

ライフサイエンスデータベース統合推進事業  
(統合化推進プログラム 2017年採択課題)

## 研究開発終了報告書

「データサイエンスを加速させる微生物統合データベースの  
高度実用化開発」

黒川 顕 | 情報・システム研究機構 国立遺伝学研究所  
情報研究系 教授

研究開発期間： 2017年4月～2022年3月



©2022 黒川 顕 (情報・システム研究機構) Licensed under CC BY 4.0

## §1. 研究開発実施の概要

### 【目的】

我々はこれまで、微生物データの統合化、超高度化を推進し、フル RDF による微生物統合データベース(以降 DB) MicrobeDB.jp を構築してきた。統合 DB においては、全データが連結された巨大なグラフとなっており、利用者は巨大グラフの全貌を把握できないため、「巨大グラフをどのように辿ればどのような答えが出てくるのか」を想定できず、これまでの DB の利用形態を適用する事が本質的に困難となる。これは統合 DB の利用そのものが、新しい科学的手法である事を意味しており、この新しい科学的手法こそが「データサイエンス」であると言えよう。そこで本研究開発では、徹底的なデータ統合および高度化は継続しつつ、「統合化されたデータをどのように渡り歩き、どのような新規知見を得るか」という統合 DB の実用化に向けた、データサイエンスを加速させる統合 DB の利活用方法の開発に重点を置き、MicrobeDB.jp 活用による新たな科学的手法、すなわちデータサイエンス研究手法を提案するとともに、MicrobeDB.jp の実用化を目指す。

### 【方法】

本研究開発では、「統合化されたデータをどのように渡り歩き、どのような新規知見を得るか」、という統合 DB の実用化に向けた、データサイエンスを加速する統合 DB の利活用方法の開発に重点を置き、① 徹底したユーザビリティの向上、② データ品質の向上、③ キラーアプリケーションの開発、④ さらなるデータの統合、⑤ 基盤データ解析技術の高度化、⑥ 効率的運用、⑦ ホログenom対応、の7項目を中心に研究開発を実施した。

### 【結果】

- ① 徹底したユーザビリティの向上: キーワード検索に加えファセット検索可能な UI の開発、RDF モデル更新にともなう全データの更新、ユーザのメタゲノムデータの容易な登録と比較解析。SPARQL エンドポイントの公開。
- ② データ品質の向上: 微生物環境オントロジー-MEO の更新、MEO 自動アノテーションの精度向上、ヒトマイクロバイオームのメタデータの徹底的な RDF 化、新しいオーソログ構築プロトコルを用いた階層的なオーソログデータの提供。
- ③ キラーアプリケーションの開発: 系統組成推定ツール VITCOMIC2 ローカル環境版、ヒトマイクロバイオームと病気との関連性を整理したオントロジー-HMADO、系統組成・遺伝子機能組成解析パイプライン MeGAP の Docker コンテナ版、真菌のメタ ITS 配列解析パイプライン MycoTAP、微生物群集統合可視化ツール LEA、超高速メタゲノム配列相同性検索ツール PZ LAST、の開発。
- ④ さらなるデータの統合: 真菌類のゲノム、ITS、RNA-Seq データの統合。
- ⑤ 基盤データ解析技術の高度化: メタゲノム解析パイプラインの高速化、オーソログ解析パイプラインの効率化。
- ⑥ 効率的運用: RDF の再設計、MEO および HMADO 自動アノテーション技術開発、DB システム全体のコンテナ化。
- ⑦ ホログenom対応: MicrobeDB.jp、Plant GARDEN、MetaboBank の3つの統合化推進プログラムの連携による、ホログenom研究の基盤となる微生物メタゲノム-植物ゲノム-化合物データの統合化。

## §2. 研究開発実施体制

### 1. 研究グループ

#### (1) 「遺伝研」グループ(研究代表者グループ)

- ・ 担当項目

プロジェクト全体の研究総括を行うとともに、① 徹底したユーザビリティの向上、② データ品質の向上、③ キラーアプリケーションの開発、④ さらなるデータの統合、⑤ 基盤データ解析技術の高度化、⑥ 効率的運用、⑦ ホログenom対応、の 7 項目の研究計画について、主体的に研究開発を行なった。

#### (2) 「基生研」グループ(主たる共同研究者グループ(1))

- ・ 担当項目

オーソログに関する② データ品質の向上、③ キラーアプリケーションの開発、⑤ 基盤データ解析技術の高度化、の研究計画について、主体的に研究開発を行なった。

#### (3) 「東工大」グループ(主たる共同研究者グループ(2))

- ・ 担当項目

ヒトメタゲノムデータに関する② データ品質の向上、⑦ ホログenom対応、の研究計画について、主体的に研究開発を行なった。

#### (4) 「千葉大」グループ(主たる共同研究者グループ(3))

- ・ 担当項目

真菌類のデータに関する③ キラーアプリケーションの開発、④ さらなるデータの統合、の研究計画について、主体的に研究開発を行なった。

### 2. 有識者会議等

有識者会議の設置なし。

### §3. 研究開発の目的、実施内容及び成果

#### 1. 研究開発の背景

本研究開発 DB が対象とするのは、主に微生物および微生物叢(マイクロバイオーム)に関する研究分野である。本分野では、H28 年 3 月に JST CRDS 戦略プロポーザル「微生物叢(マイクロバイオーム)研究の統合的推進 ～生命、健康・医療の新展開～」が発表され、特にヒト微生物叢の論文発表数が世界的に急増していること、ヒト微生物叢関連の研究分野について、我が国が世界トップレベルの強みを有するものが多く含まれていること、などが報告された。また、文科省から H28 年度研究開発目標として「宿主と微生物叢間のクロストーク・共生の解明と健康・医療への応用」が発表され、ヒトマイクロバイオーム研究が強く推進されはじめている。ヒトマイクロバイオーム研究以外においては、CRDS 発表の「研究開発の俯瞰報告書 研究開発の新しい動向(2016 年)」にて、ライフサイエンス・臨床医学分野においてアカデミアや企業の期待感が高まっている技術領域として 3 つが挙げられており、その中の一つに、土壌・環境マイクロバイオームの評価・制御基盤技術が、今後大きく注目すべきテーマとして挙げられている。さらに、宿主とマイクロバイオームの相互作用、動態、共進化などをゲノムレベルで解析するホログenom研究も重点化されようとしている。

我々は、第一期および第二期の統合化推進プログラムを通じて、微生物データ統合化の基盤作り、超高度化を推進してきた。統合化推進プログラムの中では先陣を切ってセマンティックウェブ技術によるデータ統合化を推進し、DBCLS の強力なバックアップを得ながら、多様な RDF モデル、オントロジー、ボキャブラリ、Stanza を開発して、フル RDF による微生物統合 DB「MicrobeDB.jp」を構築してきた。

一方で、未だ十分な数の研究者が MicrobeDB.jp を利用しているとは言い難く、ユーザの獲得が喫緊の課題となっている。土壌やヒト腸内メタゲノム研究における応用論文の発表のみならず、日本語総説の発表、学会や展示会などでの積極的な出展を通して、きめ細かくユーザニーズを拾い上げ、それらを可能な限りシステムに反映させてきた。その結果、多数の研究者および企業から共同研究や事業化に関する機会を頂いたが、MicrobeDB.jp のユーザ数の爆発的な増加を得るに至っていない。これには様々な原因が考えられるが、開発してきた微生物統合 DB のもつ、全データが連結された巨大なグラフになっているという統合 DB 特有の性質に起因する、と考えるに至った。レガシーな DB を検索する際、利用者は得られる結果をある程度想定した上で検索を実行する。一方で、統合 DB を検索する際、利用者は巨大なグラフの全貌を把握できないため、「巨大グラフをどのように辿ればどのような答えが出てくるのか」を想定する事が困難となる。例えるならば、全体像がわからない迷路で、出口に辿り着く方法を探す、という問題と同様である。すなわち、統合 DB に対しては、データ統合化のメリットを存分に活かすという点において、これまでの DB の利用形態を適用する事が本質的に困難となる。これは統合 DB の利用(統合 DB に対する検索行為)そのものが、新しい科学的手法である事を意味しており、この新しい科学的手法こそが「データサイエンス」であると言えよう。今まさに、データサイエンスを切り拓くため、統合 DB 利用による新たな科学を実証すべき時が来たと考えている。

## 2. 研究開発対象のデータベース・ツール

### (1) データベース

- 主要なもの

正式名称	略称	概要
MicrobeDB.jp		微生物に関する様々な情報をゲノム情報を核として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DB <a href="https://microbedb.jp/">https://microbedb.jp/</a>

- 上記以外のもの

正式名称	略称	概要
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 DB。MicrobeDB.jp において、ゲノムデータ間を統合するために必須の DB である。 <a href="https://mbgd.nibb.ac.jp/">https://mbgd.nibb.ac.jp/</a>

### (2) ツール等

正式名称	略称	概要
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。 <a href="https://bioportal.bioontology.org/ontologies/MEO">https://bioportal.bioontology.org/ontologies/MEO</a>
Microbial Culture Collection Vocabulary	MCCV	JCM や NBRC などの、菌株保存機関に蓄積されている菌株情報を記述するための統制語彙。 <a href="https://bioportal.bioontology.org/ontologies/MCCV">https://bioportal.bioontology.org/ontologies/MCCV</a>
Human Microbiome Associated Disease Ontology	HMADO	発症にヒトマイクロバイオームの関連が疑われるヒトの病気のオントロジー

※データベース、ツールの詳細は別紙参照。

## 3. 達成目標及び実施計画

### (1) 当初の実実施計画・達成目標

#### 1. 遺伝研グループ

##### 研究開発実施項目(1): MicrobeDB.jp ポータルサイト構築

MicrobeDB.jp のユーザーインターフェース拡張によるユーザビリティの向上を図るため、これまでのシンプルなキーワード検索に加え、MicrobeDB.jp のフル RDF DB の利点を活用した、新しいポータルサイトを構築する。また、藻類、菌類等のデータセットや MBGD、MeGAP による解析データセットなどを中心にユースケースを積み上げて提示する。さらに、GenomeRefine を認証付きの API サービスとして改変し、新しいポータルサイトから呼び出すことで、ユーザのデータ登録環境もポータルサイト上に配置する。このポータルサイト構築にあたっては既存のユーザおよびユーザ会、また講習会におけるユーザからのフィードバックを受け、継続的に改善を図って行く。登録されたデータについては、公共 DB サービスとの間で、ユーザ認証および登録インターフェースを含めユーザ利便性向上のための連携を図る。

## 研究開発実施項目(2): MicrobeDB.jp ユーザデータアーカイブサービス拡張

GenomeRefine で開発した機能を、ユーザ・グループ認証サーバーを介して認証する機能を付与した API 化を実施する。メタデータと API 入力パラメータにセマンティクスを紐づけることによって、RDF 変換の汎用化を行う。本 API は構築予定のポータルサイトから呼び出すことにより、ユーザは MicrobeDB.jp 上で直接データ登録が可能になり、比較解析 Stanza との接続によるデータ解析機能への誘導がスムーズになる。また、この API を介することで、当グループが提供している微生物ゲノム解析パイプライン DFAST(<https://dfast.nig.ac.jp>)などの解析支援アプリケーションとの広範な連携を拡張していく。さらに、プライベートな API 等を開発し、協力体制を確立しながら NBD C ヒト DB のヒト資源由来のメタデータでの運用も目指していく。

## 研究開発実施項目(3):メタゲノムや RNA-Seq 等のマルチオミックスデータを用いた Stanza およびアプリケーションの開発

マルチオミックスデータを中心に、対応トピックモデル等の高度な統計解析手法やインタラクティブな視覚化手法を用いた様々な Stanza およびアプリケーションを開発し、ユースケースに応じた DB データの様々な解析結果を利用者に提供できるようにする。

## 研究開発実施項目(4):オミックスデータ RDF 拡張とアノテーション統合

配列 ID ならびに配列アライメント・マッピング情報を集積し、配列位置情報オントロジー FALDO に基づく変換系を構築することで、ゲノムアノテーションの統合と再利用をさらに広い範囲で可能とするゲノム RDF の変換系を実現する。また、RDF を介して統合されたアノテーションを新たなゲノムリファレンスとする GFF 形式等の、ゲノム解析の現場で汎用的に用いられているファイル形式に変換し、ユーザの利便性を図るとともに、RDF を起点とした新しい統合データの枠組みを提供することを目指す。

## 研究開発実施項目(5):メタゲノム解析パイプラインの速度・精度向上と真核メタ ITS 解析への対応

MicrobeDB.jp で用いているメタ 16S・メタゲノム解析パイプラインを最新の解析手法を取り入れたものへとアップデートする。今後増加が予想される真核メタ ITS 解析データについては、高速かつ高精度な系統組成推定のための解析パイプラインを新たに開発し、第 3 年次末までに、公共の真核メタ ITS 解析データを網羅的に収集、再解析し、真核微生物についても環境中の分布を MicrobeDB.jp 上で検索、比較解析できるようにする。上記の開発した解析パイプラインを、ユーザのデータの解析にも利用できるようにするために、MeGAP を改良する。

