

2024 年度 研究開発実施報告

概要

研究開発課題名	マイクロバイオーム研究を先導するハブを目指した微生物統合データベースの特化型開発
開発対象データベースの名称(URL)	Microbiome Datahub (https://mdatahub.org/)
研究代表者氏名	森 宙史 (40610837)
所属・役職	情報・システム研究機構 国立遺伝学研究所 情報研究系 准教授 (2025 年 3 月時点)



□目次

概要	1
§1. 研究実施体制	4
§2. 研究開発対象とするデータベース・ツール等	5
(1) データベース一覧	5
(2) ツール等一覧	5
§3. 実施内容	6
(1) 本年度に計画されていた研究開発項目・タスク	6
(2) 進捗状況	6
【研究開発実施項目 1-1】MAG データを用いたゲノム・メタゲノムの統合化	6
【研究開発実施項目 3-2】MAG データに関する国際連携	7
【研究開発実施項目 1-2】微生物 Phenotype 情報の RDF 化	7
【研究開発実施項目 4】微生物 Phenotype 情報のデータ構築	7
【研究開発実施項目 1-3】マイクロバイオー姆解析パイプラインの大幅な速度向上	7
【研究開発実施項目 2】MAG に対する機能オーソロググループアノテーション機能の開発	7
【研究開発実施項目 1-4】効率的な逐次 DB 更新・運用システムの"再"構築	8
【研究開発実施項目 3-1】ヒトマイクロバイオー姆のメタデータ整備とキラータセットの開発	8
【研究開発実施項目 1-5】Microbiome Datahub のデータを用いたキラアプリケーションの開発	8
【研究開発実施項目 1-6】植物やメタボローム・DDBJ 等の他の DB との連携	9
§4. 成果発表等	10
(1) 原著論文発表	10
① 論文数概要	10
② 論文詳細情報	10
(2) その他の著作物(総説、書籍など)	10
(3) 国際学会および国内学会発表	10
① 概要	10
② 招待講演	11
③ 口頭講演	11
④ ポスター発表	11
(4) 知的財産権の出願 (国内の出願件数のみ公開)	12
① 出願件数	12
② 一覧	12
(5) 受賞・報道等	12
① 受賞	12
② メディア報道	12
③ その他の成果発表	12
§5. 主要なデータベースの利活用状況	13
1. アクセス数	13
① 実績	13

② 分析	13
2. データベースの利用状況を示すアクセス数以外の指標	13
3. データベースの利活用により得られた研究成果(生命科学研究への波及効果).....	13
4. データベースの利活用によりもたらされた産業への波及効果や科学技術のイノベーション(産業や科学技術への波及効果)	13
§6. 研究開発期間中に主催した活動(ワークショップ等)	14
(1) 進捗ミーティング	14
(2) 主催したワークショップ、シンポジウム、アウトリーチ活動等	15

§1. 研究実施体制

グループ名	研究代表者・研究分担者氏名	所属機関・役職名	研究題目
遺伝研グループ	森 宙史	情報・システム研究機構 国立遺伝学研究所	Microbiome Datahubのマイクロバイオーーム研究の国際的なデータハブへの発展に向けた研究開発と運用
基生研グループ	内山 郁夫	基礎生物学研究所	微生物オーソログデータの構築と、そのゲノム機能推定への応用
東京科学大グループ	山田 拓司	東京科学大学	ヒトマイクロバイオーームのメタデータ整備とキラーデータセットの開発
京大グループ	松井 求	京都大学	微生物Phenotype情報のデータ構築とRDF化

§2. 研究開発対象とするデータベース・ツール等

(1) データベース一覧

【主なデータベース】

No.	名称	別称・略称	URL
1	Microbiome Datahub		https://mdatahub.org/

【その他のデータベース】

No.	名称	別称・略称	URL
1	Microbial Genome Database for Comparative Analysis	MBGD	https://mbgd.nibb.ac.jp/

(2) ツール等一覧

No.	名称	別称・略称	URL
1	Metagenome and Microbes Environmental Ontology	MEO	https://bioportal.bioontology.org/ontologies/MEO
2	Latent Environment Allocation	LEA	http://leamicrobe.jp
3	PZLAST		https://pzlast.nig.ac.jp/
4	PZLAST-MAG		https://pzlast.nig.ac.jp/pzlast/mag
5	EMBERS		http://palaeo.nig.ac.jp/embers_app/
6	Bac2Feature		2025 年 6 月公開予定

