

公開

研究終了報告書

「微生物の新規代謝物—酵素遺伝子の統合オミクス推定法の開発」

研究期間：2020年12月～2023年3月

研究者：岡橋 伸幸

加速フェーズ期間：2023年3月～2024年3月

1. 研究のねらい

微生物は、新規生理活性物質や新しい代謝酵素を見出す探索源として有用な資源である。昨今の微生物培養技術の革新により、単離可能な微生物は増えつつある一方、単離した菌株をハイスループットに評価する必要性が生じている。微生物の評価には、網羅的なデータ取得が可能なオミクス解析の活用が期待されており、全遺伝子配列を解読するゲノム解析や代謝物を包括的に計測するメタボローム解析は最もよく用いられるアプローチである。しかし、オミクス解析は、機能未知遺伝子や構造未知代謝物などデータベースに収載のないデータのアノテーションが難しい。この問題を解決するには、従来法の既知だけに注目した解析法ではなく、未知情報を読み解くことを志向したデータ解読法を開発する必要がある。そこで、本研究では、未利用のまま山積している微生物資源の利活用加速を目指し、未知を含む代謝物データとゲノムデータを組み合わせ、特定の菌株のみに存在するユニーク代謝物とユニーク遺伝子を推定する *in silico* プラットフォームを開発する（図1）。このプラットフォームでは、ある基質に対して、様々な酵素反応を起こした際の産物構造を網羅的に計算することができる。これによって仮想の分子や反応機構も含めて、酵素反応（EC番号）—産物の組み合わせを *in silico* で生成することができる。これを、実際に取得した菌株のメタボロームデータやゲノムデータと照らし合わせ、特定の菌株にのみ存在する酵素遺伝子や産物を探索し、基質-酵素-生成物のペアの推定を目指す。

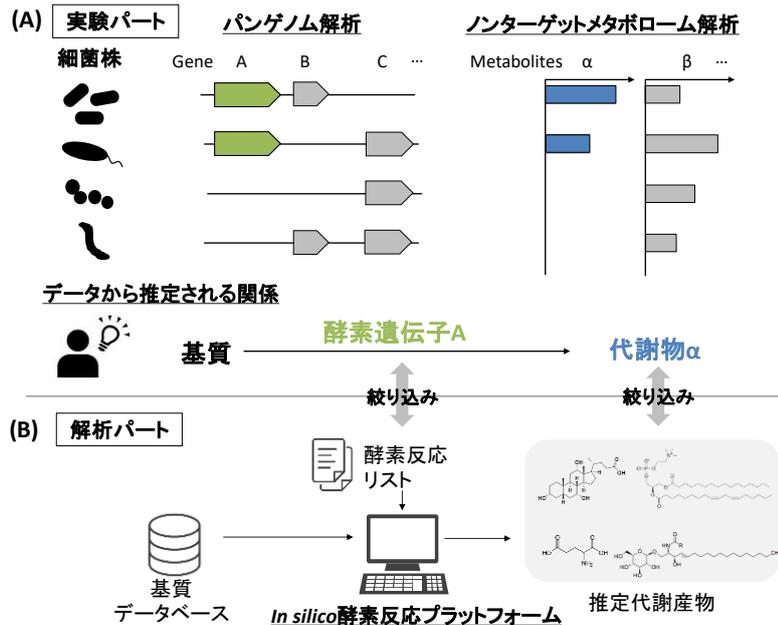


図1 本研究で開発する解析法の俯瞰図

加速フェーズでは、これまでにノンターゲットメタボロミクスを用いて見出した構造未知の代謝物の構造推定、および、基質-酵素遺伝子-代謝産物のペアを推定する解析法の精度向上を目指した。前者については、マススペクトルの類似性に基づくネットワークを作成することで、類縁体と推定される代謝物をあらかじめグループ化し、効率的な構造推定に取り組んだ。後者については、RNAseq 解析を行い、Wet 実験と Dry 解析の両面から解析精度の向上に取り組んだ。これまでの解析法では、注目する代謝物を産生する菌株のみが有する遺伝子に注目して生合成酵素遺伝子の推定を行ってきたが、スクリーニングを通過する候補遺伝子数が多く、期待したように候補を減らすことができていなかった。そこで、菌株の RNAseq 解析を行い、遺伝子発現量の観点から候補を絞り込むことで、本解析法の精度向上を試みた。

