

# 隠れマルコフモデルを用いたマウス状態の 自動判定と2値マルコフモデルによる コンソミックマウス系統の特徴付け

荒川 俊也<sup>1</sup>・高橋 阿貴<sup>2,3</sup>・田邊 彰<sup>3</sup>・柿原 聡<sup>4</sup>・木村 真吾<sup>5</sup>・杉本 大樹<sup>6</sup>  
城石 俊彦<sup>7</sup>・富原 一哉<sup>8</sup>・小出 剛<sup>2,3</sup>・土谷 隆<sup>9,10</sup>

(受付 2011年8月17日;改訂 12月14日;採択 2012年1月25日)

## 要 旨

近年、マウスの遺伝的背景と社会的行動の関係を解明するために、C57BL/6JJcl (B6)とMSM/Ms (MSM)のコンソミック系統 B6-ChrN<sup>MSM</sup>を用いた研究が進められている。コンソミック系統とは、マウス染色体21種類について、それぞれ1対の染色体のみが他系統の染色体と置き換えられた系統群のことであり、B6の1対の染色体のみがMSMのものとして置き換えられており、残りの染色体は全てB6と同一である。研究にあたっては、遺伝的背景が同等の2匹のマウスをオープンフィールドに入れ、専門家が目視でマウスの行動状態を観察および判別し、コンソミック系統毎の特徴を検討することが行われているが、多大な時間と労力を要するという問題がある。そこで、本研究では隠れマルコフモデルを用いて、マウスの行動状態の自動判定を行い、その結果に基づいて社会行動に関わる遺伝的基盤の探索を試みた。一番単純な場合である2匹のマウスが“互いに無関心”か“社会的行動”をしているかを2つの状態とした場合についてモデルを作成し、自動判定の結果の妥当性を実証する。さらに判定結果より各コンソミック系統ごとのマルコフ遷移確率を求め、これを各コンソミック系統の社会的行動を特徴づける尺度として提案する。この尺度は2次元量であり平面に表示して検討することができる。その結果、コンソミック系統としての特性を反映して各コンソミック系統は概ねB6とMSMの間を結ぶ直線の近傍に分布すること、特にChr 6C系統が特異な振る舞いをしていることなどが観察された。

キーワード：マウス、隠れマルコフモデル、無関心状態、社会的行動、マルコフ遷移

<sup>1</sup> 総合研究大学院大学 複合科学研究科統計科学専攻：〒190-8562 東京都立川市緑町 10-3

<sup>2</sup> 国立遺伝学研究所 マウス開発研究室：〒411-8540 静岡県三島市谷田 1111

<sup>3</sup> 総合研究大学院大学 生命科学研究所遺伝学専攻：〒411-8540 静岡県三島市谷田 1111

<sup>4</sup> 政策研究大学院大学：〒106-0032 東京都港区六本木 7-22-1

<sup>5</sup> 九州日本電気ソフトウェア株式会社：〒814-8567 福岡市早良区百道浜 2-4-1

<sup>6</sup> 自治医科大学 分子病態治療研究センター細胞生物研究部：〒329-0498 栃木県下野市薬師寺 3311-1

<sup>7</sup> 国立遺伝学研究所 哺乳動物遺伝研究室：〒411-8540 静岡県三島市谷田 1111

<sup>8</sup> 鹿児島大学 法文学部：〒890-0065 鹿児島市郡元 1-21-30

<sup>9</sup> 政策研究大学院大学 政策研究科：〒106-0032 東京都港区六本木 7-22-1

<sup>10</sup> 統計数理研究所 新機軸創発センター：〒190-8562 東京都立川市緑町 10-3

確率.

## 1. はじめに

本論文では、マウスの遺伝的背景と社会的行動の関係を解明するための新たな手法として、隠れマルコフモデルを用いたマウスの行動状態の自動判定と、その結果を用いた、マウス行動の2値マルコフモデルによる各コンソミック系統の特徴づけについて論ずる。

近年、マウスの遺伝的背景と社会的行動の関係を解明するために、C57BL/6Jcl (B6) と MSM/Ms (MSM) のコンソミック系統 B6-ChrN<sup>MSM</sup> を用いた研究が進められている。ここで、コンソミック系統とは、マウス染色体 21 対について、それぞれ 1 対の染色体のみが他系統の染色体と置き換えられた系統群である。この場合、B6 の 1 対の染色体のみが MSM のものと置き換えられており、残りの染色体は全て B6 と同一である。研究にあたっては、遺伝的背景が同等の同性のマウスのペアをオープンフィールドに入れ自由に探索させ、専門家が目視でマウスの行動状態を観察して判別し、コンソミック系統毎の特徴を検討することが行われているが、多大な時間と労力を要するという問題がある。そのため、人間の目視に基づく一連のプロセスが、機械により自動化できることが望ましい。

そこで、本研究では、隠れマルコフモデルを用いて、専門家が判定した少数の結果を教師データとし、マウスペアのトラッキングデータより抽出した特徴量(相対距離, 距離の変化, 相対角度, 速さの平均)を用いることによって、各時点でのマウスの行動状態が“無関心状態”と“社会的行動”の2状態のいずれであるかを自動判定することを試みてその妥当性を確認した。さらに、各ペアについて自動判定した結果を用いて、無関心状態から社会的行動に遷移するマルコフ遷移確率と、社会的行動から無関心状態に遷移するマルコフ遷移確率を計算し、遷移確率を通じて系統雌雄毎のマウスの特徴を捉え、いくつかの興味深い知見を得た。

本研究では 60 cm×60 cm の正方形のフィールド内で様々な行動を呈している、同系統および同性のマウスペアについて、10 分間の行動を撮影した、530 ペアのマウスに関する動画像を用いた(図 1)。これらの動画像は 1 秒あたり 3 フレームからなるため、1 つの動画像は 1800 フレームから構成されている。本研究では、各フレーム単位で状態を評価するために、java による動画像観察用ソフトウェアを作成した。

このソフトウェアを用いて、530 ペアの動画像のうち、50 ペアの動画像について、専門家が目視で行動を観察し、フレーム単位で状態を評価した。隠れマルコフモデルの教師データでは、

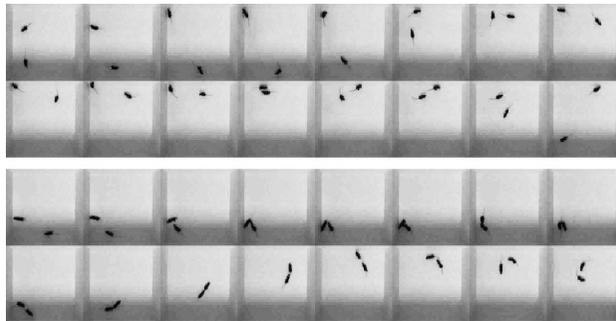


図 1. マウス動画像 (B6f01n), 上: 無関心状態の例, 下: 社会的行動の例.

この 50 ペアの評価結果を教師データとして使用し、マウスの行動状態の自動判定については 530 ペア全てについて行った。

なお、マウスの行動観察の自動化に隠れマルコフモデルを用いる研究は最近いくつか見受けられ (Carola et al., 2011; Vetrov et al., 2008), 具体的な観察対象は異なるものの, 本研究もその流れの中に位置づけることができよう。本来的には状態の判定は人間が主観的に行う部分があり, 観察者によって判定結果に違いがあることも否めない。動物観察用商用ソフトウェアの中にもそのような機能を組み込んだと主張しているものが現れつつあるようであるが (Cleversys, Inc., 2011), 観察作業を機械化するにあたっては, 判定結果の客観性・妥当性をどのように吟味して十分に担保するかは難しい問題であることに留意すべきであろう。

なお, 本論文の構成は以下の通りである。第 2 章で 2 値マルコフモデルによるマウス行動の特徴づけについて議論し, 第 3 章ではマウス行動の特徴抽出に用いた解析ソフトウェアについて説明する。第 4 章で状態推定のための隠れマルコフモデルについて扱う。第 5 章では隠れマルコフモデルによる推定の妥当性について検証する。第 6 章では, 先行研究として, Takahashi et al. (2009) における検討内容と, 隠れマルコフモデルによる判定を比較検討し, 本論で述べる手法の妥当性について述べる。第 7 章では隠れマルコフモデルによる推定結果を用いて 2 値マルコフモデルによるコンソミック系統の特徴づけを行う。第 8 章は結論である。

## 2. マウス行動のモデル化

時系列毎にマウスの行動が逐一遷移することに着目し, 本研究では行動形態をマルコフモデルで捉えることを試みる。一般的に, コンソミック系統マウスのペアの状態は, 無関心状態 (indifference), 社会的匂い嗅ぎ行動 (social sniffing), 性器嗅ぎ行動 (genial grooming), 追従行動 (following), 攻撃的追従行動 (aggressive chasing), 攻撃行動 (attacking) といった状態に分離される。この中で, 無関心状態以外の行動, すなわち, 匂い嗅ぎ行動 (sniffing), 追従行動, 攻撃的追従行動, 攻撃行動にて表されるものを社会的行動と称することとする。なお, ここで, 社会的匂い嗅ぎ行動と性器嗅ぎ行動は, 動画像観察においては区別が極めて難しいため, 両者をまとめて, 匂い嗅ぎ行動とした。本論では, マウスの行動を, 無関心状態と社会的行動の 2 状態を遷移するものと考え, 検討する。

### 2.1 2 値マルコフモデル

マウスの行動を, 無関心状態と社会的行動の 2 状態に分類した場合の状態遷移について, 2 値マルコフモデルに適合させた場合を考える。この場合, マウスは, 時系列毎に, 無関心状態と社会的行動の 2 状態を相互に遷移するか, もしくは, 同じ状態を維持するものと考えられる。いま,  $p_{00}$ ,  $p_{01}$ ,  $p_{10}$  および  $p_{11}$  を, それぞれ, 次のように定義する。

$p_{00}$  : 無関心状態から無関心状態に遷移する確率

$p_{01}$  : 無関心状態から社会的行動に遷移する確率

$p_{10}$  : 社会的行動から無関心状態に遷移する確率

$p_{11}$  : 社会的行動から社会的行動に遷移する確率

このモデルにおいては, 独立変数は 2 個となる。したがって, このモデルにおいては, コンソミックマウス毎のマルコフ遷移確率を, 2 次元平面上に描写することが可能である。いま, このモデルにおける独立変数を  $p_{01}$  と  $p_{10}$  に選ぶ。また, このマルコフ過程の定常確率を  $\mu_0$  と  $\mu_1$  とすると,  $\mu_0$  と  $\mu_1$  は, 式 (2.1) で表される。

$$(2.1) \quad \mu_0 = \frac{p_{10}}{p_{01} + p_{10}}, \quad \mu_1 = \frac{p_{01}}{p_{01} + p_{10}}$$

