

生物学的制約に基づく 5菌種バイオフィルムモデルの パラメータ削減と推定精度向上

Biologically Constrained Parameter Reduction for 5-Species Biofilm Models

IKM Hiwi Research Group

Institute for Continuum Mechanics, Leibniz University Hannover

February 13, 2026

Abstract

口腔バイオフィルムの形成メカニズムを解明するために、5菌種（*S. oralis*, *A. naeslundii*, *Veillonella* spp., *F. nucleatum*, *P. gingivalis*）の相互作用モデルが提案されている。しかし、このモデルは20個の未知パラメータを含み、限られた実験データからの推定において識別可能性の問題（Identifiability Problem）を抱えている。本稿では、実験的に確認された菌種間相互作用ネットワークに基づいてパラメータ空間を削減する手法「生物学的制約に基づくパラメータ削減法」を提案する。本手法は、生物学的根拠に基づいて5つの相互作用パラメータをゼロに固定し、さらに4つの実験条件（Commensal/Dysbiotic × Static/HOBIC）に応じた階層的な事前分布設定を行う。これにより、自由パラメータ数を20から15に削減し、計算コストの低減と推定精度の向上を実現する。初期の数値実験結果は、本手法が標準的な全パラメータ推定と比較して、より生物学的に妥当な解へ収束することを示唆している。

キーワード: ベイズ推定, バイオフィルム, パラメータ削減, TMCMC, 生物学的制約, 多菌種相互作用, 逆問題, インプラント周囲炎, 不確実性定量化

Contents

1 序論	4
2 生物学的根拠	4
2.1 モデルに含まれる菌種	4
2.2 相互作用ネットワーク (Figure 4C)	4
2.3 活性相互作用	4
2.4 不在相互作用 (ロック対象)	5
3 数学的定式化	5
3.1 支配方程式	5
3.2 対称行列の仮定	6
3.3 完全パラメータマッピング	6
4 ベイズ推定の枠組み	6
4.1 順問題 (Forward Model)	6
4.2 ベイズ逆問題	7
4.3 尤度関数	7
4.4 生物学的制約を含む事前分布	7
5 Transitional Markov Chain Monte Carlo (TMCMC)	8
5.1 アルゴリズムの概要	8
5.2 適応的テンパリングスケジュール	8
5.3 リサンプリングとMCMC変異	8
5.4 TMCMCアルゴリズム	9
5.5 モデルエビデンスの推定	9
6 提案手法：生物学的制約に基づくパラメータ削減	10
6.1 事前分布の境界 (Heineらと提案手法の統合戦略)	10
6.2 条件別詳細ロックロジック	10
6.2.1 1. Commensal Static (ロックモード)	10
6.2.2 2. Dysbiotic Static (標準 + 制約モード)	11
6.2.3 3. Commensal HOBIC (高流速モード)	11
6.2.4 4. Dysbiotic HOBIC (発見モード)	11
6.3 4段階推定構成	11
6.4 逐次推定アルゴリズム	11
7 実装	13
7.1 コアモジュール	13
7.2 推定スクリプト	13
8 比較：標準手法 vs 提案手法	13
9 数値実験結果と考察	13
9.1 実験条件と設定	13
9.2 モデル適合度と予測精度の評価	14
9.2.1 MAP推定値による適合確認	14
9.2.2 事後予測分布の不確実性評価	14
9.2.3 残差分析	14
9.3 パラメータ推定結果	15
9.4 菌種間相互作用の推定	15

1 序論

多菌種バイオフィルムの動態を理解することは、歯周病やインプラント周囲炎などの口腔疾患の予防・治療において極めて重要である。Heineら[1]は、インプラント周囲炎に関連する5つの主要な口腔細菌種（*S. oralis*, *A. naeslundii*, *Veillonella* spp., *F. nucleatum*, *P. gingivalis*）の相互作用について実験的な検討を行い、そのダイナミクスを明らかにした。

これらの生物学的知見に基づき、JunkerとBalzani[2]によって提案された拡張ハミルトン原理（Extended Hamilton Principle）の理論的枠組みを用いて、Klemptら[3]は新規な相互作用スキームを有する多菌種バイオフィルムの連続体モデルを構築した。さらに、Fritschら[4]は、サロゲートモデルを用いたハイブリッドな不確実性下での細菌マイクロフィルムのベイズ更新手法について論じている。

本稿では、これらの先行研究を基礎としつつ、5菌種モデルのパラメータ推定における課題に対処するための新たなアルゴリズムを提案する。このモデルは20個の未知パラメータ（増殖率や相互作用係数）を含み、限られた実験データからの推定において識別可能性の問題（Identifiability Problem）を抱えている。標準的なアプローチでは、全てのパラメータを自由変数として推定するため、以下の問題が生じる：

- ・ 限られた実験データによる識別可能性の低下
- ・ 生物学的に妥当でないパラメータ推定値への収束
- ・ 不必要なパラメータ空間の探索による計算効率の低下

本研究では、これらの問題を解決するために「生物学的制約に基づくパラメータ削減法」を提案する。本手法の核心は、実験生物学の知見を数理的に定式化し、推定プロセスに「ハード制約」として組み込むことである。具体的には、Heineらによって明らかにされた相互作用ネットワーク（Figure 1）に基づき、直接的な相互作用が存在しない菌種ペアのパラメータをゼロに固定する。

2 生物学的根拠

2.1 モデルに含まれる菌種

本モデルは口腔バイオフィルムに一般的に見られる5つの細菌種を含む（表1）。

ID	菌種	略称	役割
0	<i>Streptococcus oralis</i>	S.o	初期定着菌
1	<i>Actinomyces naeslundii</i>	A.n	初期定着菌
2	<i>Veillonella</i> spp.	Vei	代謝ブリッジ
3	<i>Fusobacterium nucleatum</i>	F.n	ブリッジ微生物
4	<i>Porphyromonas gingivalis</i>	P.g	後期定着菌（病原菌）

