

機関番号： 82401  
 研究種目： 基盤研究(B)  
 研究期間： 2008 年度～ 2010 年度  
 課題番号： 20310118  
 研究課題名（和文）百万年以上にわたり組換えが抑制された染色体ゲノムの構造・動態解析  
 研究課題名（英文）Structural analysis of genomic regions where genetic recombination was suppressed for over million years  
 研究代表者  
 阿部 訓也 (ABE KUNIYA )  
 独立行政法人理化学研究所・動物変異動態解析技術開発チーム・チームリーダー  
 研究者番号： 40240915

## 研究成果の概要（和文）：

本研究ではマウス 17 番染色体に位置する“数百万年の間、組換えが抑制された染色体領域”に着目し、その染色体ゲノムがどのように進化するかを調べることを目的とした。解析の結果、この領域には大規模な逆位が存在しており、その境界点近傍には重複、欠失等のゲノム再編成が認められた。また、この領域のゲノムクローンを単離し、解析した結果、逆位を持つゲノムの配列には、標準系統に比べ、非常に高頻度の一塩基遺伝子多型が検出されることがわかった。

## 研究成果の概要（英文）：

Goal of this study is to elucidate the structure and divergence of t-haplotype genome in which genetic recombination has been suppressed for over million years. Large scale DNA rearrangements including deletions and duplications were revealed by Array-CGH method, inferring the break points of inversions associated with t-haplotypes. BAC genomic library was constructed from  $t^{w5}$  haplotype, and clones corresponding to the t-complex region were isolated and sequenced. Exceptionally high incidence of SNPs (single nucleotide polymorphisms) has been detected between the  $t^{w5}$  haplotype and the standard laboratory mouse strain.

## 交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2008 年度	6,300,000	1,890,000	8,190,000
2009 年度	5,000,000	1,500,000	6,500,000
2010 年度	3,600,000	1,080,000	4,680,000
総計	14,900,000	4,447,000	19,370,000

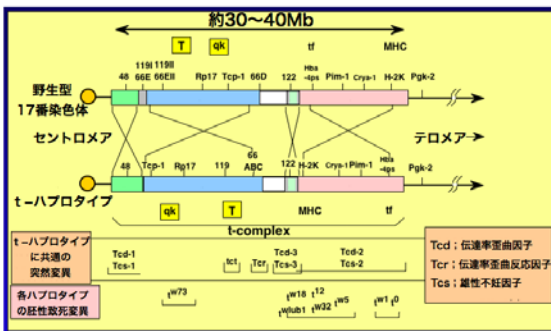
研究分野： 複合新領域

科研費の分科・細目： ゲノム科学・基礎ゲノム科学

キーワード： ゲノム進化・再編

## 1. 研究開始当初の背景

生物は周囲の環境・生態に適応するため自身が持つゲノムを変化させ、進化してきた。染色体は遺伝情報のキャリアとして、次代への正確な情報伝達を担うと同時に、複数のメカニズムにより遺伝的多様性を生み出していく。特に染色体組換え（交叉）は多様性を生み出す原動力であり、ゲノムの再編・進化にも重要な役割をもつものと考えられている。しかし、組換えが実際にどの程度、ゲノム進化に貢献しているかについての明確な回答は得られていない。



t-コンプレックスはマウス 17 番染色体近位に位置する遺伝領域であり、マウス自然集団にはこの領域に 4 つの逆位を持つ、通常の 17 番染色体とは異なる構造を持った染色体が存在する（上図参照）。この逆位を持つ染色体（t-染色体、あるいは t-ハプロタイプと呼ぶ）と通常の染色体の間では減数分裂時の染色体組換えが強く抑制されている。t-染色体は約 300 万年前に生じたと推測されているので、t-染色体は数百万年単位の進化的時間のなかで、他の染色体と組換えを起こすことなしに独自の進化を遂げてきたということができる。したがって、染色体組換えとゲノム進化の関係を追究するためには、好個の材料ということができる。

また、染色体組換えの抑制の他にも、精子伝達率を歪曲する複数の因子や、発生分化などの生命現象に関係する突然変異も多くマップされており、t-コンプレックスは様々な生物学的に重要な問題を含んでいる。t-ハプロタイプの研究は、1920 年代に端を発し、70 年代には多くの研究が行われたが、その構造の複雑さ、遺伝解析が（組換えが起きないので）困難なこと、材料の特殊性などの理由から、現在では国内はもとより、海外でも t-ハプロタイプの研究は殆ど行われていないのが現状である。申請者は長年、この分野の