§3. 実施内容

(1) 本年度に計画されていた研究開発項目・タスク

【研究開発実施項目 1-1】MAG データを用いたゲノム・メタゲノムの統合化

【研究開発実施項目 3-2】MAG データに関する国際連携

- ・ MAG データの冗長性の軽減
- ・ Microbiome Datahub の論文化

【研究開発実施項目 1-2】微生物 Phenotype 情報の RDF 化

【研究開発実施項目 4】微生物 Phenotype 情報のデータ構築

- ・ JCM と NBRC からの更新版の菌株データの取得
- ・ Bac2Feature を単離菌ゲノムに用いた結果の Microbiome Datahub での表示
- ・ MAG に対する Bac2Feature の実行

【研究開発実施項目 1-3】マイクロバイオーーム解析パイプラインの大幅な速度向上

【研究開発実施項目 2】MAG に対する機能オーソロググループアノテーション機能の開発

- ・ マイクロバイオーーム配列データの遺伝子機能組成解析パイプラインの高速化を行い大量サンプルで実行した結果の Microbiome Datahub への格納
- ・ 大量のゲノムや MAG データに対する高速なオーソログアサインメントパイプラインの開発および実行

【研究開発実施項目 1-4】効率的な逐次 DB 更新・運用システムの"再"構築

【研究開発実施項目 3-1】ヒトマイクロバイオーームのメタデータ整備とキラーデータセットの開発

- ・ LLM を用いた論文からのメタデータの抽出と、データセット 論文作成
- ・ LLM を用いたメタデータ抽出の生息環境情報等への拡張

【研究開発実施項目 1-5】Microbiome Datahub のデータを用いたキラーアプリケーションの開発

- ・ PZLAST MAG 等のアプリケーションの公開と論文化
- ・ アドバイザリー委員会の開催によるユーザの幅広い意見の収集

【研究開発実施項目 1-6】植物やメタボローム・DDBJ 等の他の DB との連携

- ・ プロテオームグループ等の他の統合化推進プログラムの DB との連携の推進
- ・ MAG データに対する DDBJ との連携強化

(2) 進捗状況

【研究開発実施項目 1-1】MAG データを用いたゲノム・メタゲノムの統合化

INSDC 由来の約 21 万件の MAG データと RefSeq 由来の約 2.6 万件のリファレンス単離菌ゲノムおよび関連する BioProject 情報を Microbiome Datahub に収録し、Elasticsearch によって系統や MAG のクオリティ、宿主の系統等で検索できるようにし、配列データやアノテーション情報等を Web ページおよび API 経由でダウンロードできるようにした。また、各 MAG についてアサインされた MBGD オーソログ情報や予測された Phenotype 情報等を収録して公開した。EMBL-EBI で開発し公開している MAG データベースである MGnify から MAG 配列データ約 31 万個を取得し、Microbiome Datahub 用のアノテーションを行いメタデータの JSON ファイルも記述形式を定義し、データ変換を行なった。2025 年度早期に Microbiome Datahub に MGnify のデータも収録予定である。また、INSDC の MAG と MGnify の MAG 合計 52 万 MAG について skani で高速にゲノムの配列類似性を計算し、配列類似性の閾値にもよるが、3-40%ほどの配列のオーバーラップがあることを確認した。それらの解析結果を含めた Microbiome

Datahub の論文を現在執筆中である。

【研究開発実施項目 3-2】MAG データに関する国際連携

ChatGPT API を用いた主にヒトメタゲノム論文からのメタデータ抽出ツールである EMBERS 論文を bioRxiv で 2024 年 10 月公開後[論文 1]、メタゲノムの系統組成推定ツールで有名な SingleM とそのツールを数十万サンプルのメタゲノムデータに適用した結果の系統組成の DB である Sandpiper を作っているオーストラリアの Queensland 大学のグループから共同研究の打診があり、何度か web 会議を遺伝研グループおよび東京科学大グループの共同で行なった。30 万サンプル以上の 16S rRNA 遺伝子のアンプリコン解析の結果および約 15 万サンプルのメタゲノム解析の結果の系統組成情報と、こちらで論文から抽出したサンプルごとのメタデータの対応付けの作業を現在共同で進めている。

【研究開発実施項目 1-2】微生物 Phenotype 情報の RDF 化

菌株保存機関の NBRC から Phenotype 情報付きの菌株データについて、2024 年末に新たに提供いただいた。それらの菌株データを、既に関済済みの RDF 形式への変換プログラムを用いて変換し、NBRC 側の菌株データベースである DBRP と共通の RDF フォーマットを使いつつ NBRC と連携して RDF 化を進めた。

