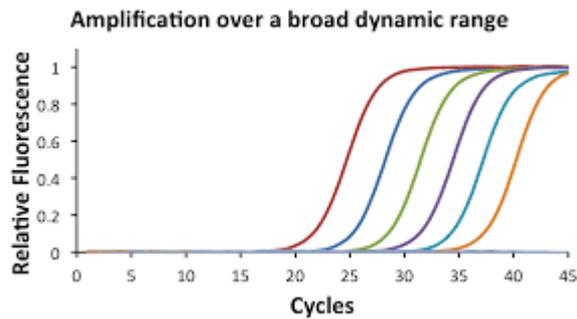
PCR inhibitors presence



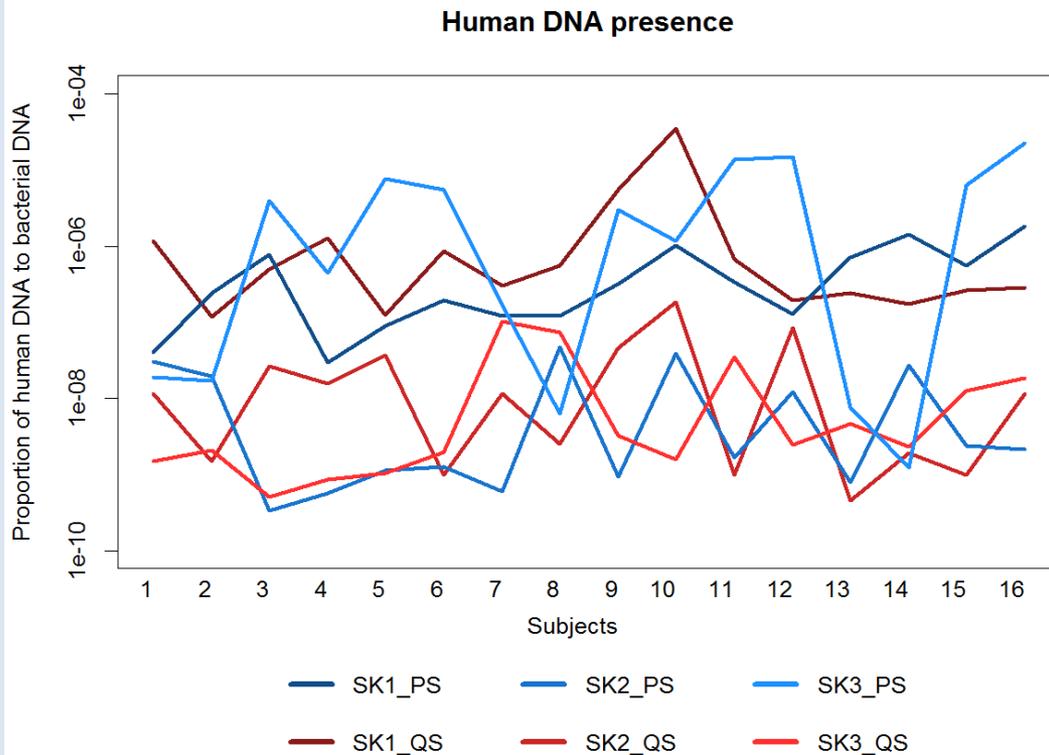
Control
DNA
(*E. coli*)

* Sample groups x control DNA

- qPCR → primers targeted bacterial DNA (16S rDNA)
- series of gDNA dilution
- Calculated slope and efficiency ($= 10^{(-1/\text{slope})} - 1$) of the reaction



