

2022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
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



目次

§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) 本年度の研究開発計画と達成目標	6
① MAG データを用いたゲノム・メタゲノムの統合化、MAG データに関する国際連携	6
② 微生物 Phenotype 情報の RDF 化、微生物 Phenotype 情報のデータ構築	6
③ マイクロバイオーム解析パイプラインの大幅な速度向上、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	6
④ 効率的な逐次 DB 更新・運用システムの"再"構築、ヒトマイクロバイオームのメタデータ整備とキラードータセットの開発	7
⑤ Microbiome Datahub のデータを用いたキラードアプリケーションの開発	7
⑥ 植物やメタボローム・DDBJ 等の他の DB との連携	7
(2) 進捗状況	7
① MAG データを用いたゲノム・メタゲノムの統合化、MAG データに関する国際連携	7
② 微生物 Phenotype 情報の RDF 化、微生物 Phenotype 情報のデータ構築	8
③ マイクロバイオーム解析パイプラインの大幅な速度向上、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	8
④ 効率的な逐次 DB 更新・運用システムの"再"構築、ヒトマイクロバイオームのメタデータ整備とキラードータセットの開発	9
⑤ Microbiome Datahub のデータを用いたキラードアプリケーションの開発	9
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	10
③ 口頭講演	10
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
(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. 研究開発期間中に主催した活動(ワークショップ等)	12
(1) 進捗ミーティング	12
(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	14

§1. 研究実施体制

グループ名	研究代表者または主たる共同研究者氏名	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
東工大グループ	山田 拓司	東京工業大学	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
東大グループ	松井 求	東京大学	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.riken.jp/meta

§3. 実施内容

(1) 本年度の研究開発計画と達成目標

新たなデータの統合化の範囲をマイクロバイームに特化して拡大すると共に検索機能と解析機能の追加および洗練化を行うことで、微生物分野のエンサイクロペディア MicrobeDB.jp からマイクロバイーム研究の国際的なデータハブ Microbiome Datahub (仮)へと発展させる。具体的には、①ゲノム・メタゲノムの統合化、②リソース(モノ)との連結、③データ爆発への対応、④効率的な逐次 DB 更新・運用システムの"再"構築、⑤キラーデータセット・アプリケーションの開発、⑥マルチオミックス連携の 6 項目を中心に研究開発を実施する。

① MAG データを用いたゲノム・メタゲノムの統合化、MAG データに関する国際連携

メタゲノム由来の Metagenome Assembled Genome (MAG)について、GTDB や INSDC DDBJ/ENA/NCBI、IMG/M などの主要なメタゲノムの DB から配列およびメタデータを取得し、DB 間のメタデータの形式の違いや各 DB の MAG データの総数等を把握した上で、最適な MAG データ収集方法を設計すると共に、MAG データの再利用に必要なメタデータ項目の選定を行う。また、高精度な MAG のみを選択的に抽出するための配列データの品質評価手法を Nicola Segata 博士のグループと相談しつつ確立する。

② 微生物 Phenotype 情報の RDF 化、微生物 Phenotype 情報のデータ構築

既に継続的に菌株データを提供いただいている菌株保存機関の JCM と NBRC から Phenotype 情報付きの菌株データを今年度新たに提供いただく。提供いただいた菌株データ中に存在する Phenotype データについて、項目と値をリスト化して把握した上で、本 DB で用いる微生物の Phenotype オントロジーについて、既存のオントロジーを調査し有用な Phenotype オントロジーがあればアノテーションを試行する。また、BacDive 等の他の菌株保存機関の Phenotype データも利用し、株レベルだけではなく種や属などのより上位の系統分類群についても、その系統に共通した Phenotype の特徴を自動でアノテーション可能なツールである Bac2Feature について、本プロジェクトでも円滑に使用可能なように入力データや出力データ等の形式を調整する。

