



The *Drosophila* Gene *Sulfateless* Modulates Autism-Like Behaviors

Kevin A. Hope^{1,2}, Daniel Flatten³, Peter Cavitch¹, Ben May⁴, James S. Sutcliffe^{5,6}, Janis O'Donnell⁷ and Lawrence T. Reiter^{2,8,9*}

¹ Integrated Program in Biological Sciences, The University of Tennessee Health Science Center, Memphis, TN, United States, ² Department of Neurology, The University of Tennessee Health Science Center, Memphis, TN, United States, ³ Christian Brothers University, Memphis, TN, United States, ⁴ Rhodes College, Memphis, TN, United States, ⁵ Department of Molecular Physiology and Biophysics, Vanderbilt University School of Medicine, Nashville, TN, United States, ⁶ Department of Psychiatry, Vanderbilt University School of Medicine, Nashville, TN, United States, ⁷ Department of Biological Sciences, University of Alabama, Tuscaloosa, AL, United States, ⁸ Department of Anatomy and Neurobiology, The University of Tennessee Health Science Center, Memphis, TN, United States, ⁹ Department of Pediatrics, The University of Tennessee Health Science Center, Memphis, TN, United States

OPEN ACCESS

Edited by:

Erica E. Davis,
Duke University, United States

Reviewed by:

Santhosh Girirajan,
Pennsylvania State University,
United States
Annette Schenck,
Radboud University Nijmegen Medical
Centre, Netherlands

*Correspondence:

Lawrence T. Reiter
lreiter@uthsc.edu

Specialty section:

This article was submitted to
Genetic Disorders,
a section of the journal
Frontiers in Genetics

Received: 18 January 2019

Accepted: 31 May 2019

Published: 19 June 2019

Citation:

Hope KA, Flatten D, Cavitch P,
May B, Sutcliffe JS, O'Donnell J and
Reiter LT (2019) The *Drosophila* Gene
Sulfateless Modulates Autism-Like
Behaviors. *Front. Genet.* 10:574.
doi: 10.3389/fgene.2019.00574

Major challenges to identifying genes that contribute to autism spectrum disorder (ASD) risk include the availability of large ASD cohorts, the contribution of many genes overall, and small effect sizes attributable to common gene variants. An alternative approach is to use a model organism to detect alleles that impact ASD-relevant behaviors and ask whether homologous human genes infer ASD risk. Here we utilized the *Drosophila* genetic reference panel (DGRP) as a tool to probe for perturbation in naturally occurring behaviors in *Drosophila melanogaster* that are analogous to three behavior domains: impaired social communication, social reciprocity and repetitive behaviors or restricted interests. Using 40 of the available DGRP lines, we identified single nucleotide polymorphisms (SNPs) in or near genes controlling these behavior domains, including ASD gene orthologs (*neurexin 4* and *neuroligin 2*), an intellectual disability (ID) gene homolog (*kirre*), and a gene encoding a heparan sulfate (HS) modifying enzyme called *sulfateless* (*sfl*). SNPs in *sfl* were associated with all three ASD-like behaviors. Using RNAi knock-down of neuronal *sfl* expression, we observed significant changes in expressive and receptive communication during mating, decreased grooming behavior, and increased social spacing. These results suggest a role for HS proteoglycan synthesis and/or modification in normal social communication, repetitive behavior, and social interaction in flies. Finally, using the DGRP to directly identify genetic effects relevant to a neuropsychiatric disorder further demonstrates the utility of the *Drosophila* system in the discovery of genes relevant to human disease.