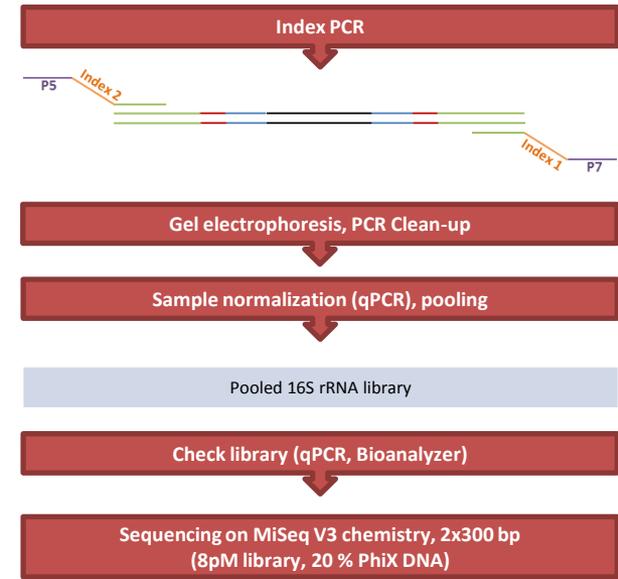
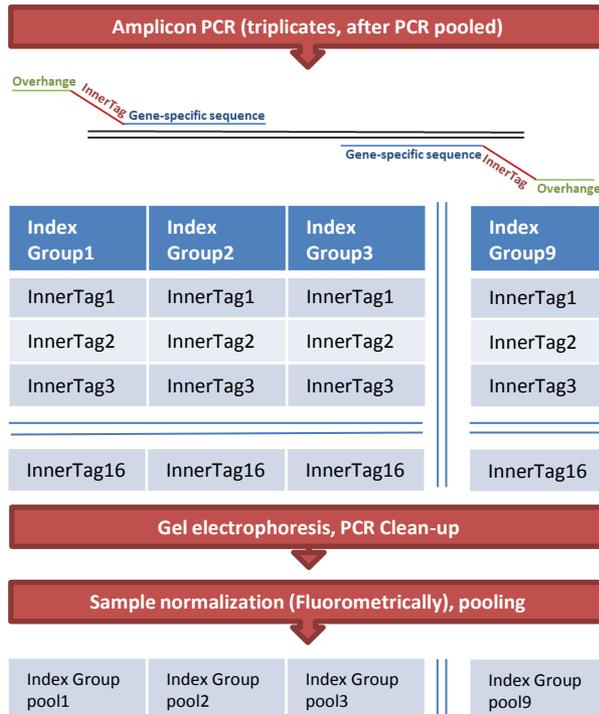
Human DNA presence



- qPCR → primers targeted bacterial DNA (16S rDNA) and human DNA (protein kinase)
- Relative quantification – proportion of human DNA to bacterial DNA (2^{dCt})

