

# A Markov Transition Score for Characterizing Interactive Behavior of Two Animals and its Application to Genetic Background Analysis of Social Behavior of Mouse

Toshiya Arakawa<sup>1</sup>, Aki Takahashi<sup>2</sup>, Akira Tanabe<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>9</sup>, Takashi Tsuchiya<sup>10</sup>

<sup>1</sup>Department of Statistical Science, The Graduate University for Advanced Studies (SOKENDAI), Tokyo, Japan.

<sup>2,3,9</sup>Mouse Genomics Resource Laboratory, National Institute of Genetics (NIG), Shizuoka, Japan.

<sup>2,3,9</sup>Department of Genetics, SOKENDAI, Shizuoka, Japan. <sup>4,10</sup>National Graduate Institute for Policy Studies, Tokyo, Japan.

<sup>3</sup>NEC Software Kyushu, Ltd., Fukuoka, Japan.

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

<sup>7</sup>Mammalian Genetics Laboratory, NIG, Shizuoka, Japan.

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

<sup>10</sup>The Institute of Statistical Mathematics, Tokyo, Japan.

<sup>1</sup>arakawa@ism.ac.jp, <sup>2</sup>aktakaha@lab.nig.ac.jp, <sup>3</sup>atanabe@lab.nig.ac.jp, <sup>4</sup>skakihara@gmail.com

<sup>5</sup>shingo.k-m-r.0715@voice.ocn.ne.jp, <sup>6</sup>hisugimo@lab.nig.ac.jp, <sup>7</sup>tshirois@lab.nig.ac.jp

<sup>8</sup>tomihara@leh.kagoshima-u.ac.jp, <sup>9</sup>tkoide@lab.nig.ac.jp, <sup>10</sup>tsuchiya@grips.ac.jp

## 1. Introduction

Characterizing interactive behavior of two animals is an important and interesting subject of research in social behavior study of animals. In this paper, we propose a score called Markov transition score to characterize interactive behavior of two animals, and apply this score to analyze social behavior of mouse.

Takahashi et al. [1][2] conducted a genetic study of mice to understand a mechanism by which individuals develop differences in social behavior. The study was based on (C57BL/6J(B6)-based) inter-subspecific consomic strains established from MSM/Ms (MSM) and C57BL/6J. Here, consomic strain is a strain where one chromosome of C57BL/6J is replaced with the same chromosome of MSM/Ms. Comparing social interactive behaviors of each consomic strain with C57BL/6J and MSM/Ms, we may well speculate the role of each chromosome in determining social interacting nature of mouse. In their study, two littermates with the same consomic strain, sex, aged 10 weeks were introduced into a novel new open field for 10 minutes and their behavior was recorded through a video tracking system (Image SI; O'Hara & Co., Tokyo, Japan; see Figure 1), which was based on a public domain image processing and analysis program, NIH Image, from the National Institute of Health (USA). The location of two mice is recorded every 1/3 second so that a video tracking data consists of 1800 frames. As shown in Table 1, a total of 530 video data was collected, where approximately 10 video data are available for each consomic strain and sex. In Takahashi et al. [1][2], a genetic mapping was conducted based on total duration time and frequency of contact of each consomic strain and sex.

We will analyze the same set of data from a different viewpoint utilizing a two-state Markov model[3]. We consider that the pair is in either “indifference” state or “interactive” state, and that the state sequence obeys to a Markov process. The model has two parameters, i.e., the transition probabilities from “indifference” to “interactive”, and from “interactive” to “indifference.” Therefore, we can represent each pair or each consomic strains as a point on a 2D plane. Looking the behavior of the consomic mice pairs through this model, we will demonstrate that a mixture nature of consomic strains are “quantitatively” clearly seen with the scattering plots of B6, MSM and consomic strains on the 2D plane. In the analysis, we need to identify the state of the pair at each tracking time point. This is an intelligent task which ideally is to be done by a human expert. In this study, we employed a hidden Markov model to automate this task. This paper is organized as follows. We introduce the method of application of Markov model and hidden Markov model in Section 2. In Section 3, we demonstrate the analyzed result. Section 4 is conclusion.

**Ethical statement** : In this study, All procedures were conducted in accordance with the NIG (National Institution of Genetics) guidelines and approved by the Institutional Committee for Animal Care and Use.

## 2. Markov Transition Score

In this section, we explain the two-state Markov model. The location of the two mice is recorded every 1/3 second in the duration of 10 minutes and therefore we have a total of 1800 tracking time points for each video data. We consider that a mice pair is either in one of the two states, namely, “indifference” and “interactive” at each tracking time point. We mean by the state “interactive” that the two mice are interacting each other like sniffing, genital grooming, following, etc., whereas by the state “indifference” that the two mice pay little attention to each other. We consider that, when moving from one tracking time point to the next, the transition between the two states occurs according to the Markov process described as below.

