

# A New Data-Mining Method to Search for Behavioral Properties That Induce Alignment and Their Involvement in Social Learning in Medaka Fish (*Oryzias Latipes*)

Takashi Ochiai<sup>1</sup>, Yuji Suehiro<sup>1,2</sup>, Katsuhiko Nishinari<sup>3</sup>, Takeo Kubo<sup>1</sup>, Hideaki Takeuchi<sup>1\*</sup>

**1** Department of Biological Sciences, Graduate School of Science, The University of Tokyo, Tokyo, Japan, **2** Department of Physiology, Tokyo Women's Medical University School of Medicine, Tokyo, Japan, **3** Research Center for Advanced Science and Technology, The University of Tokyo, Tokyo, Japan

## Abstract

**Background:** Coordinated movement in social animal groups via social learning facilitates foraging activity. Few studies have examined the behavioral cause-and-effect between group members that mediates this social learning.

**Methodology/Principal Findings:** We first established a behavioral paradigm for visual food learning using medaka fish and demonstrated that a single fish can learn to associate a visual cue with a food reward. Grouped medaka fish (6 fish) learn to respond to the visual cue more rapidly than a single fish, indicating that medaka fish undergo social learning. We then established a data-mining method based on Kullback-Leibler divergence (KLD) to search for candidate behaviors that induce alignment and found that high-speed movement of a focal fish tended to induce alignment of the other members locally and transiently under free-swimming conditions without presentation of a visual cue. The high-speed movement of the informed and trained fish during visual cue presentation appeared to facilitate the alignment of naïve fish in response to some visual cues, thereby mediating social learning. Compared with naïve fish, the informed fish had a higher tendency to induce alignment of other naïve fish under free-swimming conditions without visual cue presentation, suggesting the involvement of individual recognition in social learning.

**Conclusions/Significance:** Behavioral cause-and-effect studies of the high-speed movement between fish group members will contribute to our understanding of the dynamics of social behaviors. The data-mining method used in the present study is a powerful method to search for candidate factors associated with inter-individual interactions using a dataset for time-series coordinate data of individuals.

**Citation:** Ochiai T, Suehiro Y, Nishinari K, Kubo T, Takeuchi H (2013) A New Data-Mining Method to Search for Behavioral Properties That Induce Alignment and Their Involvement in Social Learning in Medaka Fish (*Oryzias Latipes*). PLoS ONE 8(9): e71685. doi:10.1371/journal.pone.0071685

**Editor:** Johan J. Bolhuis, Utrecht University, Netherlands

**Received:** December 4, 2012; **Accepted:** July 6, 2013; **Published:** September 6, 2013

**Copyright:** © 2013 Ochiai et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This work was supported by the NIBB (National Institute for Basic Biology) Cooperative Research Program (10–104) and the Ministry of Education, Culture, Sports, Science, and Technology Grant-in-Aid for Scientific Research (B). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: takeuchi@biol.s.u-tokyo.ac.jp

## Introduction

Forming a group and living gregariously enable group members to more efficiently find food and escape predation [1]. Social animals utilize adaptive information associated with food or predators rapidly and economically via “social learning”, which is defined as learning that is influenced by observing or interacting with other members [2]. Social learning is observed in a variety of animal species, including colonial and noncolonial insects [3,4], fish [5], birds [6], and mammals [7].

The key process in social learning is “information transfer” from informed members to other members [8–10]. The information transfer process has been studied extensively in social insects such as honeybees and ants. Honeybee workers perform waggle dancing to inform other workers in the hive of the distance and direction of a food source [11,12]. Ants communicate with each other using pheromones as chemical signals [4,13]. Some fish species that do not seem to have symbolic or chemical communication also undergo social learning. In guppies and

golden shiners, untrained individuals (observers) that observe and learn from trained fish (demonstrators) can learn the route to a feeding area by shoaling with the demonstrator [14,15]. In these species, a small number of demonstrators can lead a larger group to food, either through social facilitation of foraging movements or by eliciting coordinated behavior [5,16]. Few studies, however, have examined the types of movement demonstrator fish use to induce alignment of the observer fish that facilitates social learning. To study this issue, we used small fish, medaka (*Oryzias latipes*), which exhibit prominent coordinated and cohesive movement [17]. Here, we first demonstrated that the formation of a group by medaka fish facilitates visual associative learning with food information.

In the present study, we used our newly established and versatile data-mining method that enabled us to search for candidate explanatory variables that may affect an objective variable using a large dataset to search for candidate behaviors that induce inter-individual interaction. The results of the data-mining using time-series coordinate data of individual medaka fish led us to

hypothesize that high-speed motion induces the alignment movement of other fish. Finally, we examined this hypothesis in a behavioral paradigm for visual food learning and demonstrated that high-speed movements of the demonstrators facilitate social learning by eliciting alignment by the observers.

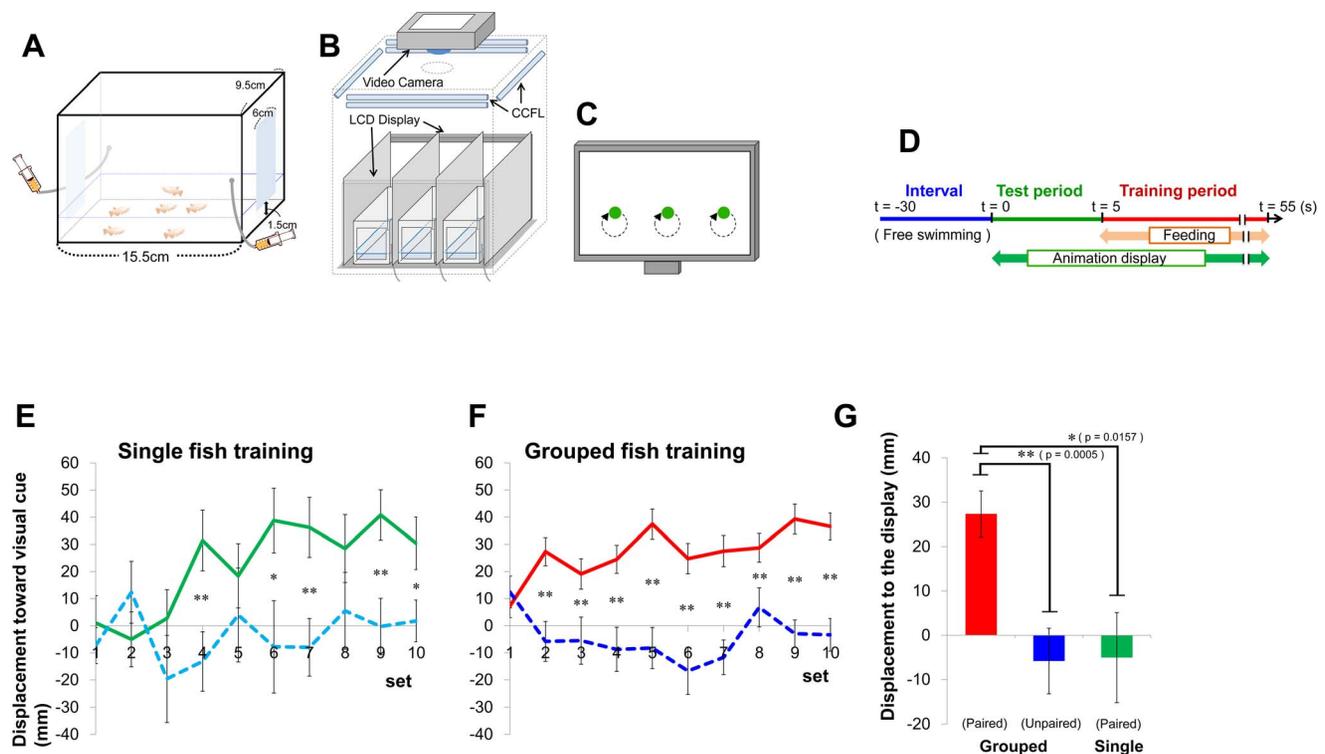
## Results

### Facilitation of learning acquisition in grouped medaka fish

To determine whether medaka fish are able to learn to associate a visual cue with a food reward, we established a new behavioral paradigm. An animated movie was displayed as a visual cue using an LCD monitor (an LCD monitor was located on each side of the tank, Fig. 1). During a 5-s test period, the fish were exposed to the visual cue at one of the two sides of the tank (Fig. 1B). Following the test period, food was supplied to the side of the tank at which the visual cue was displayed (training period). To prevent fish from forming spatial memory based on the location of the food dispensers, we alternated the sides used for the test-training trials.

We defined “one set” as two test-trainings, one trial at one side and the other trial at the other side.

First, we compared the learning acquisition between the paired group and unpaired group using a single fish. In the paired group, for which visual cues were paired with food, the visual cue was immediately followed by food supplied to the side of the tank displaying the visual cue. In the unpaired group, for which visual cues were not paired with food, the food was supplied during the interval (Fig. 1D) and the visual cue and food were not temporally associated. An unpaired group is generally used as a negative control in an associative learning paradigm. Figure 1E shows the learning acquisition of both groups in terms of displacement towards the visual cue during the test period (5 s). In the paired group, the displacement was significantly increased and reached a mean of ~40 mm. In the unpaired group, the displacement did not increase. The difference in the displacement between the two groups was statistically significant within 4 training sets (7 and 8 trials). Thus, medaka fish can learn to associate visual cues with food in this behavioral paradigm.



**Figure 1. The novel behavioral assay to assess the training efficiency between grouped and solitary medaka fish. (A) The test-training tank.** The tank comprises opaque walls with two transparent windows permitting the fish to see the LCD displays on each side of the tank. The two feed dispensers and two LCD displays are fixed at each side of the tank. **(B) Apparatus used for visual learning.** The whole apparatus is covered with a black box whose ceiling has a video camera that records the fish. Cold cathode fluorescent lamps were used for lighting. Three test-training tanks were placed in each apparatus. Brine shrimp ( $n = 120$ ) were supplied as a reward for each training trial. **(C) Visual stimulus.** An animation (circular motion of a 2.5-cm diameter green circle) was presented for each tank. **(D) Experimental procedure for testing and training.** Each procedure comprised a test and training period. The inter-trial interval was 245 s. The control group was fed the same number of shrimp as the training group, but they were fed during the inter-trial interval. **(E-G) Quantitative analysis of fish motion in response to the visual presentation.** We measured how long the individual fish moved towards the display during the test period (5-s). We calculated the distance moved (Mean  $\pm$  s.e.m.) for each set, which comprised two successive training trials. The total number of the sets was 10 (20 trials). **(E) Single fish training.** The training group is represented by the green line ( $n = 16$ ) and the control group is represented by the blue dashed line ( $n = 7$ ). Statistically significant differences were detected in the 4<sup>th</sup>, 6<sup>th</sup>, 7<sup>th</sup>, 9<sup>th</sup>, and 10<sup>th</sup> training sets (t-test,  $p < 0.05$ , see Table S1 for actual p values). **(F) Grouped fish training.** The training group is represented by the red line (12 groups, 72 individuals) and the control group is represented by the blue dashed line (6 groups, 36 individuals). Statistically significant differences were detected in all training sets except the 1<sup>st</sup> set (t-test,  $p < 0.01$ ). **(G) Comparison between grouped and solitary fish.** We compared the distance moved of grouped and single fish training for the 2<sup>nd</sup> training set. A significant difference was detected (Dunnnett method,  $p < 0.05$ ). The paired group: visual cues were paired with food. The unpaired group: visual cues were not paired with food.

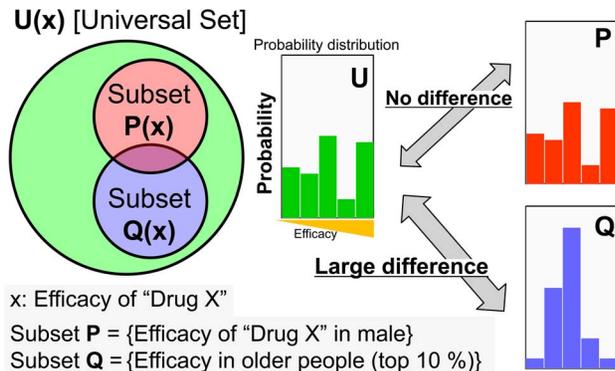
doi:10.1371/journal.pone.0071685.g001

To investigate the effect of group formation on learning acquisition in this paradigm, we performed the behavioral test using groups of six fish (Fig. 1F). The distance moved reached a similar level as when using a single fish in the paired group. A significant difference in the displacement was detected between the paired and unpaired groups within 2 training sets (3 and 4 trials). The displacement of the grouped fish during set 2 was significantly greater than that of the single fish (Fig. 1G), indicating that grouped medaka fish learn to respond to the visual cue more rapidly than single fish, suggesting that social learning provides a process whereby medaka more rapidly learn adaptive behavior.

