

課題名	ミヤコグサリソースの活用に向けた Gifu 系統の高精度ゲノム情報整備
課題管理者	佐藤 修正 東北大学大学院 生命科学研究科
実施期間	2015 年度
概要・実施体制	<p>ミヤコグサは NBRP で第 1 期から対象とされているマメ科のモデル植物で、国際的にも広く利用されている。ミヤコグサのゲノム配列解析は宮古島で採取された最も早咲きの MG-20 系統を代表系統として進められ、昨年度の本プログラムで実施した組換え自殖系統 (RILs) のリシーケンス情報を基にして構築した高解像度の遺伝地図情報と高密度の SNP マーカー情報を利用して、代表配列情報の高度化が実現された。</p> <p>本年度のプログラムでは、整備した高密度のマーカー情報を活用して、MG-20 系統と並んでミヤコグサの実験系統として広く世界で利用され、多くの NBRP 収集リソースの親株となっている Gifu 系統のゲノム情報を整備することを目的として、PacBio シーケンサーを用いたゲノム配列情報収集と、NBRP 収集リソースである Gifu 系統の BAC ライブラリの末端配列情報収集を実施する。平均長で 10 kb を越える配列情報が得られる PacBio シーケンサーの特徴を活かして構築する長いコンティグ情報を基にして、昨年度整備した高密度のマーカー情報と、本年度の解析する BAC クローンの末端配列情報を組み合わせることにより Gifu 系統の高精度ゲノム配列情報を構築し、NBRP リソースの活用に向けた基盤情報として整備することを目指す。</p> <div data-bbox="528 730 1254 1447" data-label="Image"> </div> <p style="text-align: center;">ミヤコグサ Gifu 系統</p>
成果	<p>【概要】</p> <p>NBRP の第 1 期から収集対象とされてきたミヤコグサについて、実験系統として広く利用されている Gifu 系統のゲノム情報を整備し NBRP リソースの活用を図ることを目的として、第三世代の PacBio RSII シーケンサーを用いた配列情報収集を行うとともに、Gifu 系統 BAC ライブラリの末端配列を収集し、平成 26 年度の本事業で収集した高密度の SNP マーカー情報と合わせて高精度のゲノム情報の構築を行なった。デンマーク、オーストラリアのグループと共同でミヤコグサ Gifu 系統の PacBio シーケンサーを用い、最終的に 100 倍のゲノムカバーの配列情報を収集した。これらの配列情報を基に、大規模データに対応できるように調整した新規アセンブルプログラムを用いてアセンブルを行なった結果、最終的に N50 が 810 kb の 1685 コンティグにアセンブルすることができた。これらのコンティグを BAC ライブラリの末端配列情報や高密度の SNP マーカー情報を用いて整列化することにより、サイズベースで全コンティグの 99.4% に相当する 1561 コンティグで構成される pseudomolecules を構築することができた。これらの pseudomolecules にはテロメア配列や反復配列リッチなペリセントロメア領域の配列も含まれており、当初計画通りの高精度ゲノム配列情報を得ることができた。</p> <p>【詳細】</p>

- (1) PacBio シーケンサーを用いた配列情報収集
計画通りデンマーク、オース大のグループと共同でミヤコグサ Gifu 系統の PacBio シーケンサーを用いた配列情報収集を行った。アセンブル時の混乱を避けるため、SSR マーカーによる遺伝的背景を確認した Gifu 系統の同一個体から調整した DNA サンプルを用いて双方で配列解析を実施した。本プログラムでの PacBio シーケンサーによる配列解析は、20 kb ライブラリを用いて 10 SMART cells で実施し、平均リード長 10.7 kb, 総塩基数 12.4 Gb のデータが得られ、当初目標を超えるゲノムの 28 倍に相当する配列情報が収集できた。
- (2) Gifu 系統 BAC ライブラリの末端配列解
NBRP ミヤコグサ・サイズ中核機関からレプリカの提供を受けた全 96 プレーットの 3D プールを作製した。各プール DNA サンプルから、BAC クローンの末端配列領域のみを増幅して作製したシーケン斯拉イブラリを MiSeq シーケンサーで解析し配列情報を収集した。各プールから得られた配列情報を、リファレンスゲノム配列にマップし、プール間で共通するマップ位置の情報からクローンを特定する作業を行った。情報を整理して得られた Gifu 系統 BAC ライブラリの末端配列情報は、アセンブルの精度確認作業に応用するとともに、リソースの利用促進に役立てるために NBRP ミヤコグサ・サイズの web サイト、LegumeBase から公開する準備を進めている。
- (3) Gifu 系統のゲノムアセンブリ
デンマーク、オース大のグループで最終的に 75 倍の PacBio シーケンサーを用いた Gifu 系統のゲノム配列情報が収集できたため、日本で取得した 28 倍の配列情報と合わせて目標の 100 倍を超える配列情報を収集することができた。これらの配列情報を基に、大規模データに対応できるように調整した新規アセンブルプログラムを用いてアセンブル作業を進めている。データ量が多いためプログラムの実施に試行錯誤している状況であるが、現時点で、最長コンティグのサイズが 4 Gb、N50 が 700 kb のアセンブル情報が得られている。アセンブルプログラムのパラメータと解析データ量の調整を行うことにより、コンティグ状況の向上を図るとともに、マーカー情報、末端配列情報との対応の情報解析を行い、高精度ゲノム情報の構成を完成させる計画である。