Keywords: DGRP, autism, *Drosophila*, heparin sulfate, GWAS

INTRODUCTION

Major hurdles to understanding the pathophysiology of autism spectrum disorder (ASD) include extreme genetic heterogeneity combined with small effect sizes for most risk genes, and the inherent difficulty in identifying non-genetic factors (e.g., environmental and epigenetic) that may increase autism risk. Approximately 20–30% of cases of non-syndromic ASD correspond to individually

rare, highly penetrant, often *de novo*, variants. These variants range from “pathogenic” to significant risk variants of large effect and can show incomplete penetrance, pleiotropy or both (De Rubeis et al., 2014; Sanders et al., 2015). *Rare* genic variants include both copy number variants (CNVs), a loss or gain of DNA ≥ 1000 bp and more discrete, single nucleotide variants (SNVs), or small indels detected from whole exome or whole genome sequencing. Genome wide association studies (GWAS) use genotype arrays to capture common genetic variation across the genome and test for association with disease using family-based or case-control designs. This approach has been remarkably successful in schizophrenia for which combined samples in the psychiatric genomics consortium (PGC) have surpassed 50,000 and have identified well over 150 genome-wide significant loci (Li et al., 2017; Pardiñas et al., 2018). By comparison, the most recent ASD GWAS from the PGC identified only a few genome-wide significant loci (Autism Spectrum Disorders Working Group of the Psychiatric Genomics Consortium, 2017), but amalgamated ASD cohorts are not yet of the size to reach the critical “inflection point” previously observed for schizophrenia (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014).

In contrast, a combination of whole exome and genome sequencing and CNV discovery in family and case-control samples has identified a number of genes disrupted by rare, damaging, *de novo* variants, and highlighted biology and function that resonates across psychiatric disorders, like RNAs regulated by Fragile X mental retardation 1 (FMRP) (Iossifov et al., 2012), synaptic scaffolding proteins, chromatin regulating machinery, and other genes related to overlapping neuropsychiatric disorders. A major hindrance to the success of GWAS in identifying associations to ASD and other complex behavioral disorders is the small effect sizes conferred by common variants (Devlin et al., 2011). Mouse models of single gene or contiguous gene models of autism have been useful in some cases, for example, Rett syndrome (*Mecp2*) (Liao et al., 2012) and duplications of human 15q11.2-q13 (Nakatani et al., 2009; Smith et al., 2011). In general, however, these models have a limited utility for understanding the multigenic nature of ASD, since they model only single, highly penetrant mutations. Here we propose a new approach to analyze common variants contributing to autism-like behaviors at multiple loci that conspire to confer ASD risk in humans by using the powerful genetic tools available in the model organism *Drosophila melanogaster*.

In this study, we identified a set of behaviors in flies that may be analogous to the human behaviors associated with ASD. Specifically, we assay (1) social communication during mating (or courtship behavior); (2) repetitive behavior during grooming; and (3) social reciprocity, in a social spacing test. We used these analogous behaviors to query the *Drosophila* genetic reference panel (DGRP), a community resource consisting of 205 sequenced inbred lines that have been derived from a natural population (Mackay et al., 2012), in order to identify shared common alleles across behaviors. The DGRP lines have already been used successfully to identify new candidate genes that influence a number of *Drosophila* behaviors and adaptations

to environmental stressors (Mackay and Huang, 2018). Here we assessed both individual behavior domains and combined behaviors in order to identify genes that could influence multiple ASD-like fly behaviors. Ultimately, we seek to use this approach to tease apart the complex genetics that lead to autism phenotypes in humans.

RESULTS

Selection of Behavior Assays Analogous to ASD Behaviors

We sought to identify measurable fly behaviors that map to the three core ASD phenotypic domains: communication, social interaction and repetitive behaviors, and restricted interests. The current Diagnostic and Statistical Manual for Mental Disorders, 5th edition (DSM-V; American Psychiatric Association, 2013) criteria combines social interaction with communication, creating a category of *social communication*. Although many naturally occurring behaviors can be measured in the fly model, we felt that there were only a few behaviors which resembled human behaviors closely enough to use in this study. The most obvious behavior is a recently published assay for fly social spacing (analogous to human social reciprocity). Humans with ASD often interact with other individuals awkwardly. This includes inappropriate responses to social cues which can cause them to either violate another persons’ “personal space” or to over-react when an individual touches them or invades their personal space. When flies are housed as a group they also settle into a “comfortable” social spacing that can be quantified using the social space triangle (Simon et al., 2012).

