

NGS EXPO 2024 招待講演 4： データ利活用の促進

ベイズ統計モデルに基づく微生物叢データ解析手法の開発

早稲田大学 細田 至温

# 自己紹介

細田 至温（ほそだ しおん）

博士（工学）

博士論文：「統計モデリングによる腸内微生物の関係性の推定」

経歴

2017～2022 早稲田大学 先進理工学研究科 電気・情報生命専攻 浜田研究室

2022～現在 株式会社日立製作所 研究開発グループ 基礎研究センタ 研究員

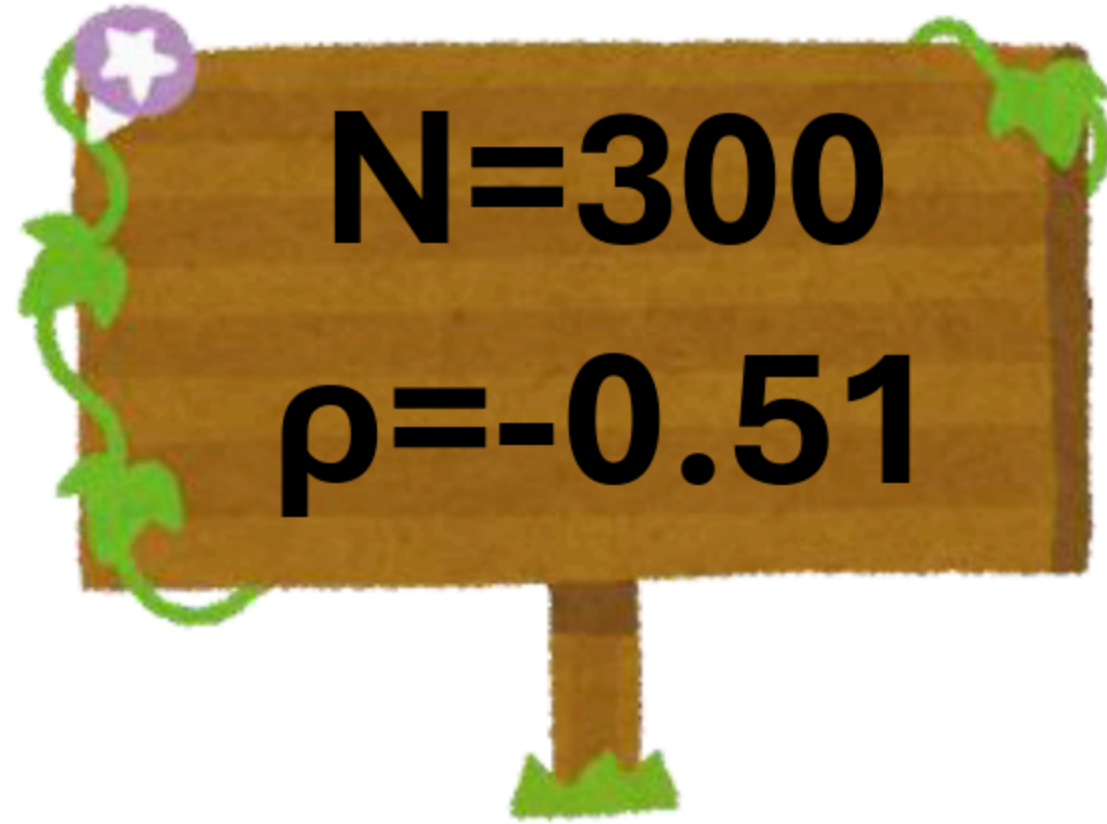
2024～現在 早稲田大学 理工学術員総合研究所 招聘研究員

興味

データに隠れた構造を使って新しい知見を抽出すること

はじめに