16S rDNA Sequencing

Library preparation

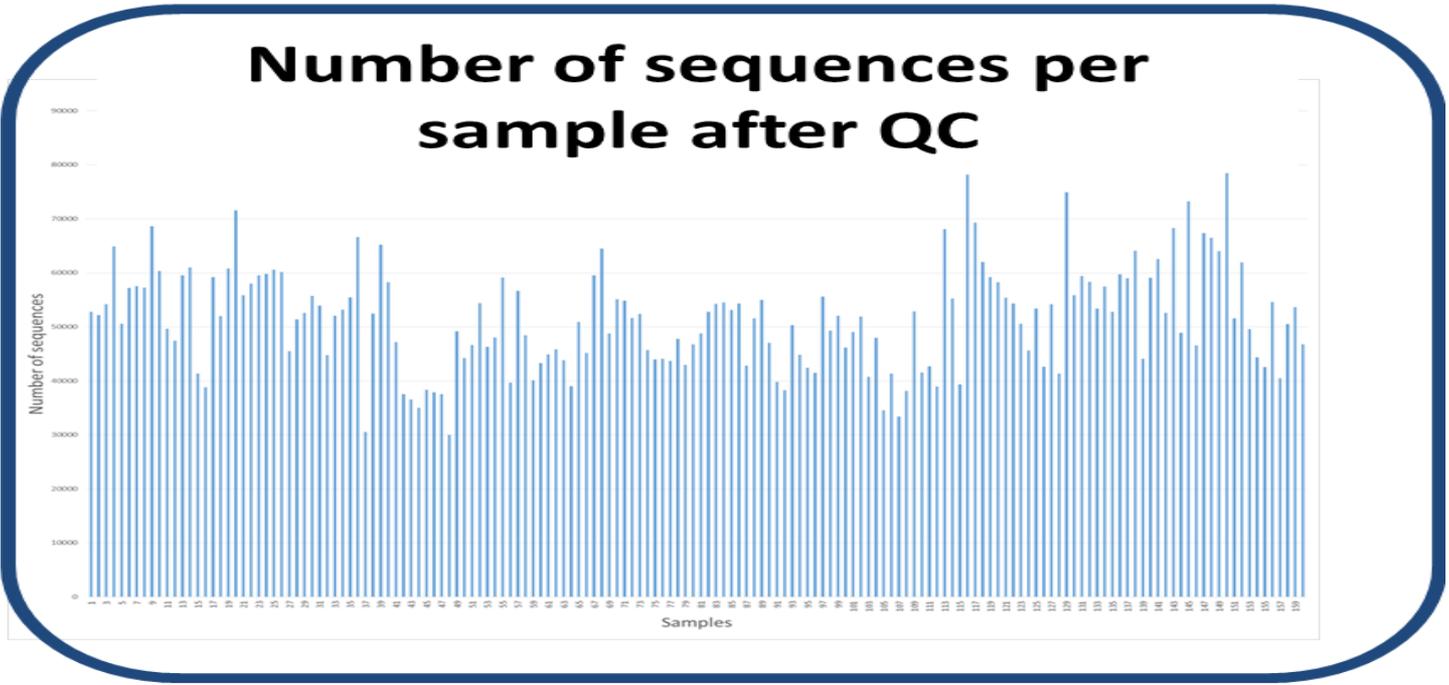


```

-AAAA02020713.1.1297.121.53.2_303789
TGGGGAAATTTCACAAATGGGGGCAAGCCCTGATGCGACGACCCGCTGATGAGAAAGCCCTCGGGTGT
AAAGCTTCTACGCGGGGAGAGAGAGGATTAAGATTAACCTTTGCTACTGACGCTACCGGAGAGAGAGCA
CCGCGCACTCCGTCACAGACCCCGGATATCGAGAGGATTCGACAGCGCTTACTGGAATCTGCGGCTTAA
GCGCAGCGAGCGGTTTAAATCAGATGTAATATCCCGGCGCACTCCGCGGACTGATGATGATGATGATG
AAGCTTGAAGTGAAGTCTGCTTTTAAAGTGGGAAATCTTGGATGATGATGATGATGATGATGATGATG
GTTATGGAATTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG
ACTGACTGTACTGACGCTGAGGCGGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG
-#AAGL03001287.3.1442.120.55.1_1130405
TGGGGAAATTGGGCAATGGGGGAAACCTTACCCAGCACGCGCCGCTGAGAGAAAGCCCTCGGGTGT
TAAATCTTACTCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
GCGCGGTAATGATGAGGGGCAAGCGTTTCCGGAATGATGGGGGTAAGGGGCTGATGATGATGATGATG
AAGCTTGAAGTGAAGTCTGCTTTTAAAGTGGGAAATCTTGGATGATGATGATGATGATGATGATGATG
GTTATGGAATTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG
ACTGACTGTACTGACGCTGAGGCGGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG
-#AAGL0300929.1.1434.113.52.2_2457824
TGGGGAAATTTCACAAATGGGGGAAACCTTACCCAGCACGCGCCGCTGATGAGAAAGCCCTCGGGTGT
AAAGCTTCTACGCGGGGAGAGAAATGACGATGATGATGATGATGATGATGATGATGATGATGATGATG
GCGGTAATGATGAGGGGCAAGCGTTTCCGGAATGATGGGGGTAAGGGGCTGATGATGATGATGATGATG
AAGCTTGAAGTGAAGTCTGCTTTTAAAGTGGGAAATCTTGGATGATGATGATGATGATGATGATGATG
GTTATGGAATTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG
ACTGACTGTACTGACGCTGAGGCGGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG
GACTGAAACTGACATGACGCGGAAAGCGTGGGGAGCAACA
-#AAG02010287.435.1901.120.82.2_3909221
TGGGGAAATTTCCGCAATGGGGGAAACCTTACCCAGCACGCGCCGCTGATGAGAAAGCCCTCGGGTGT
TAAATCTTCTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
CCGCGGTAATGATGAGGGGCAAGCGTTTCCGGAATGATGGGGGTAAGGGGCTGATGATGATGATGATG
AAGCTTGAAGTGAAGTCTGCTTTTAAAGTGGGAAATCTTGGATGATGATGATGATGATGATGATGATG
GTTATGGAATTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG
ACTGACTGTACTGACGCTGAGGCGGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG

