

# Microbial Co-occurrence Relationships in the Human Microbiome

Karoline Faust, J. Fah Sathirapongsasuti, Jeroen Raes, Curtis  
Huttenhower

2012



Detecting  
Taxa

Determining  
Frequencies

Inferring  
Interactions

# Co-occurrence



# Co-exclusion



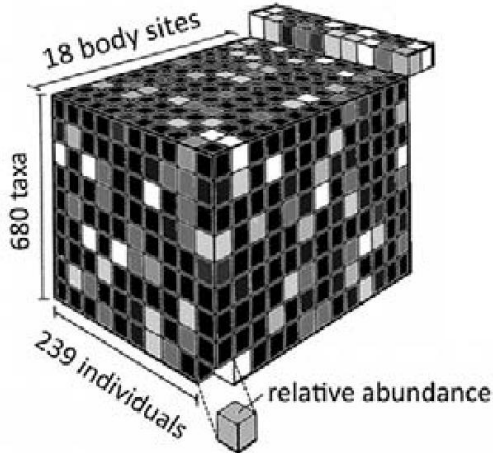
?

# Outline

- **The challenge-** detecting microbial co-occurrence & co-exclusion
- **The method-** Hybrid model
- Results
- **Validation method - “ReBoot”**– Dealing with compositional data
- Discussion & Thoughts

# Method

# The Model

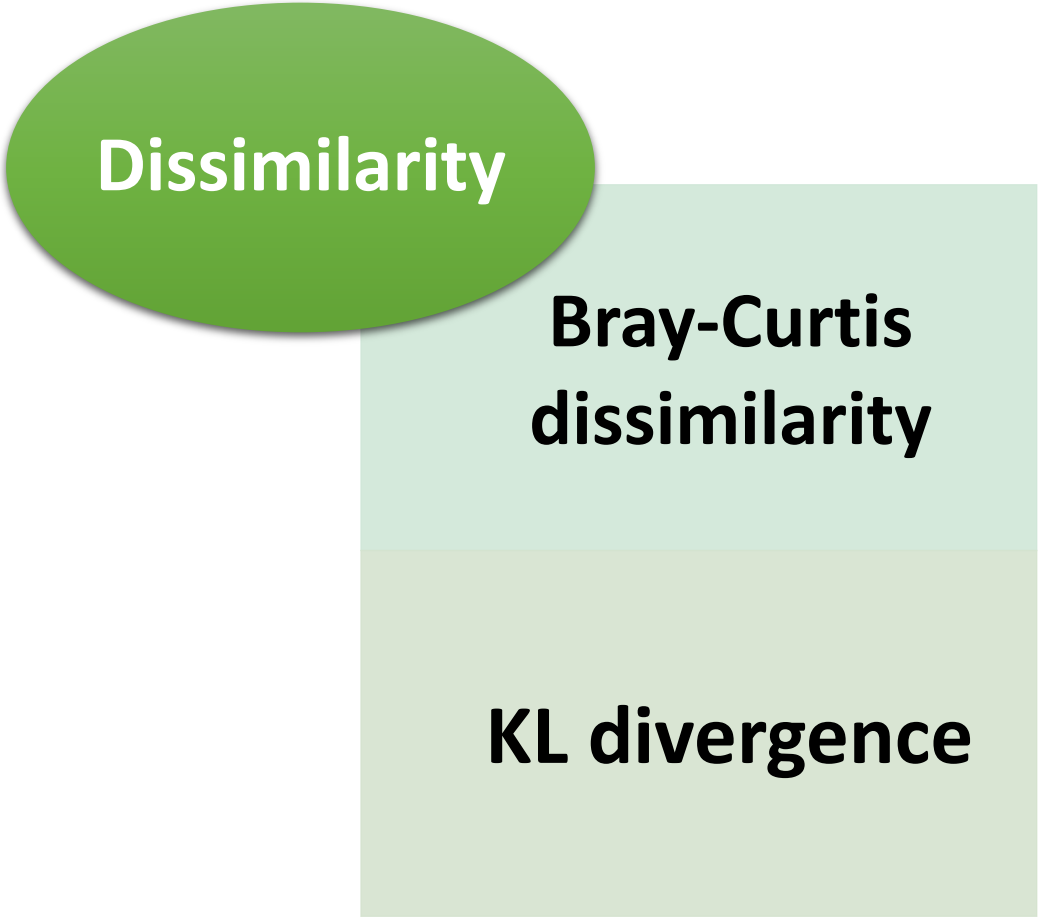
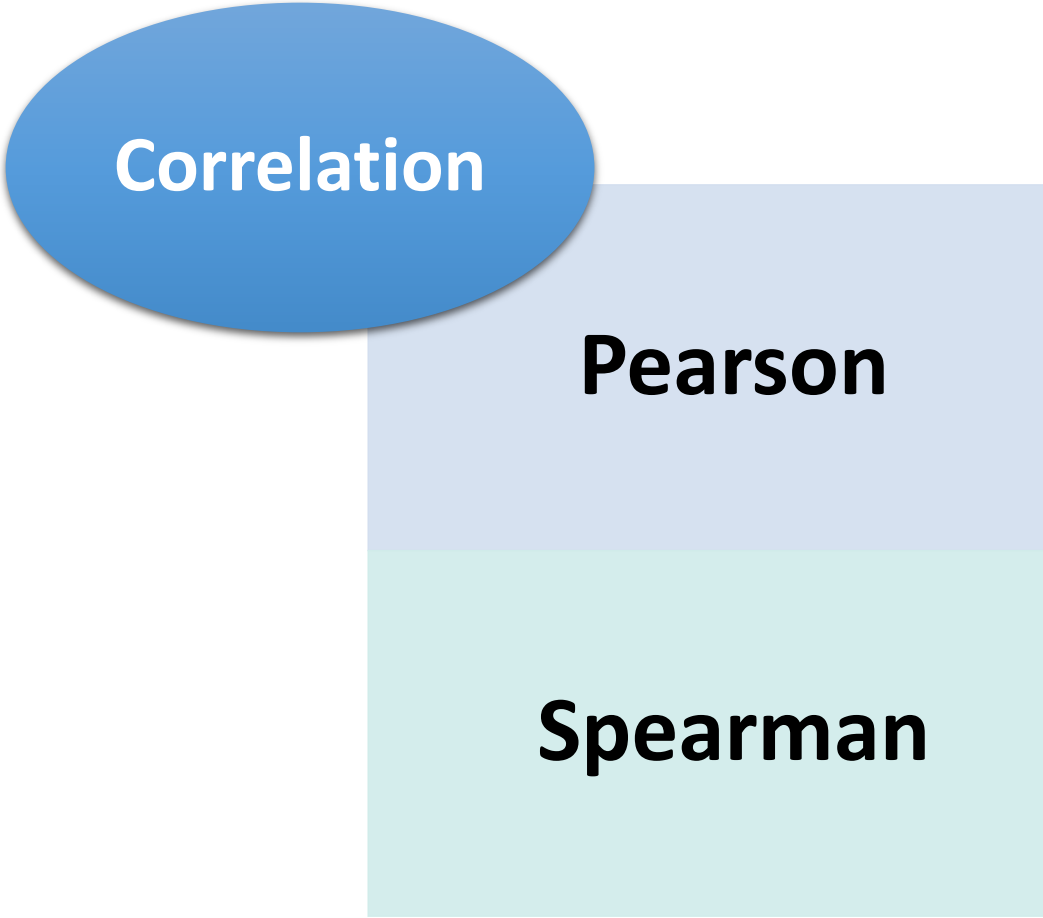


