# Quantification of Aggregation and Associated Brain Areas in *Drosophila melanogaster*

Takuto Okuno
*Graduate School of Life Sciences*
*Tohoku University*
Sendai, Japan
okuno.takuto.t5@dc.tohoku.ac.jp

Koichi Hashimoto
*Graduate School of Information Sciences*
*Tohoku University*
Sendai, Japan
koichi@m.tohoku.ac.jp

Hiromu Tanimoto
*Graduate School of Life Sciences*
*Tohoku University*
Sendai, Japan
hiromut@m.tohoku.ac.jp

***Abstract*— Social behavior requires the interaction among animals. Aggregation, spatial assembly of animals, facilitates inter-individual communication, and provides occasions of social interaction. To quantify aggregation in small animals such as insects it is necessary to detect individual animals at a high precision. By expanding the utility of recently developed machine vision software, we here track multiple individual *Drosophila melanogaster* and provide new metrics for quantifying aggregation. Flies in a circular arena immediately formed aggregation that is significantly higher than random distribution, and it developed over the course of 30 minutes. As the first step toward the identification of neural circuits underlying aggregation, we analyzed large trajectory data (18,992 videos), where 2,204 groups of neurons were genetically stimulated. Large-scale correlations of behavioral performance and labelled neurons identified brain areas associated with aggregation and isolation.**

***Keywords—aggregation, machine vision, brain areas, Drosophila melanogaster***

## I. INTRODUCTION

In order for most animals to survive and propagate, interaction among individuals is required. There are many types of inter-individual interaction, such as courtship behavior [1]. Aggregation behaviour, spatial assembly of individuals, facilitates such social interaction, and is considered improve odds of survival [2]. However, little is known about neural mechanisms underlying the formation of aggregation except the contribution of some sensory systems [3].

To address this question, we used the fruit fly *Drosophila melanogaster*, a well-established model organism with unparalleled genetic tractability. *Drosophila* genetics enables to manipulate neuronal activity in freely behaving animals, allowing the identification of neuronal correlates of behaviour. Fruit flies display local aggregation under certain conditions such as on a food source. To mimic such local aggregation, we developed a laboratory method to video record freely walking flies in a simple 2-dimensional experimental arena.

Multiple flies in laboratory experimental setups were reported to gather and form groups [3]. In the social space assay, flies form aggregation at the top of a vertically oriented triangular chamber at 20 to 40 minutes after placing the flies in the chamber [4]. Social space was then quantified based on nearest neighbor distances. This aggregation is significantly more than the sum of singly assayed flies, indicating synergistic inter-individual interaction [4]. Interestingly, aggregation in this behavioral paradigm requires visual input but not aggregation pheromones [4]. As aggregation behavior in this vertical triangular chamber heavily relies on the climbing performance of flies (i.e. negative geotaxis) [5], flies mutant for negative geotaxis also affects 'social space', so that behavioral differences cannot be attributed solely to aggregation. For instance, the dopamine system was claimed to modulate social interaction in flies [6], while it is also reported to modulate negative geotaxis [5]. In order to measure aggregation unambiguously, it is important to eliminate confounding factors from a behavioral paradigm.

Measuring crowdedness is necessary not only to measure animal aggregation, but also applied to the field of pedestrian modelling and traffic engineering. There are multiple measures to quantify the level of crowdedness in a group [7]. The principles of these measures can be categorized into two classes: distance-based and density-based measures. Distance-based metrics relies on distances of particles, while the number of particles in a given area is counted to compute density-based measures [7]. The distance-based metrics to quantify aggregation can capture grouping in most cases, but does not fully reflect aggregation, especially if pairs of particles are dispersed in an area of interest. Considering interactions between two individuals, such as courtship behavior, distance-based measures could become confounding to measure aggregation of animals in particular.

In order to improve these experimental issues, we devised a horizontally laid two-dimensional circular arena for measuring aggregation. For better quantification of aggregation, we introduced a density-based metrics that computes the population density for each fly, which takes the distance into account. With this metrics, we analyzed aggregation in a large number of video data, where effects of activating 2,204 genetically targeted neurons were measured in a circular arena [8]. Using the Browsable Atlas of Behavior-Anatomy Maps (BABAM) software, which the Branson group in Janelia Research Campus recently created [8], we further explored brain regions that explain the behavioral variability of these genetic activations.

