

# Genetic Basis of Natural Variation in Spontaneous Grooming in *Drosophila melanogaster*

Aya Yanagawa,\* Wen Huang,<sup>†</sup> Akihiko Yamamoto,<sup>‡</sup> Ayako Wada-Katsumata,<sup>‡</sup> Coby Schal,<sup>\*</sup> and Trudy F. C. Mackay<sup>§,1</sup>

\*Research Institute for Sustainable Humanosphere, Kyoto University, Uji, Japan, <sup>†</sup>Department of Animal Science, Michigan State University, East Lansing, Michigan 48824, <sup>‡</sup>Department of Entomology and Plant Pathology, North Carolina State University, Raleigh, NC 27695-7613, and <sup>§</sup>Center for Human Genetics and Department of Genetics and Biochemistry, Clemson University, Greenwood, SC 29646

ORCID IDs: 0000-0001-6788-8364 (W.H.); 0000-0002-2312-7245 (T.F.C.M.)

**ABSTRACT** Spontaneous grooming behavior is a component of insect fitness. We quantified spontaneous grooming behavior in 201 sequenced lines of the *Drosophila melanogaster* Genetic Reference Panel and observed significant genetic variation in spontaneous grooming, with broad-sense heritabilities of 0.25 and 0.24 in females and males, respectively. Although grooming behavior is highly correlated between males and females, we observed significant sex by genotype interactions, indicating that the genetic basis of spontaneous grooming is partially distinct in the two sexes. We performed genome-wide association analyses of grooming behavior, and mapped 107 molecular polymorphisms associated with spontaneous grooming behavior, of which 73 were in or near 70 genes and 34 were over 1 kilobase from the nearest gene. The candidate genes were associated with a wide variety of gene ontology terms, and several of the candidate genes were significantly enriched in a genetic interaction network. We performed functional assessments of 29 candidate genes using RNA interference, and found that 11 affected spontaneous grooming behavior. The genes associated with natural variation in *Drosophila* grooming are involved with glutamate metabolism (*Gdh*) and transport (*Eaat*); interact genetically with (*CCKLR-17D1*) or are in the same gene family as (*PGRP-LA*) genes previously implicated in grooming behavior; are involved in the development of the nervous system and other tissues; or regulate the Notch and Epidermal growth factor receptor signaling pathways. Several DGRP lines exhibited extreme grooming behavior. Excessive grooming behavior can serve as a model for repetitive behaviors diagnostic of several human neuropsychiatric diseases.

## KEYWORDS

*Drosophila*  
Genetic  
Reference Panel  
genome wide  
association  
analysis  
RNA interference  
behavioral  
genetics

Animals perform various behaviors in response to their environment and internal state, including feeding, sound production, locomotion, mating, egg laying and self-grooming (Grillner and Wallén 2004; Grillner *et al.* 2005). Almost all animals, including insects, perform self-, mutual or allo-grooming behaviors. Grooming mechanisms

have been studied in solitary insects using physical stimuli such as dust or tactile stimuli (Gratwick 1957; Hlavac 1975; Page and Matheson 2004). Aversive and microbial cues from the environment are important triggers of grooming behavior (Zhukovskaya *et al.* 2013; Yanagawa *et al.* 2017; 2019). In social insects, mutual grooming serves as a hygienic behavior (Yanagawa *et al.* 2008) and also distributes queen pheromone throughout the colony (Naumann 1991; Bahreini and Currie 2015). Insects also groom various body parts to remove noxious chemicals and maintain sensory acuity of the peripheral nervous system (Dethier 1972; Newland 1998; Page and Matheson 2004; Böröczky *et al.* 2013; Yanagawa *et al.* 2014). Recently, the output neural circuit of a mechanically induced grooming behavior, from the brain to motor neurons, has been described in *Drosophila melanogaster* (Seeds *et al.* 2014; Hampel *et al.* 2015; 2017). However, because grooming serves many functions and occurs spontaneously without extrinsic stimuli, the neural mechanisms that

Copyright © 2020 Yanagawa *et al.*

doi: <https://doi.org/10.1534/g3.120.401360>

Manuscript received May 6, 2020; accepted for publication July 28, 2020; published Early Online July 29, 2020.

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Supplemental material available at figshare: <https://doi.org/10.25387/g3.12252719>.

<sup>1</sup>Corresponding author: Self Regional Hall, Clemson University Center for Human Genetics, 114 Gregor Mendel Circle, Greenwood, SC 29646. E-mail: [tmackay@clemson.edu](mailto:tmackay@clemson.edu)

initiate and modulate spontaneous grooming behavior remain poorly understood.

Grooming is an innate behavior that follows a species-typical “stereotyped” sequence of actions by which animals clean themselves. The sequence of grooming behaviors is conserved across closely related species (Sachs 1988; Zhukovskaya *et al.* 2013), indicating that elements of this behavior are genetically based and important for survival. Indeed, mutational analyses and RNA interference (RNAi) have revealed several genes affecting grooming behavior in *D. melanogaster*, including *Hdc*, encoding *Histidine decarboxylase* (Melzig *et al.* 1996); the *Drosophila* ortholog of *Frm1*, which causes Fragile X Syndrome in humans (Tauber *et al.* 2011); *Dop1R1*, which encodes the D1 family dopamine receptor (Pitmon *et al.* 2016); *Nf1*, which encodes *Neurofibromin 1*, a Ras GTPase activating protein (King *et al.* 2016); *PGRP-LC*, which encodes *Peptidoglycan recognition protein LC*, which functions in immune response to bacteria (Yanagawa *et al.* 2017); as well as gustatory receptor genes (Yanagawa *et al.* 2019). However, little is known about the genetic basis of naturally occurring variation affecting the initiation and modulation of spontaneous grooming.

The *Drosophila melanogaster* Genetic Reference Panel (DGRP) is a population of wild-derived inbred fly lines with fully sequenced genomes, and is a community resource for genome wide association (GWA) mapping of variants associated with quantitative traits when all variants are known (Mackay *et al.* 2012; Huang *et al.* 2014). The DGRP harbors significant genetic variation for all quantitative traits assessed to date, including many behavioral traits (Mackay and Huang 2018). Here, we quantified the spontaneous grooming behavior of 201 DGRP lines and demonstrated that population variation in grooming time is genetically variable. We performed a GWA analysis of spontaneous grooming behavior and identified naturally occurring molecular polymorphisms associated with spontaneous grooming, and used RNA interference to functionally validate effects on grooming behavior for genes in which these polymorphisms were located. The candidate genes associated with spontaneous grooming are involved with neurotransmitter metabolism and transport; interact genetically with or are in the same gene family as genes previously implicated in grooming behavior; are involved in the development of the nervous system and other tissues; or regulate the Notch and Epidermal growth factor receptor signaling pathways.

