



**IPS 104, Statistics and Cosmetics**  
**Inference of microbial association networks with graphical models**

Andrei Furtuna\* <sup>1</sup> ([andrei.furtuna@rd.loreal.com](mailto:andrei.furtuna@rd.loreal.com)) - Paris Region, France  
Florent Guinot <sup>1</sup> ([florent.guinot@rd.loreal.com](mailto:florent.guinot@rd.loreal.com)) - Paris Region, France  
Alban Ott <sup>1</sup> ([alban.ott@rd.loreal.com](mailto:alban.ott@rd.loreal.com)) - Paris Region, France

<sup>1</sup> L'Oréal Research&Innovation

**Abstract:**

Microbial association networks are frequently used for understanding and comparing community dynamics from microbiome datasets between individuals with different condition (e.g. disease status, clinical sign, ...). We think network representation is probably the best method to shed light on microbial interactions in these communities. Inferring microbial correlations for such networks and obtaining meaningful biological insights, however, requires a choice of appropriate statistical methods followed by a relevant pipeline for visualizing, reporting and comparing the associations.

In this work, we assessed the performance of several types of probabilistic graphical models (PGM) to correctly infer microbial association networks. More precisely, we focused on the Graphical Lasso method to infer sparse networks, as one would expect microbial association networks to be. However, the theoretical assumptions of this method (normality of the data) imply that data should undergo appropriate transformations prior to the inference step. This led us to consider various data processing techniques and approximations that would help alleviate these assumptions.

One of the bottlenecks of microbiome statistical analysis is the lack of knowledge we can consider as certain to bench methods (sensitivity, specificity...). In order to address this limitation, we identified the most used and promising tools to simulate data from real ones.

Simulations allowed us to determine the most well suited methods for network inference in different configurations of parameters such as sample size, network topology or strength of correlations.

In addition to the simulation, we also inferred a microbial association network on a real dataset to compare the microbial dynamics between two conditions related to a cosmetic disorder.

**Keywords:** probabilistic graphical models; microbiome; data simulation