

平成 24 年度 研究開発実施報告書

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」
研究代表者

田畑哲之

公益財団法人かずさ DNA 研究所・副所長

ゲノム情報に基づく植物データベースの統合

§1. 研究実施体制

(1) かずさグループ

- ① 研究代表者: 田畑 哲之 (かずさ DNA 研究所、副所長)
- ② 研究項目
 - ・植物 DNA マーカー、連鎖地図、リソースおよび関連情報の統合

(2) 新潟大グループ

- ① 主たる共同研究者: 中谷 明弘 (新潟大学研究推進機構超域学術院、准教授)
- ② 研究項目
 - ・遺伝子オルソログ DB の構築と植物ゲノム DB の統合



§ 2. 研究実施内容

本研究では、国内に散在する植物データベース（以下、DB）の統合を行い利便性の向上を図ることを目的として、以下の4項目の研究開発を実施した。

(1) 遺伝子オルソログDBの構築とそれに基づく植物ゲノムDBの統合

モデル植物や作物についてのゲノム塩基配列の解読が精力的に進められており、それに伴い遺伝子配列も蓄積しているが、植物を主な対象としたオルソログ関係の整理は十分になされていない状況にある。そこで、植物についてのゲノム関連DB(以下、植物ゲノムDB)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 DB を作成し、遺伝子間のオルソログ関係によって植物ゲノム DB 群のエントリを相互リンクする。これによって、項目の細分化が進んでいる DB 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する。当初の統合対象 DB は、整備が進んでいるモデル植物・作物のゲノム DB(研究代表者グループ保有)や、シロイヌナズナ、イネ、ヒメツリガネゴケ、ゼニゴケ等のゲノム DB(各関連機関保有)とし、状況に応じて順次拡大する。平成 24 年度は、オルソログ DB の更新を行い(RefSeq Release57 に追随)、アミノ酸配列、配列間類似性、種間系統関係、オルソログ情報を一体化させたオルソログ DBを構築し、検索用のウェブサイトを作成した(<http://pgdbj.jp/OD/>)。DB 検索用の基本的な API (SQL と PHP で実現した検索用 URL の記述方法)を定義し(例：<http://pgdbj.jp/OD/operation.html>)、この基本的な API を組み合わせることによってより複雑な検索を実現できるようにした。現在、オルソログ DB には緑色植物 20 生物種(RefSeq に 500 以上のアミノ酸配列が登録されている生物種)の約 50 万のアミノ酸配列が含まれている。ここに RAP-DB(イネ)や TAIR(シロイヌナズナ)等のゲノム DB の他、かずさ DNA 研究所で公開しているユーカリ、ヤトロファ、ミヤコグサ、トマト等のゲノム DB 内のアミノ酸配列を BLAST によって対応付けを行い、オルソログ DB を経由してこれら複数のゲノム DB のエントリ間を遺伝子レベルでリンク(ディープリンク)する機能を作成した。また、ラン藻(111 生物種)や後口生物(44 生物種)等の系統群でのオルソログ DB も緑色植物の外群として作成し、系統的に離れた生物種群との比較解析も行えるように準備している。オルソログ DB 内に登録されたデータはタブ区切りデータとして整理されており、ウェブブラウザを用いた GUIによる検索の他、オフラインでの解析や RDF 化等への対応が可能なように管理している。

(2) DNA マーカーおよび連鎖地図情報に基づく植物ゲノム DB の統合

本研究項目では、ゲノム塩基配列の解読や国内で DNA マーカーが開発されている植物を対象として、DNA マーカーや QTL、連鎖地図関連の DB を調査し、DB に記載されていないこれらの情報を文献や総説から収集する。さらには、収集した DNA マーカーや QTL の位置をゲノム塩基配列や遺伝的地図上において特定するシステムを構築する。平成 24 年度は、前年度に選定した 55 の植物種に対して、引き続き情報の調査と収集を実施した。さらに DB と検索インターフェイスを構築し、ポータルサイトで公開した。ゲノム配列が解読された植物については、マーカー付随配列をゲノム配列上にマッピングすることで連鎖地図や QTL との対応関係を明らかにし、DB の統合化を進めた。また、DNA マーカー付随配列のゲノム配列上および遺伝的地図上の位置の対応関係を

表示するシステムについて機能を追加開発した。最終年度は、24 年度までに進めてきた情報の調査と収集、DB 化を完了させる。なお、国内外の他機関でマーカー関連 DB が整備されているイネ科などの主要な植物種については、該当 DB にリンクすることで、この作業を完了させる。また、地図情報に基づく DB の統合化を完了させる。本項目で整備・構築した DB は、項目(4)で構築したポータルサイトと連携を図って公開を開始しているが、同様に、連鎖地図および物理地図の表示システムで追加開発した機能についても、順次ポータルサイト上で公開する。