式(2.1)を書きなおすと、

$$(2.2) \quad p_{10} = \frac{1 - \mu_1}{\mu_1} p_{01} = \frac{\mu_0}{1 - \mu_0} p_{01}$$

となる。このことから、定常確率が同じとなる全てのマルコフ確率  $(p_{01}, p_{10})$  の集合は、原点を通る  $(p_{01}, p_{10})$  平面上に存在する、正の傾きを持つ直線上に存在することがわかる。

式(2.1)より、異なるマルコフ確率の組に対して、同じ定常確率  $(\mu_0, \mu_1)$  をとることがある。このマルコフ確率の組を、 $(p_{01}^*, p_{10}^*)$  および  $(p_{01}^{**}, p_{10}^{**})$  ( $p_{01}^* < p_{01}^{**}$ ) とする。この場合、 $(p_{01}^*, p_{10}^*)$  と  $(p_{01}^{**}, p_{10}^{**})$  の違いは、それぞれ、「無関心状態」と「社会的行動」の切り替わる回数に違いがある。同じ定常確率であっても、 $(p_{01}^*, p_{10}^*)$  と  $(p_{01}^{**}, p_{10}^{**})$  を比較すると、 $p_{01}^* < p_{01}^{**}$  であるため、 $(p_{01}^{**}, p_{10}^{**})$  は  $(p_{01}^*, p_{10}^*)$  よりも原点から正の方向に離れた位置にある。そのため、無関心状態と社会的行動が互いに切り替わる回数は増加する。

単位フレーム(1/3秒)毎に、コンソミックマウスの行動状態が、無関心状態と社会的行動が切り替わる確率は、次の式で与えられる。

$$(2.3) \quad \mu_0 p_{01} + \mu_1 p_{10} = \frac{2p_{01}p_{10}}{p_{01} + p_{10}}$$

従って、 $n$  個連続したフレームにおいて、無関心状態と社会的行動が切り替わる期待値は、次式のようになる。

$$(2.4) \quad n(\mu_0 p_{01} + \mu_1 p_{10}) = \frac{2np_{01}p_{10}}{p_{01} + p_{10}}$$

マウスペアのトラッキングデータより、マルコフ遷移確率を求める方法は次の通りである。

- (a) 単位フレーム(1/3秒)毎にマウスの行動状態をマーキングし、1800フレーム全部についてマーキングする。
- (b) (a)にてマーキングした結果を参照して、次の通り、遷移の回数をカウントする。

$n_{00}$  : 無関心状態から無関心状態への遷移の回数

$n_{01}$  : 無関心状態から社会的行動への遷移の回数

$n_{10}$  : 社会的行動から無関心状態への遷移の回数

$n_{11}$  : 社会的行動から社会的行動への遷移の回数

- (c) 式(2.5)に従って、 $p_{00}$ ,  $p_{01}$ ,  $p_{10}$ ,  $p_{11}$  を計算する。

$$(2.5) \quad p_{00} = \frac{n_{00}}{n_{00} + n_{01}}, \quad p_{01} = \frac{n_{01}}{n_{00} + n_{01}}, \quad p_{10} = \frac{n_{10}}{n_{10} + n_{11}}, \quad p_{11} = \frac{n_{11}}{n_{10} + n_{11}}$$

以上で述べた2値マルコフモデルについて、5つの特徴として纏めることができる。

- (1) 2値マルコフモデルは2つの独立した変数を持ち、それぞれ、無関心状態から社会的行動に遷移するマルコフ確率  $p_{01}$  と、社会的行動から無関心状態に遷移するマルコフ確率  $p_{10}$  と選ぶことができる。独立変数が2変数であり、各系統のマウスの行動が、それぞれ、2値マルコフモデルに従うことから、2匹のマウスの行動は、上記  $p_{01}$  と  $p_{10}$  を用いて、 $(p_{01}, p_{10})$  平面の座標(点)として表すことができる。
- (2) このマルコフ遷移行列の定常確率を  $\mu_0$  と  $\mu_1$  とすると、これらは、それぞれ、10分間(1800フレーム)中に、無関心状態および社会的行動を取る時間の割合を示している。
- (3) 同じ定常確率を持つマルコフモデルは、 $(p_{01}, p_{10})$  において、原点を通る直線上に存在する。この直線上では、原点から離れれば離れるほど、無関心状態と社会的行動の切り替えが頻繁に発生する。

- (4) 動画データからマルコフ遷移行列を推定するためには、まず、動画において、各フレーム毎のマウスの行動状態が、無関心状態と社会的行動のどちらであるかを目視で判定する。次いで、連続したフレーム毎に、マウスの行動状態に関して、無関心状態→無関心状態、無関心状態→社会的行動、社会的行動→無関心状態、社会的行動→社会的行動のそれぞれに遷移する個数を調べる。
- (5) (4)について、動画を元にしたマウスの行動状態判定は、理想的には、マウスの行動を熟知した専門家によって行われるべきであるが、本論では、機械学習の領域で用いられている、隠れマルコフモデルを適用することによって、マウスの行動状態を自動判定するシステムを開発し、適用させることで、判定が可能となる。

### 3. マウス行動解析ソフト

ここでは、マウスの動画を解析するために自作した、javaによるソフトウェアについて説明する。このソフトウェアの動作状況を図2に示す。なお、図2は、B6マウスを例として表示したものである。本ソフトは、マウスの動画を観察しながら、トラッキングデータにより得られたマウスの座標データから特徴量として抽出された4つの物理量(4.2節で述べる)と、マウスが無関心状態であるか、社会的行動であるかを時系列的に確認することができる。なお、図2において、右側のウィンドウに関して、各物理量を示す時系列データを示す箇所の上半分は隠れマルコフモデルによる推定結果であり、下半分は専門家の目視による判定結果を示している。

本ソフトウェアは、早送り、巻戻しが可能であると共に、再生速度を変更することが可能である。また、各フレーム毎にテンキーを押すことによって、テンキーに応じたマウスの状態をcsvファイルに書き込み、データとしてロギングすることができる。今回、専門家がマウスの行動を観察し、判定した際にも、本ソフトウェアを用いて、フレーム毎にマウスの動画を確認しながらマウスの状態をロギングした。

### 4. 隠れマルコフモデル

ここでは、530種類のコンソミック系統マウスの行動状態推定を行うために用いた、隠れマ

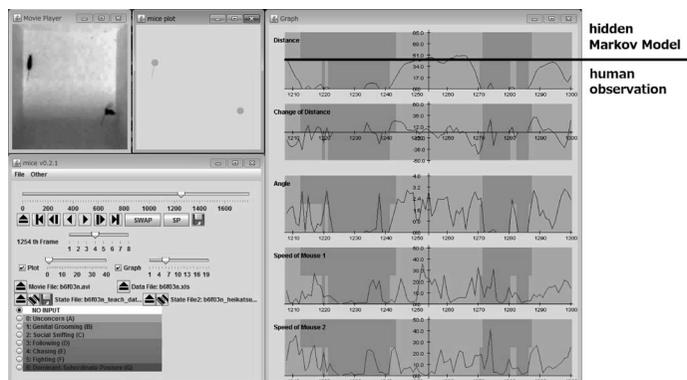


図2. 各種物理量の表示と、専門家目視判定と2値隠れマルコフモデルを比較した表示例。

ルコフモデル (hidden Markov Model) について説明する (Huang et al., 1990; Eddy, 1996, 2004; Rabiner, 1989). まず, トラッキングデータによって得られたマウスの座標を元にして, 隠れマルコフモデルを適用するために用いる, マウスの行動を表す物理量の算出方法を述べる. 次に, マウスの行動を元にして, マウスの行動状態推定をする方法, 計算式について説明する.

#### 4.1 隠れマルコフモデルの適用方法

本論で隠れマルコフモデルを導入する目的としては, コンソミック系統マウスのペアの状態を, 専門家の目視観察に基づく判定結果を学習データとした上で, トラッキングデータより抽出されたマウスの特徴的な行動を表す物理量を用いて, 観察できるシンボル系列であるマウスの行動を自動的に分類することにある. 隠れマルコフモデルに用いる教師データを準備するために, 専門家が雌雄両性の 50 種類のコンソミックマウスペアの動画像を観察し, フレーム毎に, 無関心状態, 匂い嗅ぎ行動, 追隨行動, 攻撃的追隨行動, 攻撃行動にて表される社会的行動の発生頻度を観察した. そして, トラッキングデータにより得られたマウスの座標データから特徴量として抽出された 4 つの物理量 (4.2 節で述べる) を, 隠れマルコフモデルによる解析に取り込んだ.

#### 4.2 2 値隠れマルコフモデル

いま, フレーム番号  $k(1, \dots, 1800)$  におけるコンソミックマウスの行動状態を  $s_k$  とする.  $s_k$  は, 時刻  $t$  において,

$$s_k = 0 : \text{無関心状態 (state 0)}, \quad s_k = 1 : \text{社会的行動 (state 1)}$$

の 2 状態の何れかを取ることにする. 状態  $s_k$  はマルコフ過程に基づいて更新される. 次に, 条件付き分布  $p(\mathbf{y}|s)$  について説明する.  $\mathbf{y}$  を 2 匹のコンソミックマウスの行動の特徴を示す物理量とし,  $s$  をコンソミックマウスの行動状態とする. 本論では物理量  $\mathbf{y}$  を, 次のように定義した.

##### (a) 2 匹のコンソミックマウス間の距離 (相対距離)

フレーム番号  $k$  における 2 匹のコンソミックマウス, マウス 1 とマウス 2 の重心間の距離  $L(k)[\text{cm}]$  は, 式 (4.1) に従って算出できる.

$$(4.1) \quad L(k) = \sqrt{(X_2(k) - X_1(k))^2 + (Y_2(k) - Y_1(k))^2}$$

ここで,

$X_1(k), X_2(k)$ : マウス 1 およびマウス 2 の  $x$  座標 (単位: cm)

$Y_1(k), Y_2(k)$ : マウス 1 およびマウス 2 の  $y$  座標 (単位: cm)

である.