```

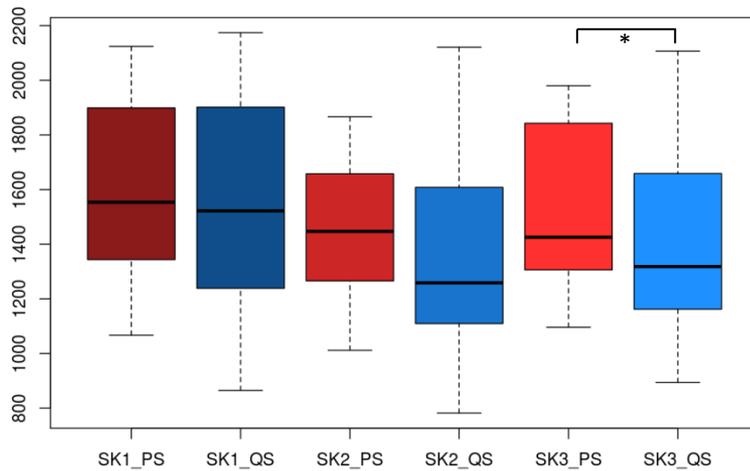
Sequencing results



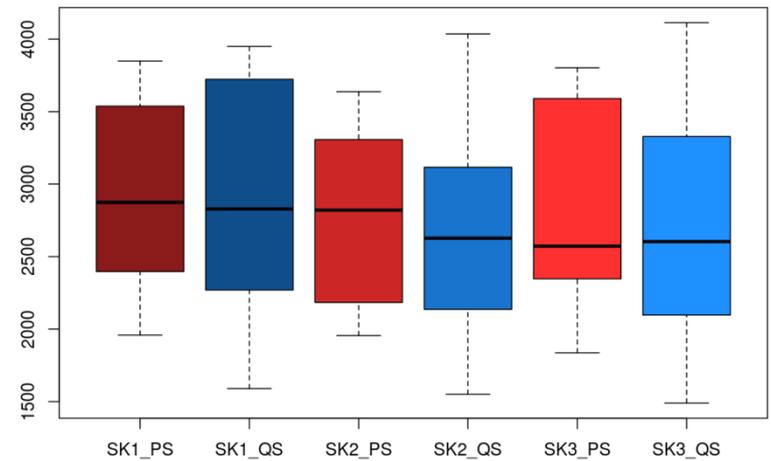
Number of sequences per sample after QC	
Average:	51 594
Max.:	78 498
Min.:	30 012