③ マイクロバイーム解析パイプラインの大幅な速度向上、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

爆発的な勢いで増加を続けるマイクロバイームデータについて、系統組成と遺伝子機能組成解析パイプラインの大幅な改良を行い、少なくとも系統組成についてはある程度の精度を保ったまま、大幅な高速化を行う。また、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構築に向けて、現状のオーソログデータ構築パイプラインに基づく MBGD のデータ更新を進めつつ、特にボトルネックとなっている DomRefine の改良に着手する。その際、機能推定に有効な分類を効率的に構築することを優先して、既存プログラムの仕様を大きく修正し、試作版を開発する。

④ 効率的な逐次 DB 更新・運用システムの"再"構築、ヒトマイクロバイオームのメタデータ整備とキラードータセットの開発

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、論文からのメタデータのマニュアルキュレーションの効率化のための手法の再検討を行った上で、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う。MEO 等のオントロジー自動アノテーションツール MDBAnnotator については、処理の高速化や効率化を目的にプログラムの改良を行う。また、DB の国際化のために MEO や HMADO 等の独自に開発しているオントロジーについては、標準として世界的に使われている EnvO や病気のオントロジーとのクラスの対応付けを行い、数年後にそれらのオントロジーへの移行を視野に入れて準備を進める。また、日々増加するマイクロバイオームデータについて最新のデータを DB に反映させるため、現存の MicrobeDB.jp の逐次 DB 更新・運用システムにおける課題の洗い出しとそれを改善するためのより効率的な逐次 DB 更新・運用システムの設計を行う。

⑤ Microbiome Datahub のデータを用いたキラード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、LEA や VITCOMIC2、PZLAST 等の、Microbiome Datahub のマイクロバイオームデータを用いた既存のツールの機能強化を行う。具体的には、各ツールのリファレンス DB の更新や解析機能の見直し等を行う。また、利用者視点に立った DB 開発を進めるために、Microbiome Datahub について幅広い層のユーザを含むアドバイザリーボードを設置し、今年度は 1 回 Microbiome Datahub の開発方針や強化すべき機能、DB のユーザインターフェース等についてのアドバイザリー委員会を開催し、DB についての幅広いユーザの意見を取り入れて開発を進める。

⑥ 植物やメタボローム・DDBJ 等の他の DB との連携

マイクロバイオームの配列データと同時に測定されることが増えつつあるメタボロームデータやヒトや植物等の宿主のオミックスデータについて、現状マイクロバイオームとセットで取得され公開されたデータがどの程度あるかを INSDC BioProject DB 等の研究プロジェクト DB のデータを調査し、優先的に取り組むべきマルチオミックスデータカテゴリのリスト化を行うと共に、データ連携の方策を立案する。

これらの研究開発を、遺伝研 G を中核とし、基生研 G、東工大 G、東大 G の計 4 グループで分担して実施する。

(2) 進捗状況

① MAG データを用いたゲノム・メタゲノムの統合化、MAG データに関する国際連携

INSDC から約 21 万個の MAG 配列データおよび対応する BioProject のメタデータを収集・整理した。BioProject のメタデータについては JSON 形式で記述を統一すると共に MEO や NCBI Taxonomy などのオントロジーアノテーションを行い、Microbiome Datahub で Elasticsearch によって環境情報等で BioProject を絞り込み検索できるようにした。MAG 配列データについては、DFAST および DFAST QC を用いて MAG ごとに遺伝子アノテ

ーションとクオリティチェックを行った結果を JSON 形式で記述し、Microbiome Datahub で各 MAG の情報を検索・閲覧できるようにした。単離菌のゲノムデータについては、NCBI RefSeq バクテリアゲノムデータと BioProject のメタデータを取得し、MAG と同様に DFAST を実行し遺伝子アノテーションを行い、メタデータとゲノムの情報を Microbiome Datahub で検索・閲覧できるようにした。

MAG データに関する国際連携については、MAG を含む細菌ゲノムデータから系統を定義し MAG の系統推定のリファレンス DB として国際的に広く使われている GTDB について、Scientific Advisory Board の委員に遺伝研グループの中村博士がなり、GTDB と連携できる基盤を整えた。