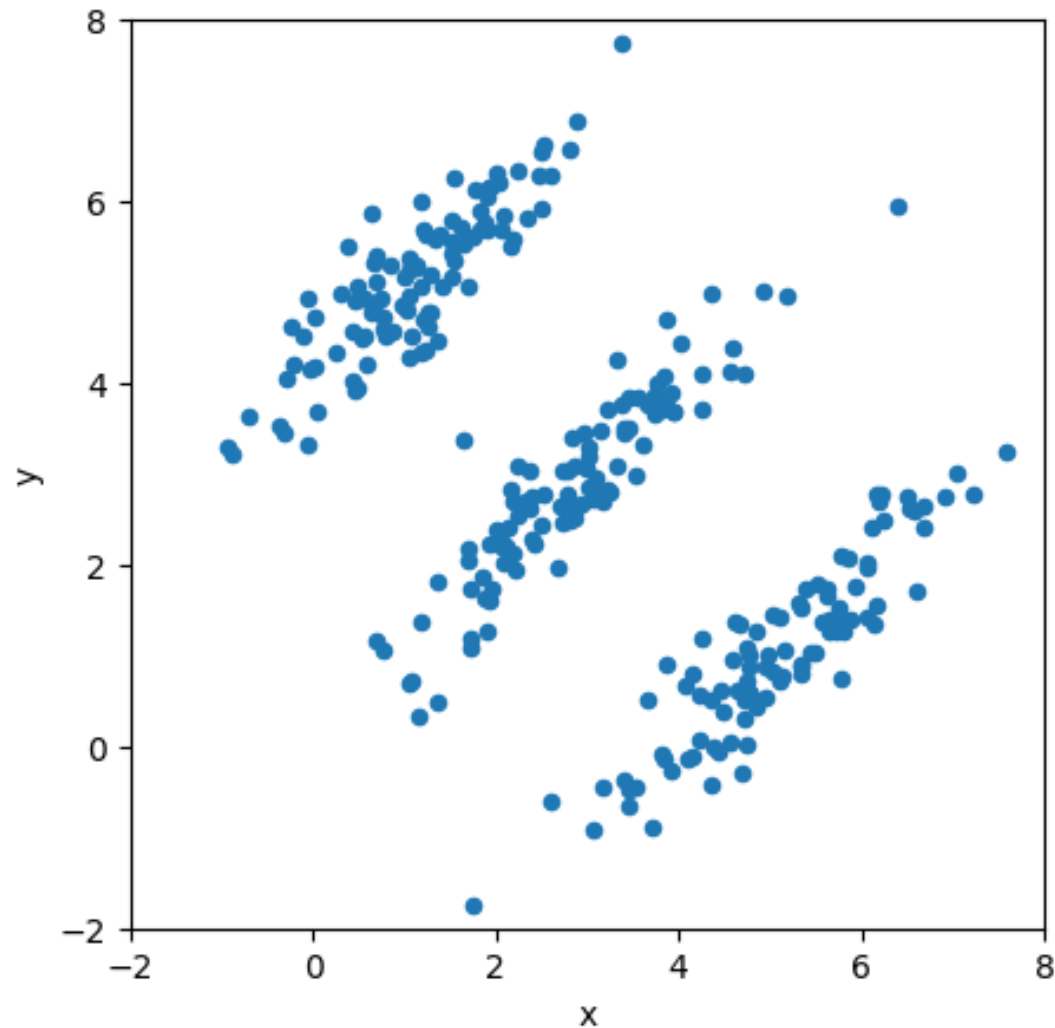
# 何を想像しますか？



はじめに

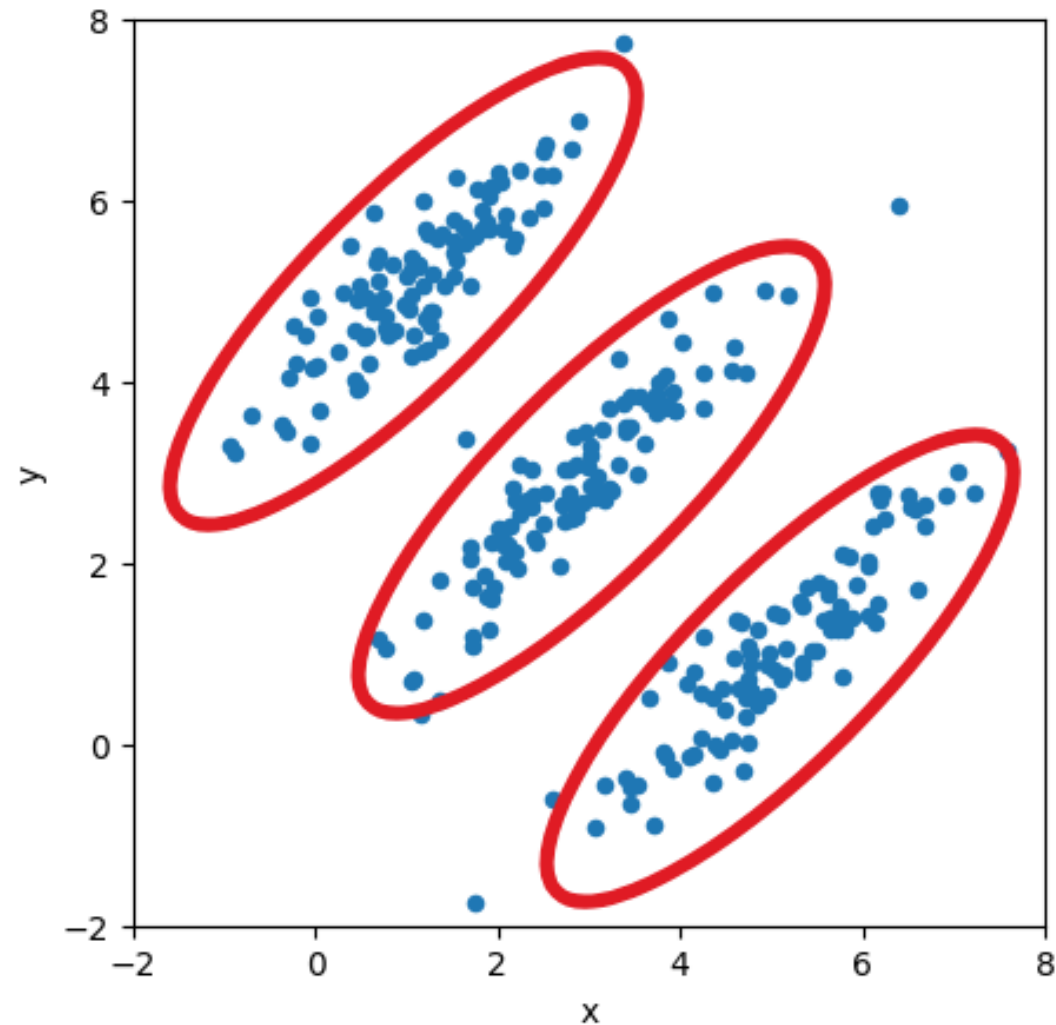
# イメージ通りでしたか？

- これはサンプルサイズ300、ピアソン相関-0.51の図です



# 変数の見落としに注意

- 見落としている変数があると間違った結論を出してしまう
- 理想的には？
  - 考えられる変数を列挙して、追加でデータを取得する



# 「隠れ変数」が今回の（隠れ）コンセプトです

- しかし、測定することが不可能な値も存在する
- 見えない変数=「隠れ変数」を導入して推定
  - e.g., 「グループ」を考慮して相関を計算
- 「どうやってできるか」ではなく「何ができるか」をお伝えできればと思います