研究を継続しており、最近劣性致死変異<sup>t<sup>w5</sup></sup>の責任遺伝子を同定したが、これは t-ハプロタイプの研究の歴史の中で初めてのことである。一旦は忘れ去られた感のある t-ハプロタイプの研究であるが、先述したように、そこに含まれる問題は全く色褪せておらず、むしろ先端的なゲノム解析技術を投入することにより、非常に学術的価値の高いゲノム情報を研究コミュニティに提供することが可能である。数百万年の間、組換えが抑制された t-ハプロタイプは*いわば生きた化石*のようなゲノム領域であり、このような状況下でゲノムがどのような変化を示すかを探るには、この t-ハプロタイプという希少な染色体ゲノムを解析する他はない。

ここから得られる情報は単に、t-ハプロタイプという特殊性を理解するだけではなく、組換えが染色体の進化、再編成にどのような働きを持つか、という一般的な命題の理解にも貢献するものと考えられる。さらに、t-ハプロタイプは多くの興味深い突然変異を含むが、連鎖解析ができない等の理由から殆どが未同定のままである。t-ハプロタイプのゲノム情報の解読が進めば、これらの変異遺伝子の同定も容易となるのは明らかである。このような理由から、t-ハプロタイプゲノム解析を着想するに至った。

## 2. 研究の目的

本研究では数百万年の間、組換えが抑制されたマウス 17 番染色体の染色体領域である t-ハプロタイプの構造・配列解析を行い、その構造情報を得ることを第一義の目的とする。さらに、構造情報をもとに、組換えが関与しない場合の染色体の進化過程を追跡し、組換えの意義を染色体レベルで改めて評価することを試みる。

## 3. 研究の方法

### 1) t-ハプロタイプ BAC ライブラリーの作製

t-ハプロタイプは通常、発生致死変異を有しているため、ホモ個体を得ることができない。しかし、ゲノム解析のためにはホモ個体からの DNA を用いた方がはるかに効率が良い。申請者らは最近発生致死変異の責任遺伝子のひとつを同定し、この遺伝子を含む約 10 kb のゲノム断片を導入することにより、ホモ個体の致死表現型をレスキューすることに成功した。そこで、このレスキューされたホモ個体から抽出したゲノム DNA を出発材料とし、BAC ライブラリーを作製する。このライブラリー

を t-コンプレックス由来のプローブを用いてスクリーニングし、クローンを単離する。

## 2) アレイCGHによるゲノム構造解析

アレイCGH (comparative genomic hybridization) を用いて、t-ハプロタイプと標準系統の間のコピー数の変化や、欠失・重複等のDNA再構成の有無を調べる。そのため、既知の標準マウスゲノム配列 (C57BL/6J 系統; 以下B6) を元にして、17番染色体のt-コンプレックスに相当する領域からプローブ配列を設計し、NimbleGen社、あるいはAgilent社に委託して、高密度ゲノムアレイを作製する。このアレイは、60merのプローブを搭載し、約30Mbと推定されている t-コンプレックス領域の全域をカバーする。B6 DNAを参照ゲノムとして、t-ハプロタイプゲノムとの比較を行う。

また、異なる t-ハプロタイプ間の比較も行う。

## 3) t-コンプレックス領域からのBACクローン単離

t-ハプロタイプに存在する逆位の境界点の詳細については不明であるが、境界領域では欠失や重複などのDNA再編成を伴う可能性が高い。そこで、アレイCGHによって、欠失や重複などの構造変化が示唆された領域に関してはBACクローニングを行う。また、4つの逆位それぞれの領域からBACクローンの単離を行う。

## 4) t-ハプロタイプゲノム配列の決定

単離したBACクローンの配列決定を行う。方法としては、サンガー法を用いたショットガンシーケンス、複数のBACクローンをプールし、ペアエンドライブラリーを作製し、次世代シーケンサーで配列を読む、という2つの方法を用いる。また、異なる t-ハプロタイプの配列取得のため、PCRにより相同領域を増幅し配列決定を行うことに加え、ハイブリッドキャプチャー法の適用も検討する。