② 微生物 Phenotype 情報の RDF 化、微生物 Phenotype 情報のデータ構築

菌株保存機関の JCM と NBRC から Phenotype 情報付きの菌株データの RDF ファイルを 2022 年度に新たに RDF ポータル経由で提供いただいた。提供いただいた菌株データ中に存在する Phenotype データについて、項目と値をリスト化し、OMP や MPO 等の既存の微生物のフェノタイプオントロジーでのカバー率を調べた。また、株レベルだけではなく種や属などのより上位の系統分類群についても、その系統に共通した Phenotype の特徴を自動でアノテーション可能なツールである Bac2Feature について、本プロジェクトでも円滑に使用可能なように入力データや出力データ等の形式を調整した。同時に、Phenotype ごとに、進化系統におけるばらつきや保存性、推移パターンを詳細に調査し、Phenotype 自動アノテーションの精度を見積もることも可能とした。これまでに、MAG から 16S rRNA 遺伝子配列を抽出することを想定したアノテーションパイプラインを完成させている。次年度以降、MAG から他の遺伝子情報や k-mer 組成情報を抽出し、その情報に基づいて Phenotype を推定できるようにパイプラインをさらに更新していく予定である。

③ マイクロバイーム解析パイプラインの大幅な速度向上、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

爆発的な勢いで増加を続けるマイクロバイームデータについて、NCBI が用いている k-mer 組成を用いた高速な系統組成推定パイプラインを参考に、系統組成解析パイプラインの大幅な改良を行い、大幅な高速化を行った。また、大量のゲノムや MAG データに対する遺伝子の機能アノテーションについては、NCBI RefSeq が 2022 年度中旬から単離菌のゲノム中の各遺伝子に Gene Ontology (GO) ID を付与するようになったため、RefSeq の GO データを参照して DFAST の結果から GO アノテーションを可能にするプログラムの試作版を作成した。MBGD におけるユーザゲノム解析機能である MyMBGD について、オーソログ対応付けを高速化したバージョンを完成させて公開した。この対応づけに基づき、GenoMaple 解析を行なって KEGG モジュールの充足性を評価するプログラムも作成しており、このパイプラインを MAG 解析にも適用できるよう準備を始めた。ドメイン単位のオーソログ解析を行う DomRefine について、マルチプルアライメントから最適なドメイン分割を決定するための DSP スコアの計算をプロファイル計算に基づいて高速化した。MBGD オーソログに基づく機能推定を改善する目的で、MBGD オーソロググループにおける遺伝子系統樹と KEGG Orthology 分類とを重ねて適切なサブグループを切り

出すプログラムのプロトタイプを作成した。また、MBGD のデータ更新にも着手し、オーソログデータの構築を進めている。次期バージョンの登録ゲノム数は 34079 件で、現バージョンから倍増する見込みである。

④ 効率的な逐次 DB 更新・運用システムの"再"構築、ヒトマイクロバイオームのメタデータ整備とキラータセットの開発

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、2023 年度末から急速に普及した ChatGPT に代表される LLM を用いた各種サービスを活用し、大量のマイクロバイオーム論文から自動的に必要なメタデータを取得するプログラムの試作版を作成した。パターンマッチを基本にしていた MEO 等のオントロジー自動アノテーションツール MDBAnnotator についても、LLM を活用して処理の高速化や効率化を行った試作版を作成した。日々増加するマイクロバイオームデータについて最新の BioProject データおよび MAG データを INSDC から自動的に取得できるように DB 更新システムの開発を行った。また、マニュアルキュレーションによる論文からのメタデータ収録は継続して行っている。本データは LLM を活用したデータ取得の正当性を検証するための検証データとしても用いることが期待できる。