(3) 植物リソース DB の統合

本研究項目では、国内に散在する植物バイオリソース情報を統合的に検索可能にすることによって、利便性を高めることをめざす。平成 23 年度は、文部科学省の植物バイオリソースが集約されている二大センターとして、ナショナルバイオリソースプロジェクト DB(以下 NBRP)ならびに理研バイオリソースセンター(以下 BRC)が提供する情報の調査と横断検索のための開発を実施した。平成 24 年度は、NBRP と BRC の双方のリソース検索機能をさらに拡充し、またこれらの拠点で維持されている植物リソース統合情報に加え、これまで未整備であったが我が国の植物リソース情報のなかで重要な位置づけをもつカンキツ類のリソースならびにゲノムデータの統合整備を目的として、情報収集と統合情報システムの作成を行った。具体的には、BRC の「SABRE」における個々の植物を対象とした利用に対応したユーザインタフェース(UI)およびアプリケーションプログラミングインタフェース(API)と並列動作し、SABER が提供する文部科学省 NBRP ならびに理研 BRC のリソース情報を、新規に収集するカンキツ類リソースならびにゲノム情報と統合的に検索するためのシステムの作成を行った。また、これを PGDBj での統合検索に組み込むための基盤整備を実施した。さらに、リソース情報を記載するコントロールボキャブラリの統合のため、NBRP, BRC および農林水産省のそれぞれの拠点にあるリソース情報にてオントロジを掲載している 28 件のデータベースについて体系的に利用オントロジの調査を実施した。

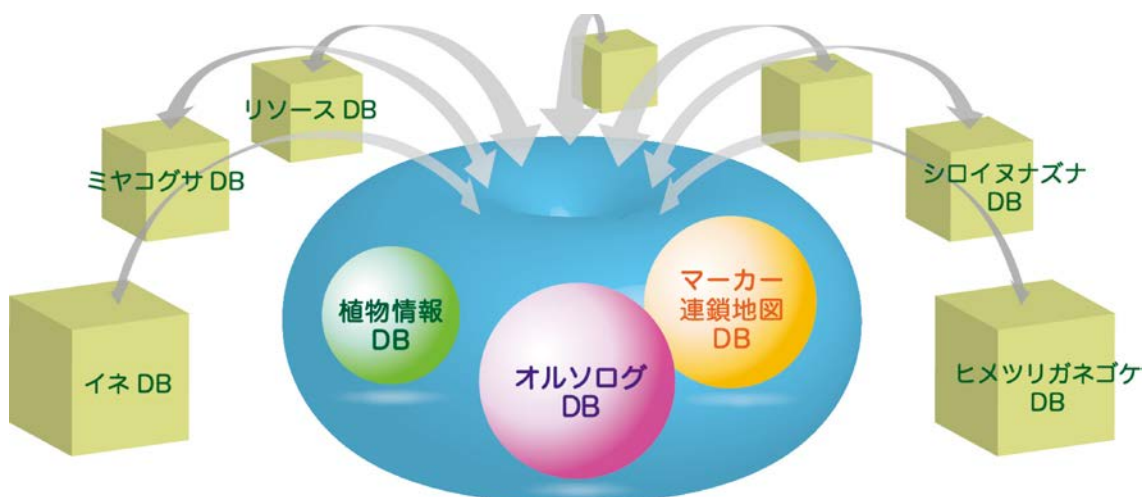
(4) 植物研究に関連する情報基盤の構築

本項目では、ゲノム塩基配列の解読や国内で DNA マーカーが開発されている植物について、植物学や農学、園芸学、育種学上の特徴、有用性、ゲノム塩基配列の決定方法、解析手法等の学術情報を文献(論文や辞典、総説)から収集する。また、ゲノムやトランスクリプトーム、メタボローム、インタラクトーム等のオミックス研究に関する国内外の DB を調査し情報を収集する。平成 24 年度は、対象とする 55 の植物種について、学術的な情報の収集と DB 化を完了した。また、これまでに収集した 1,093 件の国内外のオミックス関連 DB を精査し、重複・リンク切れした情報を除外して約 700 件にまとめた。最終年度は、精査した国内外の DB について、ゲノムやトランスクリプトーム、メタボローム、インタラクトームといった各オミックス研究分野への分類を完了させる。植物とゲノム解析手法に精通したキュレーターとアノテーターの教育も、引き続き行う。

