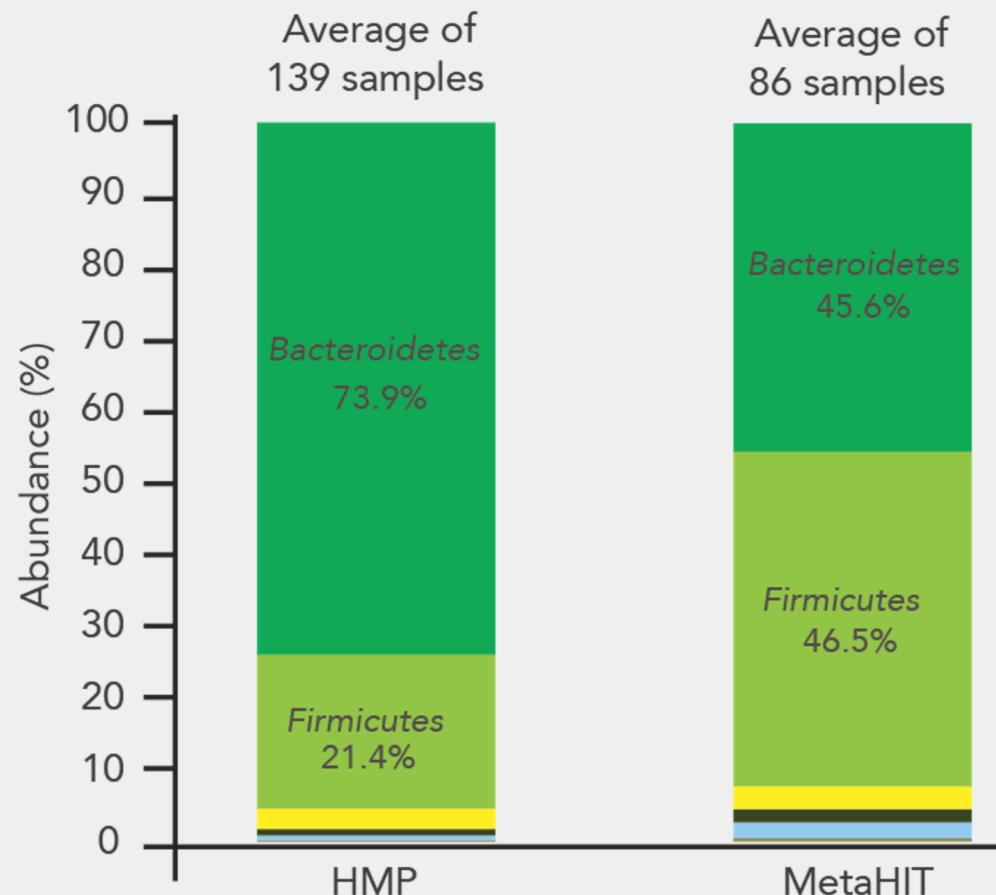


Topic

PITFALLS IN MICROBIOME ACCURACY AND REPRODUCIBILITY

Study-to-study Comparability Was Called Into Question



Wesolowska-Andersen et al. *Microbiome* 2014, 2:19
http://www.microbiomejournal.com/content/2/1/19



Microbiome

RESEARCH

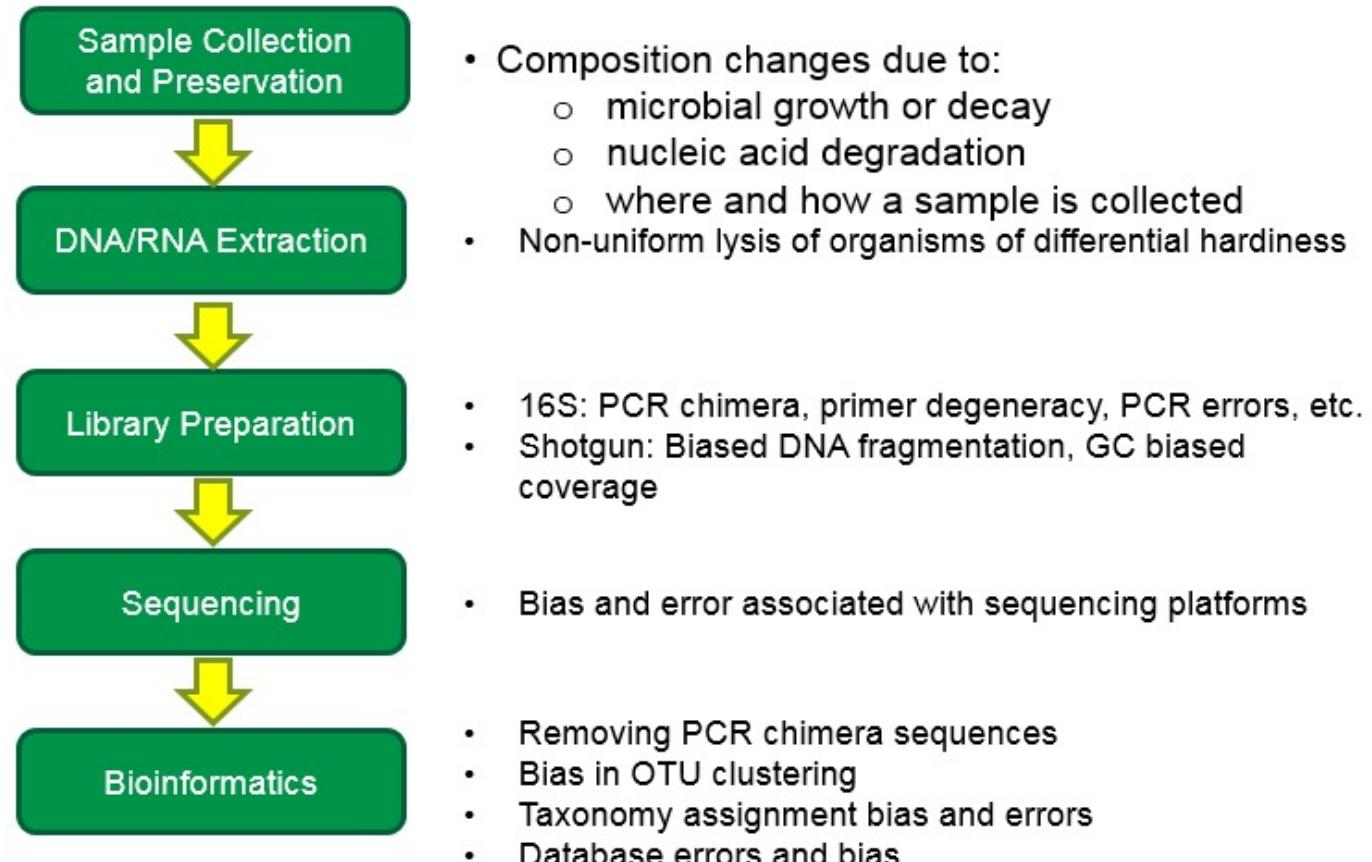
Open Access

Choice of bacterial DNA extraction method from fecal material influences community structure as evaluated by metagenomic analysis

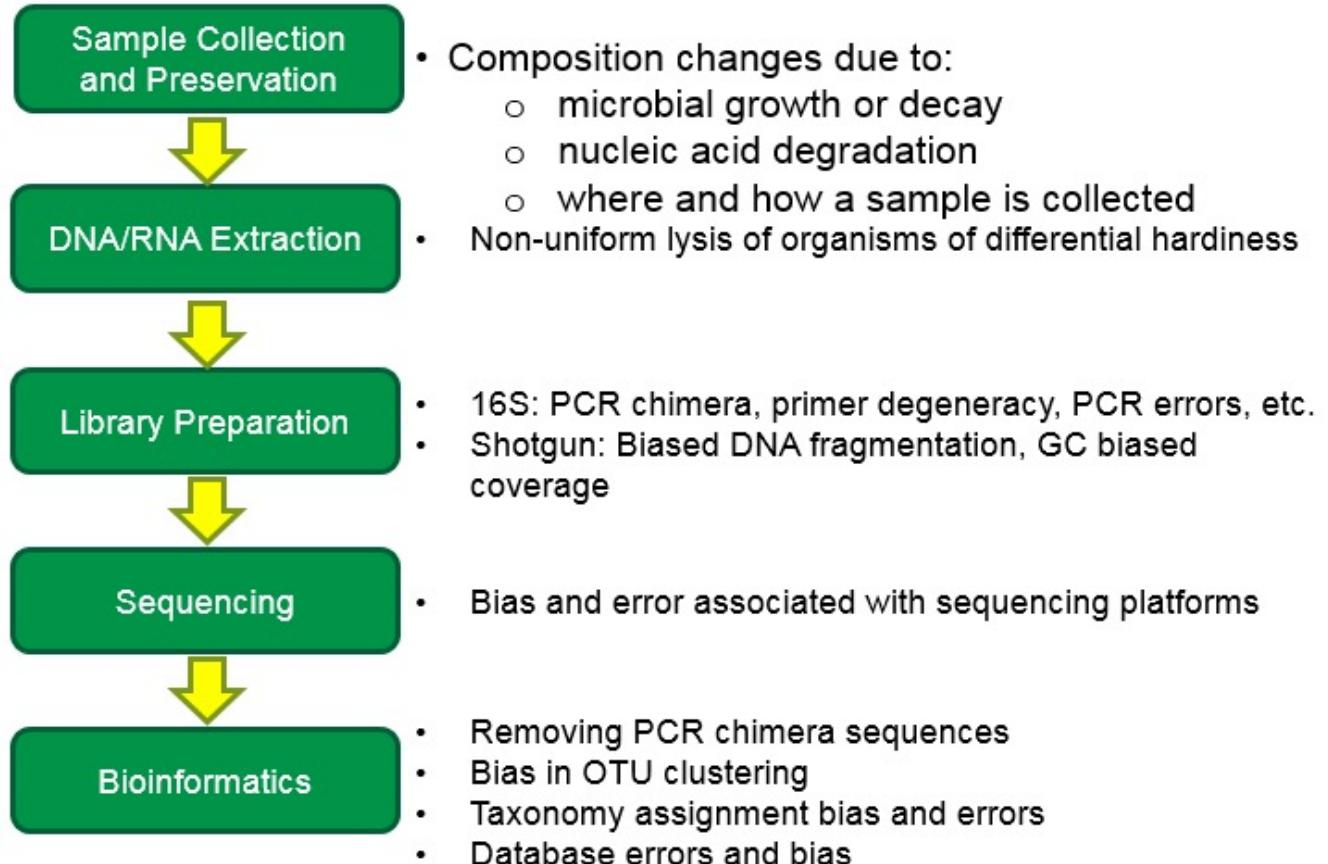
Agata Wesolowska-Andersen¹, Martin Iain Bahlo², Vera Carvalho², Karsten Kristiansen³, Thomas Sicheritz-Pontén¹, Ramneek Gupta^{1*} and Tine Rask Licht^{2*}

Conclusions: Whereas it is comforting that the inter-individual variation clearly exceeded the variation resulting from choice of extraction method, our data highlight the challenge of comparing data across studies applying different methodologies.

Data Source: Complex Workflow With Multiple Potential Sources of Bias



Bias in Microbiomics Workflows



WHAT IS BIAS?

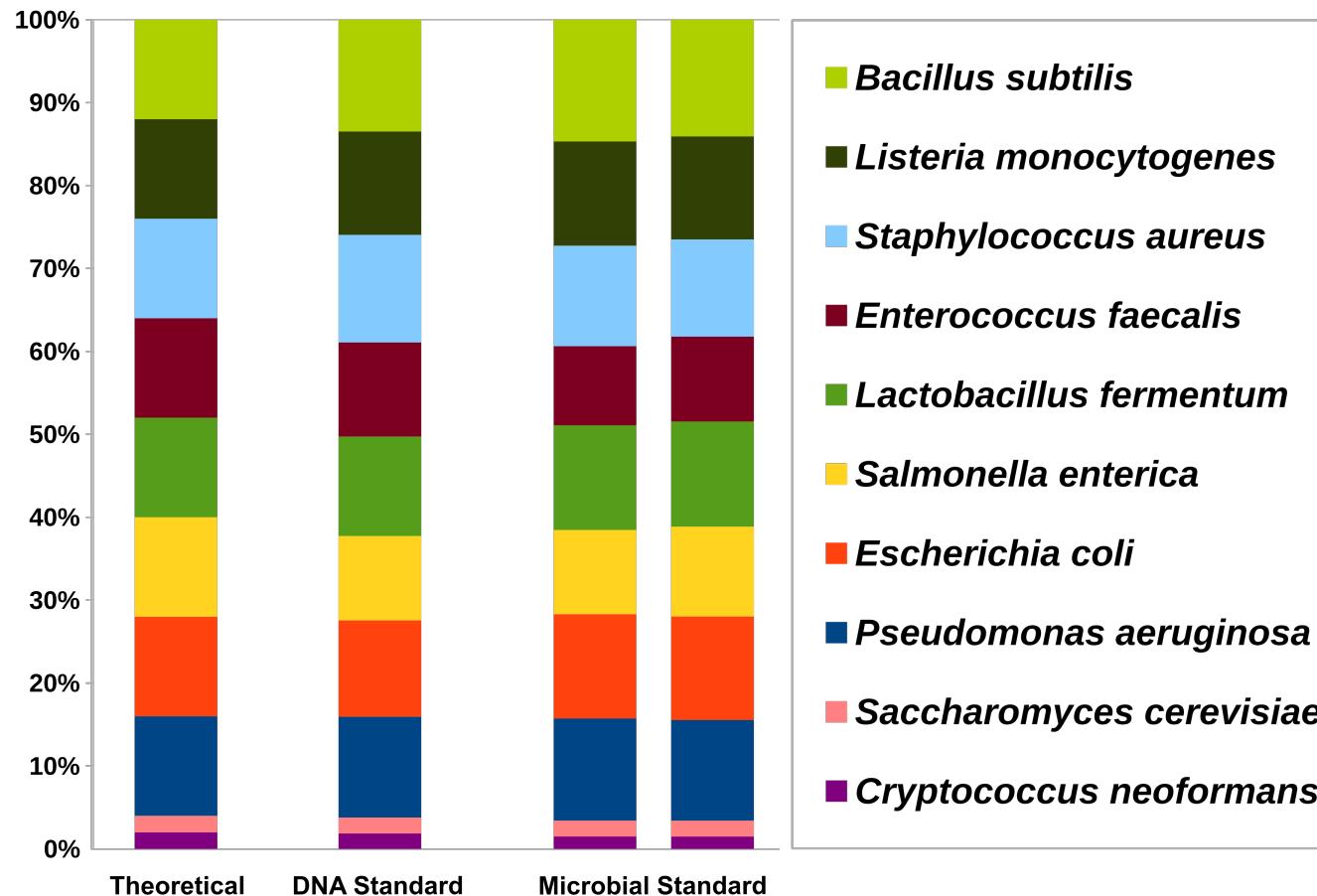
“bias occurs when systematic error is introduced into sampling or testing by selecting or encouraging one outcome or answer over others.”

Topic

THE IDEA OF A KNOWN INPUT FOR BENCHMARKING

A Standard Ground Truth For Measurement

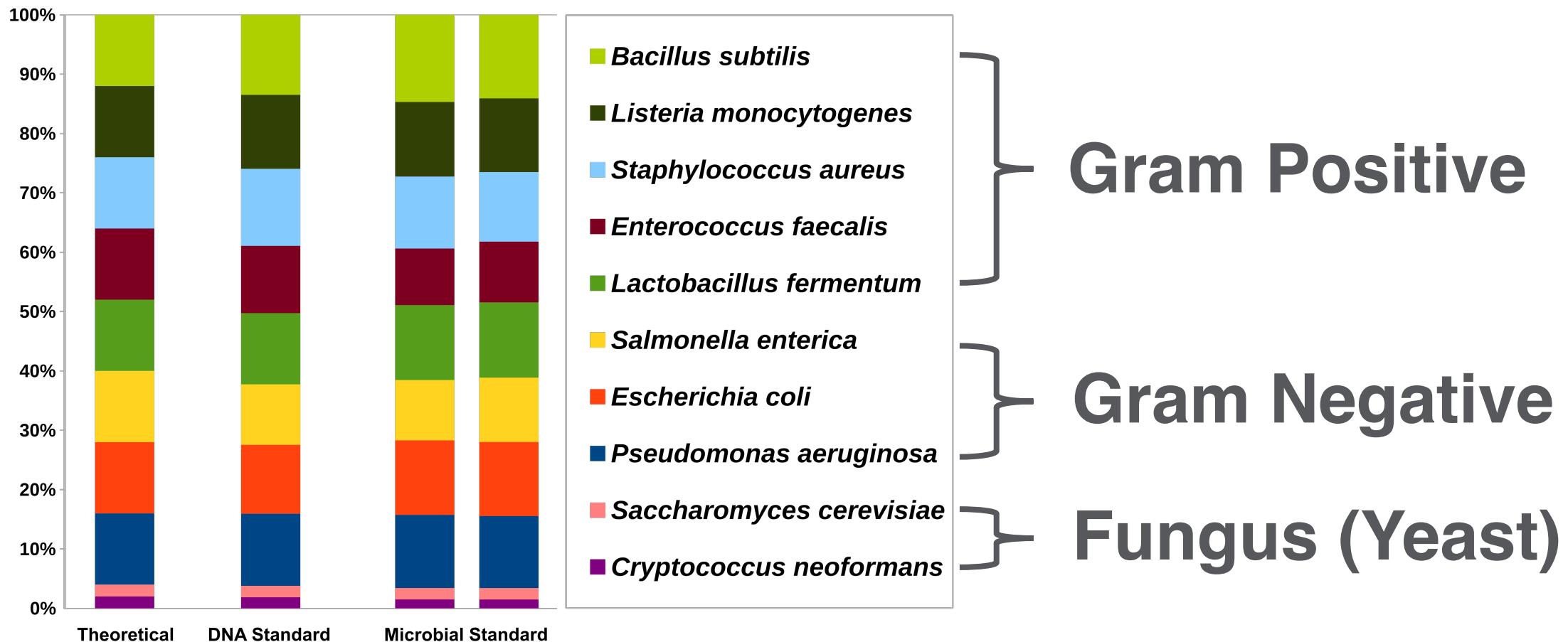
ZymoBIOMICS Mock Microbial Community



< 15% deviation from defined composition
< 0.01% foreign microbial DNA

A Standard Ground Truth For Measurement

ZymoBIOMICS Mock Microbial Community

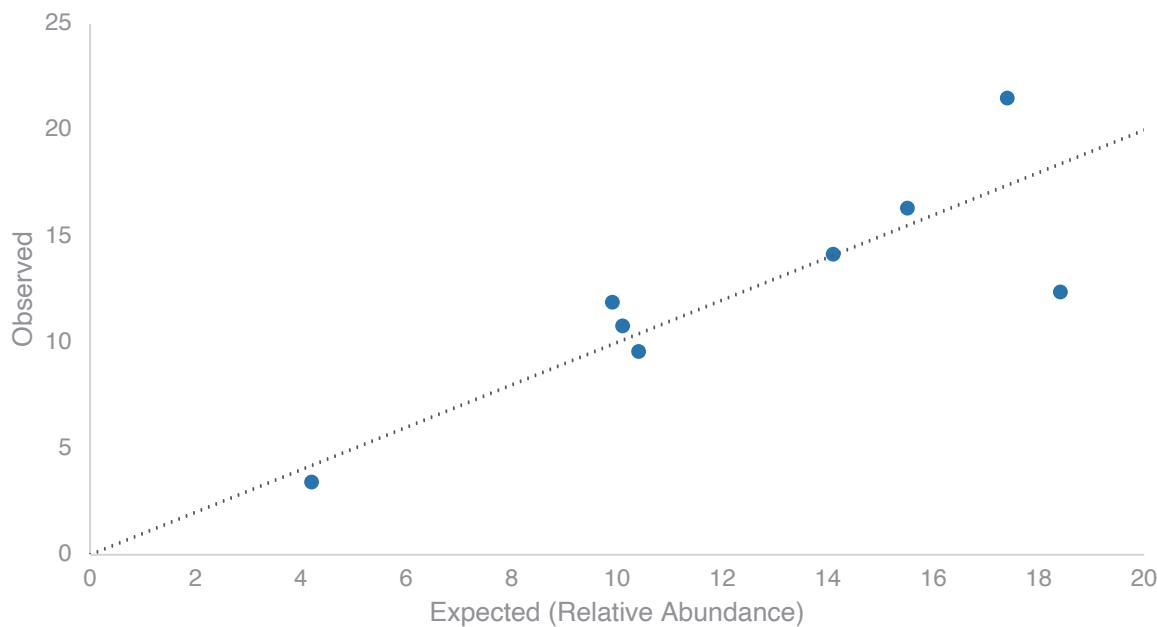


Topic

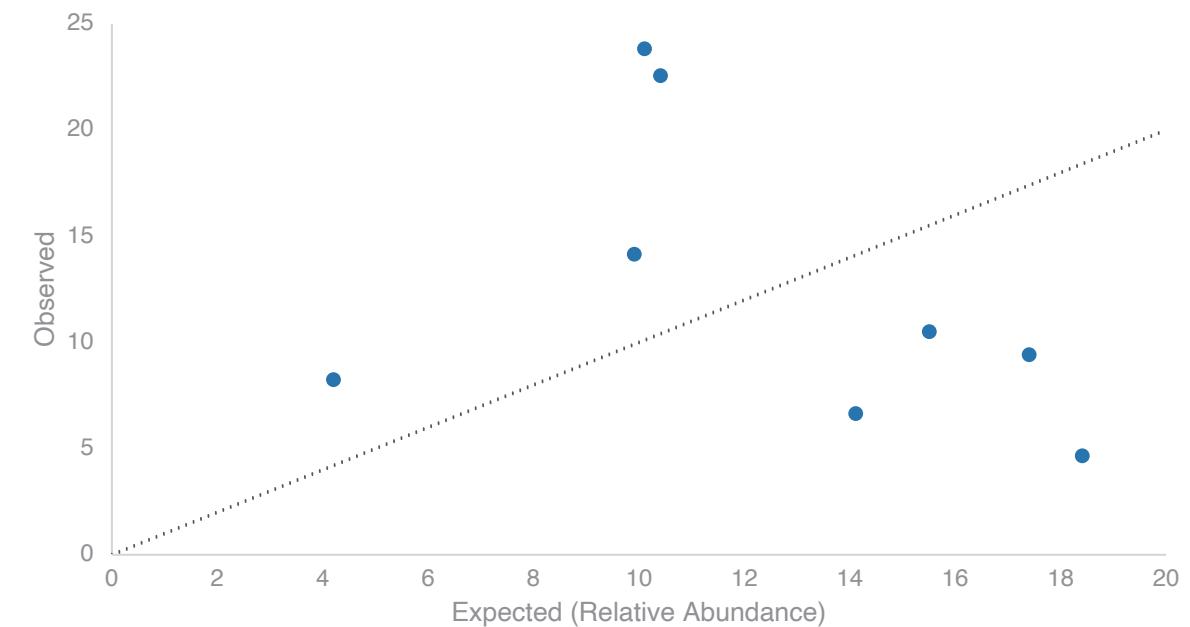
HOW TO MEASURE ACCURACY RELATIVE TO A STANDARD?