⑤ Microbiome Datahub のデータを用いたキラータ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 のマイクロバイオームデータを用いたツールの代表例である PZLAST について、今までのリードから予測された遺伝子のアミノ酸配列とは別に、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のテスト版を開発した。MAG 由来の遺伝子はリード由来の遺伝子と異なりほとんどが完全長の遺伝子配列になっており、各リファレンス配列の長さが元々の PZLAST と比べて平均的に長いため、配列類似性検索のアルゴリズムの見直しが必要であり、さらなる改良が必要である。また、利用者視点に立った DB 開発を進めるために、Microbiome Datahub について主に年齢的に中堅層の微生物系研究者から構成されるアドバイザリーボードを設置した。2023 年度から本格的にアドバイザリー委員会を開催したいと考えている。

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 と連携して遺伝研スーパーコンピュータ上でワークフロー言語を用いた大量データに対する計算を容易に実行できるように開発を進めた。また、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 データ間の連携について DBCLS の開発者と話し合いを行い、2023 年度以降の本格的な連携のための準備を行った。同じ統合化推進プログラムの糖鎖 DB グループとは、微生物の細胞表面の糖鎖構造の DB 化等について連携を開始した。

§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)	1件
	国際(欧文)	1件
未発行論文 (accepted, in press 等)	国内(和文)	0件
	国際(欧文)	0件

② 論文詳細情報

1. Takao K Suzuki, Motomu Matsui, Sira Sriswasdi, and Wataru Iwasaki, "Lifestyle Evolution Analysis by Binary-State Speciation and Extinction (BiSSE) Model", *Methods in Molecular Biology*, 2569:327-342. 2022. (DOI: https://doi.org/10.1007/978-1-0716-2691-7_16)

(2) その他の著作物(総説、書籍など)

1. 森宙史、Microbiome DatahubとPZLAST、実験医学増刊 バイオデータベース・ウェブツール、羊土社、pp101-103、2022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	7件
	国際	1件
ポスター発表	国内	7件
	国際	1件

② 招待講演

〈国内〉

1. 松井求、系統樹推定の常識を問い直す、国立遺伝学研究所研究集会、国立遺伝学研究所、2022年8月3日。
2. 松井求、ポストコッホ機能生態学が目指すもの、ポストコッホ生態研究集会、長崎大学、2022年8月28日。
3. 森宙史、マイクロバイオームからのMAG/SAGデータの取得と活用の展望、NITE スポンサーセミナー、BioJapan 2022、オンライン、2022年10月12-14日。

〈国際〉

該当なし

③ 口頭講演

〈国内〉

1. 森宙史、マイクロバイオーム研究を先導するハブを目指したMicrobiome Datahubの開発、第1回統

合微生物学研究会、ラフォーレ修善寺、2022年9月21日。

2. 森宙史、マイクロバイオーム研究を先導するハブを目指した Microbiome Datahub の開発、トーゴの日シンポジウム 2022、オンライン、2022年10月5日。
3. 鈴木誉保、松井求、森ヶ崎進、大津巖生、土肥裕希、河野祐介、林久喜、高谷直樹、岩崎渉、Trait-based アプローチでせまる 細菌の群集構造とその安定性、第35回日本微生物生態学会、2022年10月31日。
4. 森宙史、マイクロバイオームの国際的なデータハブの構築、第45回日本分子生物学会年会 フォーラム「生命科学のデータベース活用法」、2022年11月30日。
5. 内山郁夫、高柳正彦、CGB: 大規模ゲノム比較のための比較ゲノムブラウザ、第17回日本ゲノム微生物学会年会、かずさアカデミアホール、2023年3月8-10日。
6. 森宙史、マイクロバイオーム研究を先導するハブを目指した Microbiome Datahub の開発、第96回日本細菌学会総会 シンポジウム「微生物バイオインフォマティクスの最前線」、アクリエ姫路、2023年3月17日。
7. 山田拓司、ゲノム・メタゲノムデータの可視化手法の開発、第96回日本細菌学会総会 シンポジウム「微生物バイオインフォマティクスの最前線」、アクリエ姫路、2023年3月17日。

〈国際〉

1. Ikuo Uchiyama, Recent developments in MBDG and its application to genomic functional inference. Novel challenges in the quest for orthologs (7th Quest for Orthologs meeting), Sitges, Spain, Sep. 17-18, 2022.