# 目次

- はじめに
- 研究紹介
  - 「時間変化」を考慮して微生物相互作用を推定する
  - 「効果が出る期間のズレ」を考慮してプロバイオティクスの効果を推定する
  - 「細菌の所属するグループ」を考慮して細菌の共起関係を推定する
- 最後に



# 「時間変化」を考慮して微生物相互作用を推定する

*Bioinformatics*, 37, 2021, i16–i24  
doi: 10.1093/bioinformatics/btab287  
ISMB/ECCB 2021



---

## **Umibato: estimation of time-varying microbial interaction using continuous-time regression hidden Markov model**

**Shion Hosoda<sup>1,2,\*</sup>, Tsukasa Fukunaga<sup>1,3</sup> and Michiaki Hamada<sup>1,2,4,\*</sup>**

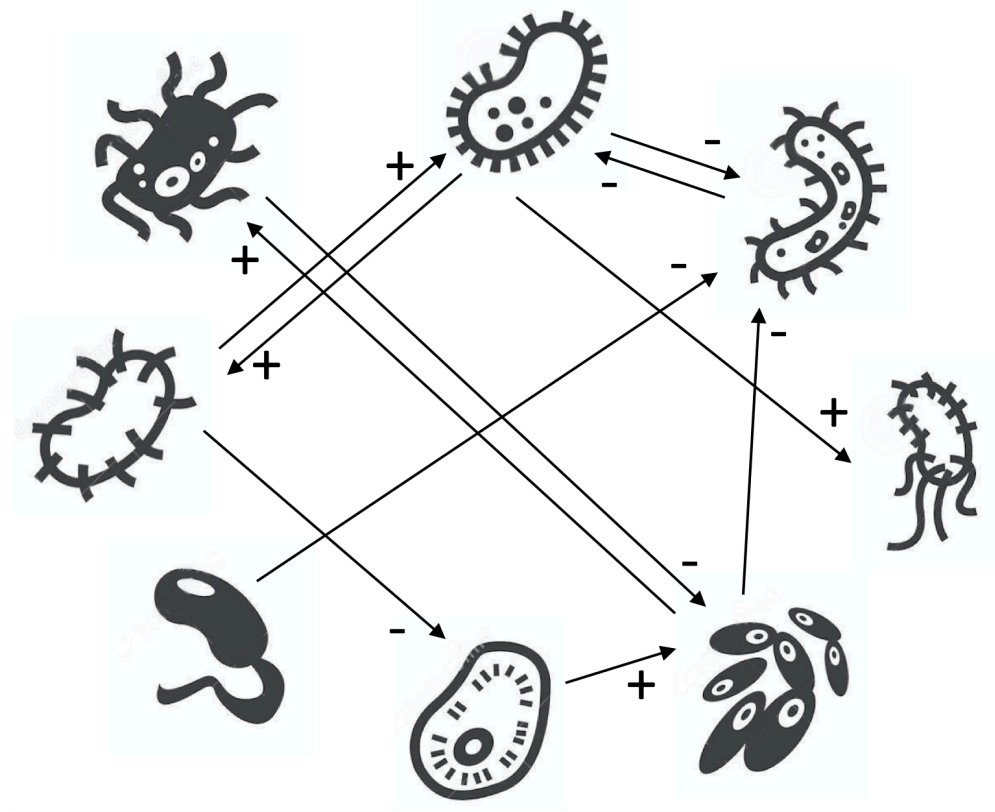
<sup>1</sup>Department of Electrical Engineering and Bioscience, Graduate School of Advanced Science and Engineering, Waseda University, Tokyo 169-8555, Japan, <sup>2</sup>Computational Bio Big-Data Open Innovation Laboratory, AIST-Waseda University, Tokyo 169-8555, Japan, <sup>3</sup>Department of Computer Science, Graduate School of Information Science and Technology, The University of Tokyo, Tokyo 113-8656, Japan and <sup>4</sup>Graduate School of Medicine, Nippon Medical School, Tokyo 113-8602, Japan

\*To whom correspondence should be addressed.