Measuring Accuracy Relative to a Standard

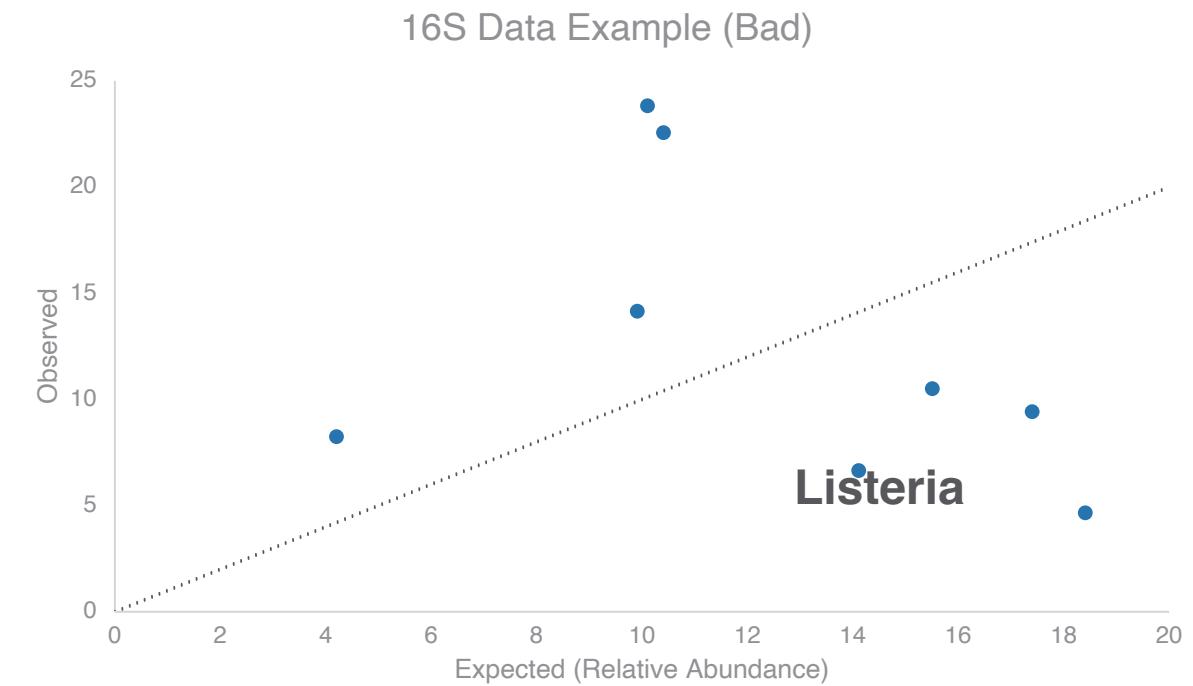
16S Data Example (Good)



16S Data Example (Bad)

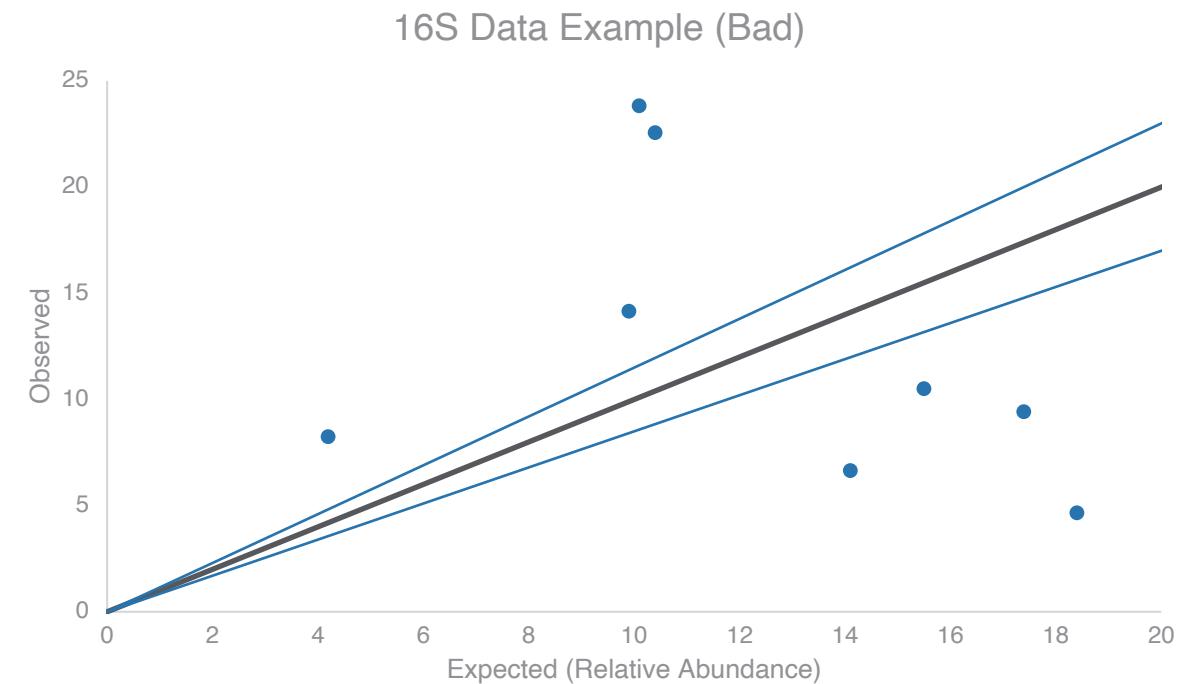
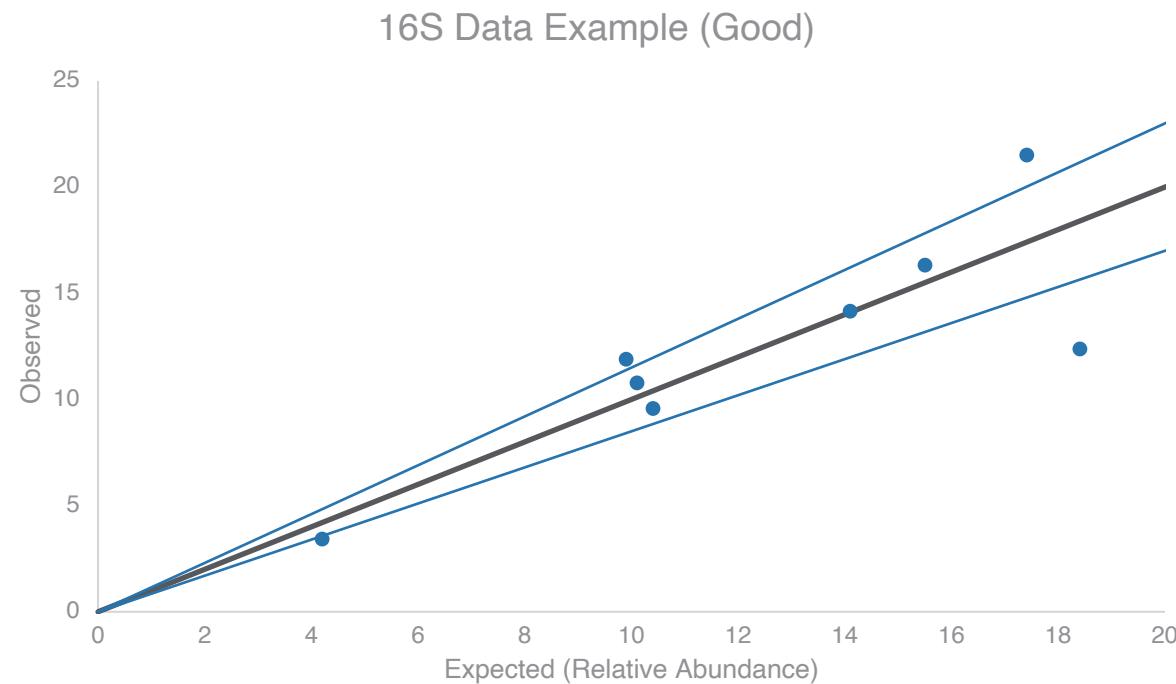


Measuring Accuracy Relative to a Standard



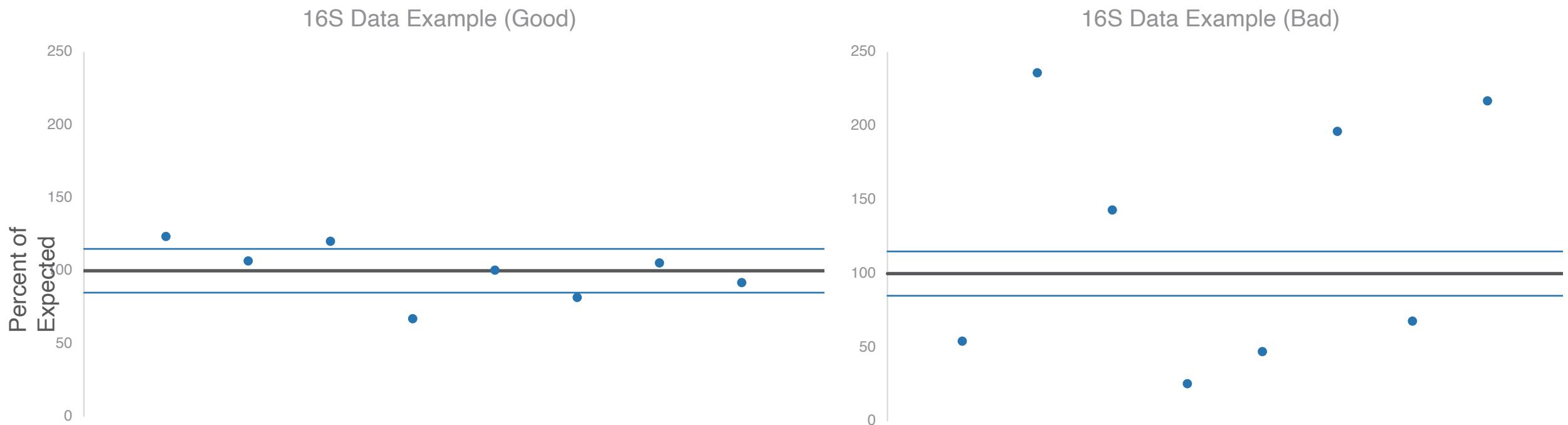
Measuring Accuracy Relative to a Standard

(adjustment for 15% tolerance)



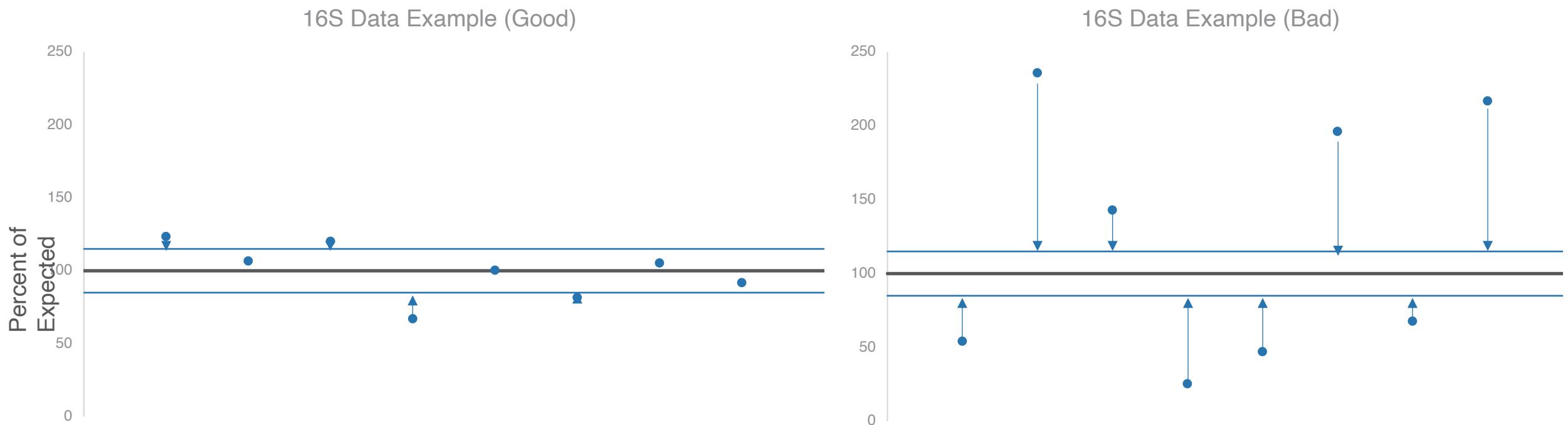
Measuring Accuracy Relative to a Standard

(Transform for Percent Expected)



Measuring Accuracy Relative to a Standard

(Determine Error Relative to Tolerances)

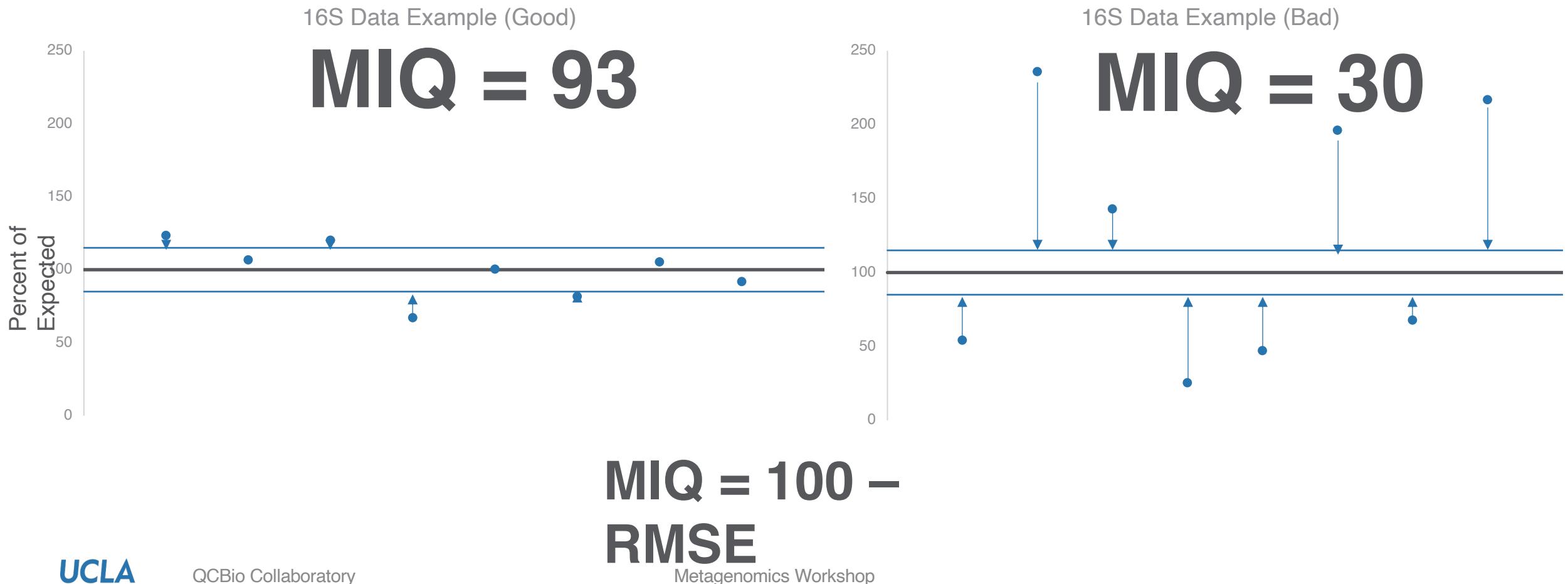


Topic

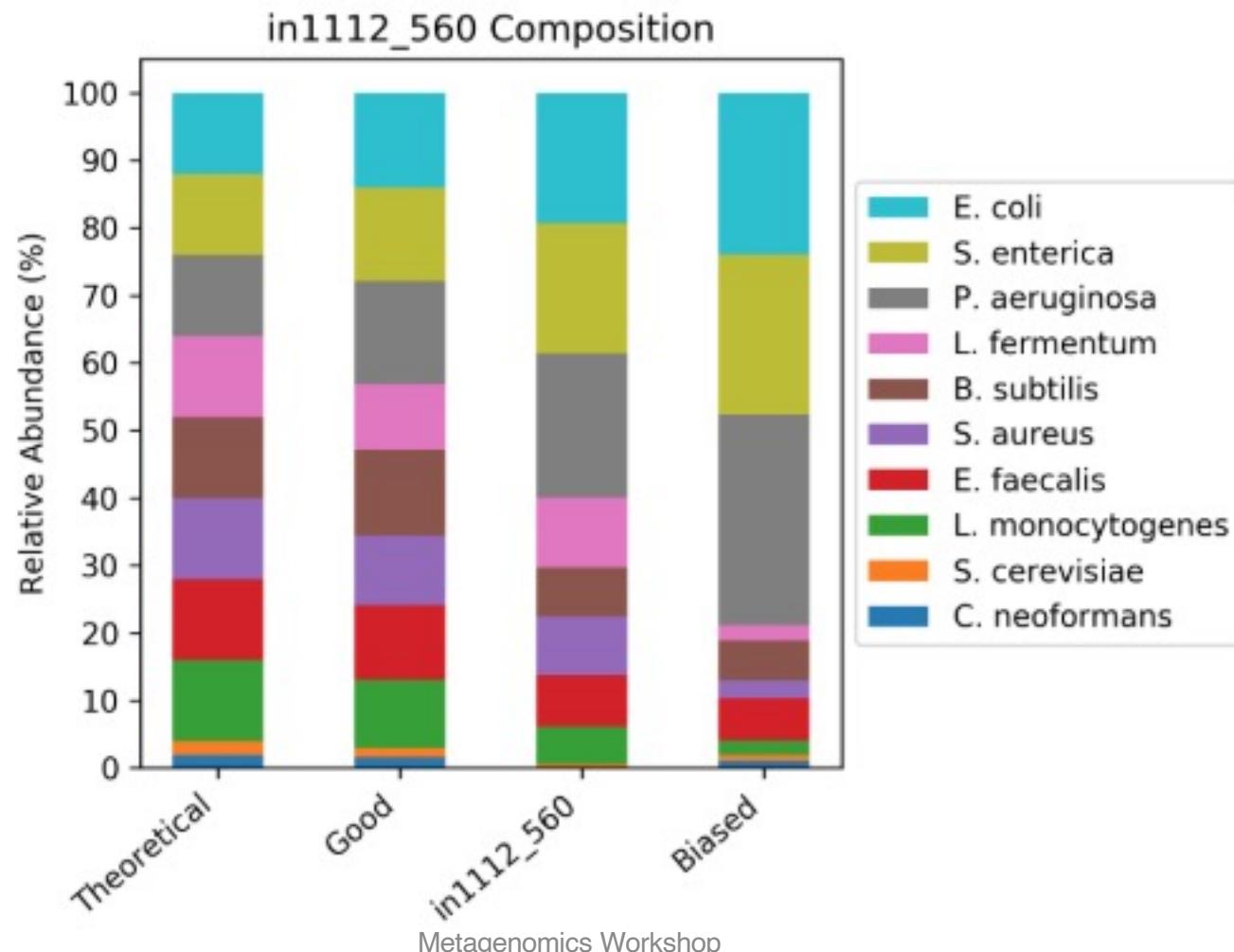
INTRODUCING THE MEASUREMENT INTEGRITY QUOTIENT (MIQ) SCORE

Measuring Accuracy Relative to a Standard

(Determine Error Relative to Tolerances)

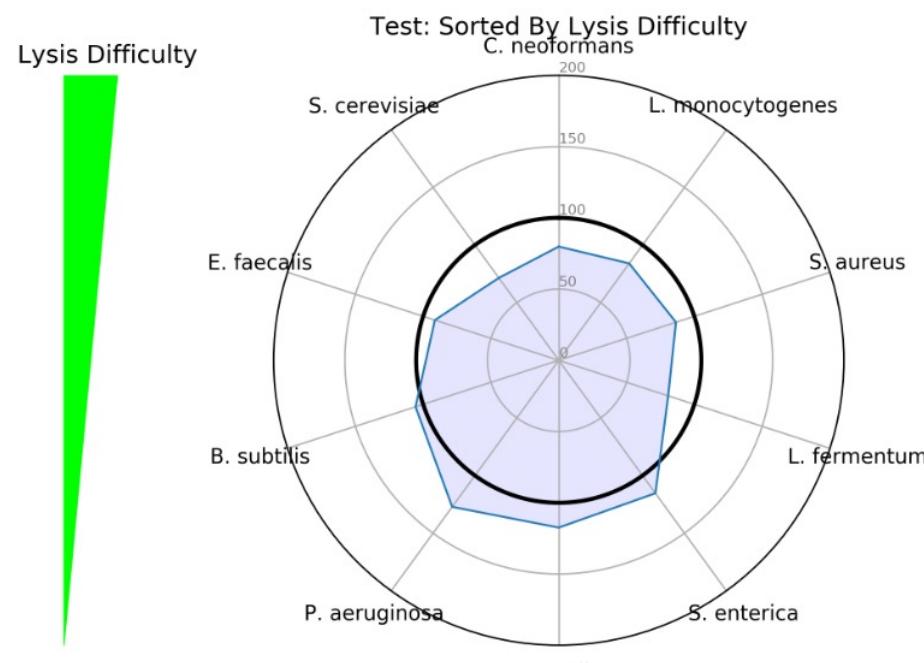


Composition/Taxa Plots Are Insensitive to Inaccuracies



MIQ Score Report (Generated Automatically)

MIQ = 93



Good (Example)

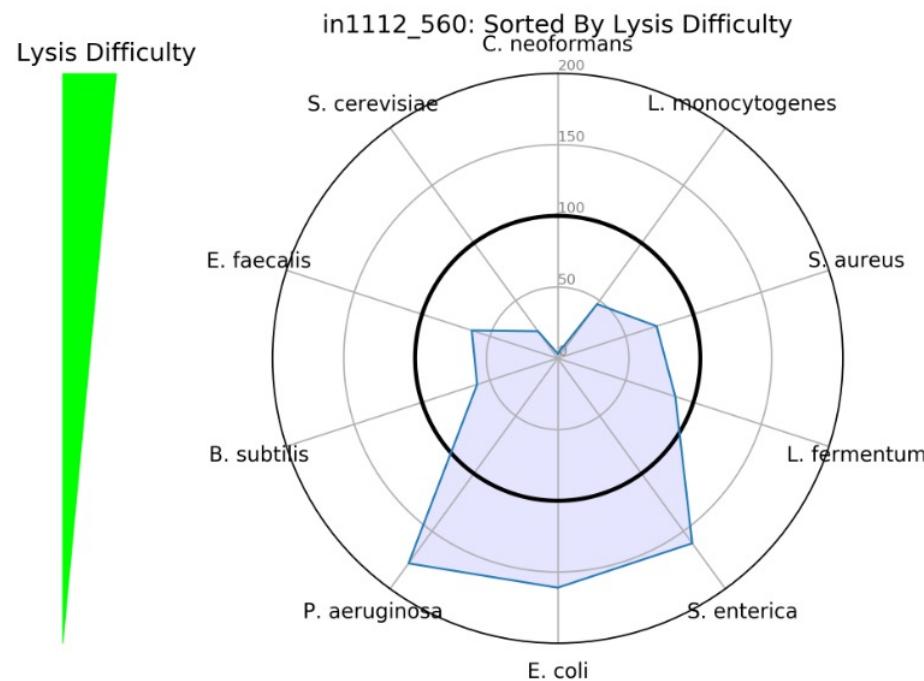
Harder to Lyse



Easier to Lyse

MIQ Score Report (Generated Automatically)

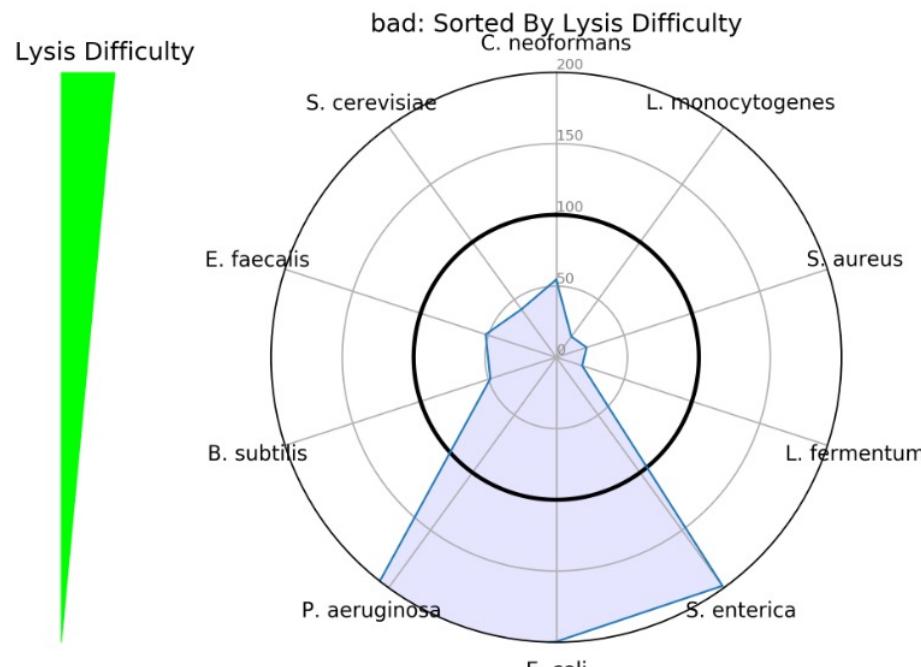
MIQ = 54



in1112_560

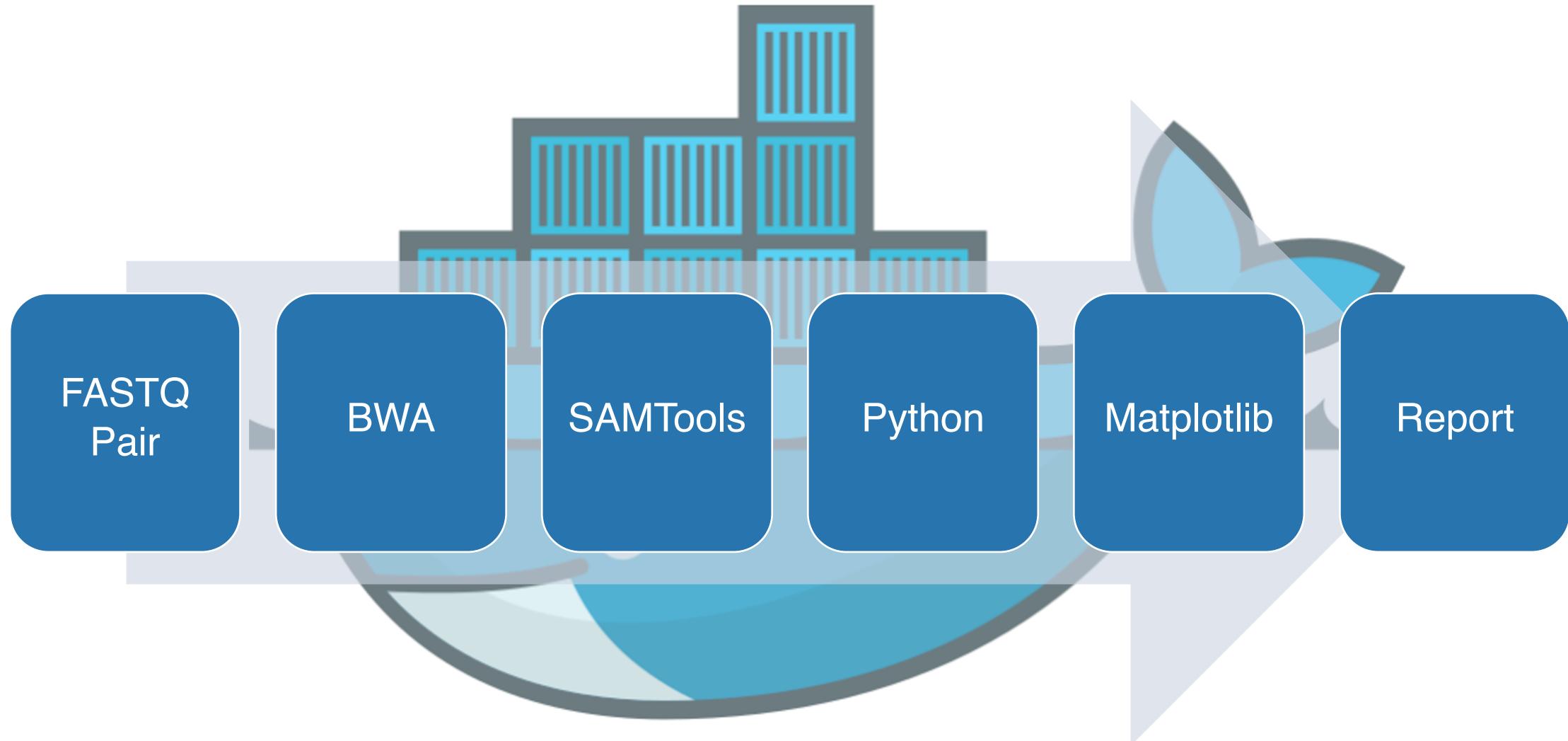
MIQ Score Report (Generated Automatically)