Table 1: 5菌種バイオフィルムモデルに含まれる菌種

2.2 相互作用ネットワーク（Figure 4C）

実験観察[1]に基づき、以下の相互作用ネットワークが確立された（図1）。

2.3 活性相互作用

以下の菌種ペアは直接的な生物学的相互作用を持つ（表2）。

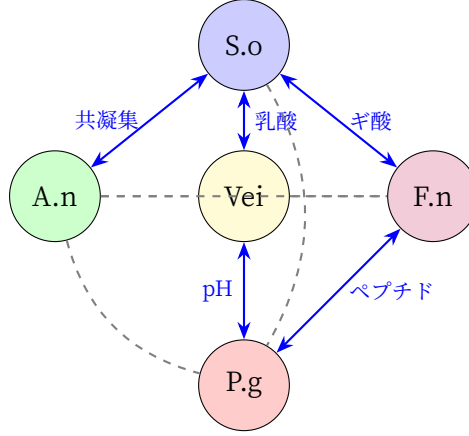


Figure 1: Figure 4Cから導出された菌種相互作用ネットワーク．青い実線は活性相互作用，灰色の破線は不在相互作用を示す．

菌種ペア	メカニズム	タイプ
S.o ↔ A.n	共凝集	双方向
S.o ↔ Ve.i	乳酸産生/消費	双方向
S.o ↔ F.n	ギ酸/酢酸共生	双方向
Ve.i ↔ P.g	pH上昇支援	正のみ
F.n ↔ P.g	共凝集, ペプチド供給	双方向

Table 2: 生物学的メカニズムを持つ活性菌種相互作用

2.4 不在相互作用（ロック対象）

以下の菌種ペアは実験的証拠に基づき直接相互作用を持たない（表3）．これらが本手法におけるパラメータ削減の根拠となる．

Index	パラメータ	菌種ペア	行列	理由
6	a_{34}	Ve.i ↔ F.n	$A[2, 3]$	代謝経路なし
12	a_{23}	A.n ↔ Ve.i	$A[1, 2]$	代謝リンクなし
13	a_{24}	A.n ↔ F.n	$A[1, 3]$	相互作用なし
16	a_{15}	S.o ↔ P.g	$A[0, 4]$	相互作用なし
17	a_{25}	A.n ↔ P.g	$A[1, 4]$	相互作用なし

Table 3: 本手法でゼロに固定される不在相互作用

3 数学的定式化

3.1 支配方程式

5菌種バイオフィームモデルは，細菌体積分率 ϕ_i と生存率 ψ_i の動態を連立常微分方程式系で記述する．菌種 i の相互作用項は次式で与えられる：

$$I_i = \sum_{j=0}^4 A_{ij} \phi_j \psi_j \quad (1)$$

ここで A_{ij} は菌種 j が菌種 i に与える影響， $\phi_j \psi_j$ は生菌体積分率である．

3.2 対称行列の仮定

重要な仮定として、相互作用行列 A は対称であると仮定する：

$$A_{ij} = A_{ji} \quad \forall i, j \in \{0, 1, 2, 3, 4\} \quad (2)$$

これにより非対角相互作用パラメータ数が20から10に削減される．例えば，S.o（菌種0）とVei（菌種2）間の乳酸ハンドオーバー相互作用は単一パラメータで表される：

$$A_{02} = A_{20} = \theta_{10} \quad (\text{コード内では} a_{13} \text{として格納}) \quad (3)$$

3.3 完全パラメータマッピング

全20パラメータベクトル θ は5ブロックに整理され（表4），本手法ではそのうち5つがロック（赤色部分）される．

Idx	名前	行列	菌種ペア	役割	状態
0	a_{11}	$A[0, 0]$	S.o自己	自己調節	自由
1	a_{12}	$A[0, 1]$	S.o \leftrightarrow A.n	共凝集	自由
2	a_{22}	$A[1, 1]$	A.n自己	自己調節	自由
3	b_1	$b[0]$	S.o	減衰率	自由
4	b_2	$b[1]$	A.n	減衰率	自由
5	a_{33}	$A[2, 2]$	Vei自己	自己調節	自由
6	a_{34}	$A[2, 3]$	Vei \leftrightarrow F.n	なし	ロック
7	a_{44}	$A[3, 3]$	F.n自己	自己調節	自由
8	b_3	$b[2]$	Vei	減衰率	自由
9	b_4	$b[3]$	F.n	減衰率	自由
10	a_{13}	$A[0, 2]$	S.o \leftrightarrow Vei	乳酸	自由
11	a_{14}	$A[0, 3]$	S.o \leftrightarrow F.n	ギ酸	自由
12	a_{23}	$A[1, 2]$	A.n \leftrightarrow Vei	なし	ロック
13	a_{24}	$A[1, 3]$	A.n \leftrightarrow F.n	なし	ロック
14	a_{55}	$A[4, 4]$	P.g自己	自己調節	自由
15	b_5	$b[4]$	P.g	減衰率	自由
16	a_{15}	$A[0, 4]$	S.o \leftrightarrow P.g	なし	ロック
17	a_{25}	$A[1, 4]$	A.n \leftrightarrow P.g	なし	ロック
18	a_{35}	$A[2, 4]$	Vei \leftrightarrow P.g	pH	自由*
19	a_{45}	$A[3, 4]$	F.n \leftrightarrow P.g	共凝集	自由

Table 4: パラメータマッピング．赤い行はロック対象．*条件により境界が異なる．

4 ベイズ推定の枠組み

4.1 順問題（Forward Model）

順モデル $\mathcal{M}(\theta)$ は、パラメータベクトル $\theta \in \mathbb{R}^{20}$ を各菌種の時系列予測にマッピングする．初期条件 $\phi(t_0)$ および $\psi(t_0)$ が与えられたとき、拡張ハミルトン原理[2, 3]から導出される連立常微分方程式系は、各菌種 i の生菌体積分率 $\phi_i \psi_i$ の時間発展を支配する：

$$\frac{d(\phi_i \psi_i)}{dt} = \left(A_{ii} + \sum_{j \neq i} A_{ij} \phi_j \psi_j \right) \phi_i \psi_i - b_i \phi_i \psi_i, \quad i = 0, \dots, 4 \quad (4)$$

ここで、第1項は自己調節（ A_{ii} ）と菌種間相互作用（ $A_{ij}, j \neq i$ ）による成長を、第2項は菌種 i の減衰率 b_i による衰退を表す．相互作用行列 A と減衰ベクトル \mathbf{b} は、表4のマッピングにより θ から構築される．順モデルの出力は、各観測時刻における予測相対存在量ベクトルである：

$$\hat{\mathbf{y}}(t_k; \theta) = \mathcal{M}(\theta)|_{t=t_k}, \quad k = 1, \dots, N_t \quad (5)$$