<https://doi.org/10.1093/bioinformatics/btab287>

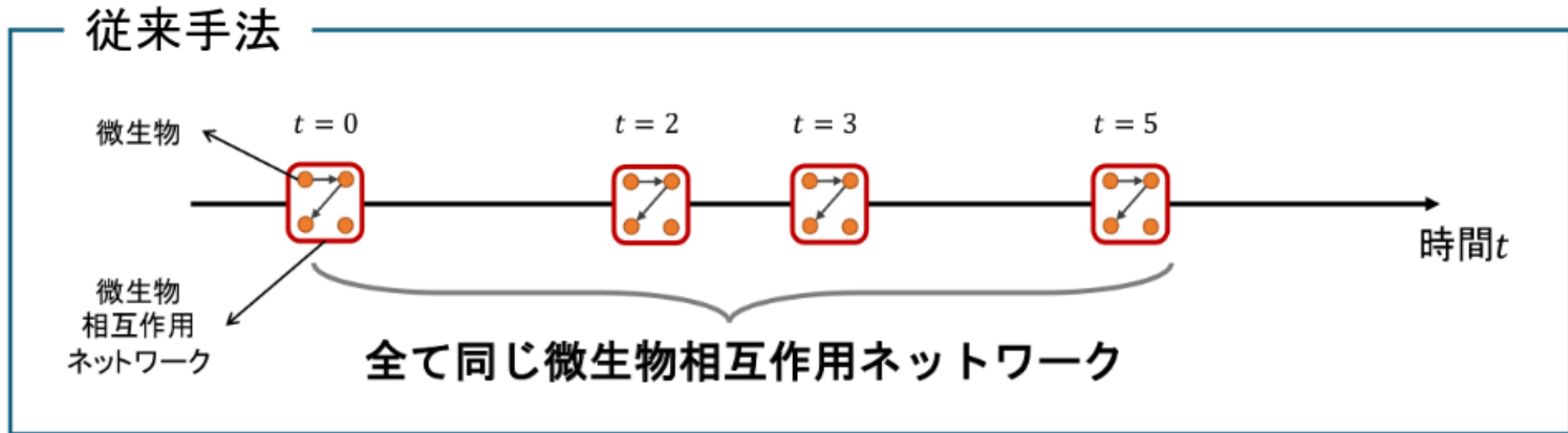
# 微生物相互作用ネットワークを推定したい

- どの微生物がどの微生物の成長に寄与しているのか？



## これまではgLVE (generalized Lotka-Volterra equation)という微分方程式に従って推定

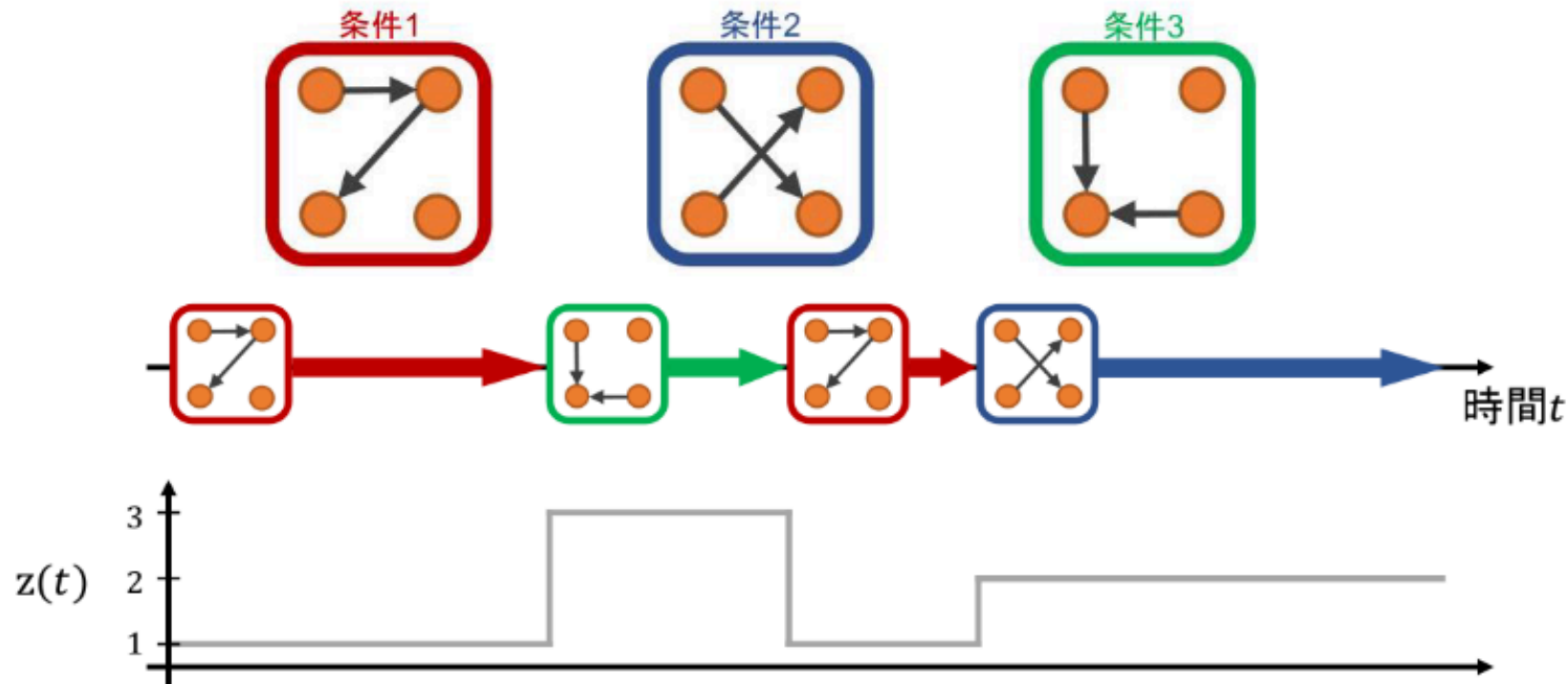
- 時系列微生物存在量データを使用
- gLVEは相互作用ネットワークの時間変化を考慮しない