## 研究開発実施項目(6):オントロジー自動アノテーションツールの精度向上

文献情報と紐づけられたサンプルについては、文献の要旨や本文等からテキストマイニングを行い、サンプルのオントロジーアノテーションに必要な情報を抽出した上で、自動アノテーションを行う機能をツールに付与することで、オントロジー自動アノテーションツールの精度向上を達成する。さらに、開発したオントロジー自動アノテーションツールを、MicrobeDB.jp における MEO 以外のオントロジーのアノテーションにも適用できるようにツールを改良する。

## 研究開発実施項目(7):藻類および植物ホロゲノム研究等を支援するサンプルメタデータおよび生物分類の再資源化

公共 DB に登録されている生物分類に関するメタデータについて、RDF 統計データを含む、データの RDF 化を実施する。さらに、RDF ガイドラインに従った既存の主要なオントロジーのマッピングによるリソース化を実施する。これらのリソースはすでに強力な連携を有する国立環境研究所や藍藻ゲノム DB CyanoBase, 根粒菌ゲノム DB RhizoBase のそれぞれの国内外のユーザコミュニティとの連携を図りながら、必要に応じて生物分類情報を拡張し、藻類および植物ホロゲノムの研究を支援できる RDF を開発する。また、RDF を起点とする微生物および植物ゲノムリファレンス情

報をホログenom解析に利用可能とする。

## 2. 基生研グループ

### 研究開発実施項目(1):オーソログデータの構築

オーソログデータは、これまでに開発した階層的なオーソログ解析プロトコルに従いつつ、効率よく差分更新を行えるように、データ構築プロセスを修正しつつ更新漏れが起きないような頑健プロトコルを開発しデータ更新を行う。新規のゲノムデータを追加する際に、既存のゲノムデータとの重複度や重要度、クオリティなどに応じて優先度をつけ、優先度の低いものを更新対象から外すことによってデータ量を減らしつつ、必要十分なデータを含むオーソログデータを構築する仕組みについて検討する。その際、解析に必要なデータを後から差分的に付け加えられるような仕組み作りも行い、長期にわたって最善のオーソログデータの維持ができるようにすることを目指す。

### 研究開発実施項目(2):比較ゲノム解析アプリケーションの開発

これまでに作成したオーソログ DB やデータ表示用 Stanza を活用しつつ、利用者がより簡便に比較解析を実行できるようなアプリケーションを開発する。特に、利用者が指定したゲノム配列に対してオーソログデータを検索してアノテーションをつける「オーソログ検索・アノテーション」、指定した生物種群に存在し、それ以外の生物種に存在しないような遺伝子を探索する「系統プロファイル解析」、および同種・近縁種ゲノム間での遺伝子の保存性や多様性を解析する「コアゲノム・パンゲノム解析」などについて、利用者が柔軟に条件を指定して、その結果を様々な角度から検討できるようなインターフェースを作成する。また、これらのアプリケーションを作成するに当たり、必要に応じて、これまでに開発したオーソログデータ用のオントロジー ORTH の拡張を行う。

## 3. 東工大グループ

### 研究開発実施項目:ヒトマイクロバイオーム関連のメタデータ整備

ヒトマイクロバイオームデータにおいて、公共 DB には登録されていない詳細な疾患関連情報に関するメタデータを論文本体から直接抽出し、MicrobeDB.jp の配列データと紐付けた RDF データを作成する。MicrobeDB.jp では、今までヒトの病気については感染症-症状オントロジーである PD O-CSSO に存在する感染症のみを扱ってきたが、本研究開発ではテキストマイニングの手法を用い、これまで感染症の病原菌としては報告されていないヒト常在菌がどのような疾患と関連するかを論文から抽出する。これらの結果をマニュアルキュレーションにより精査し、既存の標準的に用いられているヒトの疾患オントロジーを用いてアノテーションを行い、ヒトマイクロバイオームデータを解釈する上で重要になる、微生物-疾患関連性データの充実を行う。

## 4. 千葉大グループ

### 研究開発実施項目:真菌類ゲノム・菌株・オミックス情報の収集と高度化

真菌類のゲノム情報と生息環境や生理活性等のメタデータの収集および分類情報の整理を核として、MicrobeDB.jp の真菌に関するデータの基盤整備を実施する。具体的には、真菌類の完全長・ドラフトゲノムデータ、RNA-Seq データを公共塩基配列 DB から収集する。さらに、菌種同定に用いられる各種系統マーカー遺伝子(ITS 配列など)の配列データの整備を行うことで、遺伝研グループで行う公共の配列 DB 由来のメタ ITS 解析データの菌種同定に用いることが可能なリファレンス配列データを作成する。菌株については、発酵や臨床面で最も重要なアスペルギルス属を中心に生理活性や培養条件等のメタデータ収集を行い、セマンティックウェブ技術を用いたデータの記述と他のデータとの統合化を行う。

### 第3年次末までの達成目標

- MicrobeDB.jp トップページとして新たに開発するポータルサイトに、統合 DB 中のデータを駆使したユースケースを示すため、Stanza およびアプリケーションを 10 種以上設置する。
- 真核メタ ITS データに対する解析パイプラインを開発し、MicrobeDB.jp にて公共の真核メタ ITS データの検索および比較解析を可能にする。
- MicrobeDB.jp で行っている、MEO 等のオントロジーを用いたメタデータのアノテーションにかかる時間を、オントロジー自動アノテーションツールの精度向上により 60%低減する。
- オースログデータを段階的に構築するプロトコルについて、差分更新によって効率的に更新が行える体制を構築する。これを用いて、公表されたゲノムデータ全体を取り込んだオースログ DB の構築を行う。
- 更新対象とするゲノムに優先度を付けて、優先度の低い生物種を更新対象から外すことによって、データサイズを減らしつつ、必要十分なゲノムデータを取り込んだオースログ DB を構築するための汎用的な仕組みを開発する。
- オースログデータに基づいて比較ゲノム解析を行うアプリケーションとして「オースログ検索によるアノテーション」、「系統プロファイル解析」、「コアゲノム・パンゲノム解析」などを実装する。
- これら開発を通して、MicrobeDB.jp の年間訪問者数を 13,000 人以上にする。

### 最終年度末までの達成目標

- ホログenom研究に活用できるよう、植物統合 DB と密に連携する。
- MEO 以外のオントロジーアノテーションについても、MEO と同水準で自動化する。
- オントロジー自動アノテーションの高効率化を受け、MicrobeDB.jp のローコストかつ定期的なバージョンアップを実施する体制を整備する。
- より安定的な運用を目指して、DDBJ との連携を強める。
- 利用者のフィードバックに基づいてインターフェースの改良を行うとともに、「シンテニー解析」や「ゲノムアライメント」などの新たな比較ゲノム解析アプリケーションの開発を行う。
- 3 年次までに開発した、選択的にデータを取り込むことによりオースログデータを更新する仕組みを実装し、ゲノムデータの規模が計算機の処理能力を超えて増大した場合でも、規模を抑えつつ最善のオースログデータを作成できるような更新体制を構築する。
- これら開発を通して、MicrobeDB.jp の年間訪問者数を 26,000 人以上にする。

## (2) 期間中に追加・削除・変更した実施計画・達成目標

### 【第 1(2017)年次追加支援 1】

高度実用化のユースケースの一つとして、ホログenom研究の基盤の提供およびそのためのホログenom研究支援ツール開発を実施する。将来的にヒトホログenom研究 DB 化や解析支援系にも活用可能な枠組みとして、まず、植物ホログenom解析支援のための開発を共同研究として実施する。具体的には、基礎生物学研究所共生システム研究部門・川口正代司教授を生産者とするホログenomデータ連携の共同研究要請により、かずさ DNA 研究所で実施されている植物統合化推進プログラムと協力した開発として、①共生微生物ゲノム情報統合環境の構築と、②ホログenom解析支援ツールの開発を実施する。

### 【第 1(2017)年次追加支援 2】



ゲノム比較の結果をゲノムの代謝機能ポテンシャルの観点から評価するため、海洋研究開発機構の高見らによって開発された代謝ポテンシャル評価ツール MAPLE と連携させた解析機能を作成する。この目的のため、MBGD と Maple を連携させて、本システムのオーソログアノテーション結果から各ゲノムの KEGG Module の充足率を計算できるようにするとともに、その結果に基づいてゲノム間で代謝ポテンシャルの違いを評価するためのアプリケーションを開発する。このアプリケーションによって、未培養細菌から再構築したゲノムデータの機能上の特徴を、比較解析を通じてあぶりだすことを実現させて、本 DB の利用の呼び水となるような成果を導くことを目指す。

### 【第 5(2021)年次追加支援】

MicrobeDB.jp、Plant GARDEN、MetaboBank の3つの統合化推進プログラムが連携によって、さらなるデータ統合を進展させ、データ統合の価値を、セマンティックウェブ技術を意識させることなく、一般ユーザが利用できるようなサービスを新たに提供する。具体的には、(i)メタボローム統合グループにおいて、KNAPsAcK の RDF 化を実施し、代謝物と生物種および文献情報の新たなのりしろとなる関係情報を RDF によってデータ統合することによって、MicrobeDB.jp や Plant GARDEN の生物ドメイン単位で統合されたデータベースと代謝物との新たなデータ統合を実現する。(ii)微生物統合グループにおいて、SPARQL クエリを起点としたデータ変換・アノテーションのバッチ処理系を開発する。本処理系を用いて、植物統合グループが開発を実施する連携検索システムに投入する。(iii)植物統合グループにおいて、生物種、化合物や植物遺伝子単位で検索可能なインターフェースを提供する推進プログラムデータベース連携検索システムの開発を実施する。この連携検索システムから、生物種検索から各統合化推進プログラムのデータベースが扱うデータセットへの検索を実現する。

現在、Plant GARDEN では PubTator による文献アノテーションなど植物遺伝子やリソースののりしろになる RDF 化も実施している。それらのデータセットに加えて、MicrobeDB.jp が扱うメタ 16S 解析データやメタゲノム解析データおよびオントロジーアノテーションされたサンプル属性情報を用いて、植物遺伝子に対する絞り込み検索を実現する。さらには、Rhea の代謝物とタンパク質の関係情報に加えて、データ統合によって KNAPsAcK および PubTator から新たに得られた化合物と遺伝子の関係情報を用いて、化合物を検索クエリとした遺伝子・ゲノム検索可能なインターフェースを提供し、化合物から遺伝子関係情報を介して、植物・微生物のゲノム座標にリンクできることを目指す。(ii)について、微生物統合グループが実施し、連携システムに対してデータセットを生成するための基盤構築を担当する。各統合化推進プログラムのデータベースにより RDF レベルで統合されたデータセットが格納された Virtuoso コンテナから Elasticsearch コンテナへ変換・投入することで、植物統合グループが開発する連携検索システムの検索エンジンに格納する。併せて、本処理系は、MicrobeDB.jp のデータセット更新における RDF リソースから検索インデックスを生成する部分を自動化も兼ねることで、データベース運用のサステナビリティ向上を目指す。さらに、Togo DB、D2RQ Mapper、RDF ポータルの SPARQL 各種サービスやツールを有効に活用して、NBDC や DBCLS と連携しながら、植物統合グループおよびメタボローム統合グループへの技術的な支援を実施する。

## 4. 実施内容

### (1) 実施内容

本研究開発では、MicrobeDB.jp の実用化を目指し、中心となる以下の 7 項目について研究開発を実施した。

#### ① 徹底したユーザビリティの向上

ゲノム・メタゲノムデータの拡張および全ての RDF を更新したデータセットを格納した MicrobeDB.jp version 3 を、2020 年 2 月にリリースした。ユーザビリティの見直しを行った結果、これまでのシンプルなキーワード検索からなるトップページを大幅に改変し、UI を再構築することでシンプルな操作で様々な機能を提供するポータルサイトを新たなトップページとした。本リリースに合わせて、各種ドキュメントを整備し、ユーザ会、講習会、DBCLS の協力による TogoTV などポータルサイトの利用を促進するためのコンテンツ等を含む MicrobeDB.jp のユースケースを提示した。さらに、MicrobeDB.jp で公共のメタゲノム配列データから系統組成や遺伝子機能組成を推定可能な解析パイプライン MeGAP の Docker/Singularity コンテナ化を行うことで、遺伝研スーパーコンピュータ上でユーザが容易に MicrobeDB.jp の解析パイプラインと同一の方法でメタゲノム解析をできるようにした。ユーザ会、講習会等によって得られたユーザフィードバックを精査し、MeGAP で解析した結果のメタゲノムの系統組成・遺伝子機能組成データを、MicrobeDB.jp のポータルサイトからスムーズにアップロードし、解析 Stanza で他のサンプルと組成の比較や閲覧ができるようにした。大規模解析を行うバイオインフォマティクス研究者向けに、MicrobeDB.jp version 3 の RDF データの SPARQL エンドポイントを構築、公開することで、ユーザが好きな SPARQL クエリで MicrobeDB.jp のデータを検索できるようにした。