MIQ = 30

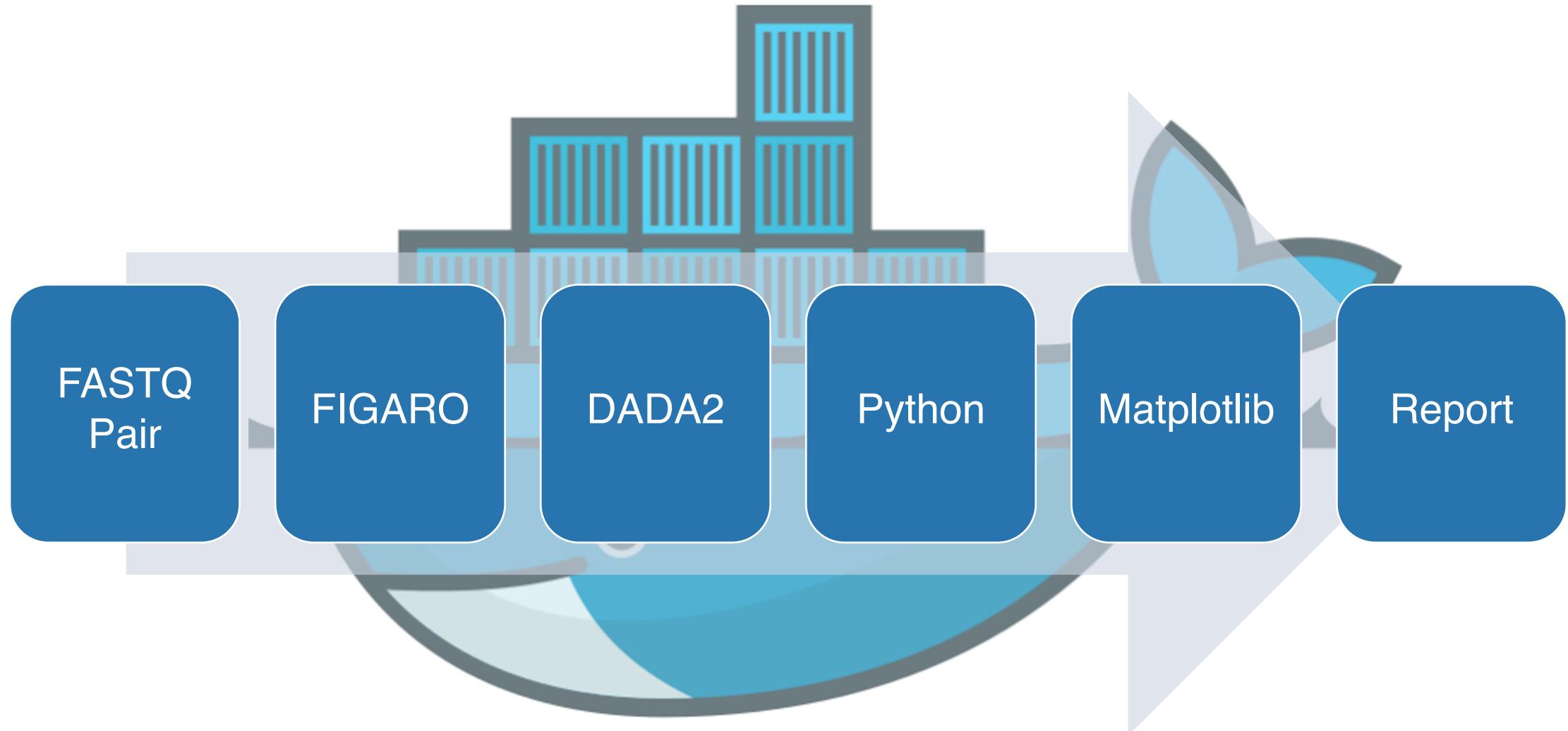


Biased (Example)

Shotgun MIQ Score is a Complete, Dockerized Pipeline



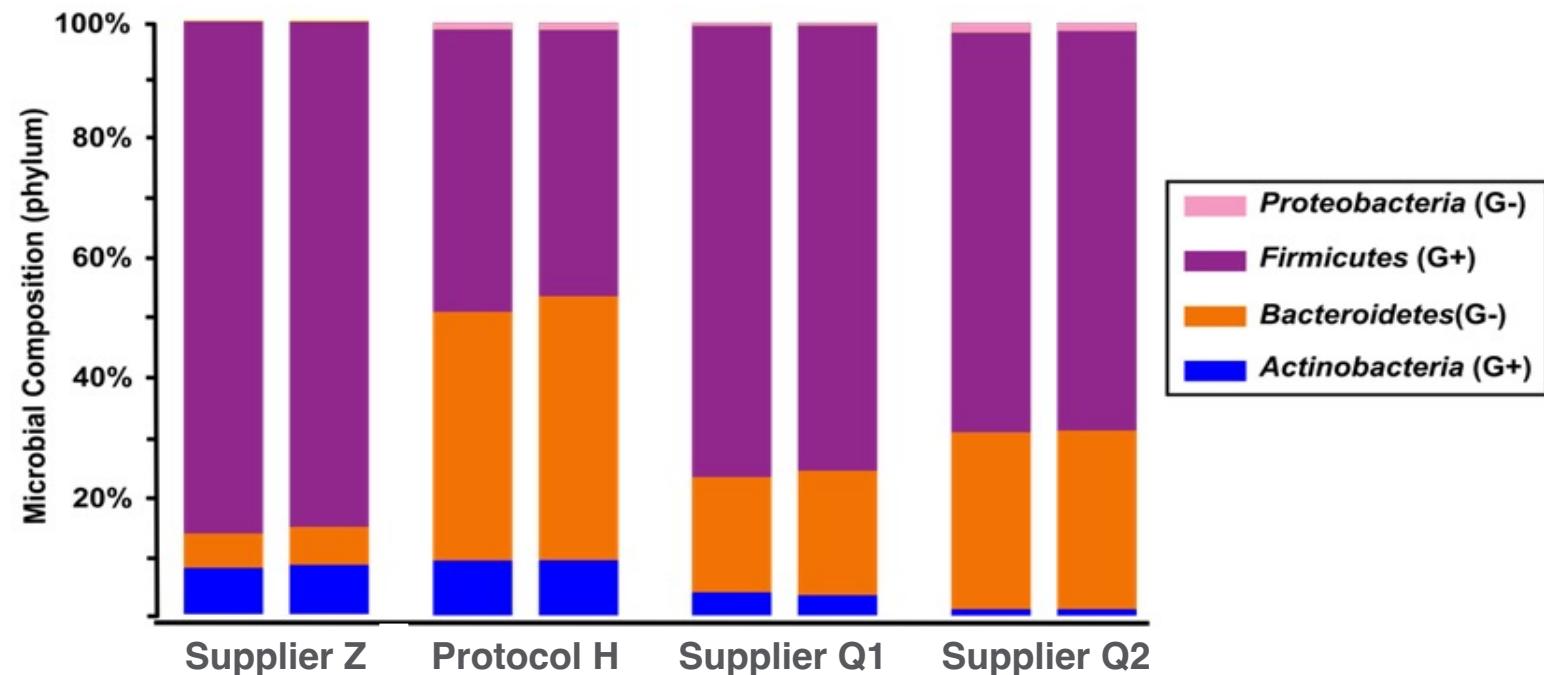
16S MIQ Score is a Complete, Dockerized Pipeline



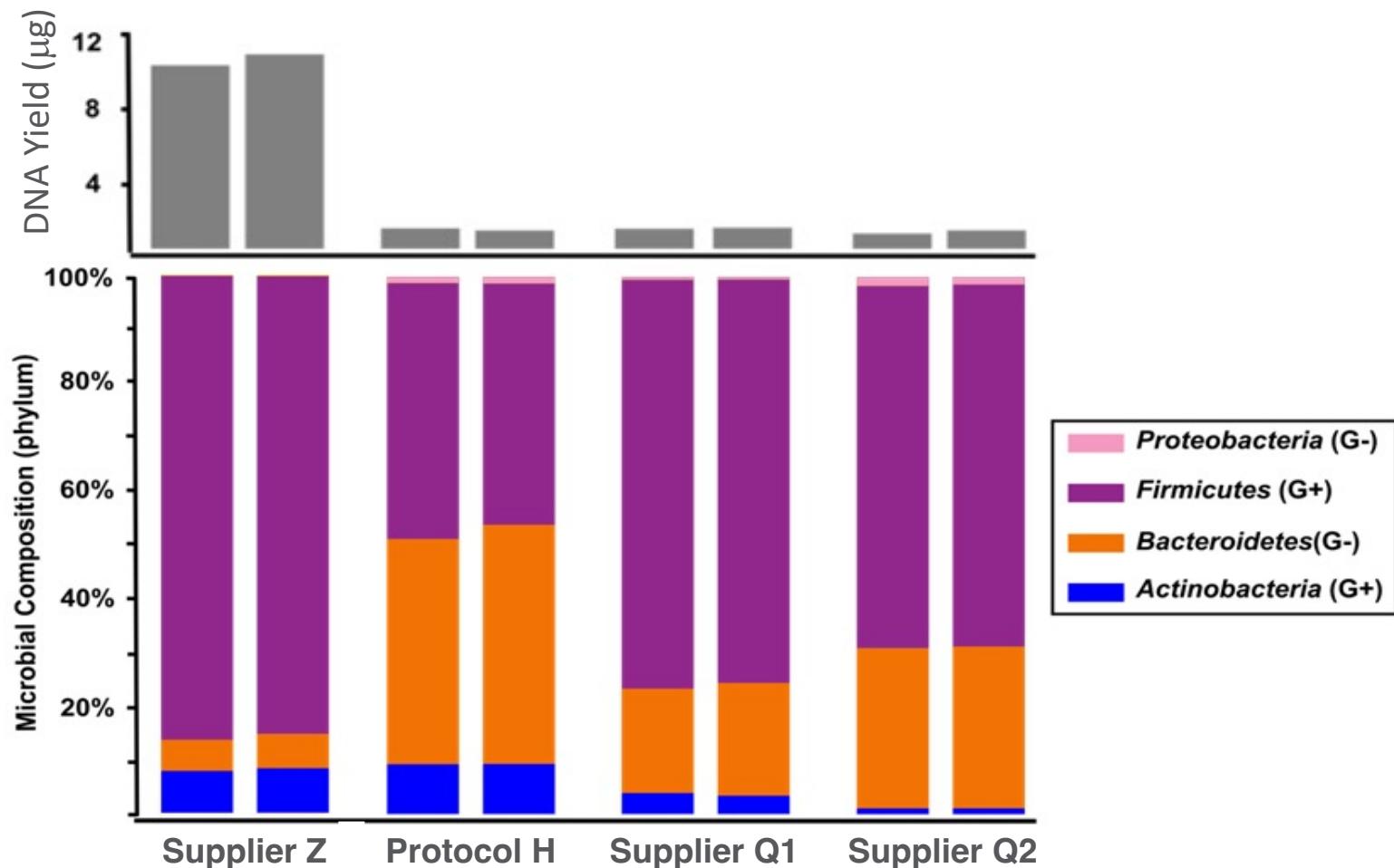
Topic

UNDERSTANDING CAUSES OF NON-REPRODUCIBILITY AND BIAS

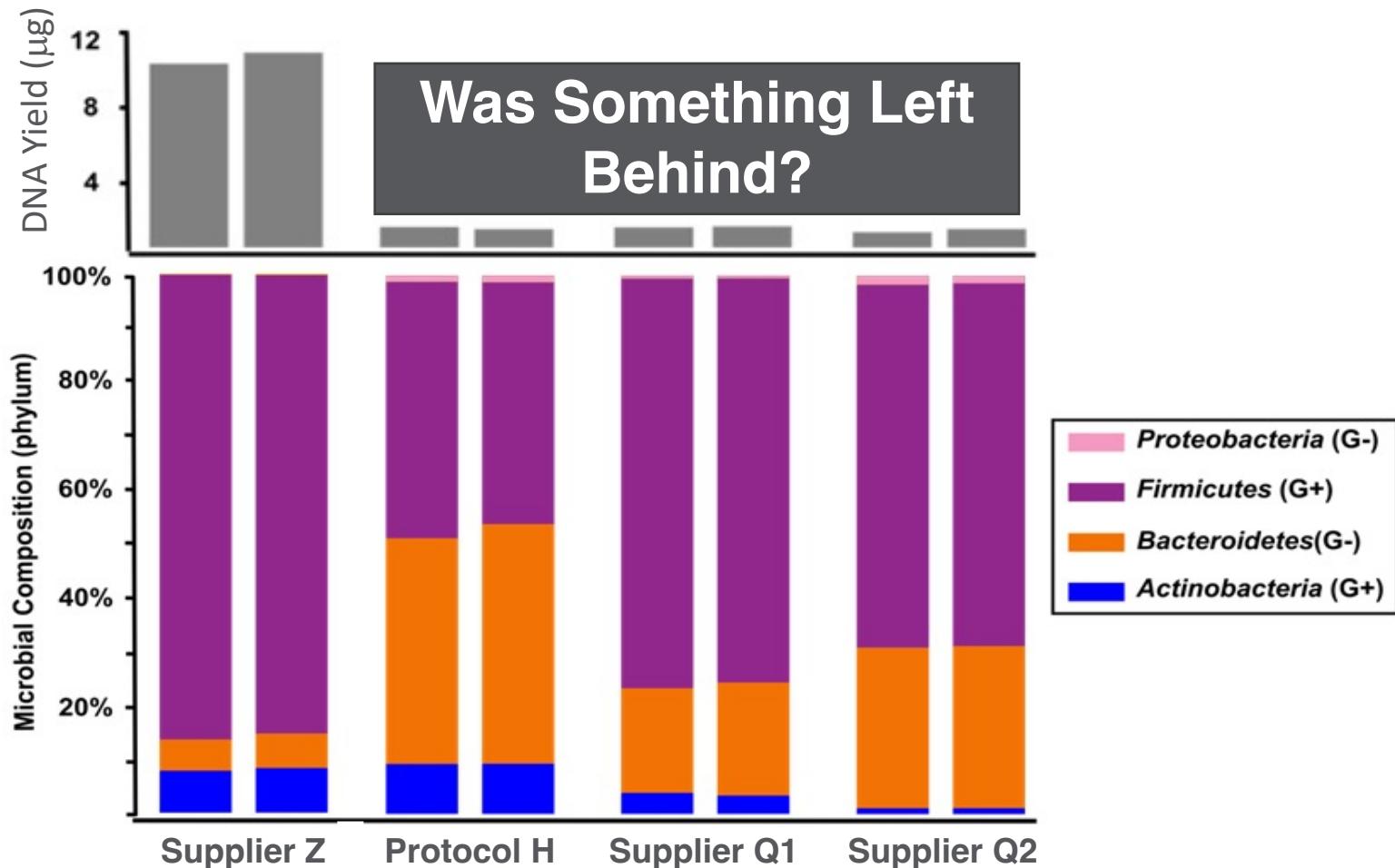
Lack of Reproducibility Between Methods



Notable Differences in Total DNA Yield



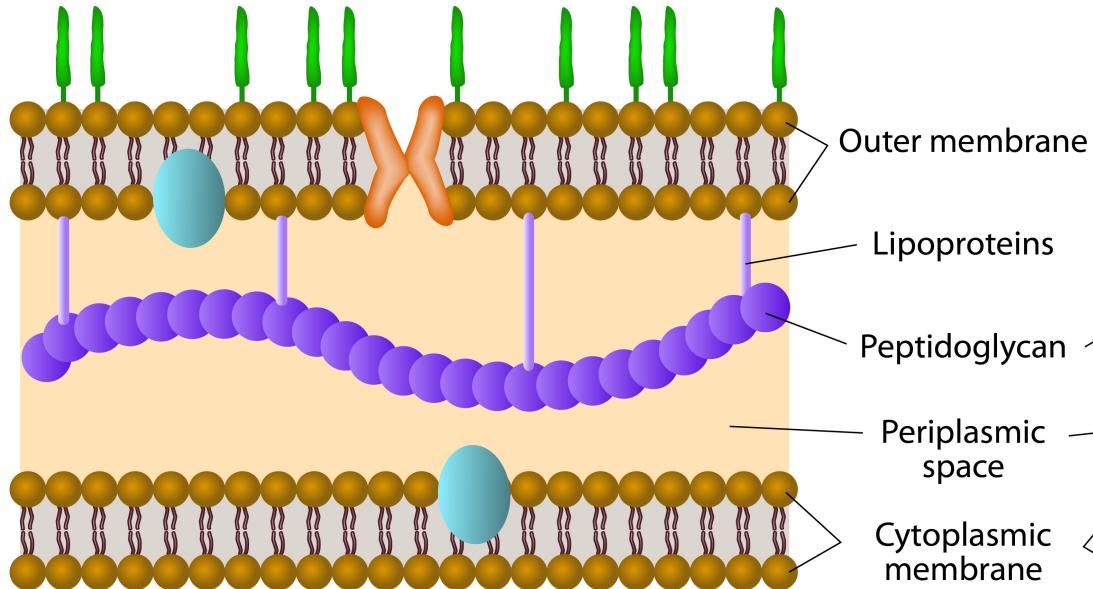
Notable Differences in Total DNA Yield



Hypothesis: Physical Differences in Microbes Cause Lysis Bias

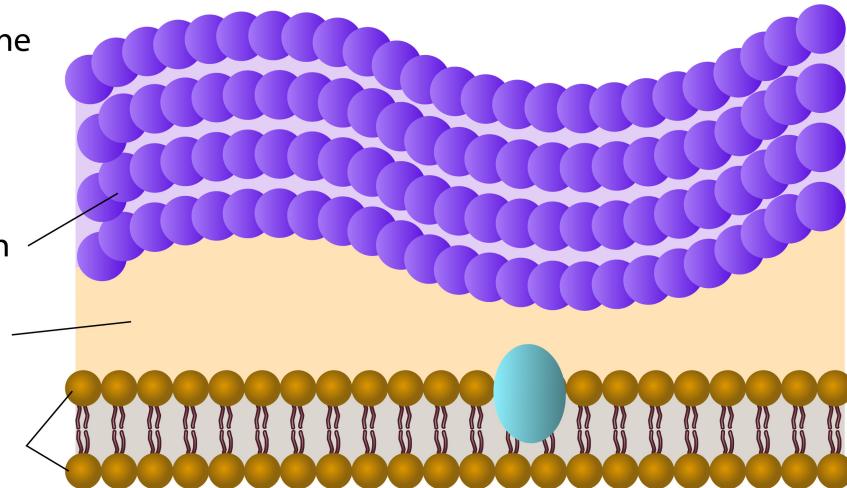
GRAM-NEGATIVE

Easy to lyse



GRAM-POSITIVE

Difficult to lyse



Lipopolysaccharides



Porin



Protein

Why Does Extraction Lead to Bias?