### A novel data-mining method to search for explanatory variables that may affect an objective variable

We then searched for candidate behaviors that may affect alignment using time-series coordinate data of individual medaka fish. For this purpose, we established a versatile and simple data-mining method to search for candidate explanatory variables that may affect an objective variable using a large dataset. Here, we show a virtual experiment as an example (Fig. 2), whose purpose is to search for candidate explanatory variables that may affect the efficacy ( $x$ : an objective variable) of “Drug X”. A dataset of the drug efficacy in all the subjects was accompanied by explanatory variables such as sex, body weight, and age.

First, we define  $\mathbf{U}(\mathbf{x})$  as the universal set whose elements are the drug efficacy ( $x$ ) of all people. Next, we prepared subsets generated by conditions, which are defined based on the explanatory variables. For example, we can prepare two subsets (male and female) based on the explanatory variable “sex”. We can also prepare any number of subsets based on the explanatory variable “age”, such as a subset with age rank in the top 10% of all, subset  $\mathbf{Q}(\mathbf{x})$ . If the probability distribution (histogram shape) of the subset of males,  $\mathbf{P}(\mathbf{x})$ , was not different from that of  $\mathbf{U}(\mathbf{x})$  [the universal set], the explanatory variable “sex” might not affect the drug



**Figure 2. Outline of the calculation procedure used for the novel data-mining method.** The purpose of data-mining is to search for candidate explanatory variables that may affect objective variables. The data-mining procedure comprises three steps; (1) Generation of  $\mathbf{U}$  [the universal set] of an objective variable ( $x$ ). This example shows that the objective variable is the efficacy of “Drug X”. (2) Generation of subsets from the universal set by the conditions that are defined based on explanatory variables. For example, subset  $\mathbf{P}(\mathbf{x})$  is generated by the condition “male” based on the explanatory variable “sex”, while subset  $\mathbf{Q}(\mathbf{x})$  is generated by the condition “older people (top 10%)” based on the explanatory variable “age”. (3) Identification of candidate explanatory variables, by measurement and visualization of the difference in the probability distribution (histogram shape) of a subset from the universal set,  $\mathbf{U}(\mathbf{x})$ . This example shows that the explanatory variable “age” may affect the drug efficacy, while “sex” does not. doi:10.1371/journal.pone.0071685.g002

efficacy (Fig. 2). If the probability distribution of the subset  $\mathbf{Q}(\mathbf{x})$  was different, the condition “age rank in the top 10%” might affect the drug efficacy (Fig. 2). Thus, by assessing the difference in the probability distribution of a subset from that of  $\mathbf{U}(\mathbf{x})$ , we can estimate the explanatory variables that may affect the drug efficacy.

Here, we applied Kullback-Leibler divergence (KLD) to assess differences in the probability distribution of subsets from the universal set. Given two datasets,  $\mathbf{U}$  and  $\mathbf{Q}$  and the discrete probability distributions,  $\mathbf{U}(i)$  and  $\mathbf{Q}(i)$ , the similarity between  $\mathbf{U}$  and  $\mathbf{Q}$  is given as the distance between  $\mathbf{U}(i)$  and  $\mathbf{Q}(i)$ . KLD is not a true metric, and the word “distance” is used in a more general sense, but it is commonly used as a measure of the similarity between two distributions [18,19]. The KLD of  $\mathbf{Q}(i)$  from  $\mathbf{U}(i)$  is defined as:

$$\text{KLD}(\mathbf{Q}, \mathbf{U}) = \sum_i \left( \mathbf{Q}_i \times \log_2 \frac{\mathbf{Q}_i}{\mathbf{U}_i} \right) \quad (1)$$

We chose the KLD from information theory, because KLD is applicable irrespective of the form of the probability distribution [20,21,22,23]. We can measure the difference of the subset from the universal set, even if the probability distribution is not a unimodal distribution and/or it has considerable variation, as shown in Fig. 2. Further, a basic property of KLD is that it is not symmetric, i.e., in general.

$$\text{KLD}(\mathbf{Q}, \mathbf{U}) \neq \text{KLD}(\mathbf{U}, \mathbf{Q}) \quad (2)$$

After preparing as many subsets as possible, which can be generated based on explanatory variables, subsets with high KLD scores can be screened. Visualization of the probability distribution (histogram shape) of subsets with high KLD scores enables us to make a reasonable assumption about the relation between explanatory and objective variables. The G-test is used to evaluate the statistical significance between  $\mathbf{Q}(\mathbf{x})$  and  $\mathbf{U}(\mathbf{x})$ . P-values were calculated from g-value and the chi-square distribution.

$$\text{G value} = 2 \sum_i n(\mathbf{Q}_i) \cdot \log_2 \left( \frac{\mathbf{Q}_i}{\mathbf{U}_i} \right) \quad (3)$$

$$n(\mathbf{Q}_i) = \mathbf{Q}_i \cdot n(\mathbf{Q}(\mathbf{x}))$$

### Search for behaviors eliciting alignment movement using a data-mining method based on KLD

To search for behaviors that induce inter-individual interactions, we investigated the movement of individual fish within a group (6 fish) under free-swimming conditions without presenting a visual cue (Fig. 1D). It is noteworthy that, under this condition, fish interacted with each other without external visual signals. We measured the x-y coordinates of six individually marked fish (Fish  $i$ :  $i = 1 \sim 6$ , individual identification number) for 30-s at 1-s intervals. The velocity vector of Fish  $i$  is defined as:

$$\begin{pmatrix} v_x(i,t) \\ v_y(i,t) \end{pmatrix} \triangleq \begin{pmatrix} x(i,t+1) \\ y(i,t+1) \end{pmatrix} - \begin{pmatrix} x(i,t) \\ y(i,t) \end{pmatrix} \quad (4)$$

To evaluate the inter-individual interactions of two fish, we measured the orientation difference (OD) between each combination of two fish (Fig. 3A). The  $OD_{i,j,t}$  of the two fish (Fish i and Fish j) is defined as:

$$od_{i,j,t} \triangleq \arccos \frac{\begin{pmatrix} v_x(i,t) \\ v_y(i,t) \end{pmatrix} \begin{pmatrix} v_x(j,t) \\ v_y(j,t) \end{pmatrix}}{\left| \begin{pmatrix} v_x(i,t) \\ v_y(i,t) \end{pmatrix} \right| \left| \begin{pmatrix} v_x(j,t) \\ v_y(j,t) \end{pmatrix} \right|} \quad (5)$$

The OD range is  $[0, 180]$ . Alignment is represented by a low OD ( $0 \sim 30^\circ$ ), while anti-alignment is represented by a high OD ( $150 \sim 180^\circ$ ).

We then searched for candidate explanatory variables that may affect the OD using the data-mining method described in the previous section. Using the OD of two arbitrary individuals ( $t = 1 \sim 30s$ , 1s interval) as an element, we prepared  $\mathbf{U(od)}$  [the universal set].  $\mathbf{U(od)}$  comprised 4950 elements for the dyadic relationship ( $11 \text{ groups} \times {}_6C_2 \times 30 \text{ frames}$ ). The probabilities of all 6 intervals in  $\mathbf{U(od)}$  were essentially the same (Fig. 3C, all combinations), indicating that the six fish did not continuously show significant inter-individual interactions as a group (6 fish).

We then searched for the conditions of fish position and movement in which a pair of fish from the 6-fish group showed transient and local inter-individual interactions. We defined eight

explanatory variables concerning absolute and relative position and movement in a focal pair of fish (Fig. 3B).

$X_{\text{large}}$ : Larger distance from X-axis.

$$x_{\text{large}} \triangleq \text{large}(|x_i(t)|, |x_j(t)|) \quad (6)$$

$X_{\text{small}}$ : Shorter distance from X-axis.

$$x_{\text{small}} \triangleq \text{small}(|x_i(t)|, |x_j(t)|) \quad (7)$$

$Y_{\text{large}}$ : Larger distance from Y-axis.

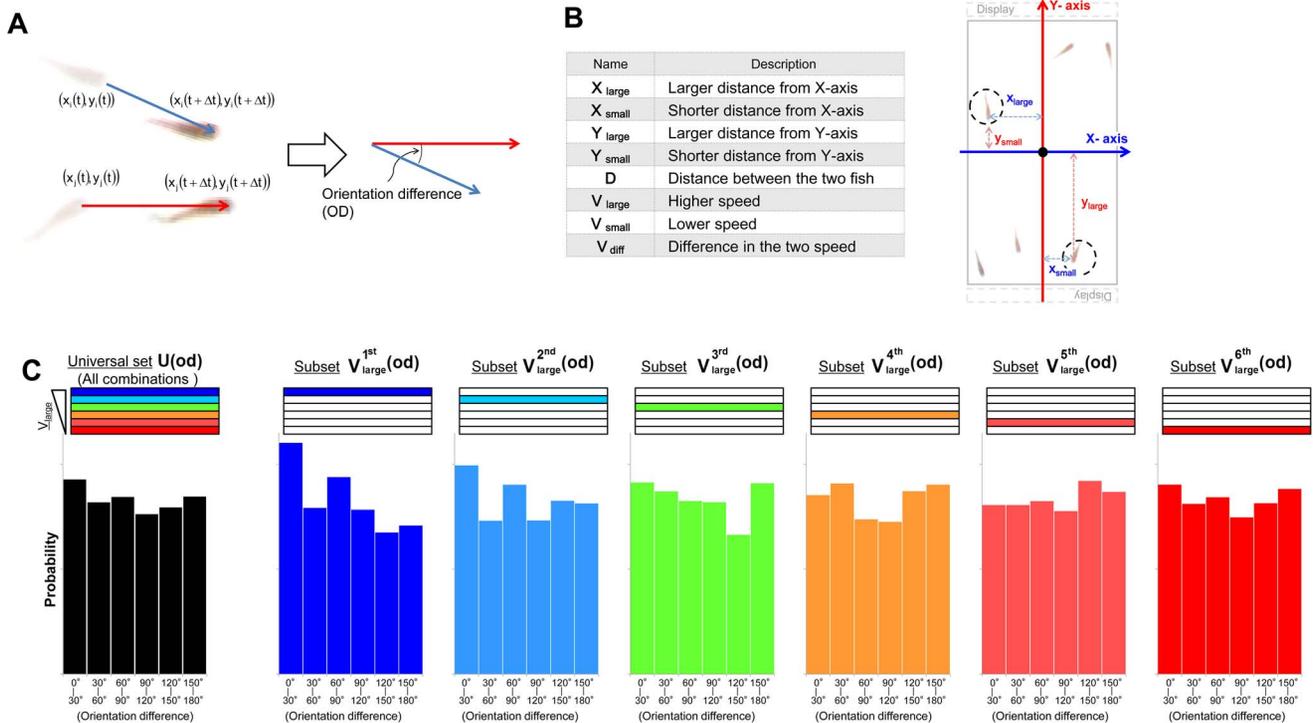
$$y_{\text{large}} \triangleq \text{large}(|y_i(t)|, |y_j(t)|) \quad (8)$$

$Y_{\text{small}}$ : Shorter distance from Y-axis.

$$y_{\text{small}} \triangleq \text{small}(|y_i(t)|, |y_j(t)|) \quad (9)$$

D: Distance between the two fish.

$$d \triangleq \sqrt{\{x_i(t) - x_j(t)\}^2 + \{y_i(t) - y_j(t)\}^2} \quad (10)$$



**Figure 3. Searching for candidate explanatory variables that may affect inter-individual interactions. (A) The difference in the orientation.** We focused on the difference between the orientations of a focal fish pair to measure the amount of inter-individual interaction. **(B) List of possible explanatory variables.** Five variables were concerned with position, and three variables were concerned with velocity. The variables concerned with position are shown in the schematic diagram. **(C) Visualization of the difference of probability distribution of six subsets generated based on  $V_{\text{large}}$ .** Each explanatory variable was stratified into 6 subsets (conditions) based on its score ranking. If the explanatory variable,  $V_{\text{large}}$  did not affect the OD, the probability distribution of the six subsets would be similar to that of the universal set. In contrast, if the condition ( $V_{\text{large}}^{1\text{st}}$  Group) affected the OD, the distribution would be different from that of the universal set. doi:10.1371/journal.pone.0071685.g003

$V_{\text{large}}$ : Higher speed.

$$v_{\text{large}} \triangleq \text{large}(|v_x(t)|, |v_y(t)|) \quad (11)$$

$V_{\text{small}}$ : Lower speed.

$$v_{\text{small}} \triangleq \text{small}(|v_x(t)|, |v_y(t)|) \quad (12)$$

$V_{\text{diff}}$ : Difference of the two speeds.

$$v_{\text{diff}} \triangleq \left\| \begin{pmatrix} v_x(i,t) \\ v_y(i,t) \end{pmatrix} - \begin{pmatrix} v_x(j,t) \\ v_y(j,t) \end{pmatrix} \right\| \quad (13)$$

Based on time-series data ( $t = -30 \sim 0$  s at 1-s intervals) of x-y coordinates of Fish  $i$  and Fish  $j$ , we calculated the eight explanatory variables and prepared datasets of OD accompanied by the eight explanatory variables.

$$\text{od}_{i,j,t} = f(i,j,t, x_{\text{large}}, x_{\text{small}}, y_{\text{large}}, y_{\text{small}}, d, v_{\text{large}}, v_{\text{small}}, v_{\text{diff}}) \quad (14)$$

To search for candidate explanatory variables that affect the OD, we divided  $\mathbf{U}(\mathbf{od})$  into six subsets based on the magnitude of each explanatory variable (score ranking) and measured the difference of subsets from the universal set using the KLD (Fig. 4). If the explanatory variables did not affect the OD, the probability distribution of all six subsets would be similar to  $\mathbf{U}(\mathbf{od})$  (Fig. 3). In the present study, we measured the difference of a total of 48 subsets (6 subsets  $\times$  8 explanatory variables; Fig. 4).