## 4. 研究成果

### (1) t-ハプロタイプBACライブラリーの作製

t<sup>w5</sup>変異マウスのレスキュー個体から繊維芽細胞を作製し、この細胞より高分子DNAを調製した。EcoRIによる部分消化を行い、pBACe3.6ベクターにクローニングし、約10万のBACクローンからなるライブラリーを作製した。インサート平均長は約100kbであった。このライブラリーはt-ハプロタイプから作られた初めてのBACライブラリーであり、t-ハプロタイプのゲノム解析に必須の材料と考えられる。

### (2) アレイCGHによるゲノム構造解析

t<sup>w5</sup>/t<sup>w5</sup>、t<sup>w2</sup>/t<sup>w2</sup>、t<sup>w8</sup>/t<sup>w8</sup>、t<sup>w5</sup>/t<sup>w12</sup>、t<sup>w18</sup>/t<sup>w18</sup>、t<sup>h49</sup>/t<sup>h49</sup>の6種類のt-ハプロタイプからゲノ

ムDNAを調製し、参照ゲノムとしてB6よりDNAを調製した。6種のt-ハプロタイプのうち、t<sup>w18</sup>とt<sup>h49</sup>は、t-コンプレックスの途中から野生型染色体と同様の構造となる部分 (partial) t-haplotype である。t<sup>w18</sup>には欠失があり、その中に発生致死変異責任遺伝子があると考えられているが、正確な欠失領域は明らかではなかった。t<sup>h49</sup>は逆位の1と2を持つが、3以降は正常タイプとなっていると考えられている。CGH解析から得られた主要な結果を以下に記す。

1) 調べた6種類すべてのt-ハプロタイプに共通して、B6との間で重複あるいは欠失が認められる領域を特定した。これはt-ハプロタイプに特徴的な逆位の境界点近傍に生じたゲノム再編成である可能性が高い。

2) 上記のゲノム再編成のうち、最もセントロメア側に存在する重複 (tのコピー数がB6に対して増加している) の位置は、セントロメア側から約5.7Mbであった。この位置からt-complexが開始している可能性が高い。また、約40Mbの位置に大きな欠失があるが、これより遠位 (B6染色体において) では、コピー数の変化は認められなくなる。したがって、この位置がt-complexの終端である可能性が高く、t-complex全体はB6ゲノム上の約34Mbの領域に相当すると推測された。

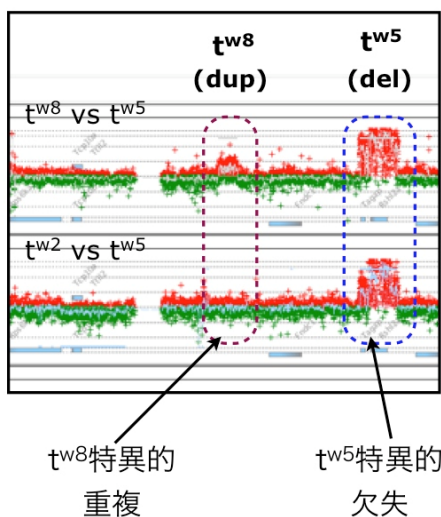
3) t<sup>w18</sup>ハプロタイプには、18.9Mbから23.6Mbの間に約4.7Mbにわたる欠失領域が存在することが確認された。この欠失領域には vomeronasal receptor 遺伝子と Zinc finger protein 遺伝子が数十回重複を繰り返すクラスター構造が存在した。

4) t<sup>h49</sup>ハプロタイプを他のt-ハプロタイプやB6と比較すると、約13.7Mbまでは、t-ハプロタイプとよく類似したパターンを示し、それ以降はt-ハプロタイプとは異なり、B6に近いパターンを示した。したがって、部分t-ハプロタイプであるt<sup>h49</sup>の組換えは、この13.7Mb近傍で起きたと考えられた。これまでの研究から、t<sup>h49</sup>ハプロタイプは逆位の1番と2番を含むとされている。したがって、これ以降が逆位の3番、4番の領域に相当する。

5) 上述したように、約40Mbの位置が、逆位4の遠位側の終端であると考えられるが、その近傍の約37Mbから39.6Mbにかけては、t-ハプロタイプとB6の間で、シグナル強度の変動が連続して認められた。この領域も嗅覚受容体遺伝子がクラスターとして存在しており、それらのコピー数、あるいは配列自体がB6とt-ハプロタイプでは有意に異なっている可能性が示唆された。

