

平成 28 年 6 月 8 日現在

機関番号：63801

研究種目：基盤研究(C) (一般)

研究期間：2013～2015

課題番号：25430097

研究課題名(和文) RNA編集様式の亜種間差を利用したエネルギー代謝関連・量的形質遺伝子の探索

研究課題名(英文) Comparative analysis of RNA editing sites in the mouse using intersubspecific consomic strains

研究代表者

高田 豊行 (TAKADA, TOYOYUKI)

国立遺伝学研究所・系統生物研究センター・助教

研究者番号：20356257

交付決定額(研究期間全体)：(直接経費) 4,000,000円

研究成果の概要(和文)：本研究は、マウス亜種由来系統およびこれらから樹立されたコンソミック(染色体置換)系統群を利用した遺伝子発現解析を行い、生体エネルギー代謝に関わる表現型多様性に関与する「遺伝子発現量差」と「RNA編集塩基の亜種間差」の探索を行い、各種の解析を通してエネルギー代謝に関わる量的形質遺伝子多型の検出を行うことを目的とした。本研究により、マウス亜種系統間の遺伝子発現差およびRNA編集受容候補塩基の座標情報を得た。また、各種の情報解析から、エネルギー代謝調節などに関する遺伝子群に系統間差を観察した。表現型の差に関わる可能性のある複数の遺伝子多型については、各種の検証による関連性の調査を継続して行っている。

研究成果の概要(英文)：Japanese wild mice-derived inbred strain MSM/Ms belongs to the *Mus musculus molossinus* subspecies. It shows highly extent of phenotypic variations in many complex traits and vast amount of genome diversity for the most commonly used inbred strain B6J, genome of which is predominantly derived from west European subspecies *M. m. domesticus*. Using B6, MSM and a number of consomic (chromosome substitution) strains, we intend to survey the intersubspecific differences of the RNA editing sites that controlling the energy metabolism-associated phenotypes by the RNA-seq analysis of liver samples. We have used the suitable genome sequences of B6 and MSM allele to generate a list of RNA editing in this tissue. We observed variation in the number of preliminary predicted editing sites with 1,439 in B6 strain and 1,302 in MSM strain being called. Currently, we are exploring causative RNA editing sites that responsible for the energy metabolism associated phenotypic differences between these strains.

研究分野：哺乳類遺伝学

キーワード：エネルギー代謝表現型 遺伝子発現 RNA編集 マウス亜種

## 1. 研究開始当初の背景

本研究の開始当初は、NGS (次世代シーケンサ) システムの急速な普及にともない、ヒトや実験動物 (系統) のゲノム多型と疾患や表現型多様性を結びつける「ゲノム機能解析」が猛烈な勢いで進行しはじめていた。例えば、ヒト 1000 人規模の全ゲノムを解読しようとする試みや、疾患患者を対象とした全エクソンを対象とした多型探索、多様なモデル生物を対象としたリシーケンス解析、RNA-Sequencing (RNA-Seq) 解析、Chromatin Immunoprecipitation-Sequencing (ChIP-Seq) 解析などであり、これらのゲノム機能解析から得られる膨大なデータは、それまで不可能であった疾患関連遺伝子の探索研究を可能にしつつあった。例えば、マウスの近交系統 17 種のリシーケンス解析は、それまでの報告になかった多くの遺伝子変異、構造多型を発見し、各種ゲノム機能解析に利用できるリファレンス配列情報を提供した (Keane et al. Nature, 2011 ; Yalcin et al. Nature, 2011)。そのような解析の中には、転写後の RNA に起きた塩基置換を高精細の RNA-Seq を使用して検出するといった「RNA 編集」に関するものも含まれる。ヒトを対象とした研究では、B 細胞を用いた RNA-seq 解析により、転写後の RNA 編集によるアミノ酸置換を複数検出していた (Li et al. Science 2011)。従来、リファレンスゲノムと RNA-Seq 解析結果については、双方の品質が問題となり、RNA 編集などのテーマを正面から論ずることはできなかった。しかしながら、データ解析や周辺の IT 環境が進歩したおかげで、小規模な解析からも RNA 編集の検出を行うことが可能になりつつあった。そのような背景から、マウスの近交系統を使用した研究でも、RNA 編集状態に系統間多様性の存在があることが推察され始めた。

