Latvian Biomedical Research and Study Centre
research and education in biomedicine from genes to human

EUROPEAN COOPERATION IN SCIENCE AND TECHNOLOGY

MICROBIOME

# Overview of the microbiome data

*Microbiome study design, sampling, wet-lab key steps.*

PhD. Ilze Elbere

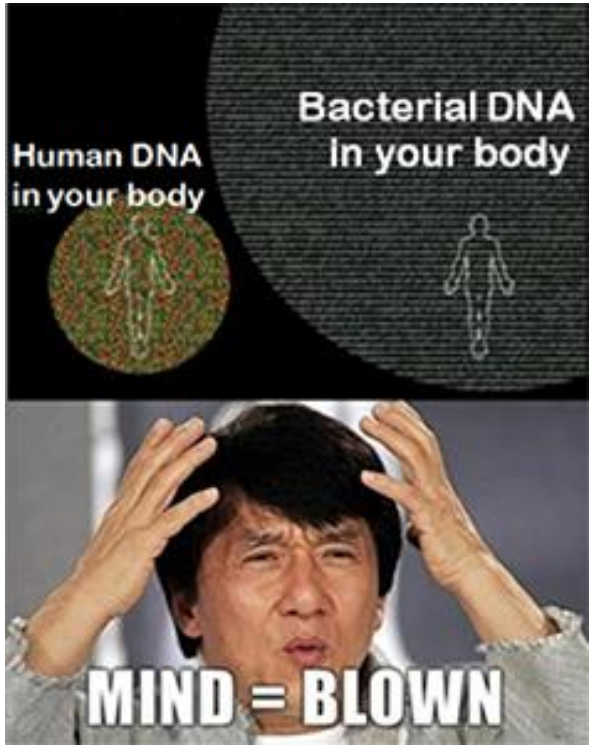
Latvian Biomedical Research and Study Centre

ML4Microbiome

Workshop on Statistical and Machine Learning Techniques for Microbiome Data Analysis

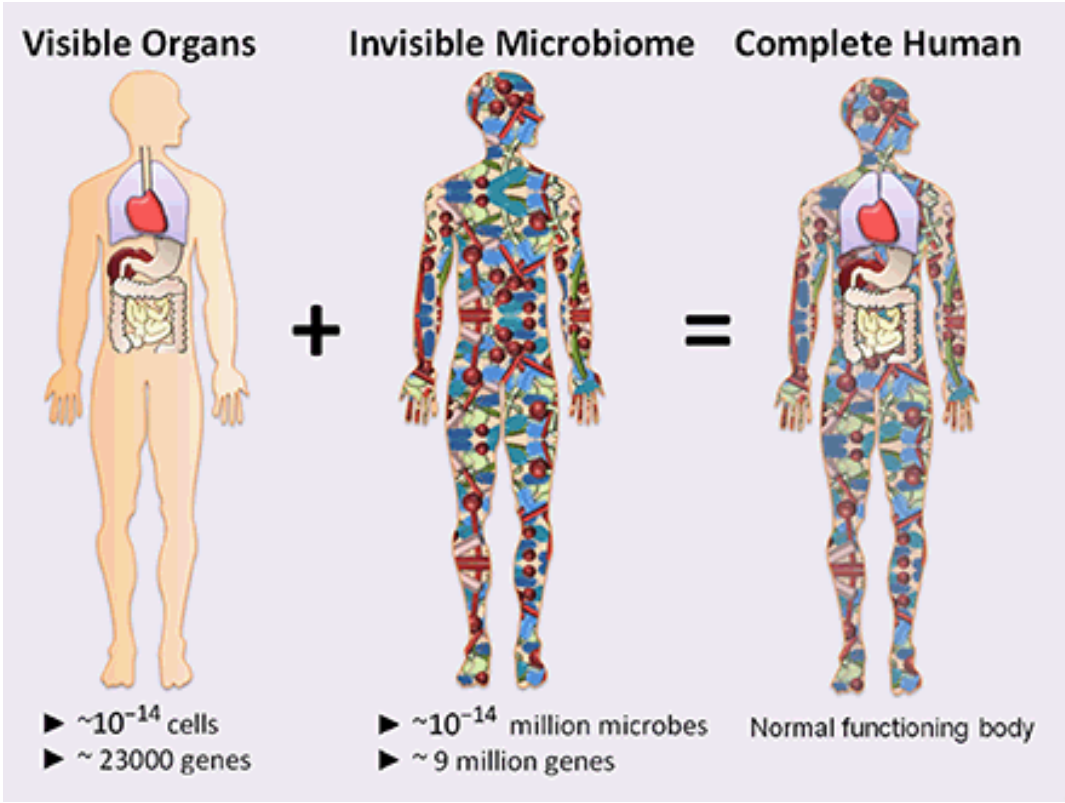
# Outline

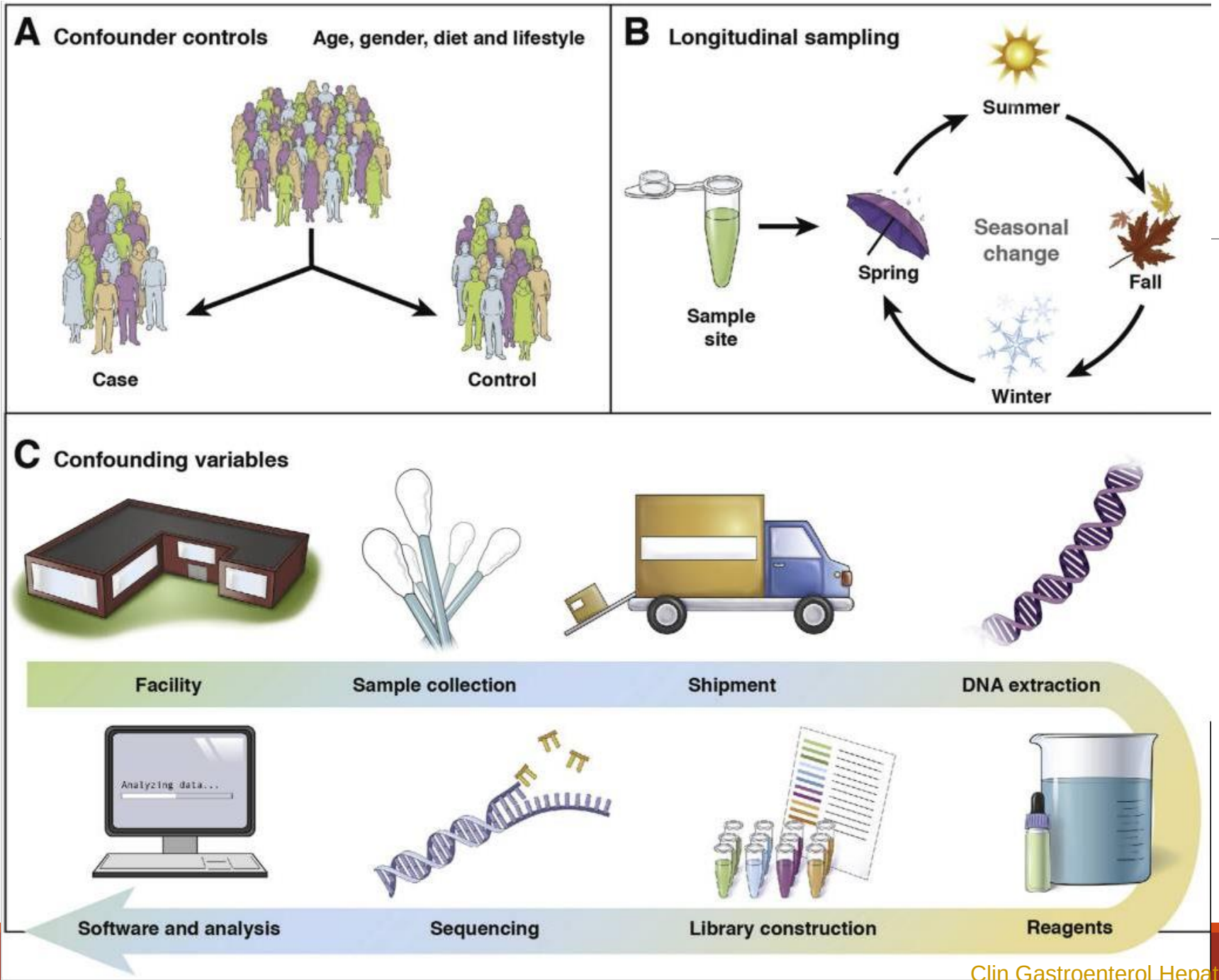
- Before sampling
  - What is the aim?
  - Study design
  - Power calculations
- Methods - overview
- Sampling
  - Logistics
  - Sample type
  - Sample transportation and storage
- DNA extraction
  - Enzymatic vs Mechanistic
- DNA processing & contamination



Appanna V.D. (2018) The Human Microbiome: The Origin. In: Human Microbes - The Power Within. Springer, Singapore. https://doi.org/10.1007/978-981-10-7684-8_1

# The Microbiome



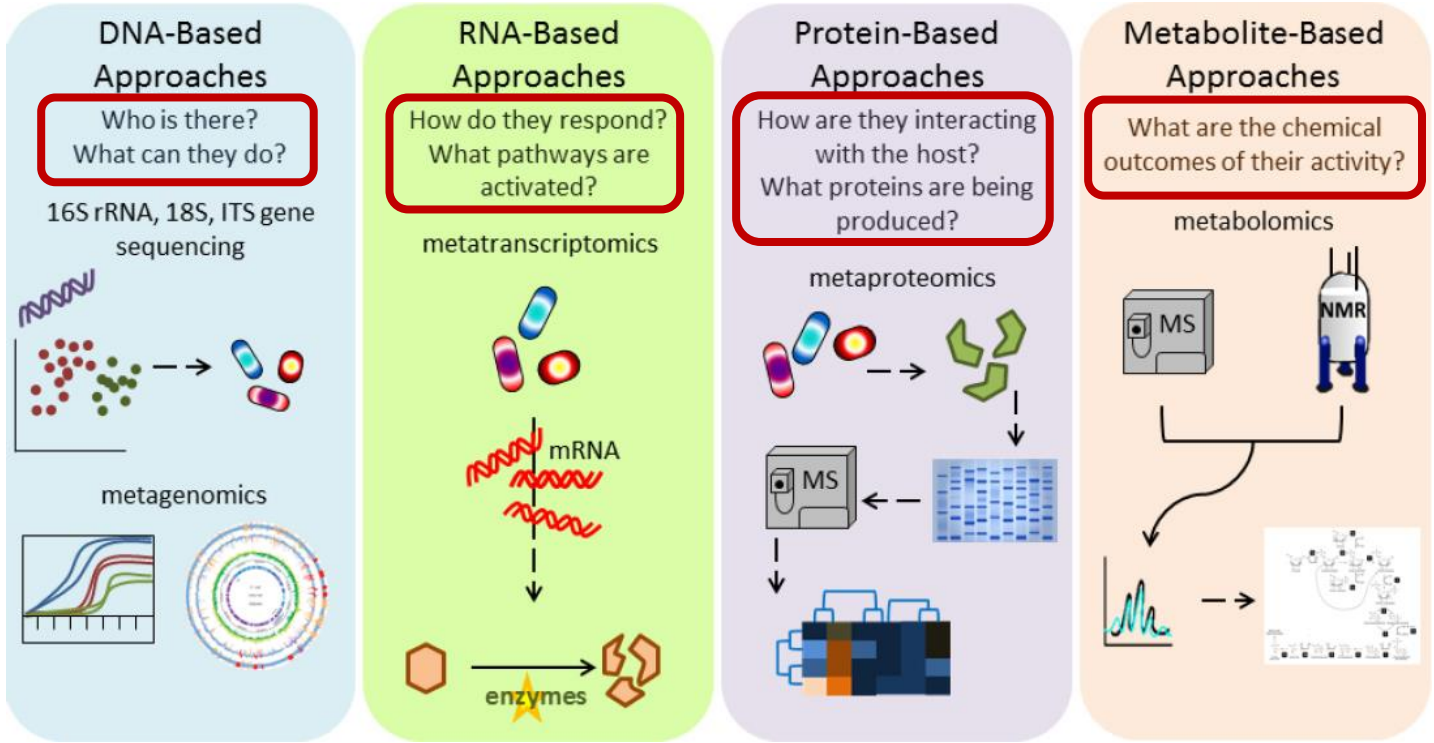


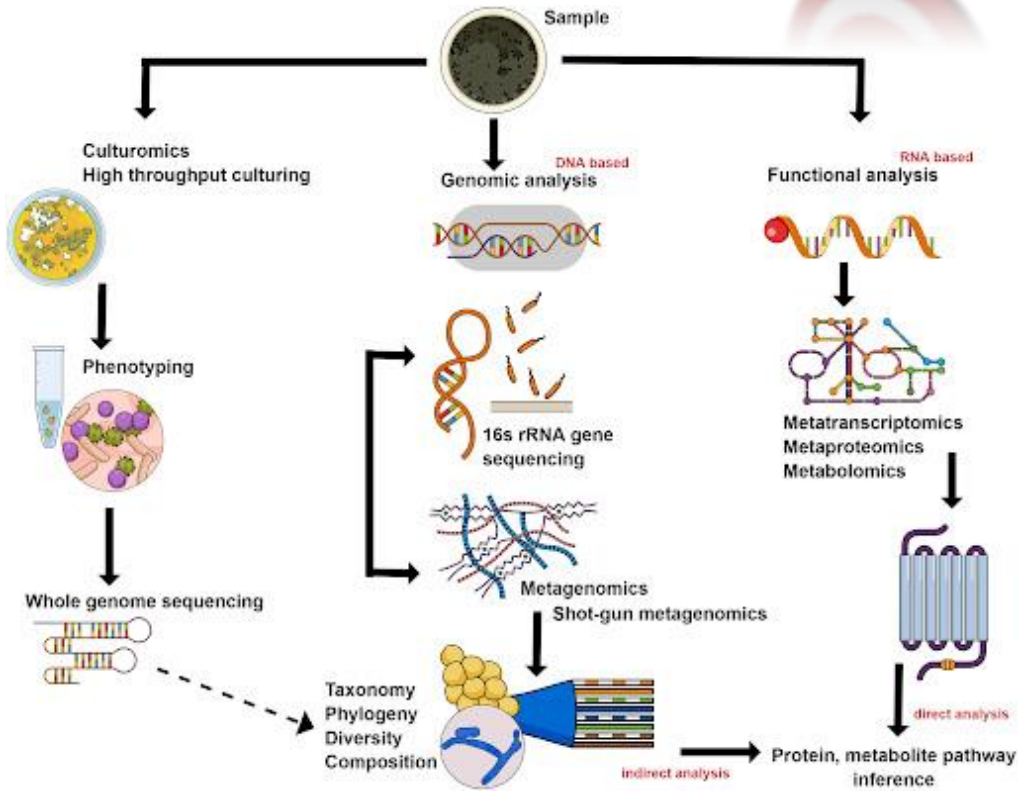


Clin Gastroenterol Hepatol. 2019 Jan;17(2): 218–230.