# 相互作用ネットワークは時間変化するはず

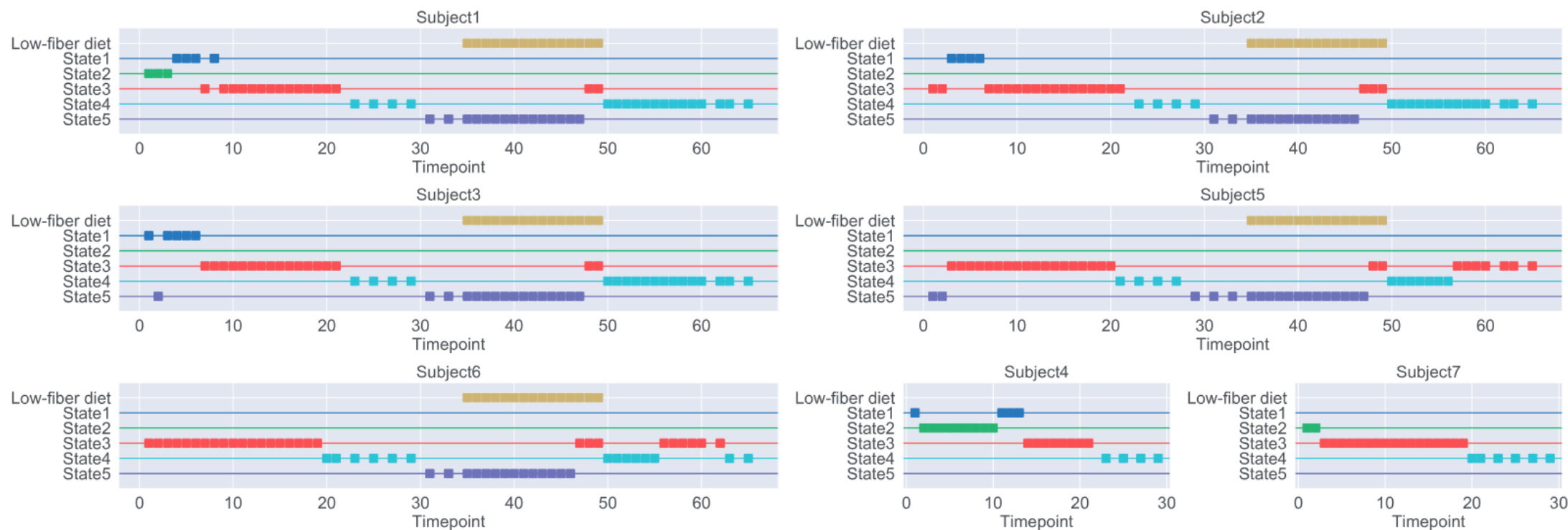
- ベイズ統計モデルを用いて推定

## 提案手法



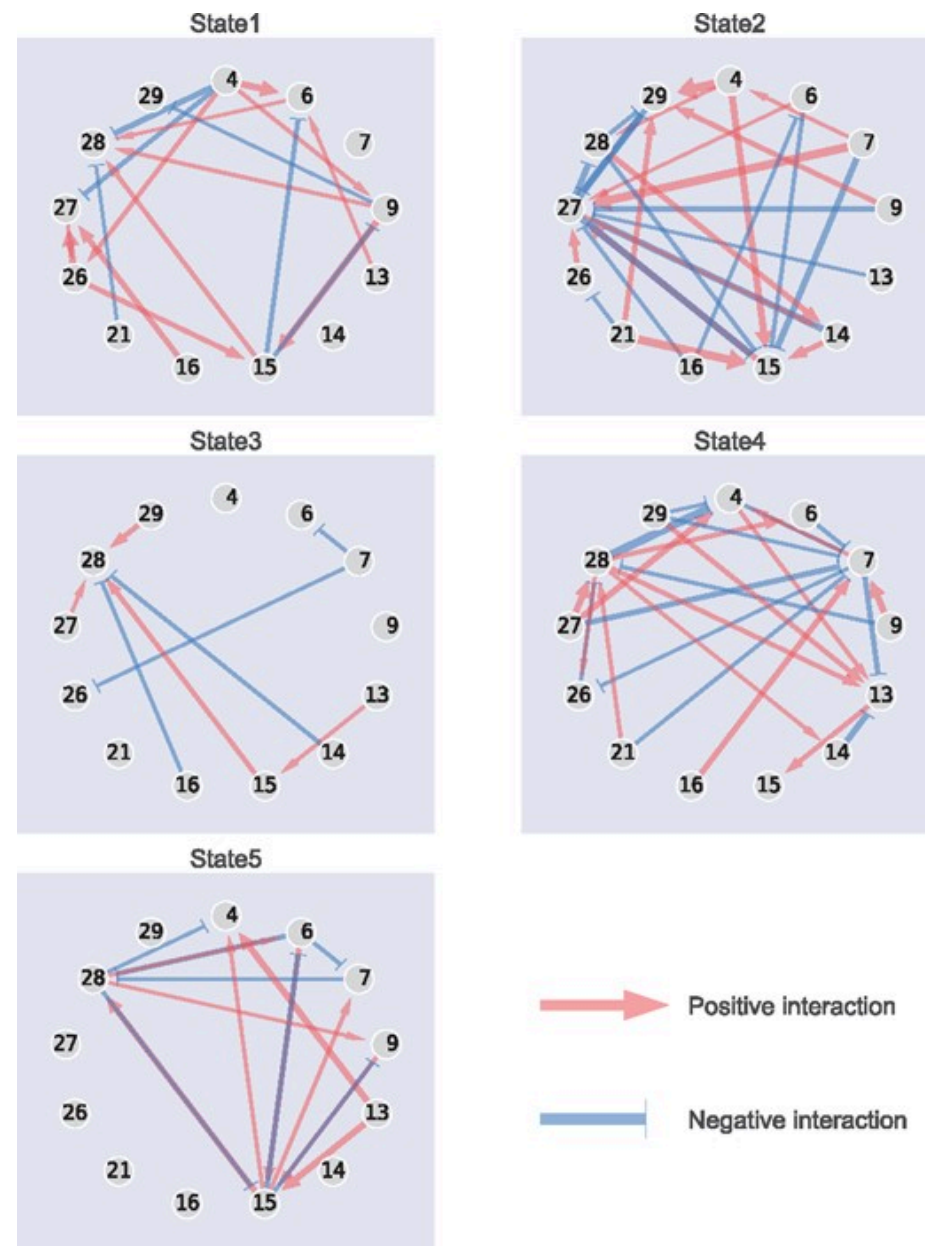
「時間変化」を考慮して微生物相互作用を推定する

# 微生物相互作用ネットワークの時間変化が推定された



「時間変化」を考慮して微生物相互作用を

# 相互作用ネットワークは大きく異なっていた



# 「効果が出る期間のズレ」を考慮してプロバイオティクスの効果を推定する

Computational and Structural Biotechnology Journal 21 (2023) 5350–5357



Contents lists available at [ScienceDirect](#)

Computational and Structural Biotechnology Journal

journal homepage: [www.elsevier.com/locate/csbj](http://www.elsevier.com/locate/csbj)



Research Article

Probiotic responder identification in cross-over trials for constipation using a Bayesian statistical model considering lags between intake and effect periods

Shion Hosoda <sup>a,b,\*</sup>, Yuichiro Nishimoto <sup>c</sup>, Yohsuke Yamauchi <sup>c</sup>, Takuji Yamada <sup>c,d</sup>,  
