

研究開発課題別事後評価結果

1. 研究開発課題名

ゲノム情報に基づく植物データベースの統合

2. 代表研究者名

かずさDNA研究所 所長 田畑哲之

3. 研究実施概要

本課題は、異なるプラットフォームに散在する植物ゲノム情報（塩基配列、転写、翻訳、代謝、形質など）や文献上にとどまっている論文のDNAマーカーや連鎖地図などの情報を統合的にまとめ、それら情報の利用利便性を高めるため、ゲノム情報に基づく植物データベースの統合を実施した。なお、統合に当たっては、Resource Description Framework (RDF) の概念を本研究開発に取り入れて実施した。

1) 遺伝子オルソログデータベースの構築とそれに基づく植物ゲノムデータベースの統合

植物についてのゲノム関連データベース（以下、植物ゲノムデータベース）を遺伝子レベルで統合するために、生物種間の進化系統関係を反映させた階層的な遺伝子オルソログデータベースを作成した。遺伝子間のオルソログ関係によって植物ゲノムデータベース群のエントリを相互リンクした。これによって、項目の細分化が進んでいるデータベースの全体像を、遺伝子IDをタグとして統一かつ俯瞰的に探索可能にした。

2) DNAマーカーおよび連鎖地図情報に基づく植物ゲノムデータベースの統合

国内外でゲノム解析やマーカー開発が行われている24科55種の植物を対象に、DNAマーカーの塩基配列情報をタグとして用いることによってマーカー情報の統合を進めた。文献からのキュレーションにより、QTL（質的形質座位）情報を収集、公開し、異なる生物種間で配列情報に基づくリンク付けと地図表示を行い、対象植物種の基本情報やゲノム解析手法の整理、関連データベースリンクの整備を行った。

3) 植物リソース情報データベースの統合

国内の主要リソースセンターが収集・配布している植物リソースのワンストップショップを構築することを目標として、各リソースセンターが持つリソース情報を横断的に検索可能にするとともに、各種植物データベースや本研究開発の他の課題で整備した情報と、本事業で開発したポータルサイト「Plant Genome DataBase Japan」(PGDBj)上で統合検索を可能にした。

4) 植物研究に関連する情報基盤の構築

本課題で構築されたデータベースを有機的に結びつけ、ユーザの利便性を高めるため、ポータルサイト「PGDBj」を構築し、2012年8月より公開を開始した。ユーザの意見を積極的に取り入れ、コンテンツの充実やユーザーインターフェースの改良を実施した。他の研究課題で開発されたデータベース「メタボローム・データベースの開発」で構築された「生物種・代謝物データベース (KNApSAcK)」と「質量分析データデータベース (MassBase)」も検索対象として、ユーザが植物のゲノム関連情報とメタボローム関連情報を取得できるように整備した。さらに、外部サイトから「PGDBj」のコンテンツに対して検索できる API を開発した。

4. 事後評価結果

4-1. 当初計画の達成度

本課題は、(1) 遺伝子オルソログデータベースの構築とそれに基づく植物ゲノムデータベースの統合、(2) DNA マーカーおよび連鎖地図情報に基づく植物ゲノムデータベースの統合、(3) 植物リソース情報データベースの統合、(4) 植物研究に関連する情報基盤の構築 という当初の研究計画を達成した。植物ゲノムデータベースの統合のため、各種データベースのオントロジーの整備や、RDF の概念を取り入れての研究開発を実施した。NBDC の生命科学系アーカイブサービスへの預け入れと公開も実現し、統合化が進んだと言える。

4-2. 研究開発成果の公開および利用の状況等

統合データベースポータルは PGDBj.jp として公開されており、月間ユニーク IP アクセス数は 1,000 件程度となっており、ユーザに利用されているといえる。

4-3. 研究開発成果によるライフサイエンス分野のデータ活用への波及効果

これまで国内各所に分散していた植物ゲノム関連情報が統合され、植物リソース情報を含めて、広範な植物研究の情報基盤として活用できるようになった。多様な植物研究関連の情報ソースとの連携がなされており、植物研究の情報基盤として非常に有用と思われる。さらなる情報の統合により、育種や有用物質生産等の応用に向けた展開が望まれる。

4-4. 広報・アウトリーチ活動等

論文発表が少ないが、学会展示における広報活動を積極的に行い、研究成果の周知に努めており評価できる。

5. 総合評価

本課題は、多種の植物の遺伝子オルソログデータベースを構築し、これをハブとして従来各所に分散していた植物ゲノムデータベースの統合化を実現した。統合された情報のための統合化データベースを構築し、その運営も順調に進捗している。今後、既に始めているメタボローム・データベースとの連携による植物ゲノム情報と生理活性物質の情報の統合や、データの RDF 化による他のデータベースとの連携強化を図り、さらなるデータベースの統合を高度化に進めることが望まれる。なお、植物ゲノムの統合データベースはこれまでは海外でも前例がなく、今後国内の植物ゲノムデータベースだけでなく、この分野を代表する国際的なデータベースへ成長していくことが期待される。