Let  $p_{00}, p_{01}, p_{10}$  and  $p_{11}$  be, respectively, the transition probability from “indifference” to “indifference”, from “indifference” to “interactive”, “interactive” to “indifference” and “interactive” to “interactive”. There are just two independent parameters in this model. We choose  $p_{01}$  and  $p_{10}$  as independent parameters. Let  $(\mu_0, \mu_1)$  be the stationary probability of this Markov process. It is easy to see that

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

This means that all Markov models with the same stationary probability lie on a ray emanating from the origin in  $(p_{01}, p_{10})$  plane. Then, what is the difference between two Markov models with the same stationary probability? They are different in the number of switching between the two states. The two models exist on the same ray on  $(p_{01}, p_{10})$  plane. Intuitively, switching occurs more often in the model farther from the origin, though the two models have the same stationary probability. The probability of occurrence of switching per step (a unit tracking interval) is given by the following formula:

$$\mu_0 p_{01} + \mu_1 p_{10} = \frac{2p_{01}p_{10}}{p_{01} + p_{10}}$$

Estimation of Markov transition probability from the tracking data of a pair is done as follows.

- Mark the state of the pair at every time step.
- Count the number of transitions  $n_{00}$ : the number of transitions of “indifference”  $\rightarrow$  “indifference”,  $n_{01}$ : “indifference”  $\rightarrow$  “interactive”,  $n_{10}$ : “interactive”  $\rightarrow$  “indifference” and  $n_{11}$ : “interactive”  $\rightarrow$  “interactive”
- Estimate  $p_{00}, p_{01}, p_{10}, p_{11}$  according to the following formulas:

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

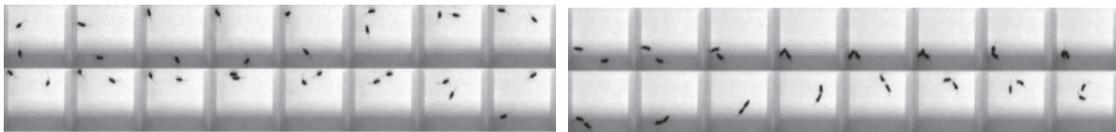


Figure 1. Sequential snapshots of a mouse pair (left: indifference, right: interactive).

Table 1.

(Consumic) Mice	B6	Chr 1	Chr 2C	Chr 2T	Chr 3	Chr 4	Chr 6C	Chr 6T	Chr 7	Chr 8	Chr 9	Chr 11
Male	14	13	10	11	11	11	10	12	13	11	11	10
Female	11	11	11	11	11	11	11	11	11	10	10	10
Total	28	24	21	22	22	22	21	23	24	21	21	20
(Consumic) Mice	Chr 12C	Chr 13	Chr 14	Chr 15	Chr 16	Chr 17	Chr 18	Chr 19	Chr XC	Chr XT	Chr Y	MSM
Male	11	11	11	11	16	11	6	12	10	10	11	10
Female	11	10	10	12	12	11	3	10	10	10	0	10
Total	22	21	21	23	28	22	9	22	20	20	11	20

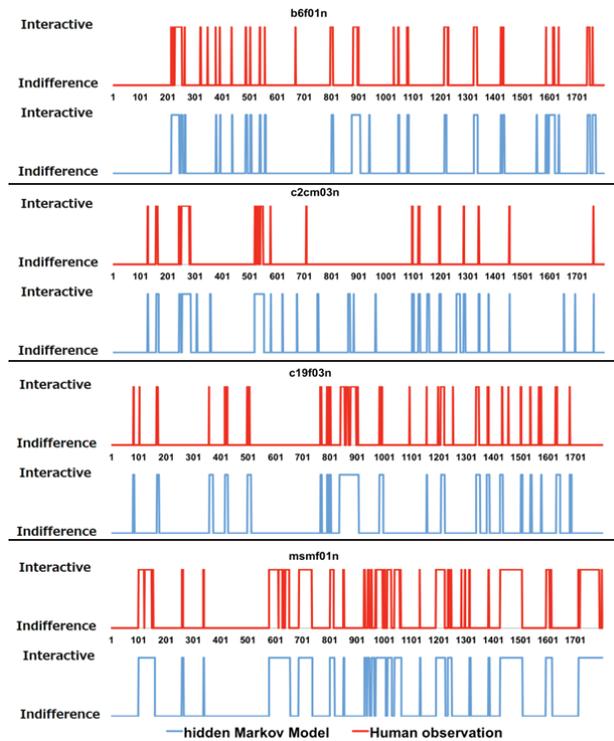


Figure 2. Comparison of results by time-series.