978-1-5386-9151-9/19/$31.00 ©2019 IEEE

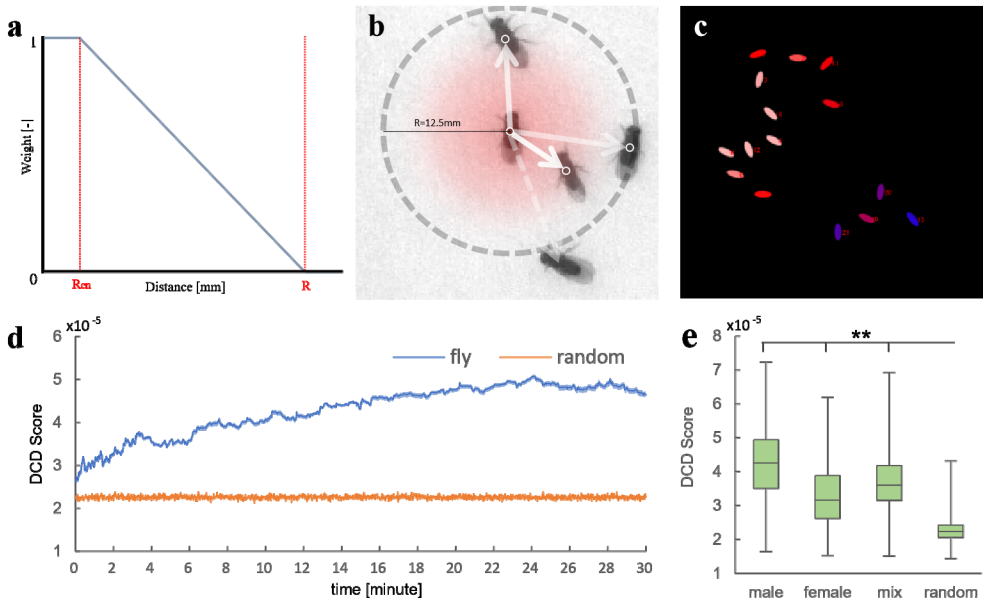


Fig. 1. Crowdedness measure based on the DCD. (a) Visualization of the weight function. (b) Visualization of DCD measure. (c) Visualization of individual DCD score. (d) Time course comparison between random points and male wild type files. The vertical axis is DCD score. (e) experimental result of 20 male, female and 10/10 gender mixed flies

## II. QUANTIFYING AGGREGATION

### A. Needs for new aggregation metrics

The most straightforward approach to calculate aggregate density is to divide the area of interest into grids and to measure the particle density of each grid [7]. When applied to a group of flies, it is possible to quantify density distribution in an experimental arena. Drawback of this method is arbitrariness of grid size selection and noisy values due to integer values (particle counts).

The range-based density computation is similar to the grid-based method, but one density value is assigned to each particle by counting the number of particles surrounding a given particle within a defined radius [7]. Although this metrics allows to estimate how each individual participates in a crowd, changes in the density value are step-wise as is the case for the grid-based metrics. Considering these constraints, we propose a new density measure assigned to each individual, and named Distance-Corrected Density (DCD).

### B. Definition of DCD

DCD substitutes the distance between individuals with the weight value of the conditional expression (Fig. 1a, b), calculates the aggregate density by summing up the values and dividing by the area of the circle (1).

$$local\ density(\overrightarrow{p_i}) = \frac{1}{\pi R^2}\sum_{k=1}^{n} w\left(\left|\overrightarrow{p_i} - \overrightarrow{p_k}\right|\right) \qquad (1)$$

where

$$w(t) = \begin{cases} 1 & \dots\ t \le R_{cn} \\ (R-t)(R-R_{cn}) & \dots\ R_{cn} < t < R \\ 0 & \dots\ R \le t \end{cases}$$

and R is outer radius, $R_{cn}$ is inner (closest neighbor) radius. Because the adult fly body length is around 3mm, empirically we choose 2.5mm as the closest neighbor distance. Otherwise, more than 12.5mm, such as around 4 body lengths, is enough distance as separated individuals.

### C. Experimental setup

The experimental apparatus consists of a horizontal circle arena, a light source from both the front LED light and the back LED light and the camera which records the arena from directly above. The arena is horizontally placed with a plastic cylinder of diameter 8 cm, height 1 cm, the bottom is covered with filter paper, and the top is covered by a glass plate. The inside wall of the cylinder was covered with Fluon (Insect-a-Slip, PTFE30, BioQuip Products) so that flies do not climb it. In addition, Sigmacote (Sigma-Aldrich, Co) was painted on the inside of the upper glass plate so that flies could not stick to it.