2. 研究成果

(1)概要

研究計画は 1.*in silico* 酵素反応プラットフォームの構築、2.微生物のメタボロームデータとゲノムデータの収集と解析の実施、3. 結果検証とプラットフォーム改善、の 3 点であり、おおよそ当初の計画通り遂行できた。

1.*in silico* 酵素反応プラットフォームの構築

本プログラムは、計算機上で基質構造と化学反応機構を与えると、産物となる構造をアウトプットする。基質には、公開データベースから 50 万種の化合物情報を、酵素反応にはデータベース BRENDA から基質特異性を緩めた EC 番号第 2~3 階層のレベルで 142 個を用いて、仮想の反応や産物も含めて約 1009 億通りの酵素反応-代謝物のパターンを生成できた。これと実際に 2.で取得したメタボロームデータやゲノムデータを比較し、代謝物の精密質量や遺伝子配列から予測される酵素の EC 番号が合致するものを絞り込みが可能となった。

2.微生物のメタボロームデータとゲノムデータの収集と解析の実施

実際の解析対象には、嫌気性菌株 45 株を選び、ゲノムデータとメタボロームデータを収集した。ゲノムデータは NCBI から取得し、パンゲノムを作成した。メタボロームデータは菌体から抽出した代謝物を液体クロマトグラフィー四重極飛行時間型質量分析装置で分析することで取得した。一部、従来法での解析が難しい分子については分析条件の最適化を行った。その結果、特定の菌株にしか存在しないユニーク代謝物を数百個オーダーで見出すことができた。加速フェーズでは、これらの分子の構造推定を行い、その一部は既報やデータベースに収載のない新規構造であることを明らかにした。

3. 結果検証とプラットフォーム改善

1.で開発した *in silico* 酵素反応プラットフォームを使って実際に特定の菌株にしか存在しない代謝物と産生酵素遺伝子の推定を行った。嫌気性菌 40 菌株のうち、Bacteroidetes 門細菌 14 種は、特徴的にスフィンゴ脂質を産生することが知られており、実際に取得したデータとも一致した。そこで、その生合成初発酵素 serine palmitoyl-CoA transferase と、その産物

3-ketosphinganine の関係をポジティブコントロールとし、1 で開発した *in silico* 酵素反応プラットフォームを用いてこのペアを言い当てることができるか検証した。MS/MS スペクトルの類似性に基づいたスコアリング機能を導入することで、最もスコアが高いペアがポジティブコントロールと合致として設定した反応となることを確認できた。

4. 解析精度の向上

加速フェーズでは、本解析手法の高精度化に向けて、RNAseq 解析を行い、絞り込まれる候補の数の削減に取り組んだ。選抜した 12 株の RNAseq 解析を行い、主要な分子の生合成酵素の発現量を調べたところ、主要な代謝酵素は FPKM 値が百から千オーダーであり、FPKM 値が 50 以下の遺伝子は候補から除外することとした。これによって、候補遺伝子を 30~50%程度減らすことができた。

(2) 詳細

「*in silico* 酵素反応プラットフォームの構築」

本プログラムは python 上で化学反応を扱うパッケージ RDKit を用いて開発した。基質には、Human Metabolome Database, PubChem, KNAPSAC など公開データベースから 50 万種の化合物情報を用いた。酵素反応リストの作成には、BRENDA データベースを参照した。特定の官能基や部分構造をもつ基質に対して、その構造の変換を起こす反応機構を記述した(図 2 は 1 級アルコールを酸化する反応 EC1.1 の例)。この際、反応部分以外の基質特異性をあえて無視することで、仮想の反応や産物も生成できるようにした。また DNA, RNA, タンパク質など今回のメタボローム解析の対象とならない分子に対する反応は除外した。最終的に、酵素反応リスト EC 番号第 2~3 階層のレベルで 142 個に集約された。実際に、仮想の反応や産物も含めて約 1009 億通りの酵素反応-代謝物のパターンを生成できるようになった。