The probability distribution of the subset with the 1<sup>st</sup> KLD rank was skewed left (Figs. 3C and 4B), and was defined by the highest magnitude of  $V_{\text{large}}$  (subset  $\mathbf{V}_{\text{large}}^{1\text{st}}(\mathbf{od})$ ). In contrast, the probability distribution of the subset that was defined by the lowest magnitude of  $V_{\text{large}}$  (subset  $\mathbf{V}_{\text{large}}^{6\text{th}}(\mathbf{od})$ ) was almost flat with low KLD rank (46<sup>th</sup>) (Figs. 3C and 4A). The G-value revealed a significant difference in the subset  $\mathbf{V}_{\text{large}}^{1\text{st}}(\mathbf{od})$  ( $p < 0.01$ ), but not in the five lower subsets ( $\mathbf{V}_{\text{large}}^{2\text{nd}}(\mathbf{od}) \sim \mathbf{V}_{\text{large}}^{6\text{th}}(\mathbf{od})$ ), indicating that high-speed movement of either fish tended to affect alignment of the other fish in a focal pair.

We also performed the same analysis based on other subdivisions (Figs. S2 and S3). In cases of 3 subdivisions, the KLD ranks of subset  $\mathbf{V}_{\text{large}}^{1\text{st}}(\mathbf{od})$  was 4<sup>th</sup> of 12 subsets and the G-test revealed a significant difference ( $p < 0.01$ ) in the subset  $\mathbf{V}_{\text{large}}^{1\text{st}}(\mathbf{od})$ , but not in the two lower subsets ( $\mathbf{V}_{\text{large}}^{2\text{nd}}(\mathbf{od})$ , and  $\mathbf{V}_{\text{large}}^{3\text{rd}}(\mathbf{od})$ ). In cases of 9 subdivisions, the KLD ranks of subset  $\mathbf{V}_{\text{large}}^{1\text{st}}(\mathbf{od})$  was 6<sup>th</sup> of 72 subsets, and the G-test revealed a significant difference ( $p < 0.01$ ) in the subset  $\mathbf{V}_{\text{large}}^{1\text{st}}(\mathbf{od})$ , but not in the eight lower subsets  $\mathbf{V}_{\text{large}}^{2\text{nd}}(\mathbf{od})$  -  $\mathbf{V}_{\text{large}}^{9\text{th}}(\mathbf{od})$ . In addition, we examined other interval widths (Fig. S2). In cases of 4 and 9 OD subdivisions (45° and 20° interval widths, respectively), the KLD ranks of subset  $\mathbf{V}_{\text{large}}^{1\text{st}}(\mathbf{od})$  were 3<sup>rd</sup> and 4<sup>th</sup>, respectively, and the G-test revealed a significant difference ( $p < 0.01$ ; Fig. S3). These findings suggested that changing the number of subdivisions does not strongly affect the results of the present data-mining method.

Here the skewed left shape (Figs. 3C and 4B) of  $\mathbf{V}_{\text{large}}^{1\text{st}}(\mathbf{od})$  led us to hypothesize that the focal pair tends to show alignment movements under conditions in which either of the focal pair fish moves with a relatively high speed (among the top one-sixth), because the probability is relatively high within these intervals [0°, 30°]. The probability distribution (histogram shapes) of the subsets with the highest, 2<sup>nd</sup>, 3<sup>rd</sup>, 4<sup>th</sup>, and 5<sup>th</sup> rank KLD were not skewed to either the left or the right (Fig. 4B). As the four subsets were defined by the magnitude of  $Y_{\text{large}}$  and  $X_{\text{small}}$ , which represented the location of the fish in the tank, the biased frequency distribution of OD might be due to the spatial restriction in the tank rather than inter-individual interactions (See Fig. S1 for details).

### High-speed access of learned fish to the food source in response to the visual cue

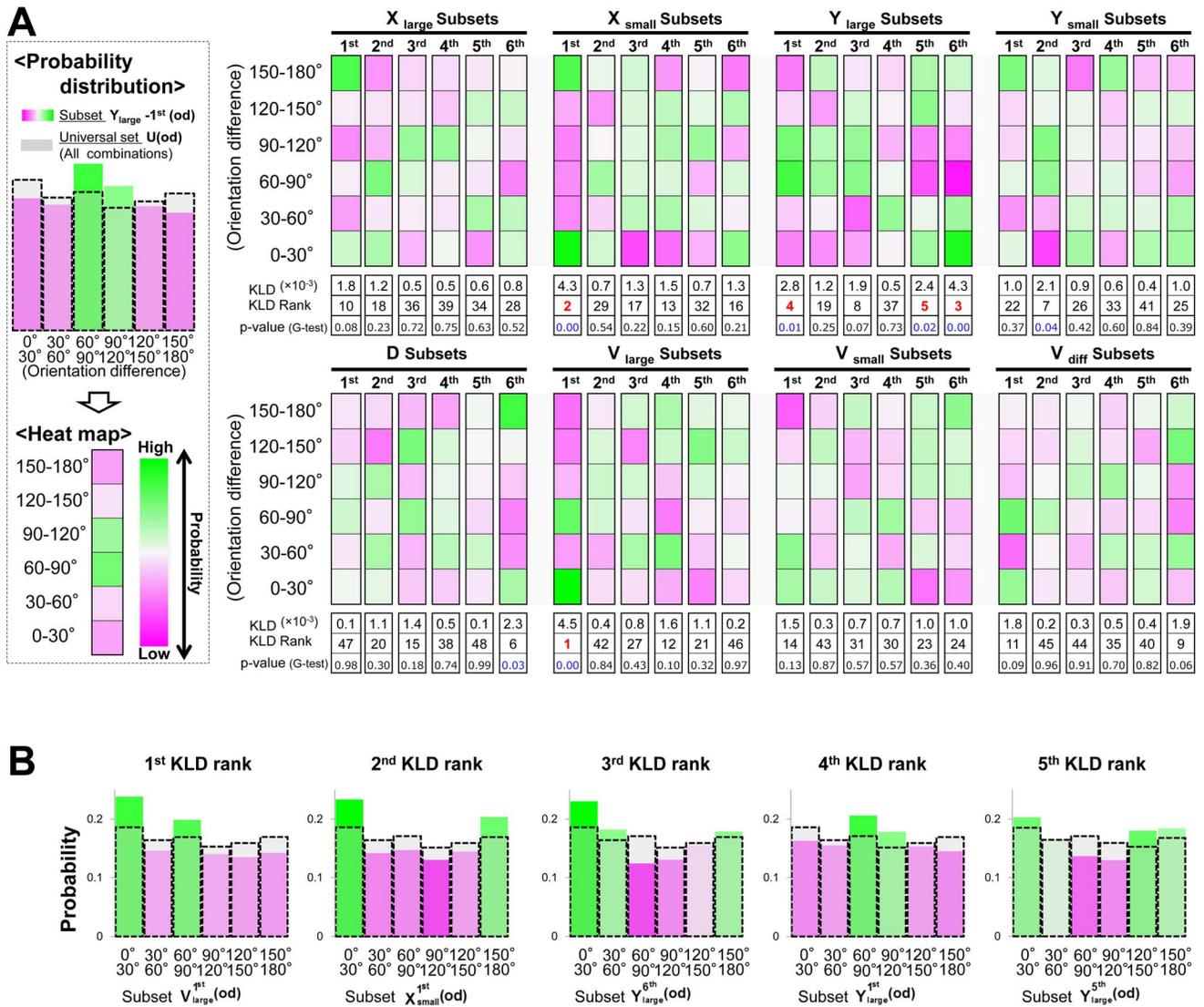
The result of the previous data-mining prompted us to examine whether the alignment elicited by high-speed movement facilitated access to a food source in response to a visual cue. First, we examined whether trained fish moved at a high speed in response to a visual cue. We defined trained fish as those that received more than 5 training sets ( $> 10$  trials) in the single-fish training paradigm in the previous experiments (Fig. 1E) and analyzed the movement of the trained fish during the test period. Then we defined the directivity to the display. When the fish moved forward toward the visual cue, directivity was defined as 0° (Fig. 5A). When they moved backward, directivity was defined as 180°. We compared the directivity of fish moving at high and low speeds (in the top 20% and the lower 80% of each individual, respectively). Directivity was significantly different at high-speed ( $59.3 \pm 4.7^\circ$ ) and low-speed ( $73.3 \pm 2.3^\circ$ , mean  $\pm$  s.e.m.; t-test,  $p = 0.0044$ ), indicating that the trained fish rapidly moved toward the visual cue (Fig. 5B).

### Demonstrator-elicited approach behavior to the visual cue in observers

We then evaluated whether the response of the trained fish (demonstrators) to the visual cue elicited approach behavior by the observers (untrained fish) to the same visual cue without supplying the food reinforcement. Here we prepared a 4-fish group comprising one demonstrator and three observers and analyzed the time-series data (200~220, 380~400, and 560~580 s at 1/3-s intervals) of x-y coordinates of the four individually marked individuals (Fish  $i$ :  $i = 1 \sim 4$ , individual identification number) in response to the visual cue (Fig. 6A). Displacement towards the visual cue of the demonstrator fish peaked at  $37.6 \pm 7.2$  mm (mean  $\pm$  s.e.m.) at 8.0 s after presentation of the visual cue, while that of the three observers was  $18.8 \pm 3.9$  mm at  $t = 11.0$  s (Fig. 6B). This finding suggests that the response of the demonstrator to the visual cue elicits approach behavior by the observers, suggesting that the demonstrator's behavior induces the observers to approach the food source during the test and training period.

