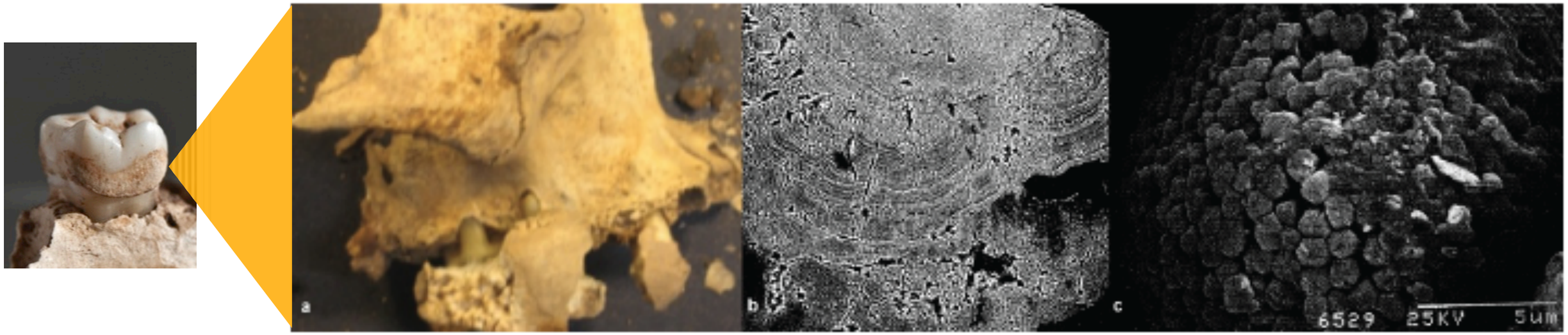
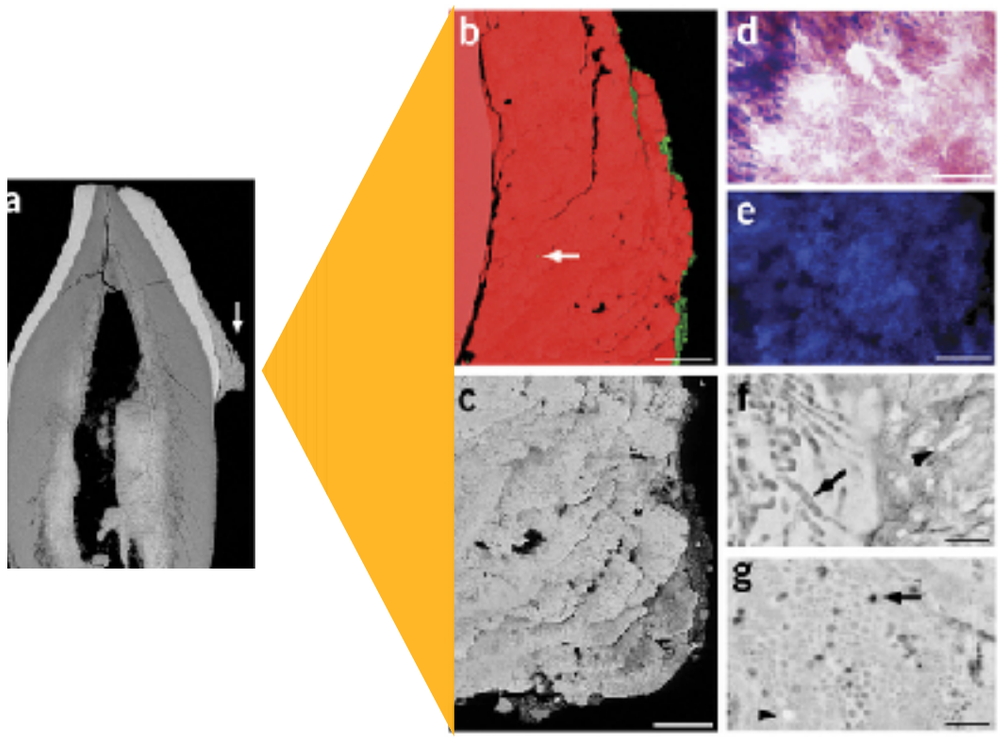
# DNA CONTAMINATION RISKS AND MITIGATION IN MICROBIOME ANALYSIS