GRAM-NEGATIVE

Easy to lyse

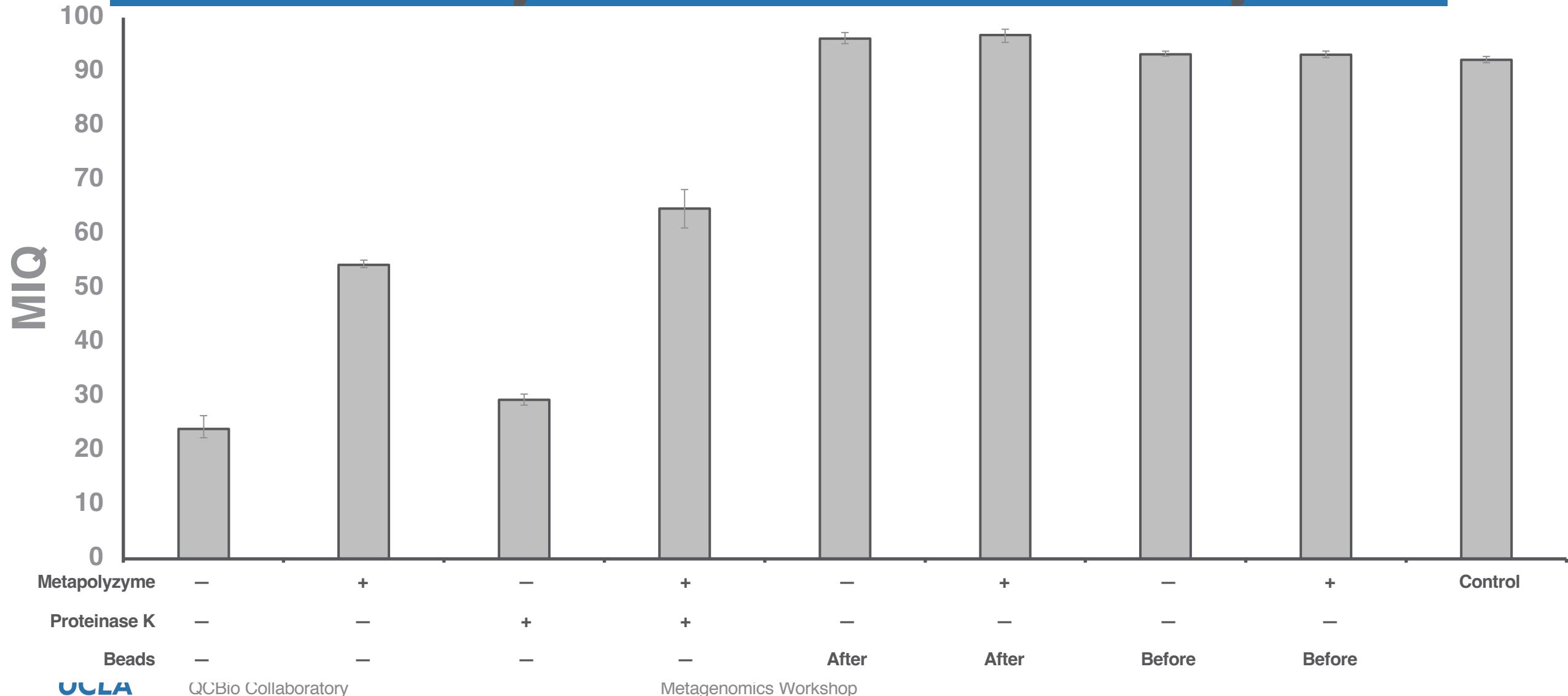


GRAM-POSITIVE

Difficult to lyse



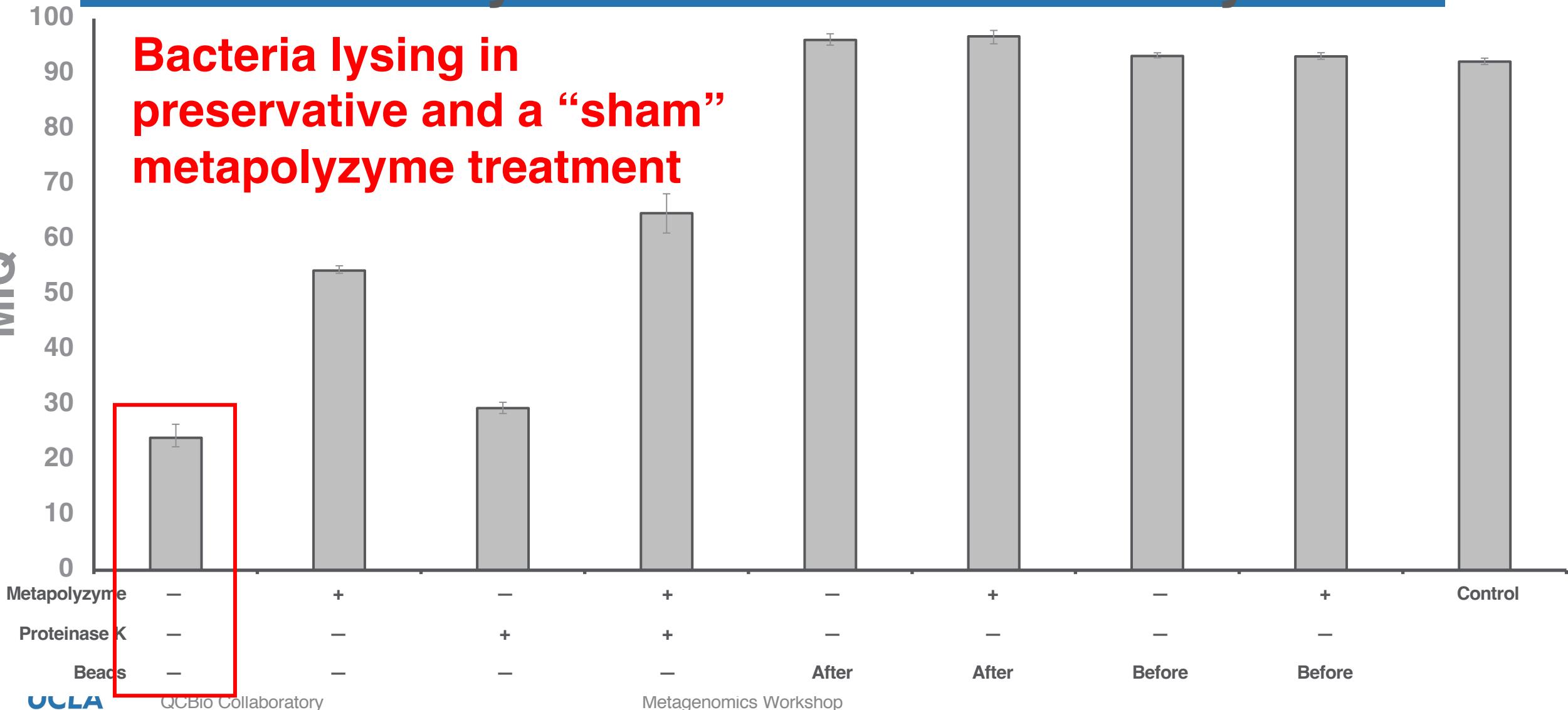
Chemical/Enzymatic and Mechanical Lysis



Chemical/Enzymatic and Mechanical Lysis

Bacteria lysing in preservative and a “sham” metapolyzyme treatment

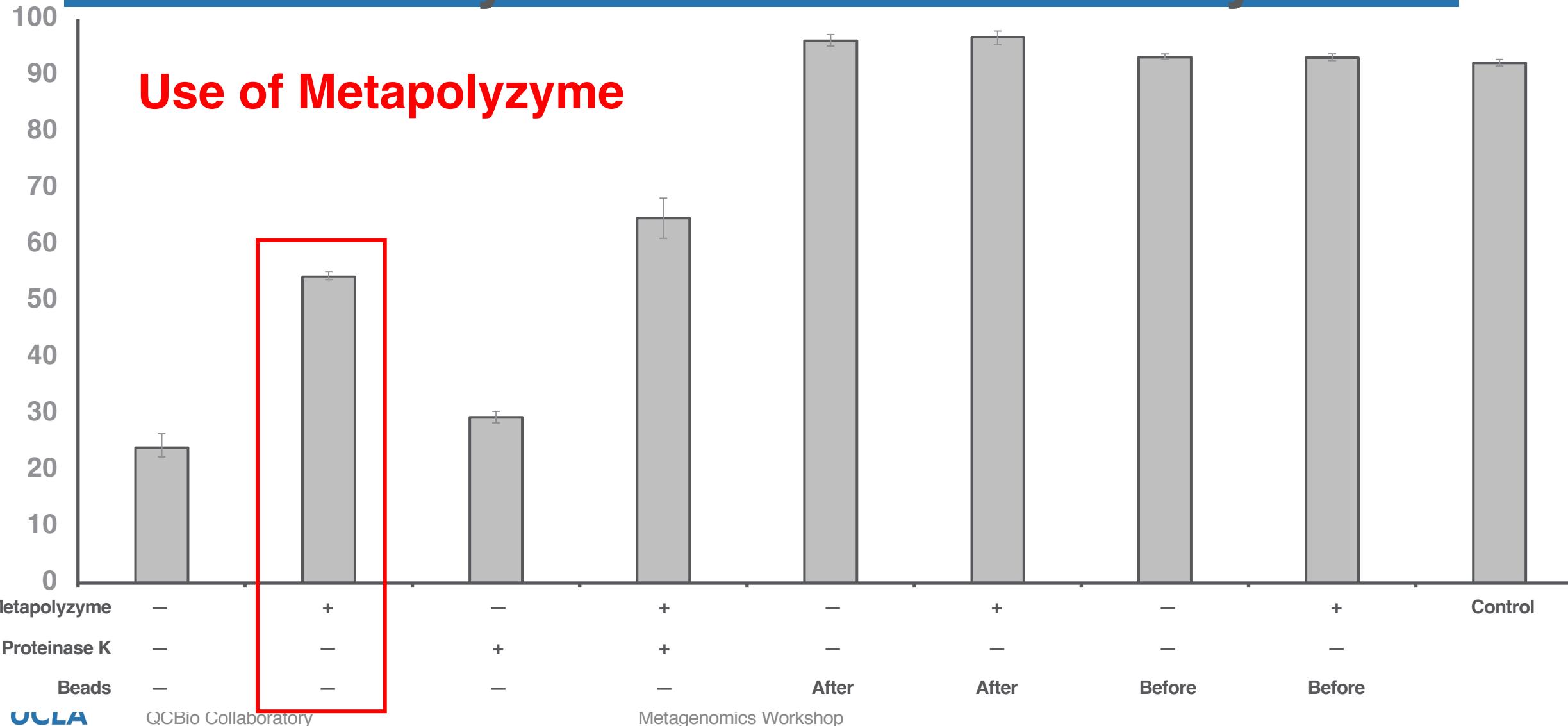
MIQ



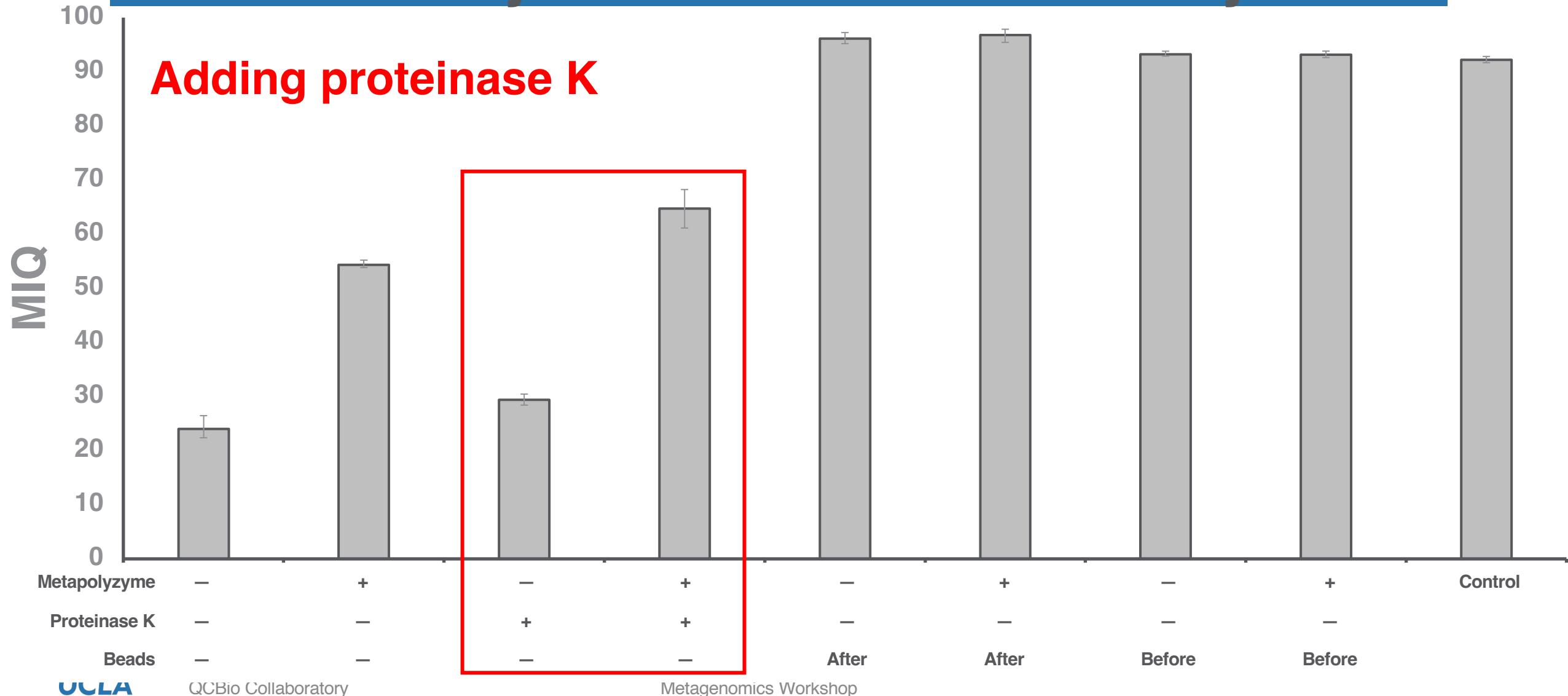
Chemical/Enzymatic and Mechanical Lysis

Use of Metapolyzyme

MIQ



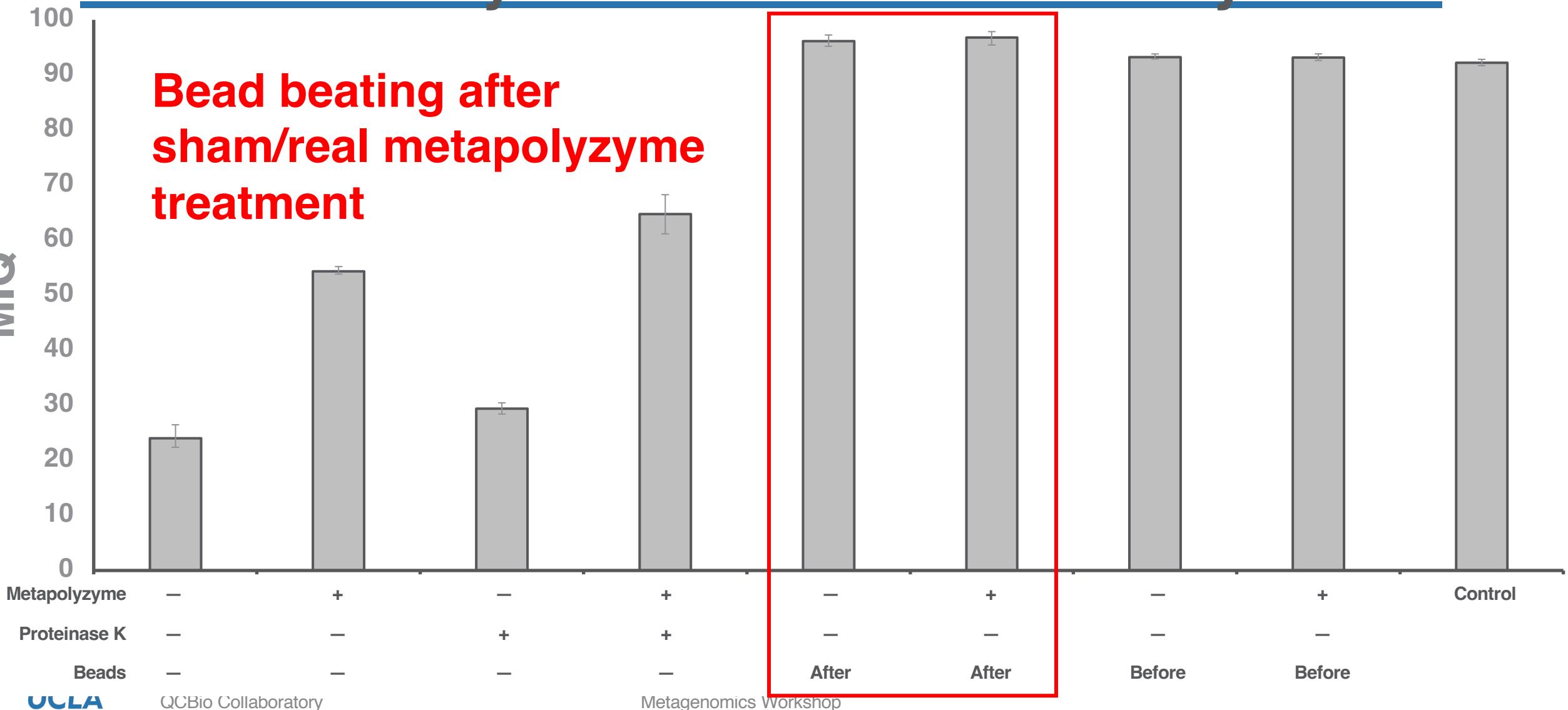
Chemical/Enzymatic and Mechanical Lysis



Chemical/Enzymatic and Mechanical Lysis

Bead beating after
sham/real metapolyzyme
treatment

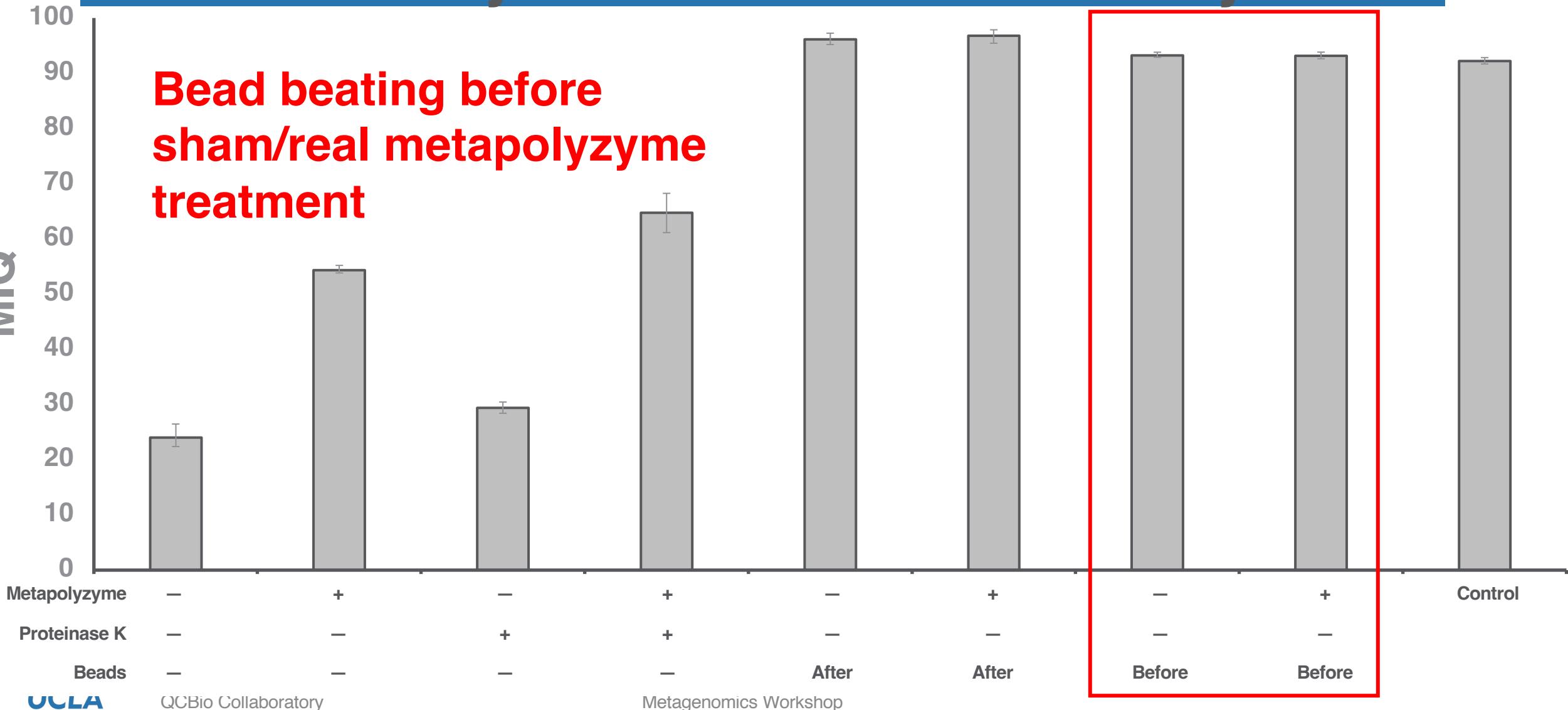
MIQ