Michiaki Hamada <sup>a,b,e,\*\*</sup>

<sup>a</sup> Department of Electrical Engineering and Bioscience, Graduate School of Advanced Science and Engineering, Waseda University, Tokyo, Japan

<sup>b</sup> Computational Bio Big-Data Open Innovation Laboratory, AIST-Waseda University, Tokyo, Japan

<sup>c</sup> Metagen Inc., Yamagata, Japan

<sup>d</sup> Department of Life Science and Technology, Tokyo Institute of Technology, Tokyo, Japan

<sup>e</sup> Graduate School of Medicine, Nippon Medical School, Tokyo, Japan



# Introduction

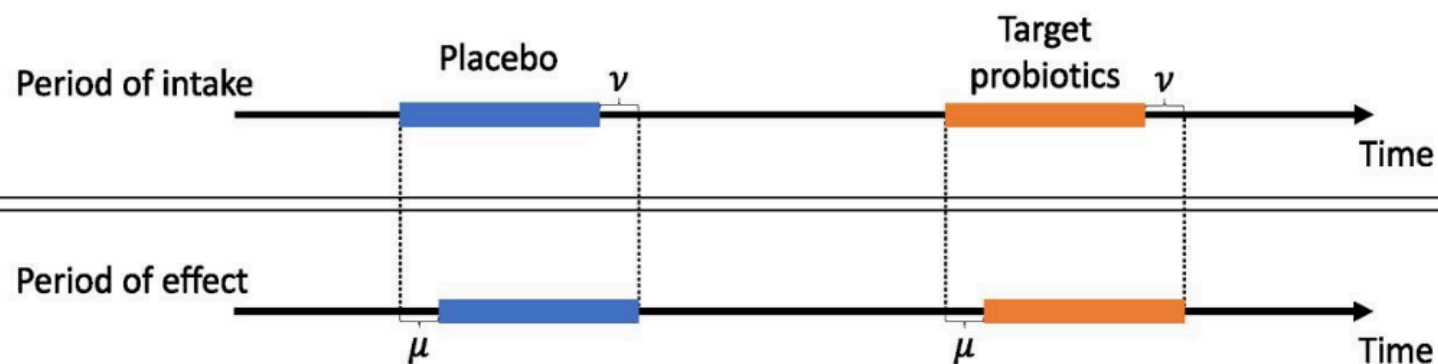
- 目的
  - 排便頻度からその被験者がそのプロバイオティクスのレスポンドー（効果がある人）かどうかを知りたい
- 従来は...
  - プロバイオティクスとプラセボを摂取している間の排便頻度の比を評価
- 問題点
  - プロバイオティクスの効果は摂取してすぐ出るのか？ 摂取をやめたらすぐ効果がなくなるのか？