The camera was Point Gray CM3-U3-13Y3M-CS and Space com 12mm 1:14 with C-CS mount adaptor was used for the lens. For the capture software, we used Point Gray FlyCap9 and took movies at grayscale 1024x1024 50fps and shutter speed at 1/1500. In this way, it is possible to record high-speed fly movements such as jumping. The temperature of the experimental environment is kept at 23.5-25 degrees Celsius, the humidity is kept around 40%, the illumination of the light source was between 403-963 lux. To minimize the influence of the circadian rhythm on fly activity, the experiments were carried out in ZT4-ZT10 (12:00-18:00).

### D. Experimental protocol for the aggregation assay

1) After eclosion, 100-150 flies of mixed genders were transferred in one food bottle for 4-5 days.
2) Twenty male, female, or mixed-gender (10 males and 10 females) flies were collected without anesthesia and kept in a plastic vial 20 minutes before each experiment.
3) Put the vials in the incubator 10 minutes before each session.
4) Video was recorded for 30 minutes from the introduction of the 20 flies into the arena.

5) After each session, flies are removed from the arena and the filter paper is exchanged.

### E. Tracking software

TPro was used as a tracking software which was originally developed by Pudith Sirigrivatanawong (2017) [9] and many functions were additionally developed. In the previous version, it took some time to split a big blob when the flies were close together, so we optimized extraction of the fly neighborhood region and splitting process. As a result, the detection speed becomes 2-3 times faster than previous version. Additional functions, such as the behavior classification of the fly from the values of velocity, acceleration, side velocity, rotational velocity and circularity of the tracking result, the wing detection which extracts the intensity of a circle whose radius is 0.4 body length from the center of gravity of fly and a function to detect fly's wings, were developed. Then, wing grooming behavior is possible to detect. Aggregate density can also be calculated from detection result. In addition, the detection setting GUI, the result GUI, the annotation GUI were developed, and batch processing of large amount of movies in the command prompt was added. This software is coded in MATLAB script and coded as open source. The source code can be downloaded from below.

https://github.com/takuto-okuno-t5/tpro

Detection and tracking of 20 flies were performed by this software, and the DCD for each fly was calculated from the trajectory data.

### F. Results

Fig. 1c shows a representative snapshot of our aggregation assay, where the DCD score of each fly-fitted ellipse is indicated a color. This result demonstrates that the DCD measure correctly characterizes the crowdedness of flies. The mean of the DCD scores of all flies thus quantifies the level of aggregation in a given video frame.

We next examined the time course of aggregation. Fig. 1d shows the development of mean DCD values over 30 minutes in a populations of 20 male flies (blue line, N=6). For comparison, the mean DCD values of 20 random points which were computer generated in the same circular arena (orange line). Due to the mechanical shock at the introduction of 20 wild-type (Canton-S) flies into the arena, flies initially displayed panic-like high locomotor activity. Correspondingly, little aggregation was formed in the beginning of measurement, which was barely above the random distribution (Fig. 1d). Subsequently, they gradually calmed down and started other behaviors such as grooming and interaction with each other. Correspondingly, flies slowly start to aggregate and form groups over the course of 30 minutes (Fig. 1d).

Fig. 1e shows the mean DCD of wild-type male, female and mixed male and female flies (N=11). All three groups showed significantly higher aggregation than the simulation control ($p<0.01$; Mann-Whitney U-test; Fig 1e). Interestingly, the group of male flies only tend to aggregate faster and more than the female group (Fig. 1e). Aggregation of the mixed males and females was between these two sexed groups (Fig. 1e).

## III. Candidate neural correlates for aggregation

### A. Trajectory data of 400,000 flies with targeted neuronal activation

The research with *D. melanogaster* is much supported by well-developed genetic techniques, with which specific nerve cells can be activated, and behavioral changes can be observed. The most popular and versatile approach is to use the GAL4/UAS modular transgene expression system [10]. The Janelia Fly Light Project generated a large collection of approximately 7000 GAL4 driver lines [11], and imaged expression patterns in the brain using a marker protein for the plasma membrane [12]. The Branson group in the Janelia Research Campus recently published the results of a large-scale behavior screen, where 400,000 flies with targeted neuronal activation were tracked [8]. They used 2,204 GAL4 strains selected from the resource to drive the expression of the temperature sensitive cation channel (dTRPA1), and video recorded the behavioral consequence of the activation of GAL4-labelled neurons in a 2D-circular arena [8]. This assay has many similarities with our aggregation assay; ten male and ten females were placed in a thin circular arena with a diameter of 127mm and a height of 3.5mm and recorded for 16 minutes (resolution 1024x1024 / 30fps) after the temperature elevation [8]. Individual flies were tracked by Ctrax to generate trajectory data [13]. Only trajectory data that contain information such as fly coordinate, direction, body major/minor length, and wing direction, reach 1.8 TB.