Another clearly analogous behavior to human ASD-like behaviors is any measure of repetitive behavior. Repetitive grooming has been used in flies previously as a measure of ASD symptomology in Fragile X mutants (Drozd et al., 2018). Here we use the flies normal grooming behavior as a read out for restrictive and repetitive interests, one of the most common ASD symptoms in humans.

Perhaps the hardest ASD feature to approximate in any animal model are defects in communication, a uniquely human trait that does not translate well to animals that lack a specific language of their own. However, researchers have made attempts to decipher the simple languages of many animals in the wild. Perhaps most well-known animal vocalizations are the various bird calls that can indicate alarm or mating calls or other specific behavioral cues, which may be learned and processed in ways similar to human language [reviewed in Mol et al. (2017)]. Flies also use a primitive version of communication during the mating process (Kamikouchi, 2013; Aranha and Vasconcelos, 2018). Male flies will vibrate their wings to generate a “song” that the female flies must respond to during the mating process. If the male fly does not produce the song correctly, or at all, there will be no mating. Similarly, if the female fly does not respond to the song or interpret the song correctly as a mating call, she may not mate with the male (Greenspan and Ferveur, 2000). Therefore, if we assay the DGRP line in a male during courtship, we are measuring a form of expressive communication (the song) and if we assay

a DGRP line in a female to a wild type male during courtship we are measuring a form of receptive communication. Here we use the latency to mating during courtship behavior to quantify these differences in expressive and receptive communication, both prominent features of ASD.

To detect differences in response to *social reciprocity*, we used a relatively new test for social spacing developed in flies. This assay exploits the natural tendency of flies to (1) be negatively geotactic (i.e., flies tend to move upward) and (2) settle into an average spacing of ~ 1.5 fly lengths apart (Simon et al., 2012). Flies were tested in groups of 25, and the average distance between flies in a group was quantified for each DGRP genotype.

To assay *repetitive behaviors/restricted interests*, we utilized a simple grooming behavior test. Flies typically groom for a few seconds when placed into a new environment, so individual DGRP flies being tested were moved into a chamber for video recording of total time spent grooming during a 5 min trial. Each of these tests produced sensitive measures that capture trait variation in DGRP lines analyzed. We conducted a GWAS with the resulting data to identify genes that potentially influence these behaviors, as in previous DGRP studies (Mackay and Huang, 2018).

GWAS and Analysis of 40 DGRP Lines Revealed Significant Associations

Forty DGRP lines were picked at random from the collection for testing in all three behavior paradigms. Each line was tested in triplicate for a given behavior, and the results averaged for input into the genome wide association (GWA) calculation. Males and females were scored separately but input together in order to identify male- or female-specific trait effects. A summary of the SNPs and genes identified for each behavior, with mixed p value < 0.001 , is provided in **Table 1**. Complete input data for each behavior in all 40 lines and a plot illustrating variation within the population for each behavior, separately in males, and females is provided in **Supplementary Data 1**. A heatmap of GWAS results for receptive and expressive communication, as measured by mating latency, is presented in **Figure 1**. We detected 159 variants (SNPs, indels or insertions) in or near several genes that showed significant association with receptive communication (females receptive to the male's song) including *Neurologin 2 (Nlg2)*, homologous to the autism associated gene *NLGN3*, and *sulfateless (sfl)*, a gene that encodes an enzyme involved in heparan sulfate residue synthesis (**Figure 1A**). Similarly, expressive communication showed strong association with 49 variants, including several in loci on chromosome 3L, such as *neurexin 4 (Nrx-IV)*, orthologous to the ASD involved

gene contactin-associated protein 2 (*CNTNAP2*) and again, *sfl*. Also associated was *numb*, homologous to *NUMBL (numb-like)*, a human gene involved in nervous system development but not implicated in ASD, called *NUMBL* in humans (**Figure 1B**).

Using social spacing as a proxy for social reciprocity, we identified significant association at 51 variants across the genome, including a single variant located within intron 1 of the *sfl* gene (**Figure 2A**). Finally, 302 variants were significantly associated with repetitive grooming behavior in the 40 DGRP lines tested (**Figure 2B**), including the gene *kin of irre (kirre)*. The homologous human gene *KIRREL3* has been associated with ID (Kalsner et al., 2018). Once again, a SNP included in intron I of the *sfl* gene was associated with repetitive grooming behavior.