# Before sampling

# What is the aim?




National Academies of Sciences, Engineering, and Medicine. 2018. *Environmental Chemicals, the Human Microbiome, and Health Risk: A Research Strategy*. Washington, DC: The National Academies Press. https://doi.org/10.17226/24960.

# Study design

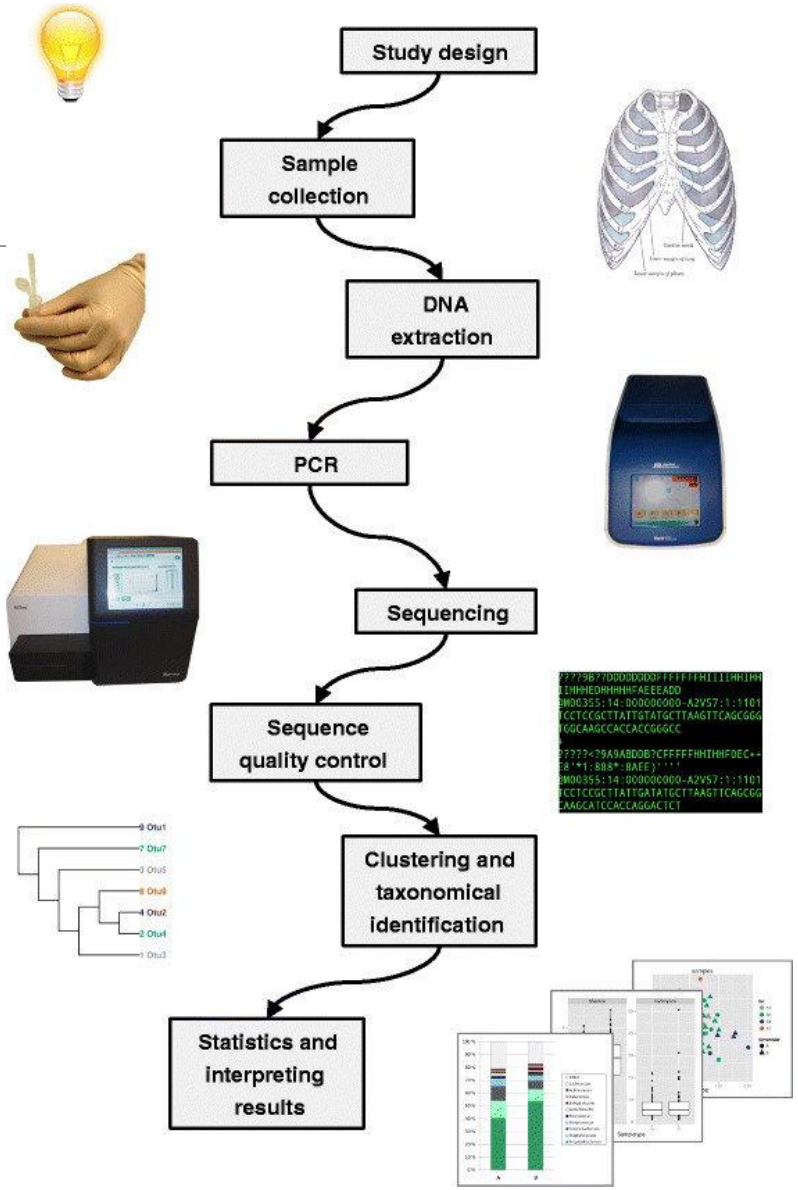
High variability of microbiome

Specific inclusion/exclusion criteria
- More homogenous group
- Increased study power

Define «healthy controls», if needed
- 18-64 years
- No chronic autoimmune, gastrointestinal, oncological diseases

Collection of metadata → information on factors that could impact the results
- Antibiotic treatment
- Other diseases
- Diet

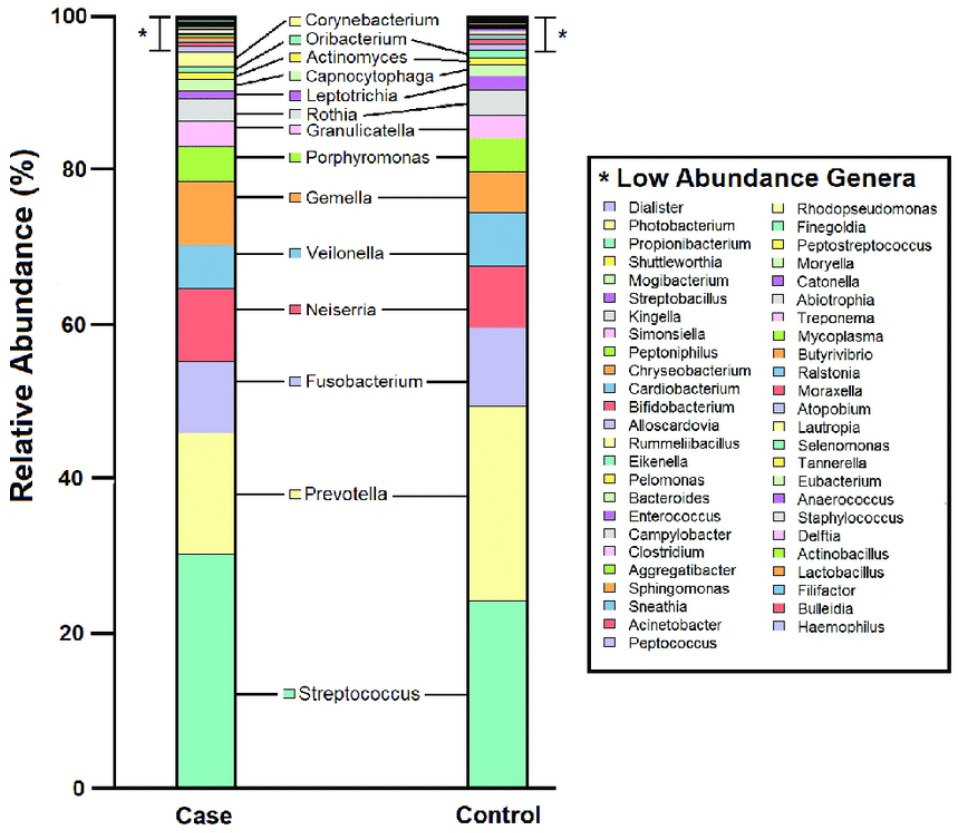
Other factors
- e.g., Skin microbiome needs ~2hours to rebalance after washing hands



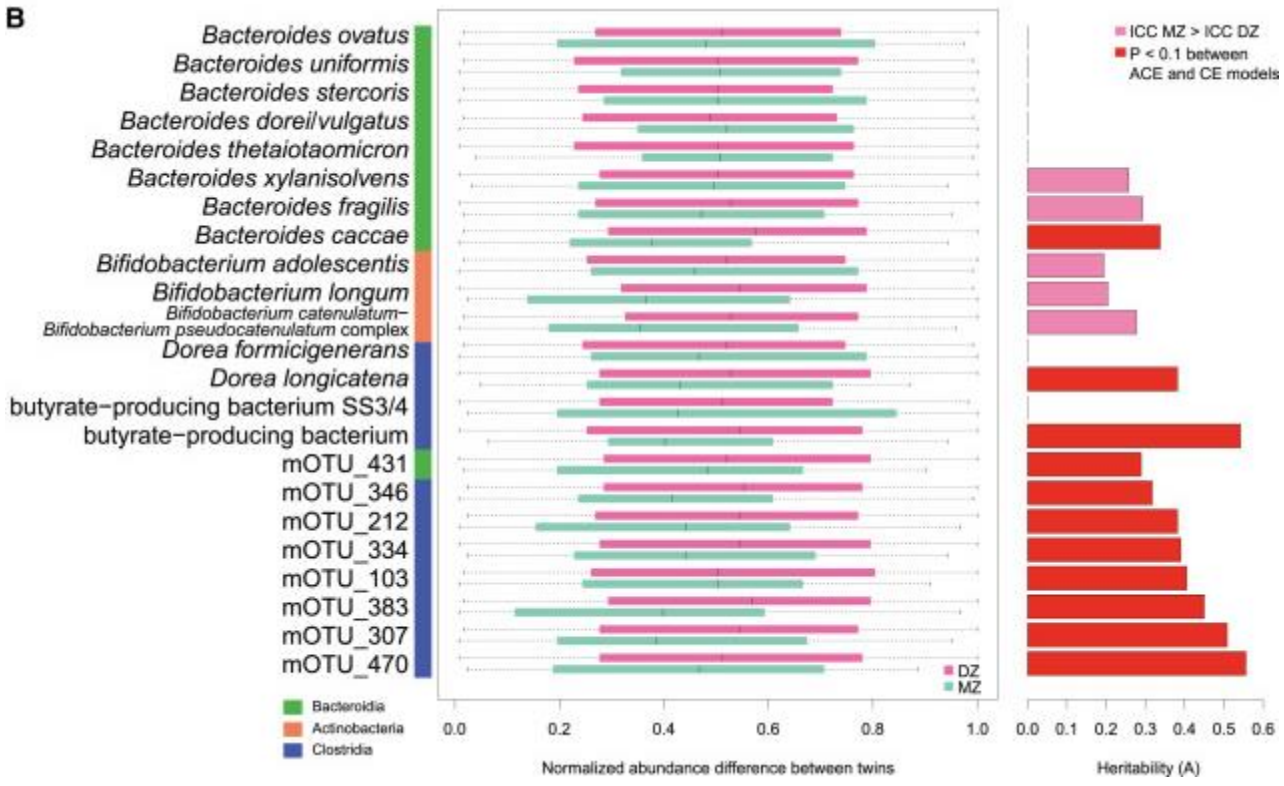

DOI: 10.1186/s40413-015-0074-z

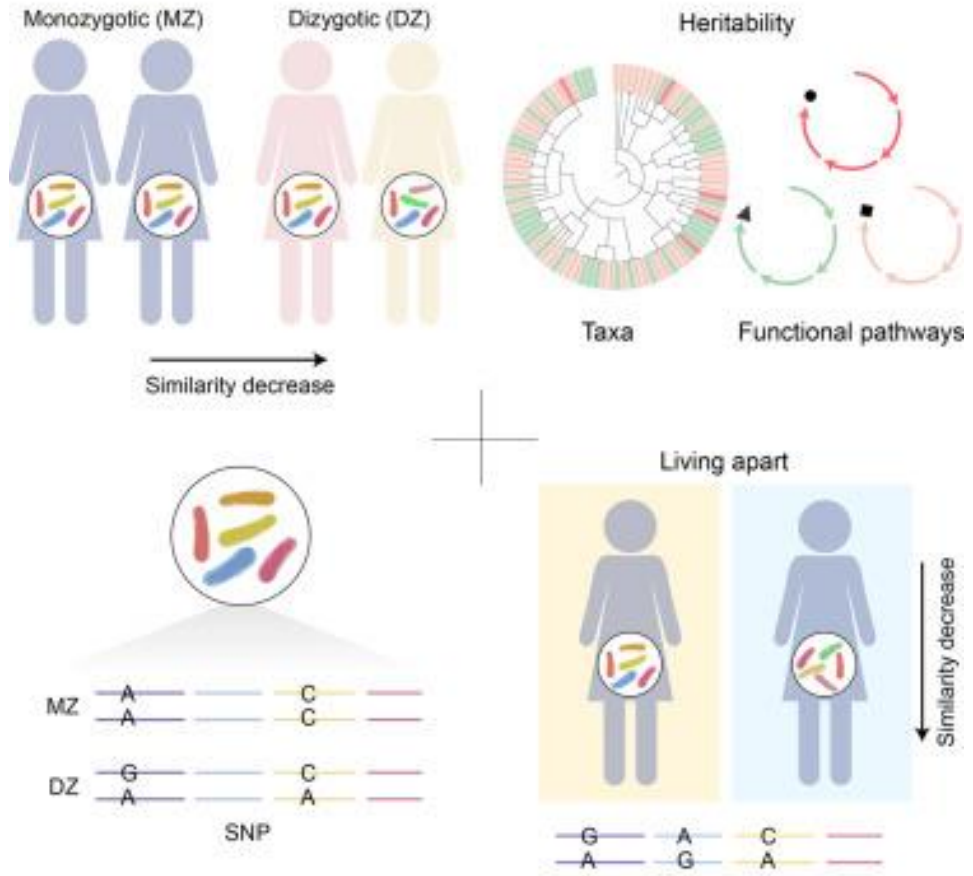
# Design – case/control

- ✓**Match study groups**
  - Age
  - Sex
  - BMI
- ✓**High intra- & inter- individual variety**
  - Lifestyle
  - Diet
  - Medication
  - Physiology
  - Geographic location
- ✓**>1 control group**
  - Healthy individuals
  - Asymptomatic patients
  - Different stages of disease/therapy response
- ✓**Time of sample collection**