LAURA S. WEYRICH
PENNSYLVANIA STATE UNIVERSITY, USA
UNIVERSITY OF ADELAIDE, AUSTRALIA

PHOTOGRAPH BY JOE MCNALLY, NATIONAL GEOGRAPHIC

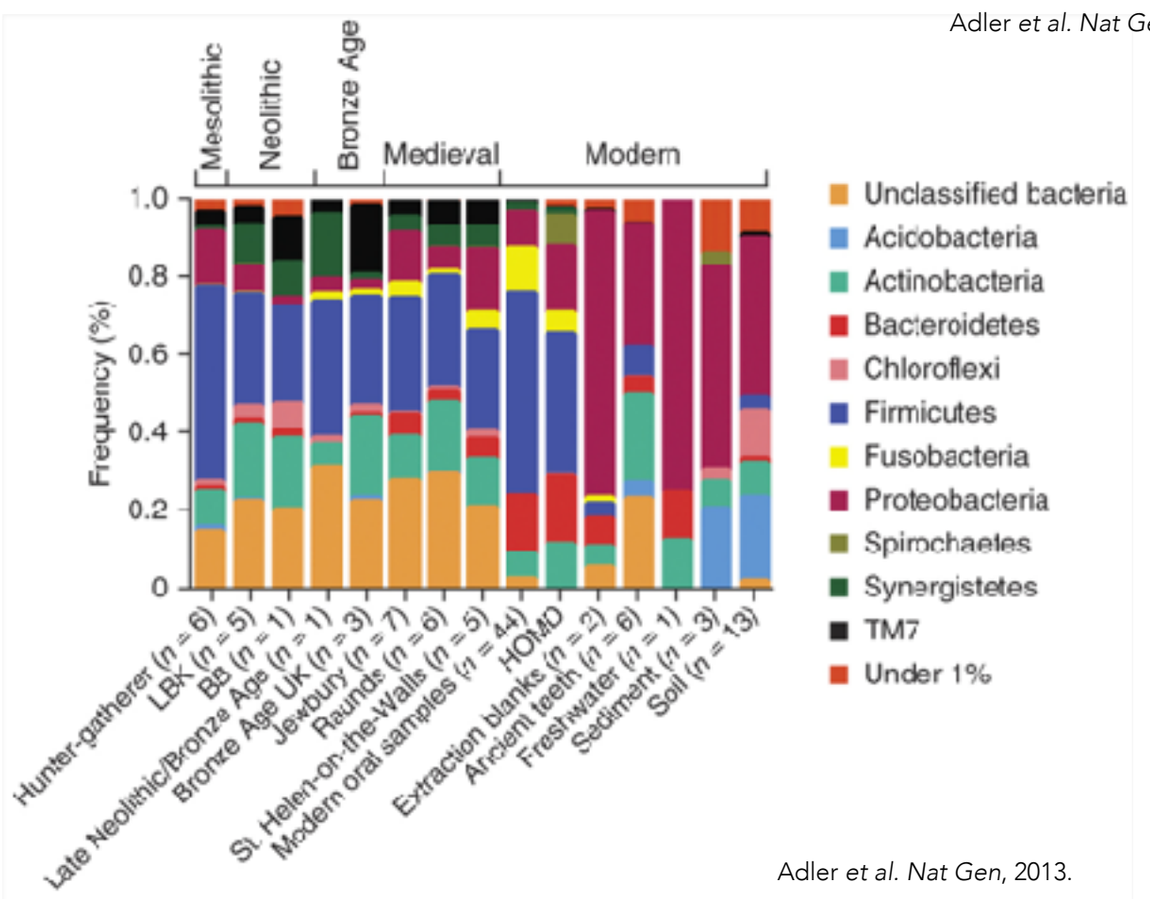
# ANCIENT DENTAL CALCULUS IS A FOSSILISED BACTERIAL RECORD




Warinner *et al. Nat Gen*, 2014.

Adler *et al. Nat Gen*, 2013.




Adler *et al. Nat Gen*, 2013.

# WHY ARE ANCIENT SAMPLES PROBLEMATIC?

**PCR**

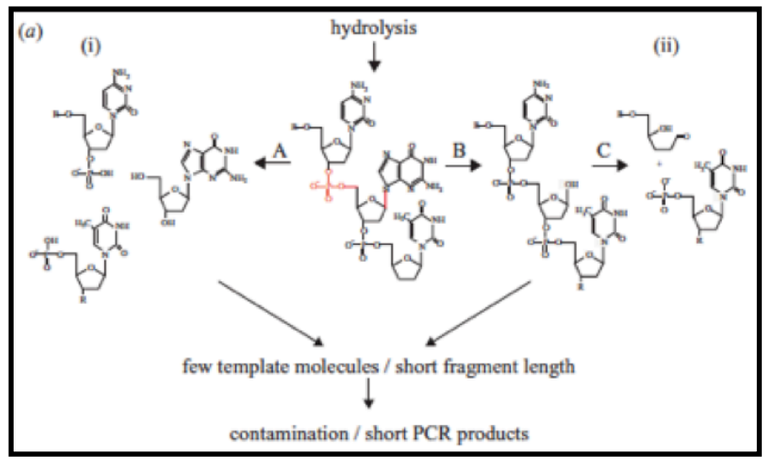
$10^6$ DNA copies per droplet (~0.005 uL)