## MATERIALS AND METHODS

### Drosophila stocks

We used 201 DGRP lines that were generated and maintained at North Carolina State University (Mackay *et al.* 2012; Huang *et al.* 2014) to quantify grooming behavior and perform a GWA analysis for this trait.

We obtained UAS-RNAi lines for 29 candidate genes implicated by the GWA analysis from the Vienna *Drosophila* RNAi Center (VDRC): *CCKLR-17D3* (*P*{KK110484}VIE-260B, v102039); *CG13229* (*P*{KK104316}VIE-260, v100433); *CG17839* (*P*{KK105048}VIE-260B, v100149); *CG31221* (*P*{KK113111}VIE-260B, v103017); *CG33181* (*P*{KK107703}VIE-260B, v103142); *CG33926* (*P*{KK112707}VIE-260B; v107542); *CG34401* (*P*{KK108172}VIE-260B, v110636); *CG6767* (*P*{KK107870}VIE-260B, v109894); *Chit6* (*P*{KK104493}VIE-260B, v105536); *Eaat1* (*P*{KK100187}VIE-260B, v109401); *for* (*P*{KK101298}VIE-260B, v108293); *Gdh* (*P*{KK107890}VIE-260B, v109499); *GEFmeso* (*w*<sup>1118</sup> *P*{GD8852}v39952, v39952), *kirre* (*P*{KK107679}VIE-260B, v109585); *Lrcl* (*P*{KK102050}VIE-260B, v107047); *mtt* (*P*{KK112979}VIE-260B, v102982); *Nrg* (*P*{KK100482}VIE-260B, v107991); *osp* (*P*{KK101464}VIE-260B,

v110701), *PGRP-LA* (*P*{KK111407}VIE-260B, v102277); *pnt* (*P*{KK100473}VIE-260B, v105390); *PsGEF* (*P*{KK103731}VIE-260B, v109769); *Ptp10D* (*P*{KK101775}VIE-260B, v110443); *Ptp99A* (*P*{KK102836}VIE-260B, v103931); *pxn* (*P*{KK108816}VIE-260B, v107180); *scat* (*P*{KK100870}VIE-260B, v108854); *SdhA* (*P*{KK101728}VIE-260B, v110440); *Smg5* (*P*{KK102117}VIE-260B, v107160); *sty* (*w*<sup>1118</sup>; *P*{GD1156}v6948/TM3, v6948); *Trf2* (*w*<sup>1118</sup>; *P*{KK107538}VIE-260B, v101318). The *P*{KK} RNAi lines are from the same genetic background and contain UAS-RNAi constructs for each candidate gene at the same locus on the second chromosome; the *P*{GD} RNAi lines are from the same genetic background, and contain *P*-element based transgenes in a random insertion site (Dietzl *et al.* 2007).

The control genotypes for the KK and GD RNAi lines are, respectively, *y, w*<sup>1118</sup>; *P*{attP, *y*<sup>+</sup>, *w*<sup>3</sup>} (v60100) and *w*<sup>1118</sup> (v60000), also obtained from the VDRC. We initially crossed all UAS-RNAi lines to an ubiquitously expressed *GAL4* driver, *Ubi-GAL4* (*P*{Ubi-GAL4}2). Lines without significant effects on grooming behavior using this driver were additionally crossed to two other ubiquitously expressed *GAL4* driver lines, *Act-GAL4* (*P*{Act5C-GAL4}25FO1) and *Ubi[156]-GAL4*. *Ubi-GAL4* and *Act-GAL4* were obtained from the Bloomington *Drosophila* Stock Center. *Ubi[156]-GAL4* was created in-house by introducing the original *Ubi-GAL4* transgene onto the third chromosome of the Canton-S B wild type strain by  $\Delta$ 2-3 transposase-mediated hopping (Garlapow *et al.* 2015). All fly stocks were maintained at 25°.

### Behavioral assay

Spontaneous grooming of single-sex groups of 20 flies per sex and genotype was measured for each line for 3 min. All flies were mated and 3-5 days old at the time of the assays. The genotypes included the 201 DGRP lines, and the F1 offspring of UAS-RNAi and *GAL4*-driver parents. The F1 offspring of the *P*{KK} or *P*{GD} control line and the *GAL4*-driver parents are the controls for RNAi knock down. Flies were lightly anesthetized by placing them on ice for 3-5 min. They were then transferred singly to each well of 96-well plates (Cell Wells, 25858, Corning, USA). Thirty minutes after the transfer, the flies were video-recorded for three minutes. The duration of spontaneous grooming was quantified during the three minutes, separately for foreleg and hindleg grooming. The fly uses its forelegs to groom the eyes, head, thorax, forelegs and midlegs, and it uses the hindlegs to groom the hindlegs, genitalia and wings. The midlegs serve to groom the forelegs and hindlegs, so midleg grooming was included in foreleg grooming and hindleg grooming, respectively.

### Statistical analyses

We performed quantitative genetic analysis of total grooming time across sexes using the factorial mixed ANOVA model  $Y = \mu + Sex + Line + Sex \times Line + e$ , where  $\mu$  is the overall mean, *Sex* is the fixed effect of sex, *Line* is the random effect of DGRP line, *Sex*  $\times$  *Line* is the interaction between sex and genotype, and *e* is the residual. Reduced models were also fitted for the two sexes separately. We estimated the broad-sense heritability,  $H^2$ , as  $H^2 = \frac{\sigma_L^2 + \sigma_{SL}^2}{\sigma_L^2 + \sigma_{SL}^2 + \sigma_e^2}$ , where  $\sigma_L^2$ ,  $\sigma_{SL}^2$  and  $\sigma_e^2$  are, respectively the among-line, sex by line and residual variance components. We estimated the cross-sex genetic correlation,  $r_{MF}$ , as  $r_{MF} = \frac{\sigma_L^2}{\sigma_L^2 + \sigma_{SL}^2}$ .