### Behavioral cause-and-effect of rapid response between group members

We then evaluated whether a focal fish moving at high speed induces the alignment of other members with a time delay during free-swimming without presentation of a visual cue (Fig. 6C). We calculated the mean OD between the focal fish (Fish  $i$ ) and the other three fish (Fish  $j_1 \sim j_3$ ) with a time difference (Td), based on time-series data ( $t = 180 \sim 200$ ,  $360 \sim 380$ , and  $540 \sim 560$  at 1/3-s intervals, Fig. 6A) of x-y coordinates of Fish  $i$  ( $t + \text{td}$ ) and Fish  $j_1 \sim j_3$  ( $t$ ). The OD with a time difference between two fish, Fish  $i$  ( $t + \text{td}$ )



**Figure 4. Top five subsets showing high-scored KLD. (A) Measurement of the difference of subsets from the universal set.** Here we represent the probability distribution using a heat map. The OD histogram-intervals with a higher proportion than that of all combinations are shown in green, while those with a lower proportion are shown in magenta. The KLD value and KLD rank are indicated below each column. Below them, the G-test p-values are shown. High KLD values indicate low similarity and we selected the top five. **(B) Visualization of the probability distributions of the top five subsets exhibiting high-scored KLD.**  
doi:10.1371/journal.pone.0071685.g004

and Fish  $j_k(t)$  is defined as:

$$\begin{pmatrix} v_x(t) \\ v_y(t) \end{pmatrix} \triangleq \begin{pmatrix} x(t + \frac{1}{3}) - x(t) \\ y(t + \frac{1}{3}) - y(t) \end{pmatrix} \quad (15)$$

$$od_{mean}(i,t,td) = \frac{1}{3} \{od(i,j_1,t,td) + od(i,j_2,t,td) + od(i,j_3,t,td)\} \quad (17)$$

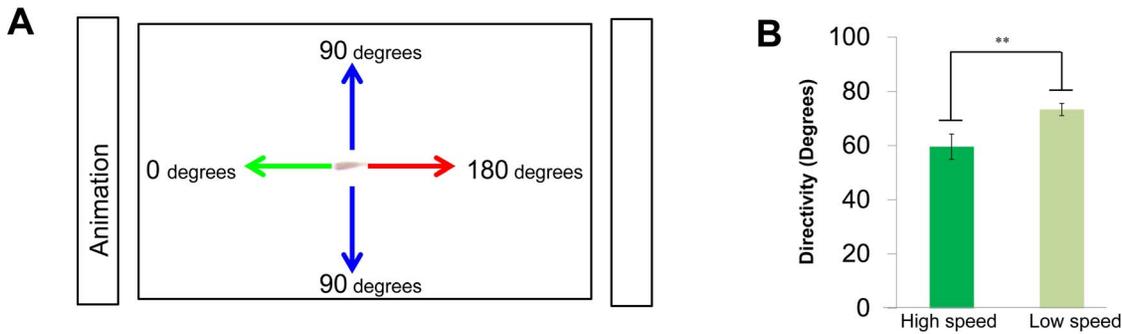
To examine the effect of fish speed on the  $OD_{mean}$ , we prepared two subsets based on fish speed.

$$od(i,j_k,t,td) \triangleq \arccos \frac{\begin{pmatrix} v_x(i,t+td) \\ v_y(i,t+td) \end{pmatrix} \begin{pmatrix} v_x(j_k,t) \\ v_y(j_k,t) \end{pmatrix}}{\left| \begin{pmatrix} v_x(i,t+td) \\ v_y(i,t+td) \end{pmatrix} \right| \left| \begin{pmatrix} v_x(j_k,t) \\ v_y(j_k,t) \end{pmatrix} \right|} \quad (16)$$

$$\mathbf{T} \triangleq \{ \text{all the concerning time, } \frac{1}{3} \text{ s interval} \} \quad (18)$$

$$\mathbf{V} \triangleq \{v(l,t) | t \in \mathbf{T} \text{ and } l = i, j_1, j_2, j_3\} \quad (19)$$

The mean OD between the focal fish (Fish  $i$ ) and the other three fish (Fish  $j_1 \sim j_3$ ) with a time difference ( $Td$ ) is defined as:



**Figure 5. The difference in the directivity orientation between high speed and low speed. (A) Definition of directivity.** The directivity value was 0 degrees if fish movement was directed to the display, and was 180 degrees if the fish moved away from the display. **(B) Difference in the directivity in the test period of solitary learning.** From the 6<sup>th</sup> to the 10<sup>th</sup> test period of solitary learning, the mean value of the directivity angle at high speed was 59.6 degrees (N=160), while that at low speed was significantly higher at 73.3 degrees (N=640; t-test,  $p=0.0044$ ). doi:10.1371/journal.pone.0071685.g005

$$\mathbf{V}^{\text{high}} \triangleq \{v(i,t) \mid \text{from the upper } \frac{1}{6} \text{ to upper the } \frac{0}{6}\} \quad (20)$$

$$\mathbf{T}(\mathbf{i})^{\text{high}} \triangleq \{t \mid v(i,t) \in \mathbf{V}^{\text{high}}\} \quad (21)$$

$$\text{od}(i,\text{td})_{\text{mean}}^{\text{high}} \triangleq \frac{1}{n(\mathbf{T}(\mathbf{i})^{\text{high}})} \sum_{t \in \mathbf{T}(\mathbf{i})^{\text{high}}} \text{od}(i,t,\text{td})_{\text{mean}} \quad (22)$$

$$\mathbf{V}^{\text{low}} \triangleq \{v(i,t) \mid \text{from the upper } \frac{6}{6} \text{ to upper the } \frac{1}{6}\} \quad (23)$$

$$\mathbf{T}^{\text{low}}(\mathbf{i}) \triangleq \{t \mid v(i,t) \in \mathbf{V}^{\text{low}}\} \quad (24)$$

$$\text{od}(i,\text{td})_{\text{mean}}^{\text{low}} \triangleq \frac{1}{n(\mathbf{T}(\mathbf{i})^{\text{low}})} \sum_{t \in \mathbf{T}(\mathbf{i})^{\text{low}}} \text{od}(i,t,\text{td})_{\text{mean}} \quad (25)$$

Figure 6D shows the  $\text{od}_{\text{mean}}^{\text{high}}$  and  $\text{od}_{\text{mean}}^{\text{low}}$  between the focal fish and other three fish with a time difference (Td). As expected, the  $\text{od}_{\text{mean}}^{\text{high}}$  (the focal fish swam at high speed; solid lines) decreased with a time delay, with a minimum value ( $76.8 \pm 0.68^\circ$ ,  $\text{mean} \pm \text{s.e.m.}$ , N=34) at  $\text{td} = -2/3$  s, indicating that the decrease in the  $\text{od}_{\text{mean}}^{\text{high}}$  followed the high-speed movement of the focal fish. In contrast, the  $\text{od}_{\text{mean}}^{\text{low}}$  (the focal fish swam at low speed; dashed lines) was approximately  $90^\circ$  at any time difference, indicating that there was no interaction between group members when the focal fish swam at low speed. Thus, high-speed movement of an individual in a group tended to elicit alignment of the other members with a 1/3- to 1-s delay.

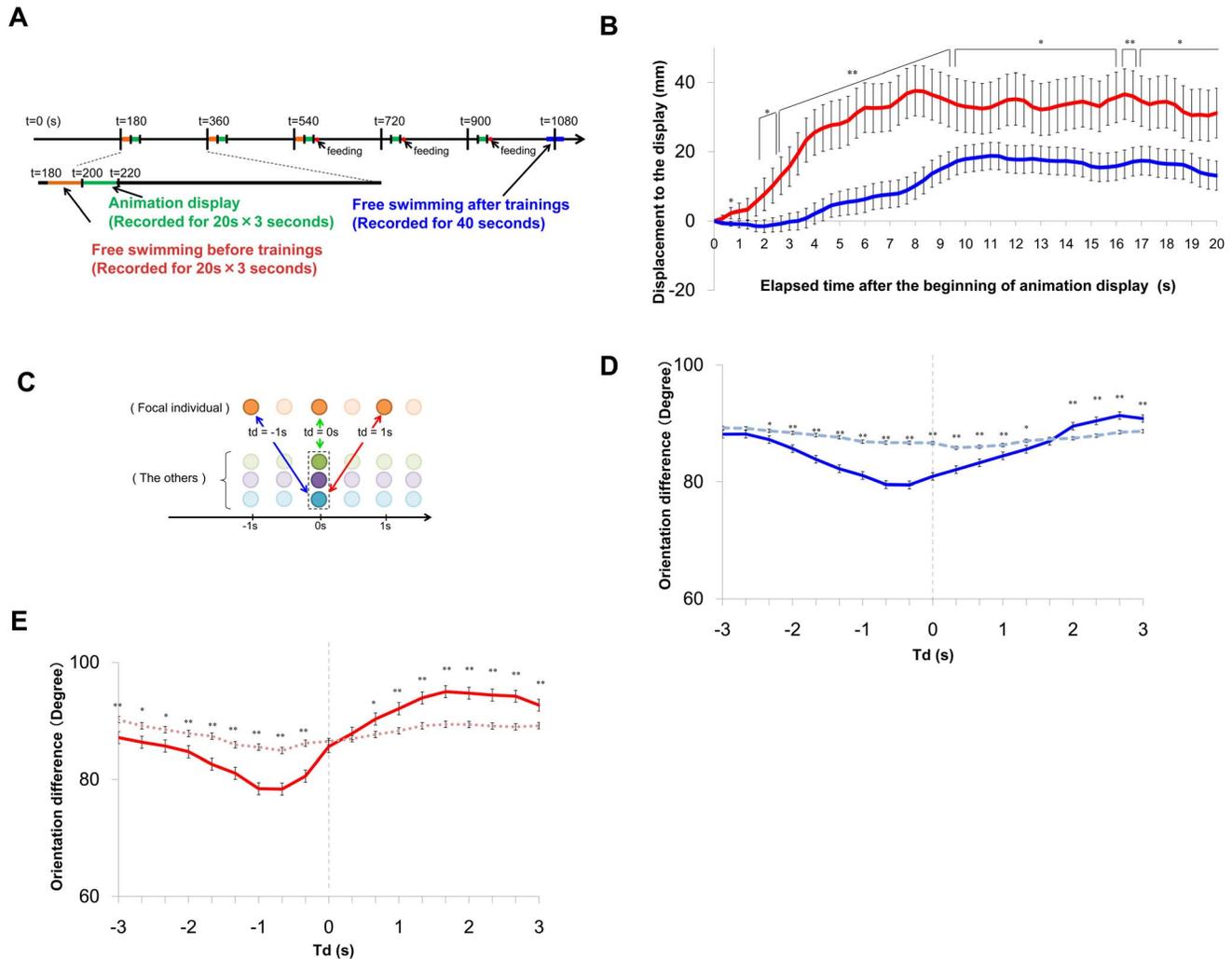
#### Elicitation of alignment in observers by high-speed movement of the demonstrator during visual cue presentation

We then examined whether the observers exhibited alignment when the demonstrators exhibit high-speed movement during the visual cue presentation. Figure 6E shows the  $\text{od}_{\text{mean}}^{\text{high}}$  and between

the demonstrator (Fish  $i$ ) and observers (Fish  $j_1 \sim j_3$ ) with the time difference (See equations 17,30,31). The time difference (Td) is the time lag of Fish  $i$  from Fish  $j_1 \sim j_3$  (as a standard). The  $\text{od}_{\text{mean}}^{\text{high}}$  (the demonstrator, as the focal fish, swam at high speed; solid lines), reached a minimum ( $78.4^\circ \pm 1.0^\circ$ ,  $\text{mean} \pm \text{s.e.m.}$ , N=34) at  $\text{td} = -1$  s. In contrast, (the demonstrator, as the focal fish, swam at low speed; dashed lines) was approximately  $90^\circ$  at any time difference. These findings strongly suggested that the observers tended to exhibit alignment in response to the high-speed movement of the demonstrator with a 1/3- to 4/3-s delay during presentation of the visual cue.