Stool diversity

Observed species

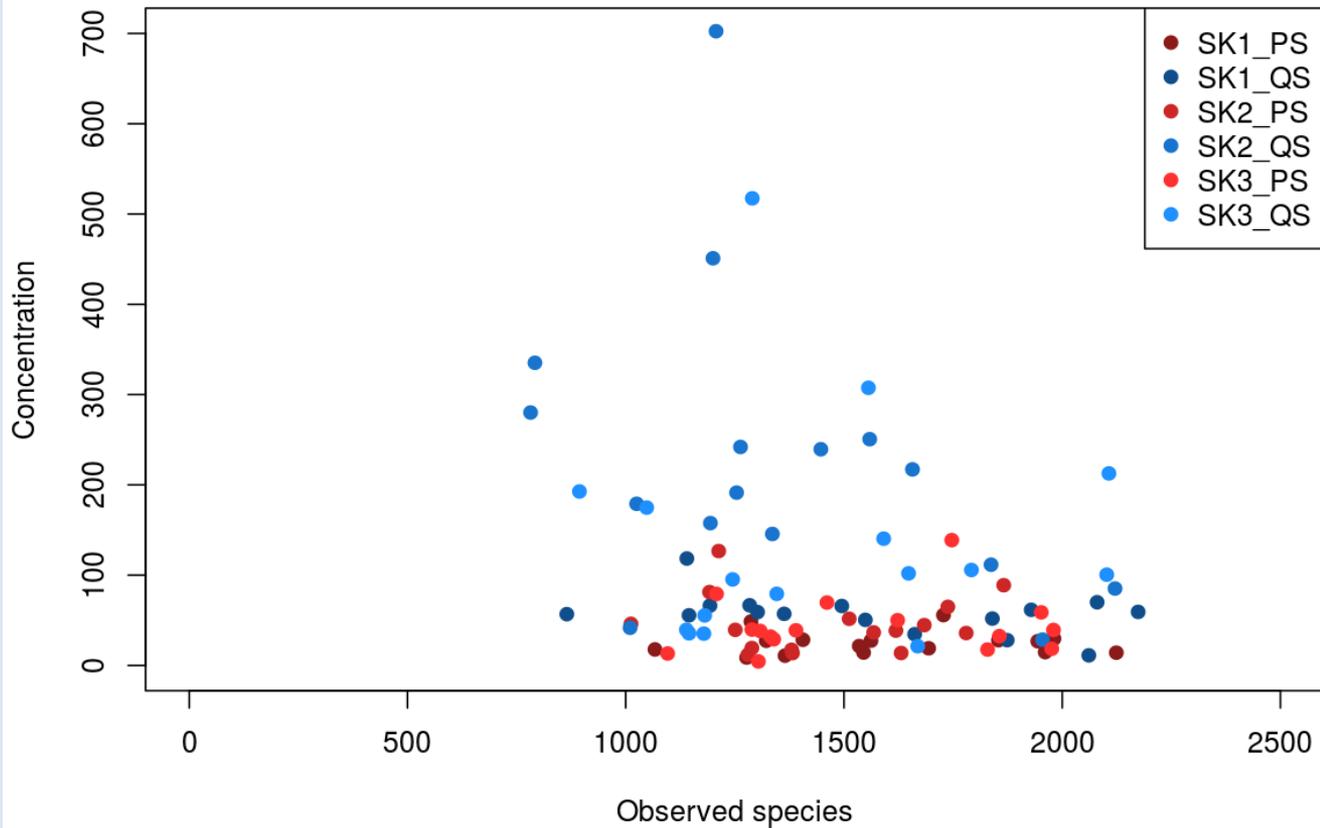


Chao 1 richness estimation



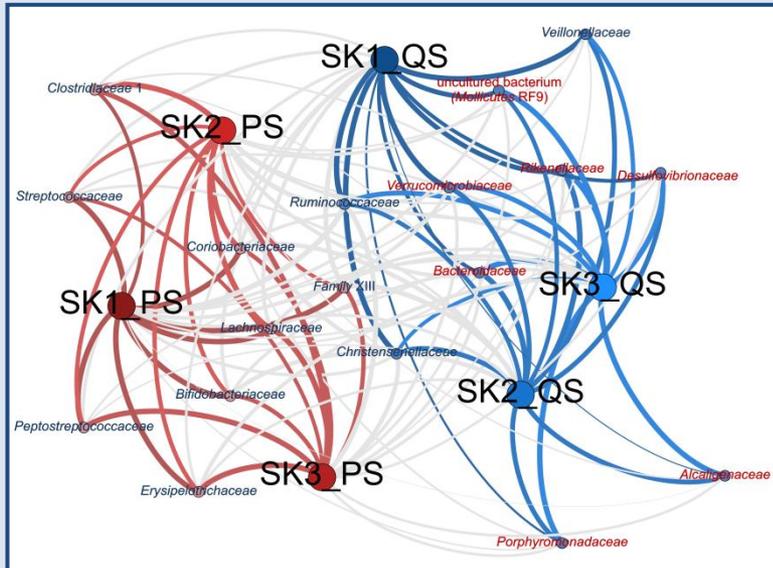
Diversity x DNA yield

Correlation of concentration with observed species

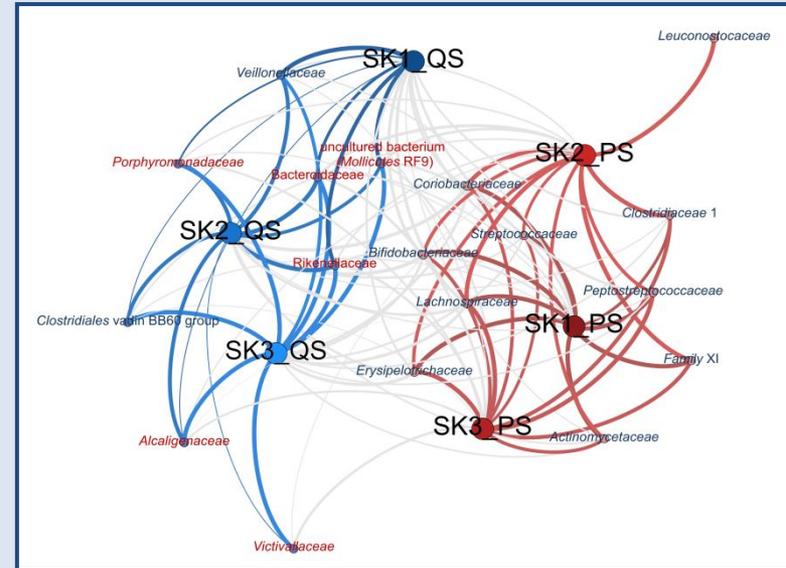