##### (b) 2 匹のマウスの距離の変化

2 匹のマウスの距離の変化  $V_r(k) [\text{cm/sec}]$  は, フレーム番号  $k-1$  から  $k$  に遷移するときの距離の変化により求めることができる. 従って, 式 (4.2) に従って算出できる.

$$(4.2) \quad V_r(k) = (L(k) - L(k-1)) \times 3$$

ここで,  $V_r$  の単位を  $\text{cm/sec}$  とするため,  $1800[\text{frames}]$  が  $600[\text{sec}]$  に相当することから, 単位換算の目的によって距離の変化  $L(k) - L(k-1)$  に 3 を掛けている.

##### (c) 2 匹のコンソミックマウスのなす角 (相対角度)

フレーム番号  $k-1$  から  $k$  に遷移する時の, 2 匹のコンソミックマウス, マウス 1 とマウス

2 が  $x$  軸となす角をそれぞれ  $\theta_1[\text{rad}]$  および  $\theta_2[\text{rad}]$  とすると、

$$(4.3) \quad \theta_1 = \arctan \frac{Y_1(k) - Y_1(k-1)}{X_1(k) - X_1(k-1)}$$

$$(4.4) \quad \theta_2 = \arctan \frac{Y_2(k) - Y_2(k-1)}{X_2(k) - X_2(k-1)}$$

となる。この時、2 匹のコンソミックマウス、マウス 1 とマウス 2 が互いになす角(相対角度)  $\theta(k)[\text{rad}]$  は、式(4.5)および式(4.6)に従って算出できる。

$$(4.5) \quad \theta^*(k) = |\theta_1(k) - \theta_2(k)|$$

$$(4.6) \quad \theta(k) = \begin{cases} \theta^*(k) (0 \leq \theta^*(k) < \pi) \\ 2\pi - \theta^*(k) (\pi \leq \theta^*(k) < 2\pi) \end{cases}$$

(d) 2 匹のマウスの速さの平均

フレーム番号  $k-1$  から  $k$  に遷移するときの、2 匹のコンソミックマウス、マウス 1 とマウス 2 の速さをそれぞれ  $V_1[\text{cm/sec}]$  および  $V_2[\text{cm/sec}]$  とすると、

$$(4.7) \quad V_1(k) = 3\sqrt{(X_1(k) - X_1(k-1))^2 + (Y_1(k) - Y_1(k-1))^2}$$

$$(4.8) \quad V_2(k) = 3\sqrt{(X_2(k) - X_2(k-1))^2 + (Y_2(k) - Y_2(k-1))^2}$$

となる。この時、2 匹のコンソミックマウス、マウス 1 とマウス 2 の速さの平均  $v[\text{cm/sec}]$  は、式(4.9)に従って算出できる。

$$(4.9) \quad v(k) = \frac{V_1(k) + V_2(k)}{2}$$

いま、物理量  $\mathbf{y}_k$  を

$$(4.10) \quad \mathbf{y} = (L(k), \theta(k), V_r(k), v(k))$$

と表すこととする。フレーム番号  $k$  における物理量  $\mathbf{y}_k$  の分布は、マウスが無関心状態であるときは  $p(\mathbf{y}_k | s_k = 0)$  に、社会的行動のときは  $p(\mathbf{y}_k | s_k = 1)$  に従うものと考えることができる。

物理量が分布  $p(\mathbf{y}_k | s_k = 1)$  に従うとき、 $s_k$  は、マルコフ遷移行列  $\mathbf{T}$  により決定されるマルコフモデルに従い更新される。

以上が、本論で提案する、専門家の観察の代替となる、マウスの行動状態を自動判定するモデルである。

このモデルを機能させるためには、マルコフ遷移行列  $\mathbf{T}$ 、条件付き分布  $p(\mathbf{y}_k | s_k = 0)$  および  $p(\mathbf{y}_k | s_k = 1)$  が必要となる。専門家が観察によって解析した、50 種類のコンソミックマウスのデータより、マウス毎に、各フレームの状態をカウントすることによって、各マウスのマルコフ遷移確率を求めることができ、従って、マルコフ遷移行列  $\mathbf{T}$  を求めることができる。また、50 種類のマウスのトラッキングデータを基にして、各マウス状態に対する物理量を表す 4 次元のヒストグラムを構築することによって、条件付き分布  $p(\mathbf{y}_k | s_k = 0)$  および  $p(\mathbf{y}_k | s_k = 1)$  を求めることができる。

本研究では、1800×50 個の教師データ(各データは 4 個の物理量とマウス状態の 5 次元からなる)を基に、各物理量につき、最大値と最小値の間を 10 階級に等分して、マウス状態に関する条件付きヒストグラムを構築し、得られたヒストグラムを正規化することにより、条件付き分布  $p(\mathbf{y}_k | s_k = 0)$  および  $p(\mathbf{y}_k | s_k = 1)$  を求めた。

さて、コンソミックマウスペアのトラッキングデータが与えられた状況下で、フレーム毎のコンソミックマウスの行動状態  $s_1, s_2, \dots, s_{1800}$  は次のように推定できる。いま、フレーム番号  $k$  までの全てのフレームに関するコンソミックマウスの物理量の集合を  $\mathbf{Y}_k = (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_k)$  とする。まず、各コンソミックマウス毎に、1番目のフレームから最後のフレーム(1800番目のフレーム)までの物理量の集合  $\mathbf{Y}_{1800} = (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_{1800})$  を、トラッキングデータより計算する。隠れマルコフモデルの標準式に適合させるために、下記に述べる通りに事後分布  $p(s_k | \mathbf{Y}_{1800})$  を計算する。なお、 $k$  番目のフレームにおける状態は、 $p(s_k | \mathbf{Y}_{1800})$  を用いて、

$$(4.11) \quad s_k = \underset{s_k}{\operatorname{argmax}} p(s_k | \mathbf{Y}_{1800})$$

として得られるとする。

$p(s_k | \mathbf{Y}_{1800})$  を求めるために、まず最初に、 $p(s_k | \mathbf{Y}_k)$  を次式に基づいて再帰的に計算する。

$$(4.12) \quad p(s_k | \mathbf{Y}_{k-1}) = \sum_{s_{k-1}} p(s_k | s_{k-1}) p(s_{k-1} | \mathbf{Y}_{k-1})$$

$$(4.13) \quad p(s_k | \mathbf{Y}_k) = \frac{p(\mathbf{y}_k | s_k) p(s_k | \mathbf{Y}_{k-1})}{\sum_{s_k} p(\mathbf{y}_k | s_k) p(s_k | \mathbf{Y}_{k-1})}$$

式(4.12)および(4.13)を繰り返し用いることで、

$$(4.14) \quad p(s_{k-1} | \mathbf{Y}_{k-1}) \rightarrow p(s_k | \mathbf{Y}_{k-1}) \rightarrow p(s_k | \mathbf{Y}_k) \rightarrow p(s_{k+1} | \mathbf{Y}_k) \rightarrow \dots$$

と、順方向再帰的に事後分布を得ることができる。式(4.12)において、 $p(s_k | s_{k-1})$  は、マルコフ遷移行列  $T$  の要素となり、

$$p(s_k = 0 | s_{k-1} = 0) = p_{00}, \quad p(s_k = 0 | s_{k-1} = 1) = p_{10} \\ p(s_k = 1 | s_{k-1} = 0) = p_{01}, \quad p(s_k = 1 | s_{k-1} = 1) = p_{11}$$

を意味している。また、 $p(\mathbf{y}_k | s_k)$  はトラッキングデータを基に構成された4次元ヒストグラムより算出できる。

$p(s_k | \mathbf{Y}_k)$  が求められた後、式(4.15)の通りに、逆方向再帰的に  $p(s_k | \mathbf{Y}_{1800})$  を計算する。

$$(4.15) \quad p(s_k | \mathbf{Y}_{1800}) = p(s_k | \mathbf{Y}_k) \sum_{s_{k+1}} \frac{p(s_{k+1} | \mathbf{Y}_{1800}) p(s_{k+1} | s_k)}{p(s_{k+1} | \mathbf{Y}_k)}$$

以上の流れを全コンソミックシステムのマウスに適用することによって、専門家が観察していないマウスについても、社会的行動の有無を推定することが可能となる。なお、上記の計算式の詳細については、Kitagawa (1987) を参照のこと。

### 4.3 2値隠れマルコフモデルを適用した具体例

ここで、4つの物理量の計測が行われた、特定のマウス(b6f01n)について、専門家が観察した場合と、隠れマルコフモデルにより判定した場合において、javaにて作成した解析ソフトにおける解析例を示す(図2)。なお、本ソフトは、専門家が目視でマウスの状態を判定する際にも用いたものである。また、本例においては、1800フレーム全てを表示することは難しいので、一部フレーム区間のみの表示となっている。

図2において、右のウィンドウには、上から、各種物理量(相対距離、2匹のマウスの距離の変化、相対角度、マウス1の速さ、マウス2の速さ)を示していると共に、各フレーム毎における、専門家の観察による判定結果と、隠れマルコフモデルによる判定結果の比較を示している。なお、隠れマルコフモデルによる状態推定においては、2匹のマウスの速さの平均を用い

たが、本ソフトウェアにおいては、マウス 1、マウス 2 の速さをそれぞれ図示していることに留意されたい。また、マウスの状態について、薄い灰色の帯で示した箇所は無関心状態を示しており、濃い灰色の帯で示した箇所は社会的行動を示している。

## 5. 観察データと隠れマルコフモデルによる推定の比較

本章では、観察に基づいて状態を判定した結果と、隠れマルコフモデルを適用して状態推定した結果を比較し、本論で述べる手法の妥当性について検討する。比較にあたって、まず考えられる基準として、1800 フレーム中、専門家が観察した場合と、隠れマルコフモデルにより判定した場合において、無関心状態と社会的行動のそれぞれが一致した割合、すなわち一致率が挙げられる。

しかしながら、マウス行動は無関心状態が多く、専門家も隠れマルコフモデルも 9 割程度は無関心状態と判定するため、一致率は必然的に高くなる傾向にあり、単純に一致率が高いからと言って、妥当な判定が得られていると判断することはできない。したがって、様々な角度からの検討が必要となる。