# The perfect control group = monozygotic twins






VOLUME 3, ISSUE 6, P572-584.E3, DECEMBER 21, 2016

# Design – longitudinal

**???**

- ✓How many time points
- ✓Observational / with intervention
- ✓Seasonal impact; Sample collection time

**+**

- ✓First sample as control
- ✓Increased study power

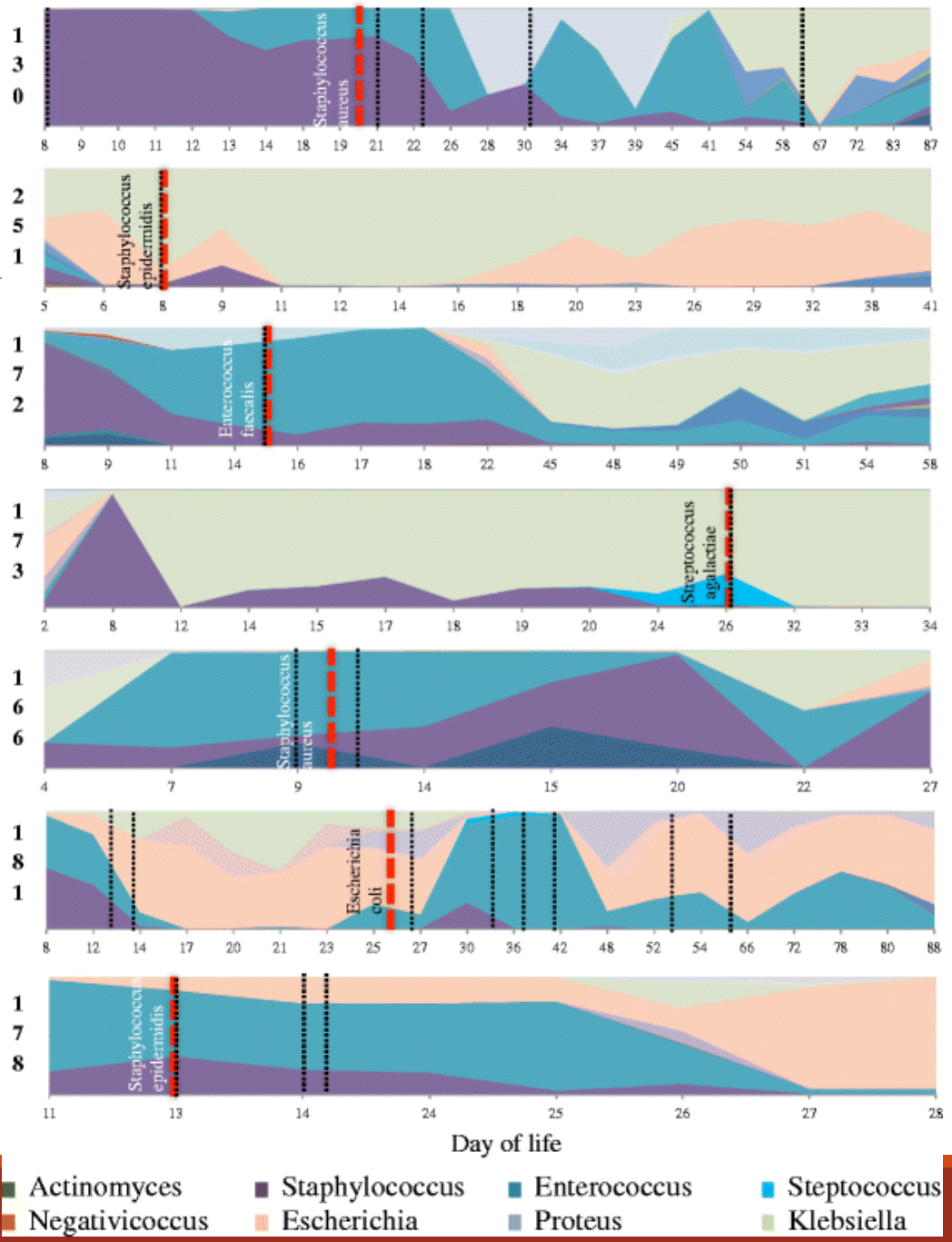
**-**

- ✓Adherence.

Published in final edited form as:
*Cell*. 2014 July 17; 158(2): 250–262. doi:10.1016/j.cell.2014.06.037.

**Conducting a Microbiome Study**

**Julia K. Goodrich**[1,2], **Sara C. Di Rienzi**[1,2], **Angela C. Poole**[1,2], **Omry Koren**[1,2,9], **William A. Walters**[3], **J. Gregory Caporaso**[4,5], **Rob Knight**[6,7,8], and **Ruth E. Ley**[1,2,*]

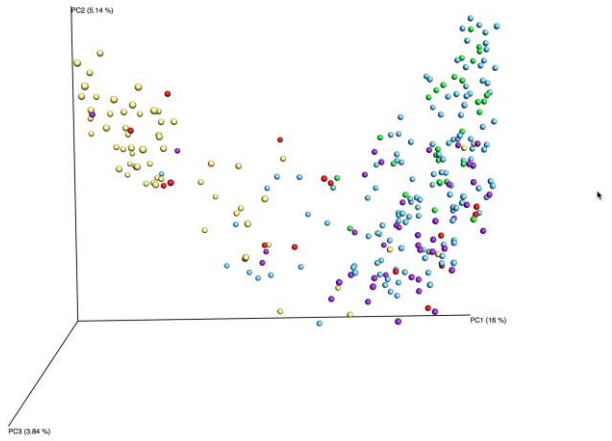


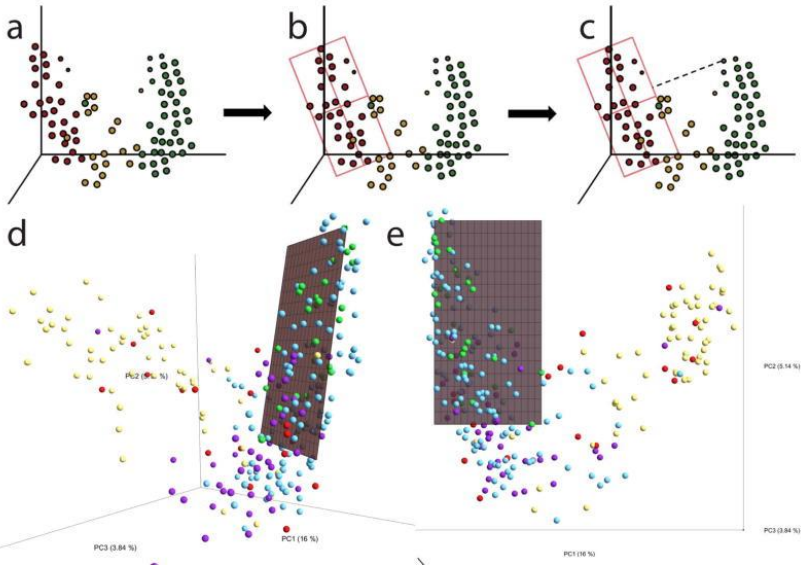

Stewart, C. J. *et al. Longitudinal development of the gut microbiome and metabolome in preterm neonates with late onset sepsis and healthy controls. Microbiome* ***5, 75, doi:10.1186/s40168-017-0295-1 (2017).***

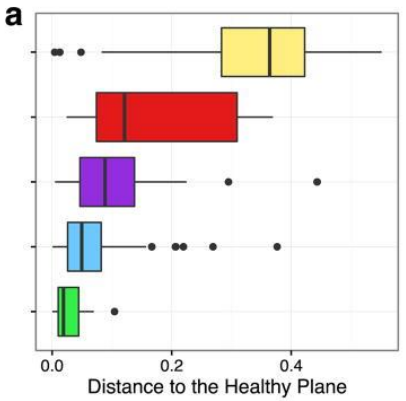
# Defining a healthy plane

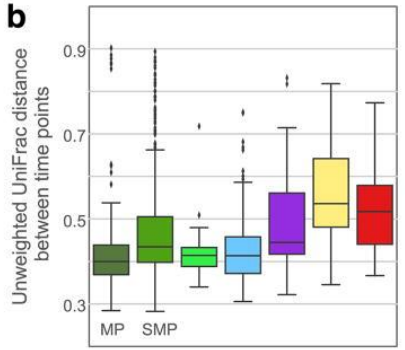
## Microbiome dynamics in IBD

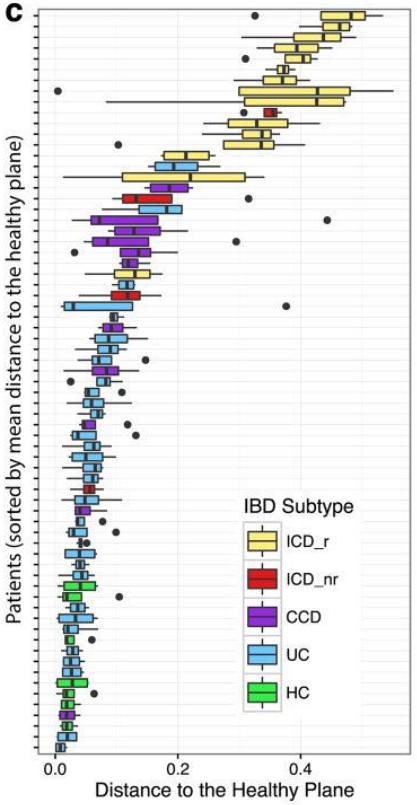
- Microbiome of IDB patients is fluctuating more
- Some IBD patient samples periodically where on the «healthy plane»







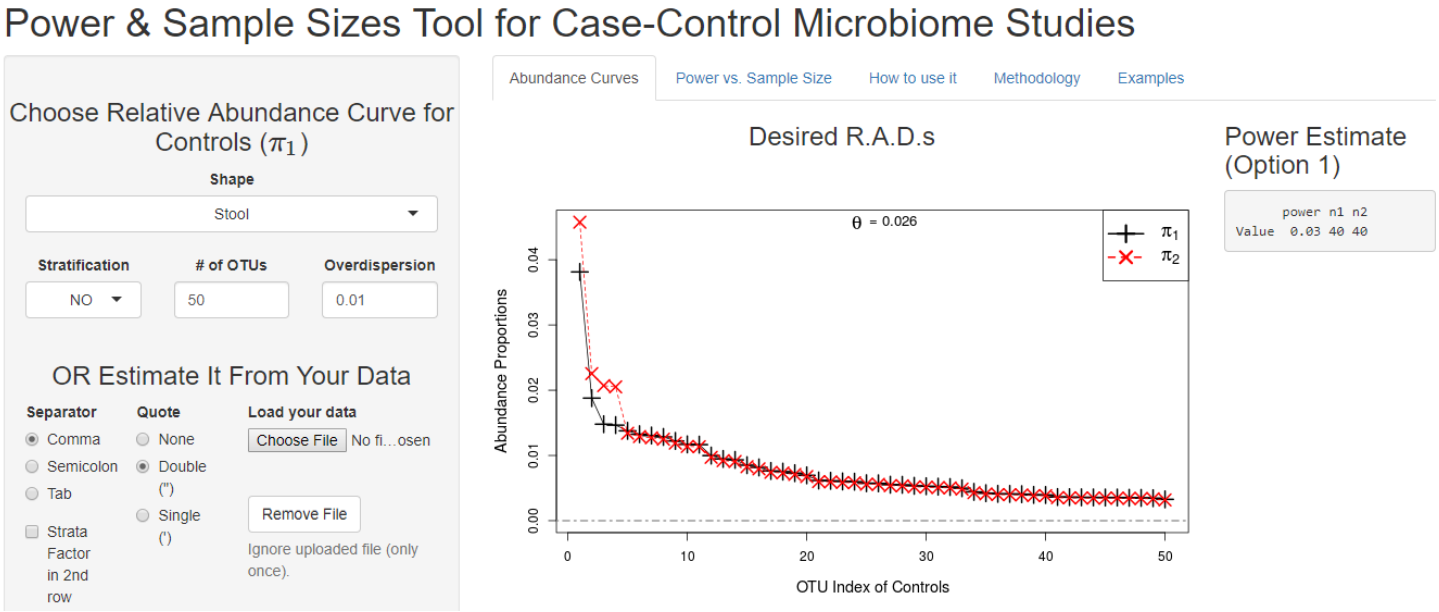

Nat Microbiol. 2017 Feb 13; 2: 17004.