## 2. 研究の目的

我々は日本産野生由来 MSM 系統のゲノムを

リシーケンスして公開しているが (Takada et al. 2015 : <http://molossinus.lab.nig.ac.jp/msmdb/index.jsp>)、B6-MSM 間では RNA 編集に関与すると考えられる複数の遺伝子にアミノ酸置換をとまなう塩基置換が存在することから、B6-MSM でこれらが機能する RNA 編集状態に差がある可能性を考えた。また、これまでに行った遺伝子発現解析から、特定の系統で RNA 編集を受けると考えられている遺伝子の発現が上昇していることも見出しており、これらの情報とコンソミック系統のエネルギー代謝関連表現型 (Takada et al. Genome Res. 2008) の間に何かしらの相関を見出すことができれば、量的形質の遺伝子探索解析にこれまで適応できなかった有効な手段を提供できると考えた。

## 3. 研究の方法

本研究課題の実施期間内に以下の解析を設定した。

(1) エネルギー代謝表現型を指標にした解析系統の選抜。

マウス亜種間コンソミック系統群の大規模表現型解析の結果を再検討して、エネルギー代謝に関わる組織を対象に RNA 編集に関する遺伝子発現解析を行う系統を選択する。RNA-Seq 解析を行う系統を決定した後は、選択した系統の試料収集個体以外に B6 および MSM を含む系統を使用して、表現型を確認する目的で、再度複数の個体を使用した表現型収集を行う。

(2) RNA-Seq 解析のための試料収集と解析結果の情報解析

RNA-Seq 解析のためのサンプルは、まず表現型データやマイクロアレイの結果が存在する 10 週令について行うこととし、研究開始年度は MSM、B6 および 9 番コンソミックを対象候補とする。RNA-Seq 解析により RNA 編

集受容塩基を検出するためには、高品質のRNA-Seq解析結果を得る必要があるが、そのためには分解を受けていない高品質なRNAを精製する必要がある。そのため、解析個体の組織から試料を採取する際は、迅速かつ温度条件などに注意し、RNA分解阻害効果のある溶液等を使用して-80度で保存する。RNA-Seq解析による結果が得られたら、B6およびMSMのリファレンス塩基配列を適宜使用して遺伝子発現情報の解析を行う。マッピングツールBowtie、BWAなどによりマッピングを行い、CufflinksあるいはCuffdiffなどにより遺伝子発現情報を得る。データ処理ではSNPの検出条件を検討し、低品質の情報を除外する。検出された発現遺伝子や、表現型に影響を及ぼす可能性のある遺伝子多型のバリデーションを行う。

### (3) RNA発現解析とRNA編集による塩基置換の確認

マッピングが終了した解析セットは、B6-MSM系統間のSNPを考慮したフィルタリングにより遺伝子の発現量をカウントする。さらにRNAの塩基配列内にリファレンスとの塩基置換が存在すれば、これをRNA編集感受座位の候補とする。確認した塩基置換をパターンごとに集計し、遺伝子発現量との相関を解析する。そのために、まずB6とMSMで異なるパターンで編集を受けている遺伝子と発現量が異なる遺伝子を選別する。

### (4) マウスモデルを用いた表現型に対する効果の検証

RNA編集による塩基置換候補は、GO解析やパスウェイ解析からエネルギー代謝に関わる遺伝子であることが確認できれば、表現型模倣(phenocopy)や表現型補完(rescue)により、遺伝子多型と表現型の相関を検証するための交配系統あるいは遺伝子改変動物の作製を行う。候補遺伝子のRNA編集状態と

