

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) 1 分子リアルタイム DNA シーケンサーによる MSM/Ms 系統のリシーケンスと  
公開  
(英語) Genome resequencing of Japanese wild mouse-derived MSM/Ms strain

補助事業担当者 (日本語) 高田 豊行  
所属 役職 氏名： (英語) Takada, Toyoyuki

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)  
分担課題名： (英語)

補助事業分担者 (日本語)  
所属 役職 氏名： (英語)

II. 成果の概要 (総括成果報告)

実験動物マウスは、高次生命現象や疾患発症のメカニズムの理解、さらには疾患治療法の開発研究に不可欠な哺乳動物のバイオリソースである。国立遺伝学研究所 (遺伝研) では、世界標準となっている C57BL/6J (B6) 系統とは遺伝的起源が異なる野生由来系統を含む「ミシマバッテリー」を樹立し、理研 BRC を通して広く研究コミュニティに提供している。それらは各種の生命・医学研究分野で幅広く活用されている。本ゲノム情報等整備事業では、「ミシマバッテリー」に属する MSM/Ms (MSM) のゲノムについて、既存の短鎖解読型ゲノム解析機器により得られた情報に加えて、新たに 1 分子リアルタイム DNA シーケンサーによるリシーケンスと公開用データの登録を行った。

まず、国立遺伝学研究所で維持している日本産モロシヌス亜種由来 MSM 系統よりゲノム DNA を取得し、品質検査を行ったのち、ゲノムライブラリを作製した。このライブラリを用いて PacBio 社製ゲノム解析機器による「1 分子リアルタイム (SMRT) シーケンス」に

よるリシーケンスを行った。ゲノム解析に関しては、最終的に総塩基数 85.54Gb のデータを得た。これはマウスゲノム (2.73Gb) の約 31.3 倍の量であり、本整備の目標 (ゲノムの 20-30 倍量のデータを得る) を超える結果を達成した。次に、整備した参照配列の品質を確認する目的で、SMRT シーケンスによる Iso-Seq 解析を行った。MSM 系統およびコントロールの B6 系統の肝臓より RNA を精製し、完全長 cDNA を合成した。完全長 cDNA はサイズ選択を行い、長さの異なるライブラリを作製してゲノム DNA と同様に解析を行い、最終的に総塩基数 54.07Gb のデータを得た。これらのデータは今後整備する MSM 系統の「プラチナゲノム」の品質確認に必要な量であると思われる。本課題は、国立遺伝学研究所の生命情報研究センター比較ゲノム解析研究室および先端ゲノミクス推進センター、さらに情報・システム研究機構データサイエンス共同利用基盤施設ゲノムデータ解析支援センターと共同で実施した。

なお、すべての生データは、DDBJ へ登録済みである (ゲノム ; DRA005749、Iso-Seq ; DRA005763)。

The laboratory mouse is one of the most important bioresources for studying principles underlying higher-order biological phenomena and etiologies of human diseases. National Institute of Genetics (NIG) has established a series of mouse experimental strains named "Mishima Battery" since 1970's. It consists of ten inbred strains originated from three or four subspecies of *Mus musculus*, which are widely distributed in the world. These strains have very remote genetic status from one another and the commonly used laboratory mouse strains such as C57BL/6J (B6), and show very unique phenotypes. The mouse strains in "Mishima Battery" are now distributed to the science research community via RIKEN BioResource center, and contribute to broad range of life science. For genome analysis of the mouse strains in "Mishima Battery", we previously performed a resequencing project of MSM, which was supported by "NBRP genome information upgrading program" and KAKENHI on Priority Areas "Comparative Genomics". Our study revealed more than ten millions of nucleotide substitutions between the B6 reference genome and the MSM genome.

In this program, to elucidate more in detail of structure and copy number variation, further resequencing by the single molecule real time (SMRT) sequencing by the PacBioRSII platform (Pacific Biosciences) was employed for the MSM genome as collaboration with Comparative Genomics Laboratory, Advanced Genomics Center, NIG, and Joint Support-Center for Data Science Research, ROIS. We resequenced genomic DNA isolated from the MSM strain, which was maintained as a pedigreed breeding stock at NIG. Approximately 8.2 million subreads encompassing approximately 85.54 billion base pairs of the mouse genome were generated using the SMRT sequencing by the PacBioRSII platform according to the manufacturer's protocols. Then, we also analyzed full-length cDNAs generated from MSM and the B6 control by the same platform for evaluation of quality of the genome sequence data. We generated

approximately 30.7 million subreads encompassing approximately 54.07 billion base pairs of the reads, which is enough number for the evaluation. These new resources will help accelerate to have the complete genome sequence, so called “platinum genome”, of the MSM strain. The raw data are deposited in a public database (DDBJ accession number, DRA005749 and DRA005763).

### 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 0件、国際誌 0件）  
なし
- (2) 学会・シンポジウム等における口頭・ポスター発表  
なし
- (3) 「国民との科学・技術対話社会」に対する取り組み  
なし
- (4) 特許出願  
なし