# Study power calculations

- ✓Current methods based on:
  - PERMANOVA
  - Dirichlet Multinomial
  - Random forest analysis

- ✓Some tools for case-control studies

https://fedematt.shinyapps.io/shinyMB/



# Designing a microbiome study

REVIEW ARTICLES

## A guide to human microbiome research: study design, sample collection, and bioinformatics analysis

Qian, Xu-Bo[1]; Chen, Tong[2]; Xu, Yi-Ping[1]; Chen, Lei[3]; Sun, Fu-Xiang[4]; Lu, Mei-Ping[1]; Liu, Yong-Xin[5,6]

Editor(s): Shi, Qiang **Author Information**

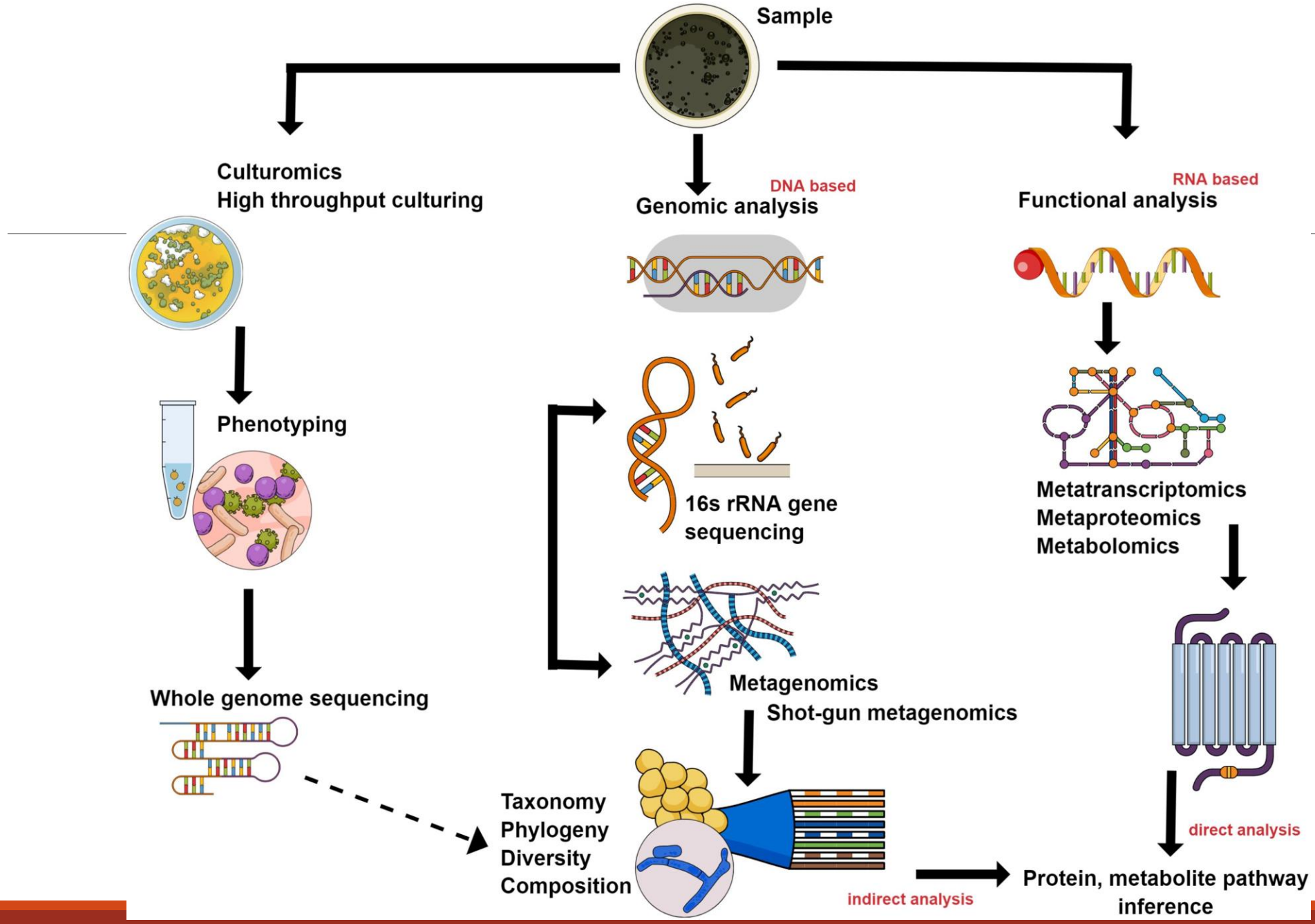
Chinese Medical Journal: August 5, 2020 - Volume 133 - Issue 15 - p 1844-1855
doi: 10.1097/CM9.0000000000000871

| Considerations | Details |
|---|---|
| Study type | ☐Cross-sectional ☐Case-control ☐Longitudinal ☐RCT ☐Other: |
| Sex | ☐Matched ☐Unmatched ☐Other: |
| Age | ☐Matched ☐Unmatched ☐Other: |
| BMI | ☐Matched ☐Unmatched ☐Other: |
| Ethnicity | ☐Matched ☐Unmatched ☐Other: |
| Geographic location | ☐Matched ☐Unmatched ☐Other: |
| Diet | ☐Monitored: detailed information<br>☐Not monitored |
| Season factor | ☐All samples in different groups are collected in the same season(s)<br>☐All samples in different groups are not collected in the same season(s) |
| Medications | What kinds of medications were used before the study?<br>How long were the medications not used before the study? |
| Inclusion criteria | ☐Defined well ☐Not defined well |
| Exclusion criteria | ☐Defined well ☐Not defined well |
| Sample size | ☐Estimated ☐Not estimated |
| Sequencing methods | ☐Amplicon ☐Metagenome |
| Negative and/or positive controls | ☐Negative controls: detailed information<br>☐Positive controls: detailed information |
| Multi-omics methods | ☐Metabolome ☐Metatranscriptome ☐Metaproteome |
| Sample types | ☐Fecal sample ☐Colonic lavage fluid ☐Luminal brush ☐Pinch biopsy<br>☐Sub-mucosal biopsy ☐Synovial fluid ☐Urinary sample ☐Dental plaque<br>☐Saliva ☐Skin ☐Other samples: |
| Animal model | ☐Results will be verified in an animal model<br>☐Results will not be verified in an animal model |

BMI: Body mass index; RCT: Randomized controlled trial.

# Methods

Overview of the most often used approaches for microbiome analysis






*Journal of Clinical and Translational Hepatology* 2019; **DOI:** 10.14218/JCTH.2019.00035

Bacteria, archaea, fungi, viruses
Small amount of data

# Culture-dependent methods

Traditional approach for characterizing bacteria, as well as some archaea and viruses

Allows to perform functional *in vitro* experiments

Basis for development of probiotics

Possible to combine with other - DNA and RNA based - methods

Complicated due to:

- Anaerobic microbiota
- Interaction between taxa

Information is limited to corresponding knowledge and skills

Relatively low % of microorganisms for which the culturing conditions are known

All microorganisms

Small amount of data

# PCR-based arrays

One or some number of specific microorganisms can be included

Fast; suitable for diagnostics, forensics

Data limited to the array

PCR artefacts

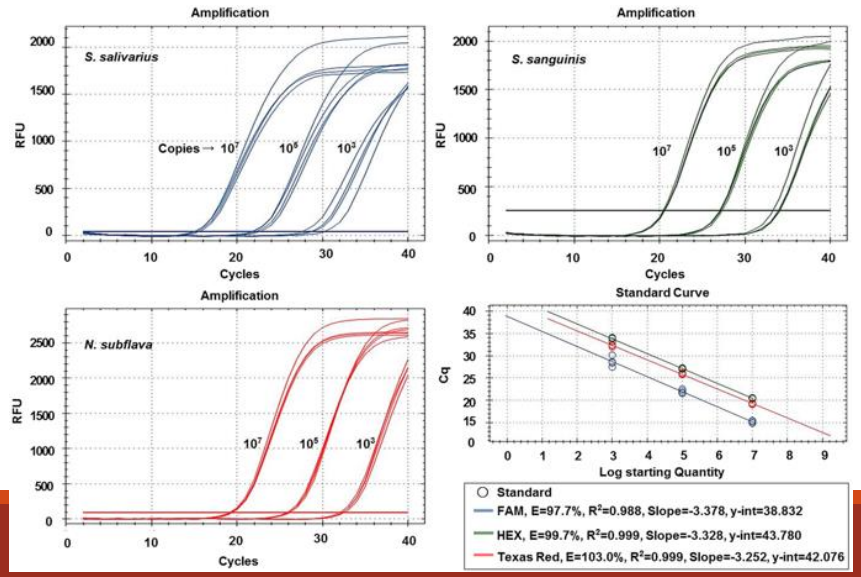
SCIENTIFIC REPORTS

OPEN

## Rapid oral bacteria detection based on real-time PCR for the forensic identification of saliva

JuYeon Jung[1], Hyun Kyu Yoon[2], Sanghyun An[3], Jee Won Lee[1], Eu-Ree Ahn[1], Yeon-Ji Kim[1], Hyun-Chul Park[1], Kyungmyung Lee[1], Jung Ho Hwang[1] & Si-Keun Lim[1]

Received: 23 November 2017
Accepted: 9 July 2018
Published online: 18 July 2018



Bacteria, archaea, eukaryotes, fungi
Large amount of data

# Analysis of marker-gene amplicons

Various types – *barcodes*:

- 16S rRNA – bacteria, some archaea
- 18S rRNA - eukaryotes
- ITS - fungi

Fast & widely used method

Can obtain information about most of the present microorganisms (corresponding to marker-gene)

Low biomass requirement

Suitable for samples with high abundance of host DNA

High number of publicly available data

Number of marker-gene copies can be variable

Mostly can classify up to genus level

Resolution depends on the chosen marker-gene and its region

No information if the detected microorganisms are dead, alive, active

PCR artefacts, primer specify

False positive results in low-biomass samples

Limited data about functions (prediction tools?)

All organisms – including the host
Large amount of data

# Shotgun metagenomics

Obtain most of the available genetic information in the sample
- Dependent of sequencing depth
- Resolution up to species and strain level

Information about taxonomic and functional profile
- Gene families
- Pathways

Possible to perform *de novo assembly*

Part of the data is taken by host DNA

Incomplete information in the reference databases

For comprehensive microbiome characterization – high sequencing depth is needed

No information if the detected microorganisms are dead, alive, active

*Assembly* artefacts

All organisms – including the host
Large amount of data

# Metatranscriptomics

Can identify live microorganisms (e.g. in combination with marker-gene analysis)

Gives information about the dynamic inter-individual differences

Gives information about microbial activity, including about response reaction to intervention or other factors

Complex sample collection and storage

Expensive and complex sample processing

Host mRNA and rRNA contamination

Data can be affected by microorganisms with high transcriptional activity

Need to be combined with DNA sequencing to distinguish transcription level from specific changes in abundance of taxonomic groups

All organisms – including the host
Large amount of data

# Metabolomics and Metaproteomics

Targeted → for detection of known metabolites/proteins (e.g. bile acids)

