

平成28年度医療研究開発推進事業費補助金
(ナショナルバイオリソースプロジェクト) 補助事業成果報告書

I. 基本情報

事業名：ナショナルバイオリソースプロジェクト

National BioResource Project

補助事業課題名：(日本語) 起源を異にするカイコ近交系のゲノムリシーケンシング (2)

(英語) Genome Re-sequencing of Diverse Strains of *Bombyx mori* and *B. mandarina* (2)

補助事業担当者 (日本語) 大学院農学生命科学研究科 教授 嶋田 透

所属 役職 氏名：(英語) Graduate School of Agricultural and Life Sciences
Professor Toru Shimada

実施期間：平成 28 年 4 月 1 日 ~ 平成 29 年 3 月 31 日

分担研究 (日本語) PacBio シーケンサーを用いたクワコゲノム解読

分担課題名：(英語) Genome Sequencing of *Bombyx mandarina* using the PacBio RSII sequencer

補助事業分担者 (日本語) 国立遺伝学研究所 特任教授 豊田 敦

所属 役職 氏名：(英語) National Institute of Genetics Project Professor Atsushi Toyota

成果の概要 (総括研究報告)

1. ゲノムシーケンスの準備

クワコ (*Bombyx mandarina*, カイコガ科) はカイコ (家蚕, *Bombyx mori*) の野生種であり、日本・中国などに分布している。現在世界中で飼育されているカイコは、今から5000年以上前に中国大陸に生息するクワコを元にして家畜化された祖先系統に由来すると考えられている。カイコについては、すでに詳細なゲノム解析が行われているが、クワコのゲノム情報は十分に得られていない。クワコのゲノム情報が得られれば、それをカイコゲノムと比較することにより、家畜化の過程で失われた形質 (たとえば成虫の飛翔能力) や、獲得された形質 (たとえば高品質のシルクを多量に生産する能力) を支配する遺伝子を解明することができると期待される。そこで、本研究ではクワコのゲノムの精密な解読を実施した。

2. ゲノムアセンブリー結果の取得及び評価

東京大学で飼育している埼玉県坂戸市由来のクワコの系統 (1982年から累代) から、雄幼虫 1 頭を配列決定に用いた。配列の決定は、分担者の豊田敦特任教授 (国立遺伝学研究所) が行なった。すなわち、1 分子リアルタイムシーケンサー (PacBio RSII) により、平均12.8kbのリードを370万リード解読し、ゲノムの被覆率95xの配列を得た。また、Illumina HiSeq 2500を用いて、600塩基長のペアエンドを大量に解読し、ゲノムの被覆率133xの配列を得た。これら 2 種類のデータセットを用いてアセンブリーを行

った結果、479Mbの総塩基長となるゲノム配列情報が得られた。500Mb弱と推定されていたクワコゲノムのゲノムサイズと一致したことから、十分な情報が得られていると考えられる。また、コンティグの塩基長はN10が17.4Mb、N50が6.4Mbであり、相当に高精度のゲノム塩基配列が得られたものと評価した。

3. トランスクリプトーム情報を利用した解析

ゲノム情報の最終的な検証に時間を要しているため、トランスクリプトームとの対応付けには、平成27年度のゲノム情報等整備プログラムで解読したゲノム情報を用いた。すなわち、クワコ（坂戸系統）ゲノムの Illumina HiSeq 2500 によるリードから *de novo* のアセンブリーを行うとともに、東大が所有しているクワコ（同系統）の RNA-seq データのアセンブリーも行った。

4. データベースの構築

今回得られた高精度のクワコゲノムの塩基配列情報を、データベース SilkBase (<http://silkbases.ab.a.u-tokyo.ac.jp>) で公開するための準備を進めている。すでに、平成27年度の NBRP ゲノム情報等整備プログラムで支援していただいたクワコ2系統（埼玉県坂戸由来、島根県隠岐の島由来）については SilkBase 上でデータを公開し、BLAST など各種検索ができるようにした。平成28年度のクワコ坂戸系統の PacBio、Illumina の各生リードのデータセットは、DDBJ の DRA アーカイブに登録した (BioProject: PRJDB5778, BioSample: SAMD00080448, DRA: DRA005795 (PacBio) + DRA005813 (Illumina))。現在進行中の確認作業が終了し次第、ゲノムアセンブリーを SilkBase へ掲載する。

1. Materials and Genome Sequencing

Bombyx mandarina (Lepidoptera: Bombycidae) is the wild species of the domesticated silkworm, *Bombyx mori*, and inhabits Japan, China, and other East Asian countries. It is hypothesized that various geographic races of *B. mori* have been originated from an ancestral strain that was domesticated from *B. mandarina* in Mainland China. The genome of *B. mori* has been already studied deeply, whereas that of *B. mandarina* is scarcely analyzed. If the genome information of *B. mandarina* would be available, we would be able to study the genes that have been lost during the domestication history (e.g. genes controlling flight ability of the adult moth) and those acquired during it (e.g. genes for efficient production of high-quality silk), respectively. This program aimed to analyze the *B. mandarina* genome precisely.