#### ② データ品質の向上

メタゲノムのメタデータへの MEO および HMADO 自動アノテーション結果を元に、MEO と Human Microbiome Associated Disease Ontology (HMADO) への語彙の追加を行うと共に自動アノテーションワークフローの見直しを行い、より詳細かつ正確に微生物の環境情報を記述可能にした。特に、ユーザからのニーズが大きいヒトマイクロバイオーム関連のメタデータ整備を重点的に行い、メタゲノム配列が紐付いた論文約 4,500 報を文献情報データベースから収集し、マニュアルキュレーションによってメタデータ抽出をおこない、既存のメタゲノム RDF データにそれらのデータを追加することで、MicrobeDB.jp 上で年齢や住んでいる地域、抗生物質の使用の有無等のメタデータでヒトマイクロバイオームデータを絞り込み検索可能にした。微生物のオーソログデータについては、新しいオーソログ構築プロトコルを用いて階層的なオーソログデータを構築し、MBGD の新バージョンとして公開した。

#### ③ キラーアプリケーションの開発

大量のメタゲノム配列データを持つユーザが、MicrobeDB.jp と同じ解析パイプラインを自らの解析環境で実行可能にするために、系統組成推定ツール VITCOMIC2 ローカル環境版、および系統組成・遺伝子機能組成解析パイプライン MeGAP の Docker/Singularity コンテナ版を開発した。MicrobeDB.jp 中のメタゲノムサンプルの遺伝子機能組成データを用いて、微生物群集の全体像を俯瞰可能な統合可視化ツール LEA を開発して公開した。データ点が増えた状態でも微生物群集構造と環境との関係性を直感的に理解できるようにするため

に、LEA の描画処理を効率化および 3D に拡張した。さらに LEA の遺伝子機能組成版のアルゴリズムも開発した。メタゲノム配列データに対して配列相同性検索を行いたいと多数のユーザから要望が寄せられたため、MicrobeDB.jp のメタゲノムアミノ酸配列データ約 2.5 TB に対して超高速にアミノ酸対アミノ酸の配列相同性検索を行う Web アプリケーション PZLAST を開発して公開した。ヒトマイクロバイオーームと疾患との関連性を整理し、疾患横断的な検索や比較解析を可能とするために、微生物の関与が疑われる疾患について、約 300 の Class で整理した HMADO version 0.1 を構築した。この HMADO でメタゲノム・ゲノムのメタデータを自動アノテーションすることで、MicrobeDB.jp 上でオントロジーを用いた統一的な検索および系統組成や遺伝子機能組成の比較解析を可能にした。真菌のメタ ITS 配列データから高速かつ正確に系統組成を推定する解析パイプライン MycoTAP を開発し、既存の他の解析パイプラインとの性能比較を行った上で、MycoTAP のコードを公開し、現在論文執筆中である。MicrobeDB.jp 上で MycoTAP を用いて既存の真菌のメタ ITS 解析結果を MEO や HMADO 等のオントロジーで整理したメタデータと共に検索や閲覧できるように、現在作業中である。

オーソログ検索・ゲノムアノテーションアプリケーションとして、ユーザのゲノム配列中の各遺伝子に対して、MBGD のオーソロググループの検索を行って対応付けを行う機能と、オーソロググループから KEGG Orthology へのクロスリファレンスを用いて、KEGG モジュール単位で機能を評価する Genomable 解析を行うツールを開発した。MBGD に登録されたゲノムについては、MBGD のオーソロググループに基づいてあらかじめ Genomable 解析まで行った結果をストアしており、直接呼び出して比較することができる。これに、MyMBGD から登録したユーザのゲノムに対して行った解析結果も併せて比較できるようにした。このツールは、MBGD におけるユーザーゲノム解析機能である MyMBGD 機能の一環として動作する一方で、MicrobeDB.jp のアプリケーションとしても動作するように、MicrobeDB.jp の認証システムとも連携するように設計した。MBGD のオーソログの系統プロファイルアプリケーションとして、系統群や表現系、生存環境などから興味のある条件をユーザが指定して、関連する系統プロファイル情報(各生物種に対するあるなしのパターン)を取得し、それに近いあるなしパターンを示すオーソロググループを検索可能な UI を作成した。

#### ④ さらなるデータの統合

本研究開発で開発した MycoTAP の参照配列データとして利用している UNITE DB の系統 ID と、MicrobeDB.jp で用いている系統 ID である NCBI Taxonomy ID 間での対応付けを行った上で、MycoTAP による真菌のメタ ITS 解析結果を RDF 化し、MicrobeDB.jp に格納できるようにした。千葉大真菌センターの菌株リソースデータの MicrobeDB.jp 上での統合を行うために、DDBJ と千葉大真菌センター間の連携を仲介し、菌株ごとに BioSample ID を付与するとともに、千葉大真菌センターが持つ菌株リソースのメタデータを一括して DDBJ の BioSample DB に登録できるようにした。MicrobeDB.jp では BioSample のメタデータの RDF 化を既に行っているため、DDBJ BioSample DB に菌株データが登録されれば、MicrobeDB.jp における菌株メタデータも自動的に RDF 化できるようになった。真菌類の RNA-Seq データについて、真核生物用の RNA-Seq 解析パイプラインを用いて各系統の遺伝子毎に発現量を算出し、メタデータと共に RDF 化することで、MicrobeDB.jp に真菌の RNA-Seq の RDF データを格納できるようにし、さらに真菌 RNA-Seq 解析データを閲覧するための Stanza も開発した。

⑤ 基盤データ解析技術の高度化

メタゲノムデータの解析パイプラインで用いる系統アサインメント用の参照 16S rRNA 遺伝子配列 DB を、Ribosomal Database Project の分類体系および配列データを基に構築し、VITCOMIC2 および MeGAP から利用できるようにした。さらに、MicrobeDB.jp で系統分類の基盤としている NCBI Taxonomy の分類体系と整合性がとれるように、両分類体系間の対応付けを行なった。メタゲノムデータの解析を高速化するために、解析パイプライン MeGAP 内で用いる各プログラムを見直し、改良前 version と同等の解析精度を保った上で、10 倍近い高速化を実現した。爆発的に増加を続けるメタゲノムデータについて、追加で解析が完了したメタゲノムサンプルのメタデータおよび解析結果の MicrobeDB.jp 上での差分更新を行うために、遺伝研のスーパーコンピュータ上で計算が終わったサンプルについて、自動的に結果をダウンロードし系統組成と遺伝子機能組成の RDF 化を行えるようにした。MBGD のオーソログデータ構築時に DomRefine において行うマルチプルアライメントの作成が、巨大なオーソロググループの場合に多大な時間がかかったため、大きなオーソロググループについてはプロフィール検索によって処理を簡略化しサブグループ化を行うことで、効率よく処理できるようにオーソログデータ構築プロトコルを改良した。

⑥ 効率的運用

ゲノム、メタゲノムのサンプル情報 RDF に関して、公開 BioSample データおよびユーザ登録サンプルメタデータの RDF を再設計し、データモデルの共通化を実施した上で、MicrobeDB.jp version 3 を 2020 年 2 月にリリースした。データの RDF 変換およびオントロジーの自動アノテーション、DB 投入、検索インデックスの作成等の一連の DB のデータ更新に必要な各作業を API 化し、DB の更新の半自動化と効率化を行なった。また、コンテナが参照する複数のコードレポジトリを跨いだ定義情報を環境変数化し、その依存関係を整理することにより、簡便に MicrobeDB.jp のミラーデータベースの構築を可能にし、環境構築を効率化した。自然言語のパターンマッチと例外処理を基本とした MEO の自動アノテーションのアルゴリズムを見直し、教師データの追加および改良を行い、MEO 自動アノテーションの精度向上を行なった。MEO 自動アノテーションの精度向上によって、自動アノテーション後のマニュアルでの修正に必要な時間と手間が激減し、MEO アノテーションに必要な時間を MicrobeDB.jp version 2 の 1/10 以下に減らすことができた。MEO 用に開発したオントロジー自動アノテーションプログラムを HMADO にも利用可能にし、HMADO アノテーションに必要な時間も大幅に短縮することができた。

⑦ ホロゲノム対応

ヒト-微生物叢ならびに植物-微生物叢の全体像解明を目的としたホロゲノム解析支援系の構築を目標に、ホロゲノム研究の基盤として MicrobeDB.jp を利用できるようにするためのユースケースとして、同様にセマンティックウェブ技術を用いて植物のデータの統合化を進めているかずさ DNA 研究所のグループと SPARQLthon 等で連携を模索してきた。その結果として、最終年度の追加支援に採択いただき、MicrobeDB.jp、Plant GARDEN、MetaboBank の3つの統合化推進プログラムの連携によって、ホロゲノム研究に関するデータ統合を大幅に発展させられるような研究開発を現在行なっている。具体的には、MicrobeDB.jp が扱う植物-微生物叢のメタゲノム解析データおよびオントロジーアノテーションされたメタゲノムサ

ンプルのメタデータを用いて、微生物と関係する植物の遺伝子に対する絞り込み検索を実現するための開発を行なっている。最終的に、データ統合によって KNApSack から新たに得られた化合物と遺伝子の関係情報を用いて、化合物を検索クエリとして遺伝子・ゲノムを検索可能なインターフェースを開発し、化合物と遺伝子の関係情報を介して、植物・微生物のゲノムにリンクできるようにする。

## (2) 実施内容のうちの特定項目の詳細

### ① 研究コミュニティを含むデータ提供者との連携・協業

日本ゲノム微生物学会および日本細菌学会と連携して、2018年に両学会年會中で MicrobeDB.jp の使い方の講習会を行うとともに、ユーザ会を立ち上げ DB へのニーズの収集を行なった。AMED CREST「微生物叢と宿主の相互作用・共生の理解と、それに基づく疾患発症のメカニズム解明」においては、支援班を組織するとともに MicrobeDB.jp の活用を促進した。第2期 SIP「スマートバイオ産業・農業基盤技術」において植物-微生物叢のホログenom研究を行なっている研究グループと共同で、2019年のトーゴの日シンポジウムにおいてワークショップを行い、ホログenomの研究者の DB へのニーズを収集した。

### ② データベース利用者への周知、利用者との連携・協業

日本ゲノム微生物学会、日本細菌学会、日本分子生物学会、日本農芸化学会にはほぼ毎年ブース出展し(延べ15回:2021年度は除く)、MicrobeDB.jpの周知およびニーズ収集を行なった。

MicrobeDB.jpをデータ基盤としたツール群 LEA および PZLAST を開発、論文発表し、MicrobeDB.jpの有用性をアピールすると共に、それらのツールから MicrobeDB.jp への導線を構築した。

### ③ 利用者にとって有用なデータ基盤の構築

統合化推進プログラムの中では先陣を切って、セマンティックウェブ技術を用いた多様なデータ(微生物のゲノム、オーソログ、メタゲノム、菌株リソース、RNA-Seq、配列に紐づく Bio Sample 等)の RDF 化および統合化を推進した。

RDF 化したデータについて、ユーザのデータ利活用を促進できるように、キーワードによる単純な SPARQL 検索だけでなく、オントロジーの階層構造や数値メタデータを用いたデータの絞り込み検索を実現し、DB のユーザビリティの向上に努めた。

DBCLS が主催する SPARQLthon や BioHackathon などの開発会議を通じて植物やメタボローム等他の統合化推進プログラムと積極的に連携し、オントロジーの選択や RDF モデルの設計等で技術提供を行うことで、微生物以外の生命科学データについてもセマンティックウェブ技術による統合化を推進し、様々なデータの統合化の成果を、セマンティックウェブ技術を意識させることなく、一般ユーザが利用できるようなサービスを新たに提供できるように開発を行なってきた。

### ④ 持続的なデータベース運用体制の構築に向けた取り組み

システムのコンテナ化によって DB の環境構築を効率化し、データ更新をより容易にできるようにすることで、メタゲノムデータを中心としたデータの逐次更新を可能にした。

データの RDF 変換およびオントロジーの自動アノテーション、DB 投入、検索インデックスの作成等、DB のデータ更新に必要な各作業を API 化し、DB の更新の半自動化と効率化

を行なった。特に時間コストの高かった MEO 等のオントロジーのアノテーションについては、自動アノテーションプログラムの改良を重ね、高い精度を維持しつつオントロジーアノテーションにかかる時間を大幅に短縮する事を可能とした。

BioSample RDF の作成や RDF データバリデーションシステムや認証システムの共有等、DDBJ との連携を強めることにより、持続的な DB 運用体制を維持できるように努めている。

⑤ 人材の育成

情報・システム研究機構および国立遺伝学研究所の事業計画に、データサイエンスやオープンデータの推進を取り入れることで、当該分野における研究者ポストの確保に努めた。その結果、遺伝研グループで研究開発に大きく貢献した若手研究者 2 名について、本プロジェクトにおける研究開発の内容も大きく評価され、それぞれ DDBJ のデータベース部門長と遺伝研の PI として昇進した。また特任研究員 1 名も助教に内定している。

⑥ 国際連携・国際貢献

微生物のゲノムやメタゲノムに対するメタデータのデータ形式等を決めている国際コンソーシアム Genomic Standard Consortium の会議に積極的に参加し、生物の環境オントロジーの国際標準である Environment Ontology の開発者とも連携して、MEO と Environment Ontology とのマッピング RDF を作り、Environment Ontology のユーザが MicrobeDB.jp のデータを MEO を介して容易に利用できるようにした。