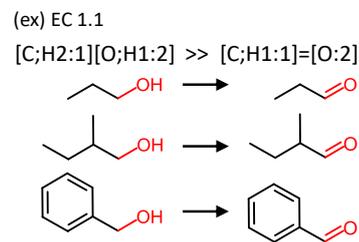


図 2 酵素反応の記載例

「菌株のデータ取得と解析の実施」

まず、メタボロームデータを取得するために、代表的なヒト腸内細菌からなる嫌気性菌 45 株を培養できる共通の培地を探索した。MRS, mGAM, BHI, YCFA などの培地を試し、BHI で 40 株を嫌気培養可能なことを見出した。回収した菌体の 1 層抽出物を液体クロマトグラフィー-四重極飛行時間型質量分析装置のデータ依存型 MS/MS 取得法で分析した。その結果、6177 個の代謝物ピークを検出することができた。それらを階層クラスタリングに供したところ、代謝物のプロファイルはおよそ菌株の分類ごとに特徴的な傾向を示した。よく知られたリン脂質などは *Escherichia coli* や *Klebsiella pneumoniae* などの Proteobacteria 門細菌や *Akkermansia* に多く見られた。*Alistipes* 属を除く Bacteroidetes 門細菌にはリン脂質に加えて、Sphingolipid が特徴的に見られ、これは Bacteroidetes 門細菌が原核生物では例外的に Sphingolipid を持つという既報の知見とも一致するものであった。Firmicutes 門に属するグラム陽性細菌では、Diacylglycerol に糖が 1 つか 2 つ結合した分子 Monoglycosyldiacylglycerol

および Diglycosyldiacylglycerol が見られ、これらの分子がグラム陽性細菌のペプチドグリカン脂質二重膜に繋ぎとめる役割を果たすという知見とも合致した。一方、既存のデータベースを使って同定できた分子は全体の 15.6%にとどまり、未同定のままとなる代謝物が多くを占めた。それらの中には、特定の菌株にのみ特徴的に存在する分子も多数見られ、これらの構造と産生酵素を *in silico* 解析プラットフォームを用いて解析するのに好適なデータが得られたことが確かめられた。

一方、一部の微生物に特徴的な分子を従来法ではとらえることができていないことが明らかとなった。そのような分子の一つでグラム陰性細菌の起炎症性分子 Lipid A はカラムへの吸着が問題となっていることが分かった。そこで溶離液の pH を検討し、アルカリ性溶媒と耐アルカリ性ポリマーベースカラムを用いることで、LC-MS での分離定量が可能となった(図 3AB)。これをグラム陰性腸内細菌株の Lipid A 分析に適用したところ、Lipid A のアシル鎖の長さや本数に多様性があることが明らかとなった(成果論文 1)。

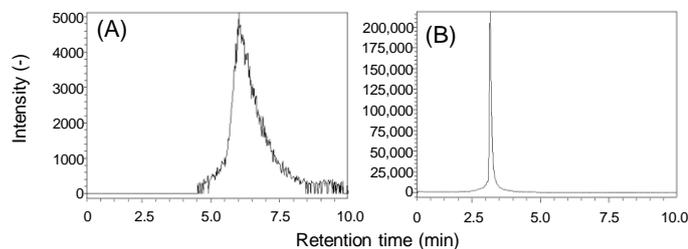


図 3 Lipid A の分析法の開発
(A) 検討前と(B)検討後の LipidA のピーク形状

加速フェーズでは、図 3 で見出した未同定分子の構造を効率的に推定するために、MS/MS スペクトルの類似性に基づくネットワークを作成する技術 Feature-based molecular networking を用いた。類縁体と推定される代謝物をあらかじめグルーピングしたところ、既存のデータベースからは同定できない未知イオンからなるクラスターを見出し、それらの構造を推定することができた。

「結果検証とプラットフォームの改善」

2.で取得したデータを使って、1 で構築した *in silico* プラットフォームの妥当性を検証した。Bacteroidetes 門細菌が産生することが既知の Sphingolipid の初発生合成酵素遺伝子 serine palmitoyl-CoA transferase とその産物 3-ketosphinganine をポジティブコントロールとした(図 4)。当初の試みでは、精密質量から産物の分子式は $C_{18}H_{37}NO_2$ であることまで絞り込むことに成功したものの、同一質量数の構造異性体とその産生反応が 300 個近く選抜されてしまうという問題が生じた。そこで、実測された代謝物と予測された産物の構造類似性を比較する絞り込みを実装した。CFM-ID (Wang et al., *Anal. Chem.*, 2021) を用いて予測された産物の構造式から MS/MS スペクトルを生成し、実測データと比較したところ、最も類似性が高い候補が、答えとなる酵素反応となることを確かめた。

