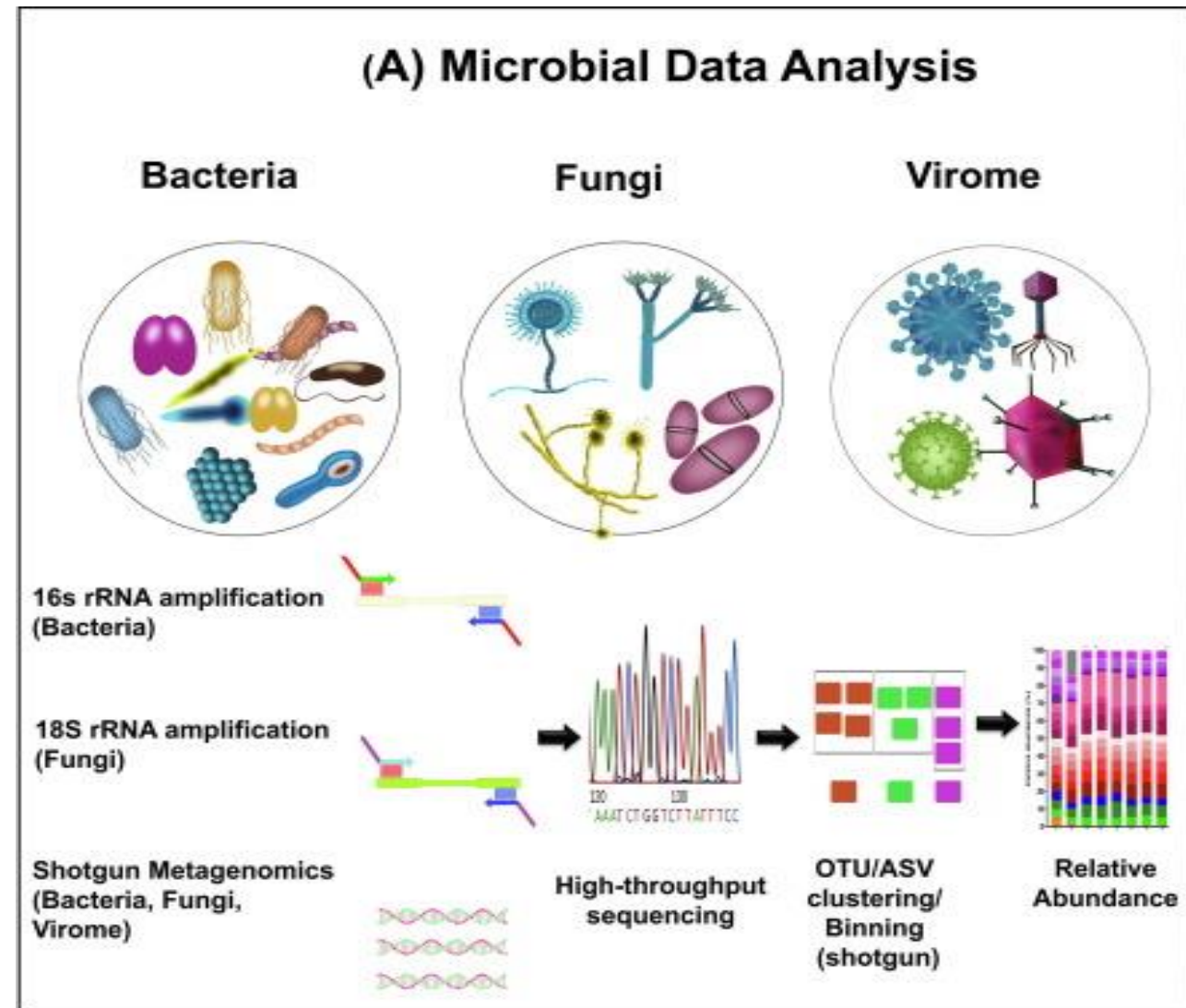


# Application of network analysis tools in understanding microbial community dynamics

Lidia Stadnik

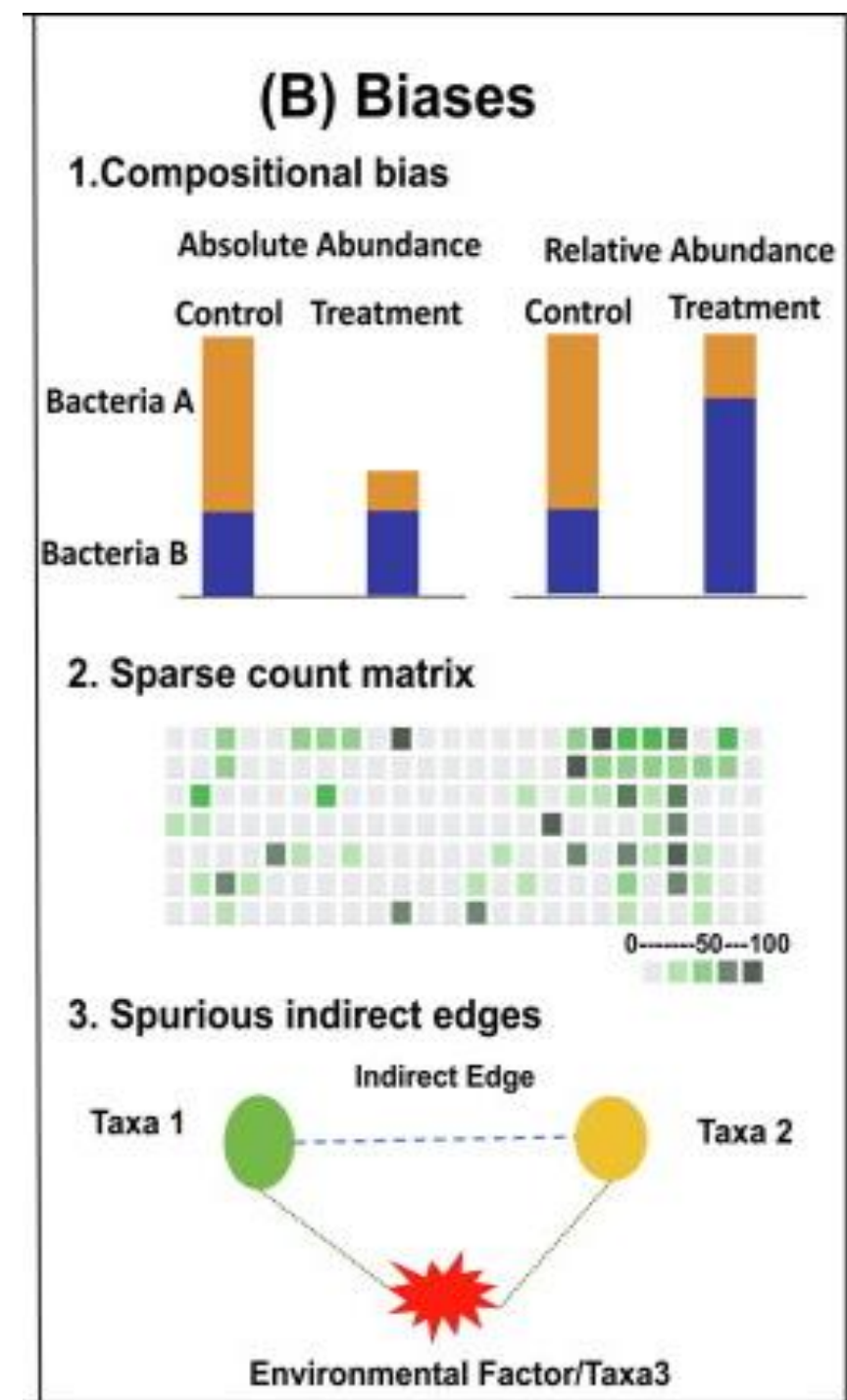
# Characterizing microbial communities

- Currently, the majority of network-based tools and models are used to study intra-kingdom interactions, mostly between bacteria.



# Network analysis

- Network-based analytical approaches are used to decipher microbial co-occurrence patterns.
- Decoding complex microbial co-occurrence relationships is associated with three main challenges.
- Most of these methods employ linear models based on correlation, regression, and probabilistic graphical models.



# Correlation Analysis

- Pearson or Spearman correlation methods are commonly used to estimate microbial interactions between pairs of taxa.
- However, this analysis do not account for compositionality, leading to false associations between low-abundant microbial members in a community.
- To address this, compositional data analysis techniques are employed.

# Overcoming Sparsity in Microbial Interaction Analysis

- Centered log ratio transformation (CLR) can be utilized as a pre-processing step (e.g., SparCC).
- Bootstrapping techniques (e.g., CoNet).
- Challenge: selecting a suitable correlation cut-off to control network sparsity.
- MENAP automates the selection of the optimal correlation threshold using a random matrix theory-based method.

# Regularized linear regression models

- It builds models where the abundance of each taxon is predicted using the abundance of all other taxa.
- The coefficient of each taxon serves as a linear measure for the interaction strength of two taxa.
- To prevent overfitting due to the large number of features, penalty terms like the  $\ell_1$ -penalty (lasso) are introduced.
- The  $\ell_1$ -penalty drives the coefficients of less influential taxa to zero, making the solution sparse (CCLasso, REBACCA).
- Compared to correlation methods there's improved detection of false relations.

# Meta-Network

- Association rule mining techniques can detect both direct and indirect relationships, including non-linear associations.
- It initially generates presence-absence indicator matrices for each sample.
- Co-occurrence frequencies of taxa pairs are then computed to derive a co-occurrence probability matrix.
- Using this matrix, a network is constructed with a predefined co-occurrence probability threshold (e.g., 80%).
- Functional Similarity Weight (FS-Weight) algorithm to identify indirect relationships and The PCA-PMI method to detect non-linear associations