発現状況の相関、また、エネルギー代謝に関連した表現型への関与を検討する。

## 4. 研究成果

(1) エネルギー代謝表現型を指標にした解析系統の選抜。

RNA-Seq解析に使用する系統を選抜するため、エネルギー代謝表現型を指標にした再検討、およびゲノム支援班との協議等の検討を行った。図1にRNA-Seq解析に使用した系統の表現型の一部を示す。これらのデータは新たに解析した。

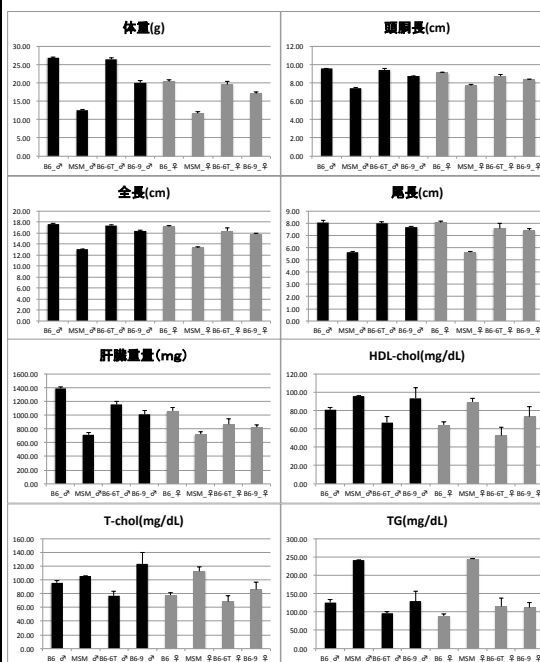


図1. RNA-Seq解析に使用したB6、MSM、コンソミック6Tおよびコンソミック9系統のエネルギー代謝関連表現型収集の結果の一部を示す。オスは黒色、メスは灰色のグラフで示してある。エラーバーは標準誤差を示す。解析に使用した個体数は各系統性別につき6~8匹。全長は鼻先から尻尾の先端まで、頭胴長は鼻先から尾の付け根まで、尾長は尾の付け根から先端までを測定。T-cho(血中総コレステロール)、HDL-cho(高比重リポ蛋白コレステロール)およびTG(トリグリセリド)。MSMのオス個体のT-cho、HDL-choおよびTGは、サンプル血液採取量が少なかったため各1個体のデータである。これら個体の肝臓からRNAを精製し、高品質なものを選択し、RNA-Seq解析を行った。

内容は、B6 および MSM 系統以外にコンソミック系統として、B6 と比較して痩せているにも関わらず、全ての血中コレステロール値が高い染色体 9 番置換系統、さらに脂肪蓄積率が高いものの、コレステロール値に変化の無い染色体 6T 領域置換系統である。RNA-Seq 解析は各系統の雌雄各 3 頭分のデータを想定したが、確認のため表現型データの再収集は 6 匹以上を目安にした。最終的に、MSM、B6 およびコンソミック系を含む計 4 種類の系統の 10 週齢個体を対象にして、各系統から雌雄を含めて合計 50 頭分の表現型データを得た。対象とする形質は、解剖学的観察（体重、肝臓重量など）および血液生化学的観察（血中コレステロールなど）を行った。この結果、いずれの系統においても表現型の再現性を確認することができ、この中の個体由来の各種組織を RNA-Seq 解析に供する試料として保存することとした。

## (2) RNA-Seq 解析のための試料収集と解析結果の情報解析

RNA-Seq 解析のためのサンプルは、MSM、B6 および 9 番コンソミックを対象候補としたが、ゲノム支援班との協議により、染色体 6T 領域置換系統も追加して解析することができた。RNA-Seq 解析により RNA 編集受容塩基を検出するためには、高品質の RNA-Seq 解析結果を得る必要があるが、そのためには分解を受けていない高品質な RNA を精製する必要がある。そのため、解析個体の組織から試料を採取する際は迅速かつ温度条件などに注意して取得し、RNA 分解阻害効果のある溶液等を使用し適切に処理したのち-80 度で保存した。RNA 鋳型調整は基本的に Strand Specific RNA kit を使用した。解析はイルミナ社の HiSeq2500 で行った。表 1 に RNA-Seq により得られた結果を示す。read 数は平均して 5,500 万程度得られた。表に示すように新学術領域研究「ゲノム支援」による支援のおか