【研究開発実施項目 4】微生物 Phenotype 情報のデータ構築

微生物の Phenotype 推定ツール Bac2Feature を用いて、Microbiome Datahub に存在する約 21 万 MAG それぞれの、細胞の形や至適 pH 等 27 種類の Phenotype を系統情報を元に推定し、Microbiome Datahub 上で閲覧可能にした。単離菌の Phenotype 推定結果については、ユーザが実際の菌株の Phenotype データと混同しないように、記述形式を変えてファイルを作成し、Bac2Feature の論文が公開され次第 Microbiome Datahub 上で閲覧可能にする。Bac2Feature については、現在論文投稿し 2025 年 5 月 30 日に accept された[論文 5]。論文が公開され次第ツールを公開し、Microbiome Datahub との間で相互リンク等行い、Microbiome Datahub のユーザ数増加を促進する予定である。

【研究開発実施項目 1-3】マイクロバイオーーム解析パイプラインの大幅な速度向上

MBGD の 34,079 ゲノム由来の 1,433,670 オースログクラスターを参照配列データベースとして、Microbiome Datahub の 21 万 MAG 由来の 4.5 億タンパク質の配列類似性検索を、遺伝研グループで開発した PZLAST-MAG をカスタマイズして行い、各 MAG が所持するホモログ情報を集計し、MBGD Ortholog ID リストと KEGG Orthology ID リストの両方を MAG ごとに Microbiome Datahub 上で閲覧およびダウンロード可能にした。

【研究開発実施項目 2】MAG に対する機能オースロググループアノテーション機能の開発

昨年度に引き続き、MBGD の更新作業をすすめた。昨年度開発した KEGG Orthology を参照した MBGD オースロググループのサブグループ化手続きを、MBGD オースロググループ全体に適用し、これに基づいて登録ゲノム全体に対して KEGG Orthology クロスリファレンスを改善し、これを用いて Genomapple による KEGG Module の充足率計算を行った。この MBGD-Genomapple パイプラインを MyMBGD インターフェースに組み込み、ユーザが登録したゲノムに対して適用する処理を完成させた。また、MBGD に含まれないドラフトゲノムのうち、MBGD に存在しない genus に属するゲノムに対してオースロググループを拡張した draft-plus バージョンも作成した。これらの変更に加え、昨年度までに開発を進めてきた系統プロファイル検索機能や、API を通じて MBGD のデータを取得できる機能についても改良を加えて完成させ、MBGD の新バージョンとして公開した。登録ゲノム数は 34079 (draft-plus のゲノム数は 36156)、オースロググループ数は 1,433,670 となった[論文 4]。

このオーソログデータを大量の MAG に高速にアサインするため、オーソログ分類された配列データをクエリとして、カスタマイズした PZLAST-MAG により、遺伝研グループが持つ約 4000 コアのメニーコアプロセッサ搭載の計算機を用いて、配列類似性検索を実施した。これにより、MBGD Ortholog および KEGG Orthology アノテーションを数億タンパク質に対して約 1 ヶ月で完了できる高速化を実現した。一方で、アサインできなかった約 3 割のタンパク質については、プロファイル検索等の感度の高い手法による検索を別途行う予定である。

【研究開発実施項目 1-4】効率的な逐次 DB 更新・運用システムの"再"構築

非常にニーズが多いヒトマイクロバイオームデータについて、優先度の高い論文からメタデータのマニュアルキュレーションを行い、ヒトマイクロバイオーム関連の約 4500 論文から約 41000 サンプルのメタデータを抽出し、マイクロバイオーム配列データの個別の BioSample ID と紐付けた。それらのマニュアルキュレーション済みのヒトマイクロバイオーム BioSample メタデータを正解データとして、Python 言語と ChatGPT API を用いた論文からのメタデータ抽出ツール EMBERS を開発し、2024 年 10 月に bioRxiv で論文を公開すると共に GitHub でツールを公開した[論文 1]。EMBERS は PubMed Central から取得した論文データからのメタデータ抽出と (EMBERS-MINE)、メタデータの名寄せと MEO や HMADO へのオントロジー自動アノテーション (EMBERS-FUSE) の二つのツールに分かれており、EMBERS を用いることで DB 更新の際に非常に手間がかかっていたメタデータ関連の作業を効率化できた。現在 MAG データに紐づく BioSample ID と MAG が由来した元々のリードデータの BioSample ID の対応付けを行っており、対応付け作業が完了し次第 EMBERS の結果も Microbiome Datahub に収録して検索可能にする。