式(4)は対称相互作用を持つ一般化Lotka-Volterra競争モデルであり、Klemptら[3]の変分定式化から自然に導かれる構造である。

4.2 バイズ逆問題

バイズパラメータ推定の目的は、観測された実験データ $\mathbf{y}_{\text{obs}} = \{y_{\text{obs},i}(t_k)\}_{i,k}$ が与えられたときのモデルパラメータ $\boldsymbol{\theta}$ の事後分布 $p(\boldsymbol{\theta} | \mathbf{y}_{\text{obs}})$ を推論することである。バイズの定理より：

$$p(\boldsymbol{\theta} | \mathbf{y}_{\text{obs}}) = \frac{p(\mathbf{y}_{\text{obs}} | \boldsymbol{\theta}) p(\boldsymbol{\theta})}{p(\mathbf{y}_{\text{obs}})} \quad (6)$$

ここで、 $p(\mathbf{y}_{\text{obs}} | \boldsymbol{\theta})$ はデータとモデルの適合度を定量化する尤度関数、 $p(\boldsymbol{\theta})$ は生物学的制約を符号化する事前分布、 $p(\mathbf{y}_{\text{obs}}) = \int p(\mathbf{y}_{\text{obs}} | \boldsymbol{\theta}) p(\boldsymbol{\theta}) d\boldsymbol{\theta}$ はモデルエビデンス（正規化定数）である。事後分布はパラメータ推定の全不確実性を捕捉し、そこから事後確率最大（MAP）推定値および事後平均などの点推定値を導出できる：

$$\hat{\boldsymbol{\theta}}_{\text{MAP}} = \arg \max_{\boldsymbol{\theta}} p(\boldsymbol{\theta} | \mathbf{y}_{\text{obs}}), \quad \hat{\boldsymbol{\theta}}_{\text{mean}} = \mathbb{E}[\boldsymbol{\theta} | \mathbf{y}_{\text{obs}}] = \int \boldsymbol{\theta} p(\boldsymbol{\theta} | \mathbf{y}_{\text{obs}}) d\boldsymbol{\theta} \quad (7)$$

4.3 尤度関数

菌種間・時刻間で独立なガウス測定誤差を仮定すると、尤度関数は以下の形をとる：

$$p(\mathbf{y}_{\text{obs}} | \boldsymbol{\theta}) = \prod_{k=1}^{N_t} \prod_{i=0}^4 \frac{1}{\sqrt{2\pi} \sigma_i} \exp\left(-\frac{(y_{\text{obs},i}(t_k) - \hat{y}_i(t_k; \boldsymbol{\theta}))^2}{2\sigma_i^2}\right) \quad (8)$$

ここで $y_{\text{obs},i}(t_k)$ は時刻 t_k における菌種 i の観測相対存在量、 $\hat{y}_i(t_k; \boldsymbol{\theta})$ はモデル予測、 σ_i は菌種 i の測定ノイズ標準偏差である。対応する対数尤度は：

$$\ell(\boldsymbol{\theta}) = -\frac{1}{2} \sum_{k=1}^{N_t} \sum_{i=0}^4 \left[\frac{(y_{\text{obs},i}(t_k) - \hat{y}_i(t_k; \boldsymbol{\theta}))^2}{\sigma_i^2} + \log(2\pi\sigma_i^2) \right] \quad (9)$$

実用上、 σ_i は反復測定から推定するか、ハイパーパラメータとして扱う。反復データが利用できない場合は、 σ_i を菌種 i の観測データ範囲の一定割合に設定することが一般的であり、これは期待される測定不確実性を反映する。

4.4 生物学的制約を含む事前分布

事前分布 $p(\boldsymbol{\theta})$ は、相互作用ネットワークに基づく生物学的制約とパラメータロック機構の双方を符号化する。事前分布の各成分間の独立性を仮定すると：

$$p(\boldsymbol{\theta}) = \prod_{k=0}^{19} p(\theta_k) \quad (10)$$

ここで各成分の事前分布は：

$$p(\theta_k) = \begin{cases} \delta(\theta_k) & k \in \mathcal{L} \text{ の場合（ディラックのデルタ：ゼロに固定）} \\ \frac{1}{u_k - l_k} \mathbf{1}_{[l_k, u_k]}(\theta_k) & k \notin \mathcal{L} \text{ の場合（自由パラメータの一樣事前分布）} \end{cases} \quad (11)$$

ここで $\delta(\cdot)$ はロックパラメータに対して $\theta_k = 0$ を強制するディラックのデルタ関数であり、 $\mathbf{1}_{[l_k, u_k]}$ は $[l_k, u_k]$ 上の指示関数である。この定式化は、事後分布を制約部分空間 $\Theta_{\mathcal{L}} = \{\boldsymbol{\theta} \in \mathbb{R}^{20} : \theta_k = 0, \forall k \in \mathcal{L}\}$ に射影することと等価である。

実装上は、自由パラメータのみからなる縮約ベクトル $\boldsymbol{\theta}_{\text{free}} \in \mathbb{R}^{n_{\text{free}}}$ に対してのみ推論が行われ、ロックパラメータは推論過程を通じてゼロに固定される。 \mathbb{R}^{20} から $\mathbb{R}^{n_{\text{free}}}$ への次元削減は、マルコフ連鎖の混合時間が一般に次元数とともに増加するため[5]、MCMCサンプリングの効率を直接的に向上させる。

5 Transitional Markov Chain Monte Carlo (TMCMC)

5.1 アルゴリズムの概要

Transitional Markov Chain Monte Carlo (TMCMC) アルゴリズムは, Ching and Chen[6]によって提案された逐次モンテカルロ法であり, 複雑かつ多峰性を持つ事後分布からのサンプリングに適している. 標準的なMCMC法 (例: Metropolis-Hastings法, Gibbsサンプリング) が高次元空間や多峰性空間で混合不良に陥る可能性があるのに対し, TMCMCは事前分布から事後分布へ, 一連の中間的な「テンパリング」分布を通じて段階的にサンプルを変換する[7].

核となるアイデアは, テンパリングスケジュール $0 = \beta_0 < \beta_1 < \dots < \beta_M = 1$ を定義し, 中間分布の列を構成することである:

$$p_m(\boldsymbol{\theta}) \propto p(\mathbf{y}_{\text{obs}} | \boldsymbol{\theta})^{\beta_m} p(\boldsymbol{\theta}), \quad m = 0, 1, \dots, M \quad (12)$$

$\beta_0 = 0$ では $p_0(\boldsymbol{\theta}) = p(\boldsymbol{\theta})$ (事前分布) に, $\beta_M = 1$ では $p_M(\boldsymbol{\theta}) = p(\boldsymbol{\theta} | \mathbf{y}_{\text{obs}})$ (完全な事後分布) に帰着する. 尤度を漸進的に導入することにより, 事前分布と事後分布が大きく離れている場合に標準的な重点サンプリングが失敗する「prior-posterior gap」問題を回避する.