6) t-ハプロタイプ同志の構造やDNA配列は、相互によく類似しているものと考えられてい

る。実際、異なる t-ハプロタイプ間でも、共通した欠失・重複が生じていることが確かめられた。しかし、そのゲノム再編成のパターンがすべての t-ハプロタイプで同じというわけではなく、ハプロタイプ間で異なる点も比較的頻繁に見つかった。例えば、右上の図に  $t^{w8}$ 、 $t^{w2}$  ハプロタイプをそれぞれ  $t^{w5}$  ハプロタイプと比較した結果を示す。図示したように、 $t^{w5}$  特異的なホモ欠失が認められ、また  $t^{w8}$  に特異的な重複が存在しており、これを見る限り、3者は相互に異なっていることがわかる。他にも、このような領域は存在し、すべての t-ハプロタイプが巨視的にみて、同一の構造を持つものではないことが明らかとなった。特に、 $t^{w5}$  ハプロタイプは、他の（今回調べた）5つの t-ハプロタイプと比較して異なる点が多いことが判明した。この結果は t-ハプロタイプの進化を考える上で重要な知見となり得るものと考えられる。



$t^{w2}$ ,  $t^{w5}$ ,  $t^{w8}$  ハプロタイプ間での構造の差異を示す CGH 結果

### (3) t-コンプレックス領域からのBACクローン単離

アレイ CGH の結果からおおよその逆位の領域、境界点の位置が推測できたので、各領域から BAC クローンの単離を試み、以下のクローンを得た。

遺伝子名	クローン数	位置
Zdhhc14	2	逆位 1
Dynlt1	6	逆位 1
Tagap1	1	逆位 2-3 末端?
Tcp10a	5	逆位 1-2、2-3
T/T2	1	逆位 2 内
Igf2r	1	逆位 2 内
Tcte3	1	逆位 3 内
Fpr1	1	逆位 4 内
H-2K	1	逆位 4 内

計 19

同数のクローンをスクリーニングしたが、Dynlt1 と Tcp10a 陽性クローンは他に比べ数多く単離された。これは、この領域のコピー数が t-ハプロタイプにおいて増加していることを示唆するものと思われる。

### 4) t-ハプロタイプゲノム配列の決定と解析

単離したクローンの配列決定を行った。各遺伝子の陽性クローンから DNA を抽出し、これをプールして、物理的に剪断し、約 8kb の断片を分画した。この分画を用いてペアエンドライブラリーを作製し、その両方の末端それぞれ約 300-400 base を Roche 社 454 FLX-Titanium を用いて決定した。約 27 万リードを解析し、アセンブルを行い、666 の contig、22 の scaffold を得た。scaffold 長の合計は 1,216,992 base であった。

これらの scaffold を B6 ゲノムにマップし、相当するゲノム領域を同定した。例えば、約 90kb の Scaffold No.5 は、Tagap1 陽性の一つの BAC クローンに由来することがわかったが、B6 ゲノム上では scaffold の前半と後半部分でマップされる場所が異なり、約 2.4Mb 離れた 2つの領域に位置することがわかった。これは、このクローン中に逆位の切断点が含まれることを意味する。また、複数のクローンが Dynlt1、Tcp10a 陽性であったが、これらのクローンに由来する配列は、B6 のゲノムでも重複していることが明らかとなった。一方、t-ハプロタイプでも重複のない領域にある Igf2r 遺伝子を含む Scaffold 配列では、B6 と同様な DNA 構造を持つことが確かめられた。このように、CGH 結果を考慮してクローンを単離・利用することで、t-ハプロタイプで起きているゲノム再編成部位の配列解析を効果的に実施できることが示された。

配列解析は当然のことながら、マクロな構造変化だけでなく、1塩基遺伝子多型 (SNP) などの塩基置換や小さな挿入・欠失 (indel) の有無を検索することが可能である。例えば、H-2K 領域の 2つの領域（それぞれ 31,732 base、34,735 base）について、 $t^{w5}$  ハプロタイプと B6 を比較したところ、前者では 634