### B. Aggregation behavior upon targeted neuronal activations using 2,204 GAL4 lines

We analyzed the trajectory data of 2,204 neuronal activations for aggregation behavior (Fig. 2) . Fig. 2a is the distribution of DCD scores upon temperature-dependent neuronal stimulation for all GAL4 lines. GAL4 lines are ranked with mean DCD scores. While most GAL4 lines show DCD scores similar to the control GAL4 line without enhancers (empty-GAL4; DCD score around 0.0006), 1-2 % of drivers marked high aggregation with double DCD scores (Fig. 2b). On the other hand, few GAL4 lines that tend not to gather showed only minor decrease in DCD scores but with much smaller variance than the control (Fig. 2c). This implies this aggregation level is the floor effect. These lines with lower scores avoid inter-individual communication by detaching from each other.

Representative video frames confirmed marked differences of aggregation levels between R49D01 and R91B01 with the highest and lowest DCD scores, respectively (Fig. 2a). Thermo-activation of neurons labelled in R49D01 induced large aggregate groups. These flies seemed to continuously interact each other within the group. There are flies that occasionally left the group and performed exploratory actions, but they soon tended to join another individual or group. These behaviors can be interpreted as enhanced sociality. In contrast, thermo-activation with R91B01-GAL4 caused clear isolation behavior. These flies constantly walk around and seemed to avoid other flies when entering they entered "personal space".

To examine our annotation, we plotted relative positions of surrounding flies in relation to an individual fly (Fig. 3d-f). These are 2D-density plots of the centroids

