

## 平成 27 年度ナショナルバイオリソースプロジェクト 成果報告書（公開）

補助事業 代表機関管理者 (所属機関・氏名)	国立大学法人 東北大学 大学院生命科学研究科 准教授 佐藤 修正
補助事業課題名	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

### 1. 補助事業の目的

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、一昨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進めている。

### 2. 補助事業の概要

#### (1) Gifu 系統の PacBio シーケンサーを用いた配列情報収集

Gifu 系統の高純度 DNA を調整し、PacBio シーケンサーによる配列解析をデンマーク・オーフス大のグループと共同で実施した。本プログラムでは 25 倍のゲノムカバーに相当する配列情報の収集を目標とした。

#### (2) Gifu 系統 BAC ライブラリの末端配列解析

Gifu 系統 BAC ライブラリの 3D-プールを作製し、MiSeq シーケンサーで末端配列情報を収集した。96 枚の 384 プレートで構成されるライブラリの末端配列を、3D-プールの作製と第二世代シーケンサーを用いたベクター隣接領域の配列解析手法を組み合わせ効率よく解析することを目標とした。

#### (3) Gifu 系統のゲノムアセンブリ

(1) で得られた配列情報をオーフス大で収集する配列情報と合わせてアセンブルし、一昨年度整備したマーカー情報と (2) で得られた情報を基に高精度ゲノム情報を整備する。

### 3. 補助事業の成果（平成 27 年度）

#### (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、オーフ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析

NBRP ミヤコグサ・ダイズ中核機関からレプリカの提供を受けた全 96 プレートの 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スライブラリを 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・ダイズの 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、オーフ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。