

研究開発課題別中間評価結果

➤ 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

➤ 中間評価結果

統合化されたデータからどのような新しい知見を得るかという統合データベースの実用化課題に向けて、新たなデータサイエンス研究手法を提案する微生物統合データベースの高度実用化を目指す研究開発課題である。

統合データベースの活用事例を示すための Stanza を設置し、真核生物のメタ ITS (Internal Transcribed Spacer) データの解析パイプラインを開発するとともに、真核生物のメタ ITS データ検索と比較解析を実装した。また、オントロジー自動アノテーションツールを改良してメタデータアノテーションの時間を低減し、効率的なオルソログデータの更新体制を構築した。これにより、ゲノムデータ全体でのオルソログデータベース構築や、必要十分なゲノムデータを取り込んだオルソログデータベース構築のための汎用的な仕組みを開発することができ、オルソログデータに基づいて比較ゲノム解析を行う「オルソ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」などを実装することができた。また、ユーザーの利便性を高めるため、語彙、ヒトメタゲノム文献、オルソログなどを強化したβ版ポータルサイトを公開した。さらに、超高速相同性検索システムを用いることで、全ショットガンゲノム予想遺伝子データのアミノ酸配列 vs. アミノ酸配列の高速解析が可能になった。また、微生物が関与する疾患情報をデータベース化した HMADO や植物ゲノムをデータベース化した PlantGARDEN とも連携させた。

マイクロバイーム研究は、近年、大きく隆盛し、微生物の専門家ではない研究者からも注目されている。MicrobeDB.jp の HMADO や PlantGARDEN との連携は、疾患と微生物、植物と微生物の相互作用を調べる統合解析の入り口にもなるだろう。データベースのコンテナ化による企業との連携強化は、基礎研究のみならず、産業界にも大きな波及効果を及ぼすものとして期待される。

進捗状況と今後の成果見込みは優れており、今後も概ね現計画通りに推進すべきだが、論文についてはできるだけ早い投稿が必要である。また、コンテンツ中のリンク切れなどの不具合が発生した場合は、早期に対処できる体制を構築して欲しい。産業界だけでなくアカデミアとの連携にも取り組むとともに統合 DB 活用の成功事例を共有するため、ユーザーからのフィードバック収集、広報活動などを通じ、さらなるユーザーの拡大につとめていただきたい。また、長期的には、臨床データとの連携についても期待したい。

以上