Untargeted → for search of new compounds

Give opportunity to evaluate functional changes

Cannot be connected to specific organisms

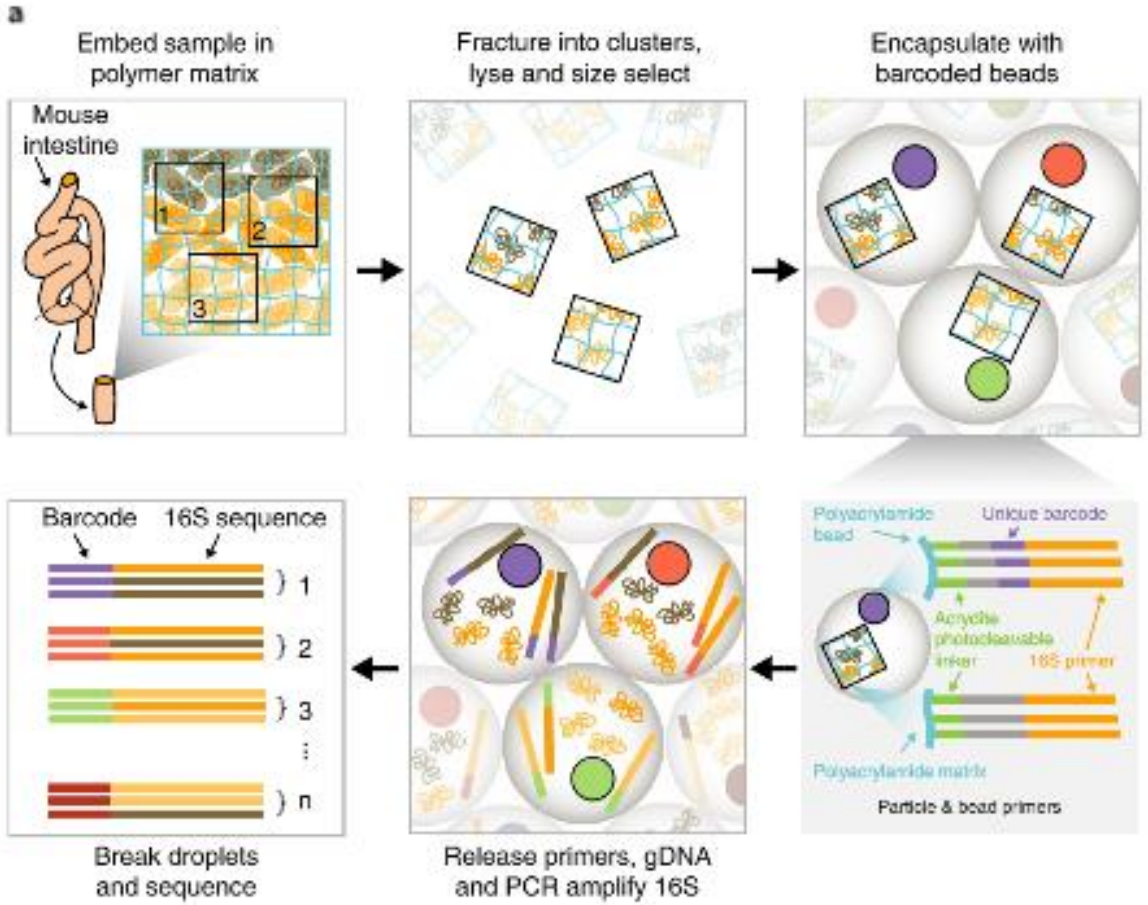
Expensive and complex

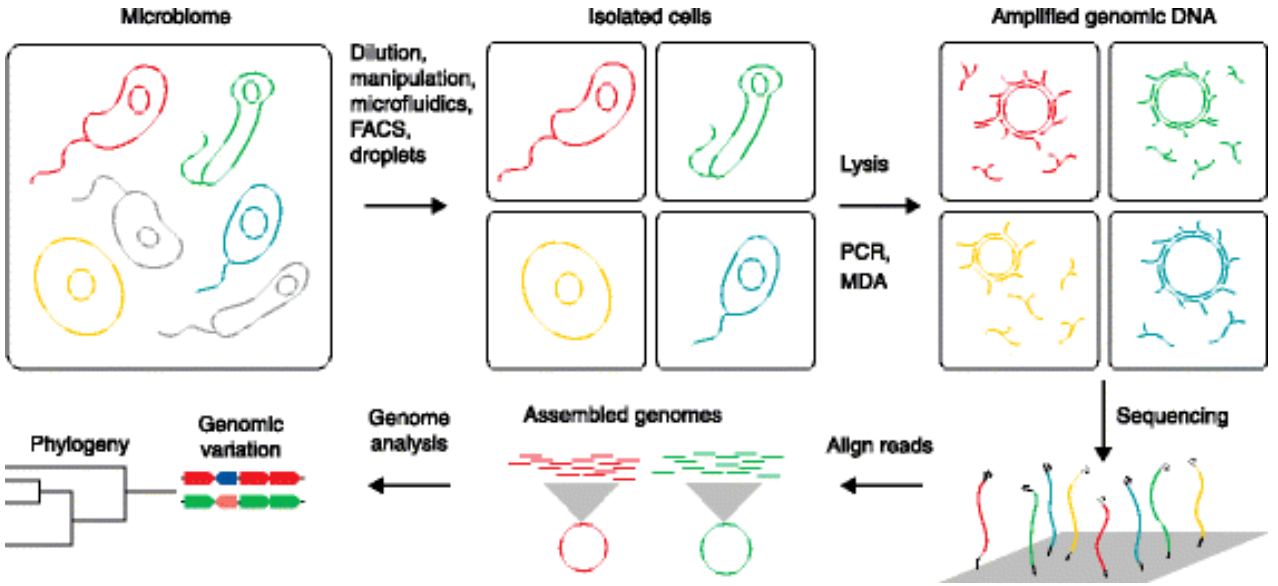
Differences in stability of study subjects

nature biotechnology
LETTERS
https://doi.org/10.1038/s41587-019-0183-2

# Spatial metagenomic characterization of microbial biogeography in the gut

Ravi U. Sheth[1,2], Mingqiang Li[3], Weiqian Jiang[3], Peter A. Sims[1,4,5], Kam W. Leong[1,3] and Harris H. Wang[1,6*]

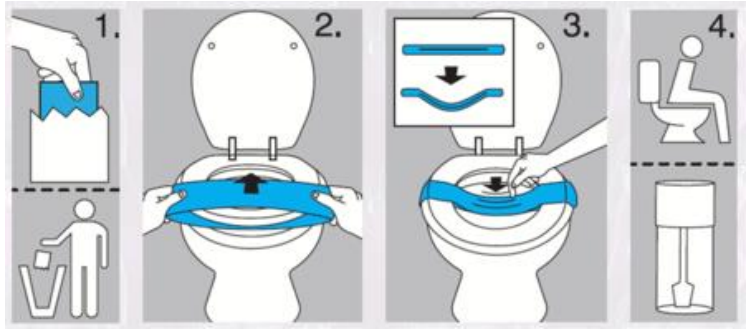


# Sampling

Factors significantly impacting the data

# Logistics




- ✓ Who will collect the sample?
- ✓ Precise instructions for collection
- ✓ Patient safety
- ✓ Assisting devices
- ✓ Transportation
- ✓ Animal studies (?)


Don't let the sample go into the toilet

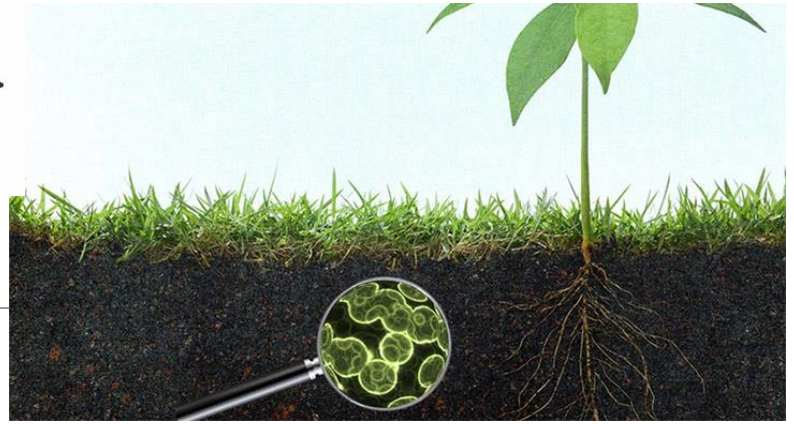

Collect stool into a clean container


Scoop a portion of the stool sample into the DNA/RNA Shield™ Fecal Collection Tube


Wash hands well

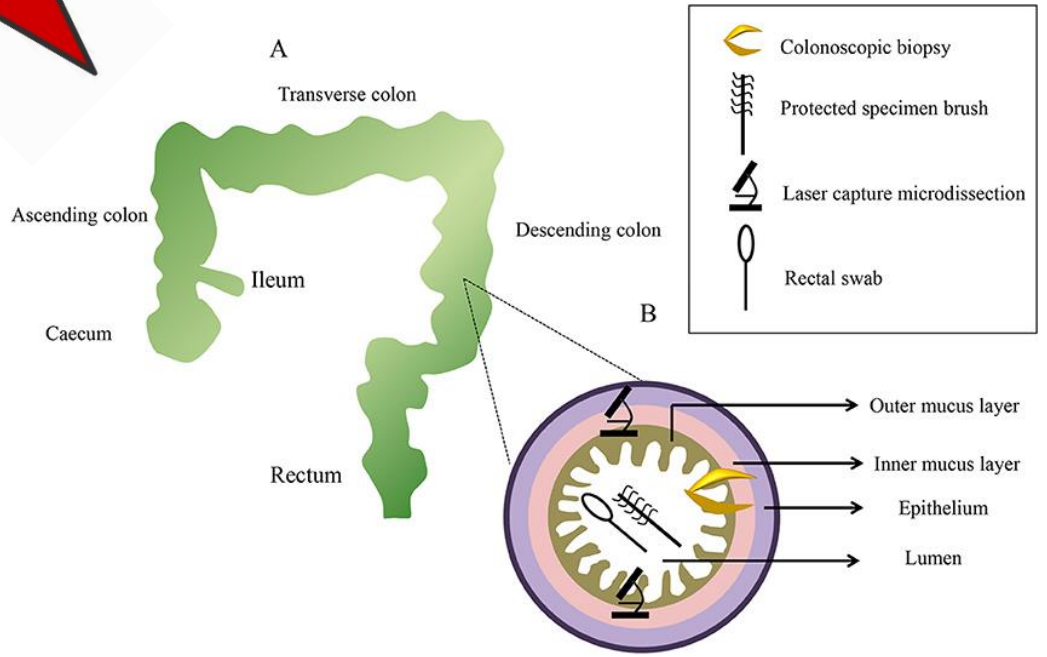

Sample collection at home

# Sample type



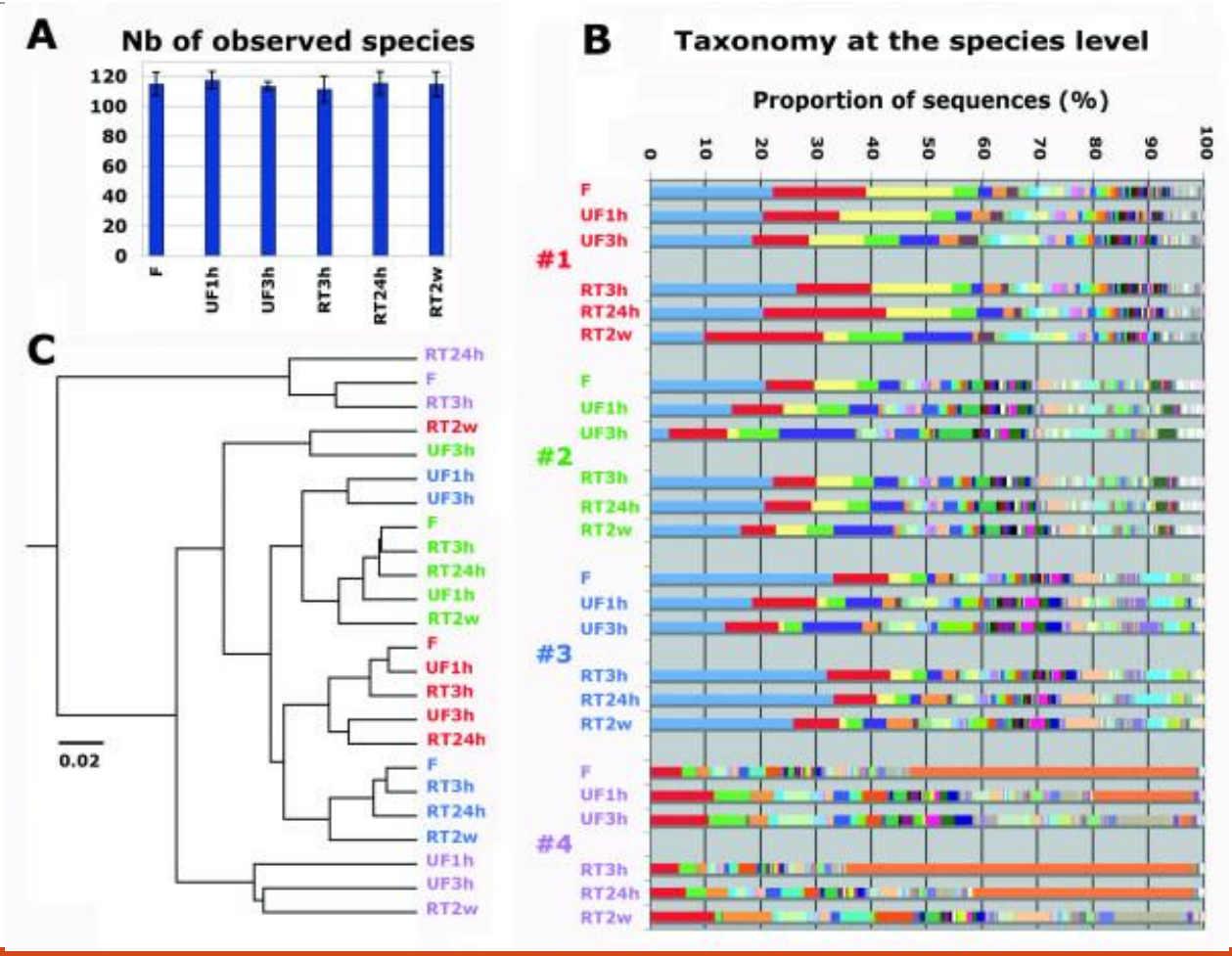
Zhang et al.Front. Cell. Infect. Microbiol., 24 February 2017





# Sample transportation and storage

- ✓Storage:
  - ✓Immediate freezing
  - ✓Long term storage at +4 C (?)
  - ✓Room temp. up to 24h
- ✓Avoid thawing/freezing cycles
- ✓Usage of stabilizers (*RNAlater, OMNIgene Gut, Eswab kit, etc.*)





Tedjo, Danyta I et al. *PloS one* vol. 10,5 e0126685. 29 May. 2015.

Cardona, Silvia et al. *BMC microbiology* vol. 12 158. 30 Jul. 2012.