アメリカにおいてメタゲノムデータの統合化を推進しているグループである、National Microbiome Data Collaborative (NMDC) から MicrobeDB.jp に接触があり、現在 NMDC と MicrobeDB.jp 間の連携を模索中である。

⑦ その他

特になし。

## §4. 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(月間平均)

名称	種別	2017年度	2018年度	2019年度	2020年度	2021年度 (9月時点)
MicrobeDB.jp	訪問者数	6,764	5,927	5,497	2,732	1,976
	訪問数	7,658	9,792	10,828	5,987	3,932
	ページ数	28,203	29,444	31,254	104,238	175,575

#### (2) 分析

- 2016 年末に MicrobeDB.jp version 2 をリリースし、2020 年 2 月に MicrobeDB.jp version 3 をリリースした。version 2 リリース後、version 3 リリースの 2019 年度末まで徐々に訪問者数は減ってきていたが、version 3 のリリース後の 2020 年度に訪問者数が半減した。この減少は、version 3 リリースによって MicrobeDB.jp の UI が大きく変わったため、検索エンジンからの人の流入が変わったことが一因として考えられる(逆 SEO 対策となってしまった)。また、COVID-19 により各国の微生物研究者のアクティビティ等に大きな影響があったことも関係していると推測される。訪問者数の減少に伴い訪問数も 2020 年度には大きく減少したが、ページ数については version 3 をリリースした後劇的に増加している。これはヘビーユーザーが誕生しメタゲノムデータの絞り込み検索等の利用が活発化したものと考えられ、論文発表等により今後 DB の有用性が広範囲に認知されれば、ユーザ数や訪問数も増加すると期待される。

### 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

特になし。

### 3. 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)

国内の複数の企業と MicrobeDB.jp を活用した共同研究を進めており(計 7 社)、企業 1 社に MicrobeDB.jp version 3 を社内専用 DB として納品済みである。また、これと異なる企業 1 社には MicrobeDB.jp のヒトマイクロバイオーム関連データを提供している。トヨタ自動車株式会社と共同で行っている「Genki 空間研究」については、空気中の微生物叢に対する VITCOMIC2 や LEA 等のツールによる情報解析がプロジェクトの中核を担っており、本プロジェクトで培った情報解析技術やメタデータの整理方法等が活用されている(特許出願済:特願 2021-005128)。この「Genki 空間研究」については、トヨタ自動車の Web サイトでも大きく取り上げられている([https://www.toyota.co.jp/jpn/tech/partner\\_robot/news/20210331\\_01.html](https://www.toyota.co.jp/jpn/tech/partner_robot/news/20210331_01.html))。この研究成果は 2021 年 9 月に開催されたバイオインフォマティクス学会主催の IIBMP2021 において優秀ポスター賞を受賞している。

革新的先端研究開発支援事業(AMED-CREST:2017 年度～)および戦略的イノベーションプログラム(SIP:2018 年度～)では、遺伝研が領域から得られるマイクロバイオームデータを集約する研

究開発拠点を担っており、そのデータ共有コアとして MicrobeDB.jp が活用されている。

本プロジェクトで開発した MEO 自動アノテーション API は、DBCLS および NITE でも菌株の分離環境のアノテーション等で用いられている。



## §5. 今後の展開

本研究開発では、ゲノム情報を核として様々な微生物学上の知識を統合し、幅広い分野での微生物学の発展に資することのできる統合 DB「MicrobeDB.jp」を、持続可能なシステムに昇華させるとともに、ユーザビリティの向上を徹底する事で、単なる統計量の検索や羅列ではなく、大規模データから新規知識発見を容易に引き出す事が可能なシステムを構築する事を目的とした。微生物データの統合化、さらには統合 DB に対する解析プロトコルの実装は、微生物の体系的な理解を促進し、これまでの仮説検証型の研究のみならず、膨大なデータの中から新たな仮説を導くデータ駆動型の研究を強力に推進する事を可能とする。新型シーケンサー等の飛躍的なハイスループット化、高精度化は今後ますます加速され、ゲノム研究やメタゲノム研究に代表されるデータ駆動型研究が本格化しつつある。これらゲノムデータは一般化してきており、多様な産業における微生物ゲノム・メタゲノム情報の応用も加速度的に推進されている。分散している膨大なデータを集約・統合した DB の運用は、工業、農業、畜産業や医薬、健康など、微生物が関与する広範な産業にとって最重要課題となった。我々は MicrobeDB.jp を基盤として、工業(廃水処理)、農業(農作物品質管理、高機能作物)、畜産業(養豚)、医薬(創薬、疾患マーカー探索)、健康(食品開発)、新規産業(ヒト環境)に関する企業との共同研究を進めている。中でもトヨタ自動車との共同研究(§4-4 参照)においては、これまでにない全く新しい分野を開拓する事もできている。

このように膨大なデータを整理・統合する事で新たな研究分野や産業の創出が可能である事を実証してきたが、爆発的に増加を続けるゲノム・メタゲノムデータをどのように統合するのかが課題となっている。事実、遺伝研のスーパーコンピュータを用いても系統組成や遺伝子機能組成の計算がデータの増加速度に全く追いつかない状況が続いている。これは世界的に共通の課題であり、MG-RAST や MGnify 等のメタゲノム DB は我々よりもさらに周回遅れの状態である。データの逐次更新や間引き等である程度は対応可能であるが、根本的に解決するにはやはりスーパーコンピュータへの投資が必須となる。我々は、惑星科学およびスーパーコンピュータ開発の専門家と密に連携し、超高速メタゲノム配列相同性検索ツール PZLAST を開発した。今後もあらゆる手を尽くして、MicrobeDB.jp の開発を継続しデータサイエンスの発展を維持し続けたいと考えている。

微生物関連情報は極めて多様であるため、微生物統合 DB の開発は与えられた研究期間内で完結するものではない。しかし、我々が開発した、基盤的な技術やアプリケーション群、コンセプトは汎用性の高いものであり、さらなる多様かつ大量のデータの統合化に向けて、将来にわたって発展させていくことが可能である。2021 年に公開されたタンパク質立体構造予測システム AlphaFold2 は、圧倒的な予測精度で研究の在り方そのものを変えようとしているが、実は予測時にメタゲノム配列情報を利用している。このような新たなサービスの発展のためにも、基盤となるデータの整備および統合化が必須であろう。

微生物にとって宿主となる動物や植物に関するホログノミックなデータや、プロテオームやメタボロームのデータについても今後情報爆発を起こす可能性があるが、MicrobeDB.jp ではそれらデータも統合化し発展していけるよう、すでに開発に着手している。このような観点から将来を展望すれば、微生物データのみならず植物や動物、タンパクやメタボロームなど他の基盤 DB とのさらなる統合化は、科学や産業の分け隔てなく社会の発展にとって極めて重要である事は容易に理解できる。発展著しいゲノム合成分野においても欧米や中国に遅れをとっており追いつくのは容易ではないが、基盤となる高品質な統合 DB は基礎から応用に至るまで必要不可欠であり、常に先を見据えつつ腰を落ち着けて開発を継続する事が重要と考えている。

## §6. 自己評価

### 【目標の達成状況】

目標はすべて達成できている。

### 【実施内容に対する研究代表者自身の評価】

我々は、統合化推進プログラムの中では先陣を切ってセマンティックウェブ技術によるデータ統合化を推進し、DBCLS の強力なバックアップを得ながら、多様な RDF モデル、オントロジー、ボキヤブラリ、Stanza を開発して、フル RDF による MicrobeDB.jp を構築してきた。一方で、DBCLS による技術の高度化に追随する必要が生じたため、RDF モデルを一新するなど、今期開発の DB システムは基盤から再構築する必要があった。この大規模な再開発にも関わらず、目標をすべて達成し、当初の目標を大幅に上回る DB アクセスページ数を獲得したことは高く評価されるべきと考えている。また、当初の予想通り、国内においてもマイクロバイオーム分野が急速に発展しており、先を見通した開発が実施できている事を大いに評価している。

### 【本プログラムの趣旨に対して、本研究開発課題がどのように貢献したか】

統合化推進プログラム内において、我々は、微生物データの RDF 化だけでなく、生命科学全体への波及効果やデータ統合による相乗効果を考え、植物、メタボローム等の他の統合化推進プログラムに積極的な技術協力をを行い、セマンティック Web 化の先駆けとして様々なデータの RDF 化を牽引した。

本 DB が対象とするのは、主に微生物および微生物叢に関する研究分野である。本分野では、2016 年 3 月に JST CRDS 戦略プロポーザル「微生物叢(マイクロバイオーム)研究の統合的推進～生命、健康・医療の新展開～」が発表後、文科省から 2016 年度研究開発目標として「宿主と微生物叢間のクロストーク・共生の解明と健康・医療への応用」が発表され、ヒトマイクロバイオーム研究が強く推進されはじめた。これを受けて、2016 年度以降 AMED では革新的先端研究開発支援事業や次世代がん医療創生研究事業、難治性疾患実用化研究事業、次世代治療・診断実現のための創薬基盤技術開発事業などで、ヒトマイクロバイオーム関連研究が開始された。さらに、PM DA 科学委員会ではマイクロバイオーム専門部会が立ち上がり、マイクロバイオーム創薬が現実化しようとしている。ヒトマイクロバイオーム研究以外においては、2014 年度から JST ACCEL「共生ネットワークの分子基盤とその応用展開」において、菌根菌と宿主植物との関係性を明らかにする研究が開始されるとともに、JST CRDS「研究開発の俯瞰報告書 研究開発の新しい動向(2016 年)」にて、土壌・環境マイクロバイオームの評価・制御基盤技術が、今後大きく注目すべきテーマとして挙げられ、宿主とマイクロバイオームの相互作用、動態、共進化などをゲノムレベルで解析するホロゲノム研究も重点化されている。さらに、2018 年度には、戦略的イノベーションプログラム SIP 第 2 期が開始され、「スマートバイオ産業・農業基盤技術」課題においては、産官学連携でのオープンイノベーション型の取り組みによる、農業からヒト健康に至るまでの情報統合を目指している。これらいずれのプロジェクトにおいても MicrobeDB.jp が少なからず活用されており、統合化推進プログラムを構成する研究グループとしてその役割を果たしていると考えている。

## §7. 外部発表等

### 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)	36件
未発行論文	国内(和文)	0件
	国際(欧文)	1件

#### (2) 論文詳細情報

1. Yohei Shimura, Yuu Hirose, Naomi Misawa, Sachiko Wakazuki, Takatomo Fujisawa, Yasukazu Nakamura, Yu Kanasaki, Haruyo Yamaguchi, Masanobu Kawachi, "Complete Genome Sequence of a Coastal Cyanobacterium, *Synechococcus* sp. Strain NIES-970", *Genome Announc.*, vol. 5 No. 14, e00139-17, 2017 (DOI: 10.1128/genomeA.00021-18).
2. Yasuhiro Tanizawa, Takatomo Fujisawa, Yasukazu Nakamura, "DFAST: a flexible prokaryotic genome annotation pipeline for faster genome publication", *Bioinformatics*, vol. 34, Issue 6, pp. 1037-1039, 2018 (DOI: 10.1093/bioinformatics/btx713).
3. Yu Kanasaki, Masaki Hirose, Yuu Hirse, Takatomo Fujisawa, Yasukazu Nakamura, Satoru Watanabe, Shigeru Matsunaga, Hiroko Uchida, Akio Murakami, "Draft Genome Sequence of the Nitrogen-Fixing and Hormogonia-Inducing Cyanobacterium *Nostoc cyadae* Strain WK-1, Isolated from the Coralloid Roots of *Cycas revoluta*", *Genome Announc.*, vol. 6, No. 7, e00021-18, 2018 (DOI: 10.1128/genomeA.00021-18).
4. Hiroshi Mori, Takayuki Maruyama, Masahiro Yano, Takuji Yamada, Ken Kurokawa, VITCOMIC2: visualization tool for the phylogenetic composition of microbial communities based on 16S rRNA gene amplicons and metagenomic shotgun sequencing, *BMC Syst Biol.*, vol. 12, 30, 2018 (DOI: 10.1186/s12918-018-0545-2).
5. 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).  
概要: MicrobeDB.jp の大量のメタ 16S サンプル由来の系統組成データから、機械学習の手法の一つであるトピックモデルを用いて系統組成および各サンプルに紐づくメタデータのパターンをもとに、世界中の微生物群集の組成のパターンを俯瞰できる Web アプリケーション LEA を開発した。
6. Youssef Darzi, Ivica Letunic, Peer Bork, Takuji Yamada, "iPath3. 0: interactive pathways explorer v3", *Nucleic acids research*, 46, W1, W510-W513, 2018 (DOI: 10.1093/nar/gky299).
7. Hikaru Watanabe, Issei Nakamura, Sayaka Mizutani, Yumiko Kurokawa, Hiroshi Mori, Ken Kurokawa, Takuji Yamada, "Minor taxa in human skin microbiome contribute to the personal identification", *PLoS one*, 13, 7, e0199947, 2018 (DOI: 10.1371/journal.pone.0199947).
8. Atsushi Ueda, Atsushi Kobayashi, Sayaka Tsuchida, Takuji Yamada, Koichi Murata, Hiroshi Nakamura, Kazunari Ushida, "Cecal microbiome analyses on wild Japanese roc