**Generalized Boosted  
Linear Model  
(GBLM)**

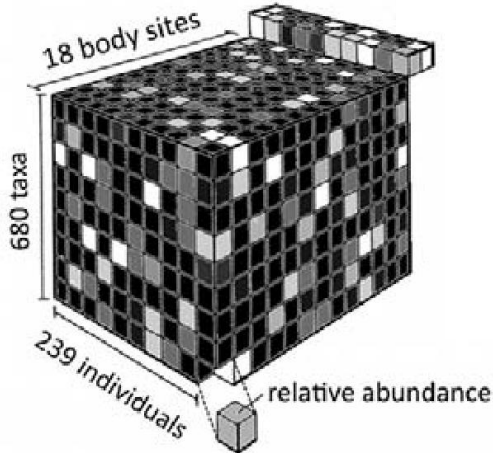
**Ensemble of  
Statistical Correlation  
measures**



# The Ensemble

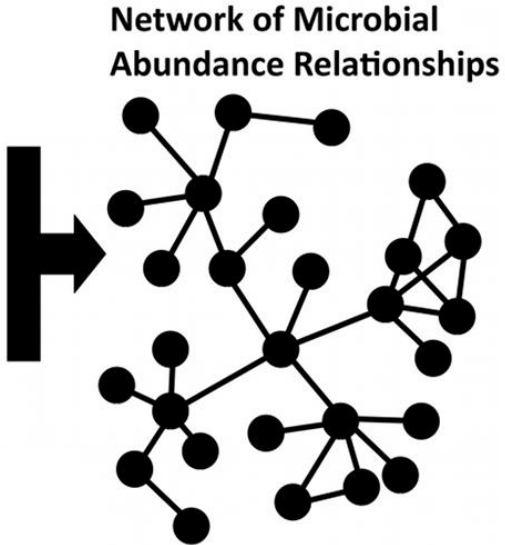
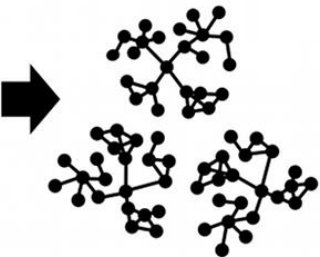
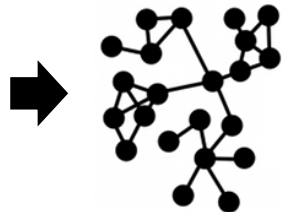


# The Model



Generalized Boosted  
Linear Model  
(GBLM)