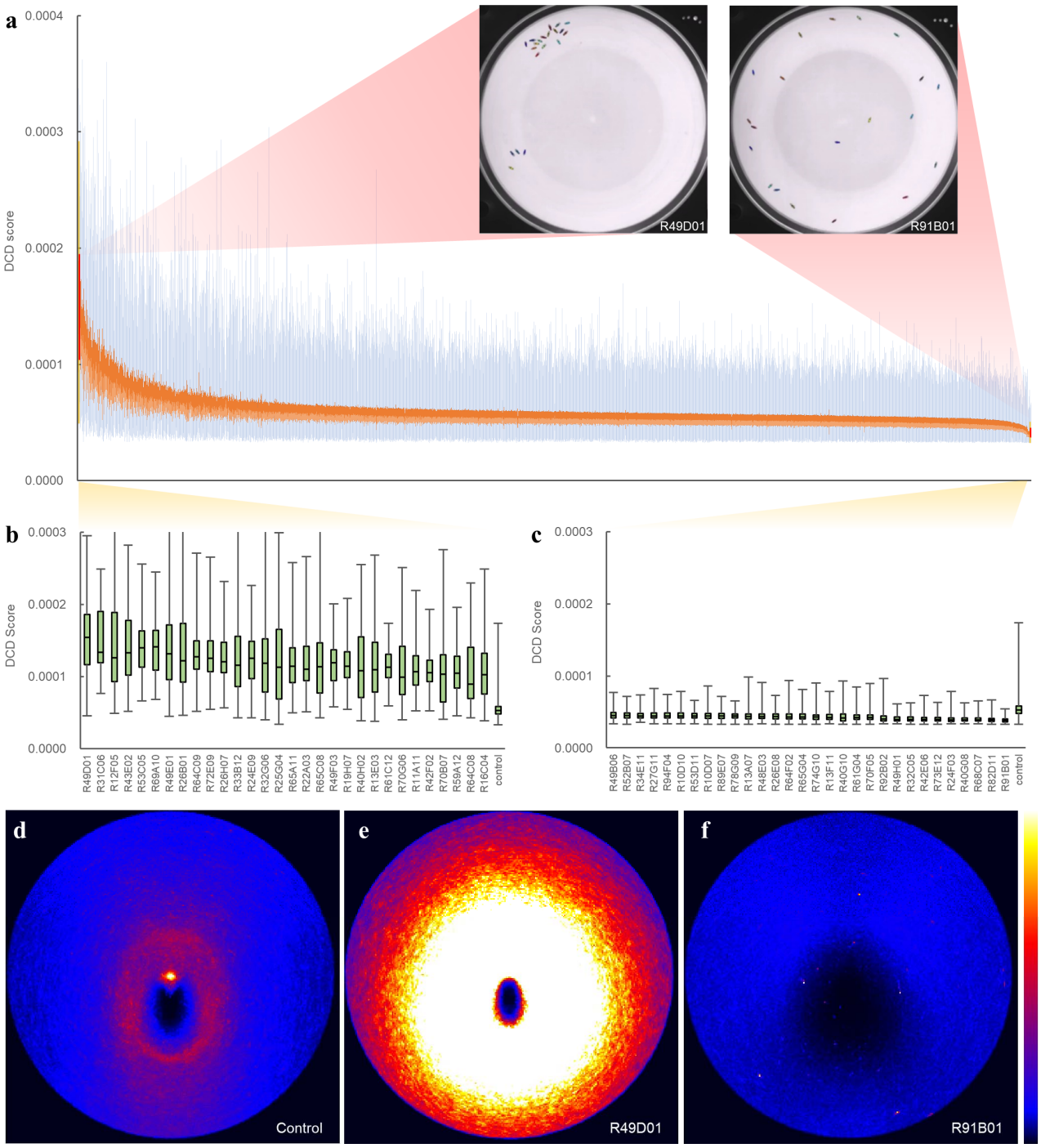


Fig. 2. Analysis result of trajectory data. (a) Whole box-and-whisker plot of 2204 GAL4 lines. Light blue lines shows more than 75% and less than 25% DCD score data. Dark orange shows 50%-75% range of data. Light orange shows 25%-50% range of data. Left picture is a sample snapshot of the R49D01 GAL4 line. Right picture is a sample snapshot of the R91B01 GAL4 line. These representative snapshots are adapted from the Fly Bowl Neural Activation Behavior Effects (http://research.janelia.org/bransonlab/FlyBowl/BehaviorResults/ index.html). (b) Box-and-whisker plot of top 30 GAL4 lines. Graph is sorted by mean value of DCD score. Compared with control (empty GAL4) (c) Box-and-whisker plot of bottom 30 GAL4 lines. (d) Porlar chart of control (empty GAL4) flies. Center is fly head. Radius is 15mm. Plot points of fly head to other files' centroid. 8 videos * 30fps * 15 * 60 * 20 flies = 4,320,000 flies result. (e) Porlar chart of the R49D01. Plot same way as image d. (f) Porlar chart of the R91B01.

of flies within the radius of 15 mm for all frames in 8 videos of each condition. The density plot of the control flies showed clear preferred positions of other flies close to a given fly (Fig. 2b). The belt of high probabilities is skewed to the head direction with the conspicuous peak in front of the head (Fig. 2b). This histogram of relative fly position is consistent with a previous report [13]. Thermo-activation with R49D01 and R91B01 induced remarkable density patterns that were clearly different from the control histogram (Fig. 2e, f). A typical *R49D01-GAL4/UAS-dTRPA1* fly is surrounded by many flies at short distances, suggesting the strong preference to be in a crowd (Fig. 2e).

On the other hand, activation with R91B01 completely eliminated the density belt around a given fly (Fig. 2f). Given a further density decrease in the head direction, avoidance of interaction may rely on visual input.