We performed the GWA analysis of total grooming time using line means for each sex as described in Huang *et al.* (2014) and implemented at <http://dgrp2.gnets.ncsu.edu>. Briefly, we adjusted the phenotypic data for the effects of *Wolbachia* infection and major polymorphic inversions, and used the adjusted line means to fit a linear mixed

model in the form of  $y = Xb + Zu + e$ , where  $y$  is the adjusted phenotypic values,  $X$  is the design matrix for the fixed SNP effect  $b$ ,  $Z$  is the incidence matrix for the random polygenic effect  $u$ , and  $e$  is the residual. The vector of polygenic effects  $u$  has a covariance matrix in the form of  $A\sigma^2$ , where  $\sigma^2$  is the polygenic variance component.

We annotated DNA variants using the gene models in Flybase release r5.57 (McQuilton *et al.* 2012). We downloaded the complete genetic interaction networks from FlyBase (release r5.57), where the genes are indicated by network nodes and the interactions between them as network edges. We mapped candidate genes from the GWA analysis ( $P < 10^{-5}$ ) to the network graph as described previously (Morozova *et al.* 2018). We extracted subnetworks from the global network whose edges were either a direct connection between candidate genes or bridged by only one gene not among the candidate gene list. We evaluated the significance of the size of the largest cluster among the subnetworks by a randomization test in which we randomly extracted subnetworks with the same number of input genes. The  $P$ -value was determined by dividing the number of instances where the size of the largest cluster exceeds the observed largest size by the total number of randomizations ( $\alpha=0.05$ ) (Antonov *et al.* 2010).

We used simple  $t$ -tests to assess differences in grooming behavior between the RNAi knock-down flies and control flies, separately for males and females. We also used Dunnett's  $t$ -tests to correct for multiple testing and control the family-wise error rate.

### Data availability

The DGRP lines and the *Ubi-GAL4* and *Act-GAL4* drivers are publicly available from the Bloomington Drosophila stock Center at Bloomington, Indiana. All RNAi lines are publicly available from the Vienna Drosophila RNAi Center in Vienna, Austria. The *Ubi[156]-GAL4* driver is available upon request. All raw data and codes used have been uploaded to the GitHub repository <https://github.com/qgg-lab/dgrp-grooming>. Supplementary Table 1 gives the ANOVA of spontaneous grooming behavior and estimates of quantitative genetic parameters. Supplementary Table 2 gives the results of the GWA analyses of spontaneous grooming behavior, including SNP ID numbers and locations. Supplementary Table 3 provides the names and gene ontology annotations of significant genes from the GWA analyses. Supplementary Table 4 provides the results of the RNAi functional assessment analyses. File S1 is a short video of males and females from a high grooming line, DGRP\_818. File S2 is a short video of males and females from a low grooming line, DGRP\_517. Full DGRP sequence data are available at the NCBI Sequence Read Archive (SRA; <http://www.ncbi.nlm.nih.gov/sra>) under accession numbers listed in Supplemental Data File S1 of Huang *et al.* 2014. All SNP genotypes, quality metrics and annotations are available at <http://dgrp2.gnets.ncsu.edu>. Supplemental material available at figshare: <https://doi.org/10.25387/g3.12252719>

## RESULTS

### Quantitative genetics of spontaneous grooming behavior in the DGRP

To understand the genetic basis of natural variation in spontaneous grooming behavior, we manually annotated video recordings of 20 females and 20 males from each of 201 DGRP lines, and quantified the total time spent grooming during a 3-minute period. There was substantial phenotypic variation both at the individual fly level (Figure 1a, b) and at the line level (Figure 1c) within each sex. Across all flies, spontaneous grooming time ranged from no grooming at all (0 s) to nearly constant grooming (176 s in a maximum of 180 s) (Files

S1, S2). Across lines, the mean grooming time ranged from 6.7 s (DGRP\_774) to 105 s (DGRP\_805) in females, and from 5.4 s (DGRP\_59) to 113 s (DGRP\_805) in males. DGRP\_805 was the most frequent groomer in both females and males (frequent grooming events result in longer cumulative grooming time). The heritability of total grooming time was 0.25 in females and 0.24 in males, with a cross-sex genetic correlation of 0.77 (Figure 1d), and significant genotype by sex interaction for this trait (Table S1). Thus, the genetic architecture of grooming behavior is partially in common and partially distinct between the sexes.

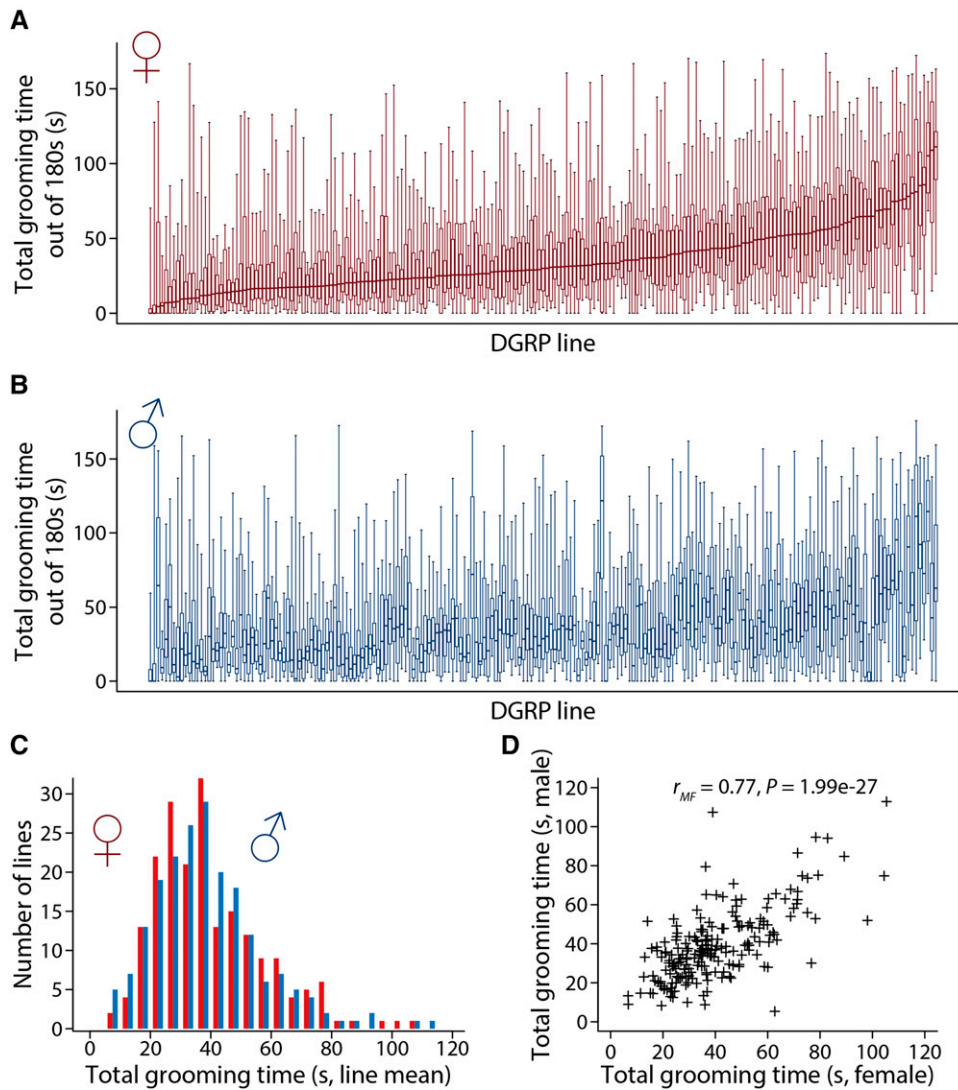
### GWA analysis for spontaneous grooming behavior