# Conditional dependence and graphical methods

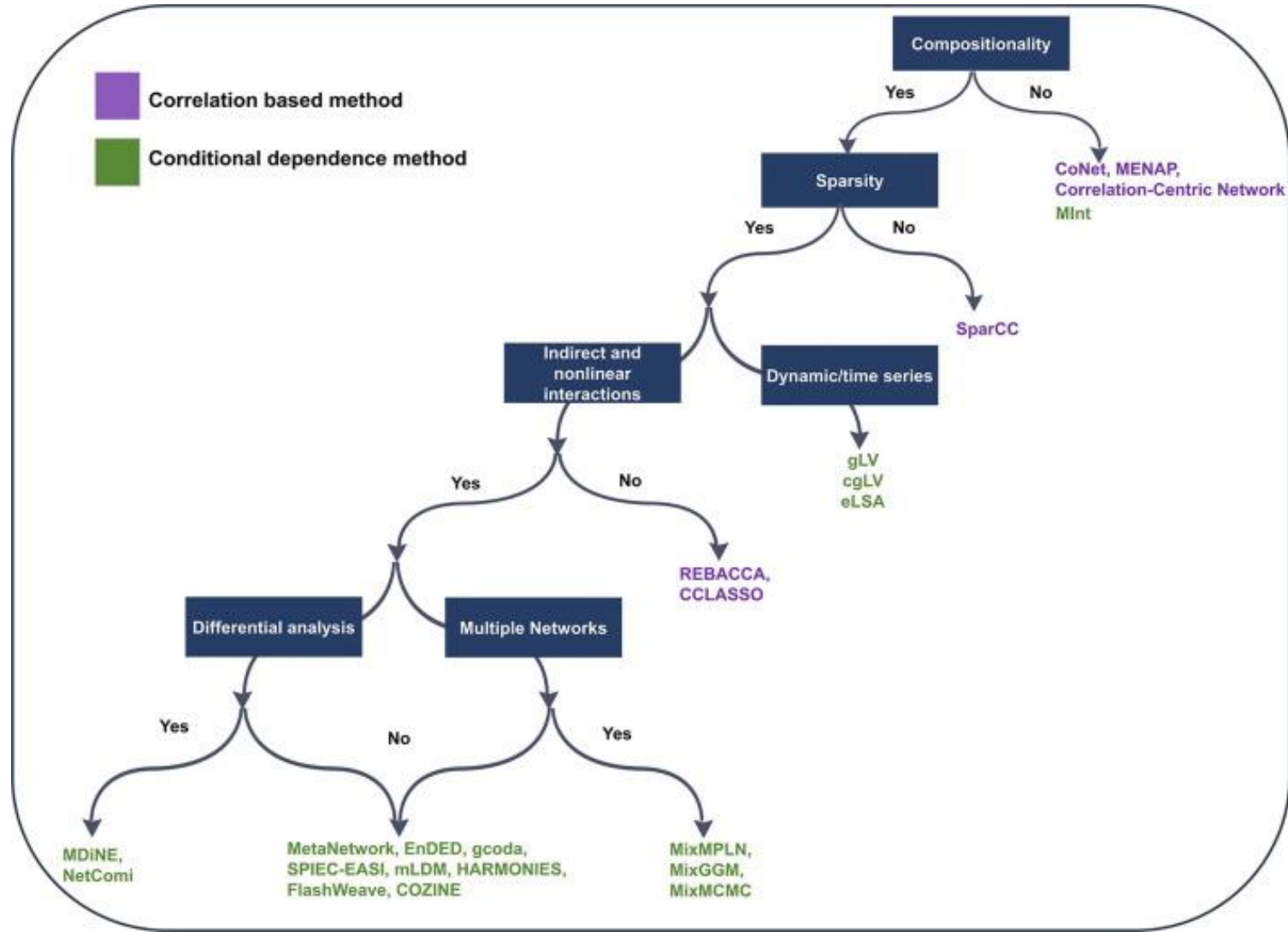
- Correlation based methods typically fail to differentiate between direct and indirect associations.
- These can be distinguished using Partial correlation, resulting in an undirected weighted graph where the edges imply the conditional dependency between two taxa.
- Most of these methods can also account for confounders such as biological covariates and technical biases.



# Conditional dependence and graphical methods

- Technique SPIEC-EASI applies a CLR transformation to the data, and then generates graphical network by estimating sparse inverse covariance matrices
- Other methods, like mLDM and HARMONIES, employ hierarchical Bayesian models and zero-inflated negative binomial distributions, respectively, to model microbial abundance and infer interactions.
- COZINE introduces a novel approach to accommodate compositionality and zero inflation, using a Multivariate Gaussian Hurdle model for interaction inference.

# Workflow indicating the suitable network approaches depending upon different challenges



# Conclusion

- Currently available methods are not able to overcome all of the challenges associated with microbiome data including compositionality bias, overdispersion, a poor sample to feature ratio and *trans*-kingdom interactions.

# Sources

- Matchado, Monica Steffi, et al. "Network analysis methods for studying microbial communities: A mini review." *Computational and structural biotechnology journal* 19 (2021): 2687-2698.
- Larsen, Peter, Yuki Hamada, and Jack Gilbert. "Modeling microbial communities: Current, developing, and future technologies for predicting microbial community interaction." *Journal of Biotechnology* 160.1-2 (2012): 17-24.