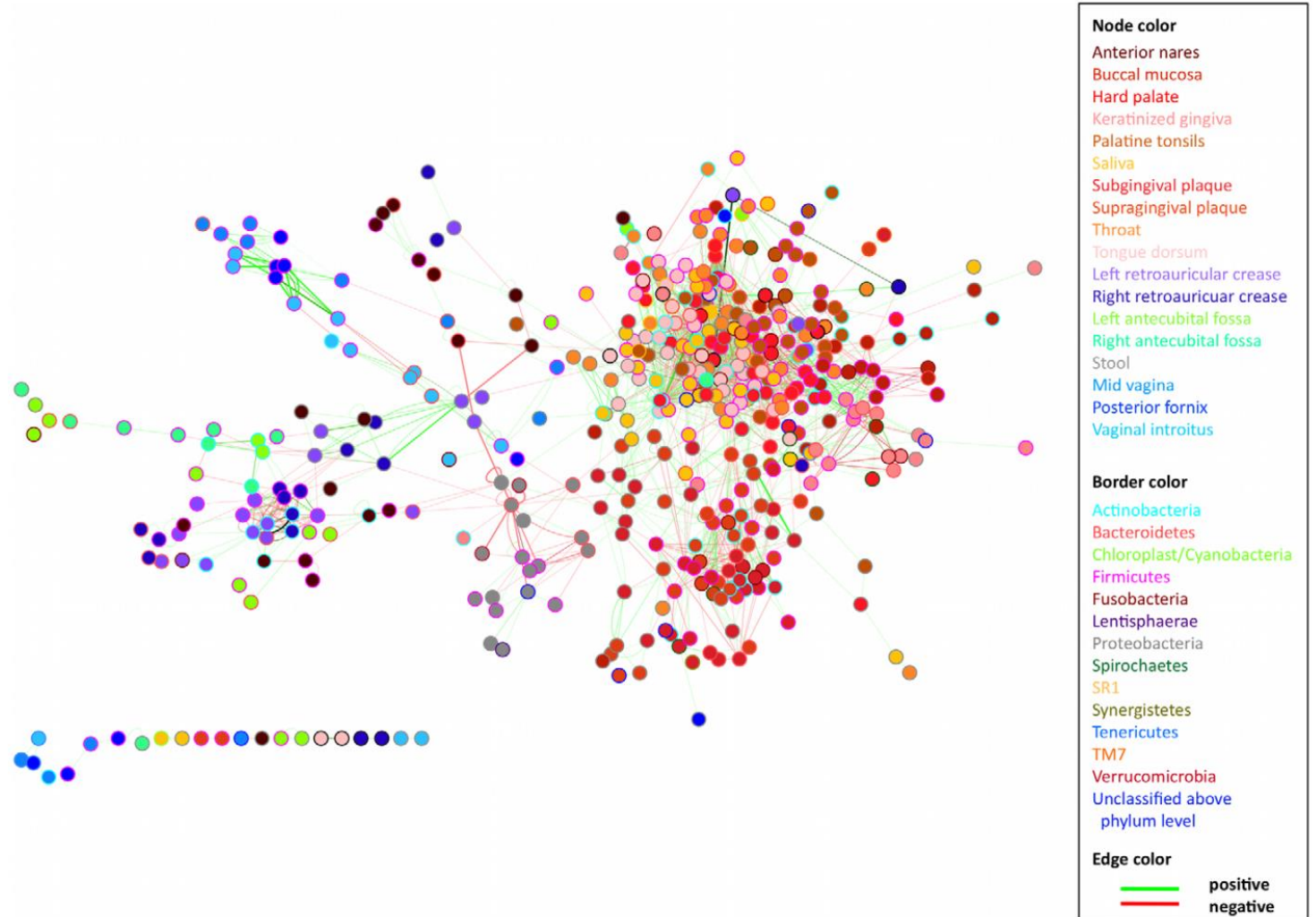
Ensemble of  
Statistical Correlation  
measures