上記4研究開発項目の成果は、平成24年度にポータルサイトで一般公開した(H24 年 8 月 8 日, <http://pgdbj.jp>)。公開後は、ポータルサイトの要望フォームやメール、各学会年会での公開展示を通じてユーザーからの意見を収集し、コンテンツの充実を図った。また、4 研究開発項目に渡る

横断的な検索システムの開発に着手した。最終年度は、横断検索システムを完成し、一般公開することにより、4項目間での情報の連携を図り、植物統合 DB を完成させる(図1参照)。

図1



§3. 成果発表等

(3-1) 原著論文発表

無し

(3-2) データベースおよびウェブツール等の構築と公開

公開中のデータベース・ウェブツール等
別紙を参照

(3-3) その他の著作物(総説、書籍など)

1. 田畑哲之、「植物ゲノムデータベースの統合」細胞工学 vol. 30, No. 12, pp.1314-1317, 2011

(3-4) 国際学会および国内学会発表

① 招待講演 (国内 1件、国際 1件)

〈国内〉

1. 市原寿子、統合化推進プログラム—ゲノム情報に基づく植物データベースの統合—、第54回日本植物生理学会大会、岡山大学、H24年3月22日

〈国際〉

1. 中村保一、Integration of databases for microbes and plants from the viewpoint of (meta-)genomics、Joint Conference on Informatics in Biology, Medicine and Pharmacology (生命医

薬情報学連合大会)、タワーホール船堀、H24 年 10 月 15 日

② 口頭講演 (国内 1件、国際 0件)

〈国内〉

1. 田畑哲之、市原寿子、中谷明弘、中村保一、平川英樹、統合化推進プログラム-ゲノム情報に基づく植物データベースの統合、植物細胞分子生物学会、奈良先端科学技術大学院大学、H24 年 8 月 4 日

〈国際〉

無し

③ ポスター発表 (国内 0件、国際 0件)

〈国内〉

無し

〈国際〉

無し

(3-5) 知財出願

無し

(3-6) 受賞・報道等

無し

(3-7) その他

学会展示での発表を通じた広報活動

1. 市原寿子、平川英樹、中谷明弘、中村保一、田畑哲之、統合化推進プログラム-ゲノム情報に基づく植物データベースの統合、第30回日本植物細胞分子生物学会大会、奈良先端科学技術大学院大学、H24 年 8 月 3日-5 日

2. 市原寿子、平川英樹、中谷明弘、中村保一、田畑哲之、統合化推進プログラム-ゲノム情報に基づく植物データベースの統合、日本植物学会第76回大会、兵庫県立大学、H24 年9月15日-17日

3. 市原寿子、平川英樹、中谷明弘、中村保一、田畑哲之、統合化推進プログラム-ゲノム情報に基づく植物データベースの統合、第35回日本分子生物学会大会、マリンメッセ福岡、H24 年12月11日-14日

4. 市原寿子、平川英樹、中谷明弘、中村保一、田畑哲之、統合化推進プログラム-ゲノム情報に基づく植物データベースの統合、第54回日本植物生理学会岡山大会、岡山大学、H25年3月21日-23日

5. 市原寿子、平川英樹、中谷明弘、中村保一、田畑哲之、統合化推進プログラム-ゲノム情報に基づく植物データベースの統合、日本農芸化学会2013年度大会、東北大学、H25年3月24日-



27日



§4. 研究開発期間中に主催した活動(主催したワークショップ等)

年月日	名称	場所	参加人数	概要
H24年4月20日	グループ合同ミーティング(非公開)	かずさDNA研究所	5人	研究計画確認、進捗報告のためのミーティング
H24年7月20日	グループ合同ミーティング(非公開)	かずさDNA研究所	5人	研究進捗報告のためのミーティング
H24年8月31日	グループ合同ミーティング(非公開)	かずさDNA研究所	4人	研究進捗報告のためのミーティング
H24年10月5日	グループ合同ミーティング(非公開)	時事通信ホール	5人	研究進捗報告のためのミーティング
H24年11月22日	アドバイザリー委員会(非公開)	東京ステーションコンファレンス	22人	外部有識者からの意見聴取のためのミーティング
H24年12月12日	グループ合同ミーティング(非公開)	マリンメッセ福岡	5人	研究進捗報告のためのミーティング
H25年3月26日	グループ合同ミーティング(非公開)	東北大学川内北キャンパス	4人	研究進捗報告のためのミーティング