Though this procedure appears simple, marking of the two states watching the recorded video of a mice pair is a highly intelligent task, and ideally this should be done by a human expert. However, it is difficult to do so if we have so many video data to analyze. We developed a system which automates this task utilizing the hidden Markov model (HMM)[4][5][6][7]. We omit the details of HMM, but show the time series of human judgement and the judgement of HMM for 4 pairs (Figure 2). It is seen that HMM judgment agree fairly well with human observation. In c2cm03n, HMM tends to take more interactive states than human observation. We checked the video and confirmed that the judgment by HMM is reasonable. A total of 530 mice pair data are available, and we calculated  $p_{01}$  and  $p_{10}$  for each pair with HMM. About ten data are available for each consomic strain and sex, and we define the *Markov transition score (MTS) of a consomic strain and sex as the average of ( $p_{01}, p_{10}$ ) over all pairs of that consomic strain and sex*. MTS of each consomic strain and sex is shown in Figure 3.

### 3. Results

From Figure 3, we see that MTS of male consomic strains are either (i) located around B6, or (ii) located along in the downward-sloping straight line segment connecting B6 and MSM, except for C6c. From this figure, we observe the following:

1. The genetic mixture nature of consomic strain is indeed reflected “quantitatively” in Markov transition probability plane. This is the most interesting and important observation we made through this analysis.
2. The chromosomes associated with the consomic strains belonging to (i) seem to have little influence on interactive behavior, while the second one do have strong effects. It is interesting that that one replacement of chromosome changes so drastically interactive behavior of mice.
3. The scattering plot also suggests that the Markov probability from “interactive” to “indifference” is determined uniquely when Markov probability from “indifference” to “interactive” (from “interactive” to “indifference”) was determined.

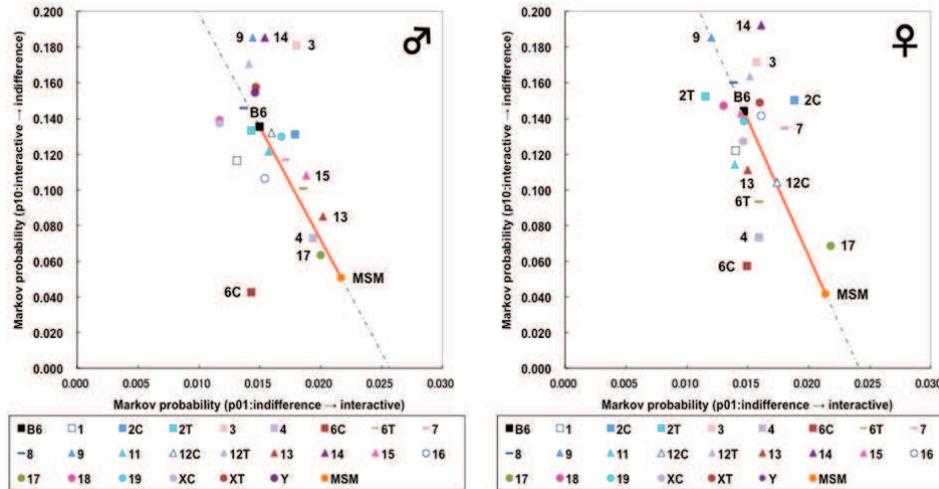


Figure 3. Markov transition score.

4. MTS of C6c is located apart from the straight line connecting B6 and MSM. This position of C6c makes a good contrast with other chromosomes in the following meaning. Basically, we observe “higher” Markov transition probability of “indifference” to “interactive” when a consomic strain has “lower” Markov transition probability of “interactive” to “indifference”. The strain C6c violates this rule: transition probability of “indifference” to “interactive” remains almost the same in spite of that transition probability of “interactive” to “indifference” is lowered. This suggests a special function of chromosome C6c in exposition of interactive behavior of mouse, and definitely deserves further study.

#### 4. Conclusion

In this study we proposed the Markov transition score to capture interactive behavior of two animals and applied the score to analyze the role of each chromosome on social behavior of mouse. We think this score can be useful in many situations including human interaction behavior.

#### References

1. Takahashi, A., Nishi, A., Ishii, A., Shiroishi, T., Koide, T. (2008) Systematic analysis of emotionality in consomic mouse strains established from C57BL/6J and wild-derived MSM/Ms. *Genes, Brain and Behavior* **7**, 849–858.
2. Takahashi, A., Tomihara, K., Shiroishi, T., Koide, T. (2009). Genetic Mapping of Social Interaction Behavior in B6/MSM Consomic Mouse Strains. *Behav. Genet.* **40**, 366–376.
3. Arakawa, T. et al. (2012). Automatic Segmentation of Mice States Using Hidden Markov Model and Characterization of Mouse Strain with 2-state Markov Model. *Proceedings of the Institute of Statistical Mathematics* **60**(1) (Japanese, in press).
4. Huang, X.D. Ariki, Y., Jack, M.A. (1990). *Hidden Markov Models for Speech Recognition*. Edinburgh University Press, Edinburgh.
5. Eddy, S.R. (1996). Hidden Markov Models. *Current Opinion in Structural Biology* **6**, 361-365.
6. Eddy, S.R. (2004). What is a hidden Markov model? *Nature Biotechnology*, **22**(10), 1315-1316.
7. Rabiner, L.R. (1989). A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. *Proc. IEEE*, **Vol. 77**(2), 257-228.