**vs**

**Ancient Sample**

$<10^6$ DNA copies per 1 g of sample

## Low 'endogenous' DNA yield




Willerslev and Cooper. Proc Biol Sci. 2005 Jan 7; 272(1558): 3–16

## Fragmented and damaged DNA

crosslinking/hydrolysis/oxidation

## Contamination (modern or environmental)

sample collection/handling & lab or reagent contamination

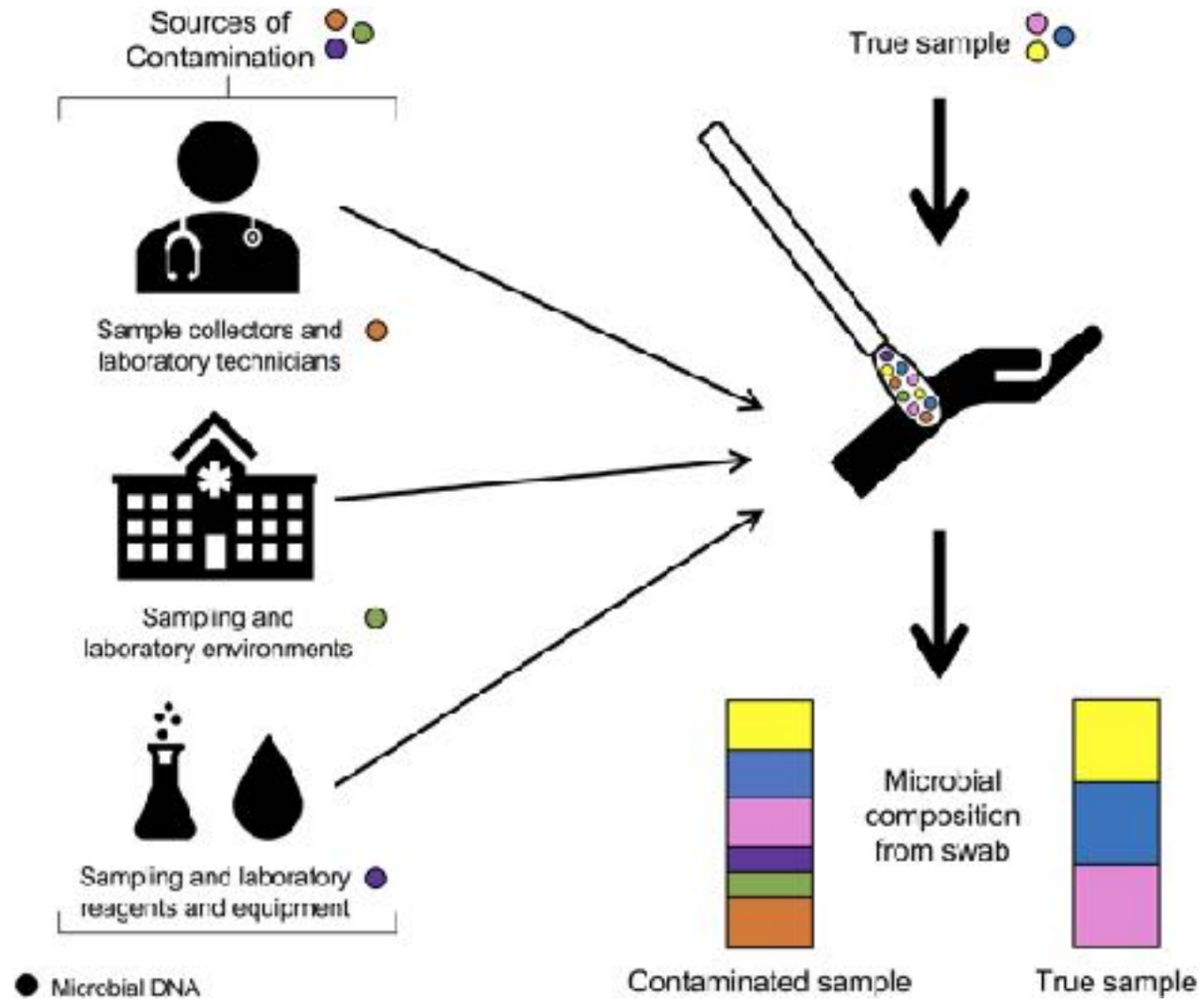
# Modern samples can also be similarly problematic

Skin swabs
Placenta and Pre-term infants
Respiratory biopsies
Blue Ice
Ancient/Historic Specimens
and more…

Are high biomass samples also susceptible to contamination?