Venn diagram analysis¹ for associated genes of all four behaviors tested confirmed that the only gene which appears to be associated with social communication, social spacing, and grooming is the sulfotransferase encoding gene, *sfl* (**Figure 3**). All variants associated with each behavior domain, the minor allele frequencies and the p -values for each variant as well as the gene names and potential regulatory effects are listed in the **Supplementary Data 2**.

Next, we combined measures from all behavioral domains into one ASD-like measure for each DGRP line. We calculated Z -scores for each DGRP line across all four behaviors tested and summed these values, producing a single value for each DGRP line. This value represents how far from the mean each DGRP line fell when considering all four behavioral measures combined, a similar approach to how ASD is diagnosed in humans as ASD testing relies on results from multiple behavioral domains. We conducted a GWAS on the combined Z -score value for each DGRP line and identified 131 significantly associated variants, including a single variant in located upstream of the *sfl* gene (**Supplementary Data 3**).

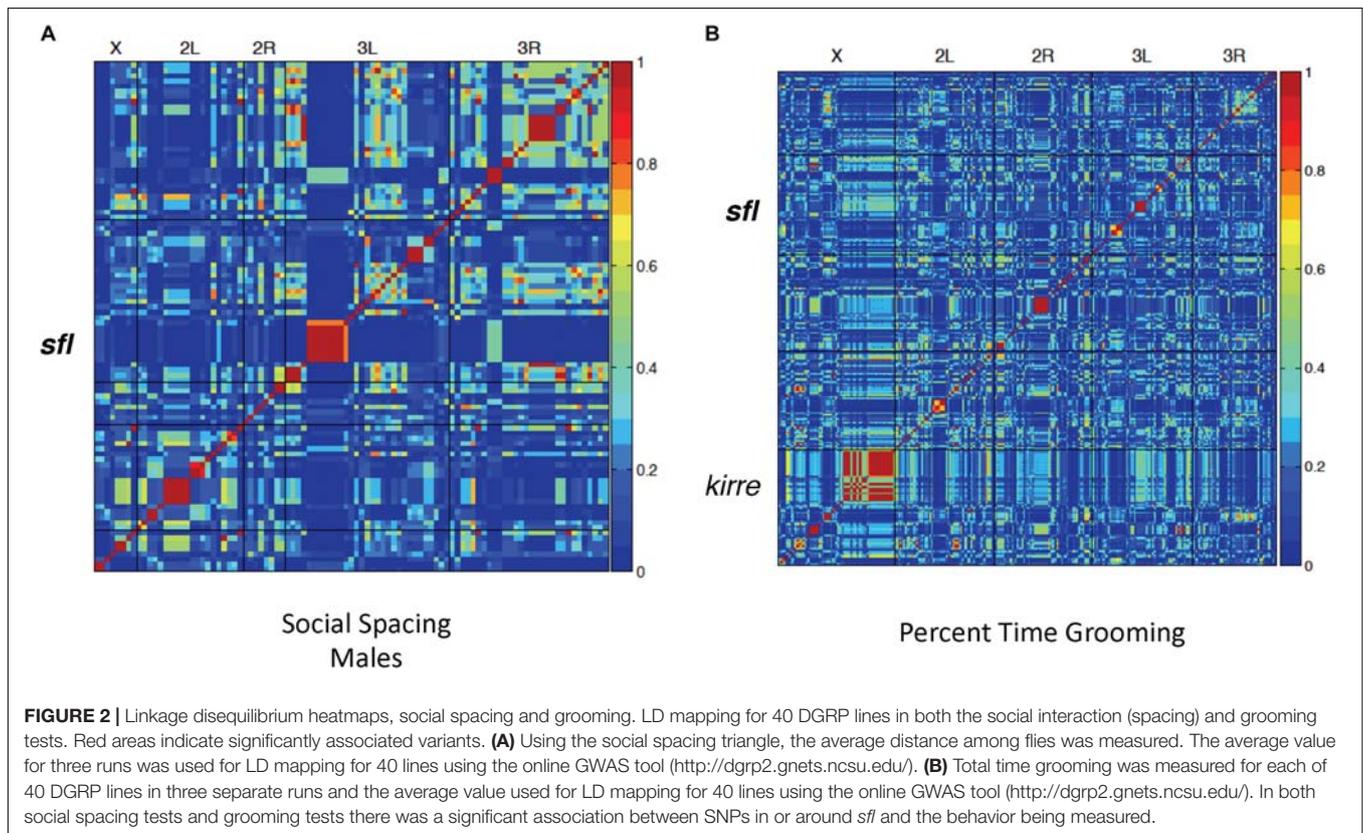
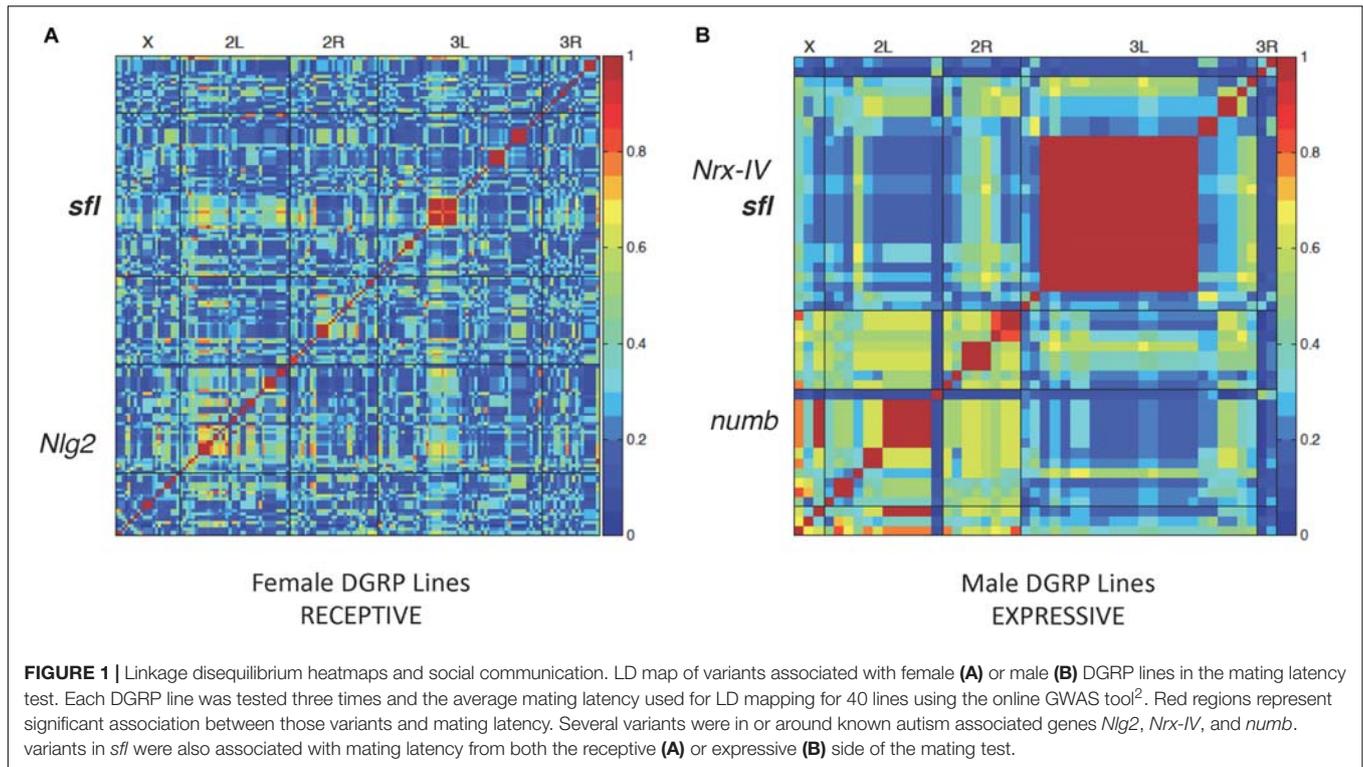
Knockdown of sfl in Neurons Significantly Affects All ASD-Like Behaviors Tested