5.2 適応的テンパリングスケジュール

各ステージ m において, 次のテンパリングパラメータ β_{m+1} は, 重要度重みの退化を制御するように適応的に選択される. Betzら[8]に従い, 重要度重みの変動係数 (CoV) が以下を満たすように β_{m+1} を決定する:

$$\text{CoV}[\{w_j^{(m)}\}_{j=1}^N] = \frac{\text{Std}[w_j^{(m)}]}{\text{Mean}[w_j^{(m)}]} = \delta_{\text{target}} \quad (13)$$

ここで重要度重みは次式で計算される:

$$w_j^{(m)} = p(\mathbf{y}_{\text{obs}} | \boldsymbol{\theta}_j^{(m)})^{\beta_{m+1} - \beta_m}, \quad j = 1, \dots, N \quad (14)$$

$\delta_{\text{target}} \in (0, 2]$ はユーザ指定の目標値である (典型的には $\delta_{\text{target}} = 1.0$). 式(13)は $(\beta_m, 1]$ 上の二分法により β_{m+1} について解かれる. この適応的スキームにより, ステージ数 M を事前に指定する必要がなくなり, 中間分布間の滑らかな遷移が保証される.

5.3 リサンプリングとMCMC変異

各ステージ m において, アルゴリズムは以下の3ステップで進行する:

1. **リサンプリング:** 正規化された重要度重み $\bar{w}_j^{(m)} = w_j^{(m)} / \sum_{l=1}^N w_l^{(m)}$ に比例する確率で, 現在の粒子集団 $\{\boldsymbol{\theta}_j^{(m)}\}_{j=1}^N$ から N 個の粒子を復元抽出する.
2. **共分散推定:** 重み付き標本共分散行列を計算する:

$$\boldsymbol{\Sigma}^{(m)} = \sum_{j=1}^N \bar{w}_j^{(m)} \left(\boldsymbol{\theta}_j^{(m)} - \bar{\boldsymbol{\theta}}^{(m)} \right) \left(\boldsymbol{\theta}_j^{(m)} - \bar{\boldsymbol{\theta}}^{(m)} \right)^T \quad (15)$$

ここで $\bar{\boldsymbol{\theta}}^{(m)} = \sum_{j=1}^N \bar{w}_j^{(m)} \boldsymbol{\theta}_j^{(m)}$ は重み付き平均である. この共分散行列はテンパリング事後分布の局所的な形状にMCMC提案分布を適応させる.

3. **MCMC変異:** リサンプリングされた各粒子に対して, ガウス提案分布を用いた Metropolis–Hastingsステップを実行する:

$$q(\boldsymbol{\theta}^* | \boldsymbol{\theta}_j) = \mathcal{N}(\boldsymbol{\theta}_j, \gamma^2 \boldsymbol{\Sigma}^{(m)}) \quad (16)$$

ここで $\gamma > 0$ はスケーリング係数 (典型的には $\gamma^2 = 0.04$) である. 採択確率は:

$$\alpha = \min\left(1, \frac{p(\mathbf{y}_{\text{obs}} | \boldsymbol{\theta}^*)^{\beta_{m+1}} p(\boldsymbol{\theta}^*)}{p(\mathbf{y}_{\text{obs}} | \boldsymbol{\theta}_j)^{\beta_{m+1}} p(\boldsymbol{\theta}_j)}\right) \quad (17)$$

この変異ステップは粒子集団を多様化し, サンプルの枯渇を防止する.

5.4 TMCMCアルゴリズム

生物学的制約の強制を含む完全なTMCMC手続きをアルゴリズム2にまとめる.

Algorithm 1: Transitional Markov Chain Monte Carlo (TMCMC)

入力: 事前分布 $p(\boldsymbol{\theta})$, 尤度 $p(\mathbf{y}_{\text{obs}} | \boldsymbol{\theta})$, 粒子数 N , ロック集合 \mathcal{L} , 目標CoV δ_{target}

出力: 重み付き事後サンプル $\{\boldsymbol{\theta}_j^{(M)}\}_{j=1}^N$

1. 初期化: $\{\boldsymbol{\theta}_j^{(0)}\}_{j=1}^N \sim p(\boldsymbol{\theta})$ を生成; $\beta_0 = 0, m = 0$ と設定
2. While $\beta_m < 1$:
 - (a) 二分法により $\text{CoV}[\{w_j^{(m)}\}] = \delta_{\text{target}}$ を満たす $\beta_{m+1} \in (\beta_m, 1]$ を求める
 - (b) 重み $w_j^{(m)} = p(\mathbf{y}_{\text{obs}} | \boldsymbol{\theta}_j^{(m)})^{\beta_{m+1} - \beta_m}$ を計算 ($j = 1, \dots, N$)
 - (c) 現在のサンプルから重み付き共分散 $\boldsymbol{\Sigma}^{(m)}$ を計算
 - (d) 確率 $\propto w_j^{(m)}$ で $\{\boldsymbol{\theta}_j^{(m)}\}$ から N 粒子をリサンプリング
 - (e) 各粒子 $j = 1, \dots, N$ に対して:
 - i. $\boldsymbol{\theta}^* \sim \mathcal{N}(\boldsymbol{\theta}_j, \gamma^2 \boldsymbol{\Sigma}^{(m)})$ を提案
 - ii. 制約の強制: $k \in \mathcal{L}$ に対して $\theta_k^* = 0$; 事前分布境界 $[l_k, u_k]$ にクリップ
 - iii. 確率 α (式17) で $\boldsymbol{\theta}_j^{(m+1)} = \boldsymbol{\theta}^*$ を採択; そうでなければ $\boldsymbol{\theta}_j^{(m+1)} = \boldsymbol{\theta}_j$ を保持
 - (f) $m \leftarrow m + 1$
3. 事後サンプル $\{\boldsymbol{\theta}_j^{(M)}\}_{j=1}^N$ を返却

Figure 2: 生物学的制約を組み込んだTMCMCアルゴリズム. ステップ2(e)(ii)により, ロックパラメータがゼロに維持され, 全ての自由パラメータが条件固有の事前分布境界を遵守することが保証される.