PostScript
Letter

# A widely used sampling device in colorectal cancer screening programmes allows for large-scale microbiome studies

Dita Gudra[1], Saeed Shoaie[2, 3], Davids Fridmanis[1], Janis Klovins[1], Hugo Wefer[2, 4], Ivars Silamikelis[1], Raitis Peculis[1], Ineta Kalnina[1], Ilze Elbere[1], Ilze Radovica-Spalvina[1], Rolf Hultcrantz[2], Ģirts Šķenders[5, 6], Marcis Leja[5, 6], Lars Engstrand[2, 4]

PDF

AMERICAN SOCIETY FOR MICROBIOLOGY | mSystems

# Preservation Methods Differ in Fecal Microbiome Stability, Affecting Suitability for Field Studies

Se Jin Song,[a,b] Amnon Amir,[a] Jessica L. Metcalf,[a,b] Katherine R. Amato,[c] Zhenjiang Zech Xu,[a] Greg Humphrey,[a] Rob Knight[a,d]

PostScript
Letter

# High stability of faecal microbiome composition in guanidine thiocyanate solution at room temperature and robustness during colonoscopy

Yuichiro Nishimoto[1],
Shinichi Yachida[3], Tal

PDF

SCIENTIFIC REPORTS

RESEARCH ARTICLE

# The Effect of Sampling and Storage on the Fecal Microbiota Composition in Healthy and Diseased Subjects

Danyta I. Tedjo, Daisy M. A. E. Jonkers, Paul H. Savelkoul, Ad A. Masclee, Niels van Best, Marieke J. Pierik, John Penders

Published: May 29, 2015 • https://doi.org/10.1371/journal.pone.0126685

# Sample storage conditions significantly influence faecal microbiome profiles

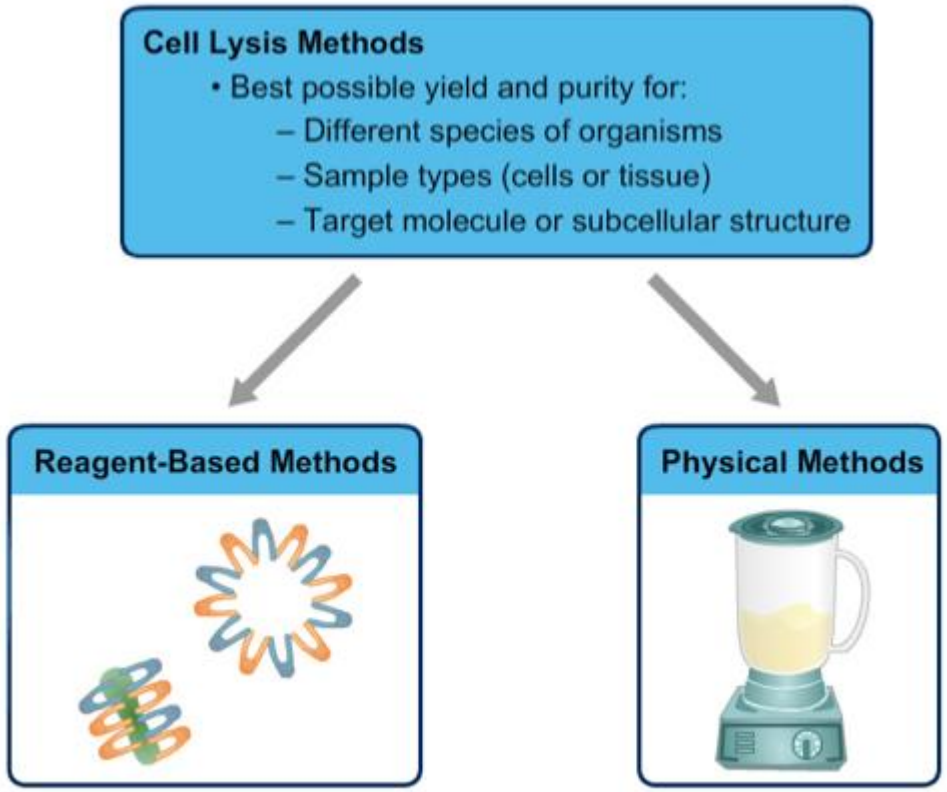
Jocelyn M Choo[1,*], Lex EX Leong[1,*] & Geraint B Rogers[1,2]

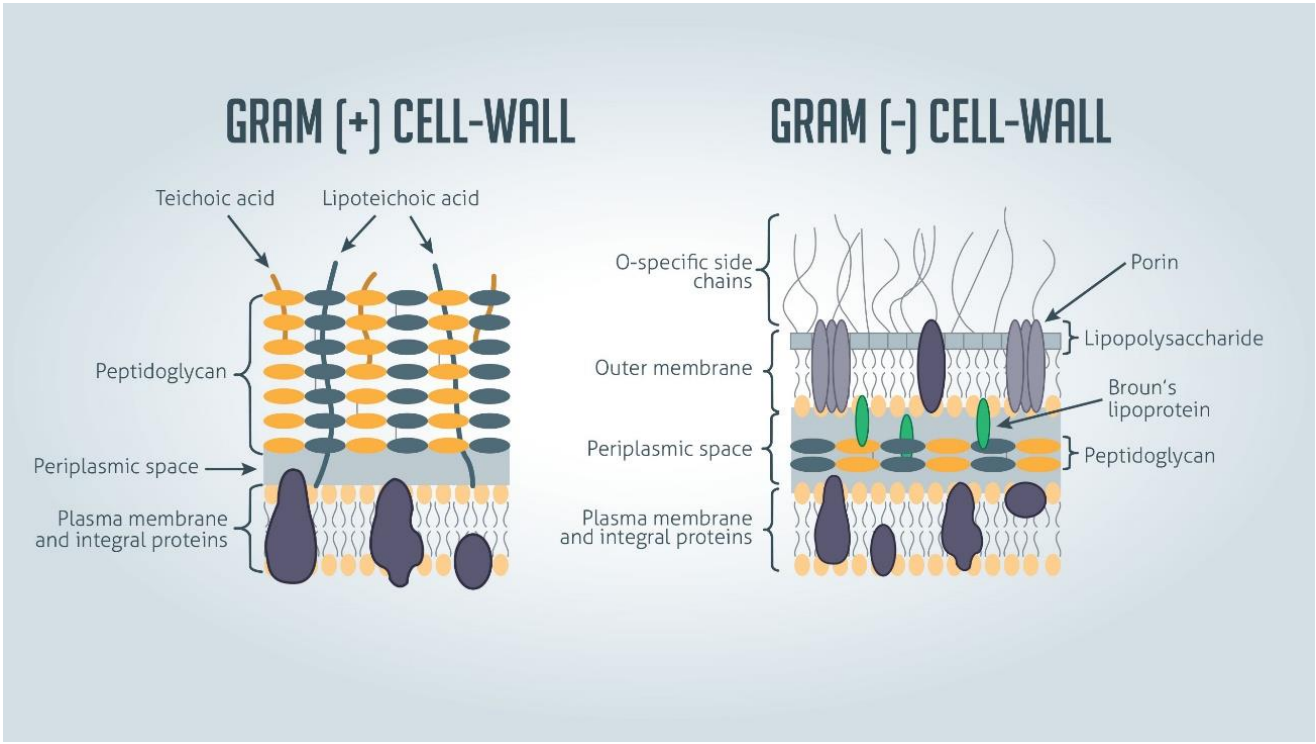
Received: 24 July 2015
Accepted: 13 October 2015
Published: 17 November 2015

# DNA extraction

Various kits available, however, the first step is the key.

# Cell degradation: Enzymatic vs Mechanistic



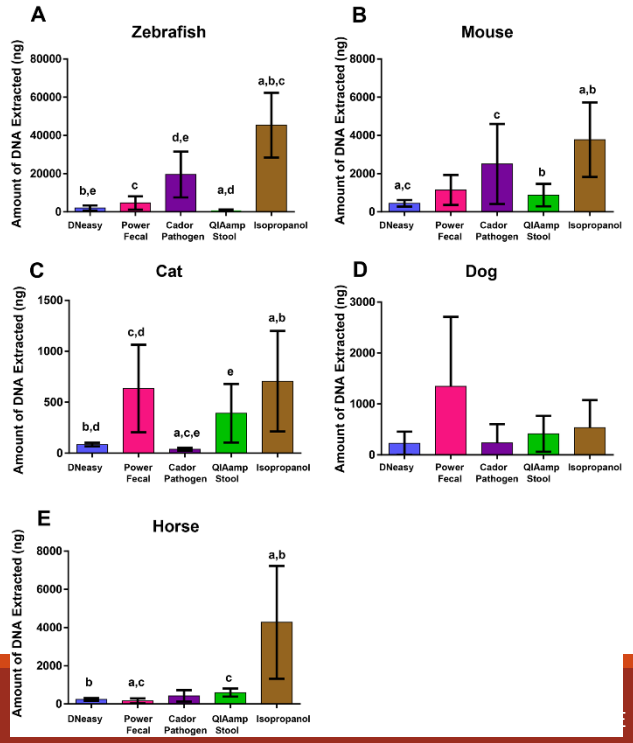

https://www.thermofisher.com

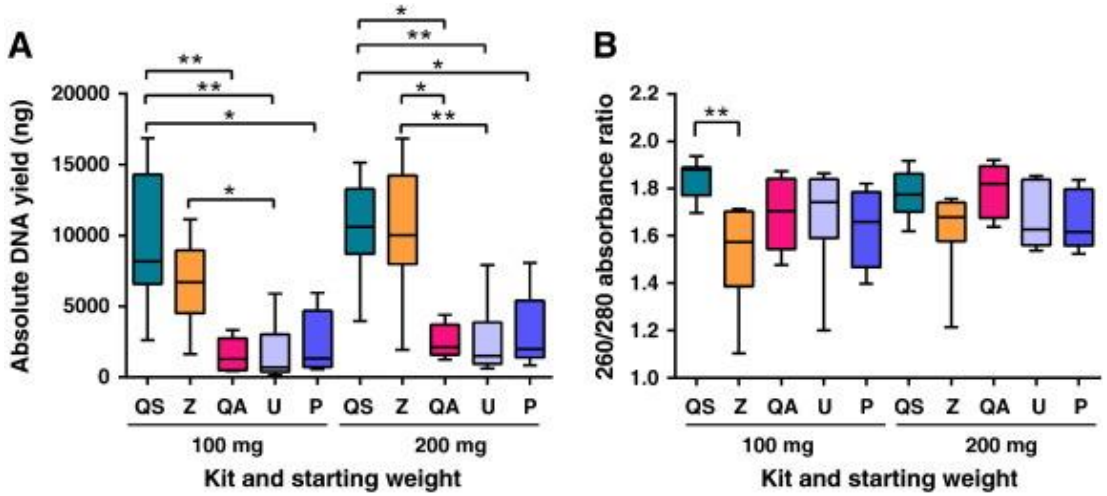
Claassen, Shantelle et al. *Journal of microbiological methods* vol. 94,2 (2013): 103-110.

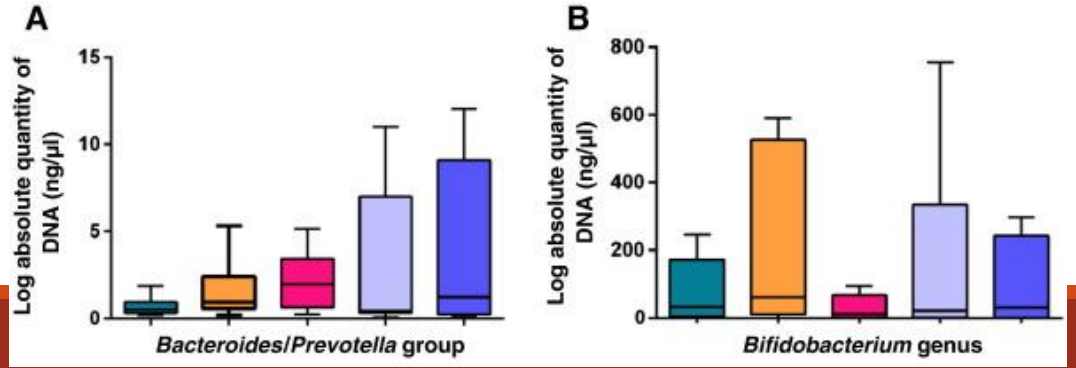
# Cell degradation: Enzymatic vs Mechanistic