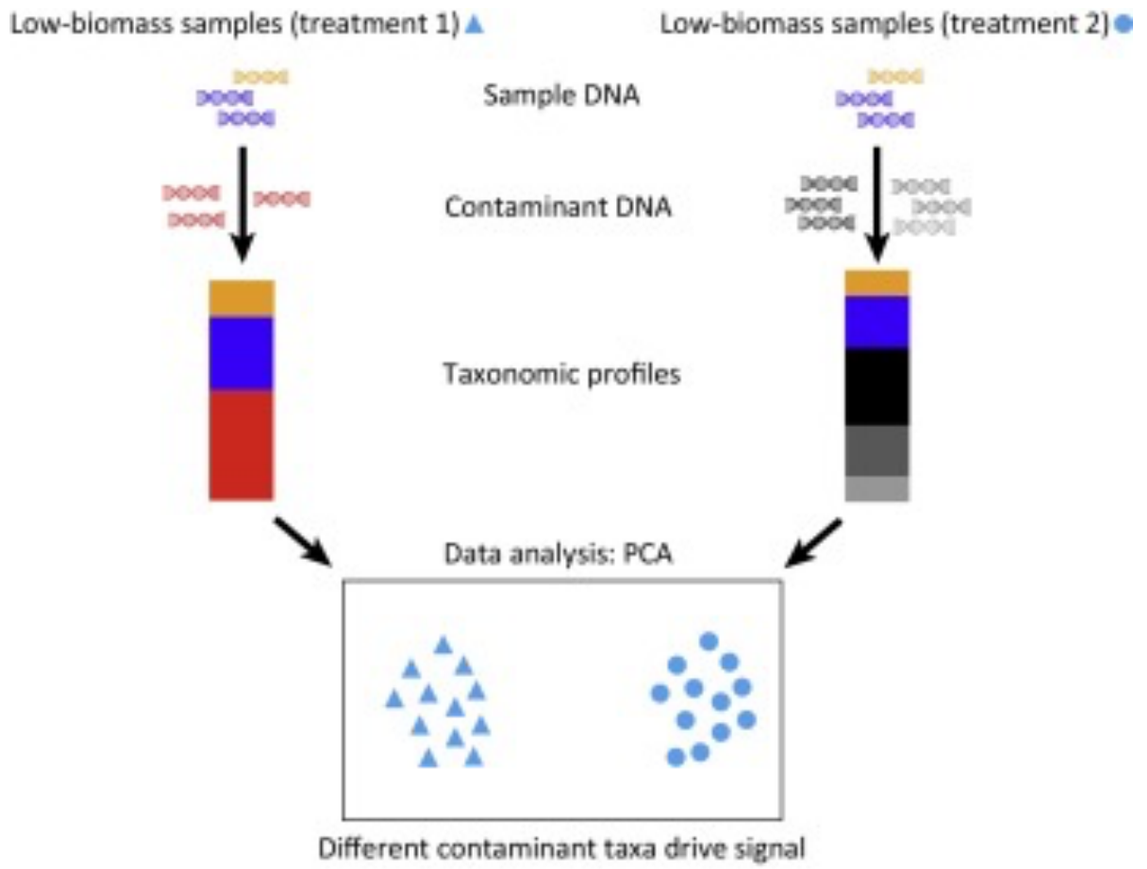
# WHERE DOES CONTAMINATION ORIGINATE?

1. Outdoor Environment
2. People - Researchers and Others
3. Lab Reagents and plasticware
4. Sequencing machines
5. Cross Contamination




Selway, et al. *Journal of Pathology,* 2020

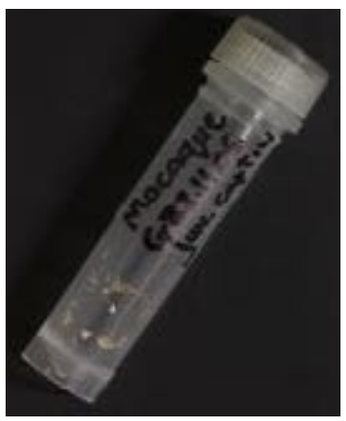
# CONCEPTS IN CONTAMINATION




1. Dedicated DNA extraction/processing facilities
2. Decontamination, if possible
3. Monitoring laboratory & environmental contamination
4. Efficient DNA extractions and library preparation
5. Accurate analysis and verification tools

Eisenhofer, et al. *Trends in Micro*. 2018

# 1. REDUCING CONTAMINATION WITH CLEAN FACILITIES



Isolated pre-PCR facility

Isolated ventilation

Glove boxes

Nightly UV irradiation

Clean -> Dirty workflow

Bleach treatment of stock

Reagents tested for DNA

Appropriate dress

Hofreiter, *et al*. 2001. Nature Reviews Genetics 2, 353-359
Willerslev and Cooper. Proc Biol Sci. 2005 Jan 7; 272(1558): 3–16