5.5 モデルエビデンスの推定

TMCMCの有用な副産物として, モデルエビデンス $p(\mathbf{y}_{\text{obs}})$ の不偏推定量が得られる:

$$\hat{p}(\mathbf{y}_{\text{obs}}) = \prod_{m=0}^{M-1} \left(\frac{1}{N} \sum_{j=1}^N w_j^{(m)} \right) \quad (18)$$

この量は、標準手法（20パラメータ）と提案手法（15パラメータ）のベイズモデル比較をベイズ因子を通じて可能にする：

$$B_{10} = \frac{p(\mathbf{y}_{\text{obs}} \mid \mathcal{M}_{\text{proposed}})}{p(\mathbf{y}_{\text{obs}} \mid \mathcal{M}_{\text{standard}})} \quad (19)$$

ベイズ因子 $B_{10} > 1$ は提案された縮約モデルを支持するエビデンスを提供し、生物学的制約が解釈性のみならず、モデル複雑度に対する予測性能も向上させることを示す。

6 提案手法：生物学的制約に基づくパラメータ削減

6.1 事前分布の境界（Heineらと提案手法の統合戦略）

Heine et al. (2025) の生物学的エビデンスと 提案アルゴリズムの特性を融合させた、条件別事前分布設定を表5に示す。各条件は生物学的仮説（病原菌の有無、流速の影響など）に基づいて個別に最適化されている。

条件 (モード)	生物学的仮説	増殖率 (b) 設定	相互作用 (a) 設定	注記
1. Commensal Static (Lock Mode)	病原菌は検出限界以下であり、健康な Commensal 菌種のみが定着する。	Lock: $b_4, b_5 = 0$ (Purple/Red完全抑制)	Lock: 病原菌関連 = 0 Est: Commensal [−1, 2]	Baseline
2. Dysbiotic Static (Restricted Mode)	病原菌は存在するが、静置環境では爆発的増殖 (Surge) に至らない。	b_5 (Red) $\in [0, 0.2]$ (低増殖に制限)	a_{33} (Orange Self) $\in [1, 5]$ (自己抑制を強化)	No Surge
3. Commensal HOBIC (High-Flow Mode)	高流速下で栄養供給が増加し、 <i>S.oralis</i> (Blue) が特異的に増殖する。	b_1 (Blue) $\in [1, 10]$ (高流速適応) b_5 (Red) = 0	標準設定 [−1, 1]	Blue Bloom
4. Dysbiotic HOBIC (Discovery Mode)	隠れた協力関係により、 <i>V.parvula</i> (Orange) と <i>P.gingivalis</i> (Red) が共生的に爆発増殖する。	b_3 (Orange) $\in [0, 5]$ (爆発的増殖)	Coop: $a_{13}, a_{35}, a_{45} \in [-3, 0]$ (負の相互作用=協力) Self: $a_{ii} \in [0, 5]$	Surge

Table 5: 実験条件別の詳細な事前分布設定と生物学的仮説（Heineらと提案手法の統合戦略）

6.2 条件別詳細ロックロジック

6.2.1 1. Commensal Static（ロックモード）

病原菌（Purple/Red）が増殖しないことが実験的に確認されている（Undetectable）。これを再現するため、病原菌に関連するパラメータを強制的にゼロに固定する。

- **ロック対象:** b_4, b_5 および $a_{4j}, a_{5j}, a_{i4}, a_{i5}$.
- **推定対象:** 健康な Commensal 菌種（Blue, Green, Yellow/Orange）の相互作用のみ。

6.2.2 2. Dysbiotic Static (標準 + 制約モード)

病原菌は存在するが、静置環境 (Static) では爆発的な増殖 (Surge) は起きない。標準的な探索範囲を採用しつつ、Redの増殖を低めに抑える。

- **設定:** Redの増殖率 b_5 を $[0.0, 0.2]$ に制限。
- **特徴:** Orangeの自己抑制 a_{33} を強め $[1.0, 5.0]$ に設定し、過剰増殖を防ぐ。

6.2.3 3. Commensal HOBIC (高流速モード)

高流速環境 (HOBIC) では、栄養供給が豊富なためBlue (S. oralis) が活発に増殖する。一方、Redは定着できず検出されない。

- **設定:** Blueの増殖率 b_1 の上限を 10.0 まで拡張。
- **ロック:** Redの増殖率 b_5 を 0.0 に固定。

6.2.4 4. Dysbiotic HOBIC (発見モード)

病原菌の爆発的な増殖 (Surge) と、それを支える「隠れた協力関係」を探索する最も困難な条件。探索空間を意図的に広げ、未知の相互作用メカニズムを発見する。この条件では本手法のロックの一部を解除する場合がある。

6.3 4段階推定構成

パラメータ推定は以下の4段階で逐次的に実行される。この構成はパラメータ間の相関 (Coupling) と探索空間の次元数を考慮して最適化されている。

段階	推定対象	パラメータ数	内容
1	M1 (菌種1-2)	5	$a_{11}, a_{12}, a_{22}, b_1, b_2$
2	M2 (菌種3-4)	5	$a_{33}, a_{34}, a_{44}, b_3, b_4$
3	M3+M4	6	$a_{13}, a_{14}, a_{23}, a_{24}, a_{55}, b_5$
4	M5 (P.gクロス)	4	$a_{15}, a_{25}, a_{35}, a_{45}$

Table 6: 4段階推定構成

6.4 逐次推定アルゴリズム

4段階逐次推定手続きをアルゴリズム3に形式化する。各段階 s において、TMCMC (アルゴリズム2) は活性パラメータ集合 $\mathcal{A}_s = \mathcal{P}_s \setminus \mathcal{L}$ のみの推定に適用され、以前の段階で推定されたパラメータはMAP推定値に固定される。

逐次分解は以下の理論的・実用的利点を提供する：

- **次元削減:** 各段階は低次元部分空間 ($|\mathcal{A}_s| \leq 6$) で動作するため、適度な粒子数でもTMCMCによる効率的な探索が可能となる。
- **条件付き識別可能性:** 以前に推定されたパラメータを条件付けることで、残りのパラメータの識別可能性が向上し、高次元ベイズ推論に固有の次元の呪いが軽減される。
- **生物学的整合性:** 段階のグループ分けは菌種の定着の時間的順序 (初期定着菌→ブリッジ微生物→後期定着菌) を反映しており、推定順序が基盤となる生態学的プロセスと整合する。

Algorithm 2: 4段階逐次パラメータ推定

入力: 観測データ \mathbf{y}_{obs} , 段階分割 $\{\mathcal{P}_s\}_{s=1}^4$, ロック集合 \mathcal{L} , 事前分布境界 $\{[l_k, u_k]\}$
出力: 完全MAP推定値 $\hat{\boldsymbol{\theta}}_{\text{MAP}}$ および段階別事後サンプル