#### Possible involvement of individual recognition in social learning

Given that medaka fish could recognize a demonstrator (trained fish) in the group, naïve fish that exhibited alignment movement with a demonstrator would increase their own foraging success through social learning about food sources. Therefore, we examined whether the high-speed movement of the demonstrator is more likely to induce alignment movement of other observers than the movement of an observer. Here we calculated the  $\text{OD}_{\text{mean}}^{\text{high}}$  during free-swimming without the visual cue before ( $t = 180 \sim 200$ ,  $360 \sim 380$ , and  $540 \sim 560$  s) and after three trainings ( $t = 1060 \sim 1100$  s). As the focal fish, we used a demonstrator or a randomly selected observer (See Equations 38 and 57, respectively). It is noteworthy that, under this condition, fish interacted with each other without external visual signals (Fig. 7A and C). Before training, irrespective of whether the focal fish was the demonstrator or observer, the  $\text{od}_{\text{mean}}^{\text{high}}$  decreased with a time delay and had minimum values ( $76.6 \pm 1.1^\circ$  and  $75.3 \pm 0.96^\circ$ , respectively,  $\text{mean} \pm \text{s.e.m.}$ , N=31,  $p=0.3667$ ) at  $\text{td} = -2/3$  s (Fig. 8A). The difference between the two minimum values of the  $\text{od}_{\text{mean}}^{\text{high}}$  was not significant (Fig. 8A). The  $\text{od}_{\text{mean}}^{\text{low}}$  in the demonstrator and observer were approximately  $90^\circ$  at any time difference (Fig. 7A). The similar tendencies were also detected during the visual cue presentation before trainings (Fig. 7B). In contrast, during free swimming after the three training trials, there was significant difference depending on whether the focal fish was the demonstrator or observer (Fig. 7C). The  $\text{od}_{\text{mean}}^{\text{high}}$  of the demonstrator were significantly lower than those of the observer at any negative time difference ( $\text{td} = 0 \sim -3/4$  s; Fig. 8B). The minimum values ( $79.1 \pm 1.2^\circ$  and  $83.7 \pm 0.93^\circ$ , respectively,  $\text{mean} \pm \text{s.e.m.}$ , N=31,  $p=0.0023$ ) at  $\text{td} = -1/3$  s were significantly different. The  $\text{od}_{\text{mean}}^{\text{low}}$  of both the demonstrator and observer were approximately



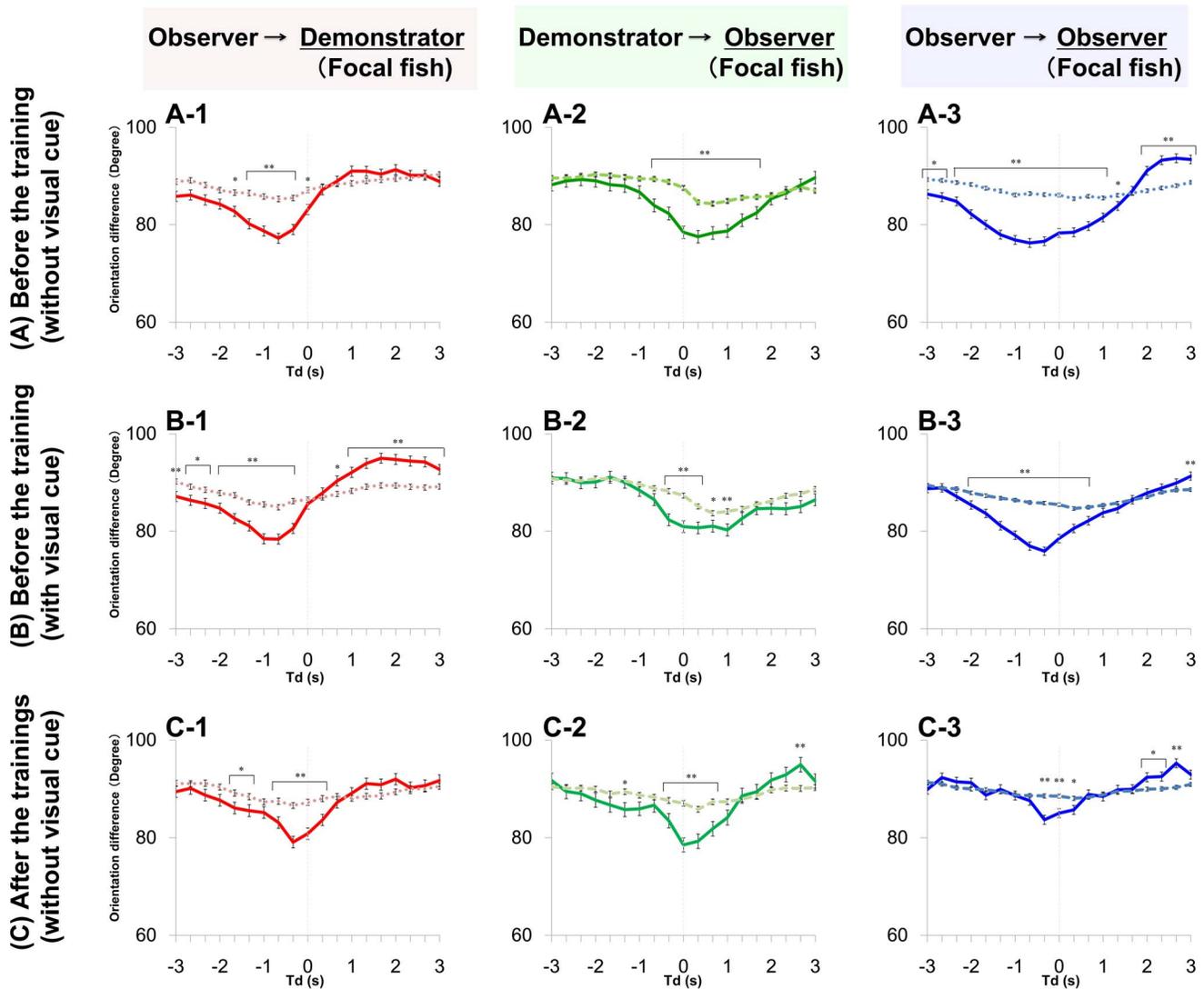
**Figure 6. Measurement of the orientation differences between focal fish and the others with a time difference. (A) Time table of Experiment 2.** In this experiment, we used 4 fish per group: one trained individual and three untrained ones. The trained fish had learned to move towards the visual cue. **(B) Change in the distance moved to the visual cue.** During the animation display before training, we measured the mean movement from the individuals to the display at 1/3-s intervals. The red line indicates the mean movement of the demonstrators, which peaked at 37.6 mm at 8.0 s. The blue line indicates the mean movement of the observers, which peaked at 18.8 mm at 11.0 s. Bars indicate s.e.m. After 2 s, a statistically significant difference was detected (t-test,  $p < 0.05$ , see Table S1 for actual p values). **(C, D) The orientation differences (OD) when the phase was shifted during free-swimming before training. (C) Definition of time difference (Td).** Td refers to the time difference between the focal fish and the other fish. **(D) The orientation difference (OD) of all the individuals at every Td.** The orientation difference  $OD_{\text{mean}}^{\text{high}}$  when the focal fish swam at high speed is indicated as the solid blue line, and the orientation difference  $OD_{\text{mean}}^{\text{low}}$  when the focal fish swam at low speed is indicated as the dashed blue line [see equation (17)]. The  $OD_{\text{mean}}^{\text{high}}$  at high speed fell from  $-7/3\text{s}$  to  $4/3\text{s}$ , but exceeded  $OD_{\text{mean}}^{\text{low}}$  from 2s to 3s (t-test,  $P < 0.01$ , Bar; S.E.). Statistically significant differences are indicated by asterisks (t-test, \*:  $p < 0.05$ , \*\*:  $p < 0.01$ ). **(E) The orientation differences versus Td during the visual cue.** The focal fish are the demonstrators. The solid red line indicates the  $OD_{\text{mean}}^{\text{high}}$  when demonstrators swam at high speed, and the dashed red line indicates the  $OD_{\text{mean}}^{\text{low}}$  when the demonstrators swam at low speed. Statistically significant differences are indicated by asterisks (t-test, \*:  $p < 0.05$ , \*\*:  $p < 0.01$ ). doi:10.1371/journal.pone.0071685.g006

$90^\circ$  at any time difference (Fig. 7C). Our findings suggested that naïve fish (observer) exhibited diminished ability to induce alignment in the other observer fish during the training trials whereas the demonstrator had a relatively high tendency to induce alignment after three trainings.

## Discussion

Here we demonstrated that grouping in medaka fish facilitated learning of visual information associated with a food source, and that naïve “observer” medaka fish would follow trained “demon-

strator” fish. Generally, fish in large groups forage more efficiently than those in small groups via socially transmitted information, which has been demonstrated in goldfish (*Carassius auratus auratus*) [24], bluntnose minnows (*Pimephales notatus*) [25], Alaska pollack (*Theragra chalcogramma*) [26], sticklebacks (*Gasterosteus aculeatus*) [27,28], golden shiners (*Notemigonus crysoleucas*), and guppies (*Poecilia reticulata*) [10,14,29,30]. Theoretical studies revealed that group coordinated dynamics can be explained by interactions between individuals, such as collision avoidance and alignment among group members [31]. The development of methods to determine the behavioral cause-and-effect between

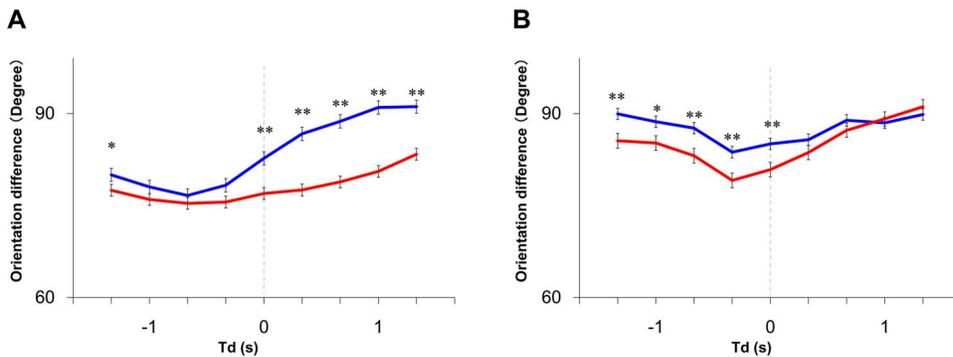


**Figure 7. Orientation differences with time difference.** The orientation difference  $OD_{\text{mean}}^{\text{high}}$  when the focal fish swam at high speed is indicated as a solid line, the orientation difference  $OD_{\text{mean}}^{\text{low}}$  when one fish swam at a low speed is indicated as a dashed line. **(A) During free-swimming before training.** (A-1) The  $OD_{\text{mean}}^{\text{high}}$  and  $OD_{\text{mean}}^{\text{low}}$  between the one demonstrator and three observers when the focal fish were demonstrators. Corresponding equations are (37) and (41). (A-2) The  $OD_{\text{mean}}^{\text{high}}$  and  $OD_{\text{mean}}^{\text{low}}$  between three observers and one demonstrator when the focal fish were observers. Corresponding equations are (47) and (50). (A-3) The  $OD_{\text{mean}}^{\text{high}}$  and  $OD_{\text{mean}}^{\text{low}}$  between one observer and the other two observers, when the focal fish were observers. Corresponding equations are (57) and (60). **(B) During the visual cue presentation before training.** **(C) During free-swimming after training.** Actual p values are shown in Table S1. doi:10.1371/journal.pone.0071685.g007

group members that mediates social learning, however, has been an obstacle [32]. In a medaka fish group under free-swimming conditions without an external stimulus, we could not detect any tendency to maintain the nearest neighbor distance (data not shown) or to exhibit alignment movement as a group, which is consistent with our previous studies showing no significant inter-individual interaction between a pair of medaka fish under free-swimming conditions [17]. Here, we demonstrated that the high-speed movement of a focal fish tends to induce alignment of the group members locally and transiently. This behavioral property can mediate a process of social learning called “observational conditioning”, which refers to the process in which “the response of a demonstrator to a stimulus elicits a matching response on the part of an observer, that simultaneously perceives the original stimulus, and effectively learns that the response is an appropriate

response to it” [33,34]. The behavioral cause-and-effect based on the high-speed movement between group members will help us to understand the dynamics of collective behaviors.

Our findings also suggest that individual recognition among a group facilitates social learning. This raises the possibility that an observer could recognize other observers as uninformed fish and choose not to respond to their movement, as the observer obtained no food by following the other observers. Theoretical models of social learning suggest that the formation of a group where social interactions occur at random can facilitate the learning process. Reader & Laland (1999), however, demonstrated that sex, age, and hunger level strongly influence foraging performance in the guppy [35], implying “directed social learning”, where transmitted information is restricted to a subset of individuals [36]. Our findings suggested the existence of “directed social learning” in



**Figure 8. Comparison of the orientation difference of the demonstrator and the observers.** We compared the  $OD_{\text{mean}}^{\text{high}}$  between a demonstrator and an observer with the  $OD_{\text{mean}}^{\text{high}}$  between an observer and another observer. **(A) During free-swimming before training.** The blue line indicates the orientation difference  $OD_{\text{mean}}^{\text{high}}$  between the observer at high speed and the other two observers. Corresponding equation is (57). The orientation difference was smaller than the red line indicating the difference  $OD_{\text{mean}}^{\text{high}}$  between the demonstrator at high speed and the three observers. Corresponding equation is (38). Statistically significant differences are indicated by asterisks (t-test, \* :  $p < 0.05$ , \*\* :  $p < 0.01$ , see Table S1 for actual p value.). **(B) During free-swimming after training.** The red line was smaller than the blue line, and statistically significant differences are indicated by asterisks (t-test, \* :  $p < 0.05$ , \*\* :  $p < 0.01$ ). doi:10.1371/journal.pone.0071685.g008

medaka fish, where information is efficiently transmitted from trained fish to naive fish by individual recognition.