# 2. DECONTAMINATION MATTERS!

## Decontamination Procedures

A. Remove surface

B. Decontaminate sample

   -UV or washing, if able

C. Include controls every time

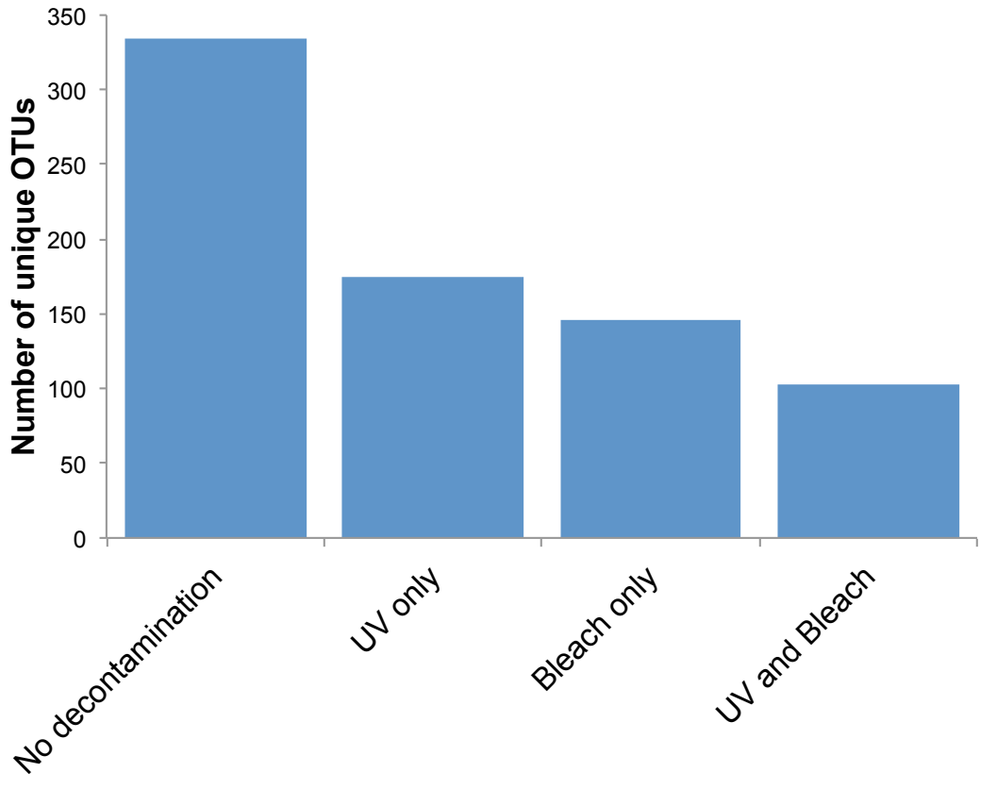
   -PCR negatives

   -Extraction blank controls

D. Duplicate samples

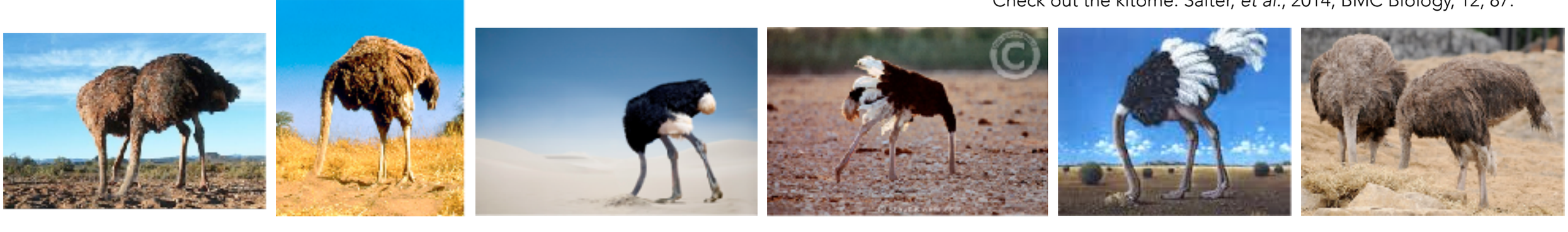
E. Remove contaminants with bioinformatic methods