2. Genome Assembly and Its Evaluation and Analysis

The genomic DNA was prepared from one male larva of the "Sakado" strain of *B. mandarina*, which was collected in Sakado, Saitama Prefecture in 1982, and thereafter had been reared in the University of Tokyo. The genome sequencing was performed by the collaborator, Dr. Atsushi Toyoda (Project Professor) at the National Institute of Genetics, Mishima. Dr. Toyoda obtained 370 million reads by a single molecule real-time sequencer, PacBio RS II, and obtained the sequences covering approximately 98x of the genome. He also analyzed pair-end short reads by the Illumina HiSeq 2500 sequencer, and obtained the result covering 133x of the genome. These two datasets were assembled together, and consequently the genome sequences of 479 Mb in total were reconstructed. It was nearly identical to the expected genome size of *B. mandarina*, indicating that we obtained the sufficient genome information.

3. Transcriptomic Analysis

We utilized the previous *B. mandarina* (Sakado strain) genome information obtained through

the Genome Information Upgrading Program, 2015 for mapping of transcriptome onto the genome because the completion of genome assembly was a little delayed. The Illumina HiSeq pair-end reads were *de novo* assembled to obtain the draft genome sequences, whereas the RNA-seq reads from the same strain of *B. mandarina* were assembled.

4. Database Construction

The genome information of two *B. mandarina* strains (Sakado and Oki) obtained through the Genome Information Upgrading Program, 2015 was published on "SilkBase" (<http://silkbases.ab.a.u-tokyo.ac.jp>), where the users can do BLAST and other searches. The raw reads produced in the present study (2016) has been deposited to the DDBJ Sequence Read Archive (DRA) (BioProject: PRJDB5778, BioSample: SAMD00080448, DRA: DRA005795 (PacBio) + DRA005813 (Illumina)). We will soon publish the assembly on SilkBase after the ongoing validation.

成果の外部への発表

- (1) 学会誌・雑誌等における論文一覧（国内誌 件、国際誌 件）

- (2) 学会・シンポジウム等における口頭・ポスター発表
 1. 日本発信のカイコバイオリソースの魅力, ポスター, 藤井告, 伴野豊, 梶浦善太, 嶋田透, 瀬筒秀樹, 日本分子生物学会年会 (横浜市), 2016/11/30-12/02, 国内.

- (3) 「国民との科学・技術対話社会」に対する取り組み

- (4) 特許出願
なし

平成 28 年度 医療研究開発推進事業費補助金
成果報告書

I. 基本情報

事業名： (日本語) ナショナルバイオリソースプロジェクト
(英語) National Bioresource Project

補助事業課題名： (日本語) 起源を異にするカイコ近交系のゲノムリシーケンシング (2)
(英語) Genome Re-sequencing of Diverse Strains of *Bombyx mori* and *B. mandarina*

補助事業担当者 (日本語) 東京大学・教授・嶋田 透
所属 役職 氏名： (英語) The University of Tokyo, Professor, Toru Shimada

実施期間： 平成 28 年 10 月 1 日 ~ 平成 29 年 3 月 31 日

分担研究 (日本語) PacBio シーケンサーを用いたクワコゲノム解読
分担課題名： (英語) Genome Sequencing of *Bombyx mandarina* using the PacBio RSII sequencer

補助事業分担者 (日本語) 国立遺伝学研究所・特任教授・豊田 敦
所属 役職 氏名： (英語) National Institute of Genetics, Project Professor, Atsushi Toyoda

II. 成果の概要 (総括成果報告)

補助事業代表者：東京大学・大学院農学生命科学研究科・嶋田 透 総括成果報告を参照。

III. 成果の外部への発表

- (1) 学会誌・雑誌等における論文一覧 (国内誌 件、国際誌 件)
今後、本成果を積極的に発信する予定である
- (2) 学会・シンポジウム等における口頭・ポスター発表
今後、本成果を積極的に発信する予定である
- (3) 「国民との科学・技術対話社会」に対する取り組み
今後、本成果を積極的に発信する予定である
- (4) 特許出願
なし