Although advances in video-tracking and GPS systems allow for automatic location of one member of a group of animals [37], [38], we had no standard data-mining method to search for unknown behavioral rules in a hypothesis-independent manner based on the times-series coordinates of individuals. In the present study, we established a novel data-mining method based on the KLD. We first used it to analyze inter-individual interactions in a hypothesis-independent manner using the times-series coordinates of individuals. Although we used dynamics and positional information of individuals as explanatory variables in the present study, other quantitative indices such as physiologic state, body size, and age, as well as qualitative indices such as sex, rearing conditions, familiarity, and strain, can be used as explanatory variables. Generation of some subsets based on an explanatory variable of interest allows investigators to estimate whether the explanatory variable is related to the inter-individual interaction by comparing the shapes of the probability distribution of subsets based on the KLD. Combining the present data-mining method and recent advances in data tracking and GPS systems will allow us to find unknown behavioral rules in animal groups.

## Materials and Methods

### Ethics statement

The work in this paper was conducted using protocols approved by the Animal Care and Use Committee of the University of Tokyo (permit number: 12–07). All surgery was performed under cold anesthesia, and all efforts were made to minimize suffering.

### Animals

Medaka (Himedaka, *Oryzias latipes*) were obtained from a local pet shop and females (20–25 mm body length) were selected for use in the study. The fish were housed in a plastic rearing tank (12×13×19 cm) maintained at a temperature of 28°C under a 14 h:10 h light:dark cycle.

### Associative visual learning

During the experiment, the fish were transferred into the test-training tank (15.5×9.5 cm×10.8 cm, water depth 2.5 cm) with

covered opaque walls and two transparent windows (6×7 cm) to allow the fish to see the LCD displays placed at each side of the tank (Fig. 1A). The feed dispensers comprised a syringe and silicon tube fixed to the sides of the tank. The feed dispenser and two LCD displays (RDT153LM, Mitsubishi) were set beside the test-training tank. Movement of the fish was recorded from above the tank in high-definition movies using a digital single-lens reflex camera (EOS T4i and EF50mm F1.8 II, Canon). Cold cathode fluorescent lamps were placed over the tanks as light sources. Three test-training tanks were used for each set of LCD displays.

### Visual associative learning using only naïve fish

For grouped training, we labeled the right or left back of the fish with silicon ink, which allowed us to discriminate among individuals using a recorded movie. The silicon ink was injected under the skin of ice-anesthetized fish using a syringe (TERMO SV-S19EL, SS-02SZ). Using three colors, we labeled the six fish of each group as follows: right-pink, left-pink, right-blue, left-blue, right-green, and left-green. During the 3 days after the injection, groups of 3 to 4 fish were kept in a large opaque cup. Prior to the day before the training day, the fish were single-housed in opaque cups. For grouped control, single training, and single control fish, a needle was penetrated into the back of the fish but no silicon ink was injected. The other manipulations were the same as for the grouped training.

As the training-test procedure for grouped and single training, each fish group (6 fish) was transferred from the rearing tank into the training tank, and the training and test procedure was started 15 min later. We repeated the test and training procedure 20 times. Each procedure comprised both test and training periods, and the interval between each procedure was 245 s. At the beginning of the procedure, we presented the animation (visual cue), i.e., a 2-cm green circle moved in a 2.5-cm diameter circle at 60-rpm, on the display on one side of the tank. After a 5-s interval, 120 brine shrimp were placed into the tank through the silicon tube next to the LCD display showing the visual cue for a period of 55 s. The positions of the visual cue display and food presentations were alternated from one side to the other. For grouped control and single control fish, we fed the fish 120 brine shrimp during the inter-trial interval and not during the visual cue display period. All

manipulations were the same as for grouped training. The other manipulations were the same as the grouped training.

**Data collection**

We measured the distance each individual fish traveled toward the display during each test period. To analyze fish movement during free swimming, we converted the video images into image sequences using QuickTime and manually obtained time series (1 per second) position data of each fish using a touch panel PC and Image J. We entered the actual path by manually following the track of the animal, because there is no methodology currently available that allows for the automatic detection of the location of every individual in a group of freely swimming fish from a video source. Due to a technical error, data for only 11 of the 12 training groups were analyzed.

**A data-mining method based on KLD**

First, we generated the  $\mathbf{U}(\mathbf{od})$ , the universal set of an orientation difference (OD). The OD can be described as the image of explanatory variables:

$$\begin{aligned} \mathbf{od} &= f(i, j, g, t, x_{\text{large}}, x_{\text{small}}, y_{\text{large}}, y_{\text{small}}, D, v_{\text{large}}, v_{\text{small}}, v_{\text{diff}}) \\ i, j &: \text{individuals (6 fish/ experimental group)} \\ g &: \text{experimental group (11 groups),} \\ t &: \text{time (30 seconds, 1s interval)} \\ x_{\text{large}} \sim v_{\text{diff}} &: \text{explanatory variables} \end{aligned} \tag{26}$$

Next we generated six subsets from  $\mathbf{U}(\mathbf{od})$  based on score-ranking of each explanatory variable. We made six groups of each explanatory variable and then generated the six OD subsets.

$$\begin{aligned} \mathbf{P}^{1\text{st}}(\mathbf{p}) &\triangleq \{\mathbf{p} \mid \text{from the upper } 1/6 \text{ to the upper } 0/6 \text{ in each experimental group}\} \\ \mathbf{P}^{2\text{nd}}(\mathbf{p}) &\triangleq \{\mathbf{p} \mid \text{from the upper } 2/6 \text{ to the upper } 1/6 \text{ in each experimental group}\} \\ \mathbf{P}^{3\text{rd}}(\mathbf{p}) &\triangleq \{\mathbf{p} \mid \text{from the upper } 3/6 \text{ to the upper } 2/6 \text{ in each experimental group}\} \\ \mathbf{P}^{4\text{th}}(\mathbf{p}) &\triangleq \{\mathbf{p} \mid \text{from the upper } 4/6 \text{ to the upper } 3/6 \text{ in each experimental group}\} \\ \mathbf{P}^{5\text{th}}(\mathbf{p}) &\triangleq \{\mathbf{p} \mid \text{from the upper } 5/6 \text{ to the upper } 4/6 \text{ in each experimental group}\} \\ \mathbf{P}^{6\text{th}}(\mathbf{p}) &\triangleq \{\mathbf{p} \mid \text{from the upper } 6/6 \text{ to the upper } 5/6 \text{ in each experimental group}\} \\ \mathbf{P} &: \mathbf{X}_{\text{large}}, \mathbf{X}_{\text{small}}, \mathbf{Y}_{\text{large}}, \mathbf{Y}_{\text{small}}, \mathbf{D}, \mathbf{V}_{\text{large}}, \mathbf{V}_{\text{small}}, \mathbf{V}_{\text{diff}} \end{aligned} \tag{27}$$

$$\begin{aligned} \mathbf{P}^{1\text{st}}(\mathbf{od}) &\triangleq \{\mathbf{od}(\mathbf{p}) \mid \mathbf{p} \in \mathbf{P}^{1\text{st}}\} & \mathbf{P}^{2\text{nd}}(\mathbf{od}) &\triangleq \{\mathbf{od}(\mathbf{p}) \mid \mathbf{p} \in \mathbf{P}^{2\text{nd}}\} \\ \mathbf{P}^{3\text{rd}}(\mathbf{od}) &\triangleq \{\mathbf{od}(\mathbf{p}) \mid \mathbf{p} \in \mathbf{P}^{3\text{rd}}\} & \mathbf{P}^{4\text{th}}(\mathbf{od}) &\triangleq \{\mathbf{od}(\mathbf{p}) \mid \mathbf{p} \in \mathbf{P}^{4\text{th}}\} \\ \mathbf{P}^{5\text{th}}(\mathbf{od}) &\triangleq \{\mathbf{od}(\mathbf{p}) \mid \mathbf{p} \in \mathbf{P}^{5\text{th}}\} & \mathbf{P}^{6\text{th}}(\mathbf{od}) &\triangleq \{\mathbf{od}(\mathbf{p}) \mid \mathbf{p} \in \mathbf{P}^{6\text{th}}\} \\ \mathbf{P} &: \mathbf{X}_{\text{large}}, \mathbf{X}_{\text{small}}, \mathbf{Y}_{\text{large}}, \mathbf{Y}_{\text{small}}, \mathbf{D}, \mathbf{V}_{\text{large}}, \mathbf{V}_{\text{small}}, \mathbf{V}_{\text{diff}} \end{aligned} \tag{28}$$

We generated 48 OD subsets (8 explanatory variables pla) and created discrete probability distributions. The discrete probability distributions had 6 intervals  $[0^\circ, 30^\circ]$ ,  $[30^\circ, 60^\circ]$ ,  $[60^\circ, 90^\circ]$ ,  $[90^\circ, 120^\circ]$ ,  $[120^\circ, 150^\circ]$ , and  $[150^\circ, 180^\circ]$ . Finally, we measured the

difference of the distribution of subsets from that of the universal set using the KLD.

$$\text{KLD}(\mathbf{P}(\mathbf{k}), \mathbf{U}(\mathbf{k})) = \sum_k \left( p_k \times \log_2 \frac{p_k}{u_k} \right)$$

$k$  :  $[0^\circ, 30^\circ]$ ,  $[30^\circ, 60^\circ]$ ,  $[60^\circ, 90^\circ]$ ,  $[90^\circ, 120^\circ]$ ,  $[120^\circ, 150^\circ]$ , and  $[150^\circ, 180^\circ]$ .

$$\mathbf{P}(\mathbf{k}) \triangleq \begin{cases} P_{[0^\circ, 30^\circ]} \\ P_{[30^\circ, 60^\circ]} \\ P_{[60^\circ, 90^\circ]} \\ P_{[90^\circ, 120^\circ]} \\ P_{[120^\circ, 150^\circ]} \\ P_{[150^\circ, 180^\circ]} \end{cases} \quad \mathbf{U}(\mathbf{k}) \triangleq \begin{cases} u_{[0^\circ, 30^\circ]} \\ u_{[30^\circ, 60^\circ]} \\ u_{[60^\circ, 90^\circ]} \\ u_{[90^\circ, 120^\circ]} \\ u_{[120^\circ, 150^\circ]} \\ u_{[150^\circ, 180^\circ]} \end{cases}$$

$$\begin{aligned} \mathbf{P}_{[0,30]}(\mathbf{od}) &\triangleq \{0 \leq \mathbf{od} \leq 30, \mathbf{od} \in \mathbf{P}(\mathbf{od})\} & p_{[0,30]} &\triangleq \frac{n(\mathbf{P}_{[0,30]}(\mathbf{od}))}{n(\mathbf{P}(\mathbf{od}))} \\ \mathbf{P}_{[30,60]}(\mathbf{od}) &\triangleq \{30 \leq \mathbf{od} \leq 60, \mathbf{od} \in \mathbf{P}(\mathbf{od})\} & p_{[30,60]} &\triangleq \frac{n(\mathbf{P}_{[30,60]}(\mathbf{od}))}{n(\mathbf{P}(\mathbf{od}))} \\ \mathbf{P}_{[60,90]}(\mathbf{od}) &\triangleq \{60 \leq \mathbf{od} \leq 90, \mathbf{od} \in \mathbf{P}(\mathbf{od})\} & p_{[60,90]} &\triangleq \frac{n(\mathbf{P}_{[60,90]}(\mathbf{od}))}{n(\mathbf{P}(\mathbf{od}))} \\ \mathbf{P}_{[90,120]}(\mathbf{od}) &\triangleq \{90 \leq \mathbf{od} \leq 120, \mathbf{od} \in \mathbf{P}(\mathbf{od})\} & p_{[90,120]} &\triangleq \frac{n(\mathbf{P}_{[90,120]}(\mathbf{od}))}{n(\mathbf{P}(\mathbf{od}))} \\ \mathbf{P}_{[120,150]}(\mathbf{od}) &\triangleq \{120 \leq \mathbf{od} \leq 150, \mathbf{od} \in \mathbf{P}(\mathbf{od})\} & p_{[120,150]} &\triangleq \frac{n(\mathbf{P}_{[120,150]}(\mathbf{od}))}{n(\mathbf{P}(\mathbf{od}))} \\ \mathbf{P}_{[150,180]}(\mathbf{od}) &\triangleq \{150 \leq \mathbf{od} \leq 180, \mathbf{od} \in \mathbf{P}(\mathbf{od})\} & p_{[150,180]} &\triangleq \frac{n(\mathbf{P}_{[150,180]}(\mathbf{od}))}{n(\mathbf{P}(\mathbf{od}))} \end{aligned} \tag{29}$$