### 5.1 社会的行動に着目した隠れマルコフモデルの判定の妥当性の検討

マウスの振る舞いの解析においては、社会的行動が適切に判定されることが特に重要である。そこで、マウスの社会的行動に着目して、隠れマルコフモデルの妥当性を検討する。

表 1 に、様々な観点からの、一致率の割合を示す。表 1 の意味する所は次の通りである。

- |       |  |
|-------|--|
| 第 1 列 | 専門家が観察し、社会的行動と判定したフレーム総数のうち、隠れマルコフモデルによっても社会的行動と判定されたフレーム数の割合  |
| 第 2 列 | 隠れマルコフモデルによって、社会的行動と判定されたフレーム総数のうち、専門家によっても、社会的行動と判定したフレーム数の割合 |
| 第 3 列 | 専門家によって社会的行動と判定されたフレーム数の割合                                     |
| 第 4 列 | 隠れマルコフモデルによって社会的行動と判定されたフレーム数の割合                               |
| 第 5 列 | 専門家と隠れマルコフモデル両方によって、共に社会的行動と判定されたフレーム数の割合                      |

第 1 列目と第 2 列目に着目すると、共に値が 100 であることが望ましい。しかし、実際には、第 1 列目の数値は大きく、90[%] 前後を推移しているが、第 2 列目の数値はばらつきが大きく、30[%] から 90[%] 程度の値を取っている。実際の画像を観察した結果、以下のような傾向が観察された。

- (1) 専門家が観察して社会的行動と判定した場合を基準として見ると、専門家が社会的行動と判定したフレームについては、隠れマルコフモデルによる判定においても、概ね、社会的行動と判定できている(図 3, 例は Chr 19 マウスのうち 1 種類)。
- (2) 一方、隠れマルコフモデルにより社会的行動と判定した場合を基準として見ると、隠れマルコフモデルでは社会的行動と判定されているが、専門家は社会的行動と判定していない場合がある(図 4, 例は Chr 2C マウスのうち 1 種類)。図 4 の状況においては、明らかに追隨行動を行っているが目視で判断できるが、専門家の基準では、かなりマウス間の距離が近くない場合でなければ追隨行動と判定していないため、結果的に社会的行動とは判定されない。
- (3) 専門家も隠れマルコフモデルも社会的行動として判断するようなイベントであっても、開始時点・終了時点がずれている場合があり、これが、社会的行動自体が比較的稀で

表 1. 様々な観点からの一致率の比較.

	Concordance rate of hMM when interactive behavior with human observation is regarded as true	Concordance rate of human observation when interactive behavior by hMM is regarded as true	Ratio of interactive frames judged with human observation	Ratio of interactive frames judged by hMM	Ratio of interactive frames judged with both human observation and hMM
b6f01n	83.94	56.10	7.61	11.39	6.39
b6f02n	94.44	35.66	3.00	7.94	2.83
b6f03n	91.72	68.28	9.39	12.61	8.61
b6m01n	72.58	49.72	6.89	10.06	5.00
b6m02n	86.21	41.32	3.22	6.72	2.78
b6m03n	80.56	38.41	4.00	8.39	3.22
c12cf03n	86.67	42.76	4.17	8.44	3.61
c12cf04n	89.87	54.83	8.78	14.39	7.89
c12cf05n	95.40	45.60	4.83	10.11	4.61
c12cm02n	85.71	45.09	5.06	9.61	4.33
c12cm03n	87.80	55.81	4.56	7.17	4.00
c12cm04n	82.72	50.00	4.50	7.44	3.72
c14f02n	81.94	43.38	4.00	7.56	3.28
c14m03n	82.18	51.55	5.61	8.94	4.61
c14m04n	89.06	43.85	3.56	7.22	3.17
c15f01n	94.55	48.60	6.11	11.89	5.78
c15f02n	95.24	48.00	3.50	6.94	3.33
c15f03n	89.21	43.51	7.72	15.83	6.89
c15m01n	91.14	59.75	8.78	13.39	8.00
c15m02n	94.38	65.46	13.83	19.94	13.06
c15m03n	91.94	31.67	3.44	10.00	3.17
c19f01n	93.75	44.55	5.33	11.22	5.00
c19f02n	98.54	71.81	7.61	10.44	7.50
c19f03n	94.37	55.60	7.89	13.39	7.44
c19m01n	90.98	61.11	7.39	11.00	6.72
c19m02n	96.83	82.92	36.78	42.94	35.61
c19m03n	95.79	63.08	11.89	18.06	11.39
c2cf01n	96.95	59.11	9.11	14.94	8.83
c2cf02n	92.86	37.96	6.22	15.22	5.78
c2cf03n	90.82	46.11	5.44	10.72	4.94
c2cm01n	84.14	43.88	8.06	15.44	6.78
c2cm02n	100.00	36.00	4.50	12.50	4.50
c2cm03n	95.70	50.57	5.17	9.78	4.94
c6tf01n	97.44	36.54	4.33	11.56	4.22
c6tf02n	98.50	75.14	14.83	19.44	14.61
c6tf03n	98.96	48.72	5.33	10.83	5.28
c6tm01n	93.39	61.90	19.33	29.17	18.06
c6tm02n	98.64	75.48	24.44	31.94	24.11
c6tm03n	99.20	46.79	6.94	14.72	6.89
c7tf01n	87.30	41.67	3.50	7.33	3.06
c7tf02n	92.56	75.42	13.44	16.50	12.44
c7tf03n	96.20	51.70	8.78	16.33	8.44
c7tm01n	99.35	47.68	8.61	17.94	8.56
c7tm02n	95.68	59.39	9.00	14.50	8.61
c7tm03n	92.86	58.33	5.44	8.67	5.06
msmf01n	99.18	83.07	26.94	32.17	26.72
msmf02n	99.34	65.51	16.89	25.61	16.78
msmf03n	98.64	76.95	32.72	41.94	32.28
msmm01n	96.33	73.55	21.17	27.72	20.39
msmm05n	95.27	70.54	25.83	34.89	24.61

あることと相俟って、一致率を下げる原因となる。

## 5.2 状態時系列による比較検討

前節で検討した一致率のばらつきにも拘わらず、実際、目視では、専門家と隠れマルコフモデルの判定はよく合致していることが観察された。それを確認するために、数種類のマウスペアについて、専門家と隠れマルコフモデルの判定を時系列として描いた(図 5)。この図からも、

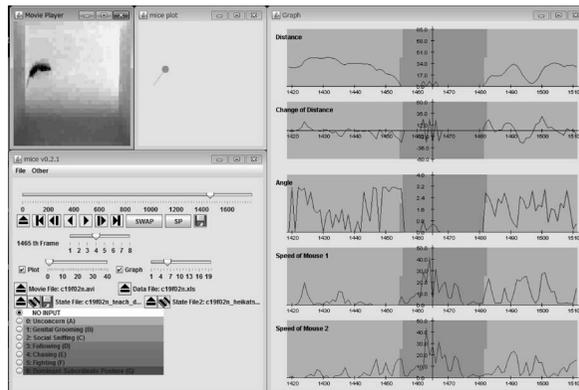


図 3. Chr 19 マウスのうちの 1 種類の判定例.

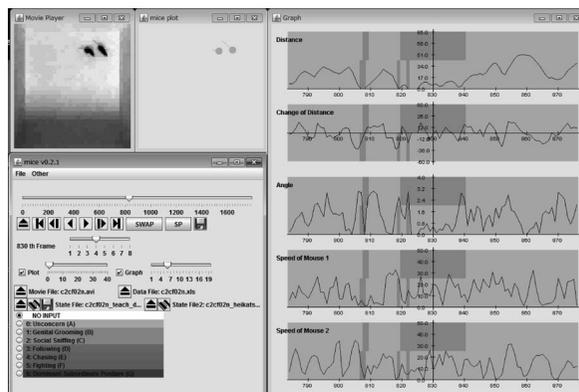


図 4. Chr 2C マウスのうちの 1 種類の判定例.

専門家と隠れマルコフモデルの判定が良く一致していることが見てとれる。両者の判定が一番異なるように見えるのは c2cm03n であるが、このマウスペアについては、隠れマルコフモデルのみが社会的行動と判定しているイベントが目立つものの、実際に目視して確認した結果、この判定は、概ね妥当であると判断した。

### 5.3 統計量の比較

教師データとして用いた 50 種類のコンソミックマウスにおいて、専門家の観察によって社会的行動と判定された時間と、隠れマルコフモデルによって社会的行動と判定された時間の相関を調べた。その結果を図 6 に示す。相関係数は  $r=0.9827$  となり、専門家の観察による判定結果と隠れマルコフモデルによる判定結果の間には高い相関関係が見られた。

### 5.4 検討結果のまとめ

本章では、隠れマルコフモデルによる社会的行動判定の妥当性を、専門家の判定との時系列グラフの比較や統計量の観点から吟味した。特に図 5 に示したように、隠れマルコフモデルに

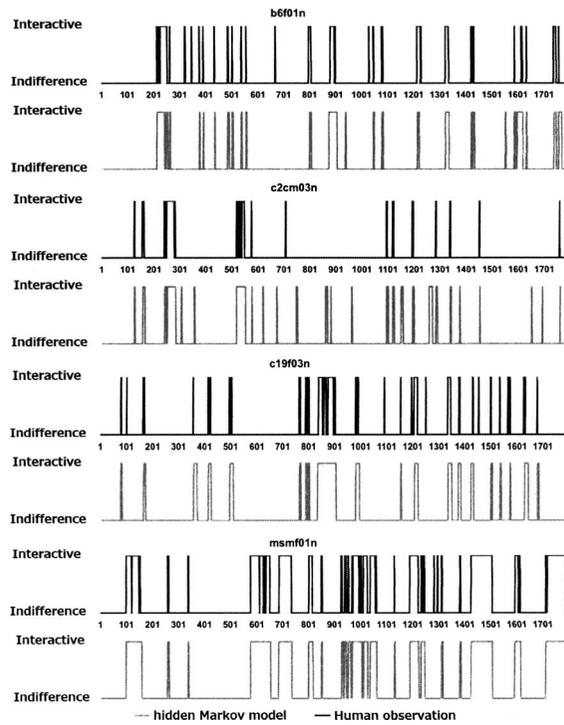


図 5. 専門家と隠れマルコフモデルの判定の時系列比較.