DNA isolation vs bacterial cell wall type

Abundant bacteria

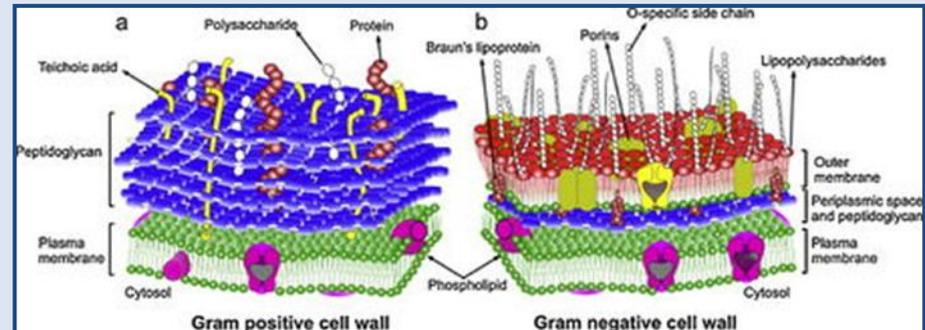


Statistically significant different bacteria



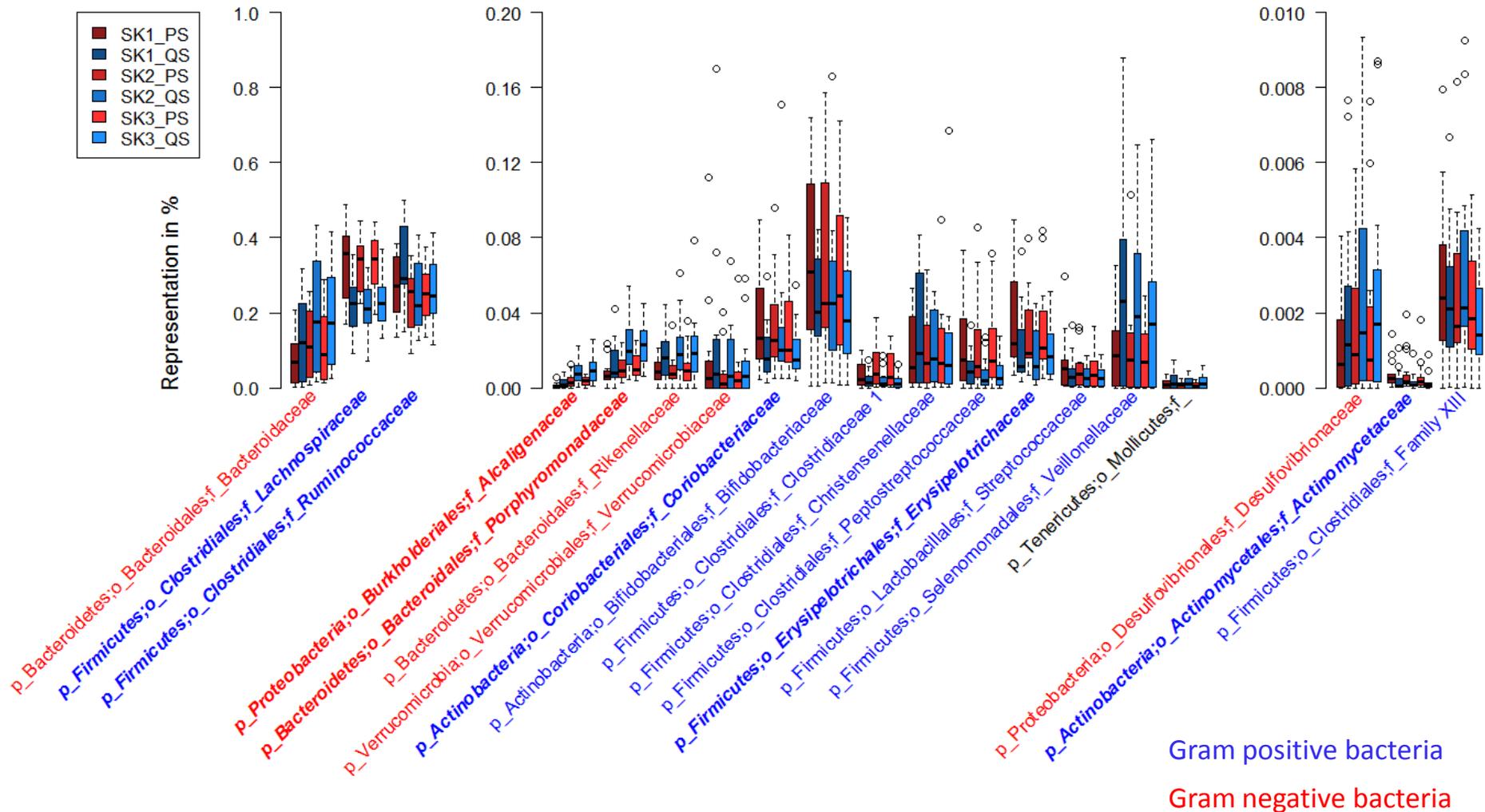
Gram positive bacteria

Gram negative bacteria