$$\begin{aligned} \mathbf{U}_{[0,30]}(\mathbf{od}) &\triangleq \{0 \leq \mathbf{od} \leq 30, \mathbf{od} \in \mathbf{U}(\mathbf{od})\} & u_{[0,30]} &\triangleq \frac{n(\mathbf{U}_{[0,30]}(\mathbf{od}))}{n(\mathbf{U}(\mathbf{od}))} \\ \mathbf{U}_{[30,60]}(\mathbf{od}) &\triangleq \{30 \leq \mathbf{od} \leq 60, \mathbf{od} \in \mathbf{U}(\mathbf{od})\} & u_{[30,60]} &\triangleq \frac{n(\mathbf{U}_{[30,60]}(\mathbf{od}))}{n(\mathbf{U}(\mathbf{od}))} \\ \mathbf{U}_{[60,90]}(\mathbf{od}) &\triangleq \{60 \leq \mathbf{od} \leq 90, \mathbf{od} \in \mathbf{U}(\mathbf{od})\} & u_{[60,90]} &\triangleq \frac{n(\mathbf{U}_{[60,90]}(\mathbf{od}))}{n(\mathbf{U}(\mathbf{od}))} \\ \mathbf{U}_{[90,120]}(\mathbf{od}) &\triangleq \{90 \leq \mathbf{od} \leq 120, \mathbf{od} \in \mathbf{U}(\mathbf{od})\} & u_{[90,120]} &\triangleq \frac{n(\mathbf{U}_{[90,120]}(\mathbf{od}))}{n(\mathbf{U}(\mathbf{od}))} \\ \mathbf{U}_{[120,150]}(\mathbf{od}) &\triangleq \{120 \leq \mathbf{od} \leq 150, \mathbf{od} \in \mathbf{U}(\mathbf{od})\} & u_{[120,150]} &\triangleq \frac{n(\mathbf{U}_{[120,150]}(\mathbf{od}))}{n(\mathbf{U}(\mathbf{od}))} \\ \mathbf{U}_{[150,180]}(\mathbf{od}) &\triangleq \{150 \leq \mathbf{od} \leq 180, \mathbf{od} \in \mathbf{U}(\mathbf{od})\} & u_{[150,180]} &\triangleq \frac{n(\mathbf{U}_{[150,180]}(\mathbf{od}))}{n(\mathbf{U}(\mathbf{od}))} \\ \mathbf{P} &: 48 \text{ OD subsets (8 explanatory variables pla)} \end{aligned}$$

In addition to the discrete probability distributions with 6 intervals, we also generated probably distributions with 4 or 9 intervals (See Fig. S2 for details).

**Measurement of the difference of the directivity angle between high speed and low speed**

The index of the directivity angle was defined as the angle between the direction of movement of a focal fish and the angle perpendicular to the display. The index of directivity was 0 if the fish moved directly towards the visual cue, and it was 180 if the fish moved backwards. We determined the high speed-movement for each training fish in each set. Each set comprised 10 periods (1 per second), and high speed was defined as the 2 periods with the fastest movement out of the 10 periods. Low speed was defined as the speed of movement in the remaining 8 periods. We calculated the mean directivity for high-speed movement and that for low-speed movement.

**Visual associative learning using a four-fish group (1 demonstrator and 3 observers)**

To generate and identify a demonstrator, we labeled the fish with silicon ink as described above. During the 3 days after the injection, each fish was kept in a large opaque cup. We then trained each fish separately to learn the association between the visual cue and food reward for 3 days. For the training, the inter-trial interval was 195 s and each fish received 14 training trials. The other manipulations were the same as in the first experiment. We trained 24 fish and selected 12 fish that demonstrated approach behavior at least 6 times in response to the visual cue as demonstrators. Observers were kept singly in a large cup and fed the same amount of food as the trained fish.

The day before the group training trials began, we trained the trained fish at least twice with 14 trials and a 195-s inter-trial interval. The observers were kept singly in a large cup and fed as much food as the demonstrators.

The next day, the trained fish and untrained fish were placed in the training-test tank. To select naive fish as observers, we presented the visual cue 6 times for 20 s with 160-s intervals. We selected the fish that did not shift to the visual cue based on observation of the demonstrator fish. Demonstrator fish were selected based on the size of the untrained fish to make the size of all four individuals in a group as uniform as possible. Before putting the four fish into the training-test tank, we trained the trained fish 4 times with a 195-s inter-trial interval. The one demonstrator fish and three observer fish were placed into the training-test tank together. At 180 s after the fish were placed into the tank, the visual cue was presented twice for 20 s with a 160-s inter-trial interval. Next, the visual cue was presented for 20 s, and then 120 brine shrimp were immediately supplied to the tank. This procedure was repeated three times with a 160-s inter-trial interval (Fig. 6A).

**Calculation procedure of the mean of the orientation difference (OD) when the focus fish is changed**

(A) The OD among the four fish. The time difference (Td) is the time lag of Fish i from Fish j<sub>1</sub>~j<sub>3</sub> (as a standard). After the calculation of metrics (15)–(21), (23)–(24), we calculated the two metrics,  $od(td)_{mean}^{high}$  and  $od(td)_{mean}^{low}$ .

$$od(td)_{mean}^{high} \triangleq \frac{\sum_{t \in T(i_1)^{high}} od(i_1, t, td)_{mean} + \sum_{t \in T(i_2)^{high}} od(i_2, t, td)_{mean} + \sum_{t \in T(i_3)^{high}} od(i_3, t, td)_{mean} + \sum_{t \in T(i_4)^{high}} od(i_4, t, td)_{mean}}{n(T(i_1)^{high}) + n(T(i_2)^{high}) + n(T(i_3)^{high}) + n(T(i_4)^{high})} \quad (30)$$

$$od(td)_{mean}^{low} \triangleq \frac{\sum_{t \in T(i_1)^{low}} od(i_1, t, td)_{mean} + \sum_{t \in T(i_2)^{low}} od(i_2, t, td)_{mean} + \sum_{t \in T(i_3)^{low}} od(i_3, t, td)_{mean} + \sum_{t \in T(i_4)^{low}} od(i_4, t, td)_{mean}}{n(T(i_1)^{low}) + n(T(i_2)^{low}) + n(T(i_3)^{low}) + n(T(i_4)^{low})} \quad (31)$$

(B) The focal fish (Fish i) is the demonstrator and the other three fish are Fish j<sub>1</sub>~j<sub>3</sub>. The time difference (Td) is the time lag of Fish i from Fish j<sub>1</sub>~j<sub>3</sub> (as a standard). We calculated the two metrics,  $od(td)_{mean}^{high}$  and  $od(td)_{mean}^{low}$ .

$$od(i, j_k, t, td) \triangleq \arccos \frac{\begin{pmatrix} v_x(i, t + td) \\ v_y(i, t + td) \end{pmatrix} \cdot \begin{pmatrix} v_x(j_k, t) \\ v_y(j_k, t) \end{pmatrix}}{\left| \begin{pmatrix} v_x(i, t + td) \\ v_y(i, t + td) \end{pmatrix} \right| \left| \begin{pmatrix} v_x(j_k, t) \\ v_y(j_k, t) \end{pmatrix} \right|} \quad (32)$$

$$od_{mean}(i, t, td) \triangleq \frac{1}{3} \{od(i, j_1, t, td) + od(i, j_2, t, td) + od(i, j_3, t, td)\} \quad (33)$$

$$T \triangleq \{\text{all the concerning time, } \frac{1}{3} \text{ s interval}\} \quad (34)$$

$$V \triangleq \{v(l, t) | t \in T \text{ and } l = i, j_1, j_2, j_3\} \quad (35)$$

$$V^{high} \triangleq \{v(l, t) | \text{from the upper } 1/6 \text{ to the upper } 0/6\} \quad (36)$$

$$T(i)^{high} \triangleq \{t | v(i, t) \in V^{high}\} \quad (37)$$

$$od(td)_{mean}^{high} \triangleq \frac{1}{n(T(i)^{high})} \sum_{t \in T(i)^{high}} od(i, t, td)_{mean} \quad (38)$$

$$V^{low} \triangleq \{v(l, t) | \text{from the upper } 6/6 \text{ to the upper } 1/6\} \quad (39)$$

$$T(i)^{low} \triangleq \{t | v(i, t) \in V^{low}\} \quad (40)$$

$$od(td)_{mean}^{low} \triangleq \frac{1}{n(T(i)^{low})} \sum_{t \in T(i)^{low}} od(i, t, td)_{mean} \quad (41)$$

(C) The focal fish is the observer (either of Fish i<sub>1</sub>~i<sub>3</sub>) and the other fish is the demonstrator Fish j. The time difference (Td) is the time lag of Fish i<sub>k</sub> from Fish j (as a standard). We calculated the two metrics,  $od(td)_{mean}^{high}$  and  $od(td)_{mean}^{low}$ .

$$od(i_k, j, t, td) \triangleq \arccos \frac{\begin{pmatrix} v_x(i_k, t + td) \\ v_y(i_k, t + td) \end{pmatrix} \cdot \begin{pmatrix} v_x(j, t) \\ v_y(j, t) \end{pmatrix}}{\left| \begin{pmatrix} v_x(i_k, t + td) \\ v_y(i_k, t + td) \end{pmatrix} \right| \left| \begin{pmatrix} v_x(j, t) \\ v_y(j, t) \end{pmatrix} \right|} \quad (42)$$

$$T \triangleq \{\text{all the concerning time, } \frac{1}{3} \text{ s interval}\} \quad (43)$$

$$V \triangleq \{v(l, t) | t \in T \text{ and } l = i_1, i_2, i_3, j\} \quad (44)$$

$$V^{high} \triangleq \{v(l, t) | \text{from the upper } 1/6 \text{ to the upper } 0/6\} \quad (45)$$

$$\mathbf{T}^{\text{high}}(i_k) \triangleq \{t | v(i_k, t) \in \mathbf{V}^{\text{high}}\} \quad (46)$$

$$\mathbf{V}^{\text{high}} \triangleq \{v(l, t) | \text{from the upper } 6/6 \text{ to the upper } 5/6\} \quad (58)$$

$$\text{od}(\text{td})_{\text{mean}}^{\text{high}} \triangleq \frac{\sum_{t \in \mathbf{T}(i_1)^{\text{high}}} \text{od}(i_1, t, \text{td})_{\text{mean}} + \sum_{t \in \mathbf{T}(i_2)^{\text{high}}} \text{od}(i_2, t, \text{td})_{\text{mean}} + \sum_{t \in \mathbf{T}(i_3)^{\text{high}}} \text{od}(i_3, t, \text{td})_{\text{mean}}}{n(\mathbf{T}(i_1)^{\text{high}}) + n(\mathbf{T}(i_2)^{\text{high}}) + n(\mathbf{T}(i_3)^{\text{high}})} \quad (47)$$

$$\mathbf{T}(i)^{\text{low}} \triangleq \{t | v(i, t) \in \mathbf{V}^{\text{low}}\} \quad (59)$$

$$\mathbf{V}^{\text{low}} \triangleq \{v(l, t) | \text{from the upper } 6/6 \text{ to the upper } 1/6\} \quad (48)$$

$$\text{od}(\text{td})_{\text{mean}}^{\text{low}} \triangleq \frac{\sum_{t \in \mathbf{T}(i_1)^{\text{low}}} \text{od}(i_1, t, \text{td})_{\text{mean}} + \sum_{t \in \mathbf{T}(i_2)^{\text{low}}} \text{od}(i_2, t, \text{td})_{\text{mean}} + \sum_{t \in \mathbf{T}(i_3)^{\text{low}}} \text{od}(i_3, t, \text{td})_{\text{mean}}}{n(\mathbf{T}(i_1)^{\text{low}}) + n(\mathbf{T}(i_2)^{\text{low}}) + n(\mathbf{T}(i_3)^{\text{low}})} \quad (60)$$

$$\mathbf{T}^{\text{low}}(i_k) \triangleq \{t | v(i_k, t) \in \mathbf{V}^{\text{low}}\} \quad (49)$$

$$\text{od}(\text{td})_{\text{mean}}^{\text{low}} \triangleq \frac{\sum_{t \in \mathbf{T}(i_1)^{\text{low}}} \text{od}(i_1, t, \text{td})_{\text{mean}} + \sum_{t \in \mathbf{T}(i_2)^{\text{low}}} \text{od}(i_2, t, \text{td})_{\text{mean}} + \sum_{t \in \mathbf{T}(i_3)^{\text{low}}} \text{od}(i_3, t, \text{td})_{\text{mean}}}{n(\mathbf{T}(i_1)^{\text{low}}) + n(\mathbf{T}(i_2)^{\text{low}}) + n(\mathbf{T}(i_3)^{\text{low}})} \quad (50)$$