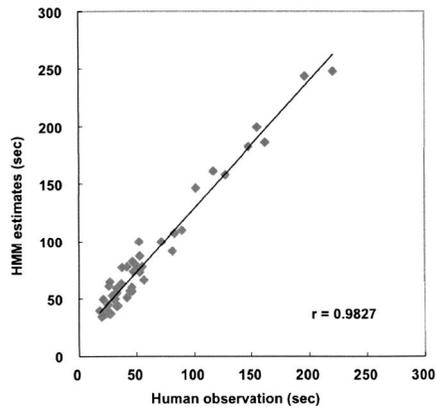


図 6. 50 種類のコンソミックマウスの社会的行動判定時間の相関.

よる社会的行動判定は、観察者による判定と傾向として十分によく一致しており、これを用いてさらに自動解析を行うに足るだけの妥当性を有するものと判断する。

## 6. 先行研究との比較

先行研究の一つである Takahashi et al. (2009)では、2匹のマウスの相対距離が12cm以下になった場合に、2匹のマウスが接触しているものと判定し、「接触時間」や「接触頻度」に基づいて各コンソミック系統の特徴を解析している。社会行動解析の観点からは、これは、2匹のマウスがお互いに近づいているときには、お互いに関心を持っているとして、社会的行動と考えると見なしていることに相当する。本節では、隠れマルコフモデルによって、距離による判定に比べてより人間の判断に近い判定が可能となることを示し、隠れマルコフモデルを用いて Takahashi et al. と同様の解析を行うことを試みる。

### 6.1 時系列による比較

図5に示したマウスと同じマウスについて、Takahashi et al. (2009)にて示されている、2匹のマウスの相対距離が12cm以下になった場合に社会的状態と判定する場合と、隠れマルコフモデルによって社会的状態と判定した場合を比較したものを図7に示す。図7を見ると、hmm01nを除いた3種類である b6f01n, c2cm03n, c19f03n, 特に b6f01n, c2cm03n については隠れマルコフモデルによるものよりも距離によるものの方が、ずっと頻繁にスパイク的に社会的状態と判定していることが見て取れる。さらに、図5の専門家による判定と比較することにより、隠れマルコフによる判定の方がより専門家の判定に近いと判断できる。実際、動画を注意深く目視観察して検証したところ、明らかに無関心状態であっても偶然距離が閾値以下となったために接触と見なされる場合などが頻繁に見受けられ、b6f01nなどで距離による判定によって



図7. Takahashi et al. (2009) の判定基準と隠れマルコフモデルの判定の時系列比較。

接触状態がスパイク状に頻繁に現れるのは、偶然、距離が近づいたためだけであることが確認された。これは、12cm という距離が、社会的行動と判定するに当たっては、やや大きめであることに起因しているものと推察される。一方、隠れマルコフモデルでは距離が離れた2個体が追跡状態にある場合をも社会的状態として判定しており、この点も単純な距離による接触判定に比して優れた点であるといえる。

これを受けて、以下では Takahashi et al. (2009) と同様の解析を隠れマルコフモデルによる判定によって行う。すなわち「社会的行動の長さ」と「社会的行動の頻度」によるコンソミック系統の特徴付けを試みる。前者は Takahashi et al. (2009) においては接触時間による解析に対応し、後者は接触頻度による解析に対応する。

## 6.2 コンソミック系統毎の接触時間の総和に関する比較検討

Takahashi et al. (2009) では、コンソミック系統毎の接触時間の総和を、図8(a)のように示している。なお、Takahashi et al. (2009) においては、Chr 18 のデータは欠損している。一方、隠れマルコフモデルを用いて判定した結果より、コンソミック系統毎に社会的行動と判定された時間の総和を、図8(b)に示す。図8より、Chr 4, Chr 11, Chr 13, Chr 14 を除いて類似した傾向があることが見て取れる。なお、図8において、\*は、B6 に比べた場合、コンソミック系統毎の接触時間の総和および社会的行動と判定された時間の総和に関して、5%有意であることを示している。

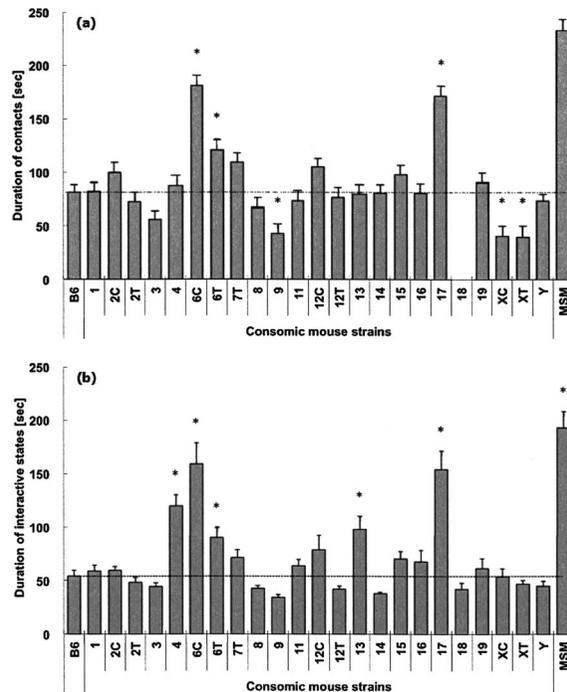


図8. 接触時間の総和 (a) と社会時間判定時間の総和 (b) の比較。

### 6.3 コンソミック系統毎の接触回数に関する比較検討

次に, Takahashi et al. (2009)で示されている, 2匹のマウスの接触回数(図9(a))についても, 社会的行動の観点から検討する. なお, Takahashi et al. (2009)においては, Chr 18のデータは欠損している. 本論では, 無関心状態と社会的行動が切り替わる期待値が, 2匹のマウスの接触回数と関連するものと考え. 式(2.4)と, 無関心状態から社会的行動に切り替わる期待値は, 無関心状態と社会的行動が切り替わる期待値の丁度半分になることを考えると, 1800フレームから構成される1個のマウス動画画像において, 無関心状態と社会的行動が切り替わる期待値は, 式(6.1)で表される.

$$(6.1) \quad \frac{1}{2} \frac{2 \times 1800 p_{01} p_{10}}{p_{01} + p_{10}} = \frac{1800 p_{01} p_{10}}{p_{01} + p_{10}}$$

式(6.1)によって, 1800フレーム中における, 無関心状態と社会的行動が切り替わる期待値が計算でき, コンソミック系統および雌雄毎の期待値の平均と分散を計算することができる. その結果を図9(b)に示す. なお, 図9において, \*は, B6に比べた場合, 2匹のマウスの接触回数および無関心状態と社会的行動が切り替わる期待値に関して, 5%有意であることを示しており, #は, 同じコンソミック系統の雌雄について比較した場合, 5%有意であることを示している. 図8より, 無関心状態と社会的行動が切り替わる期待値は, 接触回数の約半分であることがわかる. これは, 6.2節の図7からも推察されることである. 図8における無関心状態と社会的行動が切り替わる期待値の間には, 社会的行動時間の場合ほど, コンソミック系統ごとに大きな違いがないように見受けられる.

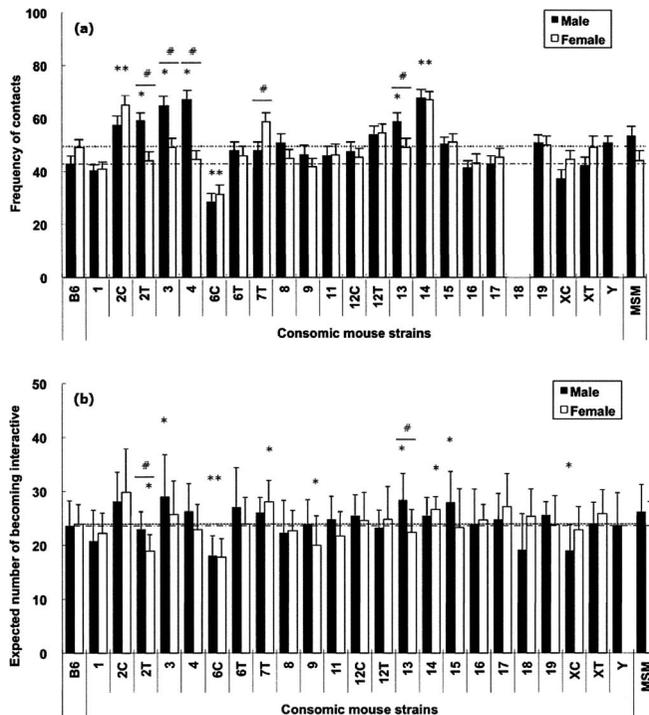


図9. 接触頻度 (a) と無関心状態・社会的行動が切り替わる期待値 (b) の比較.

## 7. マルコフ遷移確率に基づくコンソミックマウスの特徴抽出

隠れマルコフモデルにより、各ペアの動画像について、各時点において無関心状態と社会的行動のいずれであるかを判定した時系列が得られる。これを人間の観察者が同定した状態であるとみなして、各ペアについて(2.5)式にしたがってマルコフ遷移確率を計算することができる。以下、マルコフ遷移確率に基づいてコンソミック系統雌雄別の差異を検討する。なお、各ペアのマルコフ遷移確率を推定するにあたっては Baum-Welsh 法を用いることも考えられるが、ここでは、隠れマルコフモデルが人間の観察者を模倣的に実現しているという立場に立って標記のような形でマルコフ遷移確率を推定した。

### 7.1 マルコフ遷移確率によるコンソミックマウスの特徴の視覚化について

530 種類のコンソミックマウスについて、上記の方法で得られたマルコフ推定確率を 2 次元平面に図示したものを図 10 に示す。なお  $x$  軸は、無関心状態から社会的行動に遷移する確率 ( $=p_{01}$ )、 $y$  軸は、社会的行動から無関心状態に遷移する確率 ( $=p_{10}$ ) である。

図 10 から、 $p_{10}$  は  $p_{01}$  に比べて約 10 倍の値を取ることがわかる。また、Chr 6C, Chr 17, MSM は、 $p_{01}$  は広く分布しているが、 $p_{10}$  は低い値に集中していることがわかる。 $p_{10}$  に関して、530 種類の平均値は 0.016 であり、 $p_{01}$  に関しては平均値は 0.126 である。

### 7.2 コンソミックマウスの系統別のばらつき方

次に、雌雄系統別でマウスのマルコフ遷移確率の分布の特徴について検討する。いま、 $\mathbf{p} = (p_{01} \quad p_{10})^T$ 、平面上の任意の点を  $\mathbf{q} = (q_{01} \quad q_{10})^T$ 、さらに、 $D(\mathbf{p}|\mathbf{q})$  を、 $\mathbf{p}$  によるマルコフ列と  $\mathbf{q}$  によるマルコフ列の Kullback-Leibler ダイバージェンスレートとすると、

