

## 研究開発課題別事後評価結果

### ▶ 課題情報

研究開発課題名	ゲノム・メタゲノム情報統合による微生物DBの超高度化推進
研究代表名	黒川 顕

### ▶ 事後評価結果

1.実施計画の達成度	<p>MicrobeDB.jp を主な研究開発対象とし、1)各種オミックスデータへの対応、2)真菌類および藻類を対象とした拡張、3)各種オントロジー、ボキャブラリの開発、4)解析プロトコルを実装した各種 Stanza の開発、5)データの収集および QC、更新の自動化など持続可能なシステムの構築、6)データ共有、公開におけるアクセスレベルの制御システムの構築、7)構築したシステムを幅広い分野の研究者に活用してもらうためのユーザビリティの向上の7つを研究開発項目に掲げて実施された。</p> <p>1)について、公共のメタ 16S・メタゲノムデータからメタゲノム配列データと付随するメタデータを取得し、3)で開発したオントロジーを用いてアノテーションし、RDF 化して MicrobeDB.jp へ収録した。また、コミュニティとの連携をもとに主要モデル微生物のデータを収集し、データ更新システムの整備、RNA-Seq 解析の投入フローの整備などを行うことにより、ゲノム以外の各種オミックスデータの取得と MicrobeDB.jp への取り込みが自動化された。</p> <p>2)について、真菌類のデータの整理、各種オントロジーの開発などを行い、データの RDF 化を行った。藻類については、オーソログ遺伝子データベースを整備するとともに、植物ゲノム統合データベースを開発・運用する統合化推進プログラムの田畑チームと連携して、オントロジー体系の共用化、整備を行った。</p> <p>3)について、研究者コミュニティと連携し、各種オントロジー、ボキャブラリの開発を推進した。</p> <p>4)について、検索結果の表示システムのソースコードを大幅に見直すことにより、表示速度やユーザビリティを改善し、また、目的に応じた様々な表示システムを新たに整備した。</p> <p>5)について、更新自動化の最も大きな障害となっていた、環境情報オントロジー・アノテーションについて自動化ツールを開発し、作業時間の大幅な削減に成功した。</p> <p>6)について、データの登録・公開を促進するため、それまでのログイン認証システムを拡張し、グループ単位でアクセスレベルを制御できるようにすることで、MicrobeDB.jp のデータと各ユーザーが有するデータを、アクセス制限下で比較、解析し、研究グループ内で共有できる環境を構築した。</p> <p>7)について、ゲノムアノテーションの高品質化・解析支援ツールである Genome Refine を拡張し、微生物ゲノムのアノテーション・パイプライン MiGAP、微生物メタゲノム解析パイプライン MeGAP との連携を強化した。これにより、Genome Refine から MiGAP/MeGAP の出力データを処理し、MicrobeDB.jp および INSDC への登録・公開を支援するシステムが構築された。また、MicrobeDB.jp において、検索した専門用語からユーザーが期待する検索結果を提示できるよう、検索語と検索結果の表示システムの対応関係を明確化し、検索した</p>
------------	--

	<p>用語ごとに多数ある表示システムのなかから適切なものが自動選択されるようにした。</p> <p>収録データについて、当初対象としていた DB からの配列取得を除外したが、将来的な統合の可能性を考慮して生育環境関連語彙を抽出して収録する等、適切な対応がなされた。</p> <p>上記から、実施計画はほぼ全て達成した。</p>
<p>2.生命科学研究への波及効果</p>	<p>MicrobeDB.jp の月間訪問ユニーク IP アドレスは、研究開発開始前の年度と比べて約 25%増加し、期間中、約 550 件から約 600 件の間で推移した。最終年度のみ前年度に比べて一時的に落ち込んでいるが、研究代表者の異動やデータベースのメジャーバージョンアップ等に伴う影響と推察される。また、MBGD の月間訪問ユニーク IP アドレスは、初年度から最終年度にかけて約 1.9 倍(約 2,300 件から約 4,300 件)増加した。</p> <p>マウスにおける免疫抗体による腸内細菌叢の制御を介した腸炎抑制メカニズムの解明など、本データベースを利用した研究成果も報告されている。多くの企業と共同研究や業務提携を実施あるいは実施を検討している。今後、微生物分野の基盤として、保健医療・食品産業・農業など幅広い分野への波及効果が期待される。</p>
<p>3.研究開発課題の運営</p>	<p>3 つの研究グループで構成された実施体制について、分担および協業が適切に運営されており、本研究開発課題の遂行に貢献している。基礎的な部分から実用まで広く開発が行われており、研究代表者のリーダーシップに基づいた運営がなされた。</p>
<p>4.その他特記すべき事項</p>	<p>統合化推進プログラムの他の研究チームや DBCLS との連携、データ産出を行う研究組織や研究室、プロジェクトとの連携を活発かつ効果的に推進した。</p> <p>若手研究者を関連分野へプロモーションするなど、スキルを高める機会を積極的に与えており、実際に学位取得やポストの獲得に結び付いた。</p>
<p>5.総合評価</p>	<p>微生物は医療、食品、環境とさまざまな分野に重要な関わりをもっている。本研究開発課題は、研究機器の飛躍的な発展によって爆発的に産み出されるデータを活用することが困難になりつつある状況をデータベースシステムの問題と捉え、様々なデータを統合し、効果的に利活用する手段を提供することを目的として実施された。本研究開発の実施によってゲノム情報と環境を含む関連情報が融合し、多様な情報解析が可能なデータベースが構築されたことは高く評価できる。研究者コミュニティとの連携を通じ、セマンティック Web 技術を実装し、巨大・多様な生物情報を格納し、なおかつ生命科学研究者が利用できるようなデータベースを構築できることが示された。本データベースは幅広い研究分野の進展に役立つデータソースを提供するものとして大いに期待される。</p> <p>ただし、利用者数は増加傾向にあるとはいえ、本研究分野の重要性や近年の隆盛を踏まえるとさらに多くの利用が望めるものと思われる。今後の研究開発にあたっては、外部有識者を含む会議体を設置し、定期的に意見聴取する体制を構築することも今後の発展の一助となろう。また、データベースを利用することでどのような知識を獲得し、研究発展のヒントを得られるかという事について示して頂きたい。さらに、データ量の急増については、超高度化とどのように両立させるかをふまえて最適解を検討し、さらなる価値の向上に努めて頂きたい。</p>