【研究開発実施項目 3-1】ヒトマイクロバイオームのメタデータ整備とキラデータセットの開発

1-4 で言及した EMBERS を活用したメタデータ抽出結果を元に、ヒトマイクロバイオーム論文から抽出し整理したメタデータと BioSample との対応関係を整理し、BioSample をクラスタリングすることでヒトマイクロバイオームのメタ解析に使用できるサンプルメタデータ俯瞰ツールを開発した。それらの結果も bioRxiv で公開した EMBERS 論文に含めてあり[論文 2]、公開されている大量のヒトマイクロバイオームデータについてメタ解析を行なった論文として投稿論文も執筆中である。EMBERS の技術をヒトマイクロバイオーム以外の土壌や海洋等の環境メタゲノム論文にも適用するために、環境メタゲノム用にプロンプトや Python プログラム等を改良した EMBERS プログラムのテスト版を開発した。ヒトマイクロバイオームに比べて環境が多様であり各環境に対応する論文数が少ないため、元々の EMBERS ほどのメタデータ抽出精度は現状実現できていないが、ある程度の効率化は実現できた。

【研究開発実施項目 1-5】Microbiome Datahub のデータを用いたキラアプリケーションの開発

Microbiome Datahub のユーザ数向上を目的として、Microbiome Datahub のデータをユーザが自由に取得できるように、データ取得 API を開発し公開した。現状メタゲノム BioProject のメタデータや系統組成データ、MAG のメタデータ、MAG の配列データ、MAG が持つ MBGD Ortholog ID リストおよび KEGG Orthology ID リストが取得可能である。大量データのダウンロードには別途ダウンロードサイトを用意する予定であるが、個々のデータのダウンロードは API および Microbiome Datahub のサイトからマウスクリックのみでデータダウンロード API を呼び出せる機能を開発し、プログラム経由で API アクセスできない利用者にも API を利用できるようにしている。配列の検索については、マイクロバイオームの配列データに対する検索ツールである PZLAST について、約 21 万 MAG 由来の遺伝子のアミノ酸配列データに対する配列類似性検索を行える機能 (PZLAST-MAG) を開発し、公開した (<https://pzlast.nig.ac.jp/pzlast/mag>)。現在検索したタンパク質と類似したタンパク質の系統分布を表示する機能等を追加しつつ、論文執筆中である。ユーザの幅広い意見を収集するため、マイクロバイオーム

領域の専門家によるアドバイザリー委員会を、いくつかの微生物系国内学会のタイミングで併設して 2024 年度 3 度開催し、データベースの改善すべき点や、今後の機能追加や取り扱うべきデータ等についてアドバイスをいただいた。

【研究開発実施項目 1-6】植物やメタボローム・DDBJ 等の他の DB との連携

MAG 配列データの遺伝子機能アノテーション等について、DDBJ で開発している DFAST と連携して遺伝研スーパーコンピュータ上で高速に MAG のアノテーションをできるようにし、ゲノム・MAG のクオリティチェックを行えるツール、DFAST_QC を論文化し公開した[論文 3]。また、DBCLS とは NITE NBRC や bitBiome 社等と共に一昨年度から NEDO グリーンイノベーション基金事業のプロジェクトを遺伝研グループと共同で進めており、その一環としてゲノムや MAG・SAG から化学合成独立栄養細菌が持つ 7 種類の CO₂ 固定経路の有無を高精度に推定するツール AutoFixMark を開発し、論文投稿準備中である。DBCLS が開発しているゲノム DB TogoGenome や培地 DB TogoMedium と Microbiome Datahub の微生物ゲノムや MAG データ間で API を主とした連携を進めると共に、LLM を用いたメタデータの自動アノテーションについても DBCLS と毎月の Togothon で定期的に情報交換を行い、技術的な連携を行なっている。

§4. 成果発表等

(1) 原著論文発表

① 論文数概要

種別	国内外	件数
発行済論文	国内(和文)	0 件
	国際(欧文)	3 件
未発行論文 (accepted, in press 等)	国内(和文)	0 件
	国際(欧文)	2 件

② 論文詳細情報

1. Hirotsugu Shiroma, Youssef Darzi, Etsuko Terajima, Zenichi Nakagawa, Hirotaka Tsukikura, Naoki Tsukuda, Yuki Moriya, Shujiro Okuda, Susumu Goto, Takuji Yamada. Enteropathway: the metabolic pathway database for the human gut microbiota. Brief Bioinform (2024) 25: 5 (10.1093/bib/bbae419)
2. Koichi Higashi, Zenichi Nakagawa, Takuji Yamada, Hiroshi Mori. Automated Harmonization and Large-Scale Integration of Heterogeneous Biomedical Sample Metadata Using Large Language Models. bioRxiv (2024) (10.1101/2024.10.26.620145)
3. Mohamed Elmanzalawi, Takatomo Fujisawa, Hiroshi Mori, Yasukazu Nakamura, Yasuhiro Tanizawa. DFAST_QC: Quality Assessment and Taxonomic Identification Tool for Prokaryotic Genomes. BMC Bioinformatics (2025) 26 (10.1186/s12859-024-06030-y)
4. Ikuo Uchiyama, Motohiro Mihara, Hiroyo Nishide, Hirokazu Chiba, Masahiko Takayanagi, Mikihiro Kawai, Hideto Takami. MBGD: Microbial Genome Database for Comparative Analysis Featuring Enhanced Functionality to Characterize Gene and Genome Functions Through Large-scale Orthology Analysis. J Mol Biol. (2025) 437: 15 (10.1016/j.jmb.2025.168957)