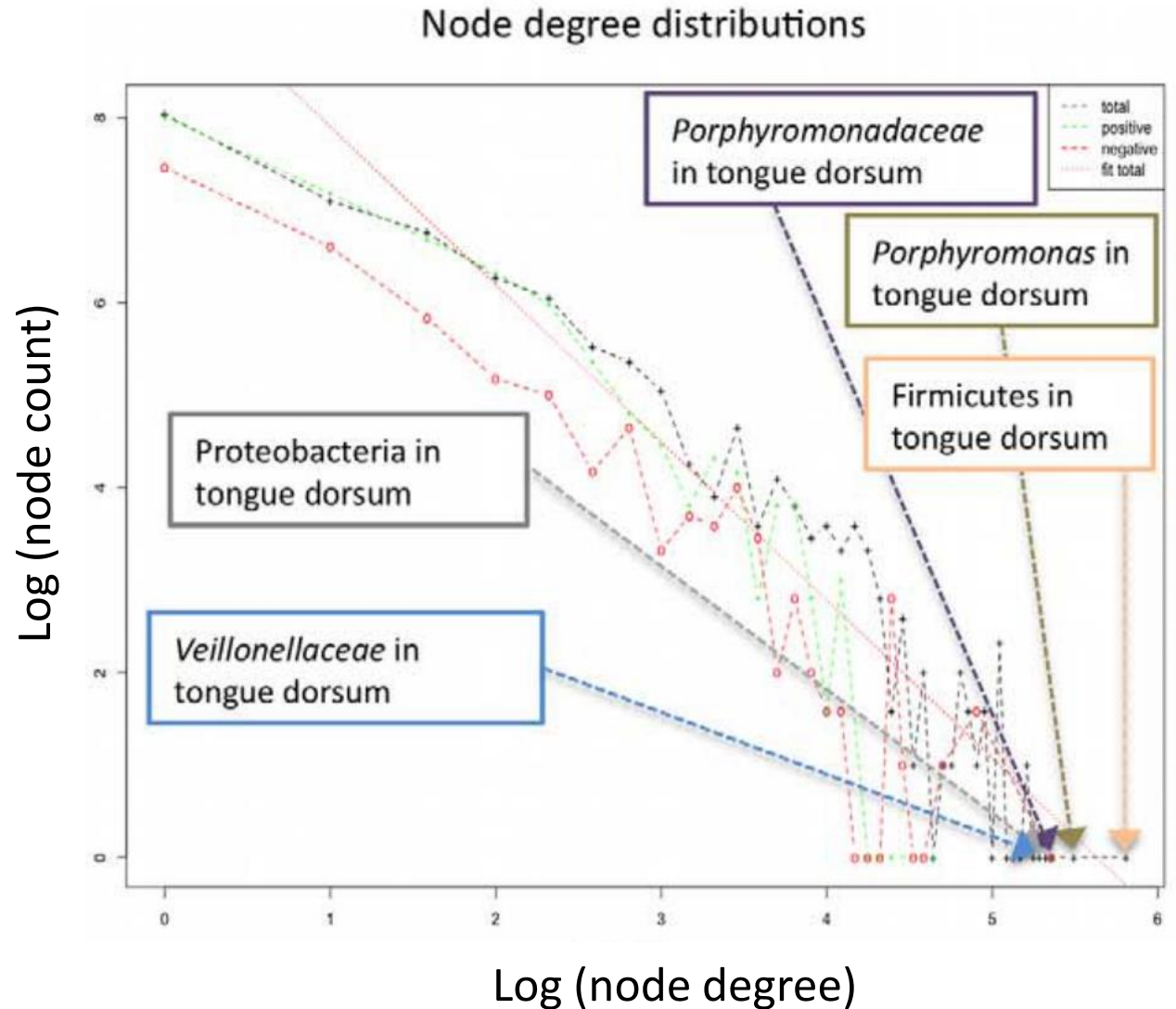
# Results

- High Modularity
- Variance in interaction patterns



# “Hubs”