- k ptarmigans (*Lagopus muta japonica*) reveals high level of coexistence of lactic acid bacteria and lactate-utilizing bacteria", *Microorganisms*, 6, 3, 77, 2018 (DOI: 10.3390/microorganisms6030077).
9. Naofumi Yoshida, Takuo Emoto, Tomoya Yamashita, Hikaru Watanabe, Tomohiro Hayashi, Tokiko Tabata, Namiko Hoshi, Naoya Hatano, Genki Ozawa, Naoto Sasaki, Taiji Mizoguchi, Hilman Zulkifli Amin, Yushi Hirota, Wataru Ogawa, Takuji Yamada, Ken-ichi Hirata, "*Bacteroides vulgatus* and *Bacteroides dorei* reduce gut microbial lipopolysaccharide production and inhibit atherosclerosis", *Circulation*, 138, 22, 2486-2498, 2018 (DOI: 10.1161/CIRCULATIONAHA.118.033714).
  10. Tomohiro Hayashi, Tomoya Yamashita, Hikaru Watanabe, Kenjiro Kami, Naofumi Yoshida, Tokiko Tabata, Takuo Emoto, Naoto Sasaki, Taiji Mizoguchi, Yasuhiro Irino, Ryuji Toh, Masakazu Shinohara, Yuko Okada, Wataru Ogawa, Takuji Yamada, Ken-ichi Hirata, "Gut microbiome and plasma microbiome-related metabolites in patients with decompensated and compensated heart failure", *Circulation Journal*, 83, 1, 182-192, 2018 (DOI: 10.1253/circj.CJ-18-0468).
  11. Sachiko Naito, Noriko Takeuchi, Misako Ohkusu, Azusa Takahashi-Nakaguchi, Hiroki Takahashi, Naoko Imuta, Junichiro Nishi, Keigo Shibayama, Mayumi Matsuoka, Yuko Sasaki, Naruhiko Ishiwada, "Clinical and bacteriologic analysis of nontypeable *Haemophilus influenzae* strains isolated from children with invasive diseases in Japan from 2008 to 2015", *J Clin Microbiol*, 56, 7, e00141-18, 2018 (DOI: 10.1128/JCM.00141-18).
  12. 高橋 (中口) 梓, 萩原大祐, 高橋弘喜, 酒井香奈江, 豊留孝仁, 渡辺哲, 亀井克彦, 五ノ井透, "次世代シーケンサーを活用したオミクス解析による *Aspergillus fumigatus* 病原因子探索", *Med Mycol J*, 59, 2, J35-J40, 2018 (DOI: 10.3314/mmj.18.005).
  13. Kenzo Matsuo, Akira Haku, Beibei Bi, Hiroki Takahashi, Nobuhiko Kamada, Takashi Yaguchi, Shinobu Saijo, Mitsutoshi Yoneyama, Yoshiyuki Goto, "Fecal microbiota transplantation prevents *Candida albicans* from colonizing the gastrointestinal tract", *Microbiol Immunol*, 63(5):155-163, 2019 (DOI: 10.1111/1348-0421).
  14. Masaaki Shimizu, Yoko Kusuya, Yikelamu Alimu, Cai Bian, Hiroki Takahashi, Takashi Yaguchi, "Draft Genome Sequence of *Aspergillus awamori* IFM 58123NT", *Microbiol Resour Announc*, 8(4): e01453-18, 2019 (DOI: 10.1128/MRA.01453-18).
  15. Takahito Toyotome, Saho Hamada, Satoe Yamaguchi, Hiroki Takahashi, Daisuke Kondoh, Masahiko Takino, Yu Kanesaki, Katsuhiko Kamei, "Comparative genome analysis of *Aspergillus flavus* clinically isolated in Japan", *DNA Res*, 6, 1, 95-103, 2019 (DOI: 10.1093/dnares/dsy041).
  16. Mohammad Vahed, Jun-ichi Ishihara, Hiroki Takahashi, "Dipartite: A tool for detecting bipartite motifs by considering base interdependencies", *PLoS One*, 14(8):e0220207, 2019 (DOI: 10.1371/journal.pone.0220207).
  17. Laure Nicolas Annick Ries, Jacob L. Steenwyk, Patrícia Alves de Castro, Pollyne Borema Almeida de Lima, Fausto Almeida, Leandro José de Assis, Adriana Oliveira Manfiolli, Azusa Takahashi-Nakaguchi, Yoko Kusuya, Daisuke Hagiwara, Hiroki Takahashi, Xi Wang, Joshua J. Obar, Antonis Rokas, Gustavo H. Goldman, "Nutritional Heterogeneity Among *Aspergillus fumigatus* Strains Has Consequences for Virulence in a Strain- and Host-Dependent Manner", *Front Microbiol*, 10:854, 2019 (DOI: 10.3389/fmicb.2019.00854).
  18. Toshiaki Katayama, Shuichi Kawashima, Shinobu Okamoto, Yuki Moriya, Hirokazu C

- hiba, Yuki Naito, Takatomo Fujisawa, Hiroshi Mori, Toshihisa Takagi, "TogoGenome/TogoStanza: modularized Semantic Web genome database", Database, vol. 2019, 2019 (DOI: 10.1093/database/bay132).
19. Ikuo Uchiyama, Motohiro Mihara, Hiroyo Nishide, Hirokazu Chiba, Masaki Kato, "MB GD update 2018: microbial genome database based on hierarchical orthology relations covering closely related and distantly related comparisons", Nucleic Acids Res., 47, D382-D389, 2019 (DOI: 10.1093/nar/gky1054).
  20. Toshiaki Katayama, Shuichi Kawashima, Gos Micklem, Shin Kawano, Jin-Dong Kim, Simon Kocbek, Shinobu Okamoto, Yue Wang, Hongyan Wu, Atsuko Yamaguchi, Yasunori Yamamoto, Erick Antezana, Kiyoko F. Aoki-Kinoshita, Kazuharu Arakawa, Masaki Banno, Joachim Baran, Jerven T. Bolleman, Raoul J.P. Bonnal, Hidemasa Bono, Jesualdo T. Fernández-Breis, Robert Buels, Matthew P. Campbell, Hirokazu Chiba, Peter J. A. Cock, Kevin B. Cohen, Michel Dumontier, Takatomo Fujisawa, Toyofumi Fujiwara, Leyla Garcia, Pascale Gaudet, Emi Hattori, Robert Hoehndorf, Kotone Itaya, Maori Ito, Daniel Jamieson, Simon Jupp, Nick Juty, Alex Kalderimis, Fumihiko Kato, Hideya Kawaji, Takeshi Kawashima, Akira R. Kinjo, Yusuke Komiyama, Masaaki Kotera, Tatsuya Kushida, James Malone, Masaaki Matsubara, Satoshi Mizuno, Sayaka Mizutani, Hiroshi Mori, Yuki Moriya, Katsuhiko Murakami, Takeru Nakazato, Hiroyo Nishide, Yosuke Nishimura, Soichi Ogishima, Tazro Ohta, Shujiro Okuda, Hiromasa Ono, Yasset Perez-Riverol, Daisuke Shinmachi, Andrea Splendiani, Francesco Strozzi, Shinya Suzuki, Junichi Takehara, Mark Thompson, Toshiaki Tokimatsu, Ikuo Uchiyama, Karin Verpoor, Mark D. Wilkinson, Sarala Wimalaratne, Issaku Yamada, Nozomi Yamamoto, Masayuki Yarimizu, Shoko Kawamoto, Toshihisa Takagi, "BioHackathon series in 2013 and 2014: improvements of semantic interoperability in life science data and services", F1000Research, 8:1677, 2019 (DOI: 10.12688/f1000research.18238.1).
  21. Rutger A Vos, Toshiaki Katayama, Hiroyuki Mishima, Shin Kawano, Shuichi Kawashima, Jin-Dong Kim, Yuki Moriya, Toshiaki Tokimatsu, Atsuko Yamaguchi, Yasunori Yamamoto, Hongyan Wu, Peter Amstutz, Erick Antezana, Nobuyuki P Aoki, Kazuharu Arakawa, Jerven T Bolleman, Evan Bolton, Raoul J P Bonnal, Hidemasa Bono, Kees Burger, Hirokazu Chiba, Kevin B Cohen, Eric W Deutsch, Jesualdo T Fernández-Breis, Gang Fu, Takatomo Fujisawa, Atsushi Fukushima, Alexander García, Naohisa Goto, Tudor Groza, Colin Hercus, Robert Hoehndorf, Kotone Itaya, Nick Juty, Takeshi Kawashima, Jee-Hyub Kim, Akira R Kinjo, Masaaki Kotera, Kouji Kozaki, Sadahiro Kumagai, Tatsuya Kushida, Thomas Lütke, Masaaki Matsubara, Joe Miyamoto, Attayeb Mohsen, Hiroshi Mori, Yuki Naito, Takeru Nakazato, Jeremy Nguyen-Xuan, Kozo Nishida, Naoki Nishida, Hiroyo Nishide, Soichi Ogishima, Tazro Ohta, Shujiro Okuda, Benedict Paten, Jean-Luc Perret, Philip Prathipati, Pjotr Prins, Núria Queralt-Rosinach, Daisuke Shinmachi, Shinya Suzuki, Tsuyosi Tabata, Terue Takatsuki, Kieron Taylor, Mark Thompson, Ikuo Uchiyama, Bruno Vieira, Chih-Hsuan Wei, Mark Wilkinson, Issaku Yamada, Ryota Yamanaka, Kazutoshi Yoshitake, Akiyasu C Yoshizawa, Michel Dumontier, Kenjiro Kosaki, Toshihisa Takagi, "BioHackathon 2015: Semantics of data for life sciences and reproducible research", F1000Research, 9:136, 2020 (DOI: 10.12688/f1000research.18236.1).
  22. Yachida S, Mizutani S, Shiroma H, Shiba S, Nakajima T, Sakamoto T, Watanabe H, Masuda K, Nishimoto Y, Kubo M, Hosoda F, Rokutan H, Matsumoto M, Takamaru H, Yamada M, Matsuda T, Iwasaki M, Yamaji T, Yachida T, Soga T, Kurokawa K, Toyo

- da A, Ogura Y, Hayashi T, Hatakeyama M, Nakagama H, Saito Y, Fukuda S, Shibata T, Yamada T, “Metagenomic and metabolomic analyses reveal distinct stage-specific phenotypes of the gut microbiota in colorectal cancer”, *Nature Medicine*, 25, 968–976, 2019 (DOI: 10.1136/gutjnl-2019-319188).
23. Thomas AM, Manghi P, Asnicar F, Pasolli E, Armanini F, Zolfo M, Beghini F, Manara S, Karcher N, Pozzi C, Gandini S, Serrano D, Tarallo S, Francavilla A, Gallo G, Trompetto M, Ferrero G, Mizutani S, Shiroma H, Shiba S, Shibata T, Yachida S, Yamada T, Wirbel J, Schrotz-King P, Ulrich CM, Brenner H, Arumugam M, Bork P, Zeller G, Cordero F, Dias-Neto E, Setubal JC, Tett A, Pardini B, Rescigno M, Waldron L, Naccarati A, Segata N, “Metagenomic analysis of colorectal cancer datasets identifies cross-cohort microbial diagnostic signatures and a link with choline degradation”, *Nature Medicine*, 25, 667-678. 2019 (DOI: 10.1038/s41591-019-0663-4).
  24. Wirbel J, Pyl PT, Kartal E, Zych K, Kashani A, Milanese A, Fleck JS, Voigt AY, Pallaja A, Ponnudurai R, Sunagawa S, Coelho LP, Schrotz-King P, Vogtmann E, Habermann N, Niméus E, Thomas AM, Manghi P, Gandini S, Serrano D, Mizutani S, Shiroma H, Shiba S, Shibata T, Yachida S, Yamada T, Waldron L, Naccarati A, Segata N, Sinha R, Ulrich CM, Brenner H, Arumugam M, Bork P, Zeller G, “Meta-analysis of fecal metagenomes reveals global microbial signatures that are specific for colorectal cancer”, *Nature Medicine*, 25, 679-689, 2019 (DOI: 10.1038/s41591-019-0406-6).
  25. Erawijantari PP, Mizutani S, Shiroma H, Shiba S, Nakajima T, Sakamoto T, Saito Y, Fukuda S, Yachida S, Yamada T, “Influence of gastrectomy for gastric cancer treatment on faecal microbiome and metabolome profiles” *Gut*, 69(8):1404-1415, 2020 (DOI: 10.1136/gutjnl-2019-319188).
  26. Laure Nicolas Annick Ries, Jacob L Steenwyk, Patrícia Alves de Castro, Pollyne Borema Almeida de Lima, Fausto Almeida, Leandro José de Assis, Adriana Oliveira Manfiolli, Azusa Takahashi-Nakaguchi, Yoko Kusuya, Daisuke Hagiwara, Hiroki Takahashi, Xi Wang, Joshua J Obar, Antonis Rokas, Gustavo H Goldman, “Nutritional Heterogeneity Among *Aspergillus fumigatus* Strains Has Consequences for Virulence in a Strain- and Host-Dependent Manner”, *Front Microbiol*, 10, 854, 2019 (DOI: 10.3389/fmicb.2019.00854).
  27. Mohammad Vahed, Jun-Ichi Ishihara, Hiroki Takahashi. “Dipartite: A tool for detecting bipartite motifs by considering base interdependencies”, *PLoS One*, 14(8), e0220207, 2019 (DOI: 10.1371/journal.pone.0220207).
  28. Kenzo Matsuo, Akira Haku, Beibei Bi, Hiroki Takahashi, Nobuhiko Kamada, Takashi Yaguchi, Shinobu Saijo, Mitsutoshi Yoneyama, Yoshiyuki Goto. “Fecal microbiota transplantation prevents *Candida albicans* from colonizing the gastrointestinal tract”, *Microbiol Immunol*, 63(5), 155-163, 2019 (DOI: 10.1111/1348-0421.12680).
  29. Yuumi Nakamura, Hiroki Takahashi, Akiko Takaya, Yuzaburo Inoue, Yuki Katayama, Yoko Kusuya, Tatsuma Shoji, Sanami Takada, Seitaro Nakagawa, Rena Oguma, Nobuko Saito, Naoko Ozawa, Taiji Nakano, Fumiya Yamaide, Eishika Dissanayake, Shuichi Suzuki, Amer Villaruz, Saranyaraajan Varadarajan, Masanori Matsumoto, Tomoko Kobayashi, Michihiro Kono, Yasunori Sato, Masashi Akiyama, Michael Otto, Hiroyuki Matsue, Gabriel Núñez and Naoki Shimojo, “*Staphylococcus* Agr virulence is critical for epidermal colonization and associates with atopic dermatitis development”, *Sci Transl Med*, 12(551), eaay4068, 2020 (DOI: 10.1126/scitranslmed.aay4068).
  30. Isato Yoshioka, Hiroki Takahashi, Yoko Kusuya, Takashi Yaguchi and Kohtaro Kirimu