(2) その他の著作物(総説、書籍など)

1. 東光一, 黒川真臣, 森宙史, メタゲノム解析をはじめの前に, 誰でも再現できる NGS「前」サンプル調整プロトコール, 実験医学別冊, 羊土社, pp61-68, 2024 年.
2. 森宙史, メタゲノム解析, 実験デザインからわかる マルチオミクス研究実践テキスト, 実験医学別冊, 羊土社, pp211-223, 2025 年.

(3) 国際学会および国内学会発表

① 概要

種別	国内外	件数
招待講演	国内	3 件
	国際	0 件
口頭発表	国内	6 件
	国際	1 件
ポスター発表	国内	5 件
	国際	6 件

② 招待講演

〈国内〉

1. 松井求, 「バイオものづくり」をめぐる数理的課題, 第3回 MfIP 連携探索ワークショップ, 武蔵野大学, 2025 年 3 月 14 日.
2. 松井求, 適切な輪作体系と施肥は微生物叢を頑強にする, 2025 年度日本農芸化学会札幌大会, 札幌, 2025 年 3 月 7 日.
3. 松井求, Bioinformatics-Driven Microbiology, 東京農工大学バイオインフォマティクスセミナー, 東京農工大学, 2024 年 11 月 12 日.

〈国際〉

なし.

③ 口頭講演

〈国内〉

1. 森宙史, 細菌の多様性を解明するための大規模 MAG 解析, 第 97 回日本細菌学会総会, 札幌コンベンションセンター, 札幌, 2024 年 8 月 7 日.
2. 森宙史, 大規模メタゲノムデータから俯瞰するバクテリアの複製・転写・翻訳装置, 第 96 回日本遺伝学会大会, 高知工科大学, 高知, 2024 年 9 月 5 日.
3. 森宙史, MAG データベースからのプラスミド関連遺伝子の探索, 第 2 回プラスミド研究会, 国立遺伝学研究所, 三島, 2024 年 12 月 25 日.
4. 森宙史, 公共 DB 中の MAG を整理した DB Microbiome Datahub の開発, 第 1 回学際分野としての微生物研究会「あつまれ!びせいぶつの森」, 京都大学宇治キャンパス, 京都, 2025 年 2 月 28 日.
5. 内山郁夫, 三原基広, 西出浩世, 千葉啓和, 高柳正彦, 河合幹彦, 高見英人, MBGD 2024: 微生物オーソログデータベースの遺伝子・ゲノム機能アノテーションに向けた拡張, 第 19 回日本ゲノム微生物学会年会, 木更津, 2025 年 3 月 17-19 日.
6. 藤吉真生, 鈴木誉保, 岩崎渉, 古澤力, 松井求, 細菌の 16S rRNA 遺伝子配列から形質を予測するインターフェース Bac2Feature, 第 19 回日本ゲノム微生物学会年会, 木更津, 2025 年 3 月 17 日-19 日.

〈国際〉

1. Ikuo Uchiyama, Motohiro Mihara, Hiroyo Nishide, Hirokazu Chiba, Masahiko Takayanagi, Mikihiro Kawai, Hideto Takami, Improvement of the Microbial Genome Database (MBGD) toward genome/metagenome annotation, Quest for Orthologs 8, Montreal, July 17-18, 2024.

④ ポスター発表

〈国内〉

1. 松井求, アライメントの品と質: 系統樹推定に及ぼす影響の再検討, 日本進化学会第 26 回神奈川大会, 神奈川, 2024 年 8 月 21 日.
2. 森宙史, 藤澤貴智, 東光一, 谷澤靖洋, 中川善一, 西出浩世, 中村保一, 山田拓司, 松井求, 内山郁夫, マイクロバイーム研究を先導するハブとなる Microbiome Datahub の開発, トーゴの日シンポジウム 2024, 東京, 2024 年 10 月 5 日.
3. 内山郁夫, 三原基広, 西出浩世, 千葉啓和, 高柳正彦, 河合幹彦, 高見英人, MBGD 2024: ゲノム・メタゲノム機能解析に向けた微生物比較ゲノムデータベースの改良, トーゴの日シンポジウム 2024, 東京, 2024 年 10 月 5 日.
4. 土岐誠司, 松井求, ゲノム組成の温度適応の時間的遅れが好冷細菌を予測困難にする, 第 37 回日本