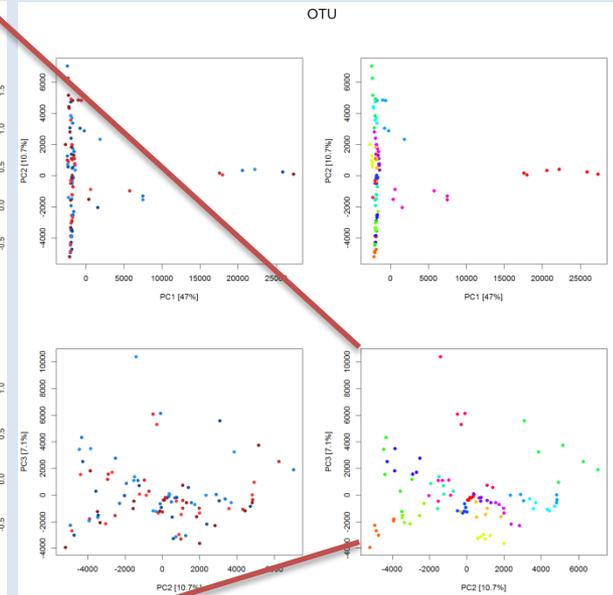
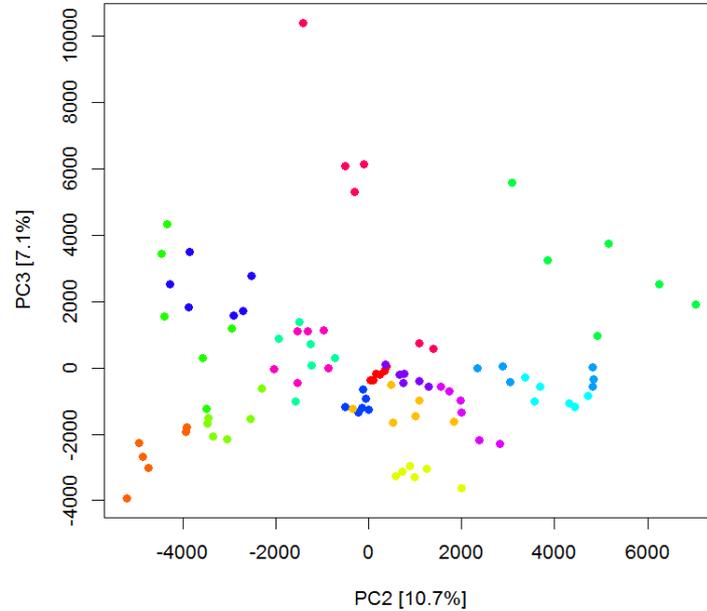
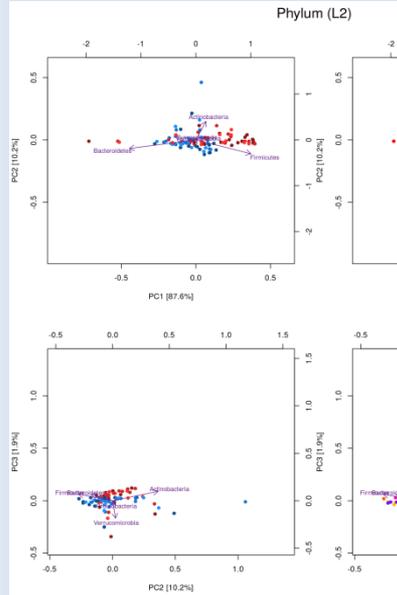


Bacterial composition

Bacterial composition



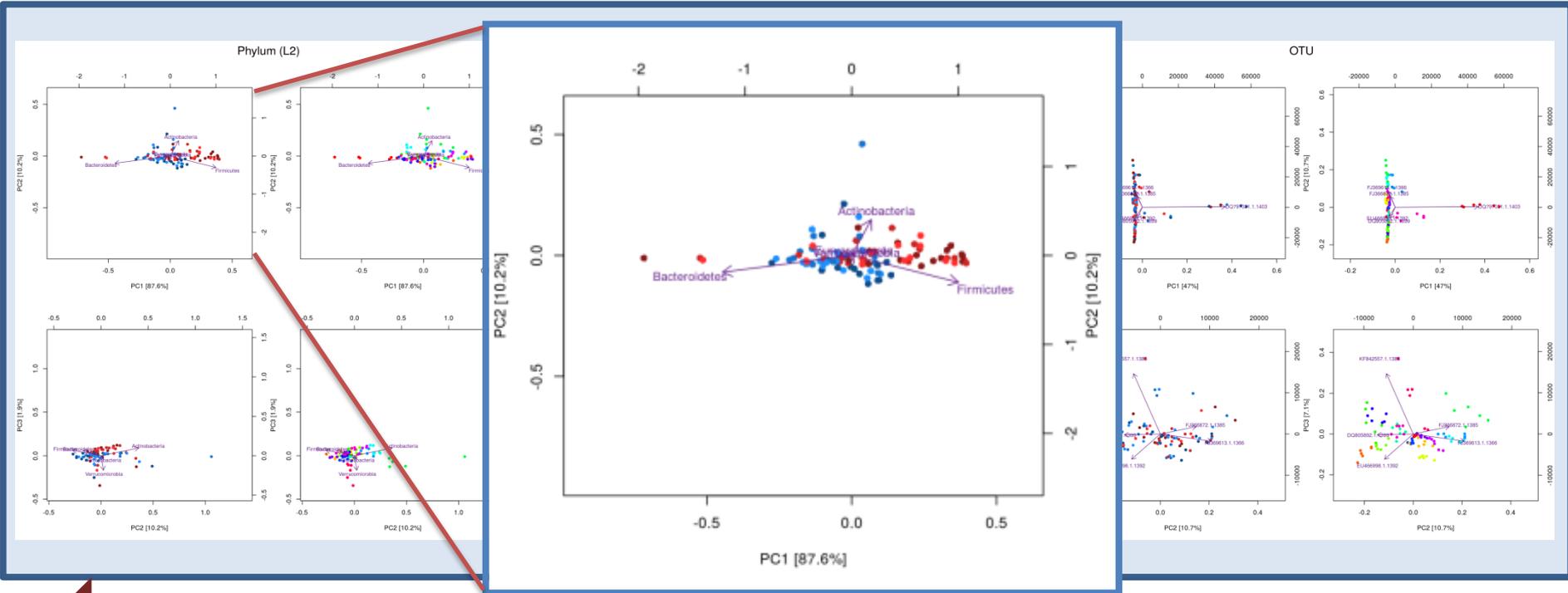
DNA isolation – Does it matter?