DNA Contamination

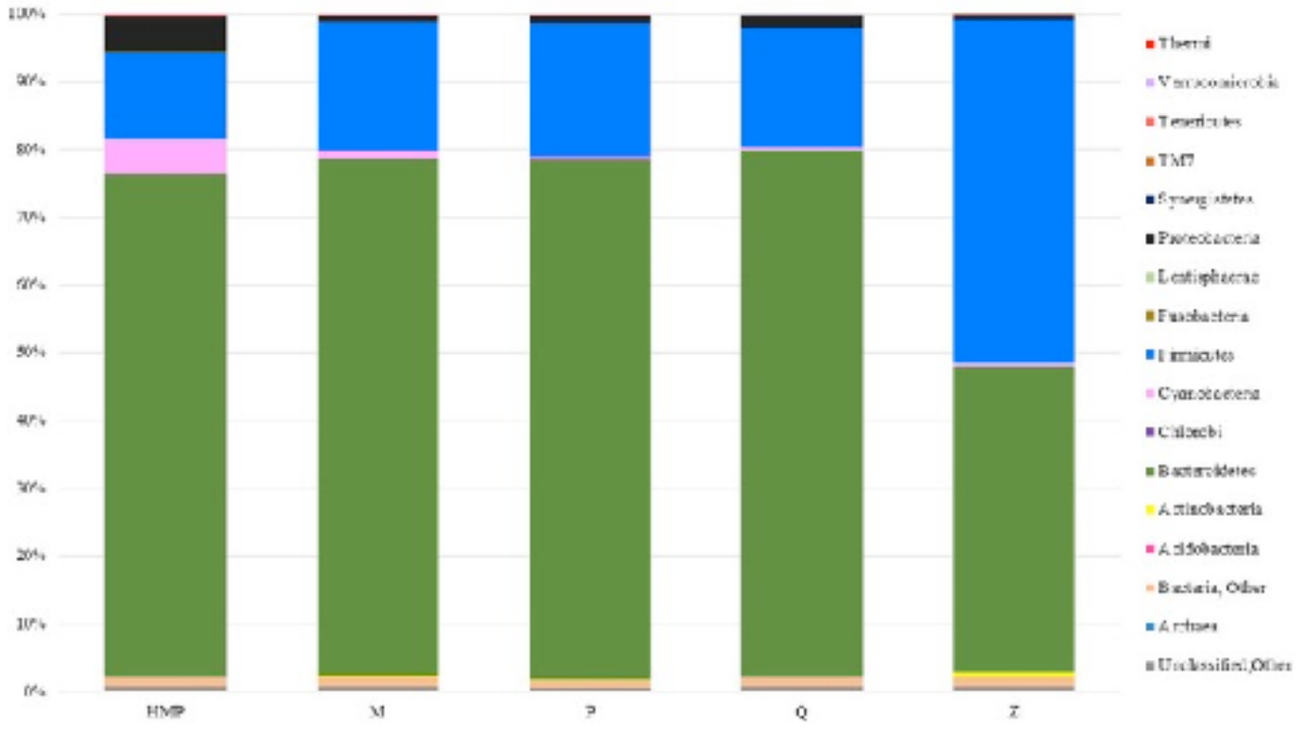



A. Farrer, in prep.

Check out the kitome: Salter, *et al.*, 2014, BMC Biology, 12, 87.



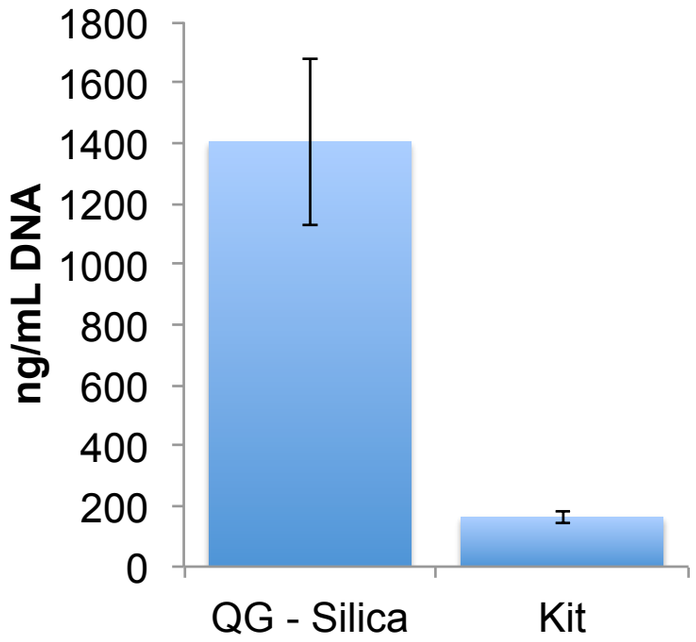
# 3. DNA EXTRACTION BIASES AND EFFICIENCY MUST BE CONSIDERED




Mackenzie, *et al*. Front Microbiol. 2015; 6: 130.