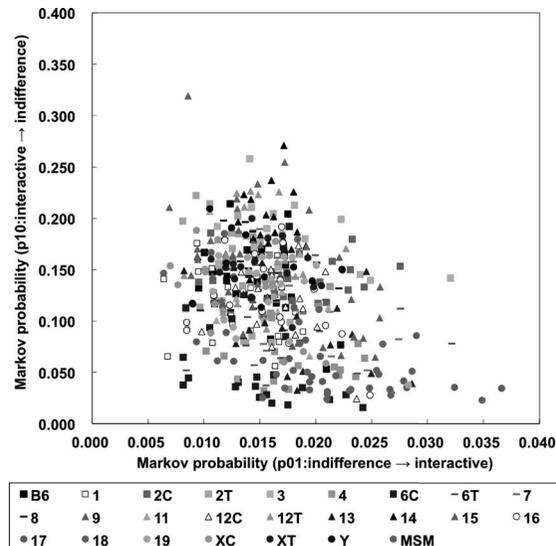


図 10. 全コンソミックマウスのマルコフ確率の分布。

$$(7.1) \quad D(\mathbf{p}|\mathbf{q}) = \mu_0 p_{00}^* \log \frac{p_{00}^*}{q_{00}^*} + \mu_0 p_{01}^* \log \frac{p_{01}^*}{q_{01}^*} + \mu_1 p_{10}^* \log \frac{p_{10}^*}{q_{10}^*} + \mu_1 p_{11}^* \log \frac{p_{11}^*}{q_{11}^*}$$

$$= \mu_0 p_{00}^* \log \frac{p_{00}^*}{1 - q_{01}^*} + \mu_0 p_{01}^* \log \frac{p_{01}^*}{q_{01}^*} + \mu_1 p_{10}^* \log \frac{p_{10}^*}{q_{10}^*} + \mu_1 p_{11}^* \log \frac{p_{11}^*}{1 - q_{10}^*}$$

が成立する (Rached et al. 2004). ここで,  $\mu_0$ ,  $\mu_1$  は, それぞれ,  $\mathbf{p}$  によるマルコフ過程で状態 0, 1 をとるマルコフ定常確率である.

フィッシャー情報量行列は,

$$\mathbf{I} = \frac{\partial^2 D(\mathbf{p}|\mathbf{q})}{\partial \mathbf{q}^2} \Big|_{\mathbf{p}} = \begin{pmatrix} \frac{\partial^2 D(\mathbf{p}|\mathbf{q})}{\partial p_{01}^2} & \frac{\partial^2 D(\mathbf{p}|\mathbf{q})}{\partial p_{01} \partial p_{10}} \\ \frac{\partial^2 D(\mathbf{p}|\mathbf{q})}{\partial p_{10} \partial p_{01}} & \frac{\partial^2 D(\mathbf{p}|\mathbf{q})}{\partial p_{10}^2} \end{pmatrix}$$

となり, 長さ  $N$  の状態系列から (2.5) 式にしたがって推定された遷移確率は, 平均 0, 分散共分散行列  $\mathbf{I}/N$  の正規分布に近似的に従うと考えられるので,

$$(7.2) \quad (\mathbf{q} - \mathbf{p})^T \mathbf{I} (\mathbf{q} - \mathbf{p}) = 4$$

にはおよそ 95% の個体が入ると考えられる. なお, フィッシャー情報量行列  $\mathbf{I}$  は対角行列となるので, (7.2) の楕円の形状は水平になる. 各系統雌雄別に,  $p_{01}$  および  $p_{10}$  の平均値を中心として (7.2) の楕円を描いたものを, 図 11 に示す. なお, 図 11 において, 各プロットは, 個体毎の  $p_{01}$  および  $p_{10}$  の値を示している. また, 図 11 の菱形のプロットおよび黒色の楕円は雄を示し, 四角のプロットおよび灰色の楕円は雌を示している.

図 11 から, 系統毎に, マウスの分布のばらつき方が異なっており, 各系統において, 大部分の個体の  $p_{01}$  および  $p_{10}$  の値が, (7.2) の楕円に収まっている場合と (7.2) の楕円外にもかなりの個体が存在する場合とに大別される. 特に, Chr 2C, Chr 2T, Chr 8, Chr 9, Chr 12T, Chr 14, Chr 16, Chr 19, Chr XT, Chr Y については, 雌雄共に (7.2) の楕円に収まっている. このことは, 上記系統の個体については系統内の個体差が小さいことを意味する. 一方, Chr 4, Chr 6C, Chr 6T, Chr 13, Chr 17 については, かなりの割合の個体が (7.2) の楕円外に存在している. 上記系統については, マルコフ確率の値の個体差が大きいと見られる.

また, 定性的な観点から, 楕円形状や各個体の  $p_{01}$  および  $p_{10}$  の値のばらつき方より, マルコフ確率について雌雄差が生じていることもわかる. 特に, Chr 13, Chr 15, Chr 16 については, 雌雄の楕円の中心間距離が他のコンソミック系統に比べて大きい上, (7.2) の楕円形状について, 雌雄で形状が異なっている ( $x$  方向が長径であるか否か) ことがわかる. このことは, マルコフ確率の観点から, 上記系統については雌雄差が存在することを示唆している.

さらに, Chr 13, Chr 15 については, 有意な性差が現れており, 雄は楕円が MSM 寄りであり, 雌は楕円が B6 寄りであるという傾向が把握できる. 目視観察の範囲では, これらの系統では雄のみが攻撃行動を示している. このように, (7.2) の楕円形状の観点から, 社会行動の質の違いを抽出することが出来たとも言える.

### 7.3 マルコフ遷移確率の平均値からみた B6 マウスおよび MSM マウスとの比較

図 12 に, 各系統に関するコンソミックマウスのマルコフ遷移確率の平均値を, 雌雄別に示す. 図 12 より, それぞれのコンソミック系統のマルコフ遷移確率の分布には以下の特徴があることが見て取れる.

- (i) 全体としては, B6 と MSM を結ぶ右下がりの直線に沿って (雄の場合) あるいは同直線に平行な帯内 (雌の場合) に分布する.
- (ii) B6 の周辺に多くが分布する.