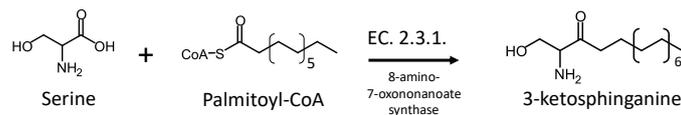


図 4 ポジティブコントロールとした反応

「RNA データの利用による解析精度向上」

加速フェーズでは、解析精度を向上させるために、菌株の RNAseq データを取得し、実際に候補となる酵素遺伝子の発現量を解析に加味することを目指した。絞り込みに重要となる菌株を 18 株選抜し、うち 12 株からクオリティーの良い RNA を回収することができた。それらを RNAseq 解析に供し、遺伝子発現量のデータを得た(図 5)。今回質量分析装置で比較的高強度に計測された代謝物の生合成酵素遺伝子の発現量に注目すると、FPKM 値は数百から数千オーダーと非常に大きいことが明らかとなり、FPKM 値が 50 以下の遺伝子は観測される代謝物の産生に関わる可能性は低いとして候補から除外した。これによって期待通り 30~50% 候補遺伝子を減らすことができた。しかし、適切な閾値は菌株ごとに異なることが考えられ、一意的な基準を置くこと難しいことも明らかとなった。

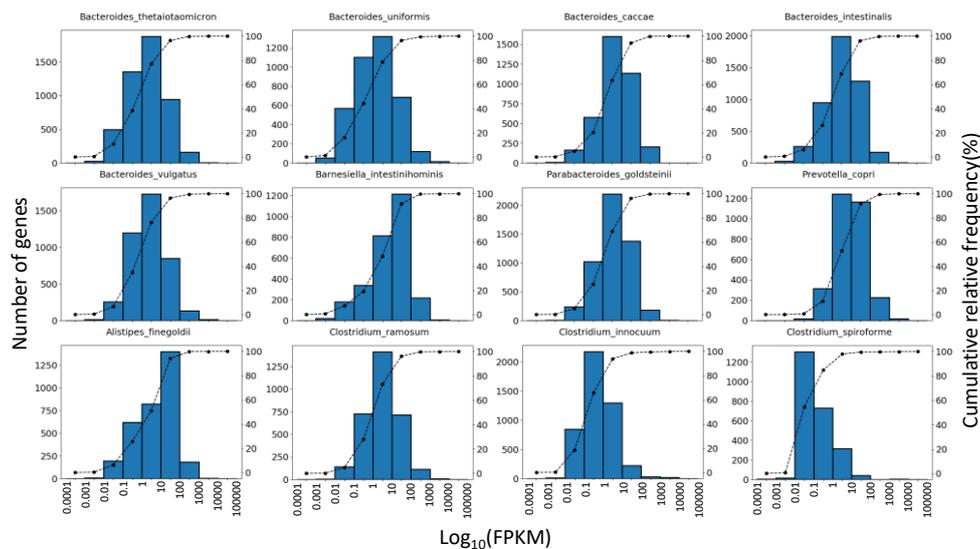


図 5 腸内細菌株の RNAseq データ

その他

- ACT-X の成果を 74 回生物工学会大会で発表したところ、話題性のある研究としてトピックス集に採択され、学会からプレスリリースされた。
(https://www.sbj.or.jp/2022/wp-content/uploads/2021/10/2022_SBJ_topics.pdf).
- 成果論文1の微生物のノンターゲットメタボローム解析手法は、質量分析メーカーである島津製作所からの依頼を受けてアプリケーションノートを執筆した。HP で公開されており、技術の社会還元も進めている。

(https://www.an.shimadzu.co.jp/sites/an.shimadzu.co.jp/files/ckeditor/aplnotes/ap_aplnote78-jp.pdf)

- 本プロジェクトで微生物の計測に最適化したノンターゲットメタボローム解析を用いて、研究室学生との共同研究で出芽酵母の脂質変化を計測した論文(原著論文 2、責任著者)や、2 期生の東大岩間先生との共同研究で、糸状菌のライフサイクルと脂質の変化を包括的に計測した論文(原著論文 3、プレスリリース https://resou.osaka-u.ac.jp/ja/research/2023/20230831_2)が受理され、計測対象を原核微生物から真核微生物にも拡大することができた。

3. 今後の展開

これまでの研究では、比較的扱いやすい単離菌株のデータを用いて技術開発を行い、本研究のコンセプトが機能することを実証できた。今後は、複合微生物系での解析を実現できるよう本技術を発展させたいと考えている。環境中の微生物のうち、培養できるものは 1%に満たないと言われており、培養の可否が微生物資源の利活用の根本的な制約になっている。単離を経ずに代謝産物とその酵素遺伝子を推定できるようなアプローチが開発できれば、この根本的な制約を解除し、微生物資源の利活用におけるパラダイムを根本から変えることができると期待される。