(D) The focal fish is an observer (either of Fish  $i_1 \sim i_3$ ) and the other two observers Fish  $j_1 \sim j_2$ . The time difference (Td) is the time lag of Fish  $i_k$  from Fish  $j_i$  (as a standard). We calculated the two metrics,  $\text{od}(\text{td})_{\text{mean}}^{\text{high}}$  and  $\text{od}(\text{td})_{\text{mean}}^{\text{low}}$ .

$$\text{od}(i_m, j_k, t, \text{td}) \triangleq \arccos \frac{\begin{pmatrix} v_x(i_m, t + \text{td}) \\ v_y(i_m, t + \text{td}) \end{pmatrix} \cdot \begin{pmatrix} v_x(j_k, t) \\ v_y(j_k, t) \end{pmatrix}}{\left\| \begin{pmatrix} v_x(i_m, t + \text{td}) \\ v_y(i_m, t + \text{td}) \end{pmatrix} \right\| \left\| \begin{pmatrix} v_x(j_k, t) \\ v_y(j_k, t) \end{pmatrix} \right\|}} \quad (51)$$

$$\text{od}(i_m, t, \text{td})_{\text{mean}} \triangleq \frac{1}{2} \{ \text{od}(i_m, j_1, t, \text{td}) + \text{od}(i_m, j_2, t, \text{td}) \} \quad (52)$$

$$\mathbf{T} \triangleq \{ \text{all the concerning time, } \frac{1}{3} \text{ s interval} \} \quad (53)$$

$$\mathbf{V} \triangleq \{v(l, t) | t \in \mathbf{T} \text{ and } l = \text{demonstrator}, i_1, i_2, i_3\} \quad (54)$$

$$\mathbf{V}^{\text{high}} \triangleq \{v(l, t) | \text{from the upper } 1/6 \text{ to the upper } 0/6\} \quad (55)$$

$$\mathbf{T}(i_k)^{\text{high}} \triangleq \{t | v(i_k, t) \in \mathbf{V}^{\text{high}}\} \quad (56)$$

$$\text{od}(\text{td})_{\text{mean}}^{\text{high}} \triangleq \frac{\sum_{t \in \mathbf{T}(i_1)^{\text{high}}} \text{od}(i_1, t, \text{td})_{\text{mean}} + \sum_{t \in \mathbf{T}(i_2)^{\text{high}}} \text{od}(i_2, t, \text{td})_{\text{mean}} + \sum_{t \in \mathbf{T}(i_3)^{\text{high}}} \text{od}(i_3, t, \text{td})_{\text{mean}}}{n(\mathbf{T}(i_1)^{\text{high}}) + n(\mathbf{T}(i_2)^{\text{high}}) + n(\mathbf{T}(i_3)^{\text{high}})} \quad (57)$$

### Supporting Information

#### Figure S1 Subsets with the five highest KLD scores. (A)

Subset  $\mathbf{V}_{\text{large}}^{\text{1st}}$  (od): If either of the two fish were swimming at high speed, they tended to swim in the same direction (0–90 degree angle). (B) Subset  $\mathbf{X}_{\text{small}}^{\text{1st}}$  (od): If two fish were near the long axis of the wall, they tended to move in the same or opposite direction, and the probability of a 0–30 degree angle was higher than that for a 150–180 degree angle. (C) Subset  $\mathbf{Y}_{\text{large}}^{\text{6th}}$  (od) and Subset  $\mathbf{Y}_{\text{large}}^{\text{5th}}$  (od): If two fish were far from the short axis of the wall, they tended to move in the same or opposite direction, and the probability of a 0–30 degree angle was higher than that for a 150–180 degree angle. (D) Subset  $\mathbf{Y}_{\text{large}}^{\text{1st}}$  (od): If either of the two fish located near the short axis of the wall, their orientation difference tended to be 90 degrees. (TIF)

#### Figure S2 KLD and p value based on other subdivisions (3 and 9).

We represent the probability distribution using a heat map. The OD histogram-intervals with a higher proportion than that of all combinations are shown in green, while those with a lower proportion are shown in magenta. The KLD value and KLD rank are indicated below each column. Below them, p values for the G-test are shown. (a) 3 subsets division based 8 explanatory variables. (b) 9 subsets division based 8 explanatory variables. (TIF)

#### Figure S3 KLD and p value when the discrete probability distribution has 4 or 9 intervals. (A) 9-intervals.

We represent the probability distribution using a heat map. The OD distribution has 9 intervals;  $[0^\circ\text{--}20^\circ]$ ,  $[20^\circ\text{--}40^\circ]$ ,  $[40^\circ\text{--}60^\circ]$ ,  $[60^\circ\text{--}80^\circ]$ ,  $[80^\circ\text{--}100^\circ]$ ,  $[100^\circ\text{--}120^\circ]$ ,  $[120^\circ\text{--}140^\circ]$ ,  $[140^\circ\text{--}160^\circ]$ , and  $[160^\circ\text{--}180^\circ]$ . The OD histogram-intervals with a higher proportion than that of all combinations are shown in green, while those with a lower proportion are shown in magenta. The KLD value and KLD rank are indicated below each column. Below them, p values for the G-test are shown. (B) 4-intervals. The probability distribution has 4 intervals;  $[0^\circ\text{--}45^\circ]$ ,  $[45^\circ\text{--}90^\circ]$ ,  $[90^\circ\text{--}135^\circ]$ , and  $[135^\circ\text{--}180^\circ]$ . (C) Comparison of p value on G-test. The p values under 0.05 are colored blue. (TIF)

#### Table S1 Summary for p values in Figs. 1, 6, 7, and 8.

A p value less than 0.01 is indicated in red, and that less than 0.05 is indicated in blue. (TIF)

## Data availability

The raw data are available to readers by contacting the corresponding author.

## Permit

The work in this paper was conducted using protocols approved by the Animal Care and Use Committee of the University of Tokyo.

## References

- Krause J, Ruxton GD (2002) *Living in groups*: Oxford University Press.
- Heyes CM (1994) Social learning in animals – Categories and mechanisms. *Biol Reviews* 69: 207–231.
- Coolen I, Dangles O, Casas J (2005) Social learning in noncolonial insects? *Curr Biol* 15: 1931–1935.
- Langridge EA, Franks NR, Sendova-Franks AB (2004) Improvement in collective performance with experience in ants. *Behav Ecol Sociobiol* 56: 523–529.
- Brown C, Laland KN (2003) Social learning in fishes: a review. *Fish and Fisheries* 4: 280–288.
- Johannessen LE, Slagsvold T, Hansen BT (2006) Effects of social rearing conditions on song structure and repertoire size: experimental evidence from the field. *Anim Behav* 72: 83–95.
- Thornton A, Clutton-Brock T (2011) Social learning and the development of individual and group behaviour in mammal societies. *Philos Trans R Soc Lond B Biol Sci* 366: 978–987.
- Hoare DJ, Krause J (2003) Social organisation, shoal structure and information transfer. *Fish and Fisheries* 4: 269–279.
- Reznikova ZI (2004) Social learning in animals: Comparative analysis of different forms and levels. *Zh Obshch Biol* 65: 136–152.
- Lachlan RF, Crooks L, Laland KN (1998) Who follows whom? Shoaling preferences and social learning of foraging information in guppies. *Anim Behav* 56: 181–190.
- Grüter C, Farina WM (2009) The honeybee waggle dance: can we follow the steps? *Trends Ecol Evol* 24: 242–247.
- Esch HE, Zhang S, Srinivasan MV, Tautz J (2001) Honeybee dances communicate distances measured by optic flow. *Nature* 411: 581–583.
- Leadbeater E, Chittka L (2007) Social learning in insects – from miniature brains to consensus building. *Curr Biol* 17: R703–R713.
- Reader SM, Laland KN (2000) Diffusion of foraging innovations in the guppy. *Anim Behav* 60: 175–180.
- Swaney W, Kendal J, Capon H, Brown C, Laland KN (2001) Familiarity facilitates social learning of foraging behaviour in the guppy. *Anim Behav* 62: 591–598.
- Reebs SG (2000) Can a minority of informed leaders determine the foraging movements of a fish shoal? *Anim Behav* 59: 403–409.
- Imada H, Hoki M, Suchiro Y, Okuyama T, Kurabayashi D, et al. (2010) Coordinated and cohesive movement of two small conspecific fish induced by eliciting a simultaneous optomotor response. *PLoS One* 5: e11248.
- Kullback S, Leibler RA (1951) On information and sufficiency. *Ann Math Statist* 22: 79–86.
- Cover TM, Thomas JA (1991) *Elements of information theory*. New York: Wiley.
- Hershey RJ, Olsen AP (2007) Approximating the Kullback Leibler divergence between Gaussian mixture models. *Proceedings, IEEE International Conference on Acoustics, Speech, and Signal (ICASSP)* 4: IV-317–IV-320.
- Ramírez J, Segura JC, Benítez C, la Torre A, Rubio JA (2004) A new Kullback-Leibler VAD for speech recognition in noise. *IEEE signal process. Lett.* 11: 266–269.
- Mei Q, Zhai C (2005) Discovering evolutionary theme patterns from text: an exploration of temporal text mining. *Proceedings of the Eleventh ACM SIGKDD International Conference on Knowledge Discovery in Data Mining*: 198–207.
- Mann RP, Perna A, Strömbom D, Garnett R, Herbert-Read JE, et al. (2012) Multi-scale inference of interaction rules in animal groups using Bayesian model selection. *PLoS computational biology* 8: e1002308.
- Pitcher TJ, House AC (1987) Foraging rules for group feeders: area copying depends upon food density in shoaling goldfish. *Ethology* 76: 161–167.
- Webster MM, Laland KN (2008) Social learning strategies and predation risk: minnows copy only when using private information would be costly. *Proc Biol Sci* 275: 2869–2876.
- Ryer CH, Olla BL (1992) Social mechanisms facilitating exploitation of spatially variable ephemeral food patches in a pelagic marine fish. *Anim Behav* 44: 69–74.
- Pike TW, Laland KN (2010) Conformist learning in nine-spined sticklebacks' foraging decisions. *Biol Lett* 6: 466–468.
- van Bergen Y, Coolen I, Laland KN (2004) Nine-spined sticklebacks exploit the most reliable source when public and private information conflict. *Proc Biol Sci* 271: 957–962.
- Reader SM, Kendal JR, Laland KN (2003) Social learning of foraging sites and escape routes in wild Trinidadian guppies. *Animal Behaviour* 66: 729–739.
- Swaney W, Kendal J, Capon H, Brown C, Laland KN (2001) Familiarity facilitates social learning of foraging behaviour in the guppy. *Anim Behav* 62: 591–598.
- Couzin ID, Krause J, James R, Ruxton GD, Franks NR (2002) Collective memory and spatial sorting in animal groups. *J Theor Biol* 218: 1–11.
- Faria JJ, Dyer JRG, Clement RO, Couzin ID, Holt N, et al. (2010) A novel method for investigating the collective behaviour of fish: introducing 'Robofish'. *Behav Ecol Sociobiol* 64: 1211–1218.
- Mineka S, Cook M (2006) Mechanisms involved in the observational conditioning of fear. *J Exp Psychol Gen* 122: 23–38.
- Olsson A, Nearing KI, Phelps EA (2007) Learning fears by observing others: the neural systems of social fear transmission. *Soc Cogn Affect Neurosci* 2: 3–11.
- Laland KN, Reader SM (1999) Foraging innovation in the guppy. *Anim Behav* 57: 331–340.
- CoussiKorbel S, Frigaszy DM (1995) On the relation between social dynamics and social learning. *Anim Behav* 50: 1441–1453.
- Cachat J, Stewart A, Utterback E, Hart P, Gaikwad S, et al. (2011) Three-dimensional neurophenotyping of adult zebrafish behavior. *PLoS ONE* 6(3): e17597.
- Nagy M, Akos Z, Biro D, Vicsek T (2010) Hierarchical group dynamics in pigeon flocks. *Nature* 464: 890–893.

## Acknowledgments

We thank K, Hashimoto and D, Kurabayashi for discussion.

## Author Contributions

Conceived and designed the experiments: TO HT. Performed the experiments: TO HT. Analyzed the data: TO. Contributed reagents/materials/analysis tools: TO YS HT. Wrote the paper: TO KN TK HT.