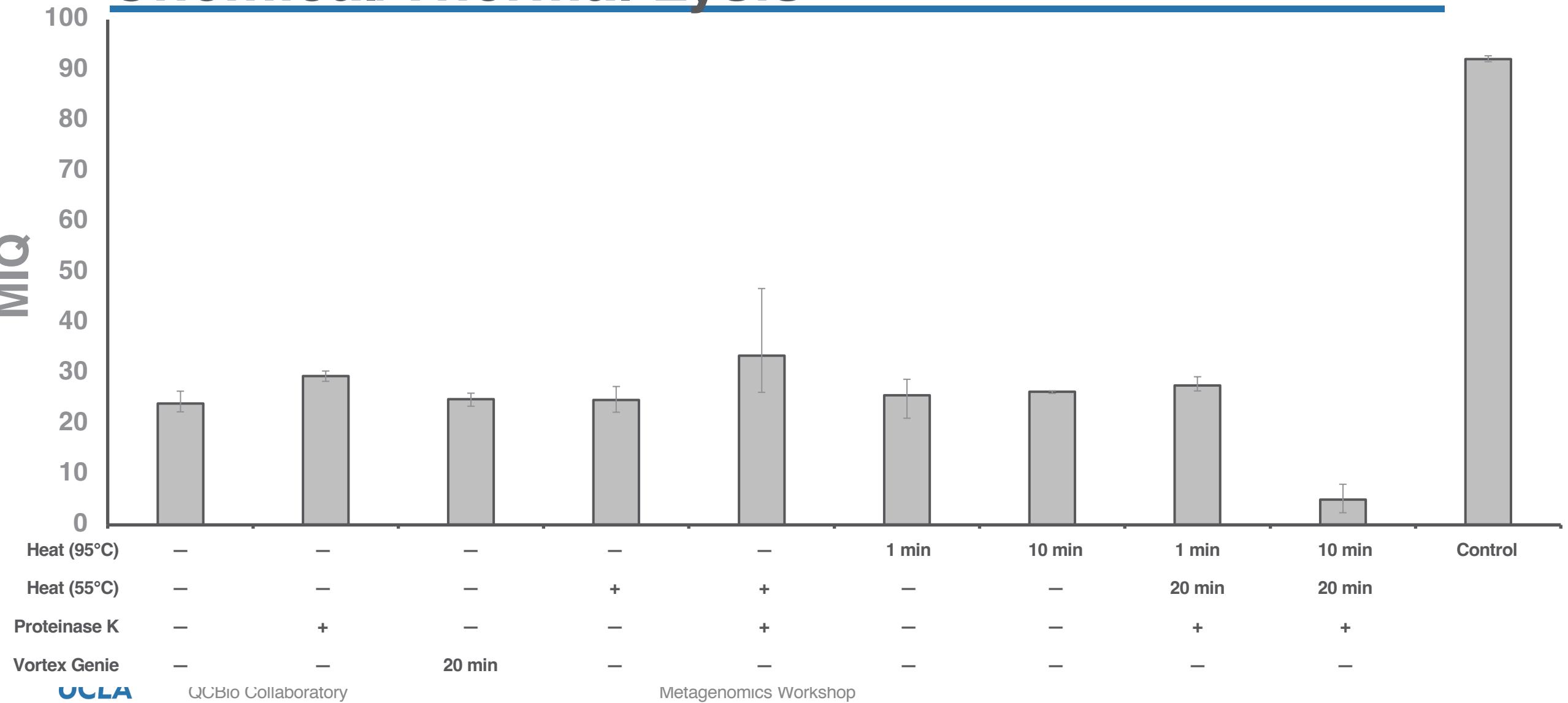
Chemical/Enzymatic and Mechanical Lysis

Bead beating before
sham/real metapolyzyme
treatment

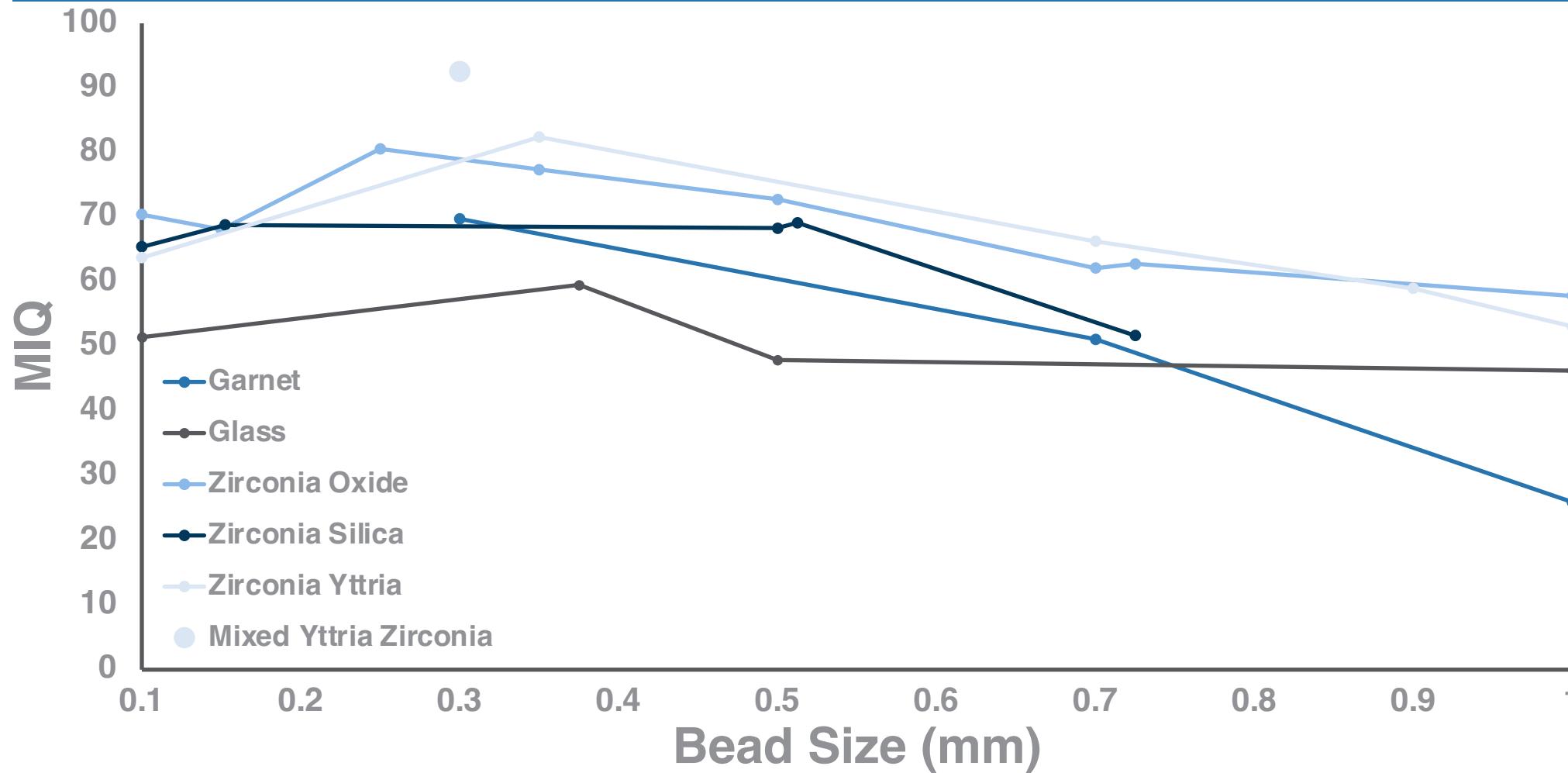
MIQ



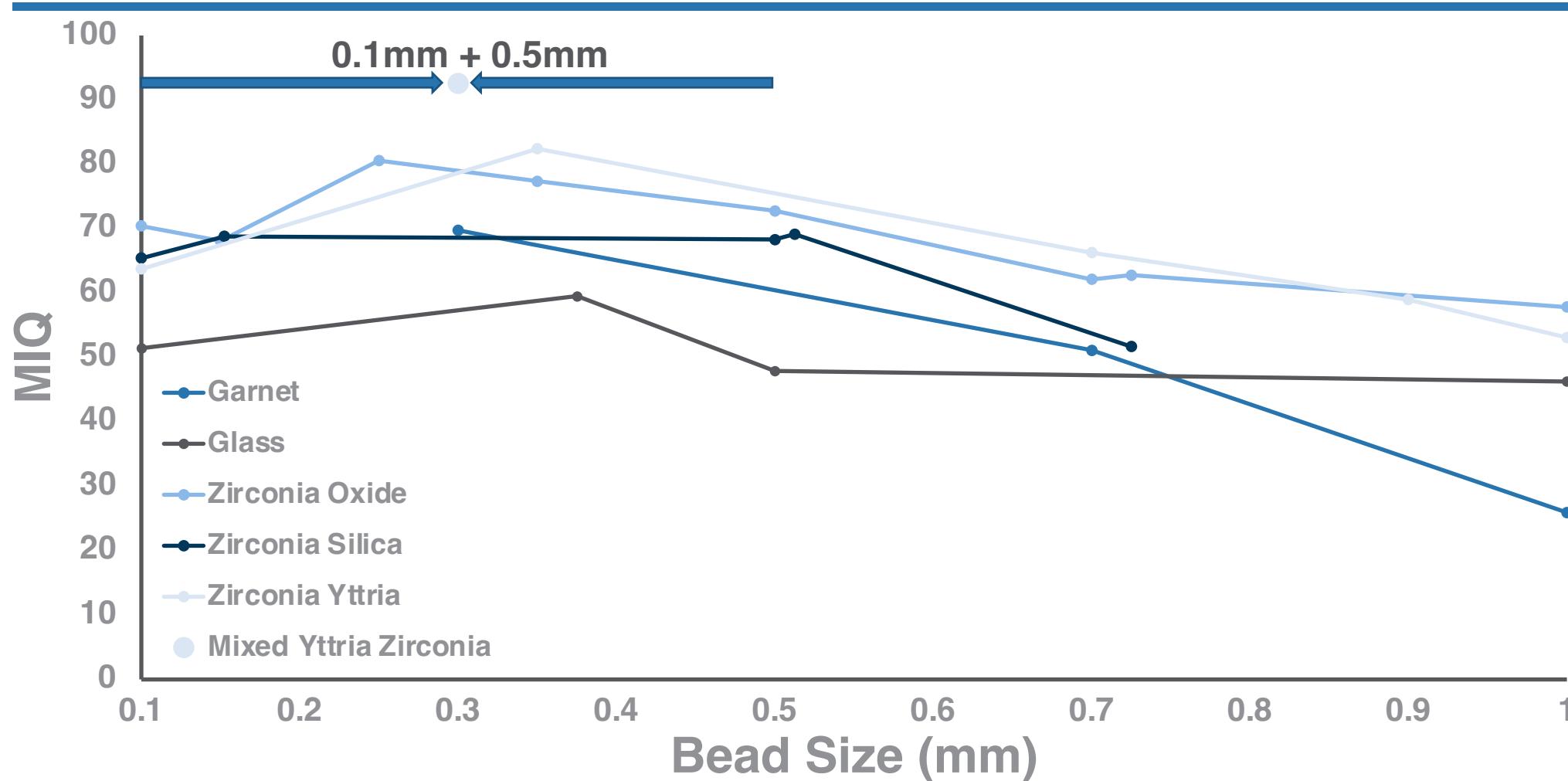
Chemical/Thermal Lysis



Bead Size Has a Strong Effect With an Optimal Range



Mixed Beads From Both Ends of the Optimal Range Are Better Than Individual Sizes

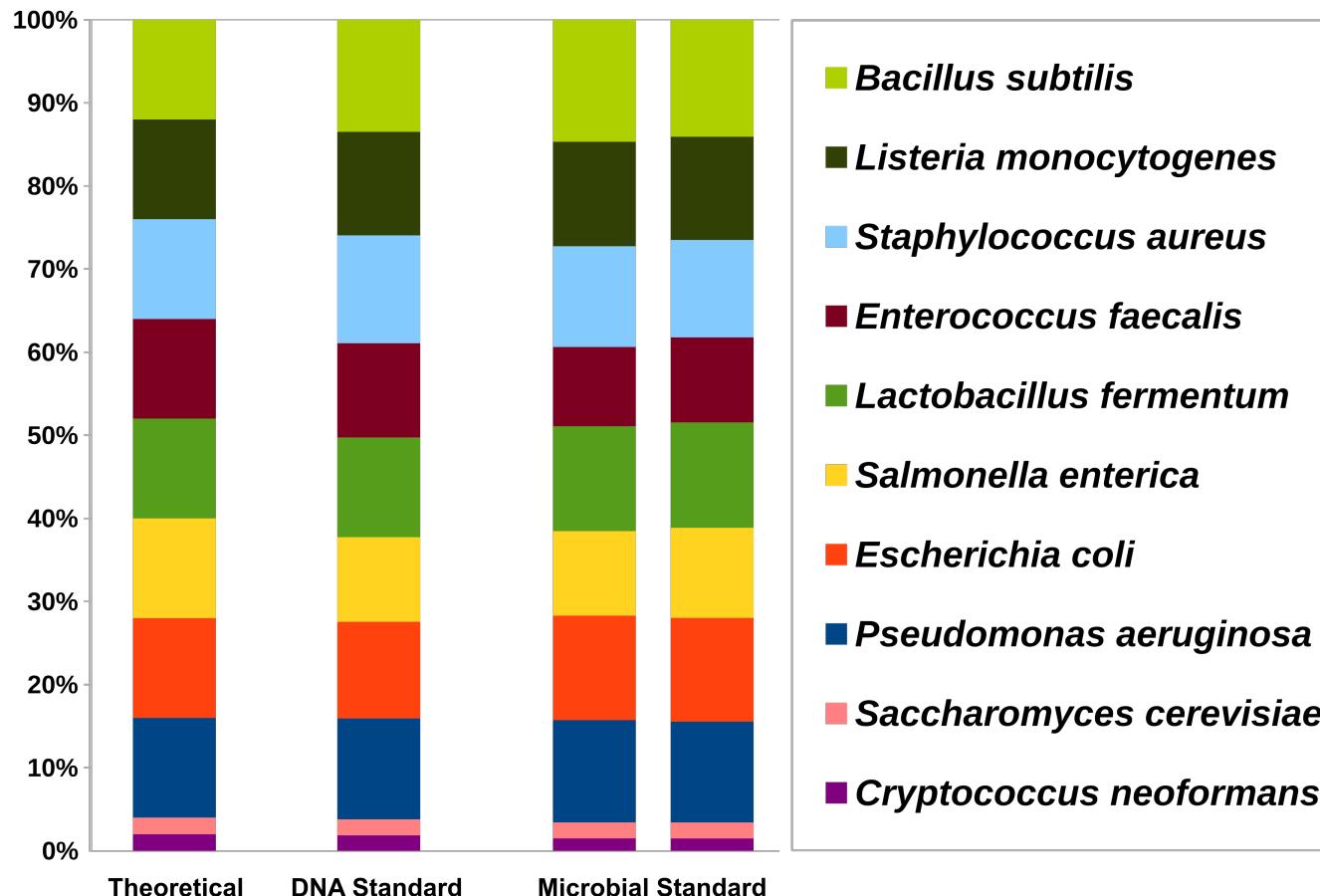


Topic

WHAT EFFECT DO BEAD BEATERS HAVE?

A Standard Ground Truth For Measurement

ZymoBIOMICS Mock Microbial Community

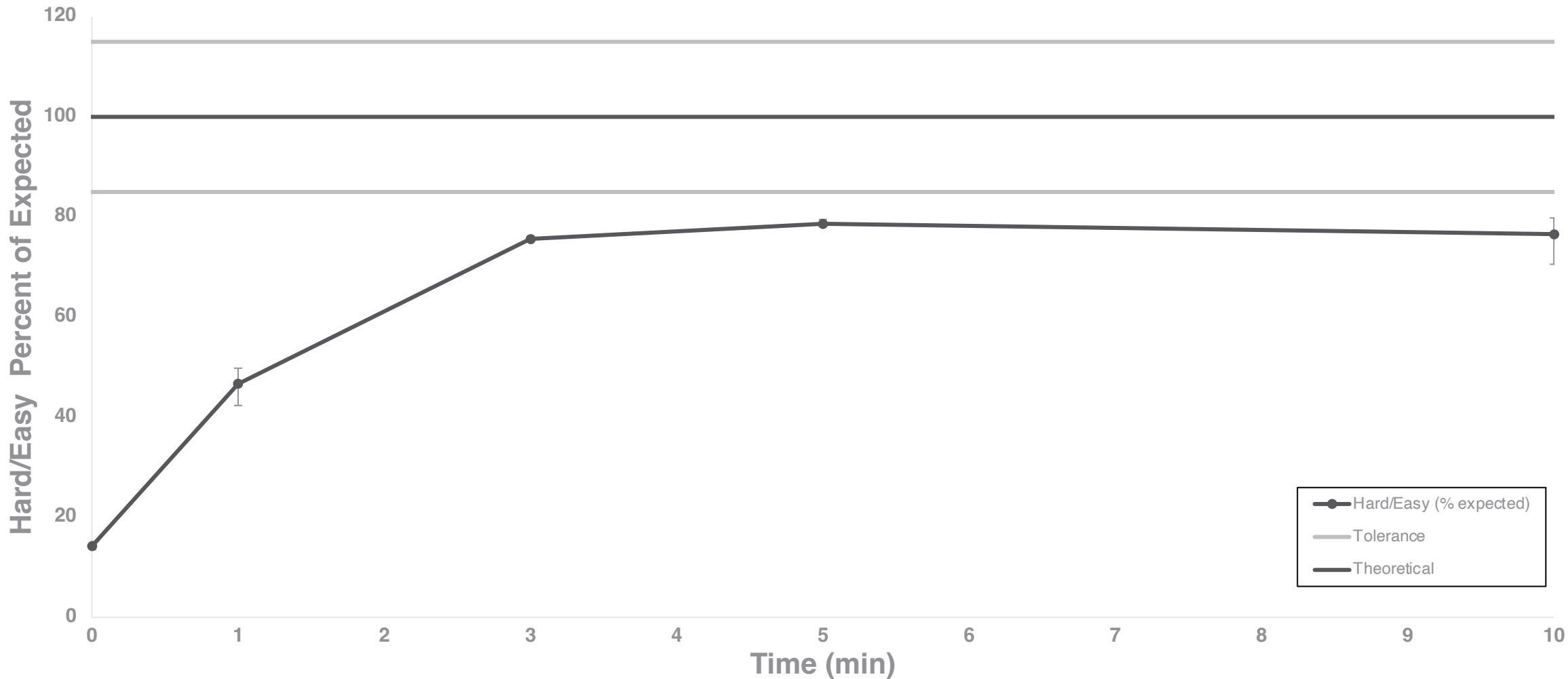


Gram Positive (Hard)

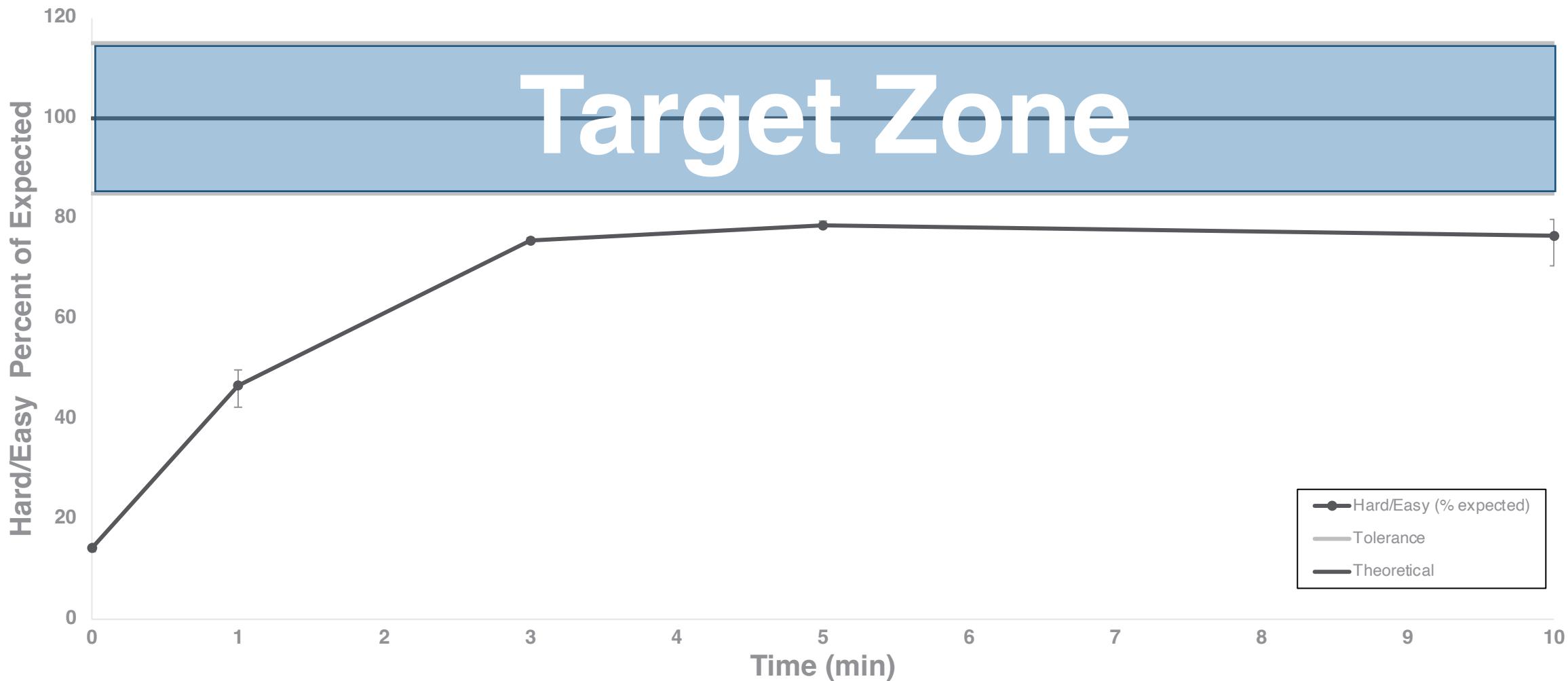
Gram Negative (Easy)

Fungus (Hard)