SNPs, 65 indels (総計 2,750 base)、後者では 273 SNPs, 28 indels (378 base)の違いが認められた。塩基置換率はそれぞれ 0.0199, 0.0076 であった。これまで報告された中では、マウス亜種の一つである *Mus musculus molossinus* と B6 の間で最も塩基置換率が高く、0.008-0.010 であることが知られている。今回得られた結果と比べると、低い方の領域でも、*molossinus* に見られる塩基置換率に匹敵し、高い方では *molossinus* の約 2 倍というこれまでにない高値を示した。この結果は、*t*-ハプロタイプのゲノムが現存のマウスのゲノムから大きく diverges していることを明確に示している。また、亜種間で見られる塩基置換率よりもさらに大きい置換頻度を示すことから、*t*-ハプロタイプがマウス亜種分化に先立って出現した、という説を支持する結果と考えられる。今後は、さらに幅広い領域について同様の解析を行うとともに、異なる *t*-ハプロタイプ間での塩基配列の差異を調べていく必要があると考えられる。

#### (5) 新しい枠組みから観る哺乳類ゲノムアイソコアの進化 (研究分担者: 太田聡史)

哺乳類のゲノムレベルの進化を考える場合、空間的な塩基組成の構造 (アイソコア) を考慮することは重要である。一塩基多型 (SNP) から得られる種内変異についてのデータと、近縁種を系統学的に比較することにより、ゲノムの GC 含量のクラスごとに突然変異パターンを推定し、MySQL データベースに格納するソフトウェアを開発した。開発には主に数式処理システムである Mathematica を用いた。さらに従来とはまったく異なる枠組みとして、AT→GC および GC→AT の変化率の (ゲノム上の位置依存的な) 一定性を仮定しない進化モデルを提案し、Variable model と名づけた。従来の AT→GC および GC→AT の変化率の一定性を仮定したモデル (Constant model) では、平衡値は必ず存在し、その値は初期値と突然変異パターンに依存する。しかし、Variable model では、平衡値は突然変異パターンのみに依存することが示された。また、Constant model では種間の置換パターンから推定した平衡値と、種内の突然変異パターン (SNP データ等を用いて推定した突然変異パターン) から推定した平衡値は、CpG サイトを除くかどうかに関わらず、一般に異なる。しかし、Variable model では、両者は近い値になることが示された。一方、Variable model では、原理的に GC 含量の平衡値が存在することは保証されない。このことは一見上記の主張と矛盾するようにも見える。われわれは突然変異パターンに応じて、GC 含量の時間的な挙動を 4 つのクラスに

分けて議論した。われわれが用いた (現在入手できる) データのほとんどは、平衡値の存在するクラスに分類される。一方、理論的にはわずかな突然変異パターンの変動により、GC 含量の時間的な挙動は大きく変化することも予測された。このことは、いまだ謎に包まれているアイソコアの進化に一石を投じるものになるかもしれない。

#### 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 21 件)

1. Oota, S., Kawamura, K., Kawai, Y. and Saitou N. (2010) A new framework for studying the isochore evolution: estimation of the equilibrium GC content based on the temporal mutation rate model. *Genome Biol Evol* 2, 558-571. (査読あり)
2. Numata, K., Kohama, C., Abe, K., Kiyosawa, H. (2010) Highly parallel SNP genotyping reveals high-resolution landscape of mono-allelic Ube3a expression associated with locus-wide antisense transcription. *Nucleic Acids Res.* [Epub ahead of print] (査読あり)
3. Kobayashi, S., Fujihara, Y. Mise, N., Kaseda, K., Abe, K., Ishino, F., Okabe, M. (2010). The X-linked imprinted gene family Fthl17 shows predominantly female expression following the two-cell stage in mouse embryos. *Nucleic Acids Res.* Feb.25. [Epub ahead of print] (査読あり)
4. Kasahara, T., Abe, K., Mekada, K., Yoshiki, A., Kato, T. (2010) Genetic variation of melatonin productivity in laboratory mice under domestication. *Proc. Natl. Acad. Sci, USA* 107, 6412-6417. (査読あり)
5. Cao, L., Shitara, H., Sugimoto, M., Hayashi, J.-I., Abe, K., Yonekawa, H. (2009) New evidence confirms that the mitochondrial bottleneck is generated without reduction of mitochondrial DNA content in early primordial germ cells of mice. *PLoS genetics* 5, e1000756. (査読あり)
6. Tomida S, Mamiya T, Sakamaki H, Miura M, Aosaki T, Masuda M, Niwa M, Kameyama T, Kobayashi J, Iwaki Y,