1. 初期化: $\boldsymbol{\theta}_{\text{base}} \in \mathbb{R}^{20}$ を設定; $k \in \mathcal{L}$ に対して $\theta_{\text{base},k} = 0$

2. 各段階 $s = 1, 2, 3, 4$ に対して:

(a) 活性インデックスを定義: $\mathcal{A}_s = \mathcal{P}_s \setminus \mathcal{L}$

(b) 縮約事前分布を構築: $p_s(\boldsymbol{\theta}_{\mathcal{A}_s}) = \prod_{k \in \mathcal{A}_s} \text{Uniform}(\theta_k; l_k, u_k)$

(c) 段階尤度を定義:

$$p_s(\mathbf{y}_{\text{obs}} \mid \boldsymbol{\theta}_{\mathcal{A}_s}) = p(\mathbf{y}_{\text{obs}} \mid \boldsymbol{\theta}_{\text{base}} \oplus_{\mathcal{A}_s} \boldsymbol{\theta}_{\mathcal{A}_s})$$

ここで $\oplus_{\mathcal{A}_s}$ は $\boldsymbol{\theta}_{\text{base}}$ のインデックス \mathcal{A}_s の成分を置換する操作

(d) TMCMC (アルゴリズム2) を事前分布 p_s と尤度 $p_s(\mathbf{y}_{\text{obs}} \mid \cdot)$ で実行し, N 個の事後サンプル $\{\boldsymbol{\theta}_{\mathcal{A}_s,j}\}_{j=1}^N$ を取得

(e) MAPを抽出: $\hat{\boldsymbol{\theta}}_{\mathcal{A}_s}^{\text{MAP}} = \arg \max_j p_s(\mathbf{y}_{\text{obs}} \mid \boldsymbol{\theta}_{\mathcal{A}_s,j}) p_s(\boldsymbol{\theta}_{\mathcal{A}_s,j})$

(f) 推定値を固定: $k \in \mathcal{A}_s$ に対して $\theta_{\text{base},k} \leftarrow \hat{\theta}_k^{\text{MAP}}$

3. $\hat{\boldsymbol{\theta}}_{\text{MAP}} = \boldsymbol{\theta}_{\text{base}}$ と全段階の事後サンプルを返却

Figure 3: 逐次推定アルゴリズム. 各段階は生物学的に整合性のあるパラメータ部分集合を推定し, 以前に推定されたパラメータを固定することで, 各ステップの実効次元数を削減する.

- 計算効率: 総計算コストは $C(n_{\text{free}})$ ではなく $\sum_{s=1}^4 C(|\mathcal{A}_s|)$ にスケーリングする. ここで $C(d)$ は d 次元でのTMCMCコストである. $C(d)$ は一般に d に対して超線形に増加するため, 逐次アプローチは大幅に効率的である.

7 実装

7.1 コアモジュール

```
1 LOCKED_INDICES = [6, 12, 13, 16, 17]
2
3 def get_constrained_bounds():
4     bounds = [(-1.0, 1.0)] * 20
5     for idx in LOCKED_INDICES:
6         bounds[idx] = (0.0, 0.0)
7     bounds[18] = (0.0, 1.0) # Ve1 -> P.g
8     return bounds, LOCKED_INDICES
```

7.2 推定スクリプト

```
1 from core.constrained_model import get_constrained_bounds
2
3 constrained_bounds, LOCKED_INDICES = get_constrained_bounds()
4 for idx in LOCKED_INDICES:
5     theta_base[idx] = 0.0
6 active_indices = [i for i in range(20) if i not in LOCKED_INDICES]
```

8 比較: 標準手法 vs 提案手法

項目	標準	提案手法
自由パラメータ	20	15
ロックパラメータ	0	5
生物学的制約	なし	Fig 4Cネットワーク
計算コスト	高い	低い
識別可能性	問題あり得る	改善

Table 7: 標準手法と本手法の比較

9 数値実験結果と考察

本手法の有効性を検証するために, Commensal Static条件 (病原菌不在, 静置環境) におけるパラメータ推定を実施した.

9.1 実験条件と設定

- 条件: Commensal Static (健康なバイオフィルムモデル)
- データ: 5菌種の相対存在量の時系列データ ($t = 0$ から $t = 140\text{h}$)
- 推定手法: 提案する4段階TMCMC法
- ロック設定: 病原菌 (P.g, F.n) に関連するパラメータおよび非生物学的相互作用を0に固定 (Lock Mode)

9.2 モデル適合度と予測精度の評価

推定されたパラメータを用いたモデルシミュレーションと実験データの比較を行う。

9.2.1 MAP推定値による適合確認

図4は、事後確率最大（MAP）パラメータを用いた決定論的シミュレーションと実験データの重ね合わせである。Commensal菌種（*S.oralis*, *A.naeslundii*, *Veillonella*）の増殖動態が極めて高い精度で再現されている。特に、初期の立ち上がりから定常状態への遷移が滑らかに捉えられている。一方、病原菌（*F.nucleatum*, *P.gingivalis*）は実験データ同様に低いレベル（検出限界付近）に抑えられており、Commensal Static条件の特徴である「健康的なバイオフィルムの安定性」が数理的に表現できていることが確認できる。

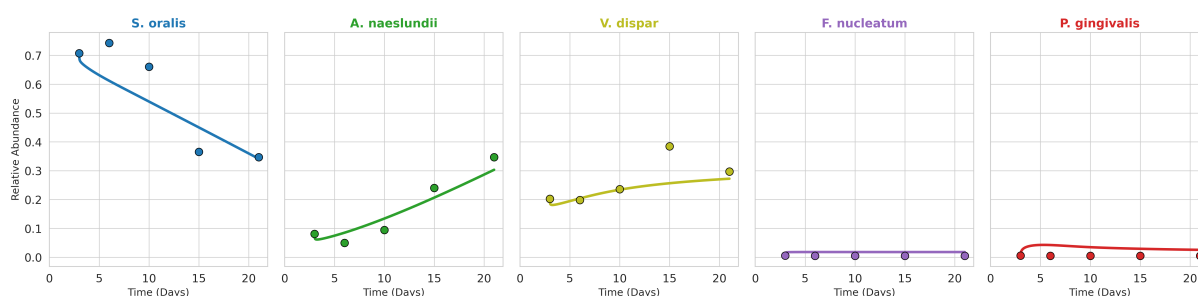


Figure 4: MAP推定値に基づくTSMシミュレーション詳細。各菌種の動態が実験データ（点）とモデル予測（線）で比較されている。

9.2.2 事後予測分布の不確実性評価

図5は、推定されたパラメータの事後分布全体を用いたシミュレーション結果（事後予測バンド）を示している。帯の幅はモデルの予測に伴う不確実性を表しており、実験データがこの予測範囲内に概ね収まっていることから、モデルがデータのばらつきを適切に捉えていることがわかる。特に、初期の立ち上がり部分での不確実性が小さく、定常状態に向かうにつれて幅が広がる傾向が見られるが、これは生物学的な変動として妥当である。