ra, "Draft Genome Sequence of *Aspergillus tubingensis* WU-2223L, a Citric Acid-Producing Filamentous Fungus Belonging to *Aspergillus* Section Nigri", *Microbiol Resour Announc*, 9(33), e00702-20, 2020 (DOI: 10.1128/MRA.00702-20).

31. Masahiko Imashimizu, Yuji Tokunaga, Ariel Afek, Hiroki Takahashi, Nobuo Shimamoto and David B Lukatsky, "Control of Transcription Initiation by Biased Thermal Fluctuations on Repetitive Genomic Sequences", *Biomolecules*, 10(9), 1299, 2020 (DOI: 10.3390/biom10091299).
32. Hiroshi Nishiwaki, Mikako Ito, Tomohiro Ishida, Tomonari Hamaguchi, Tetsuya Maeda, Kenichi Kashihara, Yoshio Tsuboi, Jun Ueyama, Teppei Shimamura, Hiroshi Mori, Ken Kurokawa, Masahisa Katsuno, Masaaki Hirayama, Kinji Ohno, "Meta-analysis of gut dysbiosis in Parkinson's disease", *Movement Disorders*, 35, 1626-1635, 2020 (DOI: 10.1002/mds.28119).
33. Amit Rai, Hideki Hirakawa, Ryo Nakabayashi, Shinji Kikuchi, Koki Hayashi, Megha Rai, Hiroshi Tsugawa, Taiki Nakaya, Tetsuya Mori, Hideki Nagasaki, Runa Fukushi, Yoko Kusuya, Hiroki Takahashi, Hiroshi Uchiyama, Atsushi Toyoda, Shoko Hikosaka, Eiji Goto, Kazuki Saito and Mami Yamazaki, "Chromosome-level genome assembly of *Ophiorrhiza pumila* reveals the evolution of camptothecin biosynthesis", *Nat Commun*, 12(1), 405, 2021 (DOI: 10.1038/s41467-020-20508-2).
34. Tatsuma Shoji, Akiko Takaya, Yoko Kusuya, Hiroki Takahashi and Hiroto Kawashima, "Ribosome Profiling in *Streptococcus pneumoniae* reveals the Role of Methylation of 23S rRNA Nucleotide G748 on Ribosome Stalling", *J Genet Genomic Sci*, 6, 024, 2021 (DOI: 10.24966/GGS-2485/100024).
35. Naoki Tsukuda, Kana Yahagi, Taeko Hara, Yohei Watanabe, Hoshitaka Matsumoto, Hiroshi Mori, Koichi Higashi, Hirokazu Tsuji, Satoshi Matsumoto, Ken Kurokawa and Takahiro Matsuki, "Key bacterial taxa and metabolic pathways affecting gut short-chain fatty acid profiles in early life", *ISME J*, 15(9):2574-2590, 2021 (DOI: 10.1038/s41396-021-00937-7).
36. Takashi Okubo, Atsushi Toyoda, Kohei Fukuhara, Ikuo Uchiyama, Yuhki Harigaya, Megumi Kuroiwa, Takuma Suzuki, Yuka Murakami, Yuichi Suwa, Hideto Takami, "The physiological potential of anammox bacteria as revealed by their core genome structure", *DNA Res*, vol. 28, issue 1, 2021 (DOI: 10.1093/dnares/dsaa028).
37. 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*, in press (DOI: 10.1093/bioinformatics/btab492).  
概要: ユーザの興味があるアミノ酸配列を、MicrobeDB.jp の大量のメタゲノムのアミノ酸配列データを参照配列 DB として超高速に配列相同性検索を行い、結果を MEO による環境カテゴリや地理的な情報等で整理して表示可能な、膨大なメタゲノム配列データに対する配列相同性検索 Web アプリケーション PZLAST を開発した。
38. Ryota Sugimoto, Luca Nishimura, Phuong Thanh Nguyen, Jumpei Ito, Nicholas F. Parrish, Hiroshi Mori, Ken Kurokawa, Hirofumi Nakaoka, Ituro Inoue, "Comprehensive discovery of CRISPR-targeted terminally redundant sequences in the human gut metagenome: viruses, plasmids, and more", *PLoS Comput Biol*, accepted.

## 2. その他の著作物(総説、書籍など)

1. 高橋弘喜、組成解析(真菌)、化学療法の領域、Vol.33、No.7、2017年
2. Ikuo Uchiyama, Ortholog identification and comparative analysis of microbial genomes using MBGD and RECOG. In “Protein function prediction: Methods and protocols”, (D. Kihara ed.), pp.147-168, Humana Press, 2017.
3. 森宙史、細菌のゲノム進化と群集ダイナミクス、生体の科学、pp155-159、Vol. 68、No. 2、2017年
4. 高橋弘喜、組成解析(真菌)、化学療法の領域、pp105-109、Vol.33、No.7、2017年
5. 森宙史、黒川顕、メタゲノム解析・ホログゲノム解析、化学療法の領域、pp111-117、Vol. 33、No. 7、2017年
6. 東光一、黒川顕、メタゲノム解析と食品開発への応用について、pp22-27、明日の食品産業 7・8月号、2017年
7. 黒川顕、森宙史、東光一、9章 メタゲノム解析、よくわかるバイオインフォマティクス入門、講談社、pp123-pp136、2018年
8. 村上匠、森宙史、黒川顕、メタゲノムデータの情報解析とデータベース、腸内細菌叢 健康と疾患を制御するエコシステム、実験医学増刊、羊土社、37、pp23-28、2019年
9. 森宙史、黒川顕、ヒトマイクロバイオーム統合データベースの構築、ヒトマイクロバイオーム Vol. 2、NTS 出版、pp111-119、2020年。
10. 森宙史、メタゲノム解析における機械学習の手法、機械学習を生命科学に使う！、実験医学増刊、羊土社、38、pp88-94、2020年。

## 3. 国際学会発表及び主要な国内学会発表

### (1) 概要

種別	国内外	件数
招待講演	国内	29 件
	国際	11 件
口頭発表	国内	22 件
	国際	8 件
ポスター発表	国内	20 件
	国際	8 件

### (2) 招待講演

〈国内〉

1. 千葉啓和、ドメインレベルのオーソログ解析でみる融合遺伝子の進化、ラン藻ゲノム交流会、東京大学駒場キャンパス、2017年6月24日
2. 山田拓司、ヒト腸内細菌と大腸がん発病モデル、日本マイクロバイオームコンソーシアム(JMBC)設立記念シンポジウム、THE GRAND HALL 品川、2017年7月28日
3. 黒川顕、メタゲノム研究の技術的進展と情報の取扱いについて、ヒトゲノム研究倫理を考える会、グランフロント大阪、2017年10月9日



4. 黒川顕、微生物ビッグデータチャレンジ: データサイエンスを加速させる微生物統合データベース "MicrobeDB.jp"、BioJapan2017 次世代微生物革命セミナー、パシフィコ横浜、2017 年 10 月 12 日
5. 黒川顕、データサイエンスを加速させる微生物データの統合化、CBI 学会講演会「創薬・ヘルスケア応用に向けた腸内細菌の最新動向」、キャンパスイノベーションセンター東京、2017 年 11 月 10 日
6. 森宙史、微生物研究開発における微生物統合データベース MicrobeDB.jp の利用法、千葉県バイオ・ライフサイエンス・ネットワーク会議 平成 29 年度講演会「バイオ産業の活性化にむけた公的データベースの活用法、ペリエホール、2017 年 12 月 22 日
7. Hiroshi Mori、Developing tools and database for microbial community analysis、マイクロバイオーム研究開発の最前線、日本橋ライフサイエンスハブ、2018 年 1 月 25 日
8. 黒川顕、微生物統合データベース「MicrobeDB.jp」、第 91 回日本細菌学会総会シンポジウム、福岡国際会議場、2018 年 3 月 27 日
9. 高橋弘喜、進歩が著しいマイクロバイオームおよび微生物ゲノム研究法、第 117 回日本皮膚科学会総会、広島市、2018 年 6 月 2 日
10. 中村保一、塩基配列データの爆発的増加と DDBJ をとりまく現状、遺伝研研究会・生命情報データの拡大と分子進化研究の進展、三島、2018 年 8 月 4 日
11. 内山郁夫、非コア遺伝子に着目した水平移動遺伝子の検出、遺伝研研究集会「DNA 水平伝搬」、遺伝研、2018 年 8 月 20-21 日
12. 内山郁夫、オーソログを探求する: Quest for Orthologs 会議について、日本進化学会第 20 回大会、東京大学駒場キャンパス、2018 年 8 月 22-25 日
13. 千葉啓和、オーソログデータベースに基づくドメイン融合解析、日本進化学会第 20 回大会、東京大学駒場キャンパス、2018 年 8 月 22-25 日
14. 中村保一、ゲノム解析手法のトレンドと DDBJ・遺伝研スーパーコンピュータ、第二回木村資生記念進化学セミナー、八王子、2018 年 9 月 3 日
15. 黒川顕、微生物統合データベース MicrobeDB.jp、JBA 発酵と代謝研究会、東京、2018 年 9 月 3 日
16. 黒川顕、メニーコアプロセッサが可能とするリアルタイムメタゲノム解析、アクセラレーション技術発表討論会、油津、2018 年 9 月 7 日
17. 森宙史、Metagenome informatics to understand microbial communities、イルミナマイクロバイオロジーワークショップ 2018『メタバーコーディングからショットガンメタゲノムによる全ゲノムへの展望』、東京コンファレンスセンター品川、2018 年 9 月 18 日
18. 高橋弘喜、病原真菌 *Aspergillus fumigatus* の環境応答能の解析、第 101 回日本細菌学会関東支部総会、東京都、2018 年 11 月 1 日
19. 黒川顕、マイクロバイオームと未来社会、第 2 回経団連・産業連携セミナー、東京、2019 年 2 月 25 日
20. 中村保一、INSDC の一員としての DDBJ 事業の紹介と塩基配列決定の国際動向と INSDC への塩基配列データの登録と管理の実際、製品評価技術基盤機構 NITE バイオテクノロジーセンターセミナー、木更津、2019 年 2 月 22 日
21. 黒川顕、微生物統合データベース MicrobeDB.jp、第 11 回日本ゲノム微生物学会、東京、2019 年 3 月 8 日
22. 高橋弘喜、病原真菌アスペルギルスフミガタスの表現型解析および遺伝型解析、感染症診断

と治療におけるゲノム解析、伊勢原市、2019年3月15日

23. 黒川顕、微生物統合データベース MicrobeDB.jp、日本農芸化学会、東京、2019年3月26日
24. 高橋弘喜、病原真菌 *Aspergillus fumigatus* の環境応答能の理解に向けて、第92回日本細菌学会総会、札幌市、2019年4月24日
25. 黒川顕、マイクロバイームデータの可能性、第32期 CAMM フォーラム本例会、東京都、2019年5月10日
26. 黒川顕、マイクロバイームデータの可能性、トヨタ自動車、豊田市、2019年6月7日
27. 森宙史、メタゲノムインフォマティクスと微生物統合データベース、産総研セミナー、産業総合技術研究所つくばセンター、2019年11月1日。
28. 高橋弘喜、次世代シーケンサー (NGS) の進展と微生物研究への応用、第56回小児アレルギー学会、千葉市、2019年11月2日
29. 高橋弘喜、病原真菌 *Aspergillus fumigatus* の環境応答能の数理モデル化による理解の試み、第93回日本細菌学会総会、ウインクあいち、2020年2月19日