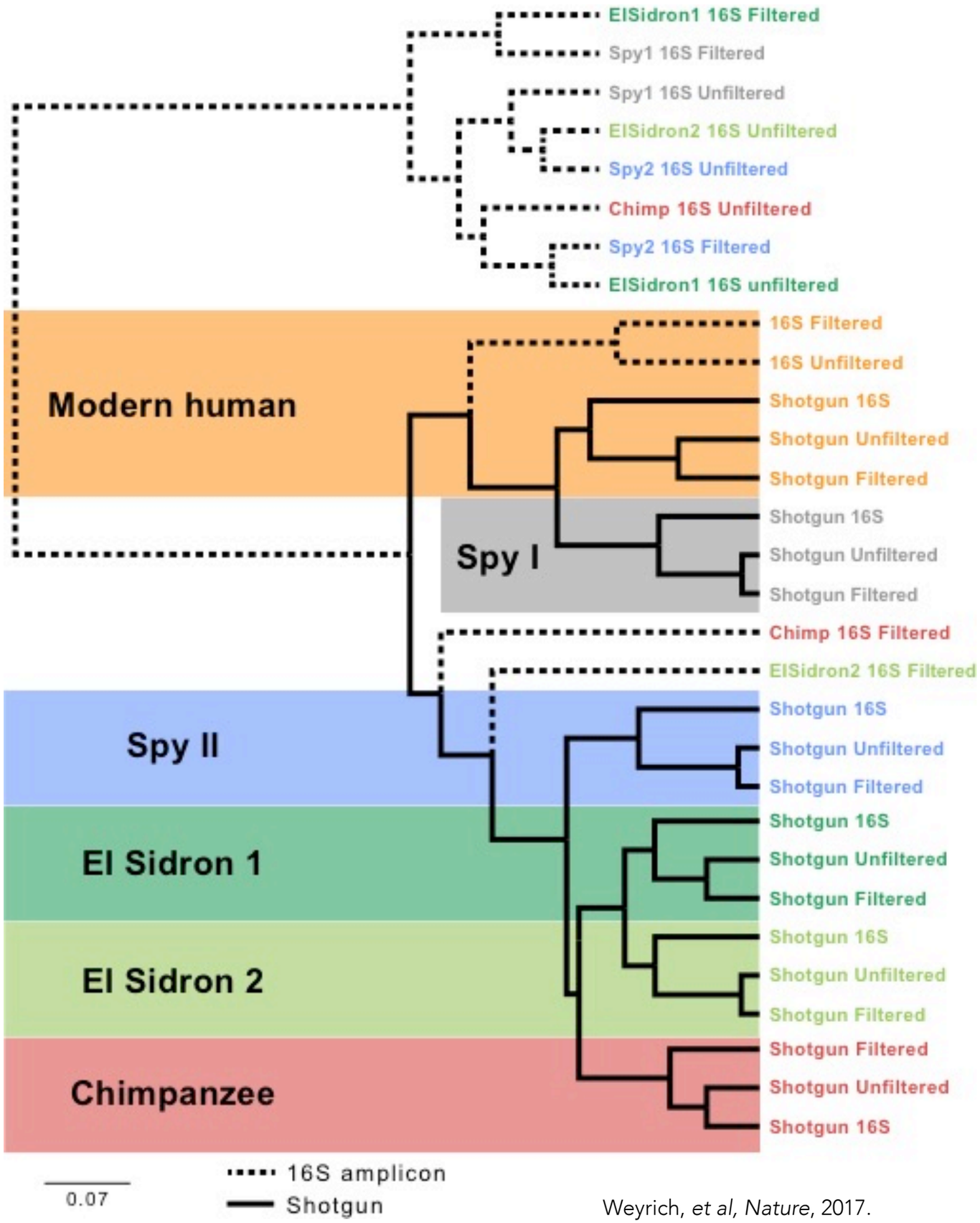
Extraction methods bias can results, and limit downstream comparisons.

QIITA!




Extraction efficiency matters!