Because there was significant genetic variation for grooming behavior (Figure 1, Table S1), we performed a GWA analysis (Huang *et al.* 2014) to identify genes that harbor common (minor allele frequency, MAF > 0.05) DNA variants associated with variation in spontaneous grooming time (Table S2). We used line means for this analysis, which has the advantage of increasing the broad-sense heritability to  $H^2 = 0.87$ , by reducing the residual variance by  $\frac{\sigma^2}{n}$ , where  $n$  is the number of individuals measured per sex and line. Only two variants were significant using a strict Bonferroni correction for nearly 2 million variants: *X\_8312234\_INS* in an intron of *Trf2* ( $P = 5.61 \times 10^{-9}$  in males and  $P = 6.52 \times 10^{-9}$  for both sexes) and the intergenic polymorphism *3L\_3396966\_SNP* ( $P = 2.18 \times 10^{-8}$ ). At the more lenient significance threshold of  $P < 10^{-5}$ , we identified 107 DNA variants associated with spontaneous grooming in either males or females, and/or the average or difference of the two sexes (Table S2). Of these variants, 34 were intergenic (> 1 kb from gene boundaries) while the remaining fell within 1 kb of 70 annotated genes (Table S2).

The genes associated with spontaneous grooming behavior fell into a wide variety of gene ontology (GO) categories, including 20 genes with GO terms associated with nervous system development and function; 11 genes with GO terms associated with organism or tissue development; seven genes associated with GO terms associated with transcriptional regulation; and seven genes with GO terms associated with epidermal growth factor, fibroblast growth factor, Notch, BMP or other signaling pathways (Table S3). We performed a GO enrichment analysis, accounting for difference in gene length as this affects the differential baseline probability of being identified in the GWA analysis. None of the tested GO terms were enriched for the grooming-associated genes, consistent with a highly polygenic genetic architecture. Next, we mapped the significant genes detected in the GWA analysis to known genetic interaction networks. We extracted subnetworks from the global network whose edges were either a direct connection between candidate genes or bridged by only one gene not among the candidate gene list, and evaluated the significance of the size of the largest cluster among the subnetworks by a randomization test in which we randomly extracted subnetworks with the same number of input genes. The subnetwork with no missing genes resulted in only two connected genes (*Ptp10D* and *Ptp99A*,  $P = 0.08$ ). However, allowing genes to be connected by at most one gene not in the GWA gene list captured a cluster of 24 genes that was significantly larger than a randomly seeded network, centered on three highly connected genes: *pnt*, which encodes a transcriptional regulator; *scrib*, which encodes a synaptic scaffold protein; and the morphogenetic gene *sty* ( $P = 0.01$ , Figure 2).

### Functional assessment of candidate genes

We assessed whether ubiquitous reduction in gene expression of a sample of candidate genes implicated by the GWA analysis affected spontaneous grooming behavior, using the bipartite *GALA/UAS-RNAi*



**Figure 1** Variation in spontaneous grooming duration (s) in 201 DGRP lines. Flies were observed for 180 s. (a) Total grooming time for females, arranged from lowest to highest grooming time. The box plot for each line indicates the distributional property of the replicates. (b) Total grooming time for males. The lines are in the same order as for females. (c) Distribution of line means for grooming time in the DGRP. Lines are grouped in 5 s bins. The red histograms indicate female line means and the blue histograms indicate male line means. (d) Relationship between male and female grooming time for each line.

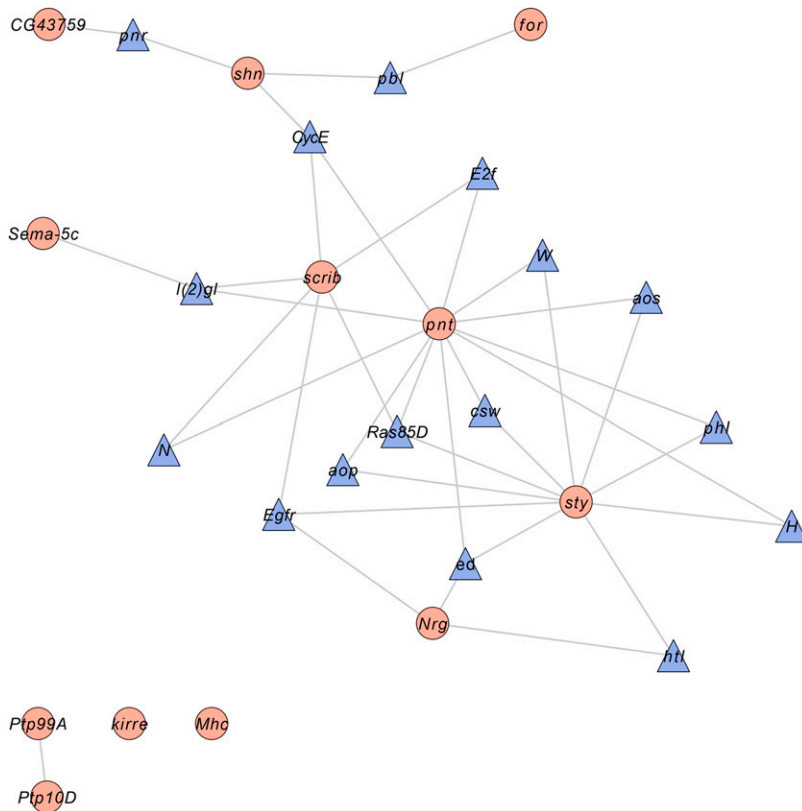
system (Brand and Perrimon 1993; Dietzl *et al.* 2007). We chose 29 candidate genes that were most significant in males, females, or the difference between the sexes, and for which RNAi lines were available. We first crossed all 29 lines to an ubiquitous driver, *Ubi-GAL4*, and assessed spontaneous grooming behavior for RNAi lines producing viable adults. For lines that did not produce viable adults or that did not have a significant phenotypic effect when crossed to *Ubi-GAL4*, we crossed them to a weak ubiquitous driver (*Ubi[156]-GAL4*) and/or a strong ubiquitous driver (*Act-GAL4*). A total of 20 candidate genes had significant effects on spontaneous grooming behavior in at least one sex and driver combination at a nominal  $P$ -value of  $P < 0.05$ , and 11 genes had significant effects in the same direction in both sexes and/or for more than one driver (*CG13229*, *CG17839*, *CG34401*, *Gdh*, *GEFmeso*, *mtt*, *osp*, *PGRP-LA*, *PsGEF*, *sty*, *Trf2*) (Table S4). A total of 11 genes remained significant ( $P < 0.05$ ) when we used Dunnett's test to account for multiple tests when several treatments are compared to a control group (*CG13229*, *CG34401*, *Gdh*, *GEFmeso*, *kirre*, *Lrch*, *mtt*, *osp*, *PGRP-LA*, *sty*, *Trf2*) (Figure 3).

