

研究開発課題別事後評価結果

➤ 課題情報

研究開発課題名	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
研究代表名	黒川 顕（国立遺伝学研究所 生命情報研究センター 教授）

➤ 事後評価結果

1. 総合評価

当初の実実施計画がすべて達成され、概ね期待通りの成果が得られた。

ホロゲノム解析のようなさまざまな生物種をまたぐ新しい研究分野が生まれてくる中で、微生物が関与する幅広い生命科学研究の課題解決を目指す研究開発課題であった。分野をまたいだ統合データベースの実用化に向け、ユーザビリティの向上やデータ統合を強く意識して開発が実施された。微生物、ゲノム、生息環境などの情報のオントロジーの整備と RDF 化を実施し、それらの情報を他の DB の情報と統合して解析できるようになった点は評価できる。研究開発期間中に複数の企業や外部プロジェクトとの共同研究が進んだことを踏まえると一定の成果はあった。

一方、MicrobeDB.jp に関する論文が期間中に公開されなかった点は残念である。認知度を高め、利用を促す観点から、早急な成果の公開を望む。企業との複数の共同研究や外部プロジェクトでの活用例は特にインパクトが大きいものと推察する。特許公開等、将来の適切なタイミングで公開していくことを検討されたい。

微生物ゲノムやマイクロバイームは日本に優位性のある分野であって、国際連携やデータ統合の中でもリーダーシップを取り、中心的な役割を果たすことが期待されている。このため、微生物分野における国外のデータサイエンスプロジェクトとの連携が望まれる。

2. 研究コミュニティを含むデータ提供者やDB利用者との連携・協業

学会等と連携して講習会や説明会を実施し、ユーザーからの要望をフィードバックすることで、ユーザビリティを考慮したデータベースが構築された点を評価する。今後、ユースケースなどを含めた成果等の発信を行い、利活用の拡大を図っていくことを期待する。

3. 利用者にとって有用なデータ基盤、持続的な DB 運用体制構築に向けた

目標としていた「膨大な微生物データからどのような情報を取り出せるか」ということを具現化するための、わかりやすく使いやすいポータルサイトが実現した。また、データベースのコンテナ化やアノテーションの自動化等、安定的な維持運用に向けた取り組みも実施された。データ収集や標準化などを適切に実施し、様々なデータを適切に RDF 化することで、宿主とマイクロバイームの相互作用、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といった、単一データベース

取り組み	では解析が容易ではなかった複合分野の研究に有用な情報基盤が構築された点を評価する。また、MicrobeDB.jp、Plant GARDEN、MetaboBank の連携は、ホログenom研究に有用な基盤になると期待する。
4. 生命科学研究やイノベーションへの波及効果（DBを利用して得られた研究成果）	複数企業との共同研究や外部プロジェクトで MicrobeDB.jp のデータが活用された。マイクロバイオーム研究の進展に伴う研究コミュニティへの貢献が期待されることから、認知度を高め、さらに利用度を上げていくことが望まれる。
5. その他特記すべき事項	期間中に研究参加者 2 名が昇任するなど、参加した若手研究者のキャリアパスに貢献した点を評価する。
6. 実施計画の達成状況	<p>MicrobeDB.jp を研究開発対象とし、1)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への対応、を目的として、研究開発が実施された。</p> <p>1)については、キーワード検索を想定した Google 検索様インターフェースに加え、ユースケースに特化した様々なアプリケーションをリスト化することで、ファセット検索が可能なユーザーインターフェースが開発された。また RDF モデル更新に伴う全データの更新、ユーザーのメタゲノムデータの容易な登録と比較解析を実施し、SPARQL エンドポイントが公開された。</p> <p>2)については、微生物環境オントロジーMEO を更新し、MEO 自動アノテーションの精度向上が図られるとともに、ヒトマイクロバイオームのメタデータの RDF 化により、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が提供された。</p> <p>3)については、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 PZLAST などが開発された。</p> <p>4)については、真菌類のゲノム、ITS、RNA-Seq データが統合された。</p> <p>5)については、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が実施された。</p> <p>6)については、RDF の再設計、MEO および HMADO 自動アノテーション技術開発、データベースシステム全体のコンテナ化が実施された。</p> <p>7)については、統合化推進プログラムの MicrobeDB.jp、Plant GARDEN、</p>

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年度採択課題 事後評価結果

MetaboBank の3つの課題の連携によるホロゲノム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が実施されている。

上記の通り、当初の実施計画はすべて達成された。