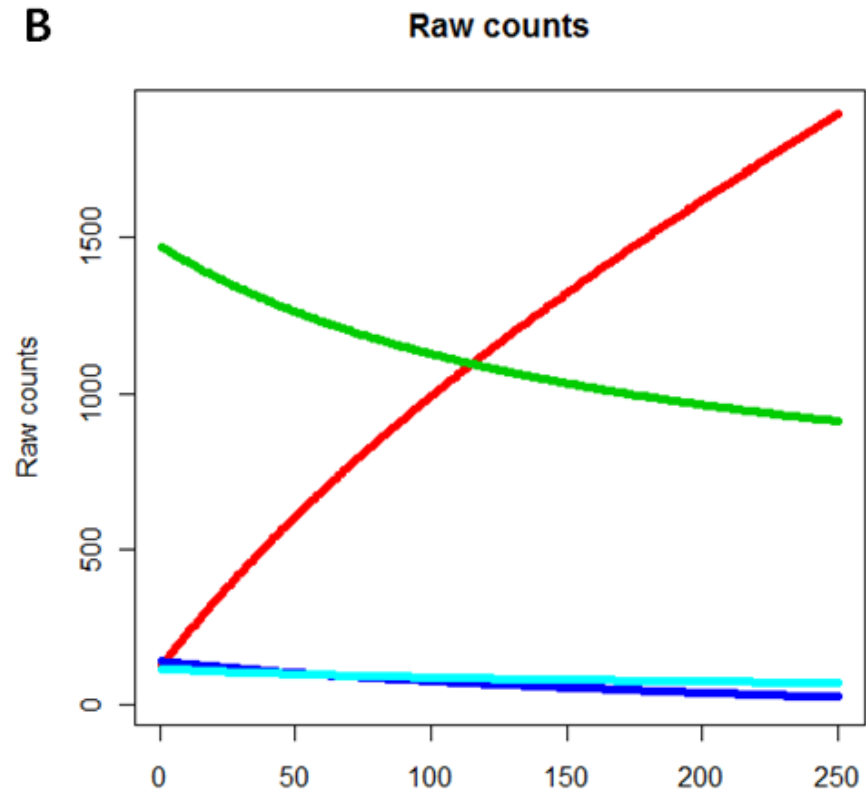
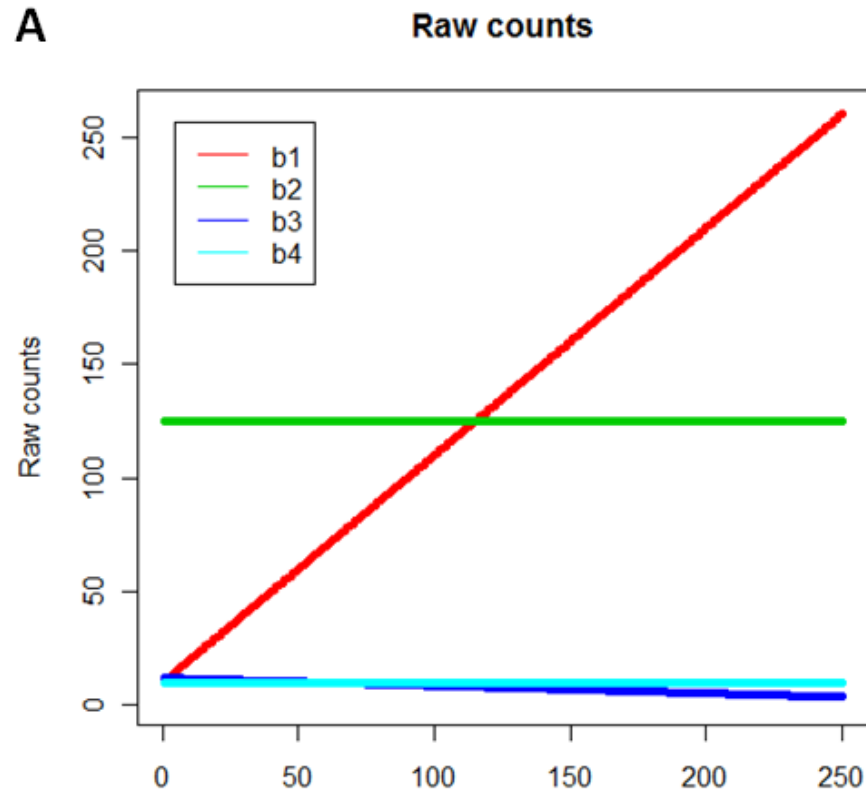
- Key OTUs acting as “coordinators” within and between body sites