## DISCUSSION

We have demonstrated that there is significant natural genetic variation in spontaneous grooming behavior in *D. melanogaster*,

and that the genetic architecture of grooming behavior is partially distinct between males and females, as has been observed for most quantitative traits in the DGRP (Mackay and Huang 2018). We performed a GWA analysis and identified 107 molecular polymorphisms in 70 genes associated with spontaneous grooming behavior at a nominal  $P < 10^{-5}$ . Only two of these polymorphisms resulted in nonsynonymous changes in protein coding genes; the rest presumably regulate gene expression, including the 34 polymorphisms located over 1,000 bp from the nearest protein coding gene.

None of the candidate genes identified in our GWA analysis has been associated previously with spontaneous grooming behavior, highlighting the complementary information to be gained from mutational analyses and naturally occurring variation due to common alleles. However, many of the candidate genes identified in this study are plausible given the trait assessed, and some are components of pathways implicated previously in grooming behavior. For example, glutamate is the primary excitatory neurotransmitter at the neuromuscular junction in insects (O'Kane 2011). *Gdh*, which encodes Glutamate dehydrogenase (*Gdh*); and *Eaat*, which encodes Excitatory amino acid transporter 1 (*Eaat1*), were candidate genes in the GWA analysis, and RNAi knockdown of *Gdh* affected grooming behavior. *Gdh* is involved in glutamate metabolism, and is the



**Figure 2** Genetic interaction network for spontaneous grooming behavior. The network consists of candidate genes identified by the GWA analysis at  $P < 10^{-5}$  (orange circles) and computationally recruited genes not identified in our study but which are known to interact with the candidate genes (purple triangles).

enzyme that catalyzes the reversible conversion of L-glutamate to  $\alpha$ -ketoglutarate (Yelamanchi *et al.* 2016), while *Eaat1* is a glutamate transporter (McQuilton *et al.* 2012).

*Nf1* is known to affect grooming behavior; *Nf1* genetically interacts with *CCKLR-17D1*, one of the candidate genes from the GWA analysis (Walker *et al.* 2013), as well as with *Ras85D* (Williams *et al.* 2001), which was computationally recruited to the significant genetic network derived from the GWA analysis candidate genes. The highly connected hub genes *scrib*, *pnt* and *sty* in the genetic interaction network also interact with *Ras85D*, suggesting Ras signaling may be involved in grooming behavior.

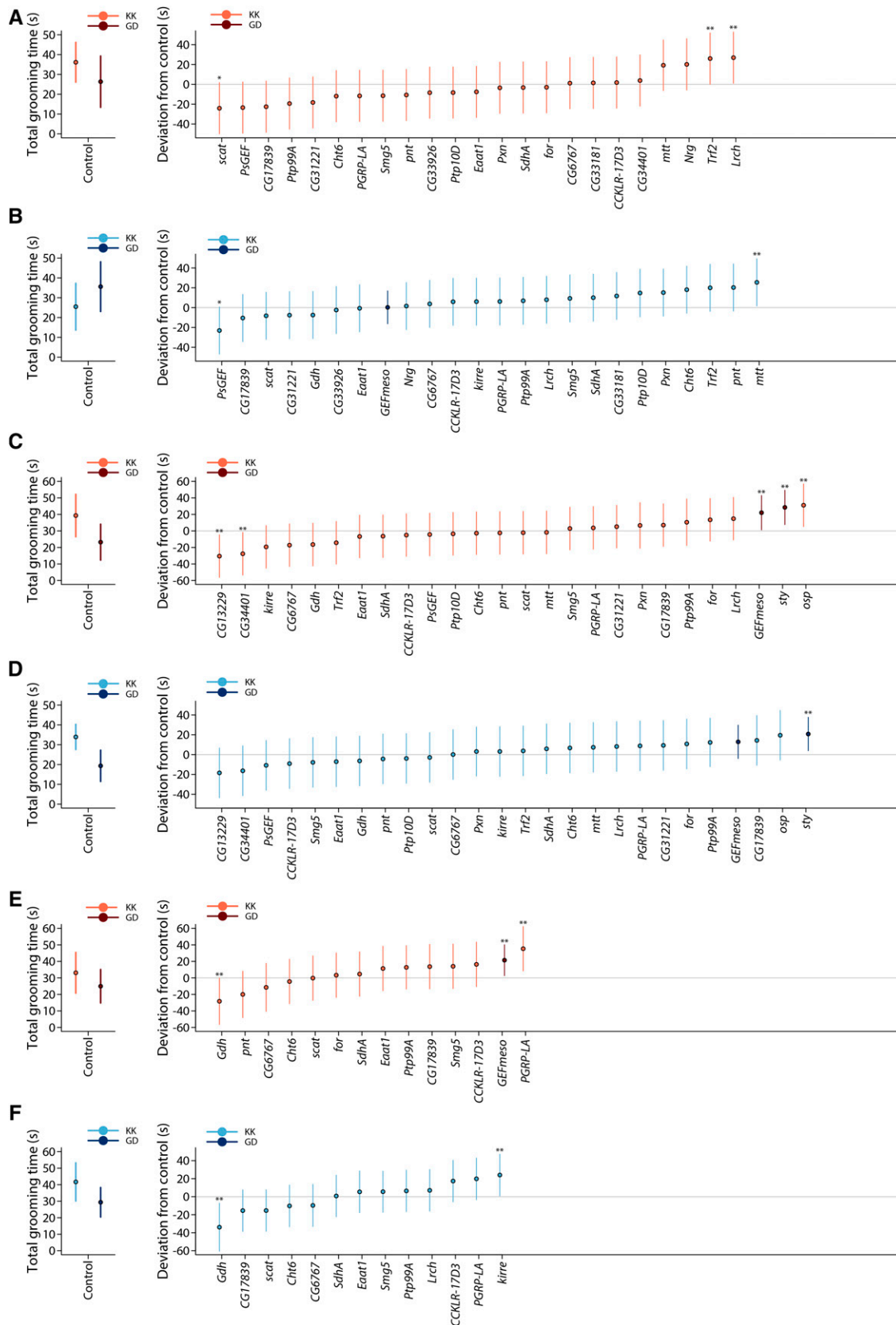
Grooming can be stimulated by mechanical (Page and Matheson 2004; Hlavac 1975; Gratwick 1957), olfactory (Becher *et al.* 2012; Stensmyr *et al.* 2012; Dweck *et al.* 2015; Kapsetaki *et al.* 2014), gustatory (Yanagawa *et al.* 2014; 2018; Soldano *et al.* 2016), and immunity-related (Yanagawa *et al.* 2014, 2017; Soldano *et al.* 2016) stimuli. We note that *scrib* (Ganguly *et al.* 2003) and *Sema-5c* (Sambandan *et al.* 2006; Rollmann *et al.* 2007) affect olfactory behavior. *PGRP-LC* has been associated with grooming behavior (Yanagawa *et al.* 2017); here, another candidate gene from this gene family, *PGRP-LA*, was a candidate gene from the GWA analysis, and RNAi knockdown of *PGRP-LA* affected grooming behavior. Peptidoglycan-recognition proteins regulate the immune deficiency pathway which controls the defense against Gram-negative bacteria (Gendrin *et al.* 2013).