④ ポスター発表

〈国内〉

1. 藤吉真生、鈴木誉保、松井求、岩崎渉、Bac2Feature: 系統情報を利用した微生物の形質推定パイプライン、第24回日本進化学会、静岡県沼津市、2022年8月5日。
2. 松井求、岩崎渉、シン・近隣結合法: PANJEP 法が問い直す系統解析の常識、第24回日本進化学会、静岡県沼津市、2022年8月5日。
3. 鈴木誉保、松井求、森ヶ崎進、大津巖生、土肥裕希、林久喜、高谷直樹、岩崎渉、微生物群集は普遍的に極めて安定な形質組成を持つ、第60回日本生物物理学会、2022年9月28日。
4. 内山郁夫、三原基広、西出浩世、千葉啓和、高柳正彦、高見英人、MBGD 2022: 大規模な微生物オーソログデータベースの効果的な活用に向けて、トーゴの日シンポジウム 2022、2022年10月5日。
5. 松井求、明石基洋、鈴木誉保、藤吉真生、土肥裕希、森ヶ崎進、林久喜、高谷直樹、岩崎渉、ポストコックホ生態学が解き明かす輪作に呼応した微生物動態、第35回日本微生物生態学会、2022年10月31日。
6. 藤吉真生、鈴木誉保、松井求、岩崎渉、Bac2Feature による微生物群集の表現型解析、第35回日本微生物生態学会、2022年10月31日。
7. 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、マイクロバイオーム研究を先導するハブを目指した Microbiome Datahub の開発、第17回日本ゲノム微生物学会年会、かずさアカデミアホール、2023年3月8-10日。

〈国際〉

1. Hiroshi Mori, Takatomo Fujisawa, Koichi Higashi, Hiroyo Nishide, Zenichi Nakagawa, Yasuhiro Tanizawa, Yasukazu Nakamura, Motomu Matsui, Ikuo Uchiyama, Takuji Yamada, Microbiome Datahub: an integrated database of microbiome samples and MAG data, IHMC Kobe 2022, Kobe Portopia Hotel, Kobe, November 8-19, 2022.

(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

(5) 受賞・報道等

① 受賞

該当なし

② メディア報道

該当なし

③ その他の成果発表

該当なし

§5. 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
2022年 4月20日	チーム内キックオフ会議 (非公開)	オンライン	9人	プロジェクト内のキックオフ会議
2022年 4月25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
2022年 5月9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
2022年 5月23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
2022年 6月11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
2022年 7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
2022年 7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
2022年 7月28日	企業との連携に関する会議(非公開)	オンライン	5人	企業との連携に関する会議
2022年 8月1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
2022年 8月18日	統合化推進プログラムの他のグループとの連携についての会議(非公開)	オンライン	3人	統合化推進プログラムの他のグループとの連携についての会議
2022年 8月22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
2022年 9月11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
2022年 9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
2022年 10月11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
2022年 10月17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
2022年 10月18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
2022年 10月27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
2022年 11月2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
2022年 11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
2022年 12月5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
2022年 12月7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
2022年 12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
2023年 1月24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
2023年 1月25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
2023年 2月9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
2023年 2月17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
2023年 3月2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
2023年 3月22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

(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
2023年 2月5-7日	チーム内 Hackathon(非 公開)	ホテルアソ シア静岡	7人	プロジェクト内のDB開発合宿
2023年 3月12-14 日	チーム内 Hackathon(非 公開)	ホテルアソ シア静岡	6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	維持・発展	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	新規	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を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, Masaki Kato, "MBGD update 2018: microbial genome database based on hierarchical orthology relations covering closely related and distantly related comparisons", <i>Nucleic Acids Res.</i> , 47, D382-D389, 2019. (DOI: 10.1093/nar/gky1054).
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", <i>PLoS Comput Biol</i> , 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.riken.jp/meta	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", <i>Bioinformatics.</i> , 37, 3944-3946, 2021. (DOI: 10.1093/bioinformatics/btab492).