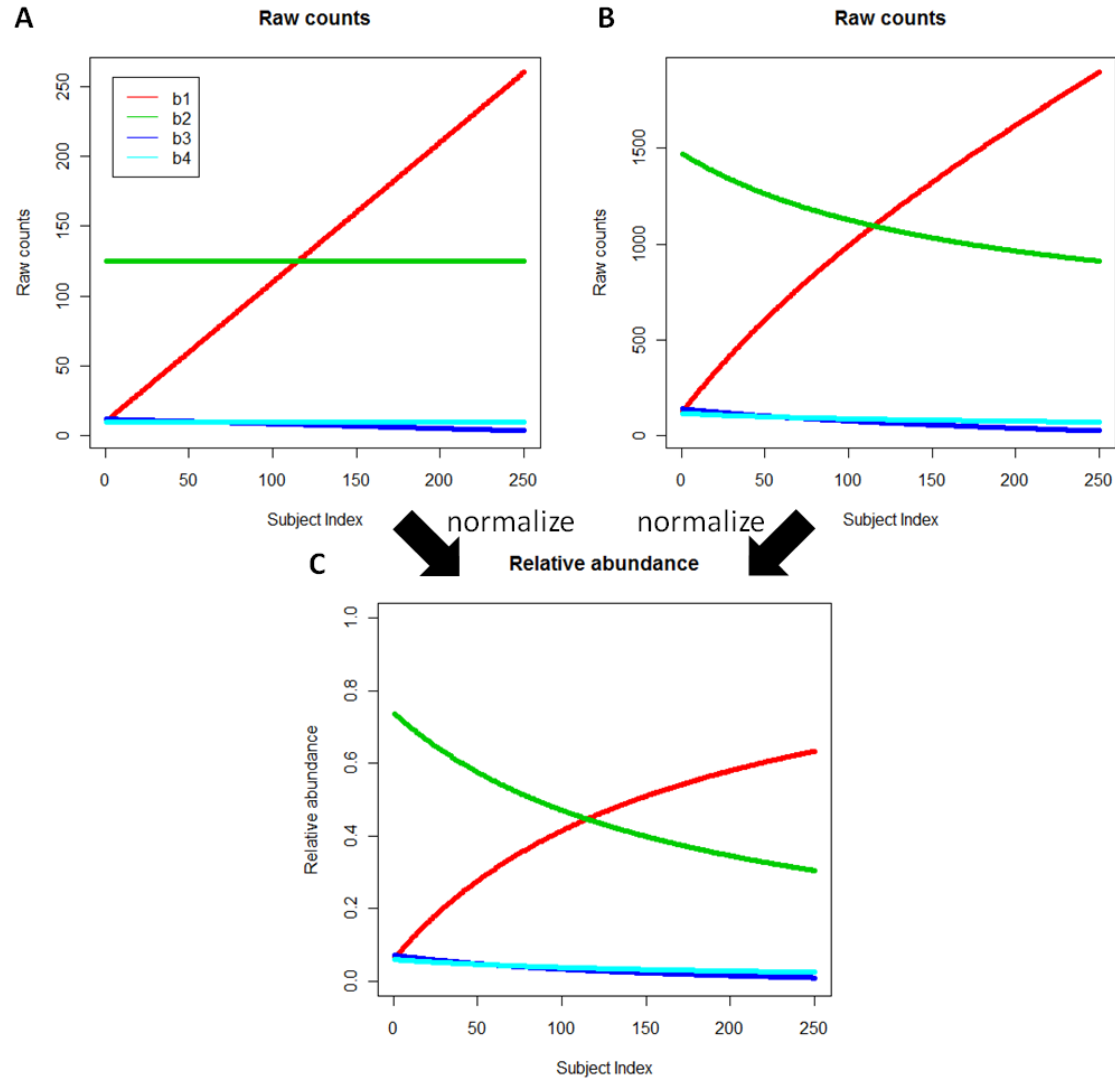
# “ReBoot”