〈国際〉

1. 中村保一、How to make bio-DB's and services sustainable?、日中韓バイオインフォマティクスシンポジウム、ソウル、2017/06/23.
2. 中村保一、DDBJ/INSDC/DFAST、GCM 2.0 workshop、北京、2018/6/26
3. Ken Kurokawa, Integrated database for microbe "MicrobeDB.jp", JSME2018&10thASM E, Okinawa, 2018/7/13.
4. 中村保一、Genome analyses using NGS for next-generation model species、The 16th CJK (日中韓) Bioinformatics Symposium、葉山、2018/8/30.
5. 中村保一、DDBJ / the INSDC is ready for GCM 2.0、8th WDCM / GCM 2.0 meeting、北京、2018/11/21.
6. 中村保一、Introduction of the INSDC, data submission & genome annotation、GCM 2.0 training course、北京、2018/11/26.
7. Ken Kurokawa, Hiroshi Mori, Integrated Database for Microbes, "MicrobeDB.jp", Laser Solutions for Space and the Earth 2019 (LSSE2019), Yokohama, April 23-26, 2019.
8. Metagenomic and metabolomic analyses reveal dynamic shifts in gut microbiota along the adenoma-carcinoma sequence in colorectal cancer, 山田拓司, OIST Workshop "A World of Microbiota", 2019/7/4.
9. Uchiyama, I. Ortholog relationships among pan-genomes: the Microbial Genome Database (MBGD) for exploring microbial diversity, The 67th NIBB Conference "Quest for Orthologs", National Institute for Basic Biology, Okazaki, Jul. 31-Aug.2, 2019.
10. Hiroshi Mori, MicrobeDB.jp: integration of microbial data from microbiome point of view, ACM17 Online, September 9, 2020.
11. Metagenomic and metabolomic analyses in colorectal cancer, 山田拓司, Microbiome Movement - Drug Development, Singapore, 2019/9/17.

### (3) 口頭講演

〈国内〉

1. 森宙史、微生物群集の系統組成推定・描画ツール VITCOMIC2 の開発と応用、NGS 現場の会第五回研究会、仙台国際センター展示棟、2017 年 5 月 23 日
2. 山田拓司、大腸がん発病に伴うヒト腸内細菌のマルチオミクス解析、第 17 回日本抗加齢医学会総会、東京国際フォーラム、2017 年 6 月 2 日
3. 森宙史、東光一、黒川顕、微生物群集ビッグデータの統合と機械学習を利用した群集予測・診断システム、BIOTech2017 アカデミックフォーラム、東京ビッグサイト、2017 年 6 月 30 日
4. 黒川顕、データサイエンスを加速させる微生物データの統合化、環境微生物系学会合同大会 2017、東北大学川内キャンパス、2017 年 8 月 31 日
5. 山田拓司、ヒト腸内細菌データベースとその利用、第 69 回日本生物工学会大会。早稲田大学西早稲田キャンパス、2017 年 9 月 11 日
6. 山田拓司、大腸がん発病に伴うヒト腸内細菌のマルチオミクス解析、第 42 回日本医用マスペクトル学会、一橋講堂、2017 年 9 月 14 日
7. 山田拓司、Human gut microbiome associated with colorectal cancer、第 76 回日本癌学会学術総会、パシフィコ横浜、2017 年 9 月 28 日
8. 中村保一、MicrobeDB.jp、日本分子生物学会年会 NBDC ConBio フォーラム「生命科学のデータベース活用法」、神戸、2017 年 12 月 9 日
9. 山田拓司、大腸がん発病に伴うヒト腸内環境のマルチオミクス解析、第45回日本毒性学会学術年会、大阪国際会議場、2018 年 7 月 19 日
10. 森宙史、黒川顕、微生物統合データベース MicrobeDB.jp の活用法、ワークショップ「生命科学のデータベース活用法 2018」、第 41 回日本分子生物学会年会、パシフィコ横浜、2018 年 11 月 28 日
11. 内山郁夫、微生物比較ゲノムデータベースを用いた水平移動遺伝子動態の網羅的解析、遺伝研研究集会「環境中の DNA 循環」、国立遺伝学研究所、2019 年 8 月 19-20 日
12. 森宙史、メタゲノムインフォマティクスとヒトマイクロバイオーム統合データベース、第 92 回日本生化学会大会シンポジウム「共生微生物叢研究の最前線」、パシフィコ横浜、2019 年 9 月 20 日
13. 森宙史、東光一、黒川顕、微生物統合データベース MicrobeDB.jp と関連するツール、ワークショップ「微生物—植物相互作用を解明して育種に繋げるには?」、トーゴの日シンポジウム 2019、日本科学未来館、2019 年 10 月 5 日
14. Cai Bian, Yikelamu Alimu, Hiroki Takahashi, Takashi Yaguchi, Characterization and proposal of two new species in *Aspergillus* section Nigri、第 63 回日本医真菌学会総会・学術集会、千葉市、2019 年 10 月 11 日
15. 森宙史、中川善一、山田拓司、Human Microbiome Associated Disease Ontology の構築と比較メタゲノム解析、第 13 回日本ゲノム微生物学会若手の会研究会、山喜旅館、2019 年 10 月 27 日。
16. 森宙史、微生物の生息環境のアノテーションとヒトコホートのメタゲノム研究、Annotathon 2019、東京大学柏の葉キャンパス駅前サテライト、2019 年 11 月 13 日。
17. 森宙史、MicrobeDB.jp version 3 の活用法、第 42 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 2019」、福岡国際会議場、2019 年 12 月 3 日。
18. 内山郁夫、微生物比較ゲノムデータベース MBDG に基づく水平移動遺伝子動態の網羅的解析、日本ゲノム微生物学会年会、ウインクあいち、2020 年 3 月 6-8 日
19. 新居鉄平、高橋弘喜、楠屋陽子、渡辺哲、亀井克彦、*Aspergillus fumigatus* のアゾール耐

性に関わる新規因子について, 第 64 回日本医真菌学会総会, オンライン開催, 2020 年 10 月 9 日-10 日

20. 萩原大祐, 老木紗予子, 楠屋陽子, 高橋弘喜, 輸入植物球根から分離されたヒト病原真菌における農薬耐性進化の可能性, 日本農薬学会第 46 回大会, オンライン開催, 2021 年 3 月 8 日-10 日
21. 前川尚輝, 板寺健悟, 石原潤一, 梶谷颯希, 田中大器, 関口哲志, 庄子習一, 石橋正己, 高橋弘喜, 高屋明子, ATP 依存型 Lon プロテアーゼによるパーシスタンス制御機構, オンライン開催, 第 94 回日本細菌学会総会, 2020 年 3 月 23 日-25 日
22. 吉岡育哲, 高橋弘喜, 楠屋陽子, 矢口貴志, 桐村光太郎, クエン酸高生産菌 *Aspergillus tubingensis* WU-2223L のドラフトゲノムの決定, 日本農芸化学会 2021 年度大会, オンライン開催, 2021 年 3 月 18 日-21 日

#### 〈国際〉

1. Ikuo Uchiyama, A new protocol for constructing the ortholog table in Microbial Genome Database for Comparative Analysis, NIG International Symposium 2017, Mishima, May 27-29, 2017.
2. Ikuo Uchiyama, A new protocol for constructing ortholog database for large-scale microbial genome comparison, Quest for Orthologs 5, University of Southern California, Los Angeles, Jun. 8-10, 2017.
3. Hirokazu Chiba, Fernández-Breis, Jesualdo Tomás Fernández-Breis, Ramón García Martínez, Yuki Moriya, Susumu Goto, Ikuo Uchiyama, Development of orthology ontology and its application to orthology meta-search, Quest for Orthologs 5, University of Southern California, Los Angeles, Jun. 8-10, 2017.
4. Hiroshi Mori, MicrobeDB.jp: integrating microbial genome and metagenome data using semantic web technology, BioHackathon 2017 symposium, JST Science Plaza, Japan, Sep. 10, 2017.
5. Hiroshi Mori, Takayuki Maruyama, Masahiro Yano, Takuji Yamada, Ken Kurokawa, VITCOMIC2: Visualization tool for the phylogenetic composition of microbial communities based on 16S rRNA gene amplicons and metagenomic shotgun sequencing, GIW 2017, Hotel Samjung, Seoul, Nov. 1, 2017.
6. Hiroshi Mori, Ken Kurokawa, VITCOMIC2: Visualization and analysis tool for the phylogenetic composition of microbial communities based on 16S rRNA gene sequences, ASM Microbe 2018, Georgia World Congress Center, Atlanta, USA, June 7-12, 2018.
7. Hirokazu Chiba, Ikuo Uchiyama, Unraveling the microbial gene repertoire by sub-gene level orthologous clustering. The 67th NIBB Conference "Quest for Orthologs", National Institute for Basic Biology, Okazaki, Jul. 31-Aug.2, 2019.
8. Hiroshi Mori, Metagenome informatics for microbiome researches, The 43rd Annual Meeting of the Molecular Biology Society of Japan, Online, December 2, 2020.

#### (4) ポスター発表

##### 〈国内〉

1. 内山郁夫, 西出浩世, 比較ゲノム解析ツール RECOG を用いた NGS データの比較解析, 第 5 回 NGS 現場の会, 仙台国際センター, 2017 年 5 月 22 日-24 日

2. 内山郁夫、三原基広、西出浩世、千葉啓和、MBGD 2018:大規模な微生物ゲノムデータのより効果的な比較に向けた改良、トーゴの日シンポジウム 2017、東京大学弥生講堂一条ホール、2017年10月4-5日
3. 森宙史、藤澤貴智、鈴木真也、千葉啓和、東光一、神沼英里、西出浩世、矢口貴志、高橋弘喜、山田拓司、内山郁夫、中村保一、黒川顕、微生物統合データベース MicrobeDB.jp の高度実用化開発、トーゴの日シンポジウム 2017、2017年10月4-5日
4. 内山郁夫、三原基広、西出浩世、千葉啓和、公開ゲノムの全遺伝子レパートリーをカバーする微生物オーソログテーブルの構築、第12回ゲノム微生物学会年会、京都大学桂キャンパス、2018年3月5-7日
5. 藤澤貴智、森宙史、谷澤靖洋、神沼英里、内山郁夫、山田拓司、高橋弘喜、中村保一、黒川顕、MicrobeDB.jp ポータル: 統合微生物データベースのポータルサイト構築、第12回日本ゲノム微生物学会年会、京都市、2018年3月5日-7日
6. 森宙史、黒川顕、微生物群集の系統組成推定ツール VITCOMIC2 の MicrobeDB.jp への応用、第12回日本ゲノム微生物学会年会、京都大学桂キャンパス、2018年3月5-7日
7. 豊留孝仁、高橋弘喜、亀井克彦、日本臨床分離 *Aspergillus flavus* 株と *A. flavus* NRRL3357 株および *A. oryzae* RIB40 株とのゲノムワイド SNP 比較解析、日本マイコキシソ学会第82回学術講演会、帯広市、2018年8月24日
8. 藤澤貴智、森宙史、谷澤靖洋、神沼英里、内山郁夫、山田拓司、高橋弘喜、中村保一、黒川顕、MicrobeDB.jp ポータル: 統合微生物データベースのポータルサイト構築、トーゴの日シンポジウム 2018、2018年10月5日
9. 森宙史、黒川顕、MicrobeDB.jp のメタゲノム解析パイプライン、トーゴの日シンポジウム 2018、日本科学未来館、2018年10月5日
10. 内山郁夫、非コア遺伝子の並び順の保存性に基づくゲノミックアイランドの抽出、第13回日本ゲノム微生物学会年会、首都大学東京南大沢キャンパス、2019年3月6-8日
11. 森宙史、藤澤貴智、西出浩世、矢口貴志、高橋弘喜、山田拓司、内山郁夫、中村保一、黒川顕、微生物統合データベース MicrobeDB.jp の更新とヒトマイクロバイオームデータの統合、第13回日本ゲノム微生物学会年会、八王子市、2019年3月6-8日
12. 内山郁夫、三原基広、西出浩世、千葉啓和、高柳正彦、高見英人、微生物比較ゲノムデータベース MBGD、トーゴの日シンポジウム 2019、日本科学未来館、2019年10月5日
13. 藤澤貴智、森宙史、谷澤靖洋、児玉悠一、内山郁夫、中川善一、山田拓司、高橋弘喜、中村保一、黒川顕、統合微生物データベース MicrobeDB.jp ポータルサイト拡張、トーゴの日シンポジウム 2019、2019年10月5日
14. 森宙史、藤澤貴智、中川善一、山田拓司、黒川顕、MicrobeDB.jp ver. 3 のマイクロバイオームデータの概要、トーゴの日シンポジウム 2019、日本科学未来館、2019年10月5日
15. 千葉啓和、オーソログ情報の統合化と利活用、トーゴの日シンポジウム 2019、日本科学未来館、2019年10月5日
16. 内山郁夫、微生物比較ゲノムデータベース MBGD を用いた水平移動遺伝子動態の網羅的解析、日本分子生物学会年会、マリンメッセ福岡、2019年12月3-6日
17. 森宙史、藤澤貴智、西出浩世、矢口貴志、高橋弘喜、中川善一、山田拓司、内山郁夫、中村保一、黒川顕、微生物統合データベース MicrobeDB.jp version 3、第14回日本ゲノム微生物学会年会、ウインクあいち、2020年3月6日-8日
18. 山崎由里子、田中麻里、楠屋陽子、青山礼華、大曾根義輝、高橋弘喜、中村悠美、高屋明