微生物生態学会広島大会，広島，2024 年 10 月 29 日。

5. Zhenbo Jiang, Ikuo Uchiyama, Integration of co-expression relationship across multiple datasets based on extended Pan- & Core-network approaches, 第 19 回日本ゲノム微生物学会年会，木更津，2025 年 3 月 17 日-19 日。

〈国際〉

1. Ikuo Uchiyama, Motohiro Mihara, Hiroyo Nishide, Hirokazu Chiba, Masahiko Takayanagi, Mikihiro Kawai, Improvement of the Microbial Genome Database (MBGD) toward genome/metagenome annotation, 32nd Conference on Intelligent Systems for Molecular Biology (ISMB 2004), Montreal, July 12-16, 2024.
2. Ikuo Uchiyama, Motohiro Mihara, Hiroyo Nishide, Hirokazu Chiba, Masahiko Takayanagi, Mikihiro Kawai, Improvement of the Microbial Genome Database (MBGD) toward genome/metagenome annotation, 1st Asia & Pacific Bioinformatics Joint Conference, Naha, October 22-25, 2024.
3. Mikihiro Kawai, Motohiro Mihara, Hiroyo Nishide, Ikuo Uchiyama, Exploration of links among phylogenetic profiles of orthologous gene groups in Microbial Genome Database for Comparative Analysis database, 1st Asia & Pacific Bioinformatics Joint Conference, Naha, October 22-25, 2024.
4. Zenbo Jiang, Ikuo Uchiyama, Extraction of the shared co-expression patterns based on extended pan-network and core-network concepts, 1st Asia & Pacific Bioinformatics Joint Conference, Naha, October 22-25, 2024.
5. Motomu Matsui, PANJEP: Pairwise alignment-based neighbor joining with reliability measurements of interior branches, 1st Asia & Pacific Bioinformatics Joint Conference, Naha, October 22-25, 2024.
6. Koki Kamada, and Motomu Matsui, Machine learning for predicting tree inference accuracies, 1st Asia & Pacific Bioinformatics Joint Conference, Naha, October 22-25, 2024.

(4) 知的財産権の出願（国内の出願件数のみ公開）

① 出願件数

種別		件数
特許出願	国内	0 件

② 一覧

1) 国内出願

なし。

(5) 受賞・報道等

① 受賞

なし。

② メディア報道

なし。

③ その他の成果発表

なし。

§5. 主要なデータベースの利活用状況

1. アクセス数

① 実績

表 1 研究開発対象の主要なデータベースの利用状況

名称	種別	2024 年度(月間平均値)
Microbiome Datahub (2024 年 12 月に正式公開)	訪問者数	1,268
	訪問数	2,181
	ページ数	8,467

② 分析

Microbiome Datahub への移行のために、2024 年 4 月に MicrobeDB.jp を閉鎖し、それ以降は Microbiome Datahub のテスト版のみ公開している状態であったが、2024 年 12 月に MAG ごとの MBGD Ortholog 組成や Bac2Feature による Phenotype 推定結果等も収録した Microbiome Datahub version 1 を正式公開した。2025 年に入ってから微生物系の若手の会等の研究会やゲノム微生物学会等で Microbiome Datahub の宣伝を行い、正式公開以降データベースの利用者数は着実に伸びてきている。しかしながら国際的な認知度がまだ低いため、Microbiome Datahub の論文出版や国際学会での宣伝等が今後急ぎで必要である。

2. データベースの利用状況を示すアクセス数以外の指標

特に無し

3. データベースの利活用により得られた研究成果(生命科学研究への波及効果)

特に無し

4. データベースの利活用によりもたらされた産業への波及効果や科学技術のイノベーション(産業や科学技術への波及効果)

特に無し

§6. 研究開発期間中に主催した活動(ワークショップ等)