# 3. LIBRARY PREPARATION ALSO MATTERS

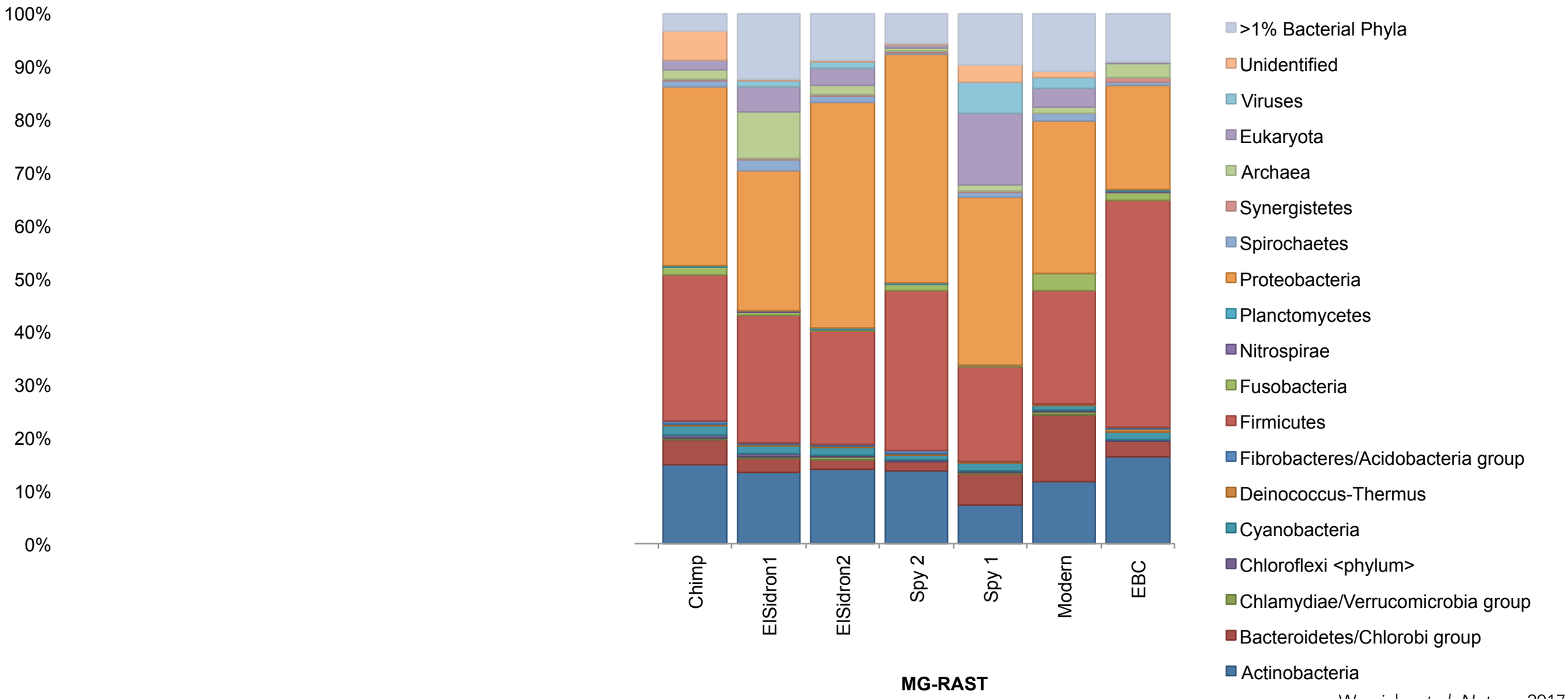



Weyrich, *et al, Nature*, 2017.

16S rRNA gene seq will likely not work for highly degraded samples.

Filtering helps (sometimes!)

# 4. BIOINFORMATIC METHODS TO IDENTIFY SPECIES FROM SAMPLES




Weyrich, *et al*, *Nature*, 2017.

Satisfies input requirements; accurate; rapid

nucleotide vs protein

Current methods applied: DIAMOND, MALTX

# 4. BIOINFORMATIC METHODS TO ACCOUNT FOR CONTAMINATION

1. Decontam — https://github.com/benjjneb/decontam
2. MEGAN6CE — https://github.com/husonlab/megan-ce
3. Direct Filtering — QIIME2, etc.
4. Contaminant Assessment — Comparison to known lists in:
   Salter, et al., *BMC Biol*, 2014 or
   Weyrich, et al., MER, 2019

What method is most stringent?

Eisenhofer, et al. *Trends in Micro*. 2018

# **RIDE** CHECKLIST FOR PERFORMING/REVIEWING LOW MICROBIAL BIOMASS MICROBIOME STUDIES

**Report** the experimental design and approaches used to reduce and assess the contributions of contamination.

**Include** controls to assess contaminant DNA. One of each type of negative control (sampling blanks, DNA extraction blanks, and no-template amplification) must be included per sampling, extraction, or amplification batch.

**Determine** the level of contamination by comparing biological samples to controls.

**Explore** contaminant taxa within each study and report their impact on the interpretation of biological samples.

Eisenhofer, et al. *Trends in Micro*. 2018

# WHAT FACTORS SHOULD BE CONSIDERED BEFORE WORKING WITH LOW BIOMASS SAMPLES?

*Before you start…*

Consider your working environment.

Assess contamination risks.

Decide on the best approach (including extraction/lib).

*During analysis…*

Use clean environments for processing.

Include extraction blank and PCR controls.

Don’t let bioinformatics be a black box.

Scrutinize and test your results!

**And relax…!**

# QUESTIONS

1.) What are the contamination risks?

2.) What steps might you take to mitigate those risks?

3.) What issues might you encounter in your own projects?