Validating Compositional Data

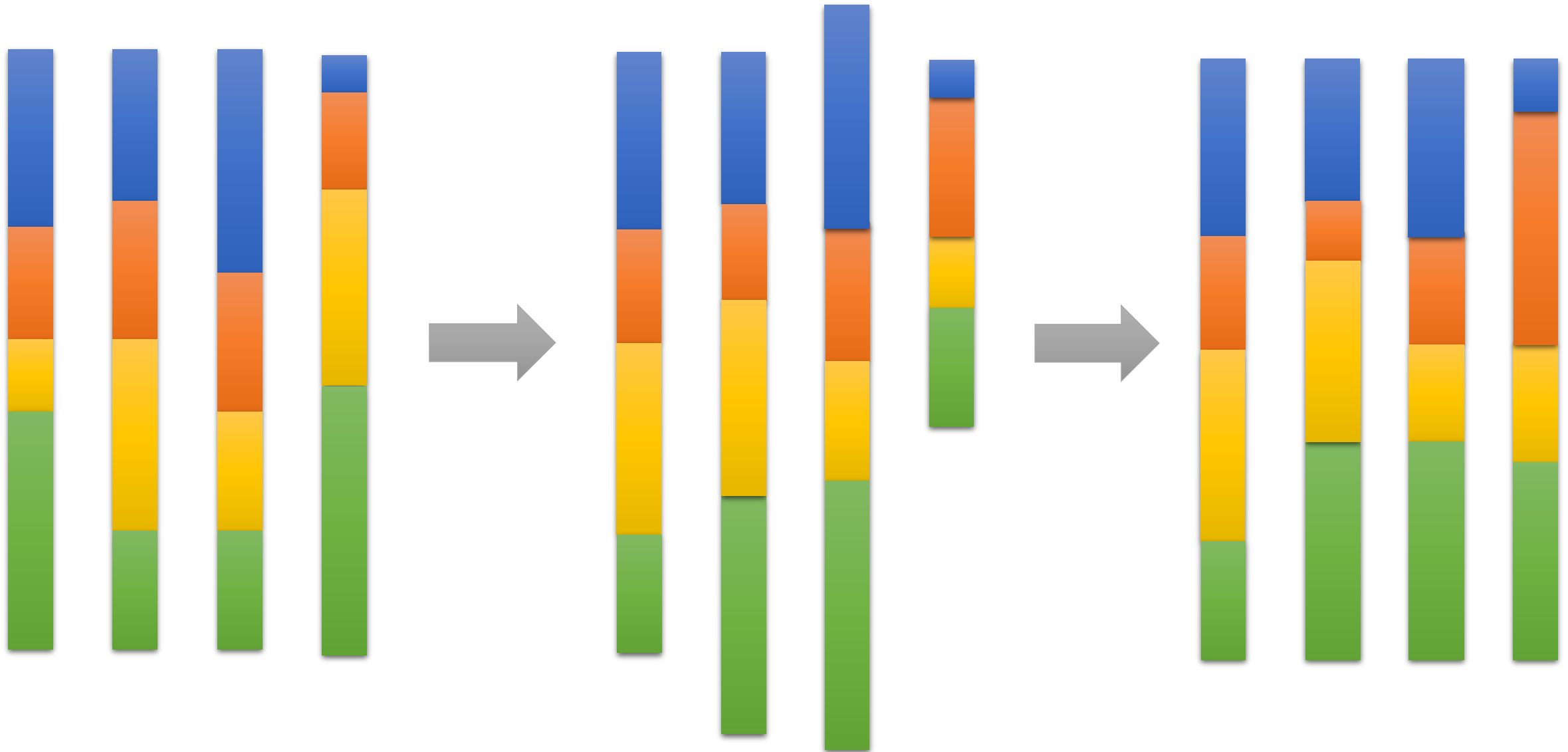
# Artifacts in Compositional data



# Artifacts in Compositional data

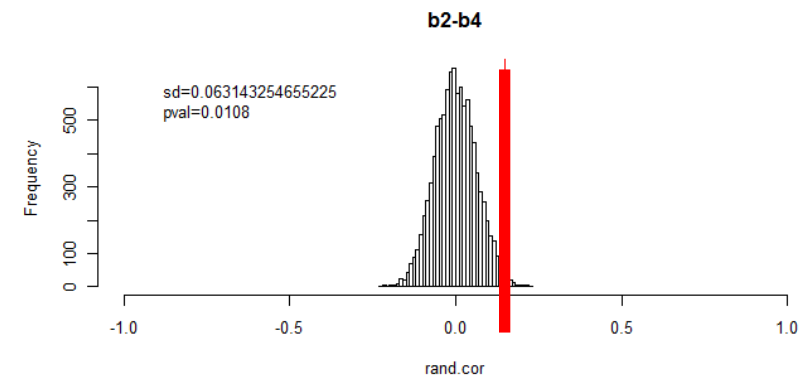
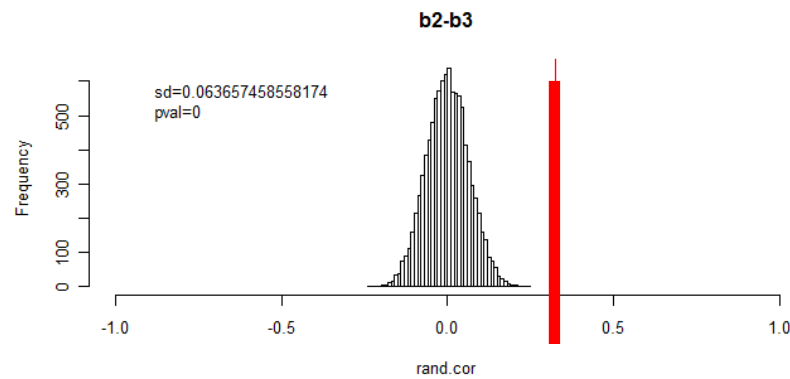
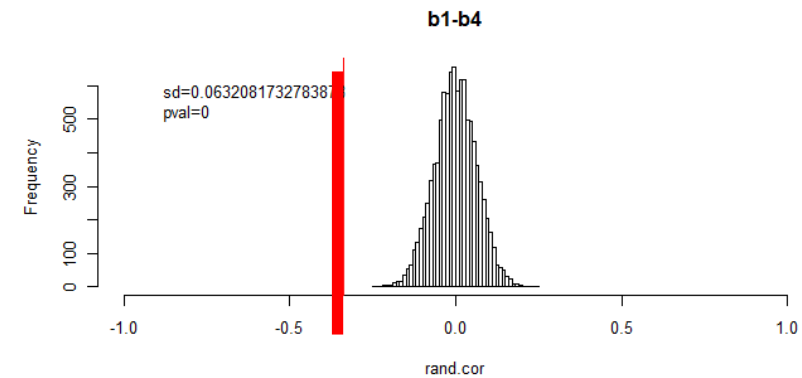
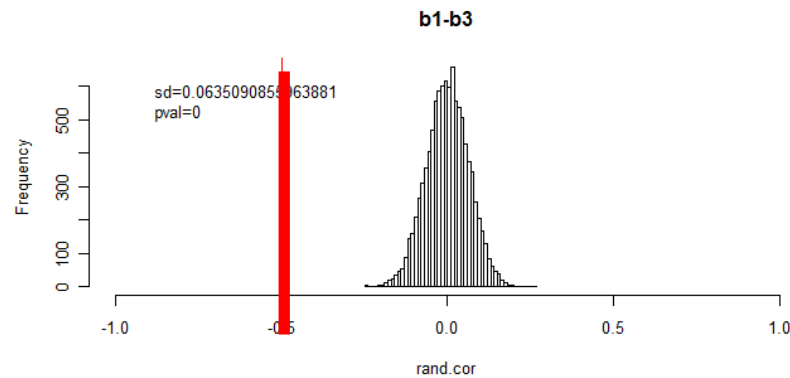