表1. RNA-Seq解析により得られた結果

系統	解析個体		RIN	Yield (Mbases)	% PF	# Reads
B6	♂	平均	8.8	5682	95	59585519
B6	♀	平均	8.7	5537	96	57939967
MSM	♂	平均	9.1	4815	96	50209967
MSM	♀	平均	9.0	5642	95	59391675
6T	♂	平均	9.2	5475	95	57397875
6T	♀	平均	9.0	6043	95	63360131
9	♂	平均	9.1	5493	96	57226483
9	♀	平均	9.1	5008	96	52251964

注) 各系統各性別3匹分の平均を示す。

RIN: RNA Integrity Number

Yield: リード長合計

% PF: %Clusters Passing Filter

# Reads: リード数合計

げで高品質な結果を得ることができたので、上述の方法によりマッピングを行い、CufflinksあるいはCuffdiffなどにより遺伝子発現情報を得た。その際のマウスゲノム情報として、B6 に関しては mm9 あるいは mm10、MSM に関しては、我々が公開している MSM 系統のリシーケンスゲノム (<http://molossinus.lab.nig.ac.jp>) をリファレンスとして利用した。MSM のゲノム情報に関しては、別プロジェクトで国立遺伝学研究所の MSM 系統のリファレンスゲノムの品質改善を行っているため、これを参考にした解析により、さらなる結果の品質向上が期待されるので、可能な限りこの情報も使用して再解析を行った。また、B6 対 B6 (mm10) でリファレンスに対して多型のある座位を RNA 編集候補塩基として抽出した。MSM の遺伝子発現は我々が公開している上述の MSMv4 を使用して雌雄別にマッピングを行ったが、この際 MSM vs MSM (MSMv4) でリファレンスに対して多型のある座位を RNA 編集候補塩基として抽出した。コンソミック 6T 系統の遺伝子発現は、染色体 6 番、座標 (0 - 30551893) および (148073078 - end) は mm10、それ以外は MSMv4 を使用してマッピングを行った。30551893 - 33584450 および 145537874 - 148073078 の遺伝子型は不明であった。それ以外は mm10 を使用して雌雄別にマッピングを行ったが、この際コンソミック 6T vs mm10(染色体 6T 領域以外)、MSMv4 (染色体 6T 領域) でリファレンスに対して多型のある

座位を RNA 編集候補塩基として抽出した。コンソミック 9 系統:染色体 9 番(座標 7280995-123427726)は MSMv 4 を使用、0-7280995 および 123427726 - end の遺伝子型は不明であった。それ以外は mm10 を使用して雌雄別にマッピングを行ったが、この際コンソミック 9 vs mm10 (染色体 9 以外)、MSMv 4 (染色体 9) でリファレンスに対して多型のある座位を RNA 編集候補塩基として抽出した。

### (3) RNA 発現解析と RNA 編集による塩基置換の確認

マッピングが終了した解析セットは、B6-MSM 系統間の SNP を考慮したフィルタリングにより遺伝子の発現量をカウントした。さらに RNA の塩基配列内にリファレンスとの塩基置換が存在すれば、これを RNA 編集感受座位の候補とした。確認した塩基置換をパターンごとに集計し、遺伝子発現量との相関を解析することとした。そのために、まず B6 と MSM で異なるパターンで編集を受けている遺伝子と発現量が異なる遺伝子を選別することにした。表 2 に示すように、亜種間で発現