Mostly – mechanistic method is used

In some cases – a combination of these methods



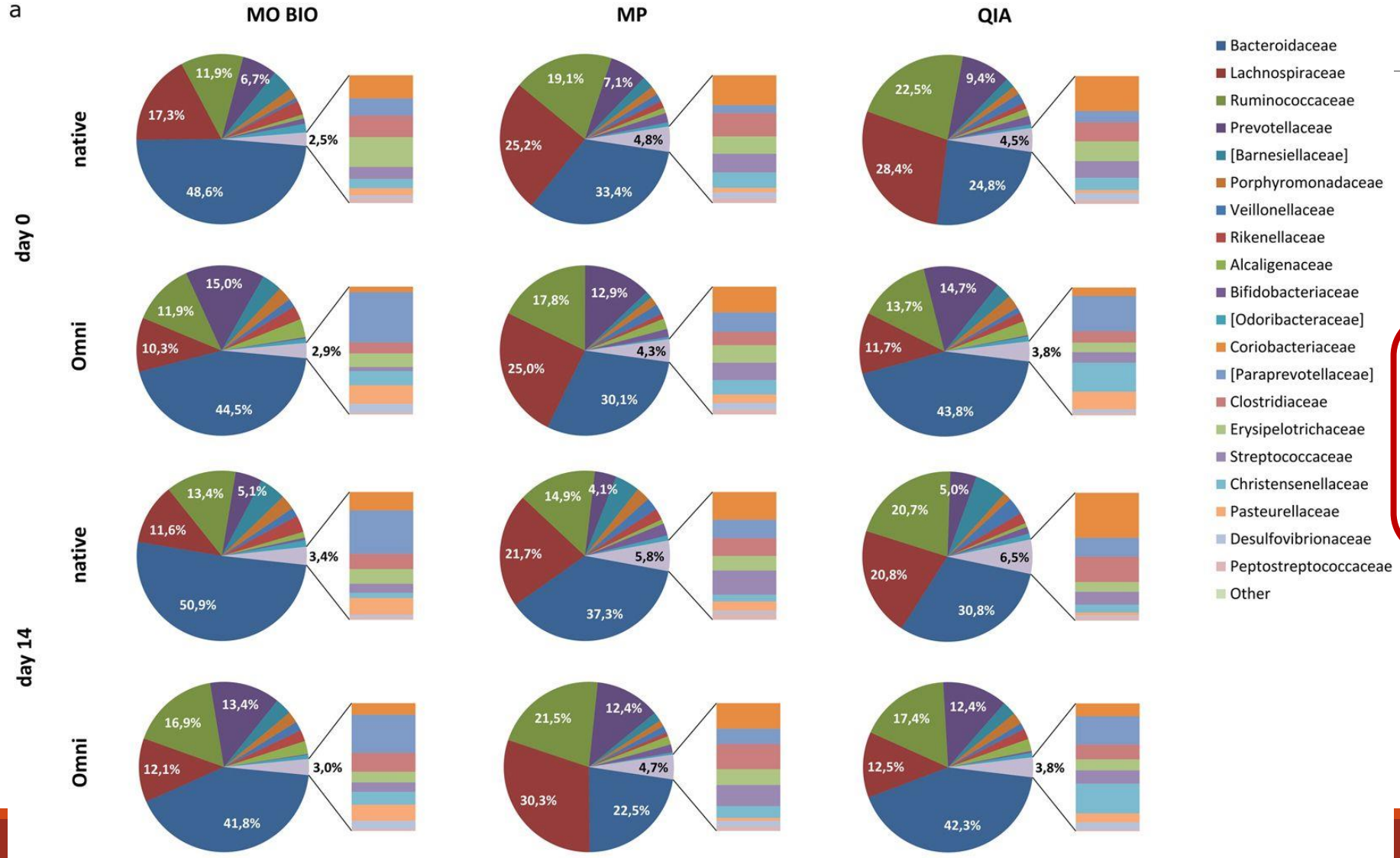




15.10.2021.

https://doi.org/10.1371/journal.pone.0143334

# Effects of sample processing on study results

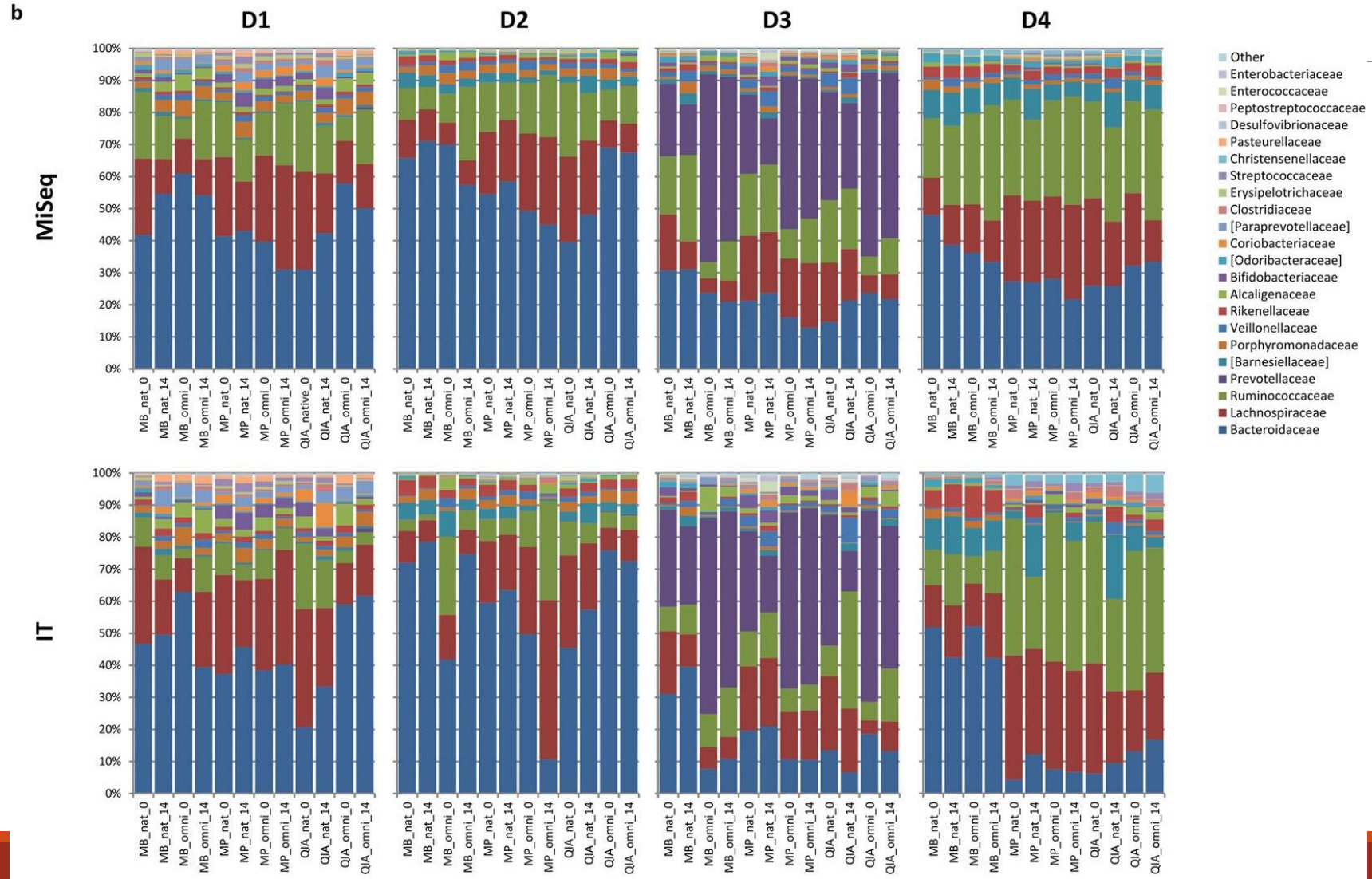



**Combination of extraction kit, preservation buffer, and time.**



*Scientific Reports* **volume 8**, Article number: 5143 (2018)

# Effects of sample processing on study results

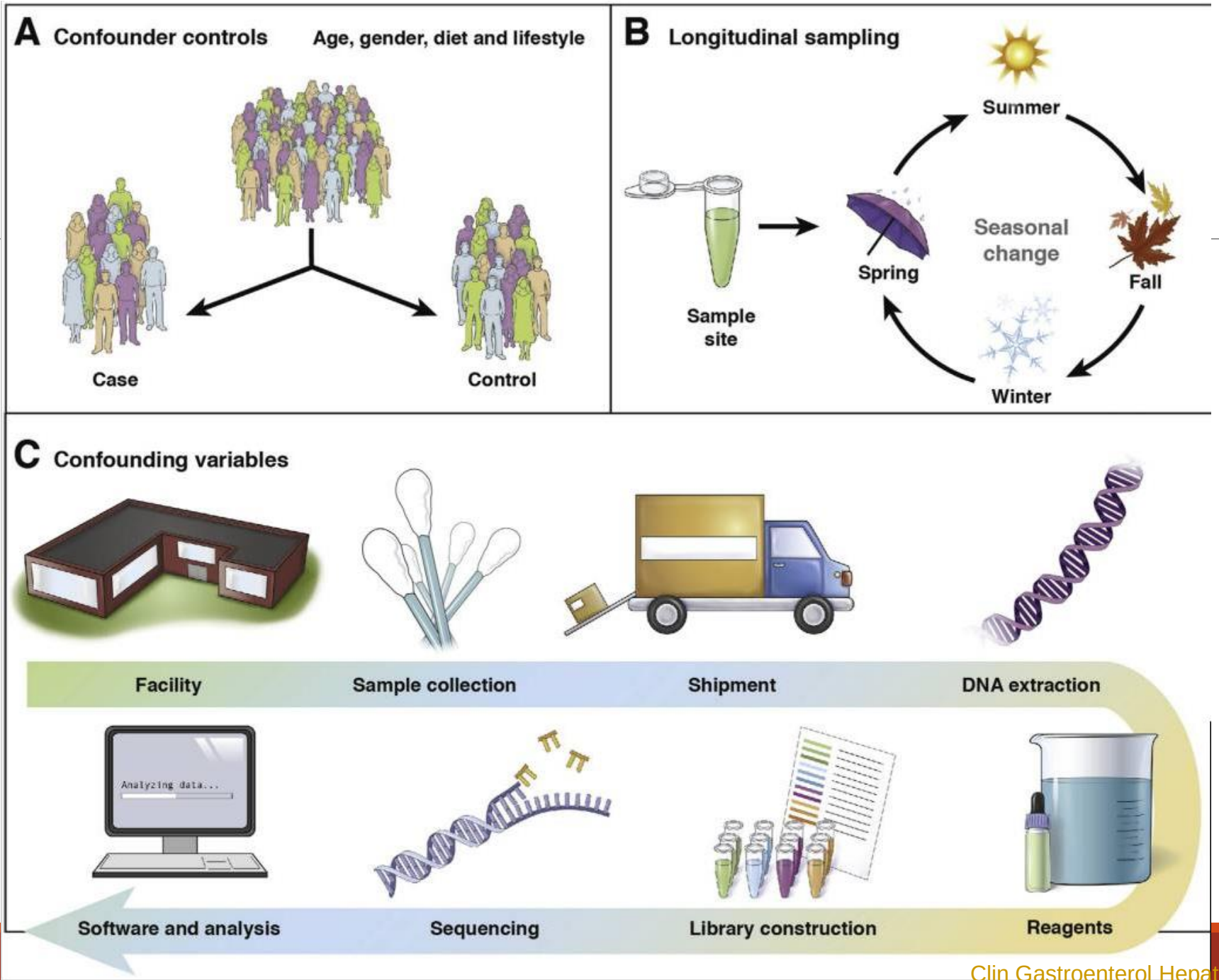



**Subject specifity can be observed independently of the workflow.**



*Scientific Reports* **volume 8**, Article number: 5143 (2018)

# DNA processing







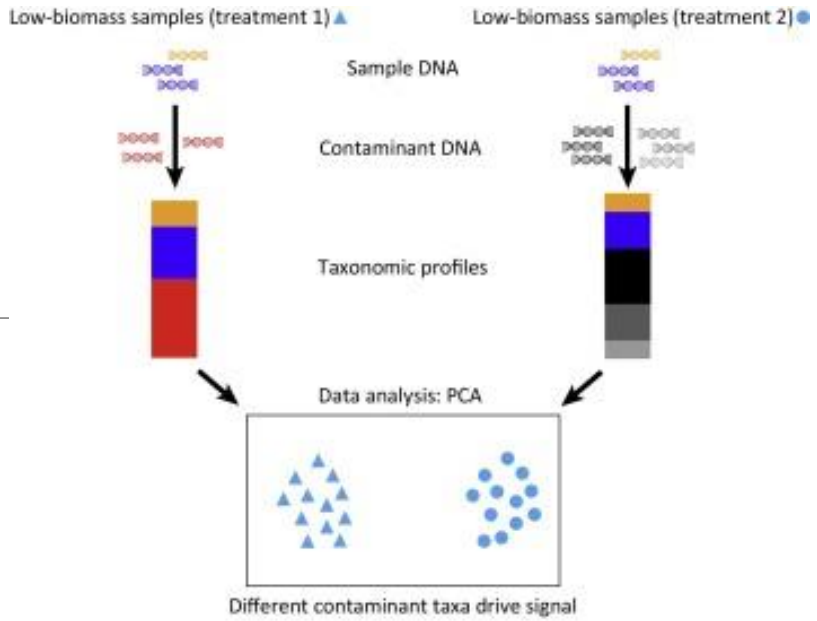

Clin Gastroenterol Hepatol. 2019 Jan;17(2): 218–230.