- Imai S, Ishikawa A, Abe K, Yoshimura T, Nabeshima T & Ebihara S. (2009) *Usp46*, encoding a ubiquitin specific peptidase, is a quantitative trait gene regulating mouse immobile behavior in the tail suspension and forced swimming tests. *Nature Genetics* 41, 688-695. (査読あり)
7. Mekada, K., Abe, K., Murakami, A., Nakamura, S., Nakata, H., Moriwaki, K., Obata, Y. and Yoshiki, A. (2009) Genetic differences among C57BL/6 substrains. *Exp. Anim.* 58, 141-149. (査読あり)
8. Yoshiki, A., Ike, F., Mekada, K., Kitaura, Y., Nakata, H., Hiraiwa, N., Mochida, K., Ijuin, M., Kadota, M., Murakami, A., Ogura, A., Abe, K., Moriwaki, K. and Obata, Y. (2009) The mouse resources at the RIKEN BioResource Center. *Exp. Anim.* 58, 85-96. (査読あり)
9. Ogonuki N, Inoue K, Hirose M, Miura I, Mochida K, Sato T, Mise N, Mekada K, Yoshiki A, Abe K, Kurihara H, Wakana S, Ogura A. (2009) A high-speed congenic strategy using first-wave male germ cells. *PLoS One* 4(3):e4943. Epub 2009 Mar 31. (査読あり)
10. Nomura-Kitabayashi A., Phoon K.L. C., Kishigami S., Rosenthal J., Yamauchi Y., Abe K., Yamamura K., Samtani R., Lo W. C., Mishina Y. (2009) Outflow tract cushions perform a critical valve-like function in the early embryonic heart requiring BMPRIA-mediated signaling in cardiac neural crest. *Am. J. Physiol. Heart Circ. Physiol.* 297, H1617-1628. (査読あり)
11. Smalheiser, N.R., Lugli, G., Torvik, V.I., Mise, N., Ikeda, R., Abe, K. (2008) Natural antisense transcripts are co-expressed with sense mRNAs in synaptoneuroosomes of adult mouse forebrain. *Neurosci Res.* 62, 236-239. (査読あり)
12. Mise, N., Fuchikami, T., Sugimoto, M., Kobayakawa, S., Ike, F., Ogawa, T., Tada, T., Kanaya, S., Noce, T. and Abe, K. (2008) Differences and similarities in the developmental status of embryo-derived stem cells and primordial germ cells revealed by global expression profiling. *Genes to cells* 13, 863-877. (査読あり)
- [学会発表] (計 75 件)  
国内会議 (55 件)
1. 太田聡史: アイソコアの進化:新しい枠組み 第82回日本遺伝学会大会、2010年9月21日、札幌。
  2. 益崎智子、山崎孝仁、阿部訓也 (10名中7番目): C57BL/6N系統由来のBACライブラリーの構築と日本発大型リソースの整備。第32回日本分子生物学会年会、2009年12月10日、横浜。
  3. Liqin CAO, Kuniya ABE: Evaluations of functional processes in live animals by in vivo imaging. 第56回日本実験動物学会総会、2009年5月16日、大宮。
  4. 津田 薫、持田慶司、吉木 淳、森脇和郎、阿部訓也: マウス垂種間の兄妹交配で生じる雑種崩壊現象の表現型解析 第55回日本実験動物学会総会、2008年5月15日、仙台。
- 国際会議 (20 件)
- [図書] 和文総説  
(計 3 件)  
太田聡史、阿部訓也、沼田興治  
「表現型解析の潮流・表現型の数値化と定量解析」『生命科学研究の戦略：疾患モデルマウスと表現型解析』中山書店 (印刷中)
- [その他]  
ホームページ等  
<http://www.brc.riken.jp/lab/mcd/mcd2/>
6. 研究組織
- (1) 研究代表者  
阿部 訓也 ( ABE KUNIYA )  
独立行政法人理化学研究所・動物変異動態解析技術開発チーム・チームリーダー  
研究者番号： 40240915
  - (2) 研究分担者  
太田 聡史 ( OOTA SATOSHI )  
独立行政法人理化学研究所・情報解析技術室・専任研究員  
研究者番号： 30391890