子, 黄色ブドウ球菌は Agr のサイレンシングにより院内感染でのパーシステンスとコンピテンシーを獲得する, 第 94 回日本細菌学会総会, オンライン開催, 2020 年 3 月 23 日-25 日

19. 森宙史, 藤澤貴智, 西出浩世, 矢口貴志, 高橋弘喜, 中川善一, 山田拓司, 内山郁夫, 中村保一, 黒川顕, 微生物統合データベース MicrobeDB.jp version 3, トーゴの日シンポジウム 2020, オンライン, 2020 年 10 月 5 日
20. 森宙史, 藤澤貴智, 西出浩世, 矢口貴志, 高橋弘喜, 中川善一, 山田拓司, 内山郁夫, 中村保一, 黒川顕, 微生物統合データベース MicrobeDB.jp version 3 とマイクロバイオーーム解析, 第 15 回日本ゲノム微生物学会年会, オンライン開催, 2021 年 3 月 4 日-6 日

〈国際〉

1. Yuumi Nakamura, Hiroki Takahashi H, Akiko Takaya, Yuzaburo Inoue, Yuki Katayama, Yoko Kusuya, Rena Oguma, Fumiya Yamade, Naoki Shimojo, Gabriel Nunez, Hiroyuki Matsue, Cutaneous acquisition of *Staphylococcus*-sensing agr mutations protects against atopic dermatitis, International Eczema Council at SID meeting/76th SID Annual Meeting, Portland, Apr. 26-29, 2017
2. Hirokazu Chiba, Ikuo Uchiyama, Applying the semantic web technology to diversified biological resources for comparative omics, NIG International Symposium 2017 Commemorating the 30th Anniversary of DDBJ, Mishima Citizens Cultural Hall, May. 27-29, 2017
3. Hiroshi Mori, Takayuki Maruyama, Masahiro Yano, Takuji Yamada, Ken Kurokawa, VITCOMIC2: visualization tool for the phylogenetic composition of microbial communities based on 16S rRNA gene amplicons and metagenomic shotgun sequencing data, NIG International Symposium 2017 Commemorating the 30th Anniversary of DDBJ, Mishima Citizens Cultural Hall, May. 27-29, 2017
4. Hiroshi Mori, Takayuki Maruyama, Masahiro Yano, Takuji Yamada, Ken Kurokawa, VITCOMIC2: visualization of the phylogenetic composition of microbial communities based on 16S rRNA gene amplicons and metagenomic shotgun sequencing, Prokaryotic Genomics 2017, Georg-August-University Göttingen, Germany, Sep. 19-22, 2017
5. Hiroshi Mori, Ken Kurokawa, VITCOMIC2 and MicrobeDB.jp: Analyzing taxonomic composition of microbial communities based on 16S rRNA gene sequences without sequence clustering, ISME17, Congress Center Leipzig, Leipzig, Germany, Aug. 13-17, 2018
6. Hiroshi Mori, Ken Kurokawa, VITCOMIC2 and MicrobeDB.jp: Analyzing and comparing microbial community sequencing data, Frontiers of Genome Science 2019, International Research Center, University of Tokyo, Japan, Jan. 9-10, 2019
7. Hiroki Takahashi, Maico Umemura, Akihiro Ninomiya, Masaaki Shimizu, Yoko Kusuya, Shun-ichi Urayama, Akira Watanabe, Katsuhiko Kamei, Takashi Yaguchi, Daisuke Hagiwara, Diversified transcriptional regulation of secondary metabolic gene clusters in closely related *Aspergillus* species, 30th Fungal Genetics Conference, Pacific Grove, Mar. 13, 2019
8. Hiroshi Mori, Development of the semantic-web based integrated database of human microbiome, The 8th Global Network Forum on Infection and Immunity: Microbi

ome, Chiba University Hospital, January 10–11, 2020.

#### 4. 知財出願

##### (1) 出願件数

種別	件数
特許出願	国内 1 件
	国外 1 件
その他の知的財産出願	0 件

##### (2) 一覧

###### ① 国内出願

1. 情報処理システム、情報処理方法、及びプログラム、黒川顕、東光一、森宙史、2018/1/19、PCT/JP2018/0015942.

###### ② 海外出願

1. INFORMATION PROCESSING SYSTEM, INFORMATION PROCESSING METHOD, AND PROGRAM, Ken Kurokawa, Koichi Higashi, Hiroshi Mori, 2018/1/19, 米国 16/480,539

###### ③ その他の知的財産権

該当なし。

#### 5. 受賞・報道等

##### (1) 受賞

1. Linked Open Data チャレンジ 2017 アイデア部門優秀賞、藤澤貴智・山本泰智、2018 年 3 月 10 日
2. Linked Open Data チャレンジ Japan 2018 プログレス賞、山本 泰智、藤澤 貴智、小林 巖生、2018 年 12 月 8 日

##### (2) メディア報道

該当なし。

##### (3) その他

特になし。

## §8. 研究開発期間中の活動

### 1. 進捗ミーティング

年月日	名称	場所	参加人数	目的・概要
2017年 5月12日	チーム内ミーティング (非公開)	遺伝研	4人	産業応用のためのミーティング
2017年 6月26日	チーム内ミーティング (非公開)	遺伝研	8人	進捗報告のためのミーティング
2018年 1月19日	チーム内ミーティング (非公開)	遺伝研	9人	同上
2018年 2月26日	チーム内ミーティング (非公開)	遺伝研	4人	講習会準備のためのミーティング
2018年 4月25日	チーム内ミーティング (非公開)	遺伝研	4人	進捗報告のためのミーティング
2018年 5月16日	チーム内ミーティング (非公開)	東工大	3人	同上
2018年 6月20日	チーム内ミーティング (非公開)	遺伝研	3人	同上
2018年 7月5日	チーム内ミーティング (非公開)	遺伝研	3人	同上
2018年 9月6日	チーム内ミーティング (非公開)	遺伝研	4人	同上
2018年 11月5日	チーム内ミーティング (非公開)	遺伝研	4人	同上
2018年 11月21日	チーム内ミーティング (非公開)	遺伝研	3人	同上
2018年 12月25日	チーム内ミーティング (非公開)	遺伝研	7人	産業応用のためのミーティング
2018年 12月26日	チーム内ミーティング (非公開)	遺伝研	4人	進捗報告のためのミーティング
2019年 1月30日	チーム内ミーティング (非公開)	遺伝研	3人	同上
2019年 2月19日	チーム内ミーティング (非公開)	遺伝研	3人	同上
2019年 4月11日	チーム内ミーティング (非公開)	遺伝研	3人	同上
2019年 5月10日	チーム内ミーティング (非公開)	遺伝研	3人	同上
2019年 5月24日	チーム内ミーティング (非公開)	東工大	3人	同上
2019年	チーム内ミーティング	遺伝研	3人	同上



年月日	名称	場所	参加人数	目的・概要
6月3日	(非公開)			
2019年 6月20日	チーム内ミーティング (非公開)	遺伝研	3人	同上
2019年 7月1日	チーム内ミーティング (非公開)	遺伝研	3人	同上
2019年 7月22日	チーム内ミーティング (非公開)	遺伝研	3人	同上
2019年 7月26日	チーム内ミーティング (非公開)	東工大	3人	同上
2019年 8月13日	チーム内ミーティング (非公開)	遺伝研	3人	同上
2019年 8月26日	チーム内ミーティング (非公開)	遺伝研	3人	同上
2019年 9月17日	チーム内ミーティング (非公開)	遺伝研	5人	ワークショップ準備のための話し合い
2019年 9月30日	チーム内ミーティング (非公開)	東工大	3人	進捗報告のためのミーティング
2019年 10月17日	チーム内ミーティング (非公開)	遺伝研	7人	進捗報告のためのミーティング
2019年 10月18日	チーム内ミーティング (非公開)	遺伝研	4人	産業応用のためのミーティング
2019年 10月31日	チーム内ミーティング (非公開)	遺伝研	3人	ホログenom連携のためのミーティング
2020年 1月23日	チーム内ミーティング (非公開)	遺伝研	4人	産業応用のためのミーティング
2020年 1月29日	チーム内ミーティング (非公開)	遺伝研	3人	進捗報告のためのミーティング
2020年 2月5日	チーム内ミーティング (非公開)	遺伝研	3人	同上
2020年 4月10日	チーム内ミーティング(非公開)	オンライン	4人	同上
2020年 6月25日	チーム内ミーティング(非公開)	オンライン	5人	同上
2020年 7月9日	チーム内ミーティング(非公開)	オンライン	3人	同上
2020年 7月14日	チーム内ミーティング(非公開)	オンライン	4人	同上
2020年 8月25日	チーム内ミーティング(非公開)	オンライン	4人	同上

年月日	名称	場所	参加人数	目的・概要
2020年 9月28日	チーム内ミーティング(非公開)	オンライン	4人	進捗報告のためのミーティング
2020年 10月15日	チーム内ミーティング(非公開)	オンライン	4人	同上
2020年 11月19日	チーム内ミーティング(非公開)	オンライン	4人	同上
2021年 2月22日	チーム内ミーティング(非公開)	オンライン	3人	同上
2021年 3月18日	チーム内ミーティング(非公開)	オンライン	4人	同上
2021年 3月30日	産業応用のためのミーティング(非公開)	企業社内	3人	産業応用のためのミーティング
2021年 4月15日	チーム内ミーティング(非公開)	オンライン	4人	同上
2021年 5月14日	チーム内ミーティング(非公開)	オンライン	3人	同上
2021年 6月16日	チーム内ミーティング(非公開)	オンライン	4人	同上
2021年 7月7日	チーム内ミーティング(非公開)	オンライン	4人	同上
2021年 7月28日	チーム内ミーティング(非公開)	オンライン	4人	同上
2021年 9月1日	チーム内ミーティング(非公開)	オンライン	4人	同上
2021年 9月29日	チーム内ミーティング(非公開)	オンライン	4人	同上

## 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
2018年 3月7日	MicrobeDB.jp 講習会 (第12回日本ゲノム微生物学会年会内)	京都大学桂 キャンパス	20人	MicrobeDB.jp のポータルサイトの紹介と使い方の講習会、およびユーザ会としてのDBへのニーズの収集
2018年 3月29日	MicrobeDB.jp 講習会 (第91回日本細菌学会総会内)	福岡国際会 議場	50人	MicrobeDB.jp のポータルサイトの紹介と使い方の講習会、およびユーザ会としてのDBへのニーズの収集
2019年7 月31日～8	Quest for Orthologs	基礎生物学 研究所・岡崎	50人	オーソログ DB や推定手法の開発者と利用者が集まっ

月 2 日		コンファレンスセンター		て隔年で開催される国際コンソーシアム会議。内山がオーガナイザーとして、基生研コンファレンスの一環として開催した。
2019 年 10 月 5 日	ワークショップ「微生物—植物相互作用を解明して育種に繋げるには？」	日本科学未来館コンファレンスルーム土星	40 人	<b>MicrobeDB.jp</b> のホログenom連携を推し進めるために、ホログenom研究を行なっている研究者および植物統合DB 開発者と共に、ホログenom研究に有用な統合 DB とはどのような DB なのか等について話し合った。

以上

別紙 研究開発対象のデータベース等

No.	正式名称	別称	概要	URL	公開日	状態	分類	生命科学系データベースアーカイブ	NBDCヒトデータベース	NBDC RDPポータル	関連文献 (論文リストに記載があれば、その番号でも可)
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。	<a href="http://microbedb.jp">http://microbedb.jp</a>	2011/12/12	継続・発展	データベース等	提供前	対象外	提供前	
2	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。	<a href="https://mbgd.nibb.ac.jp">https://mbgd.nibb.ac.jp</a>	1997/7/1	継続・発展	データベース等	提供前	対象外	公開済	Ikuko 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. Nucleic Acids Res., 47, D382-D389, 2019. (DOI: 10.1093/nar/gky1054).
3	Latent Environment Allocation	LEA	MicrobeDB.jpの数万サンプル以上の細菌群集の系統組成データと生息環境に関するメタデータを基に、機械学習の一種であるトピックモデルを用いて既存の細菌群集の系統組成の全体像を2 or 3次元で描画した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。	<a href="http://leamicrobe.jp">http://leamicrobe.jp</a>	2018/6/6	継続・発展	ツール等	対象外	対象外	対象外	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).