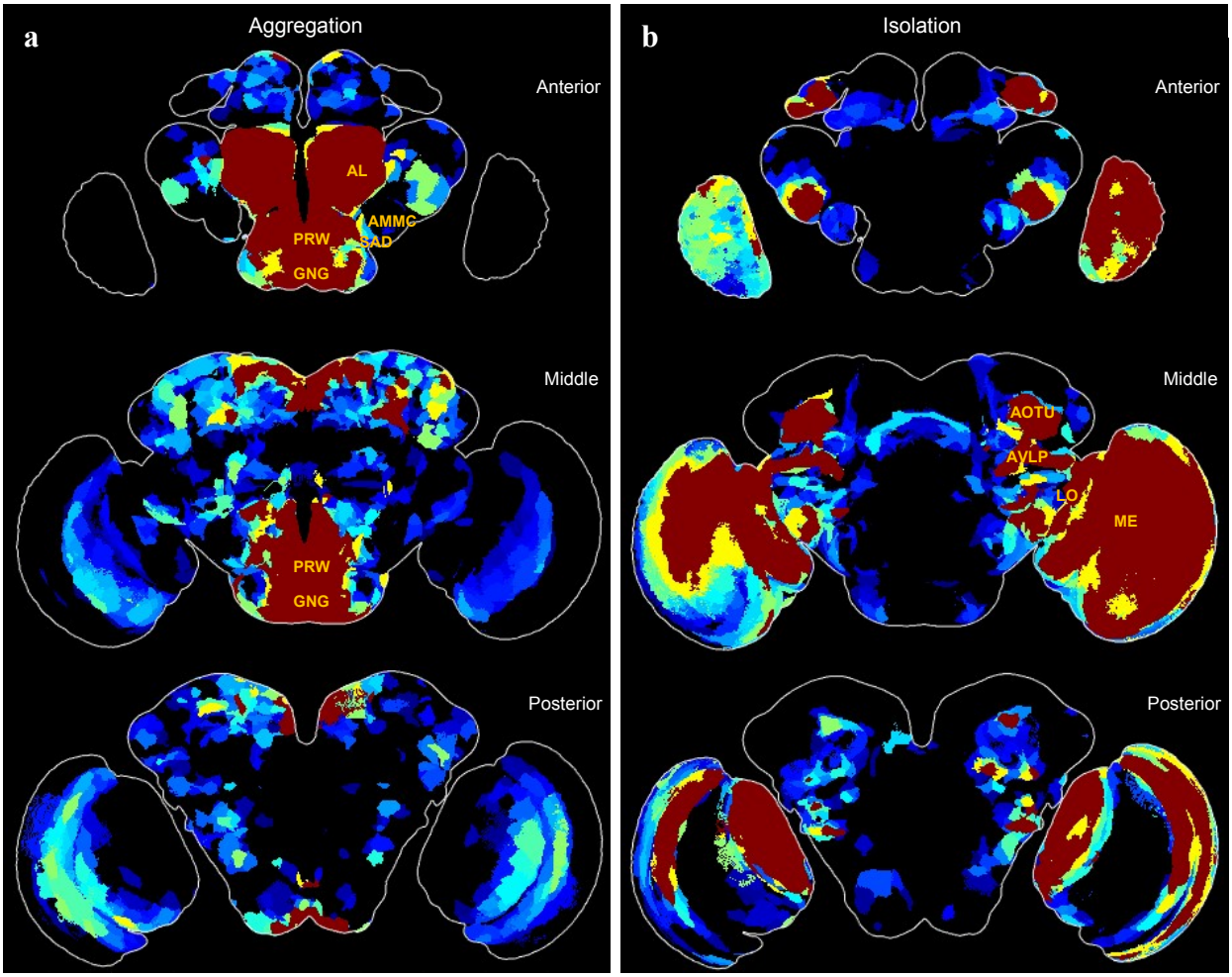


Fig. 3. Supervoxel correlation map images based on analysis result. Images are generated from the BABAM (a) Anterior, middle and posterior images of super voxel correlation map of aggregation. These voxels have bigger value than the DCD mean value of the control. Brown color means strong correlation and blue means weak correlation. (b) Anterior, middle and posterior images of super voxel correlation map of isolation. These voxels have smaller value than the DCD mean value of the control.

### C. *Correlation analysis of behavioral variance and GAL4 expression identifies candidate brain regions for aggregation*

Janelia Fly Light Project imaged the expression patterns of all the 2,204 GAL4 lines tested for behavior, and registered each confocal stack to a standard brain [12]. To quantitatively analyze these GAL4 expressing cells, the Branson group introduced the segmentation of the standard fly brain stack into 7,065 voxel clusters (supervoxels) based on co-labelled voxels [8]. The expression pattern of each GAL4 line can thus be represented by the combinations of supervoxel labels. They further developed the Browsable Atlas of Behavior-Anatomy Maps (BABAM) software to identify supervoxels that correlate with the behavioral variability by genetic activations [8]. BABAM statistically tests for a positive correlation between the behavior measure and the expression of all GAL4 lines in a given supervoxel, and visualizes correlated supervoxels in a standard brain [8].