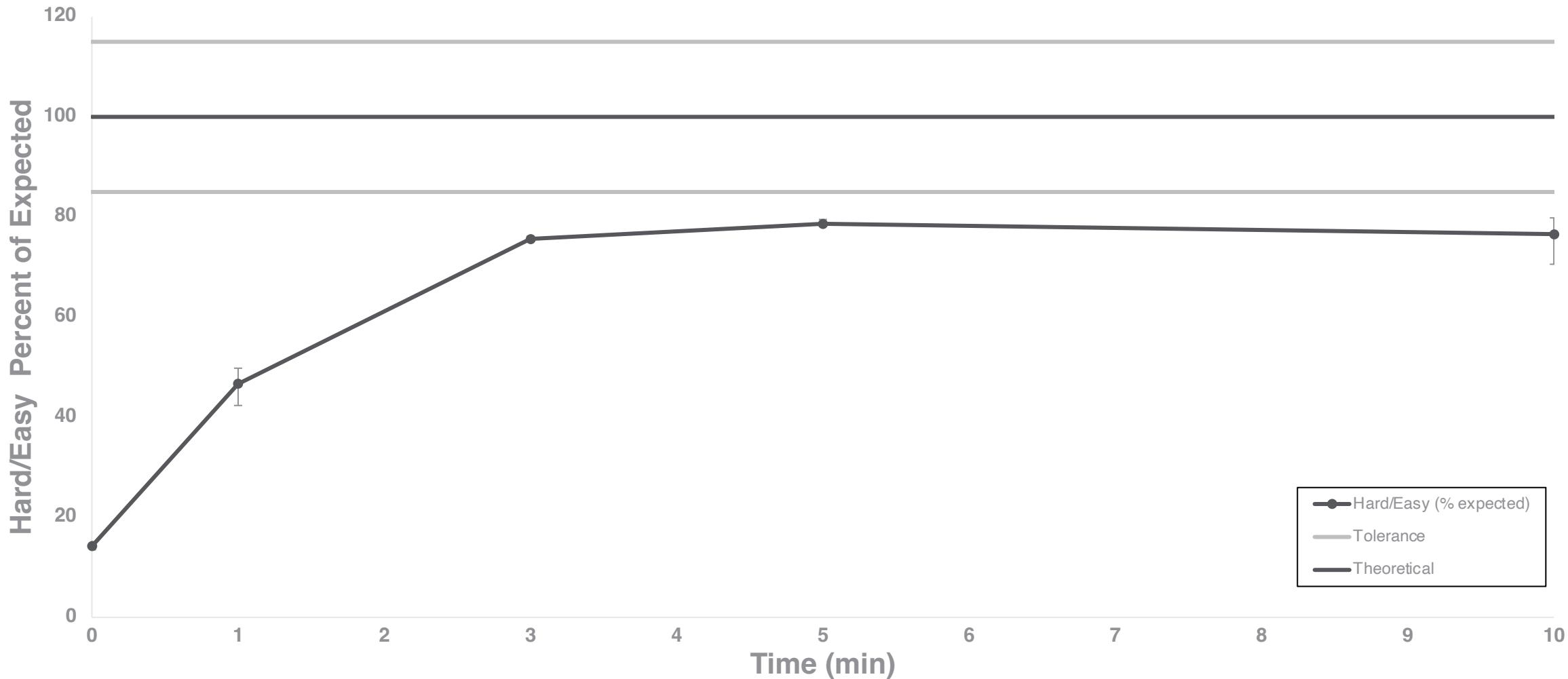
MPBio FastPrep 24



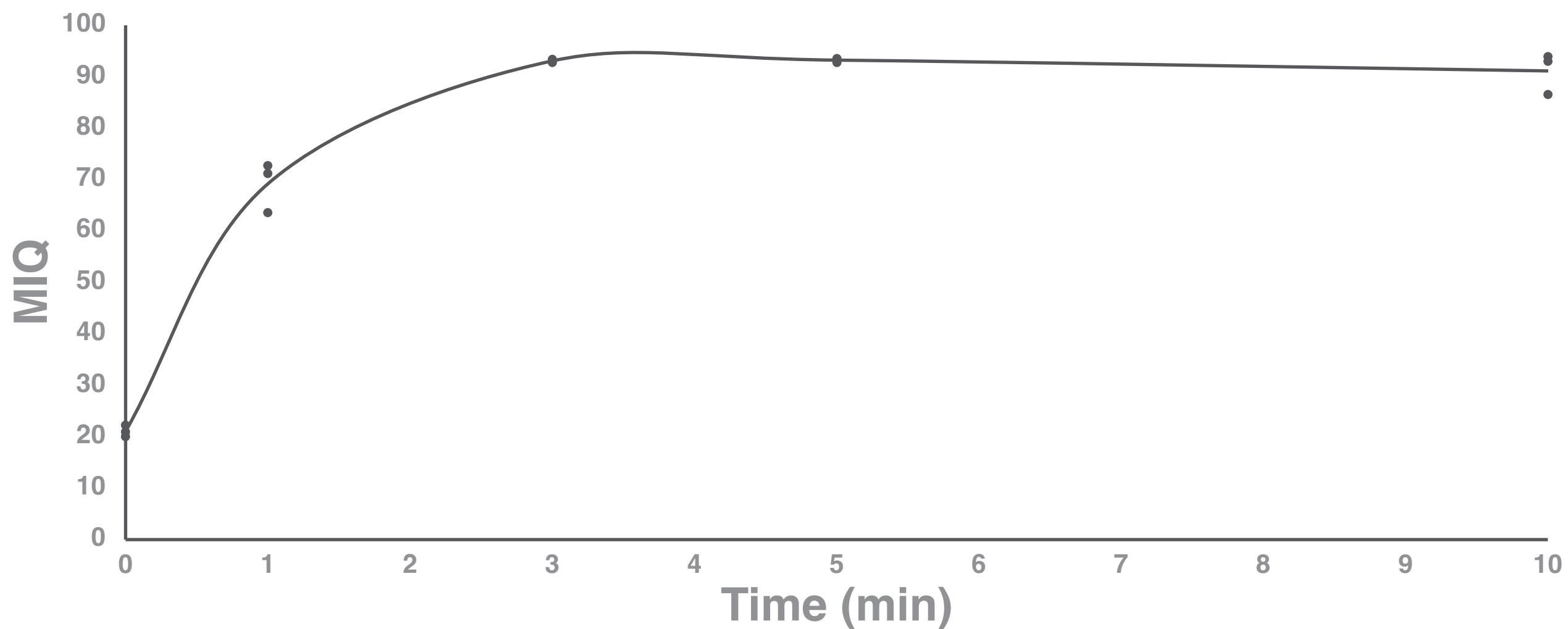
MPBio FastPrep 24



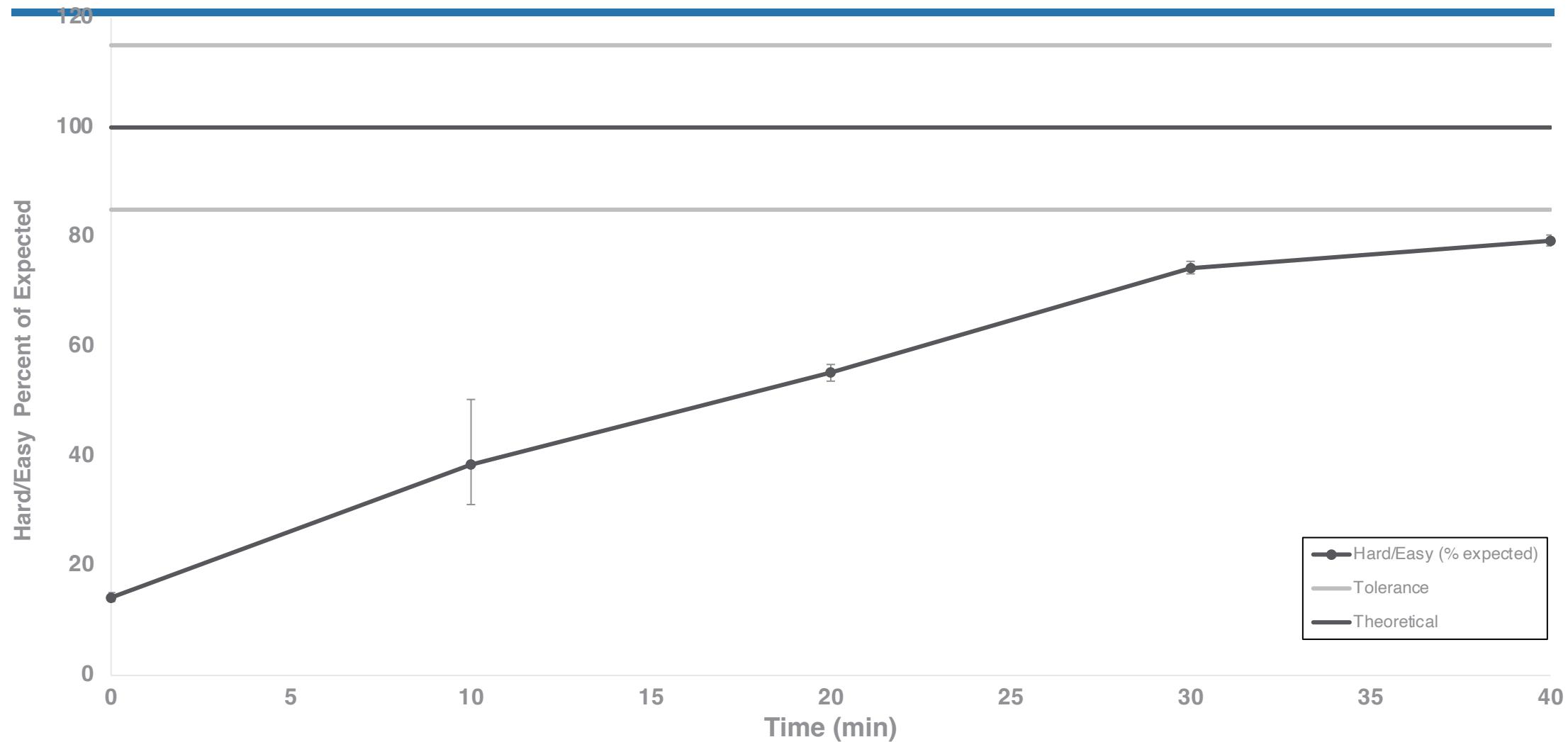
MPBio FastPrep 24



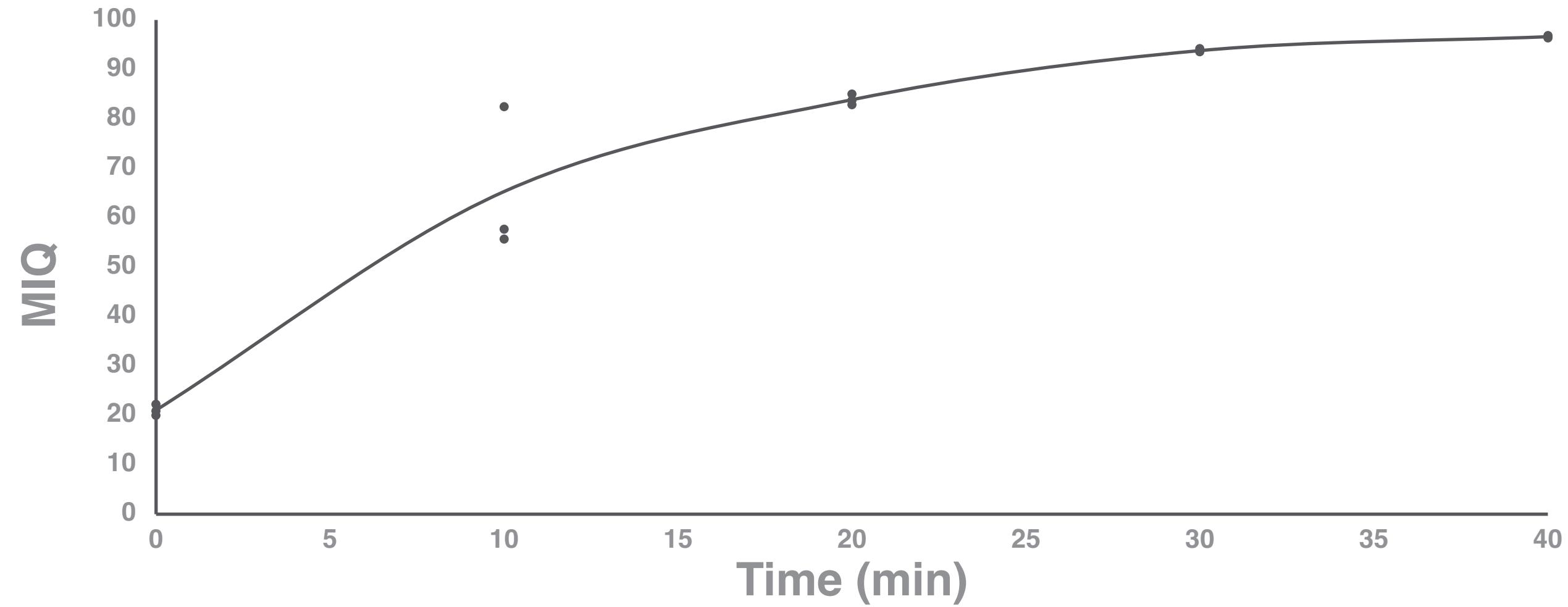
MP Bio Fastprep 24



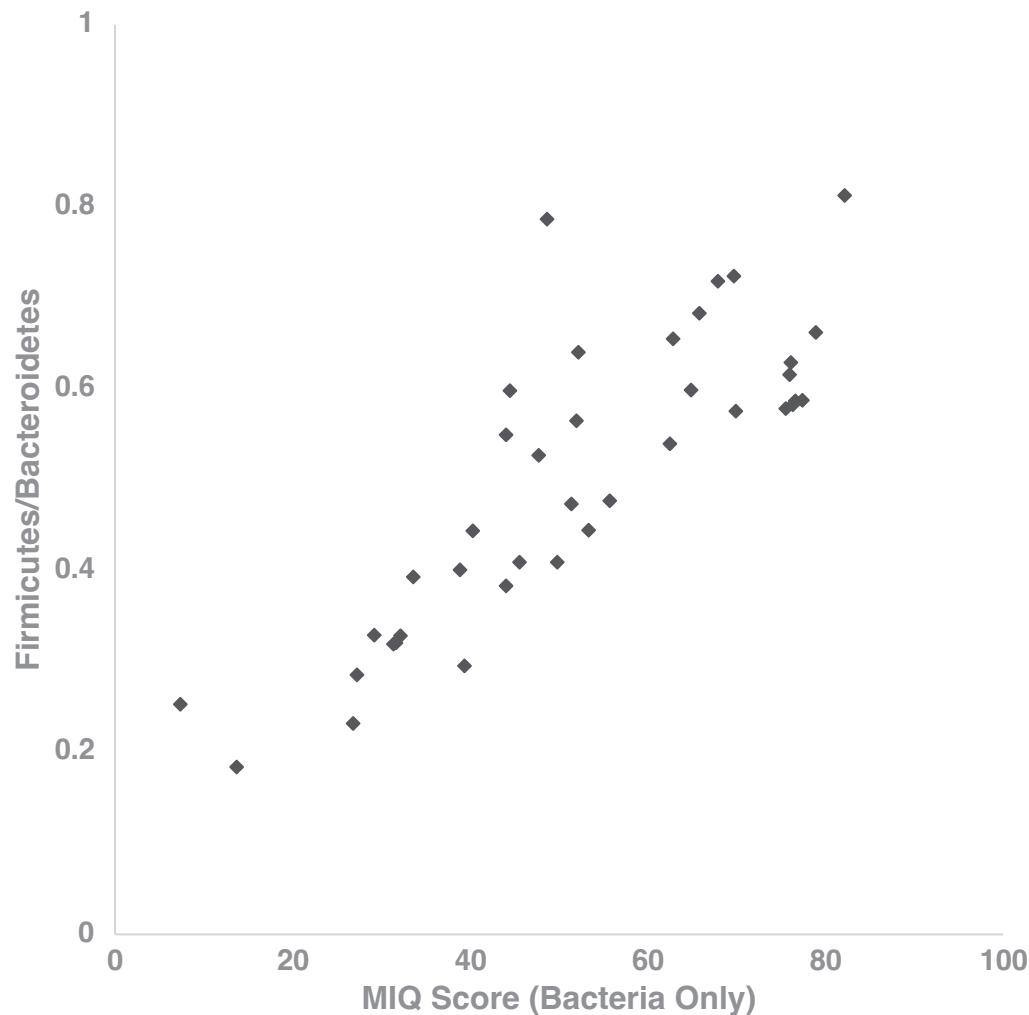
Vortex Genie + Adapter



Vortex Genie + Adapter



How Does This Affect My Results? F/B ratio at least is changed

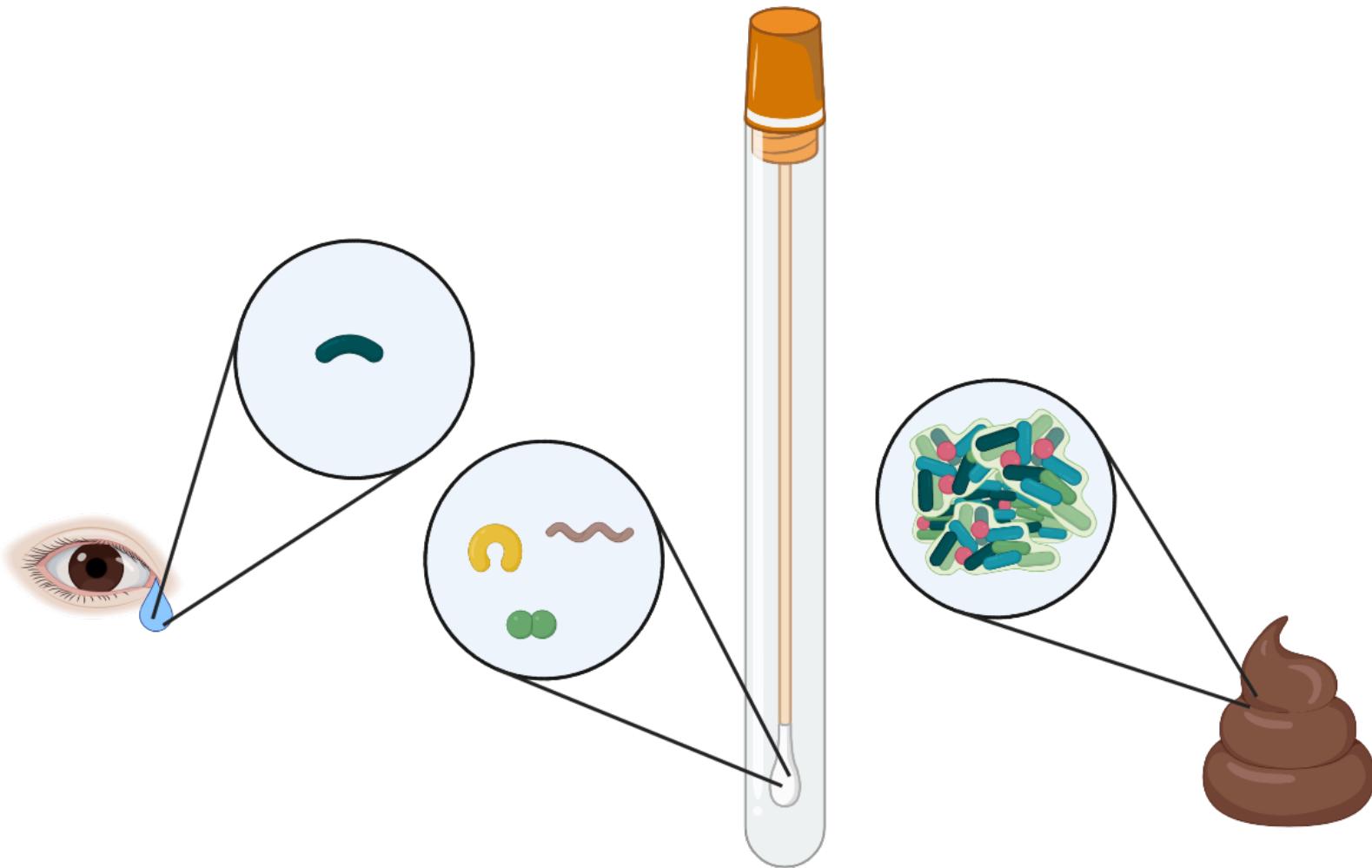


- Bacterial MIQ Score calculated using ZymoBIOMICS mock community standard
- Firmicutes/Bacteroidetes ratio calculated on a repeated stool sample with identical processing, other than bead matrix
- Each point represents the mean of at least three replicates using a different bead matrix, with other processing steps identical
- Pearson Correlation = 0.835, $p = 3.7 \times 10^{-11}$

Topic

SAMPLE COLLECTION AND PRESERVATION

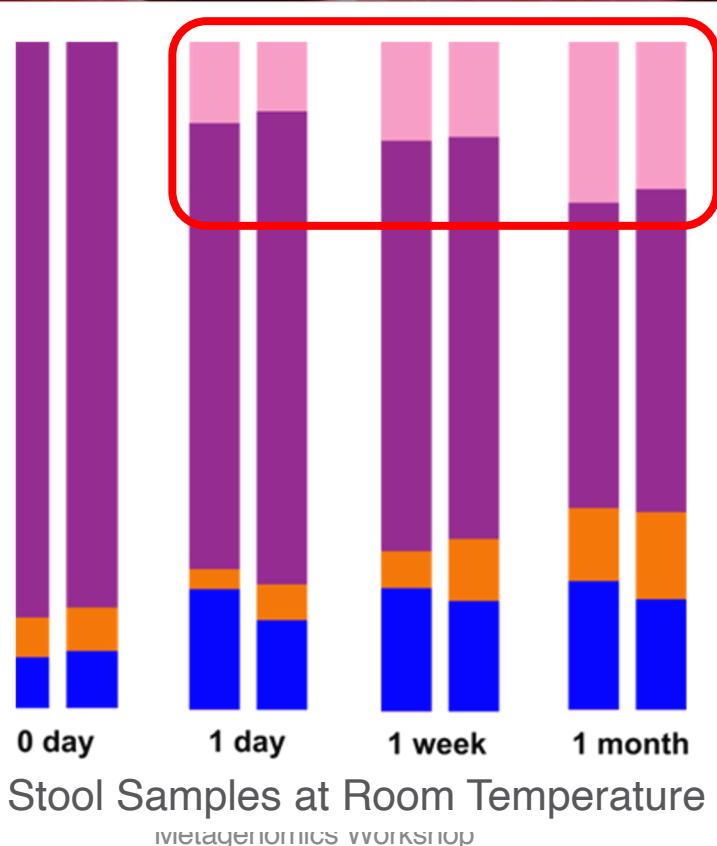
Bioburden: AKA the Kit-ome



A Microbiome Is a Living Sample, Responsive to Environmental Conditions



American Gut: an Open Platform for Citizen Science Microbiome Research



May/June 2018 Vol 3 (3) e00031-18

UCLA

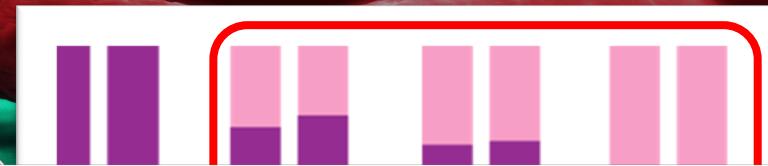
QCBio Collaboratory

metagenomics workshop

A Microbiome Is a Living Sample, Responsive to Environmental Conditions



American Gut: an Open Platform for Citizen Science Microbiome Research



"Removal of bacterial blooms. [...]

Escherichia coli and a few other taxa grow in transit, so based on data from controlled-storage studies as previously described (15), we removed sub-operational taxonomic units (sOTUs) (16) (median of 7.9% of sequences removed per sample) shown to bloom."

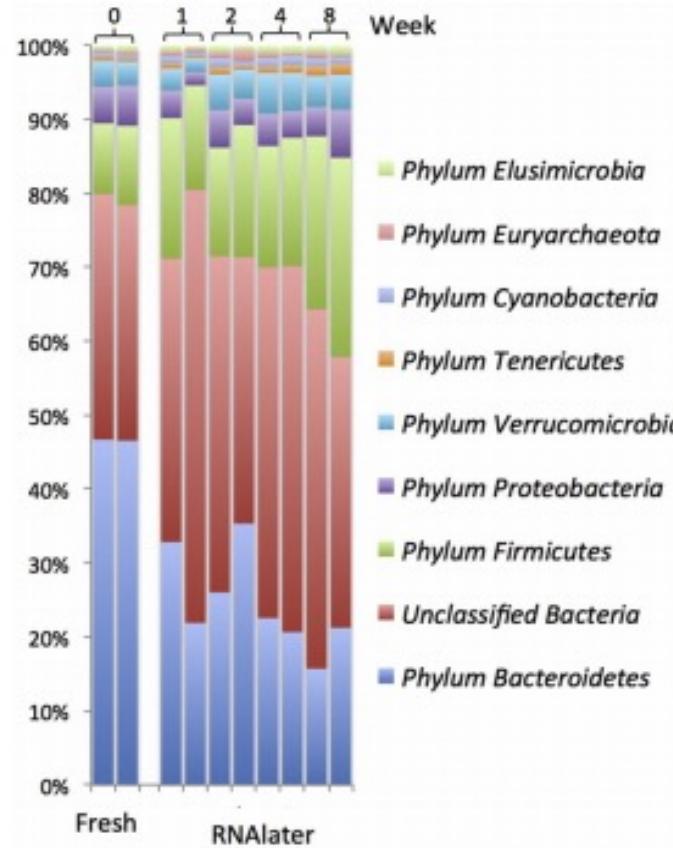
May/June 2018 Vol 3 (3) e00031-18

UCLA

QCBio Collaboratory

0 day 1 day 1 week 1 month
Stool Samples at Room Temperature
metagenomics workshop

Use of RNA Preservative Causes a Loss of Bacteroidetes



Modified from Hale, V. L., Tan, C. L., Knight, R., & Amato, K. R. *Journal of Microbiological Methods*, 113, 16-26.

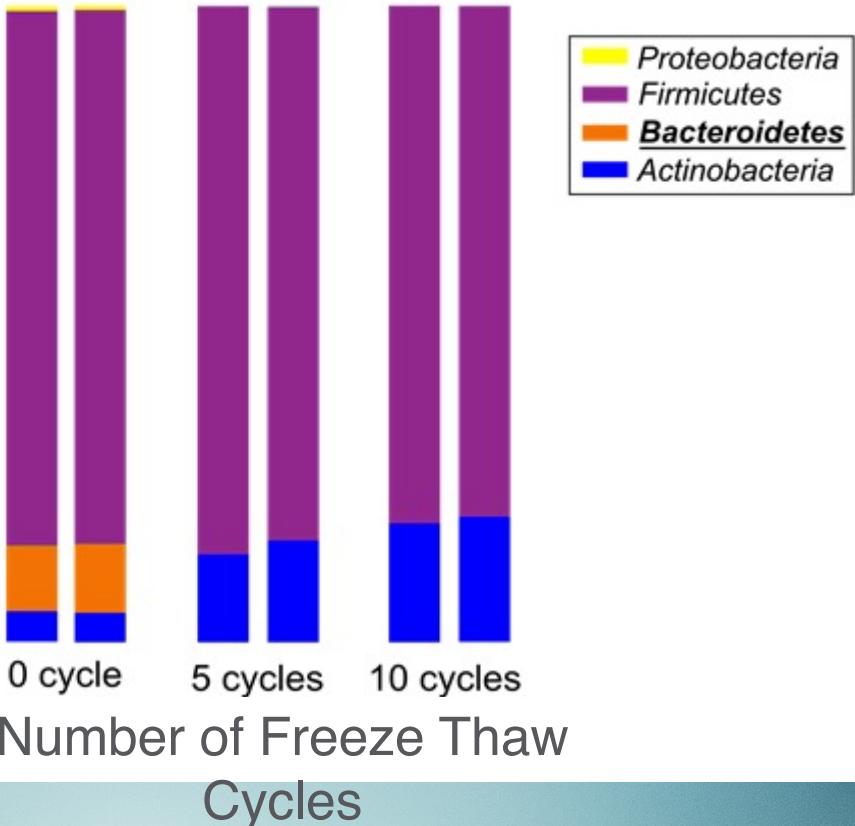


QCBio Collaboratory

Metagenomics Workshop

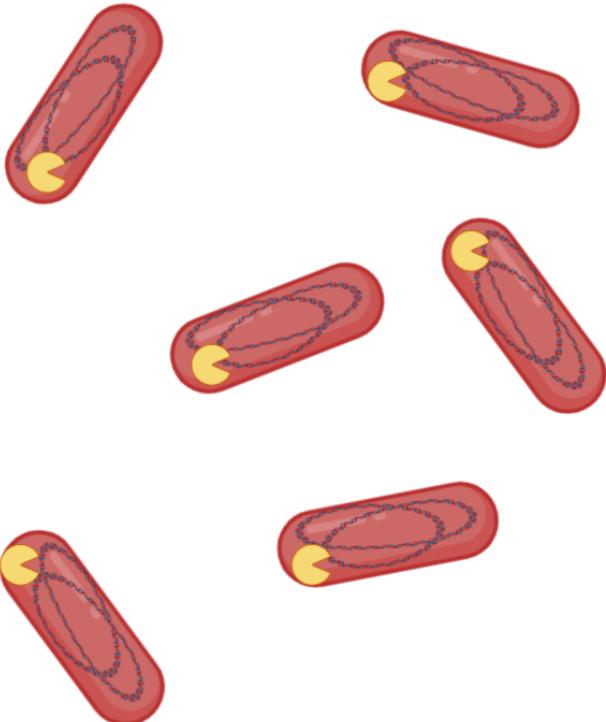
Freeze/Thaw Causes a Loss of Bacteroidetes

Stool **WITHOUT** Preservative

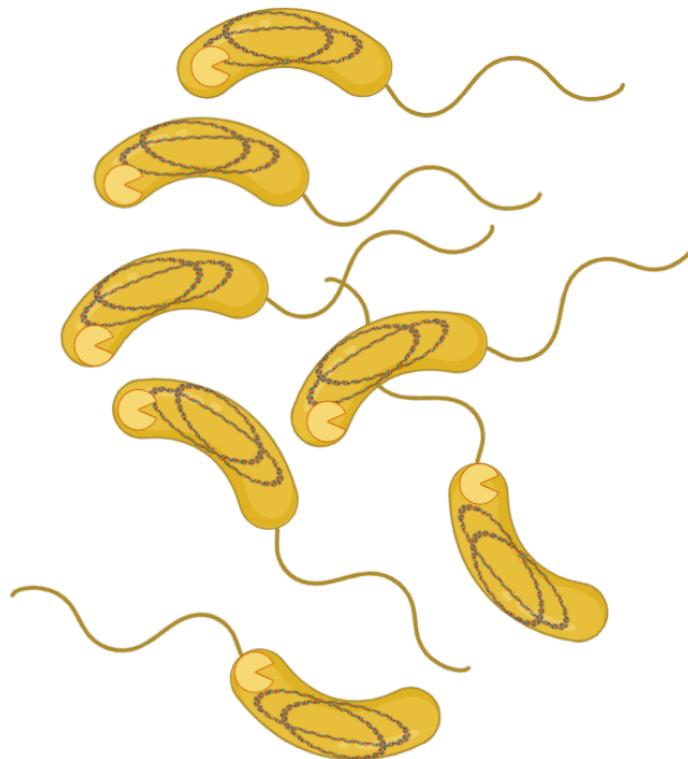


A Tale of Two Microbes

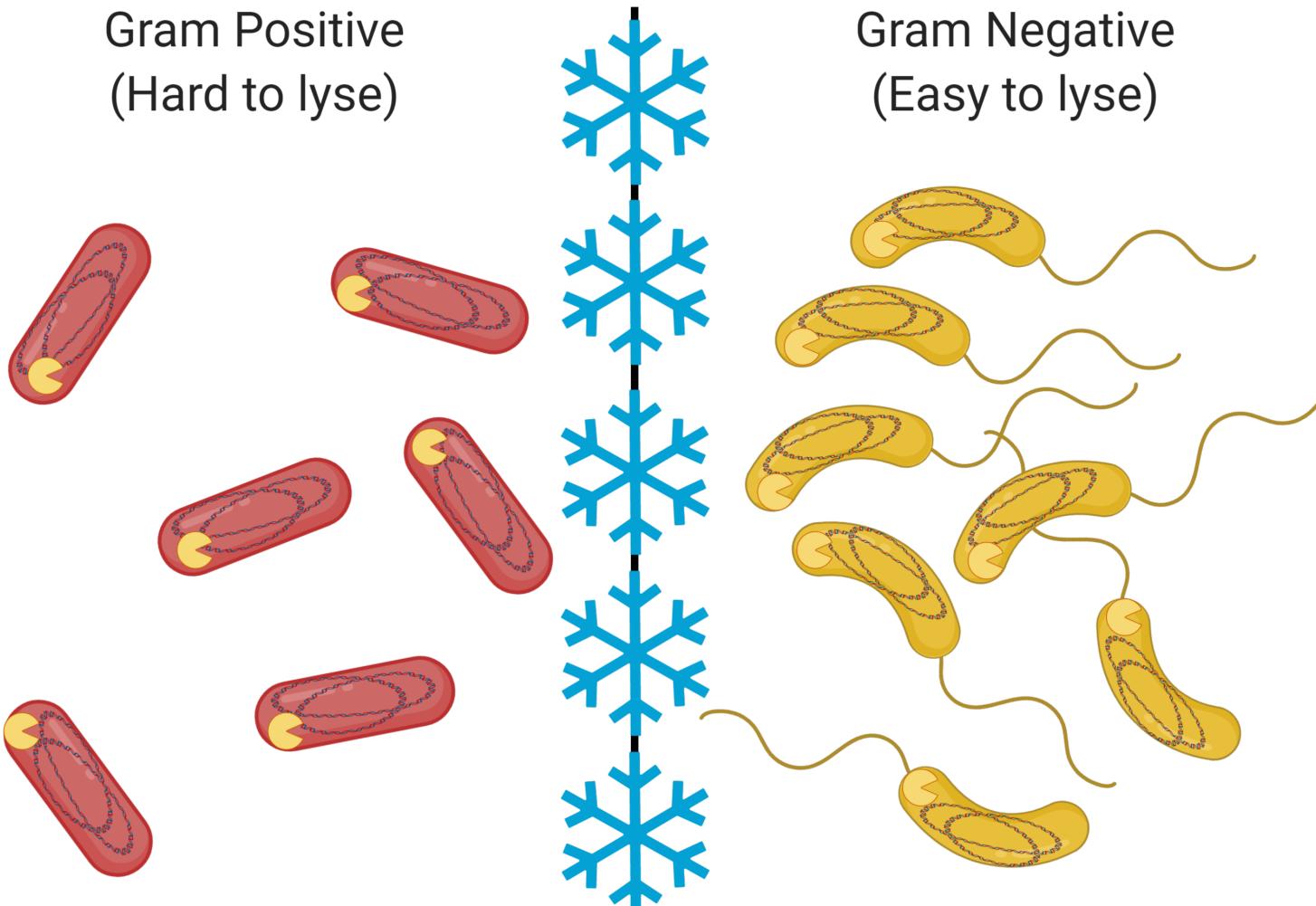
Gram Positive
(Hard to lyse)