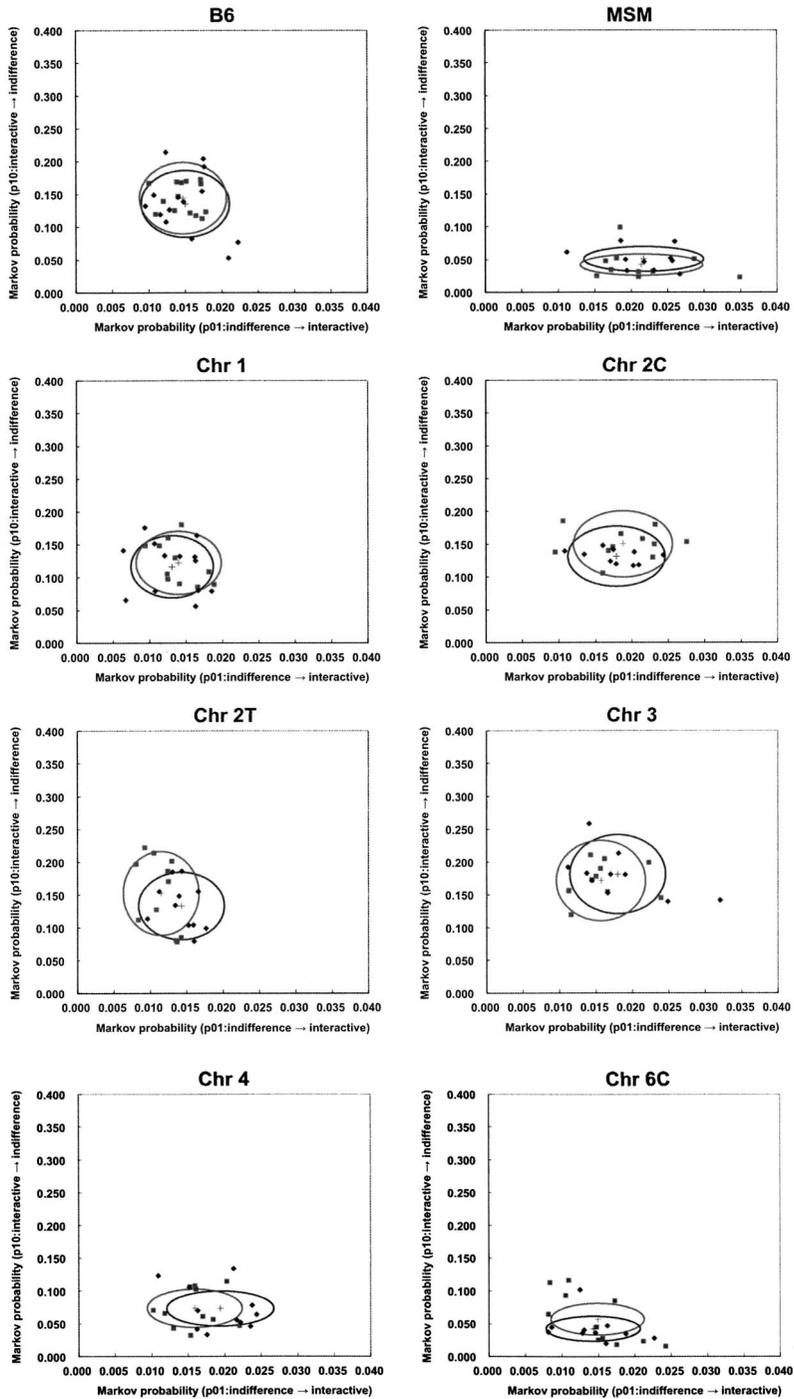


図 11. コンソミック系統毎のマルコフ確率分布とばらつきを示す楕円。

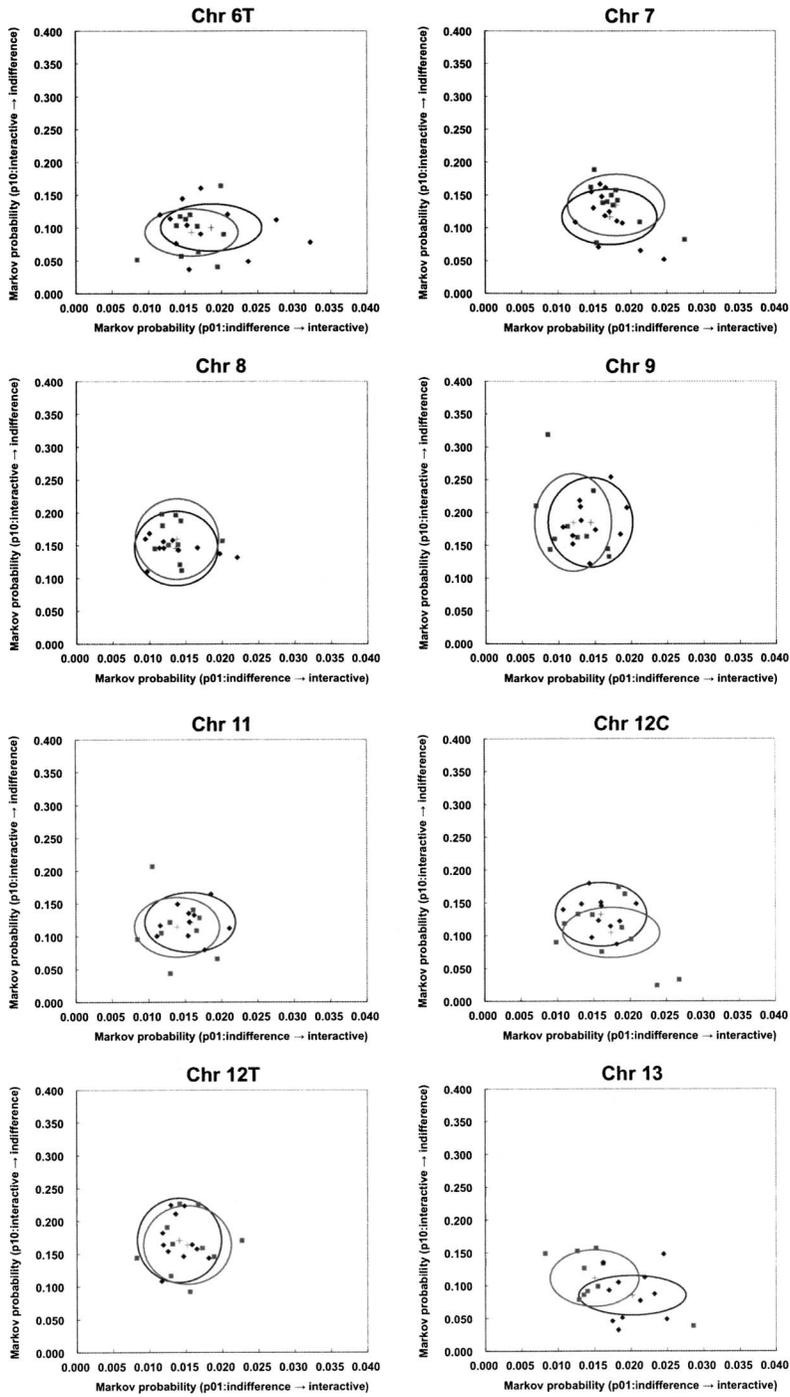


図 11. 続々.

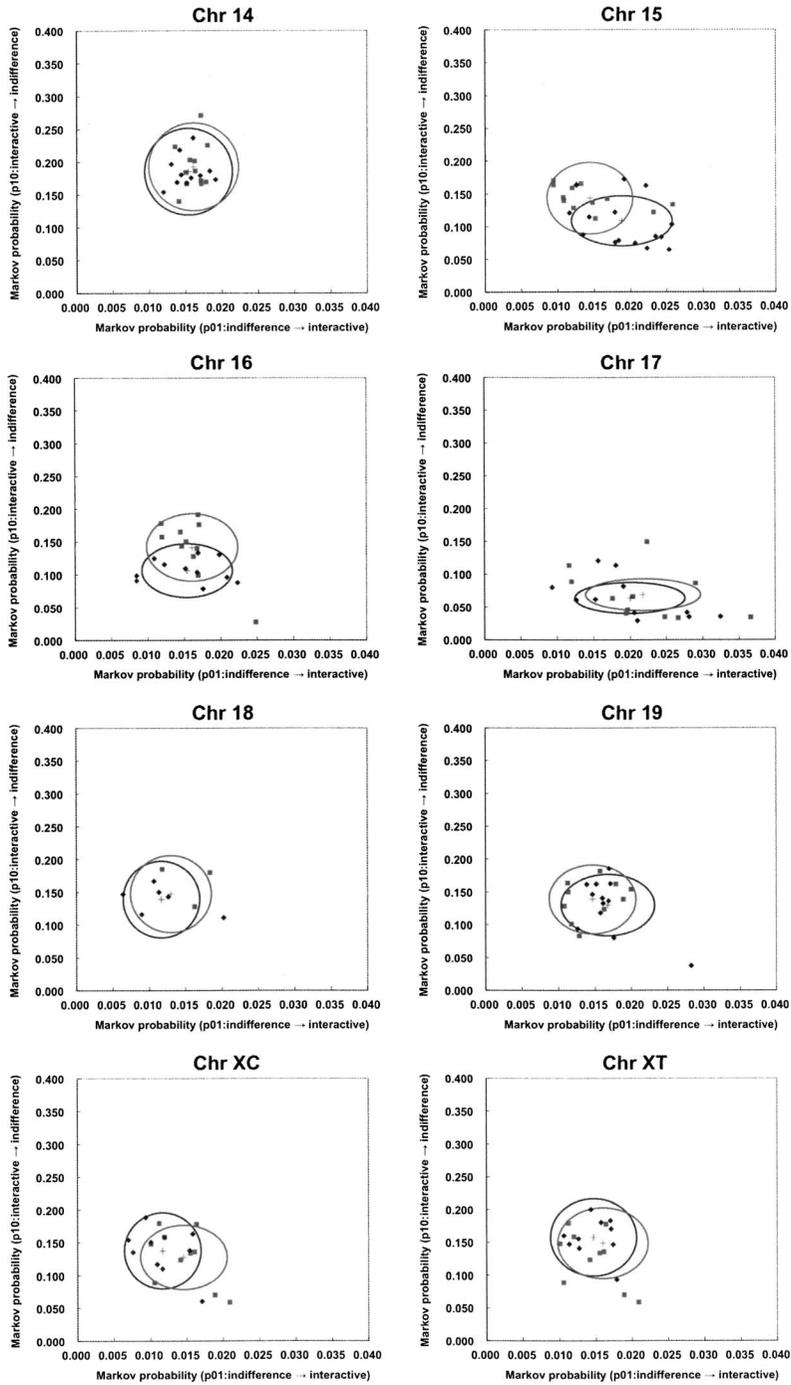


図 11. 続き.

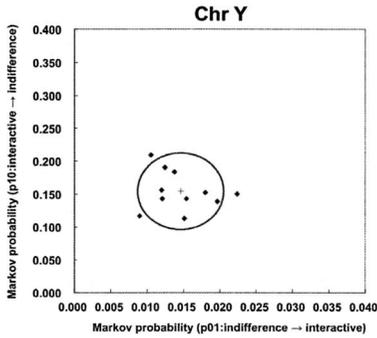


図 11. 続き.

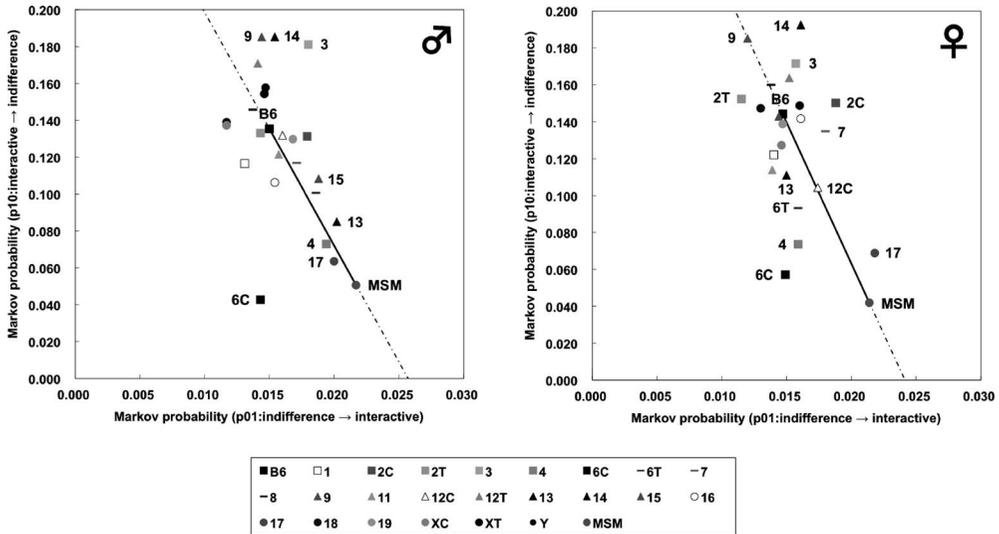


図 12. コンソミック系統雌雄別のマルコフ平均値の分布.

特に上の特性(i)は、特に雄マウスについて顕著であり、遺伝形質的にコンソミック系統が B6 と MSM の“混合”であることを考えると、その事実が定量的に現れているとも考えられ、興味深い。この関係は、 $p_{10}$  から  $p_{01}$  が(あるいは  $p_{10}$  から  $p_{01}$  が)決定されるような生物学的メカニズムが存在する可能性を示唆する。しかし、全てのコンソミック系統に関して(i)もしくは(ii)に合致してはいない。特に、雄の Chr 6C は、上の(i) (ii)の何れの場合にも合致しておらず、 $p_{01}$  については B6 とほぼ同じ値で  $p_{10}$  については MSM より若干小さい値に近い値であり、直線から大きく外れている。雌についても直線からの外れ方は大きい。この事実は Chr 6C の特別な役割を示唆するとも考えられる。

図 12 の結果を踏まえて、雌雄毎に、B6 とそれ以外の系統を比較するために、permutation test による有意差検定を行った。その結果を表 2 に示す。

表 2 より、雄マウスについては、B6 と比べて、Chr 3, Chr 9, Chr 15 では、有意水準 5% で有意差が見られ、また、Chr 4, Chr 6C, Chr 13, Chr 14, Chr 17, MSM では、有意水準 1% で

表 2. permutation test の結果.