The candidate genes associated with grooming behavior span a wide variety of gene ontology categories, suggesting novel genetic contributions to this complex behavior. Many of the genes are involved in nervous system development and/or function in organism or tissue development (including leg and muscle development, required for insect grooming) (Table S3). However, transcriptional

regulation and post-transcriptional modification are also implicated. Finally, the Notch and Epidermal growth factor signaling pathways are implicated directly by the inclusion of *N* and *Egfr* as computationally recruited genes in the enriched genetic interaction network of candidate genes associated with spontaneous grooming behavior; and by the annotated biological function gene ontology categories of candidate genes in regulating Notch (*Kul*, *scrib*) and Epidermal growth factor (*kek5*, *pnt*, *Ptp10D*, *RhoGAP5A*, *scrib*, *sty*) pathway signaling (Table S3).

Functional assessment of 29 candidate genes using RNAi showed that 11 (*CG13229*, *CG34401*, *Gdh*, *GEFmeso*, *kirre*, *Lrch*, *mtt*, *osp*, *PGRP-LA*, *sty*, *Trf2*) affect spontaneous grooming behavior. However, the outcome of RNAi depends on many factors. RNAi can have off-target effects on assessed phenotypes (although few are predicted for the constructs utilized in this study), and ubiquitous RNAi affects all development times and tissues/cell types in which the target gene is expressed. It is possible that utilizing neuronal, glial or more specific *GAL4* drivers, or *GAL4* drivers with expression restricted to particular developmental stages, may have implicated more candidate genes, and future studies are needed to assess in which tissues/cell types and stages of development knockdown of expression affects the phenotype. Furthermore, RNAi cannot capture the potentially subtle effects of the naturally occurring variants; future variant-specific tests are needed for functional confirmation.

Grooming is an important fitness trait in insects (Yanagawa *et al.* 2008; 2017; 2019; Böröczky *et al.* 2013; Zhukovskaya *et al.* 2013; Yanagawa *et al.* 2014). Grooming is also a repetitive behavior. Aberrant repetitive behaviors are common for many human neurodevelopmental disorders (autism, fragile X syndrome, Prader-Willi syndrome, nonsyndromic intellectual disability) and clinical disorders (obsessive-compulsive disorder, Rett and Tourette



**Figure 3** Results of RNAi functional assessments of candidate genes. The panels depict mean spontaneous grooming times (circles) for each of the tested *UAS-RNAi/Driver-GAL4* genotypes, expressed as a deviation from the appropriate Control/*Driver-GAL4* genotypes. The grooming times of the control genotypes are given to the left of each panel. The whiskers indicate the 95% Dunnett's confidence intervals. (a) *UAS-RNAi/Ubi156]-GAL4* females. (b) *UAS-RNAi/Ubi156]-GAL4* males. (c) *UAS-RNAi/Ubi-GAL4* females. (d) *UAS-RNAi/Ubi-GAL4* males. (e) *UAS-RNAi/Act-GAL4* females. (f) *UAS-RNAi/Act-GAL4* males. \* indicates  $P < 0.1$  and \*\* indicates  $P < 0.05$  from Dunnett's tests.

syndrome, Parkinson's disease, fronto-temporal and Alzheimer's dementia) (Lewis and Kim 2009; Whitehouse and Lewis 2015). Further analysis of the genetic underpinnings of naturally occurring variation in *Drosophila* spontaneous grooming may yield insight into these human disorders.

## ACKNOWLEDGMENTS

We thank Badar Beth Werek (Kyoto University, Japan) who conducted the behavioral analysis of flies. This work was supported by OECD CRP fellowship 2015: Biological Resource Management for Sustainable Agricultural system, the Research Institute for Sustainable Humanosphere, Kyoto University (grant no. 2016-5-1-1) and the Future Development Funding Program of Kyoto University Research Coordination Alliance.