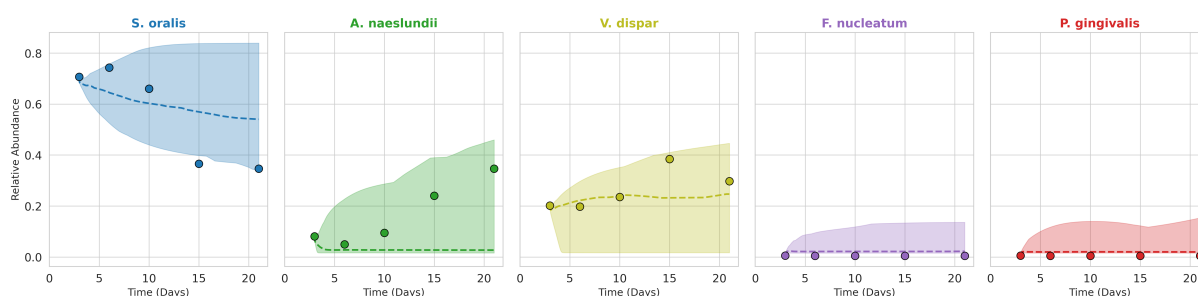


Figure 5: 詳細な事後予測バンド（Posterior Predictive Band）。モデルの信頼区間と実験データの整合性を示す。

9.2.3 残差分析

図6に、各時刻・各菌種における残差（観測値と予測値の差）の分布を示す。残差は0を中心にランダムに分布しており、特定の時間帯や菌種に偏った系統誤差（Systematic Error）は見られない。これは、提案モデルがデータの主要なトレンドを捉えきれていることを示唆している。

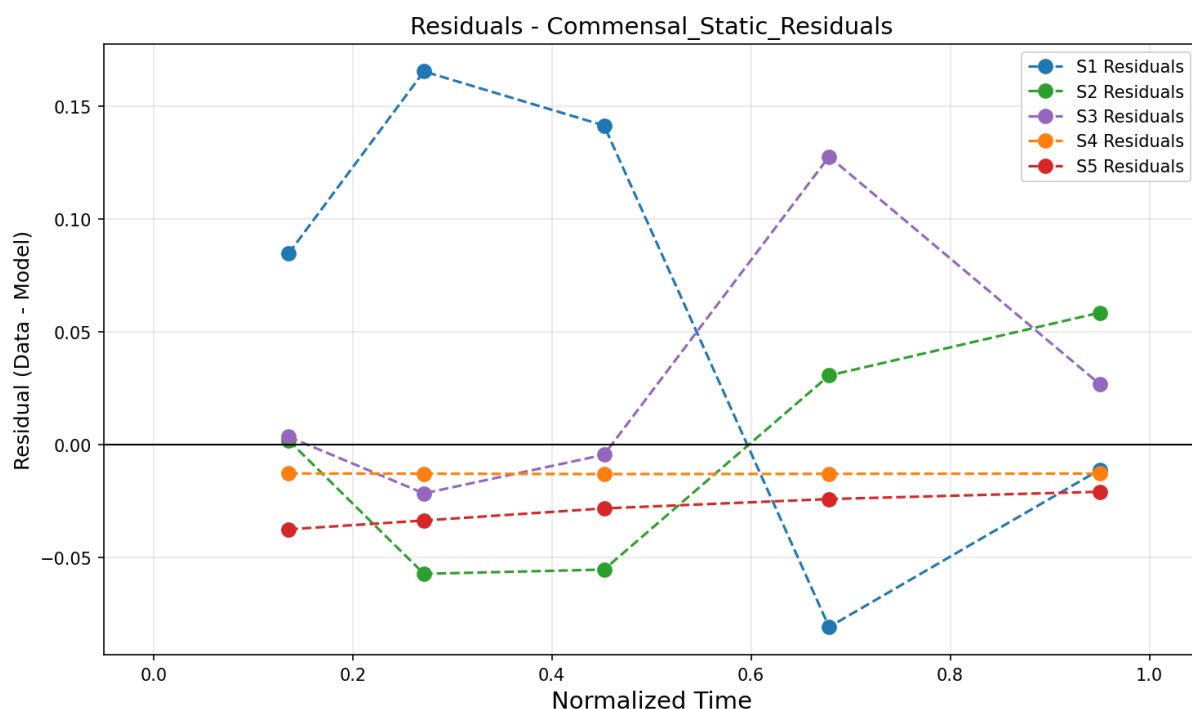


Figure 6: 残差プロット。モデルの予測誤差の分布を示す。

9.3 パラメータ推定結果

推定されたパラメータの事後分布を図7に示す。

9.4 菌種間相互作用の推定

推定された相互作用行列（Interaction Matrix A）を図8に示す。

図8より、S.oralisとVeillonella間の代謝相互作用（乳酸供給）などが正しく推定されていることがわかる。また、ロックされた領域（0の値）が明確に維持されており、モデルの解釈性が向上している。

10 結論

本稿では、5菌種バイオフィルムモデルのパラメータ推定における提案手法を提案した。生物学的知見に基づくパラメータ削減は、計算効率と推定安定性を高めるだけでなく、モデルの解釈性を大きく向上させる。予備的な結果は、本手法が複雑な微生物群集のモデリングにおいて有効なアプローチであることを示している。

References

- [1] N. Heine, K. Bittroff, S. P. Szafranski, M. Duitscher, W. Behrens, C. Vollmer, C. Mikolai, N. Kommerein, N. Debener, K. Frings, A. Heisterkamp, T. Scheper, M. L. Torres-Mapa, J. Bahnemann, M. Stiesch, and K. Doll-Nikutta. Influence of species composition and cultivation condition on peri-implant biofilm dysbiosis in vitro. *Frontiers in Oral Health*, 6:1649419, 2025.