(1) 進捗ミーティング

年月日	名称	場所	参加人数	目的・概要
2024 年 4 月 26 日	チーム内開発会議(非公開)	オンライン	7 人	プロジェクト内の開発会議
2024 年 5 月 24 日	チーム内開発会議(非公開)	オンライン	7 人	プロジェクト内の開発会議
2024 年 6 月 12 日	チーム内開発会議(非公開)	オンライン	4 人	プロジェクト内の開発会議
2024 年 6 月 19 日	チーム内開発会議(非公開)	オンライン	4 人	プロジェクト内の開発会議
2024 年 6 月 27 日	Microbiome Datahub 利 活用推進会議(非公開)	国内版 Bio Hackathon BH24.6 会場(群馬 県)	4 人	外部の専門家による Microbiom e Datahub の改善すべき点や取 り組むべき課題等のアドバイスを 得るための会議
2024 年 7 月 24 日	チーム内開発会議(非公開)	オンライン	4 人	プロジェクト内の開発会議
2024 年 8 月 8 日	Microbiome Datahub 利 活用推進会議(非公開)	日本細菌学 会総会会場 (北海道)	3 人	外部の専門家による Microbiom e Datahub の改善すべき点や取 り組むべき課題等のアドバイスを 得るための会議
2024 年 8 月 28 日	チーム内開発会議(非公開)	オンライン	4 人	プロジェクト内の開発会議
2024 年 9 月 27 日	チーム内開発会議(非公開)	オンライン	5 人	プロジェクト内の開発会議
2024 年 10 月 12 日	チーム内開発会議(非公開)	オンライン	6 人	プロジェクト内の開発会議
2024 年 10 月 21 日	チーム内開発会議(非公開)	オンライン	4 人	プロジェクト内の開発会議
2024 年 10 月 29 日	チーム内開発会議(非公開)	オンライン	4 人	プロジェクト内の開発会議
2024 年 10 月 31 日	チーム内開発会議(非公開)	オンライン	5 人	プロジェクト内の開発会議
2024 年 11 月 6 日	チーム内開発会議(非公開)	オンライン	4 人	プロジェクト内の開発会議
2024 年 11 月 8 日	チーム内開発会議(非公開)	オンライン	4 人	プロジェクト内の開発会議
2024 年 11 月 13 日	チーム内開発会議(非公開)	オンライン	4 人	プロジェクト内の開発会議
2024 年 11 月 14 日	チーム内開発会議(非公開)	オンライン	6 人	プロジェクト内の開発会議
2024 年 11 月 19 日	チーム内開発会議(非公開)	オンライン	3 人	プロジェクト内の開発会議
2024 年 11 月 26 日	チーム内開発会議(非公開)	オンライン	4 人	プロジェクト内の開発会議

年月日	名称	場所	参加人数	目的・概要
2024 年 12 月 4 日	チーム内開発会議(非公開)	オンライン	4 人	プロジェクト内の開発会議
2024 年 12 月 5 日	チーム内開発会議(非公開)	遺伝研	3 人	プロジェクト内の開発会議
2024 年 12 月 11 日	チーム内開発会議(非公開)	遺伝研	5 人	プロジェクト内の開発会議
2024 年 12 月 13 日	オーストラリアのグループと のデータベースについての 共同研究会議(非公開)	オンライン	4 人	SingleM や Sandpiper 等のツールやデータベースを開発しているオーストラリアのグループとの連携のための会議
2024 年 12 月 24 日	チーム内開発会議(非公開)	オンライン	4 人	プロジェクト内の開発会議
2025 年 1 月 7 日	チーム内開発会議(非公開)	オンライン	4 人	プロジェクト内の開発会議
2025 年 1 月 29 日	チーム内開発会議(非公開)	オンライン	5 人	プロジェクト内の開発会議
2025 年 2 月 26 日	チーム内開発会議(非公開)	オンライン	5 人	プロジェクト内の開発会議
2025 年 2 月 28 日	チーム内開発会議(非公開)	京都大学	3 人	プロジェクト内の開発会議
2025 年 3 月 12 日	チーム内開発会議(非公開)	オンライン	5 人	プロジェクト内の開発会議
2025 年 3 月 18 日	チーム内開発会議(非公開)	日本ゲノム 微生物学会 年会会場 (千葉)	5 人	外部の専門家による Microbiome Datahub の改善すべき点や取り組むべき課題等のアドバイスを 得るための会議
2025 年 3 月 28 日	チーム内開発会議(非公開)	オンライン	5 人	プロジェクト内の開発会議

(2) 主催したワークショップ、シンポジウム、アウトリーチ活動等

年月日	名称	場所	参加人数	目的・概要
2024 年 12 月 16-18 日	チーム内 Hackathon(非公開)	三井ガーデンホテル柏の葉	3 人	プロジェクト内の DB 開発合宿
2025 年 1 月 15-17 日	チーム内 Hackathon(非公開)	ホテルアソシア静岡	5 人	プロジェクト内の DB 開発合宿
2025 年 2 月 10-13 日	チーム内 Hackathon(非公開)	ホテルアソシア静岡	5 人	プロジェクト内の DB 開発合宿