## LITERATURE CITED

- Antonov, A. V., E. E. Schmidt, S. Dietmann, M. Krestyaninova, and H. Hermjakob, 2010 R spider: a network-based analysis of gene lists by combining signaling and metabolic pathways from Reactome and KEGG databases. *Nucleic Acids Res.* 38: W78–W83. <https://doi.org/10.1093/nar/gkq482>
- Bahreini, R., and R. W. Currie, 2015 The effect of queen pheromone status on Varroa mite removal from honey bee colonies with different grooming ability. *Exp. Appl. Acarol.* 66: 383–397. <https://doi.org/10.1007/s10493-015-9907-2>
- Becher, P. G., G. Flick, E. Rozpędowska, A. Schmidt, A. Hagman *et al.*, 2012 Yeast, not fruit volatiles mediate *Drosophila melanogaster* attraction, oviposition and development. *Funct. Ecol.* 26: 822–828. <https://doi.org/10.1111/j.1365-2435.2012.02006.x>
- Böröczky, K., A. Wada-Katsumata, D. Batchelor, M. Zhukovskaya, and C. Schal, 2013 Insects groom their antennae to enhance olfactory acuity. *Proc. Natl. Acad. Sci. USA* 110: 3615–3620. <https://doi.org/10.1073/pnas.1212466110>
- Brand, A. H., and N. Perrimon, 1993 Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* 118: 401–415.
- Dethier, V. G., 1972 Sensitivity of the contact chemoreceptors of the blowfly to vapors. *Proc. Natl. Acad. Sci. USA* 69: 2189–2192. <https://doi.org/10.1073/pnas.69.8.2189>
- Dweck, H. K., S. A. Ebrahim, A. Farhan, B. S. Hansson, and M. C. Stensmyr, 2015 Olfactory proxy detection of dietary antioxidants in *Drosophila*. *Curr. Biol.* 25: 455–466. <https://doi.org/10.1016/j.cub.2014.11.062>
- Dietzl, G. D., F. Chen, F. Schnorrer, K.-C. Su, Y. Barinova *et al.*, 2007 A genome-wide transgenic RNAi library for conditional gene inactivation in *Drosophila*. *Nature* 448: 151–156. <https://doi.org/10.1038/nature05954>
- Ganguly, I., T. F. C. Mackay, and R. R. H. Anholt, 2003 Scribble is essential for olfactory behavior in *Drosophila melanogaster*. *Genetics* 164: 1447–1457.
- Garlapow, M. E., W. Huang, M. T. Yarboro, K. R. Peterson, and T. F. C. Mackay, 2015 Quantitative genetics of food intake in *Drosophila melanogaster*. *PLoS One* 10: e0138129. <https://doi.org/10.1371/journal.pone.0138129>
- Gendrin, M., A. Zaidman-Rémy, N. A. Broderick, J. Paredes, M. Poidevin *et al.*, 2013 Functional analysis of PGRP-LA in *Drosophila* immunity. *PLoS One* 8: e69742. <https://doi.org/10.1371/journal.pone.0069742>
- Gratwick, M., 1957 The contamination of insects of different species exposed to dust deposits. *Bull. Entomol. Res.* 48: 741–753. <https://doi.org/10.1017/S0007485300002868>
- Grillner, S., and P. Wallén, 2004 Innate vs. learned movements—a false dichotomy? *Prog. Brain Res.* 143: 3–12.
- Grillner, S., J. Hellgren, A. Ménard, K. Saitoh, and M. A. Wikström, 2005 Mechanisms for selection of basic motor programs—roles for the striatum and pallidum. *Trends Neurosci.* 28: 364–370. <https://doi.org/10.1016/j.tins.2005.05.004>
- Hampel, S., C. E. McKellar, J. H. Simpson, and A. M. Seeds, 2017 Simultaneous activation of parallel sensory pathways promotes a grooming sequence in *Drosophila*. *eLife* 6: e28804. <https://doi.org/10.7554/eLife.28804>
- Hampel, S., R. Franconville, J. H. Simpson, and A. M. Seeds, 2015 A neural command circuit for grooming movement control. *eLife* 4: e08758. <https://doi.org/10.7554/eLife.08758>
- Hlavac, T. F., 1975 Grooming systems of insects: Structure, mechanics. *Ann. Entomol. Soc. Am.* 68: 823–826. <https://doi.org/10.1093/aesa/68.5.823>
- Huang, W., A. Massouras, Y. Inoue, J. Peiffer, M. Ràmia *et al.*, 2014 Natural variation in genome architecture among 205 *Drosophila melanogaster* Genetic Reference Panel lines. *Genome Res.* 24: 1193–1208. <https://doi.org/10.1101/gr.171546.113>
- Kapsetaki, S.-E., I. Tzelepis, K. Avgousti, I. Livadaras, N. Garantonakis *et al.*, 2014 The bacterial metabolite 2-aminoacetophenone promotes association of pathogenic bacteria with flies. *Nat. Commun.* 5: 4401. <https://doi.org/10.1038/ncomms5401>
- King, L. B., M. Koch, K. R. Murphy, Y. Velazquez, W. W. Ja *et al.*, 2016 Neurofibromin loss of function drives excessive grooming in *Drosophila*. *G3 (Bethesda)* 6: 1083–1093. <https://doi.org/10.1534/g3.115.026484>
- Lewis, M., and S.-J. Kim, 2009 The pathophysiology of restricted repetitive behavior. *J. Neurodev. Disord.* 1: 114–132. <https://doi.org/10.1007/s11689-009-9019-6>
- Mackay, T. F. C., S. Richards, E. A. Stone, A. Barbadilla, J. F. Ayroles *et al.*, 2012 The *Drosophila melanogaster* Genetic Reference Panel. *Nature* 482: 173–178. <https://doi.org/10.1038/nature10811>
- Mackay, T. F. C., and W. Huang, 2018 Charting the genotype-phenotype map: lessons from the *Drosophila melanogaster* Genetic Reference Panel. *Wiley Interdiscip. Rev. Dev. Biol.* 7: e289. <https://doi.org/10.1002/wdev.289>
- McQuilton P., S. E. St Pierre, J. Thurmond and FlyBase Consortium, 2012 FlyBase 101—the basics of navigating FlyBase. *Nucleic Acids Res.* 40: D706–D714. <https://doi.org/10.1093/nar/gkr1030>
- Melzig, J., S. Buchner, F. Wiebel, R. Wolf, M. Burg *et al.*, 1996 Genetic depletion of histamine from the nervous system of *Drosophila* eliminates specific visual and mechanosensory behavior. *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.* 179: 763–773. <https://doi.org/10.1007/BF00207355>
- Morozova, T. V., Y. Hussain, L. J. McCoy, E. V. Zhirnov, M. R. Davis *et al.*, 2018 A Cyclin E centered genetic network contributes to alcohol-induced variation in *Drosophila* development. *G3 (Bethesda)* 8: 2643–2653. <https://doi.org/10.1534/g3.118.200260>
- Naumann, K., 1991 Grooming behaviors and the translocation of queen mandibular gland pheromone on worker honey bees (*Apis mellifera* L.). *Apidologie (Celle)* 22: 523–531. <https://doi.org/10.1051/apido:19910505>
- Newland, P. L., 1998 Avoidance reflexes mediated by contact chemoreceptors on the legs of locusts. *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.* 183: 313–324. <https://doi.org/10.1007/s003590050258>
- O’Kane, C. J., 2011 *Drosophila* as a model organism for the study of neuropsychiatric disorders. *Curr. Top. Behav. Neurosci.* 7: 37–60. [https://doi.org/10.1007/7854\\_2010\\_110](https://doi.org/10.1007/7854_2010_110)
- Page, K. L., and T. Matheson, 2004 Wing hair sensilla underlying aimed hindleg scratching of the locust. *J. Exp. Biol.* 207: 2691–2703. <https://doi.org/10.1242/jeb.01096>
- Pitmon, E. G., G. Stephens, S. J. Parkhurst, F. W. Wolf, G. Kehne *et al.*, 2016 The D1 family dopamine receptor, DopR, potentiates hind leg grooming behavior in *Drosophila*. *Genes Brain Behav.* 15: 327–334. <https://doi.org/10.1111/gbb.12264>
- Rollmann, S. M., A. Yamamoto, T. Goossens, L. Zwarts, Z. Callaerts-Vegh *et al.*, 2007 The early developmental gene *Semaphorin 5c* contributes to olfactory behavior in adult *Drosophila*. *Genetics* 176: 947–956. <https://doi.org/10.1534/genetics.106.069781>
- Sachs, B. D., 1988 The development of grooming and its expression in adult animals. *Ann. N. Y. Acad. Sci.* 525: 1–17. <https://doi.org/10.1111/j.1749-6632.1988.tb38591.x>
- Sambandan, D., A. Yamamoto, J. J. Fanara, T. F. C. Mackay, and R. R. H. Anholt, 2006 Dynamic genetic interactions determine odor-guided behavior in *Drosophila melanogaster*. *Genetics* 174: 1349–1363. <https://doi.org/10.1534/genetics.106.060574>