表2. RNA編集候補SNP数のB6-MSM系統間比較

系統	RNA編集候補遺伝子(Ensembl Gene ID)数	
	発現差あり	発現差なし
B6	385	917
MSM	453	986

注) BWAマッピングから得られた結果を用いて、各遺伝子のUTR, CDS, intron領域上に存在するSNP数をカウントした。編集を受けることが判っている遺伝子もこの中に含まれる

量が顕異なる遺伝子群をリストアップすることができた。この RNA 編集受容候補塩基に関しては、既知の RNA 編集受容遺伝子も検出することができていた。また、B6 および MSM とも、RNA 編集の候補塩基は遺伝子発現差に関係なく多数観察された。

### (4) 表現型への効果の検証

RNA 編集による塩基置換候補は、GO 解析やパスウェイ解析を行い、エネルギー代謝に関わる複数の遺伝子を解析の候補とすることができた。一例として表 3 に GO の Biological

表3. RNA編集を受けていると考えられる遺伝子のGO 解析例

GO ID	Category	Term
GO:0045930	BP	negative regulation of mitotic cell cycle
GO:0048641	BP	regulation of skeletal muscle tissue development
GO:0006325	BP	chromatin organization
GO:0030182	BP	neuron differentiation
GO:0010605	BP	negative regulation of macromolecule metabolic process
GO:0045786	BP	negative regulation of cell cycle
GO:0006334	BP	nucleosome assembly
GO:0031487	BP	chromatin assembly
GO:0048634	BP	regulation of muscle development
GO:0034728	BP	nucleosome organization
GO:0065004	BP	protein-DNA complex assembly
GO:0048643	BP	positive regulation of skeletal muscle tissue development
GO:0008284	BP	positive regulation of cell proliferation
GO:0042325	BP	regulation of phosphorylation
GO:0046907	BP	intracellular transport
GO:0019220	BP	regulation of phosphate metabolic process
GO:0051174	BP	regulation of phosphorus metabolic process
GO:0010559	BP	regulation of glycoprotein biosynthetic process
GO:0045069	BP	regulation of viral genome replication
GO:0045941	BP	positive regulation of transcription

注) Biological Process (p値5%未満, 上位20件)

Process (p 値 5%未満, 上位 20 件) の結果を示す。このように、RNA-Seq 解析の結果を用いて、遺伝子発現の標記各系統間の差および RNA 編集受容候補塩基の座標に関する情報を得た。RNA 編集受容候補塩基に関しては、既知の RNA 編集受容遺伝子も検出することができた。B6、MSM およびコンソミック系統とも、RNA 編集の候補塩基は遺伝子発現差に関係なく多数観察された。さらに、遺伝子機能に関する情報解析を行った結果、転写や代謝、筋骨格組織発達の調節に関する遺伝子群に差が観察された。標記遺伝子発現差とエネルギー代謝表現型の系統間差については、各種の検証による関連性の調査を継続して行っている。一例として、染色体 6 番および 9 番に存在するコンソミック系統よりサブコンソミック系統や遺伝子改変動物を作製し、候補遺伝子の遺伝子発現差や RNA 編集と表現型とのリンクを解析中である。このように、得られた情報を指標に RNA 編集感受座位の配列情報と遺伝子発現量、表現型への影響の調査を開始することができた。本研究期間内に責任遺伝子多型の同定には至らなかったが、現在論文化を目指した解析が進行中である。なお、本研究を遂行するにあたり、一部のコンソミック系統の繁殖困難が原因で、データ生産および遺伝子発現解析試料に必要な個体の確保に遅延が生じたが、国立遺伝学研究所系統生物研究センターマウス系統維持事業スタッフの協力により、当初予定数の個体を

得ることができた。RNA-Seq 解析については、新学術領域研究「ゲノム支援」の支援を受けることができた。このおかげで、基盤(C)である本研究課題のみでは解析が困難であったデータを得ることができ、目標のデータ生産量を達成することができた。さらに新学術領域研究「ゲノム支援」には、解析結果の DDBJ への登録についてもご支援いただいた。遺伝子発現情報の 2 次・3 次解析に関しては、新潟大学工学部情報工学科、阿部研究室 阿部博士、矢坂氏の協力を得た。情報解析の一部は国立遺伝学研究所 DDBJ のスーパーコンピュータを利用して行った。