以上

別紙1 既公開のデータベース・ウェブツール等

No.	正式名称	別称・略称	概要	URL	公開日	状態	分類	関連論文
1	MicrobeDB.jp		ゲノムを軸にして微生物学の様々なデータをセマンティックウェブ技術を用いて統合した微生物統合データベース。	https://microbedb.jp	2011/12/12	休止・閉鎖	データベース等	
2	Microbiome Datahub		メタゲノム由来のゲノムであるMAGを中心に微生物のゲノムとメタゲノムデータを統合したマイクロバイオームの統合データベース。	https://mdatahub.org/	2023/5/31 (2024年12月24日version 1を正式公開)	維持・発展	データベース等	
3	Microbial Genome Database for Comparative Analysis	MBGD	オーソログ解析に基づいて微生物ゲノムの比較解析を行うためのデータベース。公開されたゲノム全体を含む標準オーソログテーブルに基づいて、各オーソロググループの系統プロファイルの比較などを行えるほか、動的なオーソログ解析機能によって、利用者自身が持つゲノム配列も含めて、興味のある生物種セットに対象を絞った比較を行うことも可能。	https://mbgd.nibb.ac.jp	1997/7/1	維持・発展	データベース等	Ikuo Uchiyama, Motohiro Mihara, Hiroyo Nishide, Hirokazu Chiba, Masahiro Takayanagi, Mikihiro Kawai, Hideto Takami. "MBGD: Microbial Genome Database for Comparative Analysis Featuring Enhanced Functionality to Characterize Gene and Genome Functions Through Large-scale Orthology Analysis", J Mol Biol. 168957, 2025. (DOI: 10.1016/j.jmb.2025.168957).
4	Metagenome and Microbes Environmental Ontology	MEO	微生物の生息環境に関するメタデータを記述し整理するためのオントロジー。	https://bioportal.bioontology.org/ontologies/MEO	2012/2/21	維持・発展	ツール等	
5	Latent Environment Allocation	LEA	Microbiome Datahubに存在する大量のマイクロバイオームサンプルの系統組成と環境を対応させて視覚化するWebアプリケーション。ユーザが自分の群集組成データをアップロードしてLEA上にマッピングすることで、系統組成から環境を大まかに推定することも可能。	http://leamicrobe.jp	2018/6/6	維持・発展	ツール等	Koichi Higashi, Shinya Suzuki, Susumu Kurosawa, Hiroshi Mori, Ken Kurokawa, "Latent environment allocation of microbial community data", PLoS Comput Biol, 14, e1006143, 2018 (DOI:10.1371/journal.pcbi.1006143).
6	PZLAST		Microbiome Datahubに存在する大量のメタゲノムサンプルのアミノ酸配列データに対して高速な配列相同性検索を行うことが可能なWebアプリケーション。	https://pzlast.nig.ac.jp	2020/4/1	維持・発展	ツール等	Hiroshi Mori, Hitoshi Ishikawa, Koichi Higashi, Yoshiaki Kato, Toshikazu Ebisuzaki, Ken Kurokawa, "PZLAST: an ultra-fast amino acid sequence similarity search server against public metagenomes", Bioinformatics., 37, 3944-3946, 2021. (DOI: 10.1093/bioinformatics/btab492).
7	PZLAST-MAG		Microbiome Datahubに存在する大量のMAGのアミノ酸配列データに対して高速な配列相同性検索を行うことが可能なWebアプリケーション。	https://pzlast.nig.ac.jp/pzlast/mag	2024/12/24	新規	ツール等	

8	EMBERS		ChatGPT APIとPython言語を用いて主にヒトマイクロバイオームの論文からサンプルのメタデータを抽出するツール。EMBERSで抽出したメタデータを整理して現在のヒトマイクロバイオーム研究を俯瞰できるwebサーバも用意した。	http://palaeo.nig.ac.jp/embers_app/	2024/10/29	新規	ツール等	Koichi Higashi, Zenichi Nakagawa, Takuji Yamada, Hiroshi Mori, "Automated Harmonization and Large-Scale Integration of Heterogeneous Biomedical Sample Metadata Using Large Language Models", bioRxiv, 2024. (DOI: 10.1101/2024.10.26.620145).
9	Bac2Feature		既知の微生物のPhenotype情報と系統樹情報を基に、16S rRNA遺伝子配列や系統名から微生物のPhenotypeを予測するツール。		2025/06/	新規	ツール等	Masaki Fujiyoshi, Takao K Suzuki, Wataru Iwasaki, Chikara Furusawa, Motomu Matsui, "Bac2Feature: an easy-to-use interface to predict prokaryotic traits from 16S rRNA gene sequences", Bioinformatics Advances (accepted).