# Overview



## Given cross-over trial dataset

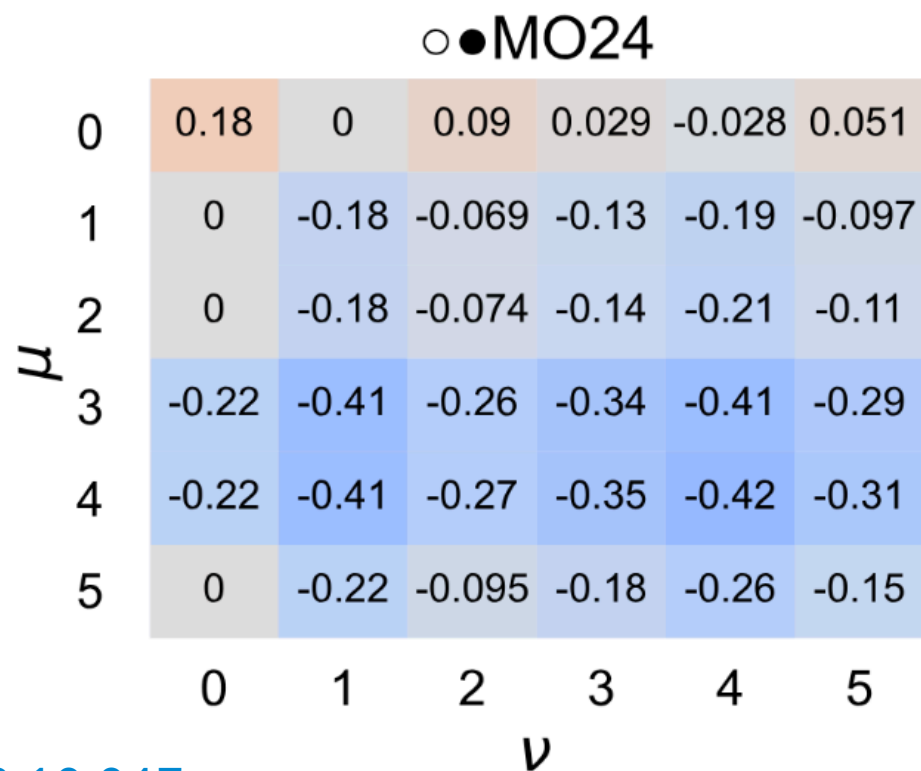


## Estimation by the proposed model

$\mu, \nu$ : Lags of the effect period

# Results

- 従来手法ではレスポnder、提案手法ではレスポnderとはいえないと判断された被験者の排便頻度改善スコア



# 「細菌の所属するグループ」を考慮して細菌の共起関係を推定する

Hosoda et al. *Microbiome* (2020) 8:95  
<https://doi.org/10.1186/s40168-020-00864-3>


Microbiome

RESEARCH

Open Access

## Revealing the microbial assemblage structure in the human gut microbiome using latent Dirichlet allocation



Shion Hosoda<sup>1,2</sup>, Suguru Nishijima<sup>1,2,3</sup>, Tsukasa Fukunaga<sup>1,4</sup>, Masahira Hattori<sup>1,3,5</sup> and Michiaki Hamada<sup>1,2,6,7\*</sup> 

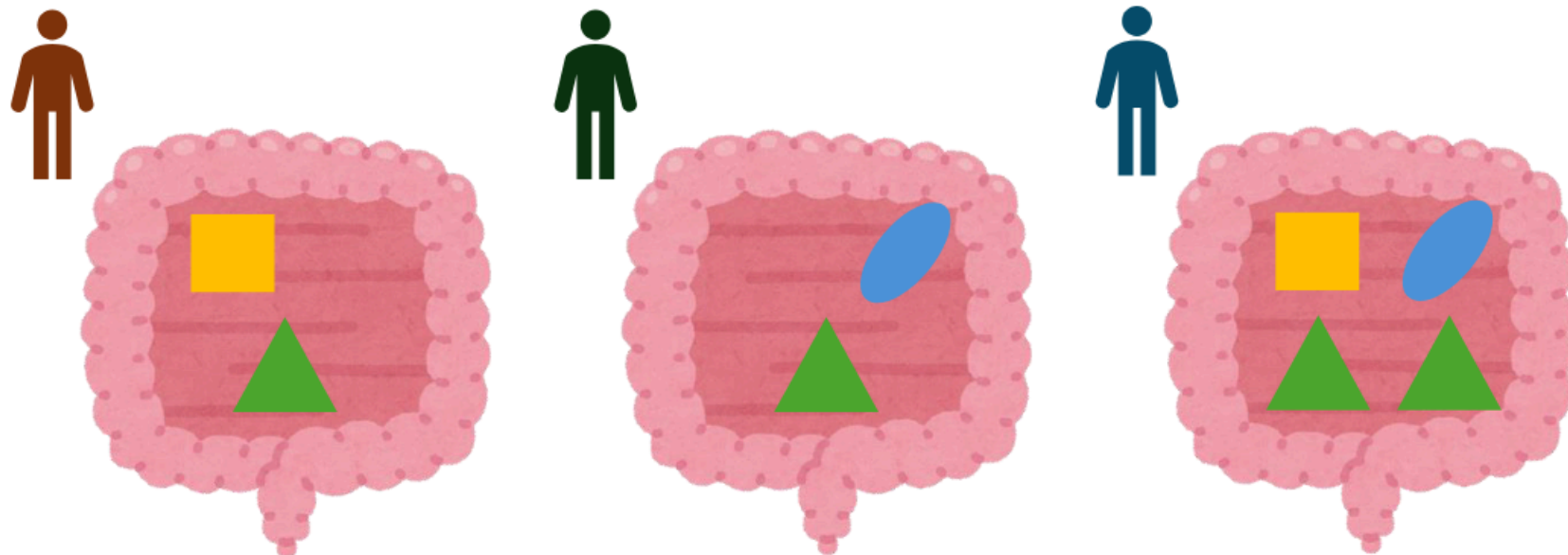
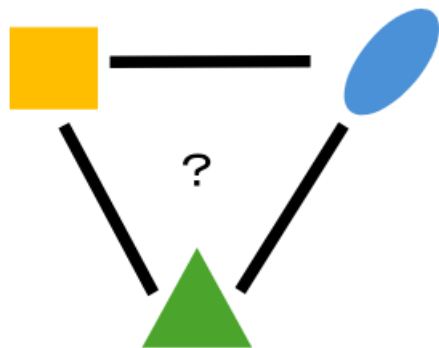
<https://doi.org/10.1186/s40168-020-00864-3>