これと同時に、既に培養可能な菌株については、ゲノムとメタボロームデータの集積を進める。土壌菌、水圏菌、極限環境菌、動植物共生菌、真核微生物などのデータを取得し、データベースを整備、公開する。これによって、他の研究者が扱う菌株をデータベースと照合できるようになり、これまでに開発してきた技術や情報を社会に還元することが可能となる。

(加速フェーズ実施後追記)

加速フェーズでは、実際に既報にない新規構造の脂質を多数見出すことに成功したことから、ノンターゲットメタボロミクスを用いて微生物の産生する未知化合物を集積していく重要性を改めて確認することができた。一方、本研究を未利用のまま山積している多くの微生物に拡大していくには、培養、分析、クローニングなど wet 実験面に多くの律速点が存在している。今後、上述のような研究を大規模かつハイスループットに展開するためには、ロボティクスや実験オートメーション技術の活用が重要となるであろう。

4. 自己評価

研究目的の達成状況

本研究の目的は、ゲノムデータとメタボロームデータを統合的に解析する *in silico* 酵素反応プラットフォームを開発し、特定の菌株に特徴的な代謝物とその産生酵素遺伝子を見出す方法論を確立することであった。本目的を達成するための 3 つのマイルストーンは上述の通りおおむね達成できた。

研究の進め方(研究実施体制及び研究費執行状況)

本プロジェクトでは、菌株の培養とメタボロームデータの収集が律速となることが予想されたため、自動分注機の導入と技術員の雇用による解決を試みた。プロジェクト開始後すぐには技術員が見つからず、雇用開始は半年ほど遅れたものの、最終的には目的としたデータをすべて取得できた



ため、今回の研究実施体制は有効であった。

研究成果の科学技術及び社会・経済への波及効果

本研究で開発した微生物のノンターゲットメタボローム解析手法は、質量分析メーカーである島津製作所からの依頼を受けて執筆したアプリケーションノートとして HP で公開されており、質量分析装置のアプリケーション作成、開発した技術の社会還元につながっている。また本手法は、微生物を扱う他の研究者とも親和性が高く、ここで確立したオリジナルの方法論を別の分野にも展開することで、微生物研究、代謝研究の発展に広く貢献できると期待される。この技術を使って有用な生理活性物質や、新規反応を触媒する酵素遺伝子が獲得されることで、それらを活用した発酵生産や世界に先駆けた特許取得など、ビジネス面での利点も期待される。

(加速フェーズ実施後追記)

加速フェーズでは、ノンターゲットメタボローム解析のデータから新規構造化合物を多数見出すことができた。また RNAseq データを用いて本解析法を高度化することができたため、おおよそ目標を達成でき、今後の研究の礎を築くことができたと考えている。研究費は、菌株の培養や質量分析装置のオペレーション、RNA サンプルの回収を行う技術員の雇用に充て、研究を加速できた。一方、今回見出した新規構造化合物が非常に多数であったため、生化学的な検証ができた候補は一部にとどまった。今後、本研究を発展させるためには、細菌の遺伝子操作や、実験のオートメーション、新規化合物の有機合成など、複数の専門家からなるグループを組織し、大きなビジョンで研究を遂行する必要がある。

5. 主な研究成果リスト

(1) 代表的な論文(原著論文)発表

研究期間累積件数: 1件

研究期間累積件数: 3件(加速フェーズ後更新)

1. **Nobuyuki Okahashi***, Masahiro Ueda, Fumio Matsuda, Makoto Arita*. Analyses of Lipid A Diversity in Gram-Negative Intestinal Bacteria Using Liquid Chromatography-Quadrupole Time-of-Flight Mass Spectrometry, *Metabolites*, 2021, 11, 197, 1-15.

