

課題名	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
課題管理者	高田 豊行 国立遺伝学研究所
実施期間	2016 年度
概要・実施体制	<p>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す。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」を樹立しました。これらは理研 BRC を通して広く研究コミュニティに提供され、生命・医学研究分野で幅広く活用されています。さらに、遺伝研では、日本産 MSM/Ms (MSM) 系統と B6 との交配により樹立したコンソミック（染色体置換）系統群を開発・提供し、疾患関連表現型などを解析するリソースとして活用されています。我々は、以前より MSM のゲノム多型情報を NBRP 等の支援を受けて整備し、その成果を論文として公表するとともに、データベースを通じて公開してきました。本ゲノム情報等整備事業では、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データの公開を行います。これにより、構造多型を考慮した「より正確な MSM のゲノム参照配列」が整備され、体質の違いや疾患発症に関連したゲノム構造多型の同定に向けた研究コミュニティの多角的な解析に貢献します。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ます。またデータの公開は、国立遺伝学研究所マウスゲノムデータベース“NIG_MoG” (http://molossinus.nig.ac.jp/msmdb/index.jsp) を通じて行います。</p> 
成果	<p>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 倍量のデータを取得) 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</p>

	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）。
--	---