# Introduction

- 目的
  - ヒト腸内細菌同士の構築する共起関係を推定したい
- 従来は...
  - 相関係数の評価
- 問題点
  - 複数の細菌と複数のパターンで共起する細菌を検知できない

# Overview

腸内細菌はどの組み合わせで存在しやすいのか？



Assemblage 1:

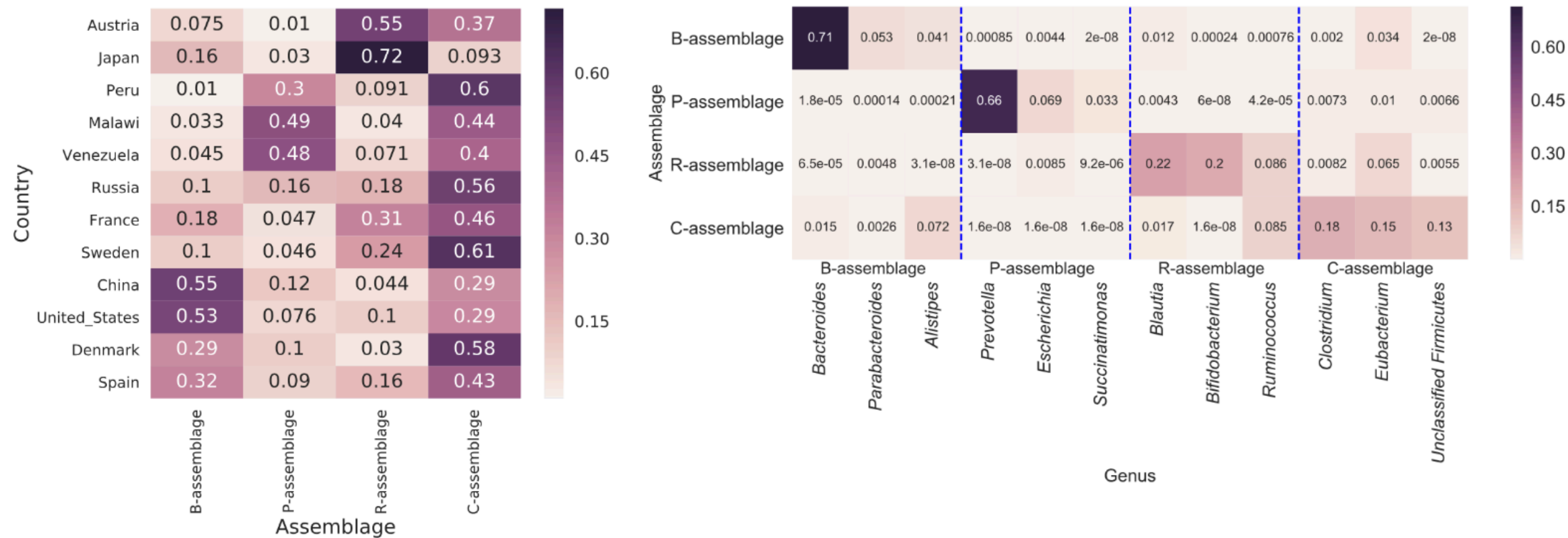


Assemblage 2:



相関係数では捉えられない( $\rho=0.5$ )

# Results



最後に

# まとめ

- 既存の解析手法をベースとして、隠れ変数を導入した統計モデルのコンセプトを紹介した
- 基本的な統計解析だけでは得られなかった意外な結果が得られることもある
  - ただし、うまく行かないことは多い
- 「どうやって」が気になる場合は以下の書籍などを参照
  - 久保拓弥『データ解析のための統計モデリング入門』



# 謝辞

こちらの皆様に感謝いたします。

- 早稲田大学 浜田 道昭 先生
- 早稲田大学 浜田研究室の皆様
- 早稲田大学 福永 津嵩 先生
- 株式会社メタジェン 山田 拓司 先生、西本 悠一郎 博士（当時）、山内 洋輔 様
- 早稲田大学 服部 正平 先生、西嶋 傑 先生（当時）

ご清聴ありがとうございました