# Null Hypothesis using Standard Permutation test



# Null Hypothesis using Standard Permutation test

- Assessing null hypothesis using simulated data-



# ReBoot

Permutation

Renormalization

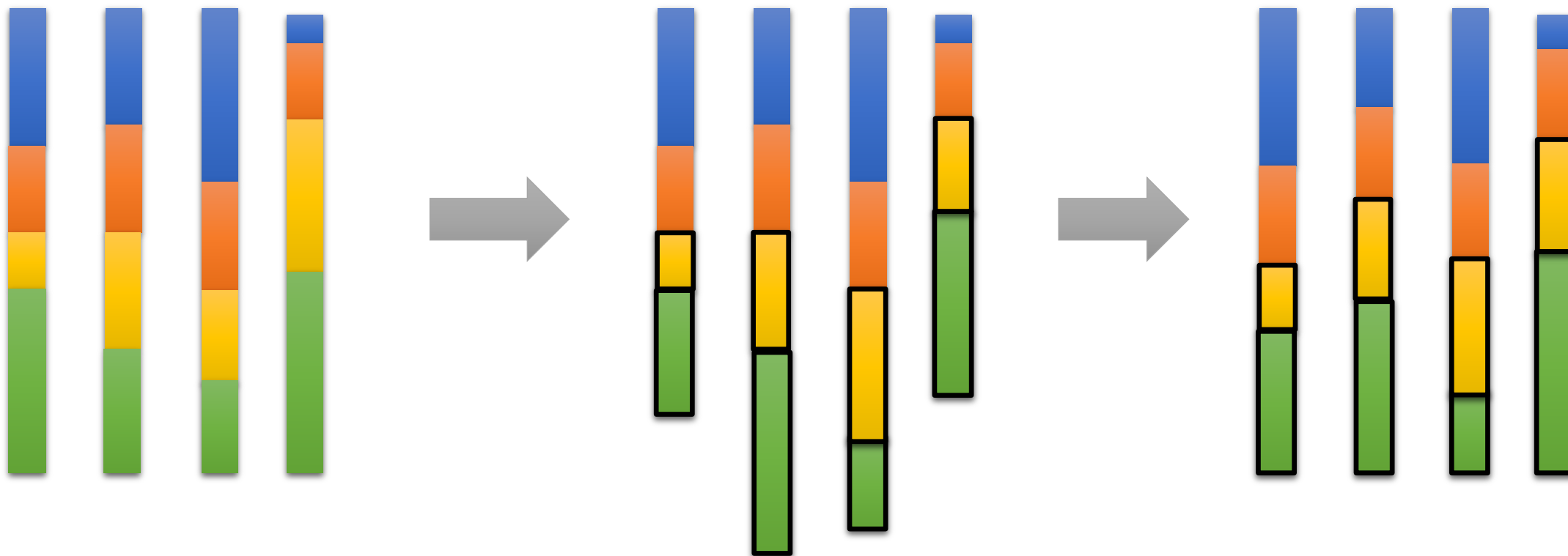
Bootstrapping



# ReBoot

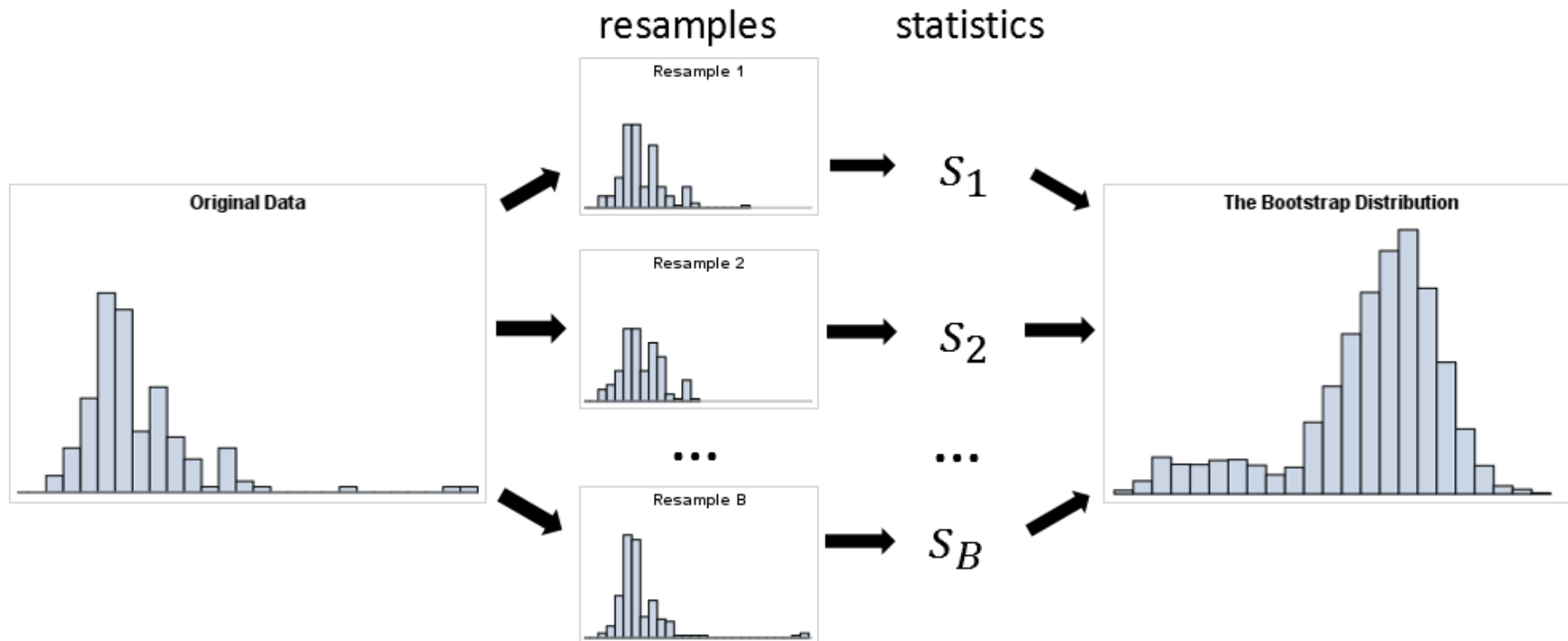
Permutation

Renormalization



# ReBoot

## Bootstrapping



# ReBoot

## Bootstrapping

- **Hints on low abundance data**
- **Bootstrap variance -> signal-to-noise ratio**

# ReBoot- results

	Pearson correlation		P-value			
	Absolute abundance (true corr.)	Relative abundance (composition)	Permutation test on the absolute abundance	Permutation test on the relative abundance	Permutation-Renormalization	Permutation-Renormalization and Bootstrap (ReBoot)
<b>b1-b2</b>	-0.00043	-0.11471	0.490	0.039	0.067	0.114
<b>b1-b3</b>	-0.37481	-0.45163	0.000	0.000	6.69E-14	1.34E-13
<b>b1-b4</b>	0.016205	-0.04366	0.403	0.247	0.203	0.167
<b>b2-b3</b>	0.022527	-0.03726	0.359	0.292	0.264	0.238
<b>b2-b4</b>	-0.0603	-0.13541	0.170	0.013	0.031	0.082
<b>b3-b4</b>	0.002101	-0.05086	0.518	0.225	0.196	0.170

# Summary

- **The challenge-** detecting microbial co-occurrence & co-exclusion
- **The method-**
  - Predictive model- GBLM
  - Ensemble of statistical correlation methods
- **“ReBoot”**– Improved Null hypothesis for compositional data

# Discussion

- What are the compositionality aspects not handled by “ReBoot”?
- How can the weights of the different measures determined?

**Thanks!**