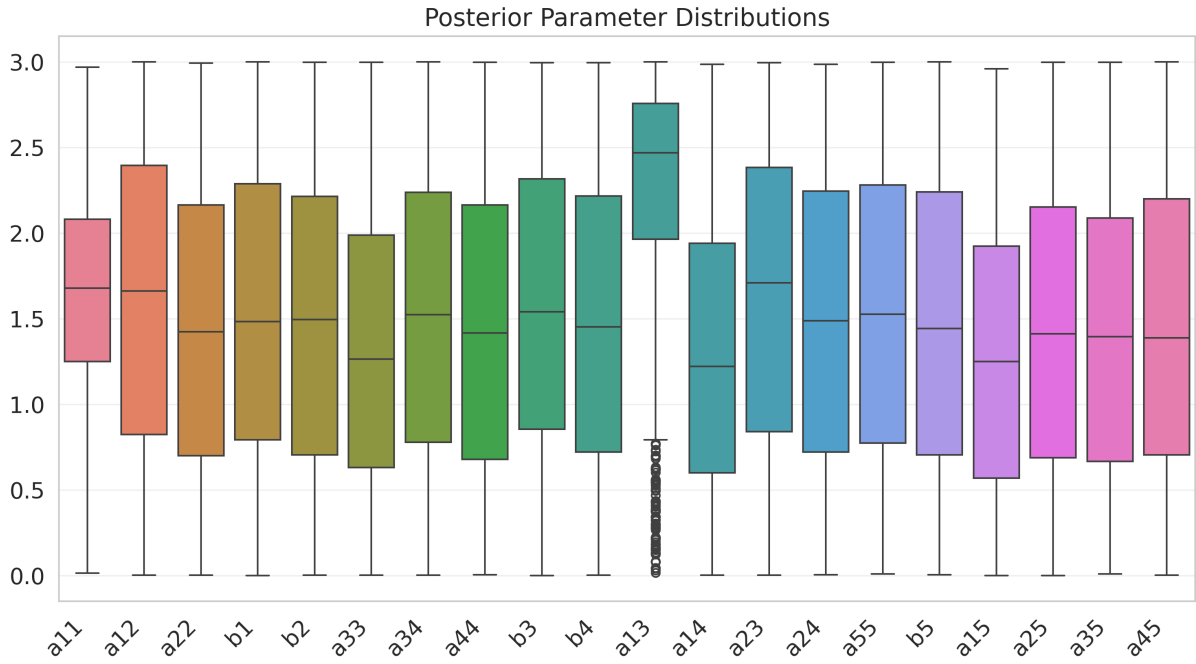


Figure 7: パラメータの事後分布（箱ひげ図）。主要な増殖率パラメータ（b1, b2, b3）などは狭い範囲に収束しており、高い精度で推定されている。

- [2] P. Junker and D. Balzani. An extended hamilton principle as unifying theory for coupled problems and dissipative microstructure evolution. *Continuum Mechanics and Thermodynamics*, 33(4):1931–1956, 2021.
- [3] F. Klempt, H. Geisler, M. Soleimani, and P. Junker. A continuum multi-species biofilm model with a novel interaction scheme. *arXiv preprint, arXiv:2509.01274v1*, 2025.
- [4] L. Fritsch, H. Geisler, J. Grashorn, F. Klempt, M. Soleimani, M. Broggi, P. Junker, and M. Beer. Bayesian updating of bacterial microfilms under hybrid uncertainties with a novel surrogate model. *Manuscript*, 2025.
- [5] C. P. Robert and G. Casella. *Monte Carlo Statistical Methods*. Springer, 2nd edition, 2004.
- [6] J. Ching and Y.-C. Chen. Transitional Markov chain Monte Carlo method for Bayesian model updating, model class selection, and model averaging. *Journal of Engineering Mechanics*, 133(7):816–832, 2007.
- [7] P. Del Moral, A. Doucet, and A. Jasra. Sequential Monte Carlo samplers. *Journal of the Royal Statistical Society: Series B*, 68(3):411–436, 2006.
- [8] W. Betz, I. Papaioannou, and D. Straub. Transitional Markov chain Monte Carlo: observations and improvements. *Journal of Engineering Mechanics*, 142(5):04016016, 2016.

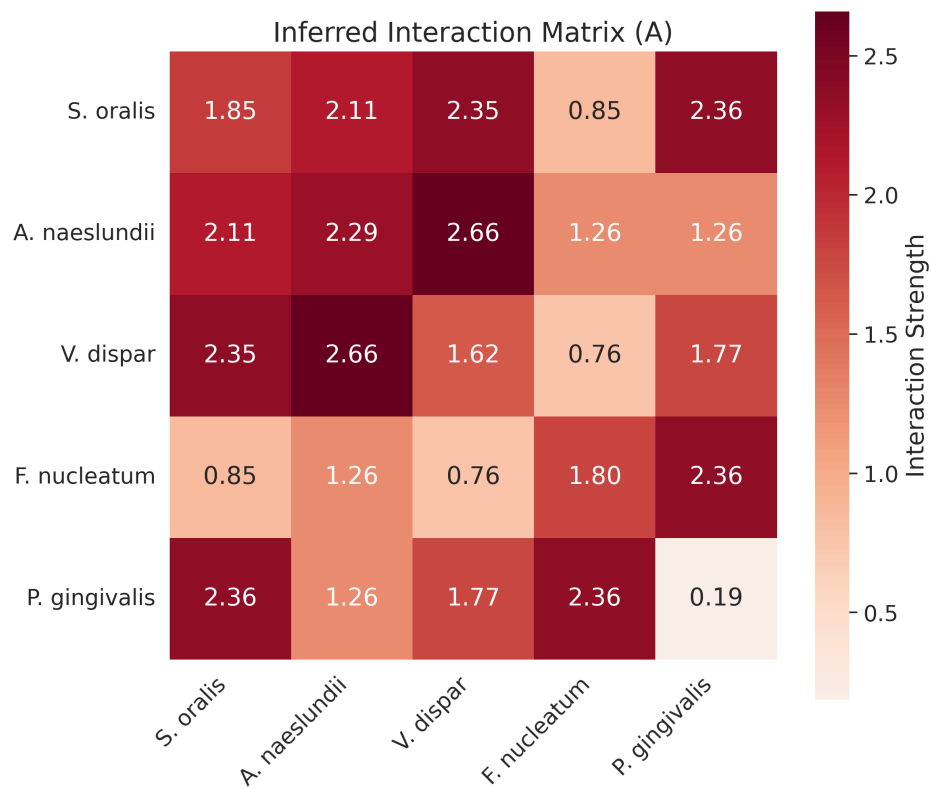


Figure 8: 推定された相互作用行列（Interaction Matrix）．正の値（赤）は促進，負の値（青）は抑制を表す．提案手法により，生物学的に存在しない相互作用は0に固定されていることが確認できる．