本論文では、これまで申請者が哺乳類細胞を対象に構築してきたノンターゲット解析法を腸内細菌に適用する多めに、分析条件を最適化し、計測法の高度化を実現できた。グラム陰性細菌がもつ分子 Lipid A はこれまでのノンターゲット解析系では分析できないことが分かったため、LC の溶離液条件を最適化することで克服した。本法を種々の腸内細菌株に適用することで、腸内細菌種ごとに Lipid A の構造に多様性があることが明らかとなった。

2. Shuka Komori, Nobuyuki Okahashi*, Junko Iida, Fumio Matsuda. Lipidome variation of industrial *Saccharomyces cerevisiae* strains analyzed by LC-QTOF/MS-based untargeted lipidomics, *Journal of Bioscience and Bioengineering*, 2023, 135, 2, 102-108.

本論文では、本プロジェクトで微生物の分析に最適化したノンターゲット解析法を用いて



様々な発酵食品酵母株の脂質多様性を明らかにした。低温や高糖などの至適条件から外れたストレス環境に対応するには、外界と細胞内を隔てる膜脂質が重要である。そこで、パン、ワイン、日本酒などの製造に利用される酵母の脂質を包括的に解析し、様々なストレス環境下での生育速度との相関解析を実施することで、ストレス適応に関係する脂質を絞り込むことができた。※加速フェーズ実施の成果

3. Ryo Iwama, Nobuyuki Okahashi, Tetsuki Suzawa, Chuner Yang, Fumio Matsuda, Hiroyuki Horiuchi*. Comprehensive analysis of the composition of the major phospholipids during the asexual life cycle of the filamentous fungus *Aspergillus nidulans*, *Biochimica et Biophysica Acta – Molecular and Cell Biology of Lipids*, 2023, 1868, 10, 159379.

本論文は、当領域 2 期生の岩間亮先生(東京大学)との共同研究であり、生活環の中で糸状菌 *Aspergillus nidulans* が大きく形態変化することに着目し、細胞膜を形成するリン脂質の量および質の変化を包括的にとらえた。特に、発芽のタイミングでリン脂質中の不飽和度が増加するという特徴的な変化が起こることを明らかにした。さらに、ノンターゲットリポミクスを利用して、中性脂質やマイナーなリン脂質にも含めた包括的な計測を行うことで、同様の変化が広い脂質クラスにわたって起きていることを示した。※加速フェーズ実施の成果

(2) 特許出願

該当なし

研究期間累積件数: 1件(加速フェーズ後更新)

(3) その他の成果(主要な学会発表、受賞、著作物、プレスリリース等)

- **岡橋伸幸** 上田政宏 松田史生 有田誠, LC-QTOF/MS を用いたグラム陰性腸内細菌の Lipid A の解析、第 69 回質量分析総合討論会、2021/05/19
- **岡橋伸幸** 松田史生 ゲノムデータとメタボロームデータを用いた酵素遺伝子と代謝産物の推定法の開発、第 74 回日本生物工学会大会、2022/10/20 トピックス集に採択され、学会からプレスリリース https://www.sbj.or.jp/2022/news/news_topics_20221017-1.html
- **岡橋伸幸** 松田史生 統合オミクス解析を用いた微生物の酵素遺伝子と代謝産物の推定法の開発、第 45 回日本分子生物学会年会、2022/12/2
- **岡橋伸幸** 微生物資源の活用に向けたノンターゲットメタボロミクス、酵素工学ニュース
- **岡橋伸幸** 松田史生 有田誠 質量分析法を使った未知代謝物の計測と同定 生物工学会誌 doi: 10.34565/seibutsukogaku.99.5_233
- **Nobuyuki Okahashi**, Untargeted analysis of microbial metabolites using mass spectrometry, YABEC2023 (Invited), 2023/7/13 ※加速フェーズ実施の成果
- **Nobuyuki Okahashi**, Comprehensive analysis and data integration for identification of unique bacterial metabolites and enzyme genes, 3rd International BMS Symposium 2023 in Kyoto (Invited), 2023/10/7 ※加速フェーズ実施の成果
- **岡橋伸幸** 微生物代謝物の網羅計測とゲノムデータとの統合、日本微生物生態学会第 36 回浜松大会(招待講演)、2023/12/6 ※加速フェーズ実施の成果

- **岡橋伸幸** 培地成分・組織で特定のアミノ酸を精確に定量する 決定版質量分析活用スタンダード、羊土社 ※加速フェーズ実施の成果
- **岡橋伸幸** 糸状菌(カビ)の包括的な生体膜変化の解明(大阪大学プレスリリース https://resou.osaka-u.ac.jp/ja/research/2023/20230831_2) 「環境とバイオテクノロジー」領域 2 期生の岩間亮先生(東京大学)との共同研究 ※加速フェーズ実施の成果