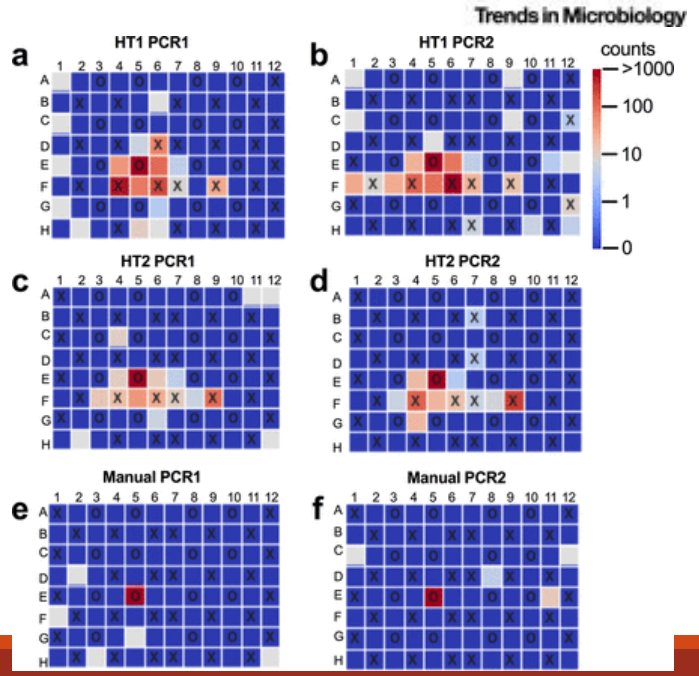
https://www.cifar.ca/cifarnews/2017/06/08/our-microbes-ourselves

# Contamination risks

- ✓Contaminant DNA/RNA or other type of sample
- ✓Cross-contamination
- ✓Environmental contamination
- ✓Reagents, e.g. polymerase

- ✓↑ sensitivity NGS techniques = ↑ contamination detection = cofound interpretation of results = controversial results?
- ✓Sequence + and – controls!



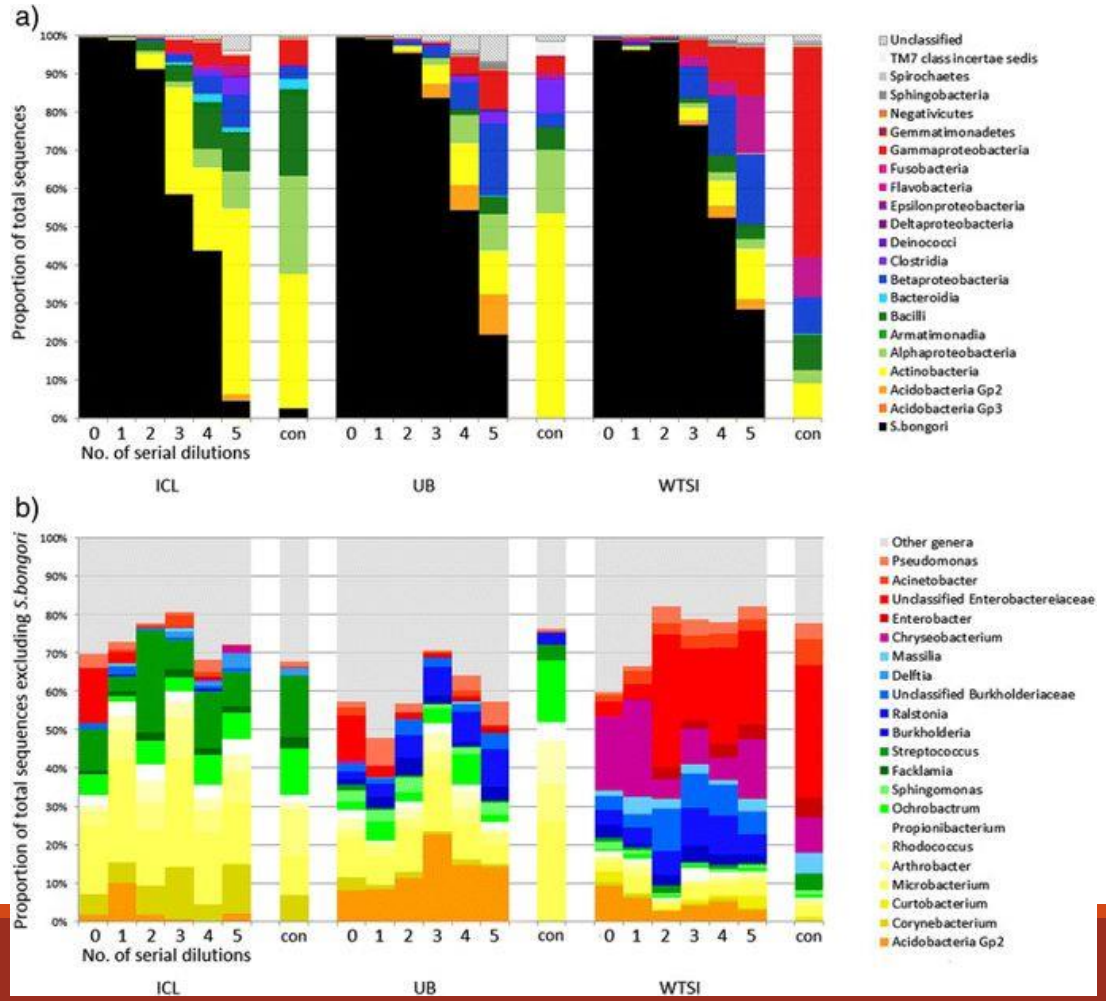



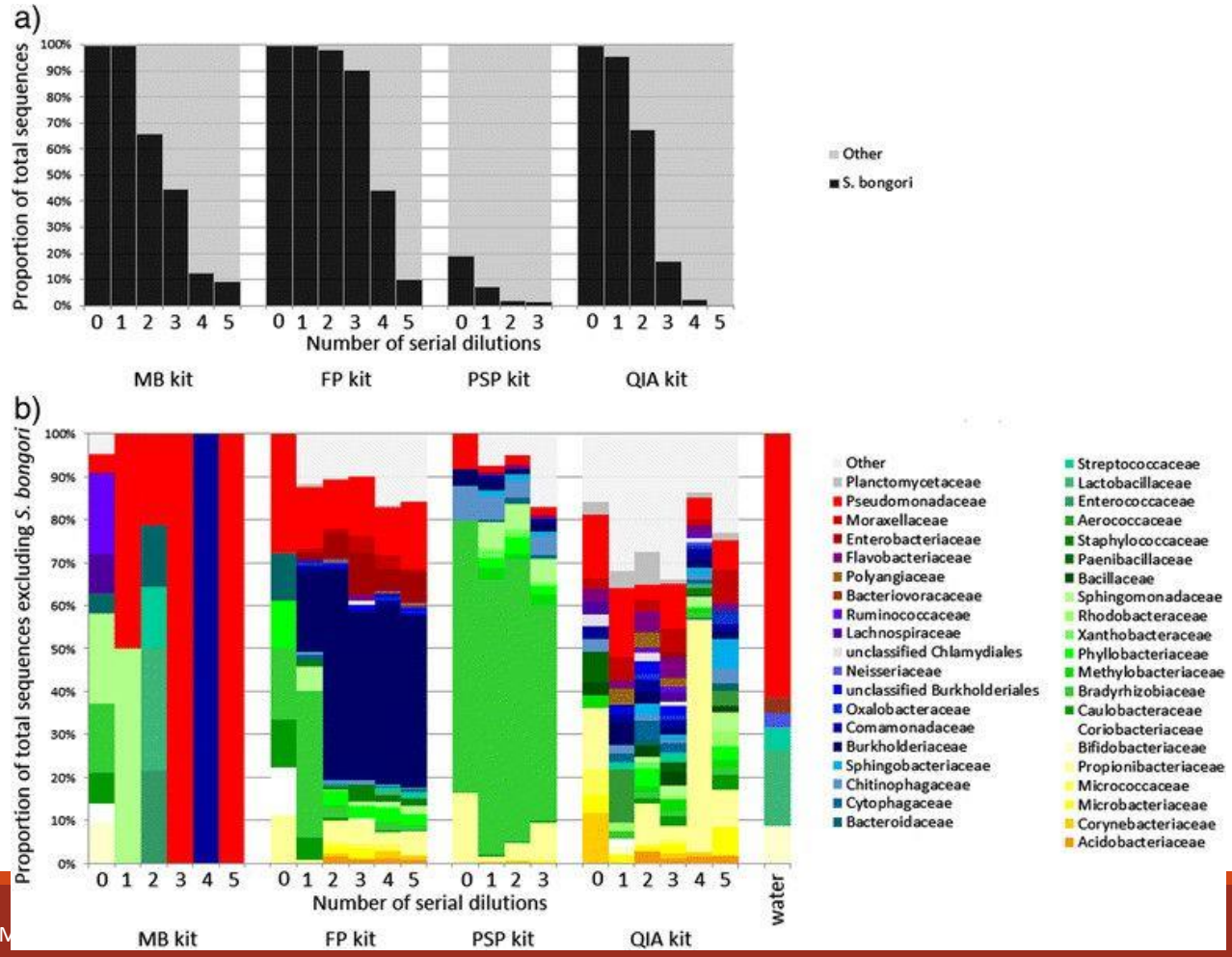
RESEARCH ARTICLE Open Access

Salter et al. BMC Biology 2014, 12:87
http://www.biomedcentral.com/1741-7007/12/87

# Reagent and laboratory contamination can critically impact sequence-based microbiome analyses

Susannah J Salter[1*], Michael J Cox[2], Elena M Turek[2], Szymon T Calus[3], William O Cookson[2], Miriam F Moffatt[2], Paul Turner[4,5], Julian Parkhill[1], Nicholas J Loman[3] and Alan W Walker[1,6*]



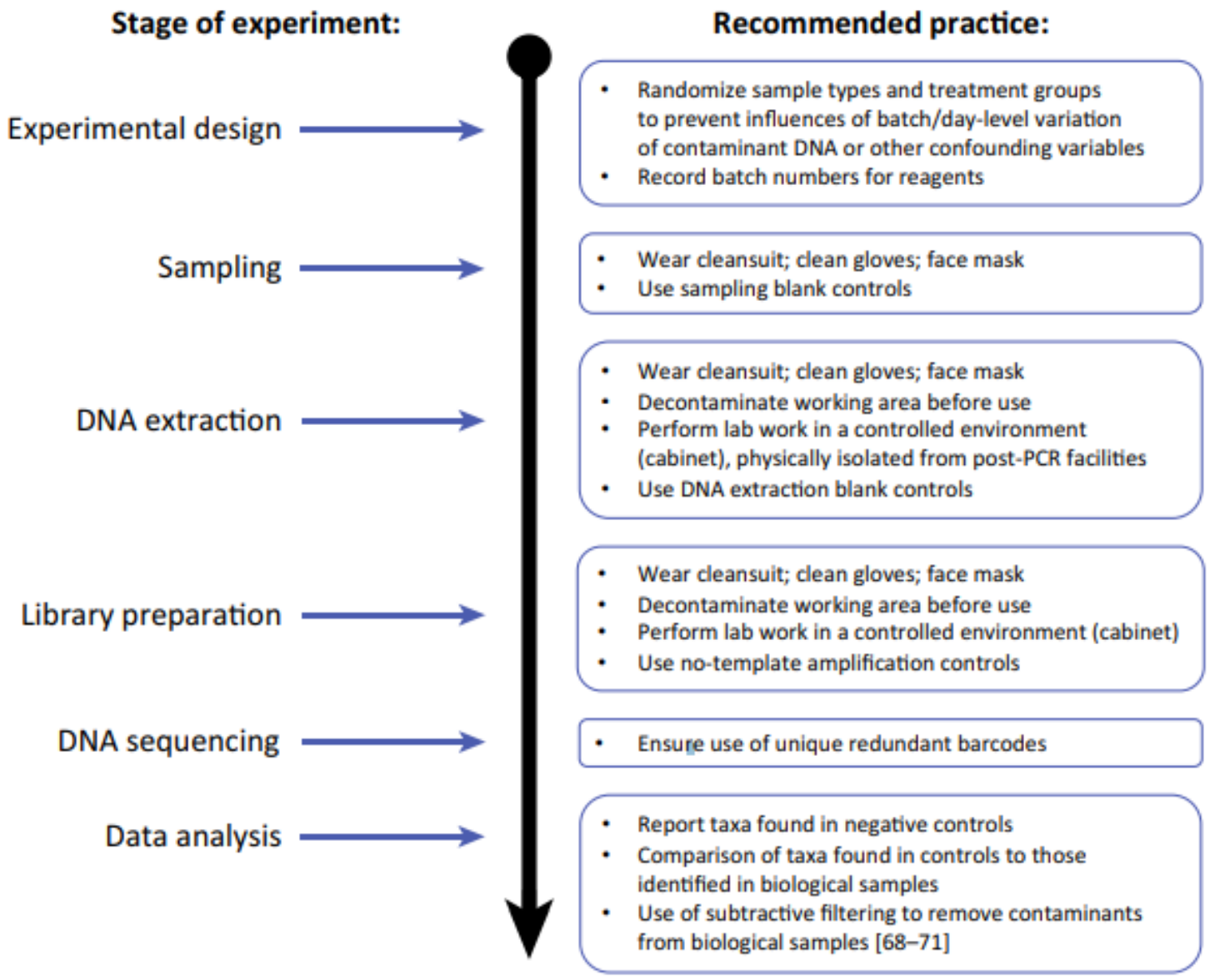



Figure 2. Flowchart of Methods To Minimize the Influence of Contaminant DNA in Low Microbial Biomass Samples. Measures to reduce experimental bias and the introduction of contaminant DNA in low microbial biomass microbiome studies.

# Thank you!

*Prof. Janis Klovins group*

ML4 MICROBIOME

cost
EUROPEAN COOPERATION
IN SCIENCE AND TECHNOLOGY

LV BMC

Latvian Biomedical
Research and Study Centre
research and education in biomedicine from genes to human