	Permutation p-value					
	Strain difference (vs B6)				Sex difference	
	Male		Female			
Parental strains						
B6	-		-			0.784
MSM	0.0001	**	0	**		0.5709
Consomic strains						
1	0.4079		0.1223			0.8007
2C	0.1582		0.0382	*		0.0515
2T	0.857		0.007	**		0.0199
3	0.0246	*	0.0322	*		0.512
4	0.0028	**	0	**		0.1111
6C	0.0001	**	0	**		0.5276
6T	0.096		0.0008	**		0.4249
7	0.2307		0.0141	*		0.4073
8	0.6563		0.3001			0.3397
9	0.0167	*	0.0226	*		0.2036
11	0.645		0.0766			0.3838
12C	0.7109		0.0204	*		0.208
12T	0.1023		0.3085			0.6589
13	0.0031	**	0.0295	*		0.0244
14	0.0083	**	0.0004	**		0.5448
15	0.0447	*	0.9816			0.0315
16	0.1273		0.4272			0.052
17	0.0007	**	0	**		0.802
19	0.4509		0.8508			0.3921
XC	0.0838		0.3997			0.1659
XT	0.359		0.4301			0.5187
Y	0.4497		-			-

For strain differences, \*\*P<0.01, \*P<0.05. For sex differences, a P<0.05.

有意差が見られる。また、雌マウスについては、B6 と比べて、Chr 2C, Chr 3, Chr 7, Chr 9, Chr 12C, Chr 13 では、有意水準 5% で有意差が見られ、Chr 2T, Chr 4, Chr 6C, Chr 6T, Chr 14, Chr 17, MSM では、有意水準 1% で有意差が見られる。

図 12 には、B6 と比べて有意差が見られたコンソミックマウスについてラベリングをしてある。permutation test により B6 マウスと有意差が見られたマウスの  $x$ - $y$  平面上の位置が、B6 マウスと明確に離れた位置関係にあることが、視覚的にも把握することができる。

## 8. まとめ

本論では、隠れマルコフモデルを用いて、マウスの行動について、無関心状態と社会的行動という 2 状態の自動判定を行うとともに、自動判定の結果よりマルコフ遷移確率を求め、2 次元平面に表してコンソミック系統毎の差異を検討した。2 匹のマウスの相対距離を基にして接触と判定しこれによって社会的行動と無関心状態を弁別する従来手法と比較した結果、隠れマルコフモデルによる状態判定がより専門家の目視による判定と近く妥当であることが示された。これは、隠れマルコフモデルによる状態推定法が、従来、専門家の目視に頼っていた判定方法に代わる方法として、有用な手法となりうることを示唆している。また、マルコフ遷移確率を平面上に図示することによって各コンソミック系統の差異を検討した結果、概ねコンソミック系統が B6 と MSM を結ぶ直線上あるいはその直線に沿った帯の中に存在すること、が示された。これは、B6 と MSM の混合性というコンソミック系統の由来を考えると興味深い。また、Chr 6C 系統は、直線近傍に存在しない特異的な例外である。この事実についてはさらに検討を重ねていきたい。本研究では 2 状態に焦点を絞ったが、無関心状態、匂い嗅ぎ行動、追隨行動の 3 状態からなる 3 状態のモデルや、更に攻撃行動等を取り入れた 4 状態以上のモデルについての隠れマルコフモデルの構築と検証は今後の課題である。

### 補足：実験に使用したマウスについて

ここでは、実験に使用したマウスについて、コンソミック系統の作成方法を説明すると共に、本論の実験に用いたマウスの育成方法について説明する。

#### A.1 コンソミック系統について

2つの系統が1つの染色体全領域について異なった由来を持ち、他の遺伝的背景が共通である場合、これらの2つの系統は互いにコンソミックな状態にあり、これらの系統はコンソミック系統と呼ばれる。マウスは、ゲノム中に19本の常染色体とXまたはY染色体が存在し、合計で21種類の染色体をもつ。コンソミック系統群は、これら21種類の各染色体について2系統間で置き換えが行われている。一本の染色体を与える系統を染色体供与系統、受け取る方を染色体受容系統と呼ぶ。

#### A.2 マウス亜種間コンソミック系統の作成とその応用

国立遺伝学研究所は、東京都臨床医学総合研究所・疾患モデル開発センターとの共同研究により、受容系統に標準的な実験系近交系であるC57BL/6Jcl(B6)系統、供与系統として日本産野生マウス由来の近交系であるMSM系統を用いたコンソミック系統の作製を行っている。この新しいマウス系統は亜種間コンソミック系統と呼ばれている。これまでの研究からB6系統とMSM系統には大きな遺伝的差異や多様な表現型の違いが観察されている。亜種間コンソミック系統を用いた解析により、疾患や様々な表現型に関連する遺伝子の探索など、医学・生物学の様々な分野に利用できると期待されている。

#### A.3 飼育方法

C57BL/6Jcl(B6)は日本クレア株式会社より購入し、国立遺伝学研究所にて飼育したものである。また、日本産亜種 *Mus musculus molossinus* に由来する近交系統であるMSM/MSM(MSM)およびB6-ChrN<sup>MSM</sup>コンソミック系統は、国立遺伝学研究所にて育成・飼育したものである。全てのコンソミック系統は、一本の染色体のみMSMより移したものである以外、B6と全く同じ遺伝的背景を持っている。しかし、Chr 2, Chr 6, Chr 7, Chr 12の4本の染色体については、染色体全体をMSM系統に置き換えることができなかつたために、2つの系統(CとT)として分けて作成した。C系統とは染色体のセントロメア側半分以上をMSMにおきかえており、T系統はテロメア側の半分以上が置き換わっている。全マウスは生後10週のものであり、適宜、食糧と水を与えた。これらのマウスは、室温が23 ± 2℃にコントロールされ、適宜、12時間周期で明暗を繰り返す室内で飼育された。これらの飼育に関しては、国立遺伝学研究所のガイドラインおよび動物実験委員会(Institutional Animal Care and Use Committee: IACUC)により是認されている手続きに沿って遂行された。

### 参 考 文 献

- Carola, V., Mirabeau, O. and Gross, C. T. (2011). Hidden Markov model analysis of maternal behavior patterns in inbred and reciprocal hybrid mice, *PLoS ONE*, **6**(3), e14753, doi:10.1371/journal.pone.0014753.
- Cleversys, Inc. (2011). SocialScanTop, <http://www.cleversysinc.com/docs/software-SocialScan.pdf>
- Eddy, S. R. (1996). Hidden Markov models, *Current Opinion in Structural Biology*, **6**, 361–365.
- Eddy, S. R. (2004). What is a hidden Markov model?, *Nature Biotechnology*, **22**(10), 1315–1316.
- Huang, X. D., Ariki, Y. and Jack, M. A. (1990). *Hidden Markov Models for Speech Recognition*,

- Edinburgh University Press, Edinburgh.
- Kitagawa, G. (1987). Non-Gaussian state-space modeling of nonstationary time series, *Journal of the American Statistical Association*, **82**, 1032–1063.
- Rached, Z., Alajaji, F. and Campbell, L. L. (2004). The Kullback-Leibler divergence rate between Markov sources, *IEEE Transactions on Information Theory*, **50**(5), 917–921.
- Rabiner, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition, *Proceedings of IEEE*, **77**(2), 257–286.
- Takahashi, A., Tomihara, K., Shiroishi, T. and Koide, T. (2009). Genetic mapping of social interaction behavior in B6/MSM consomic mouse strains, *Behavior Genetics*, **40**, 366–376.
- Vetrov, D., Kropotov, D., Konushin, A., Zarayskaya, I. and Anokhin, A. (2008). Automatic segmentation of mouse behavior using hidden Markov model, *Measuring Behavior*, **2008**, 241–242.

## Automatic Segmentation of Mouse States Using Hidden Markov Model and Characterization of Mouse Strain Using 2-state Markov Model

Toshiya Arakawa<sup>1</sup>, Aki Takahashi<sup>2,3</sup>, Akira Tanave<sup>3</sup>, Satoshi Kakihara<sup>4</sup>  
Shingo Kimura<sup>5</sup>, Hiroki Sugimoto<sup>6</sup>, Toshihiko Shiroishi<sup>7</sup>, Kazuya Tomihara<sup>8</sup>  
Tsuyoshi Koide<sup>2,3</sup> and Takashi Tsuchiya<sup>9,10</sup>

<sup>1</sup>Department of Statistical Science, The Graduate University for Advanced Studies

<sup>2</sup>Mouse Genomics Resource Laboratory, National Institute of Genetics

<sup>3</sup>Department of Genetics, The Graduate University for Advanced Studies

<sup>4</sup>National Graduate Institute for Policy Studies

<sup>5</sup>NEC Software Kyushu, Ltd.

<sup>6</sup>Division of Genetic Therapeutics, Center for Molecular Medicine, Jichi Medical University

<sup>7</sup>Mammalian Genetics Laboratory, National Institute of Genetics

<sup>8</sup>Faculty of Law, Economics and Humanities, Kagoshima University

<sup>9</sup>National Graduate Institute for Policy Studies

<sup>10</sup>The Institute of Statistical Mathematics

Research has been conducted to find the relation between genetic character and social behavior of mice based on the consomic mouse B6-ChrN<sup>MSM</sup>, which is consomic strain between C57BL/6JJcl (B6) and MSM/Ms (MSM). A pair of genetically identical mice were put in a square open field and an expert identified segments of interactive states between the pair such as “indifference”, “sniffing”, “following”, and so on. If we had numerous pairs to observe, this task would have been a major obstacle and a laborious and time-consuming process. In this study, we automated segmentation utilizing the hidden Markov model. Specifically, we developed a model that recognized whether the pair was “indifferent” or “interactive.” The plausibility of the obtained segmentation was carefully examined from various aspects. Based on the segmentation by the hidden Markov model, we calculated the Markov transition probability of each consomic strain. Markov transition probability is a two-dimensional quantity, and we proposed this quantity as a characteristic of social behavior of each consomic strain. Through a plot of Markov transition probability of consomic strains in a 2-dimensional plane, we observed the following two interesting facts: (i) consomic strains are located along the line connecting B6 and MSM reflecting the nature of consomic strains, (ii) of these, the Chr 6C consomic strain showed an exceptional singular behavior.