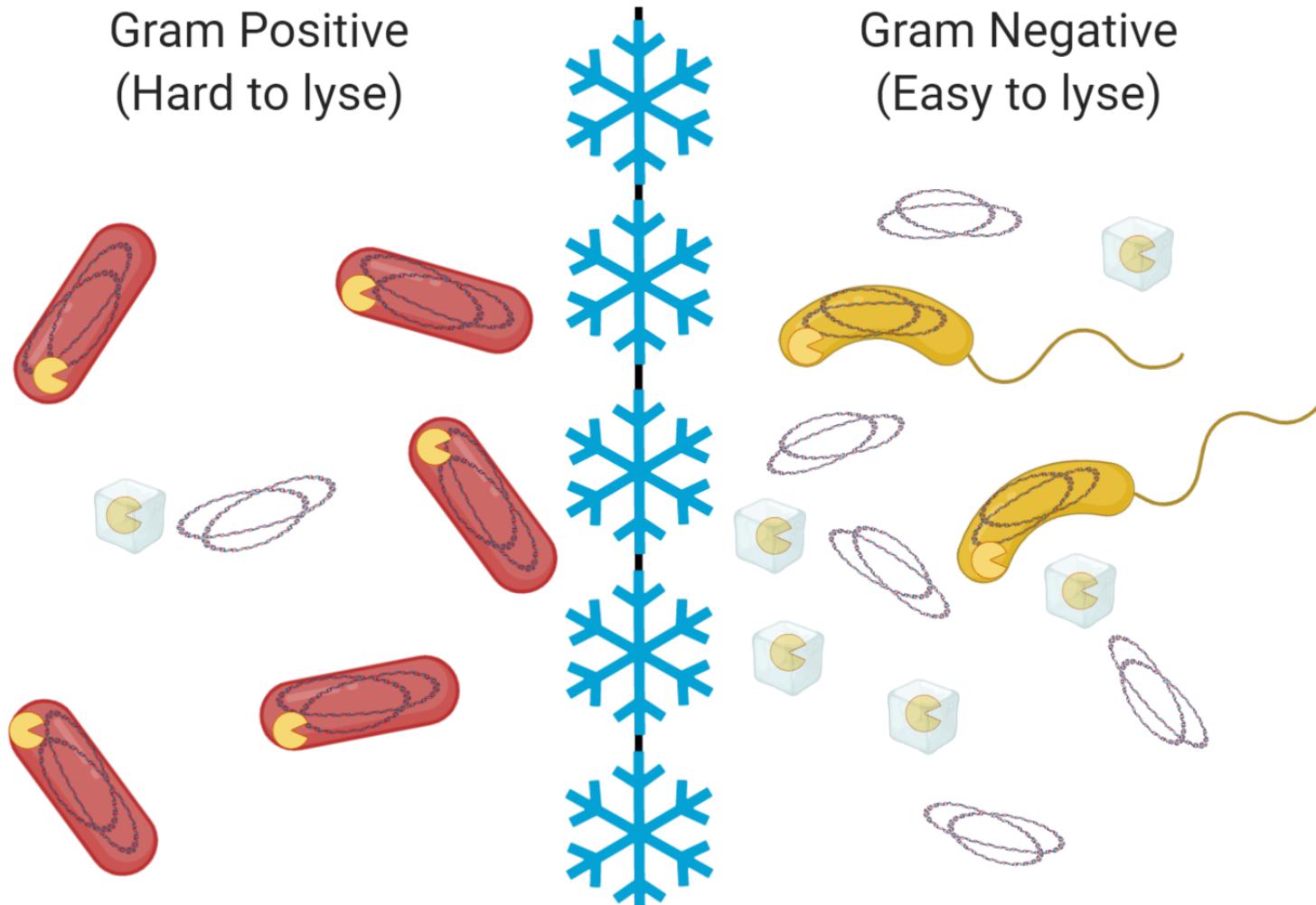
Gram Negative
(Easy to lyse)



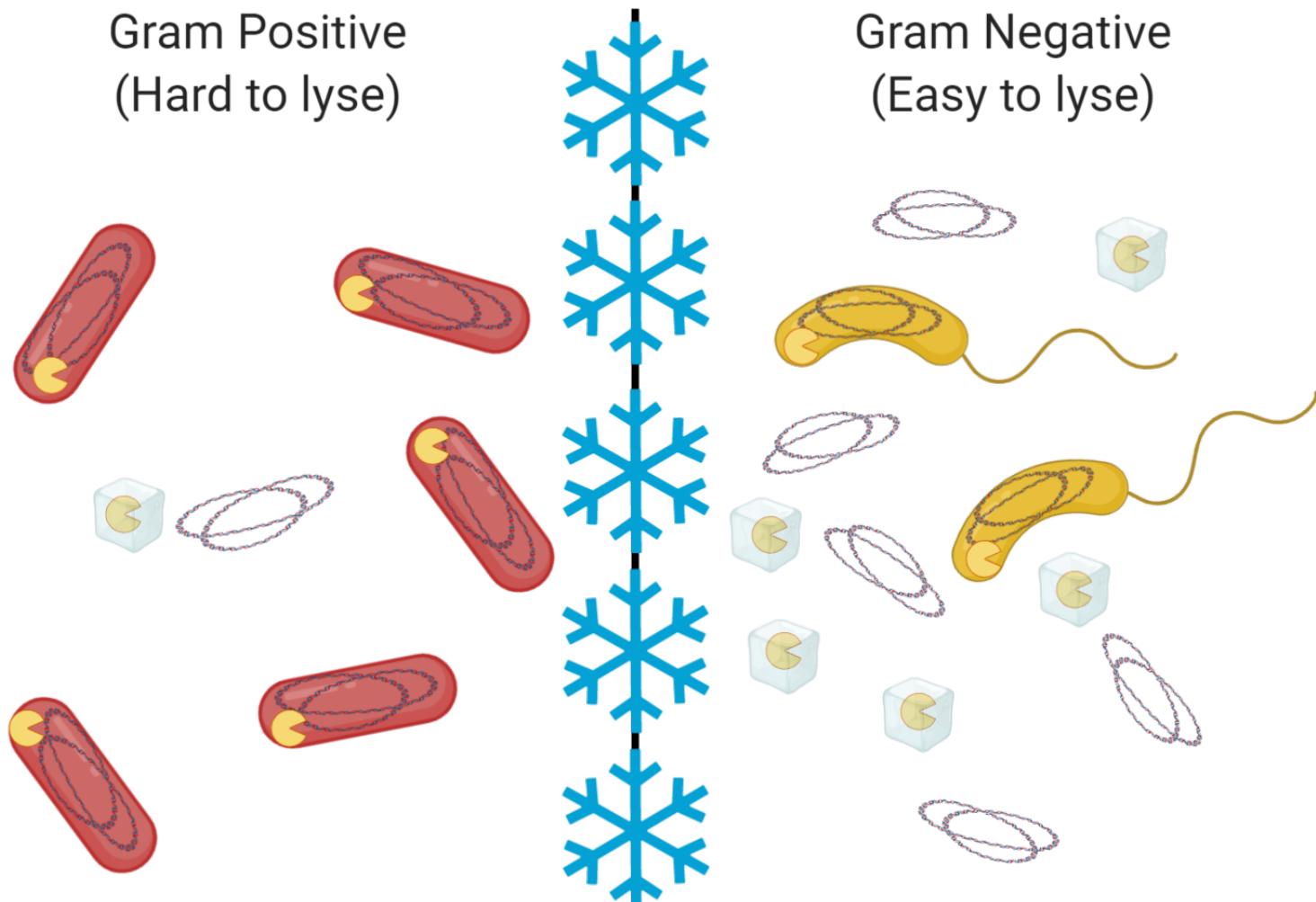
Use of Freezing as a Preservative



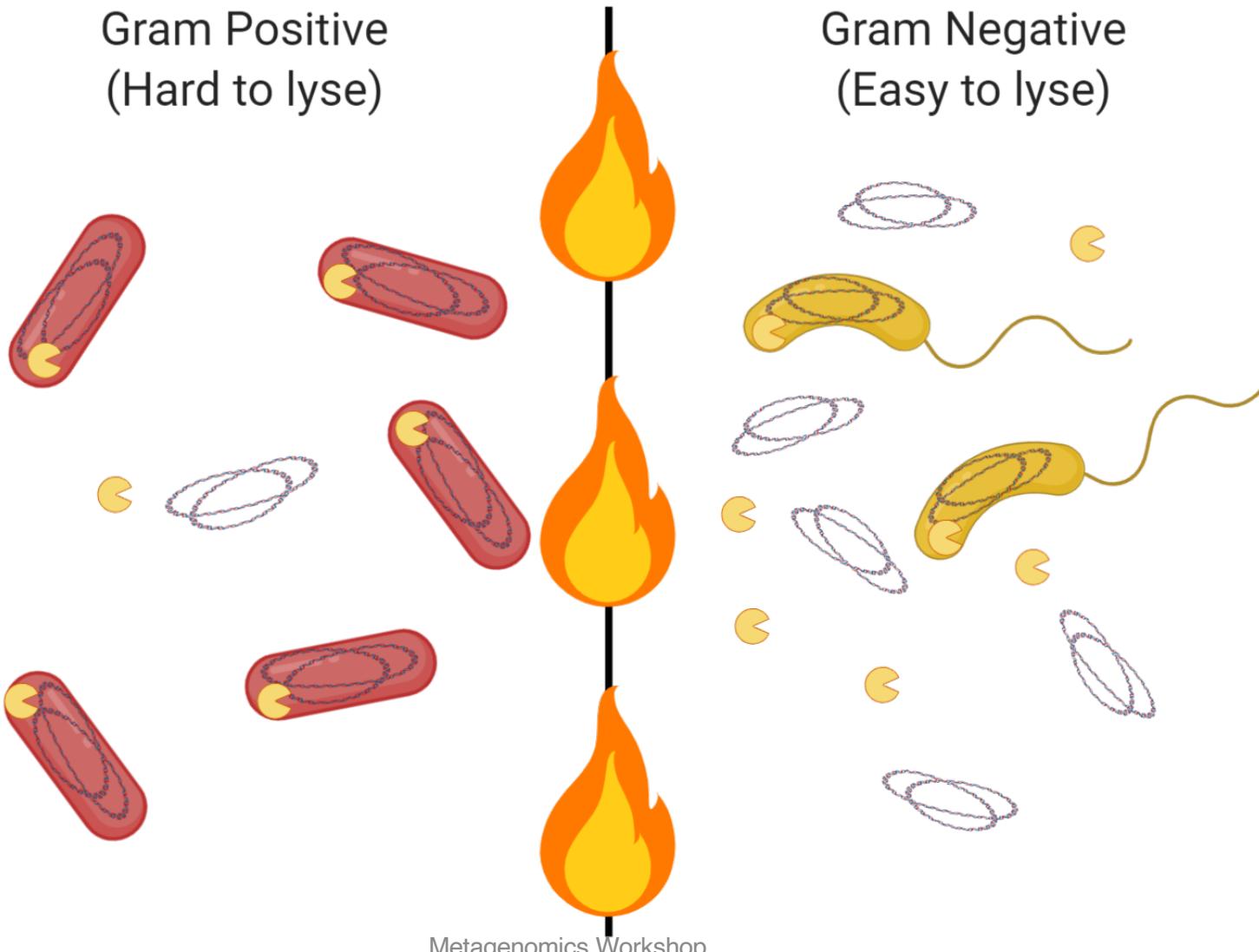
Freezing Lyses More Fragile Bacteria at a Higher Rate



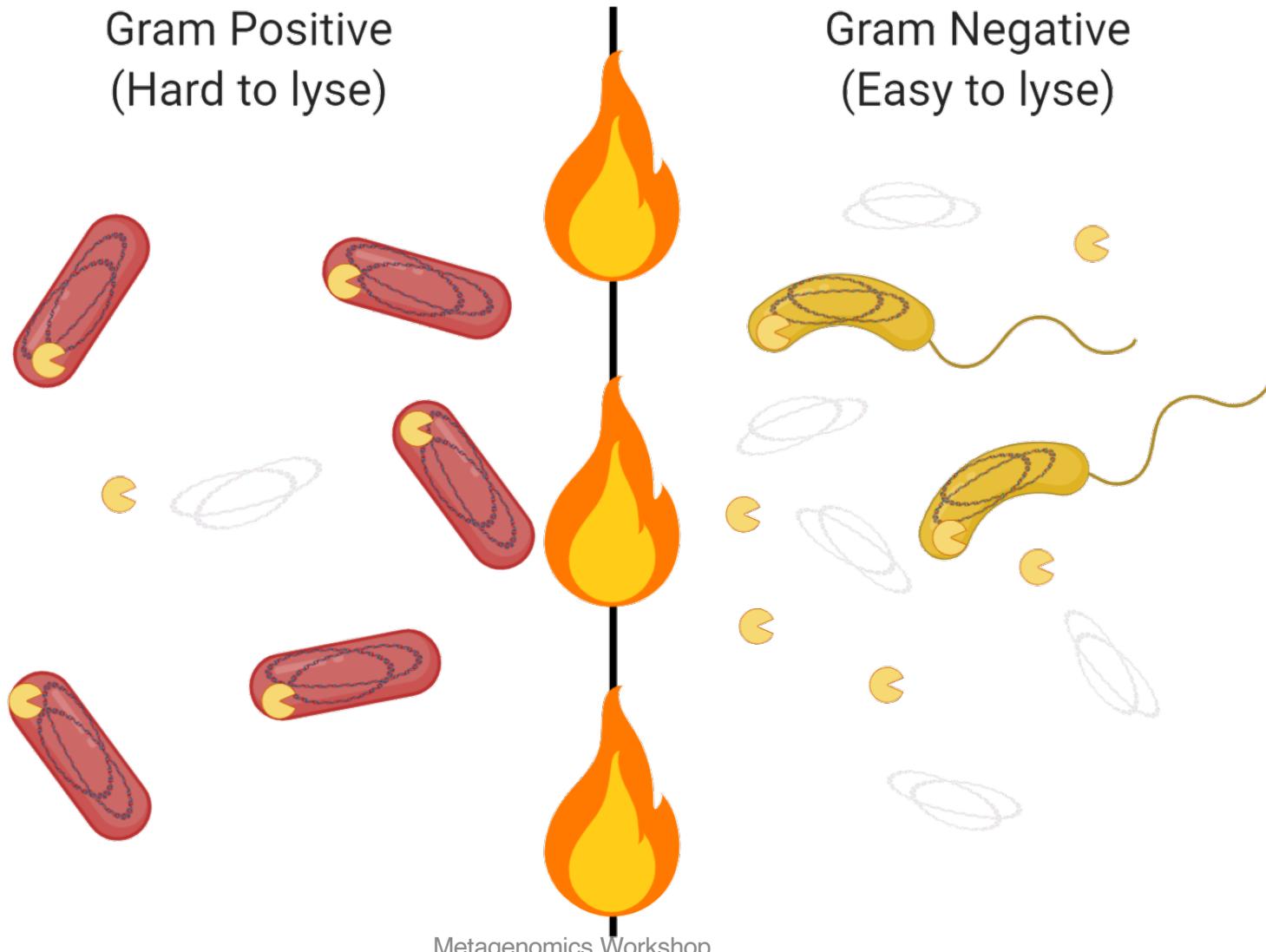
But Freezing Also Inhibits Degradative Enzymes



Thawing Restores Activity to Enzymes



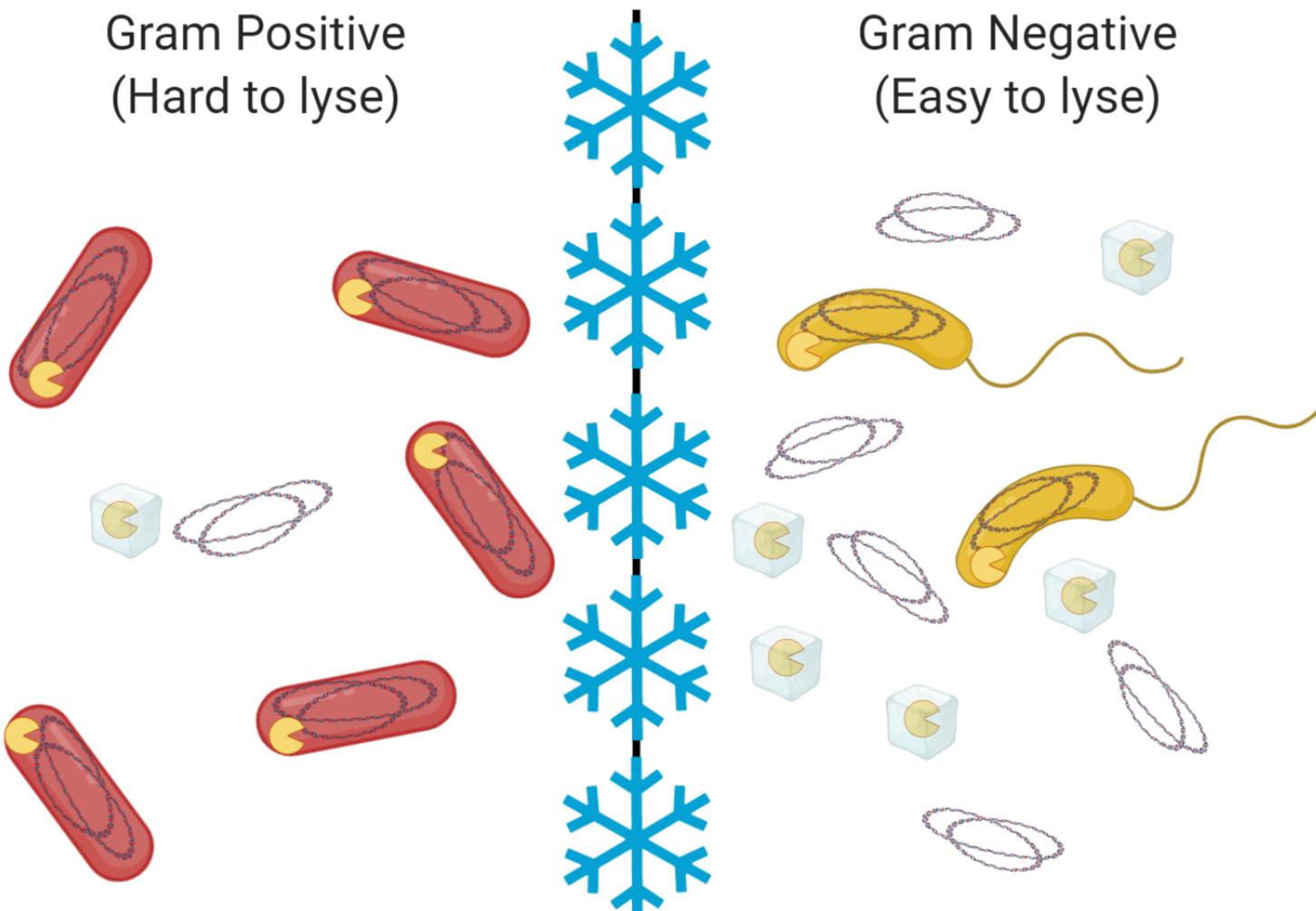
Active Enzymes Degrade Unprotected DNA



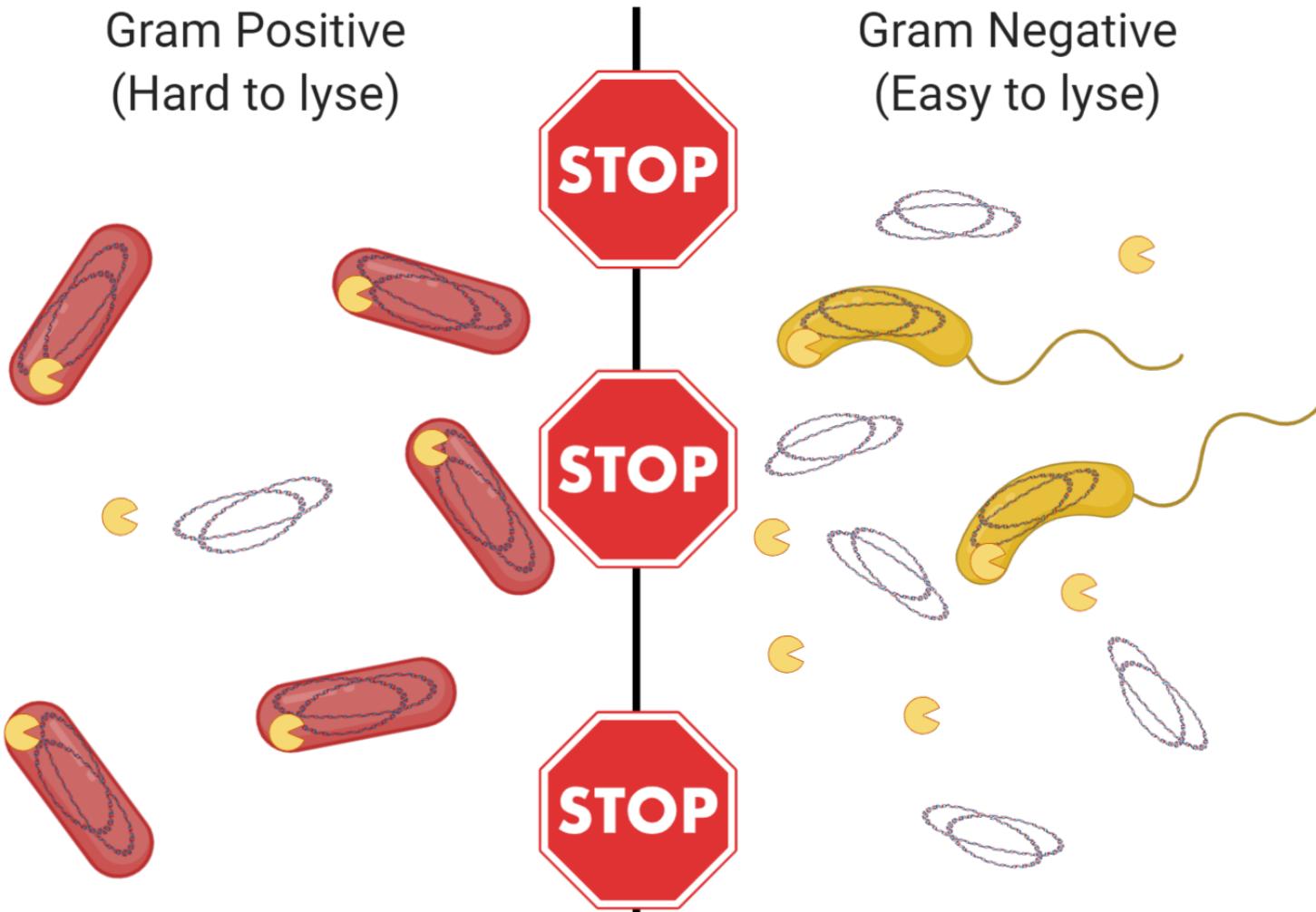
Topic

WILL RAPID DENATURING/INHIBITION OF NUCLEASES HELP?