Using BABAM, we tested supervoxel-wise correlation between aggregation scores and the expression for all 2,204 GAL4 lines (Fig. 3). The deviation of the DCD score of each GAL4 line from that of the control was calculated and added to the BABAM matrix data to finally create a correlation image with supervoxels. Fig. 3a and b show the output 3D-correlation map of BABAM for aggregation and isolation (decreased aggregation), respectively. The supervoxel-wise *P*-values of the correlation tests were plotted in sub-stacks representing the anterior, middle, and posterior part of the brain. For aggregation, GAL4 expression in brain areas including the gnathal ganglia (GNG), and antennal mechanosensory and motor center (AMMC), Prow (PRW), and the antennal lobe (AL) showed significant correlation (Fig. 3a). Indeed, GNG and AMMC received dense projections by neurons that strongly express GAL4 in R49D01.

The BABAM analysis further revealed supervoxels with positive correlation with isolation (Fig. 3b). Strikingly, many supervoxels in the optic lobes and the target areas of visual projection neurons showed high correlations. The brain regions with correlation with isolation include the anterior optic tubercle (AOTU) and anterior ventrolateral protocebrum (AVLP), as well as neuropils in the optic lobes (Fig. 3b). These results suggest that markedly decreased inter-individual association might be genetic activation of avoidance-inducing visual projection neurons [14]. Globally, correlated supervoxels of aggregation and isolation are spatially segregated. Medial brain regions had robust correlations with aggregation, whereas the more lateral neuropils were identified for isolation (Fig. 3).

## IV. DISCUSSION

Spatial aggregation is prevalent in the animal kingdom. It not only facilitates inter-individual communication, but also helps protections from predators, finding food sources, and group behavior. Swarming formation and collective behavior by millions of animals, such as gregarious locusts, can cause leave even enormous economic damages. Despite its prevalence, little is known about underlying neuronal mechanisms. We here reported a behavioral paradigm where multiple individual *Drosophila* can be detected at a high precision using machine vision techniques, and devised a metrics to quantify aggregation behavior (Fig. 1). Flies in a circular arena immediately formed aggregation, which develops over the course of 30 minutes (Fig. 1). Furthermore, we scored aggregation behavior based on previously published large trajectory data (18,992 videos), where 2,204 groups of neurons were genetically stimulated. Large-scale correlations of behavioral performance and labelled neurons identified brain areas associated with aggregation and isolation (Fig. 3).

Sensory input has been shown to be important for aggregation and social interaction. In *Drosophila*, many experiments have thus far been performed to disable sensory systems. For example, the importance of visual input has been examined by testing aggregation in darkness and by examining the performance of mutants for vision [15]. Furthermore, it has been reported that social network formation decreases in darkness [16]. The BABAM analysis for decreased aggregation (isolation) here supports the contribution of vision in social interaction (Fig. 3). Given the spatial distribution of supervoxels with significant correlation with isolation (Fig. 3b), some visual projection neurons seem to be involved in decreased inter-individual association. Visual projection neurons that are involved in collision avoidance would well explain social isolation induced by dTRPA1-activation [14].

In contrast, supervoxels with significant correlation with elevated aggregation were clustered in the medial regions of the brain (Fig. 3a). Considering that many of these regions receive sensory innervations, these results suggest that particular tactile, olfactory and taste input is triggering aggregation behavior. This is consistent with previous reports showing aggregation pheromones and social interaction through mechanosensation [17; 16]. To give an example, male cuticular hydrocarbon 9-Tricosene and its cognate receptor Or7a were shown to mediate aggregation pheromone upon the exposure to appetitive odors [17].

These supervoxels are however just a small fraction of what we have identified as candidate brain areas causing aggregation and behavioral isolation. Therefore, a meaningful next step would be to characterize other more central brain areas for controlling aggregation. Such identification of neurons would contribute to the understanding of circuit functions for social interaction.

## ACKNOWLEDGMENT

We thank A. Robie and K. Branson (Janelia Research Campus) for providing trajectory data. We thank P. Sirigrivatanawong (Tohoku University) for the initial work on TPro software and the source code. We thank T. Ichinose and V. Thoma (Tohoku University) for constructive comments and advices on the work and manuscript. K. Hashimoto (Tohoku University) acknowledges the support by JSPS KAKENHI Grant Number 16H06536.