DNA isolation

Individual variability

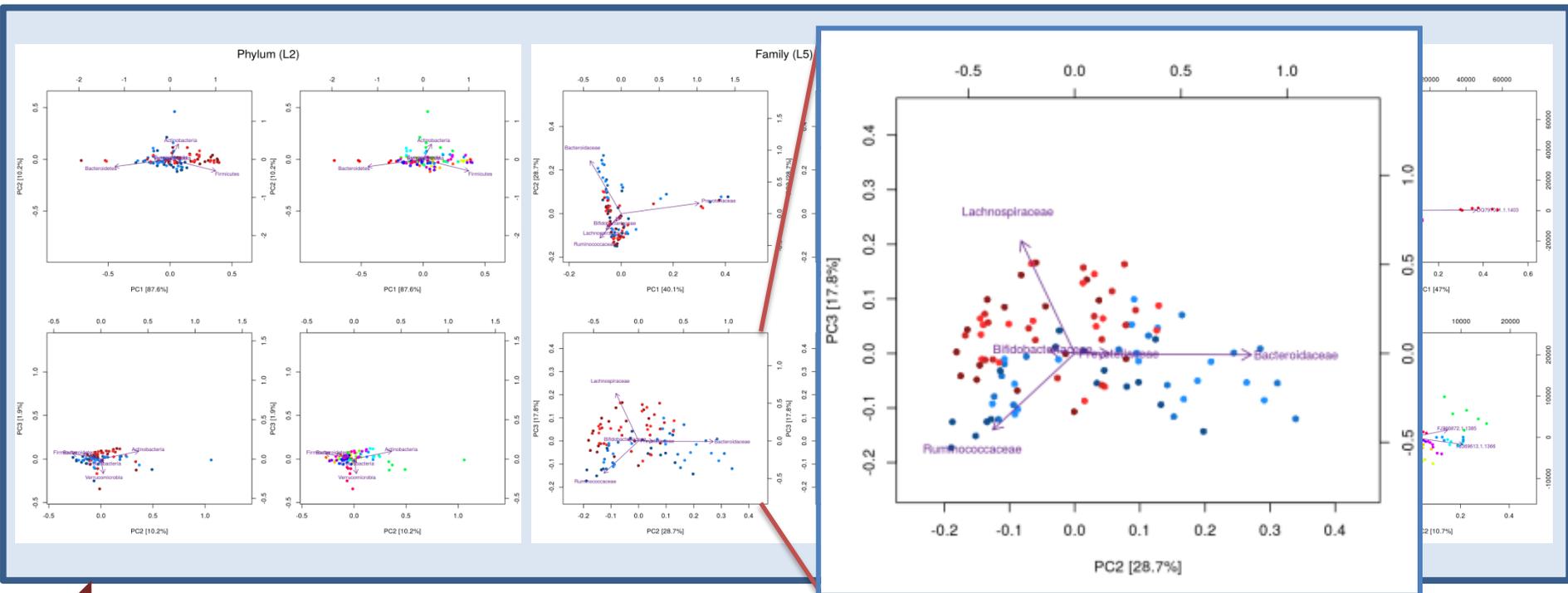
DNA isolation – Does it matter?



DNA isolation

Individual variability

DNA isolation – Does it matter?



DNA isolation

Individual variability

Summary

	PS (Powerlyzer PowerSoil DNA isolation Kit)	QS (QIAamp DNA Stool MINI Kit)
DNA yield	↓	1 ↑*
DNA purity	1 ↑	2 ↑
DNA quality	2 ↑	↓
Presence of inhibitors	↓	↓
Presence of human DNA	3 ↑	3 ↑
Observed species	4 ↑	4 ↑
G+ bacteria isolation	5 ↑	↓

*There is no statistical significant correlation between DNA yield and observed species

The effect of extraction kits increases with higher taxonomical level
!!!Complication in comparison between different studies!!!

THANKS FOR YOUR ATTENTION