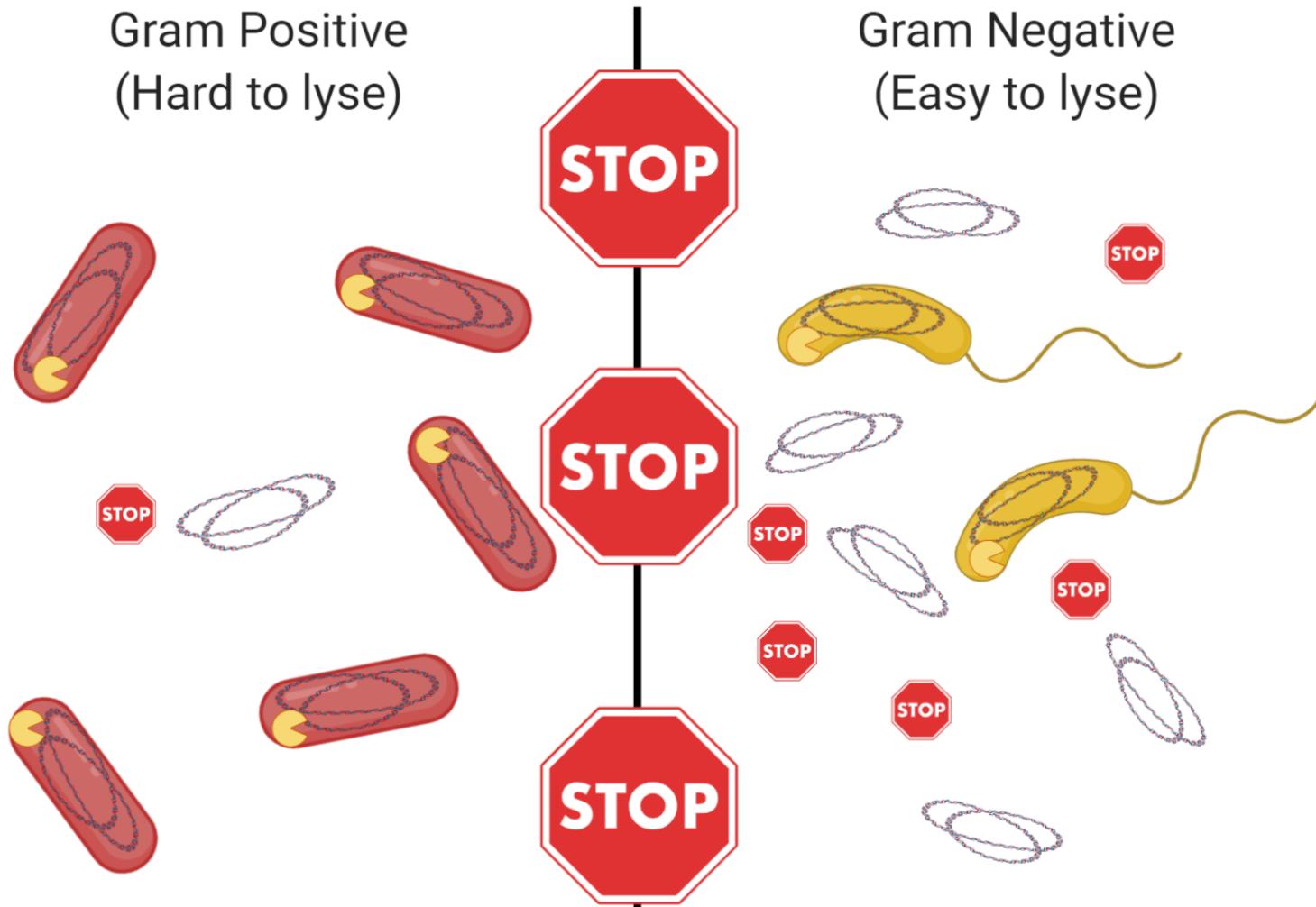
Be Ready For Degradative Enzymes with Inhibitor



Enzymes Thaw in the Presence of Inhibitor...



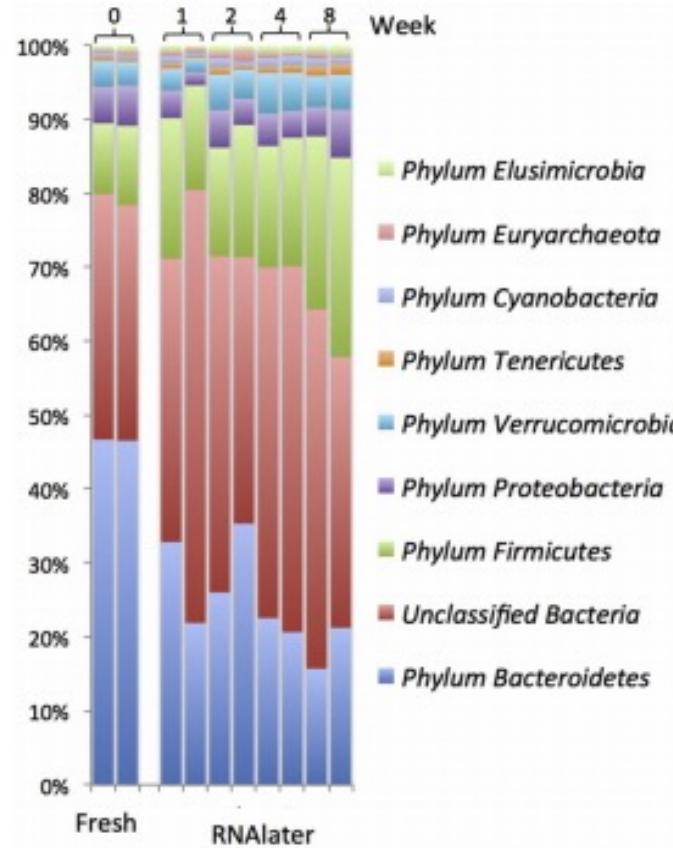
...and Are Stopped Before Significant Degradation



Topic

ARE ALL INHIBITORS SUITABLE FOR THIS?

Use of RNA Preservative Causes a Loss of Bacteroidetes



Modified from Hale, V. L., Tan, C. L., Knight, R., & Amato, K. R. *Journal of Microbiological Methods*, 113, 16-26.

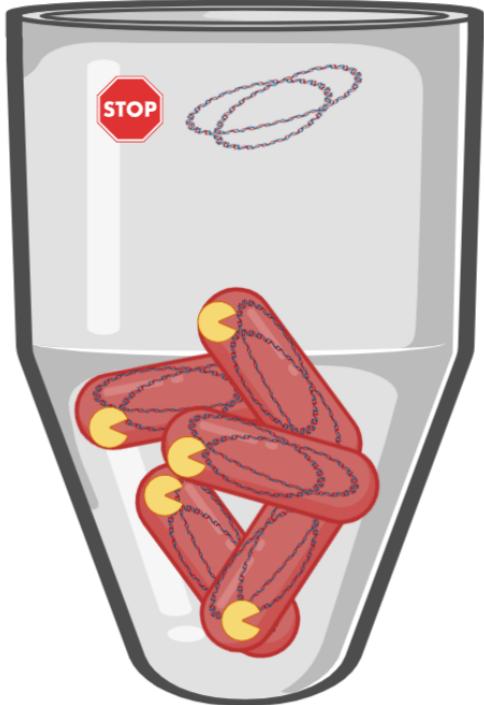


QCBio Collaboratory

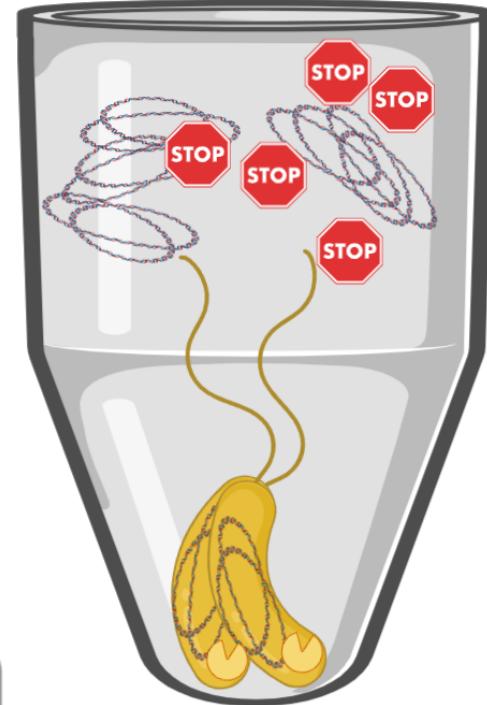
Metagenomics Workshop

If the Inhibitor Cannot Go Through Purification, DNA From Lysed Cells Is Lost

Gram Positive
(Hard to lyse)



Gram Negative
(Easy to lyse)



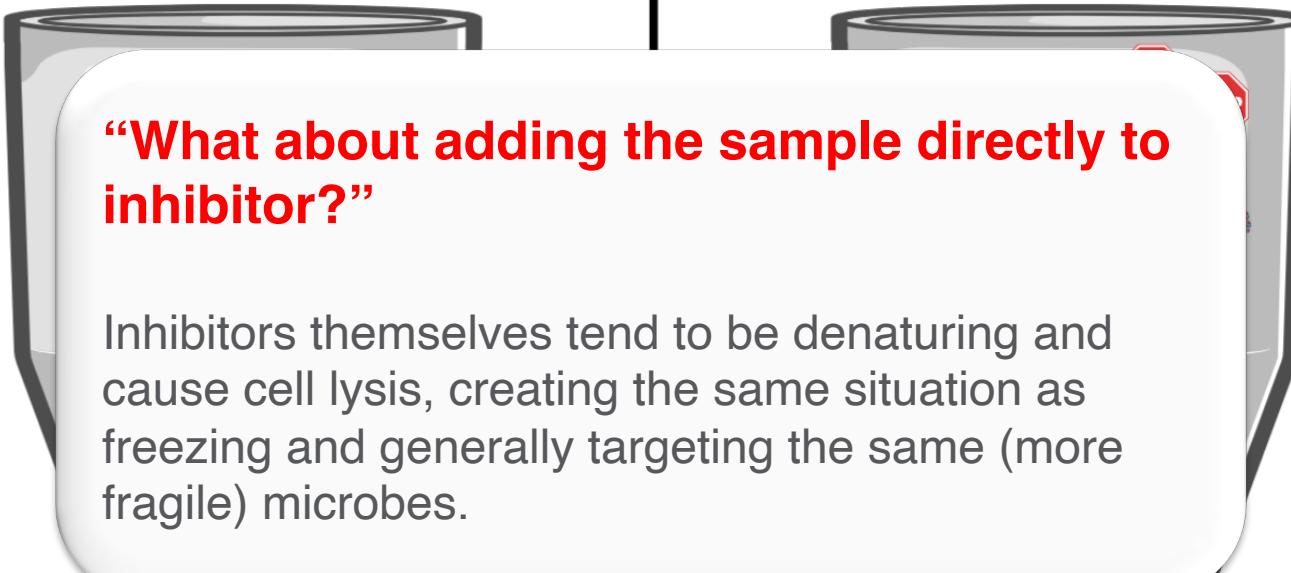
If the Inhibitor Cannot Go Through Purification, DNA From Lysed Cells Is Lost

Gram Positive
(Hard to lyse)

Gram Negative
(Easy to lyse)

“What about adding the sample directly to inhibitor?”

Inhibitors themselves tend to be denaturing and cause cell lysis, creating the same situation as freezing and generally targeting the same (more fragile) microbes.



Metagenomics Workshop

Topic

ON MICROBIAL SEQUENCE DATABASES

A Very Brief History of Microbiology

AND HOW MICROBIOLOGY WAS FUNDED THROUGHOUT IT

Because labs run on ideas and money.

- van Leeuwenhoek invents better microscope optics, discovers “animalcules” in water drops and everywhere else he looks
 - Some interest, but mostly a curiosity
- Pasteur discovers that microbes carry out commercially important fermentation.
 - Also that they carry out commercially important food spoilage
 - And (along with Koch) that microbes and not bad air (miasmas), evil spirits, or vengeful deities cause illness
- Biochemists and molecular biologists in the later half of the 20th Century want to understand the mechanics of life
 - E. coli provide an excellent model, basic mechanics of life are characterized in E. coli strains
 - E. coli becomes the workhorse of biology

A Very Brief History of Microbiology

AND HOW MICROBIOLOGY WAS FUNDED THROUGHOUT IT

Because labs run on ideas and money.

- van Leeuwenhoek invents better microscope optics, discovers “animalcules” in water drops and everywhere else he looks
 - Some interest, but mostly a curiosity
- Microbial ecology is born, interest is limited

A Very Brief History of Microbiology

AND HOW MICROBIOLOGY WAS FUNDED THROUGHOUT IT

Because labs run on ideas and money.

- Pasteur discovers that microbes carry out commercially important fermentation.
 - Also that they carry out commercially important food spoilage
 - And (along with Koch) that microbes and not bad air (miasmas), evil spirits, or vengeful deities cause illness
- Microbiology suddenly becomes of general interest due to health implications
- Microbiology suddenly becomes of great financial interest due to market for fermented products (wine, beer, etc.) or prevention of spoilage

A Very Brief History of Microbiology

AND HOW MICROBIOLOGY WAS FUNDED THROUGHOUT IT

Because labs run on ideas and money.

- Biochemists and molecular biologists in the later half of the 20th Century want to understand the mechanics of life
 - E. coli provide an excellent model, basic mechanics of life are characterized in E. coli strains
 - E. coli becomes the workhorse of biology
- Lots of E. coli strains already thoroughly characterized and in the lab, waiting to be sequenced
- Understanding the mechanics of life is a key to medical innovations that can make a lot of money

Consequences of This History

WHY OUR DATABASES ARE GOOD FOR SOME AND TERRIBLE FOR OTHERS

You get what you (and others) pay for

- If *E. coli* are found, most databases will have several strains to compare to
- If you find a pathogen (or something close to one), it will be well-represented in the database, with at least one, and probably a few strains to compare to.
- If you are sequencing microbes that are on or in humans or other well-studied animals, your microbes will probably be represented in the database
- If you are studying food, or find microbes similar to those that can carry out commercially important food spoilage or fermentation, they will probably be well-represented.
- If you are studying environmental microbiology, especially of more remote or obscure environments, the databases may confound you more than help.

Databases Are Imperfect

RESEARCHERS ARE ENCOURAGED TO SUBMIT NEW SEQUENCES

The are less-strongly encouraged to have rigorous quality control

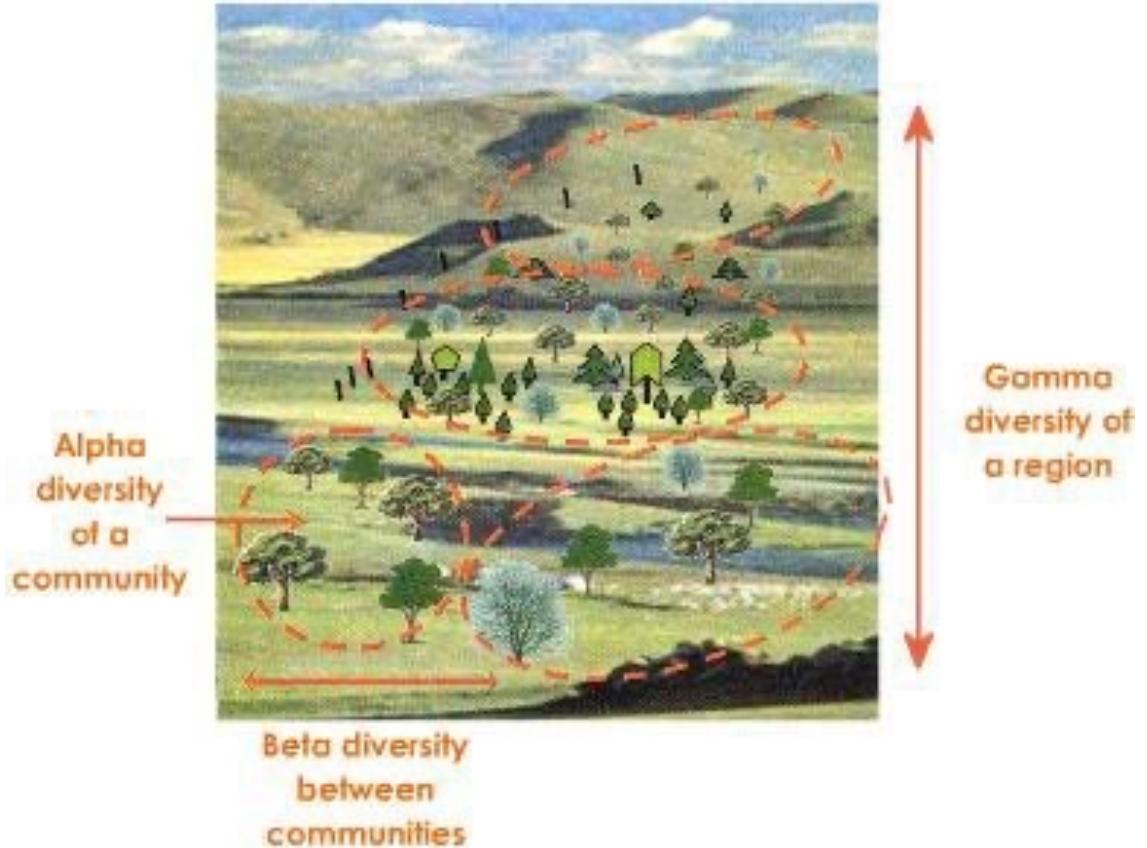
- Sequences can be found in databases that are clearly chimera in origin
 - Someone did poor chimera removal
- Sequences can be found listed under one species that clearly cluster much more closely with a different family, sometimes even a different phylum
 - Someone made an error when cataloging or recording the name or assigning taxonomy
- Sequences can be found that are identical between poorly-related species
 - Same as above
- Database updates often contain significant amounts of both material removal and material addition.
- **DON'T CONTRIBUTE TO GARBAGE IN THE DATABASES!**

Topic

DIVERSITY METRICS

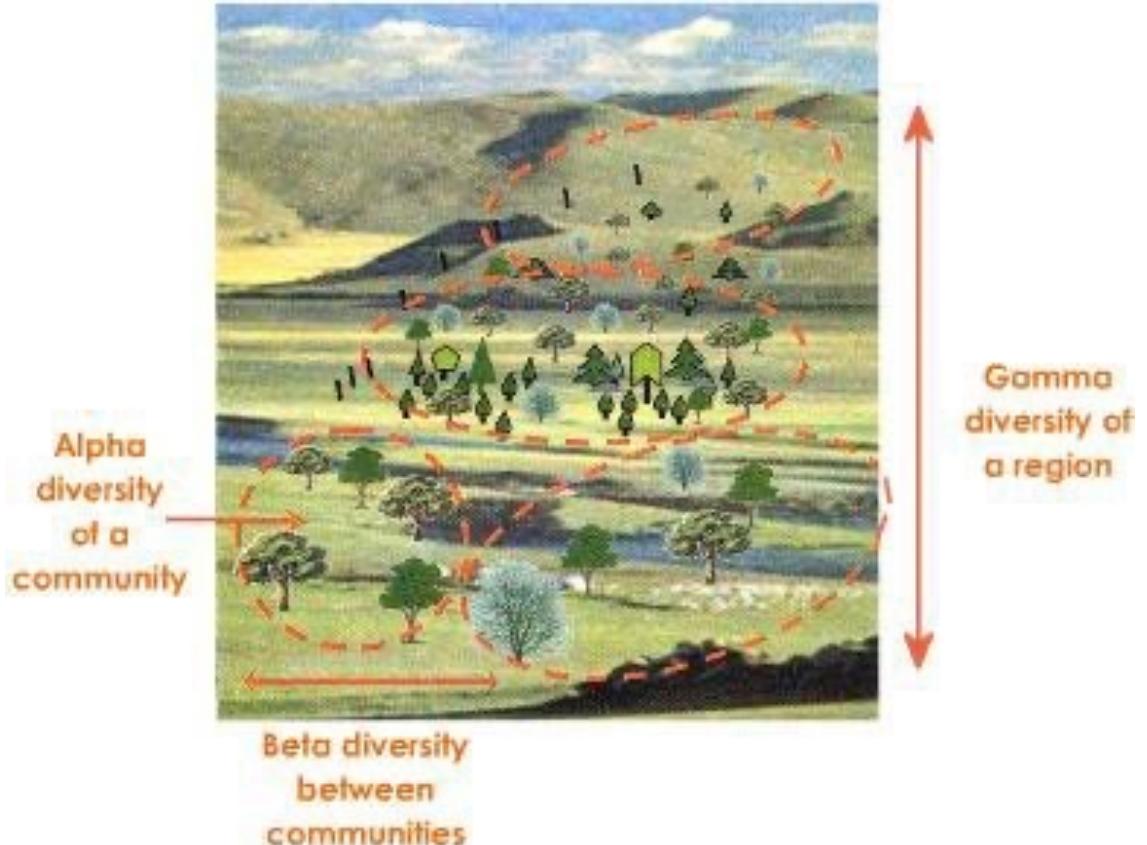
Diversity Types

- Alpha Diversity
 - How many different species are in my sample and how evenly are they distributed?
 - A monoculture has no alpha diversity. A sample with multiple species where only one is very predominant will also be very low.
 - Intrinsic property of the sample itself, no “better” or “worse” result, just more or less accurate.
 - High or low is not inherently better or worse, all depends on context.



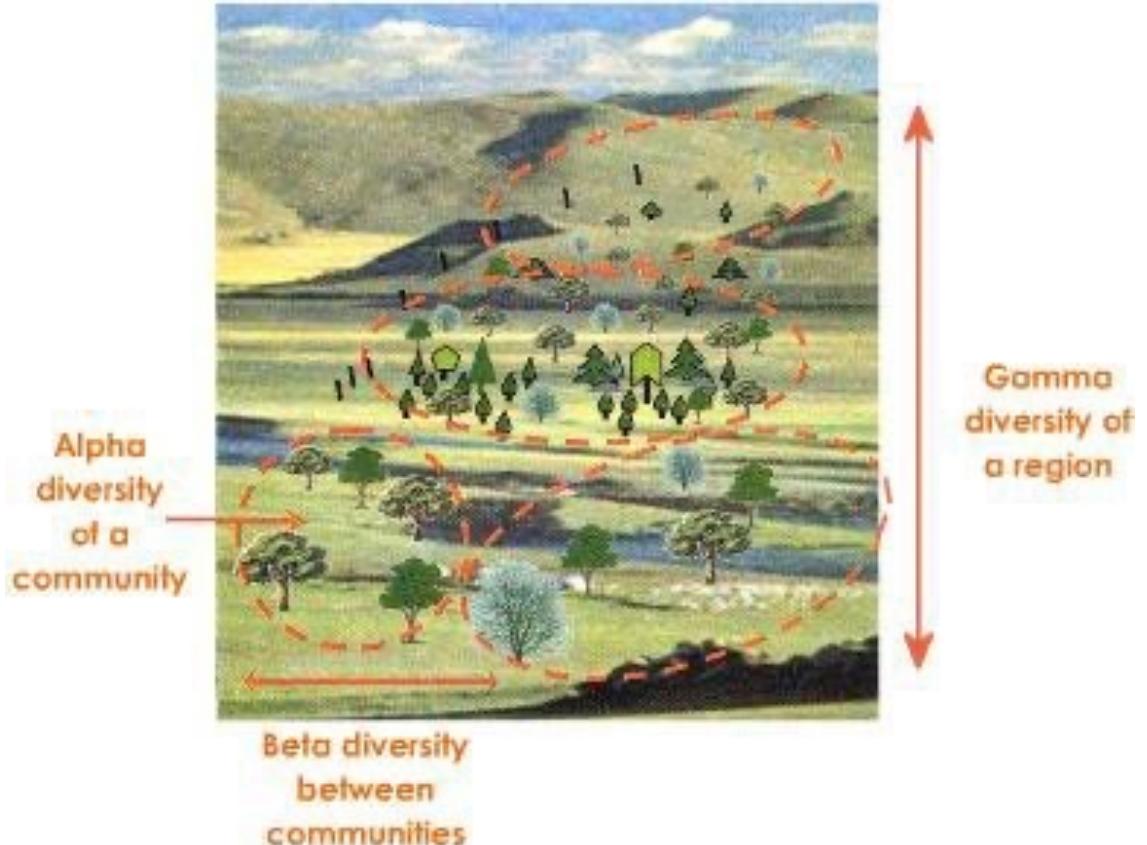
Diversity Types

- Beta Diversity
 - How similar/different are two samples?
 - The gut and oral microbiomes tend to have high alpha diversity, but may have low beta diversity between between similar individuals.
 - Beta diversity between an individual with IBD and a healthy gut will be extremely high.



Diversity Types

- Gamma Diversity
 - How diverse is this multi-site ecosystem?
 - Rarely used in metagenomics



Thank You!