## REFERENCES

[1] Spieth HT, "Courtship behavior in Drosophila," Annu Rev Entomol, 1974, pp.385-405

[2] Asa Johannesen, Alison M. Dunn, Lesley J. Morrell, "Prey aggregation is an effective olfactory predator avoidance strategy," PeerJ, 2014, e408.

[3] Pavan Ramdya, Jonathan Schneider, Joel D. Levine, "The neurogenetics of group behavior in Drosophila melanogaster," Journal of Experimental Biology, 2017, pp.35-41

[4] Anne F Simon, M‐ T Chou, Evelyn D Salazar, Terriann Nicholson, Navdeep Saini, Stanimir Metchev, David E Krantz, "A simple assay to study social behavior in Drosophila: measurement of social space within a group," Genes, Brain and Behavior, 2012 Mar, pp.243-252

[5] Bainton RJ, Tsai LT, Singh CM, Moore MS, Neckameyer WS, Heberlein U, "Dopamine modulates acute responses to cocaine, nicotine and ethanol in Drosophila," Current Biology, 2000 Feb 24, pp.187-194

[6] Fernandez RW, Akinleye AA, Nurilov M, Feliciano O, Lollar M, Aijuri RR, O'Donnell JM, Simon AF, "Modulation of social space by dopamine in Drosophila melanogaster, but no effect on the avoidance of the Drosophila stress odorant," Biol Lett., 2017 Aug

[7] Dorine C.Duives, WinnieDaamen, Serge P.Hoogendoorn, "Quantification of the level of crowdedness for pedestrian movements," Physica A: Statistical Mechanics and its Applications, Volume 427, 1 June 2015, pp.162-180

[8] Alice A. Robie, Jonathan Hirokawa, Austin W. Edwards, Julie H. Simpson, Michael B. Reiser, Kristin Branson, "Mapping the Neural Substrates of Behavior," Cell, 2017 July 13, pp.393-406

[9] Pudith Sirigrivatanawong, Shogo Arai, Vladimiros Thoma, Koichi Hashimoto, "Multiple Drosophila Tracking System with Heading Direction," Sensors 2017, 17, 96

[10] Brand AH, Perrimon N, "Targeted gene expression as a means of altering cell fates and generating dominant phenotypes," Development, 1993 Jun, pp.401-415

[11] Pfeiffer BD, Jenett A, Hammonds AS, Ngo TT, Misra S, Murphy C, Scully A, Carlson JW, Wan KH, Laverty TR, Mungall C, Svirskas R, Kadonaga JT, Doe CQ, Eisen MB, Celniker SE, Rubin GM, "Tools for neuroanatomy and neurogenetics in Drosophila," Proc Natl Acad Sci U S A, 2008 Jul 15

[12] Jenett A, Rubin GM, Ngo TT, Shepherd D, Murphy C, Dionne H, Pfeiffer BD, Cavallaro A, Hall D, Jeter J, Iyer N, Fetter D, Hausenfluck JH, Peng H, Trautman ET, Svirskas RR, Myers EW, Iwinski ZR, Aso Y, DePasquale GM, Enos A, Hulamm P, Lam SC, Li HH, Laverty TR, Long F, Qu L, Murphy SD, Rokicki K, Safford T, Shaw K, Simpson JH, Sowell A, Tae S, Yu Y, Zugates CT, "A GAL4-driver line resource for Drosophila neurobiology," Cell Rep., 2012 Oct 25, pp.991-1001

[13] Kristin Branson, Alice Robie, John Bender, Pietro Perona, Michael Dickinson, "High-throughput Ethomics in Large Groups of Drosophila," Nature Methods, 2009 Jun, pp.451-457

[14] Saskia E. J. de Vries, Thomas R. Clandinin, "Loom sensitive neurons link computation to action in the Drosophila visual system," Curr Biol., 2012 Mar 6, pp.353–362

[15] Burg ED, Langan ST, Nash HA, "Drosophila social clustering is disrupted by anesthetics and in narrow abdomen ion channel mutants," Genes, Brain and Behavior, 2013 Apr, pp.338-347

[16] Schneider J, Dickinson MH, Levine JD, "Social structures depend on innate determinants and chemosensory processing in Drosophila," Proc Natl Acad Sci U S A., 2012 Oct 16

[17] Chun-Chieh Lin, Katharine A Prokop-Prigge, George Preti, Christopher J Potter, "Food odors trigger Drosophila males to deposit a pheromone that guides aggregation and female oviposition decisions," eLife, 2015, 4