#### 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 6 件)

1. Takada T, Yoshiki A, Obata Y, Yamazaki Y, Shiroishi T: NIG\_MoG: a mouse genome navigator for exploring intersubspecific genetic polymorphisms. *Mammalian genome*, 26: 331-337, 2015, 査読有. DOI: 10.1007/s00335-015-9569-8.
2. Kawano S, Fujisawa H, Takada T and Shiroishi T: Sparse principal component regression with adaptive loading. *Computational Statistics & Data Analysis*, 89, 192-203, 2015. 査読有. DOI: 10.1016/j.csda.2015.03.016.
3. Dou X, Kuriki S, Maeno A, Takada T, Shiroishi T. Influence analysis in quantitative trait loci detection. *Biom J.* 56: 697-719, 2014. 査読有. DOI: 10.1002/bimj.201200178.
4. Oka A, Takada T, Fujisawa H and Shiroishi T. Evolutionarily diverged regulation of X-chromosomal genes as a primal event in mouse reproductive isolation. *PLoS Genet.* 10, e1004301, 2014. 査読有. DOI:10.1371/journal.pgen.1004301.
5. Takada T, Ebata T, Noguchi H, Keane TM, Adams DJ, Narita T, Shin-I T, Fujisawa H, Toyoda A, Abe K, Obata Y, Sakaki Y, Moriwaki K, Fujiyama A, Kohara Y, Shiroishi T. The ancestor of extant Japanese fancy mice contributed to the mosaic genomes of classical inbred strains. *Genome Res.* 23: 1329-1338, 2013. 査読有. DOI: 10.1101/gr.156497.113.

[学会発表] (計 8 件)

1. 矢坂拓, 阿部貴志, 城石俊彦, 高田豊行. マウスを用いた RNA 編集様式の亜種間差探索に向けた遺伝子発現解析. 第 38 回日本分子生物学会 第 88 回日本生化学大会 合同大会. 2015 年 12 月 01 日~04 日. 神戸
2. Toyoyuki Takada, Kentaro Fukuta, Hideki Noguchi, Atsushi Toyoda, Asao Fujiyama and Toshihiko Shiroishi. Genome resequencing of wild-derived inbred strains originated from four *Mus musculus* subspecies. 29th International Mammalian Genome Conference (国際学会). 2015 年 11 月 08 日~11 月 11 日. Yokohama
3. Toyoyuki Takada, Toshihiko Shiroishi. Natural history of Asian wild mice and origin of the laboratory inbred strains, inferred from wholegenome-resequencing. 5th International Wildlife Management Congress (IWMC) (国際学会). 2015 年 07 月 25 日~07 月 30 日. Sapporo
4. 高田豊行, 近藤伸二, 阿部貴志, 清澤秀孔, 豊田 敦, 藤山秋佐夫, 城石俊彦. マウス野生由来系統のゲノム情報を利用したアレル発現量解析. 日本遺伝学会第 86 回大会. 2014 年 09 月 17 日~09 月 19 日. 長浜

[図書] (計 1 件)

1. 高田豊行, 城石俊彦. 生命のビッグデータ利用の最前線 (監修: 植田充美). シエムシー出版. 2014 年. 231 (161-168)

#### 6. 研究組織

(1) 研究代表者

高田 豊行 (TAKADA, TOYOYUKI)  
国立遺伝学研究所・系統生物研究センター・助教  
研究者番号: 20356257

(2) 研究分担者

なし

(3) 連携研究者

なし

(4) 研究協力者

阿部 貴志 (ABE, TAKASHI)  
新潟大学・自然科学系・准教授  
矢坂 拓 (YASAKA, TAKU)  
新潟大学・自然科学系・学生