- Seeds, A. M., P. Ravbar, P. Chung, S. Hampel, F. M. Midgley *et al.*, 2014 A suppression hierarchy among competing motor programs drives sequential grooming in *Drosophila*. *eLife* 3: e02951. <https://doi.org/10.7554/eLife.02951>
- Soldano, A., Y. A. Alpizar, B. Boonen, L. Franco, A. López-Requena *et al.*, 2016 Gustatory-mediated avoidance of bacterial lipopolysaccharides via TRPA1 activation in *Drosophila*. *eLife* 5: e13133. <https://doi.org/10.7554/eLife.13133>
- Stensmyr, M. C., H. K. M. Dweck, A. Farhan, I. Ibba, A. Strutz *et al.*, 2012 A conserved dedicated olfactory circuit for detecting harmful microbes in *Drosophila*. *Cell* 151: 1345–1357. <https://doi.org/10.1016/j.cell.2012.09.046>
- Tauber, J. M., P. A. Vanlandingham, and B. Zhang, 2011 Elevated levels of the vesicular monoamine transporter and a novel repetitive behavior in the *Drosophila* model of fragile X syndrome. *PLoS One* 6: e27100. <https://doi.org/10.1371/journal.pone.0027100>
- Walker, J. A., J. Y. Gouzi, J. B. Long, S. Huang, R. C. Maher *et al.*, 2013 Genetic and functional studies implicate synaptic overgrowth and ring gland cAMP/PKA signaling defects in the *Drosophila melanogaster* *Neurofibromatosis-1* growth deficiency. *PLoS Genet.* 9: e1003958. <https://doi.org/10.1371/journal.pgen.1003958>
- Werner, T., K. Borge-Renberg, P. Mellroth, H. Steiner, and D. Hultmark, 2003 Functional diversity of the *Drosophila* PGRP-LC gene cluster in the response to lipopolysaccharide and peptidoglycan. *J. Biol. Chem.* 278: 26319–26322. <https://doi.org/10.1074/jbc.C300184200>
- Whitehouse, C. M., and M. H. Lewis, 2015 Repetitive behavior in neurodevelopmental disorders: clinical and translational findings. *Behav. Anal.* 38: 163–178. <https://doi.org/10.1007/s40614-015-0029-2>
- Williams, J. A., H. S. Su, A. Bernards, J. Field, and A. Sehgal, 2001 A circadian output in *Drosophila* mediated by Neurofibromatosis-1 and Ras/MAPK. *Science* 293: 2251–2256. <https://doi.org/10.1126/science.1063097>
- Yanagawa, A., F. Yokohari, and S. Shimizu, 2008 Defense mechanism of the termite, *Coptotermes formosanus* Shiraki, to entomopathogenic fungi. *J. Invertebr. Pathol.* 97: 165–170. <https://doi.org/10.1016/j.jip.2007.09.005>
- Yanagawa, A., A. Guigue, and F. Marion-Poll, 2014 Hygienic grooming is induced by contact chemicals in *Drosophila melanogaster*. *Front. Behav. Neurosci.* 8: 254. <https://doi.org/10.3389/fnbeh.2014.00254>
- Yanagawa, A., M. Chabaud, T. Imai, and F. Marion-Poll, 2018 Olfactory cues play a significant role in removing fungus from the body surface of *Drosophila melanogaster*. *J. Invertebr. Pathol.* 151: 144–150. <https://doi.org/10.1016/j.jip.2017.11.011>
- Yanagawa, A., C. Neyen, B. Lemaitre, and F. Marion-Poll, 2017 The gram-negative sensing receptor PGRP-LC contributes to grooming induction in *Drosophila*. *PLoS One* 12: e0185370. <https://doi.org/10.1371/journal.pone.0185370>
- Yanagawa, A., A. Couto, J. C. Sandoz, T. Hata, A. Mitra *et al.*, 2019 LPS perception through taste-induced reflex in *Drosophila melanogaster*. *J. Insect Physiol.* 112: 39–47. <https://doi.org/10.1016/j.jinsphys.2018.12.001>
- Yelamanchi, S. D., S. Jayaram, J. K. Thomas, S. Aafaque, A. Khan *et al.*, 2016 A pathway map of glutamate metabolism. *J. Cell Commun. Signal.* 10: 69–75. <https://doi.org/10.1007/s12079-015-0315-5>
- Zhukovskaya, M., A. Yanagawa, and B. T. Forschler, 2013 Grooming behavior as a mechanism of insect disease defense. *Insects* 4: 609–630. <https://doi.org/10.3390/insects4040609>

Communicating editor: D.-J. de Koning