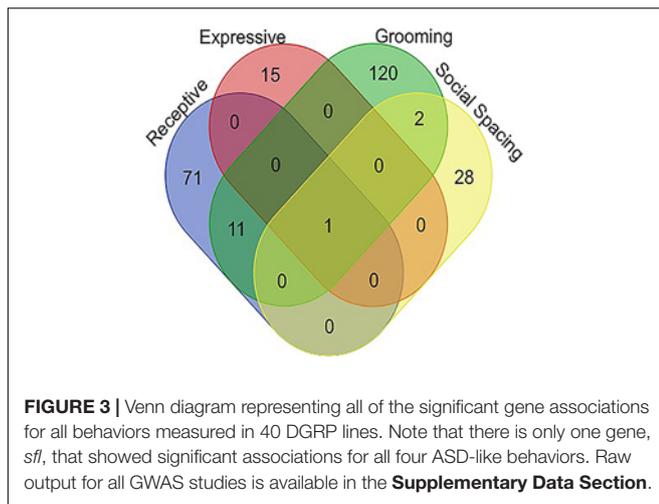
Since all of the individual behaviors tested, as well as the combined Z -score analysis, showed at least one association to *sfl*, we focused functional analysis on experiments designed to test involvement of *sfl* in ASD-relevant behaviors. First, we looked at each variant for a potentially direct role in the regulation of *sfl*. Several transcription factor binding sites for critical developmental factors (*delta*, *hairless*, *twist*, *zeste*, *giant*, *trithorax-like*, *hunchback*, *medea*, and *daughterless*) are located in the upstream region containing SNPs associated with expressive communication. However, none of these transcription factors have been shown to regulate mating behavior previously. There were no other transcription factor binding sites disrupted by *sfl* SNPs identified in this study, and the majority of variants occurred in intragenic or intronic regions and did not obviously disrupt protein coding or gene regulation. All of the SNPs for *sfl* were mapped back to the UCSC genome browser

TABLE 1 | Summary of GWAS results.

Behavior	SNPs (mixed p value < 0.001)	Genes associated
Mating (receptive)	159	83
Mating (expressive)	46	16
Grooming	302	134
Social spacing	51	31

¹<http://bioinformatics.psb.ugent.be/webtools/Venn/>





for Flybase Release August 06, 2014 and are illustrated in **Supplementary Figure 1**.

Given the association between all behaviors tested and *sfl*, we used the GAL4/UAS system (Duffy, 2002) to knock down *sfl* expression in neurons using UAS-RNAi and the pan-neuronal driver, *elav*-GAL4. We then looked for significant changes in any of these behaviors vs. control [empty attP recombination site for the TRiP project (Hu et al., 2017)]. For each behavior we measured males and females separately to account for sex specific effects. We found a significant decrease in the time spent grooming in *elav > sfl*-RNAi males, but not females (two-tailed *t*-test p value ≤ 0.05) (**Figure 4A**). Similarly, we saw a decrease in the number of grooming events in males, but not females with knockdown for *sfl* in neurons (two-tailed *t*-test p value ≤ 0.05) (**Figure 4B**). We also saw a significant increase in social spacing for both males and females in *elav > sfl*-RNAi flies (two-tailed *t*-test p value ≤ 0.05) (**Figure 4C**). Finally, we found that both expressive and receptive communication during mating behavior is disrupted by knock down of *sfl* in neurons (**Figures 4D,E**). Male *elav > sfl*-RNAi flies had significantly longer copulation latencies compared to *elav > 36303* control males when paired with *Canton-S* wildtype virgin female flies (*t*-test, $p = 0.0317$). Similarly, female *elav > sfl*-RNAi flies displayed significantly longer copulation latencies compared to *elav > 36303* flies when paired with virgin male *Canton-S* flies (*t*-test, $p = 0.0157$). These results show a clear causation between decreased *sfl* expression in neurons and all three ASD-related behavior domains tested.

DISCUSSION

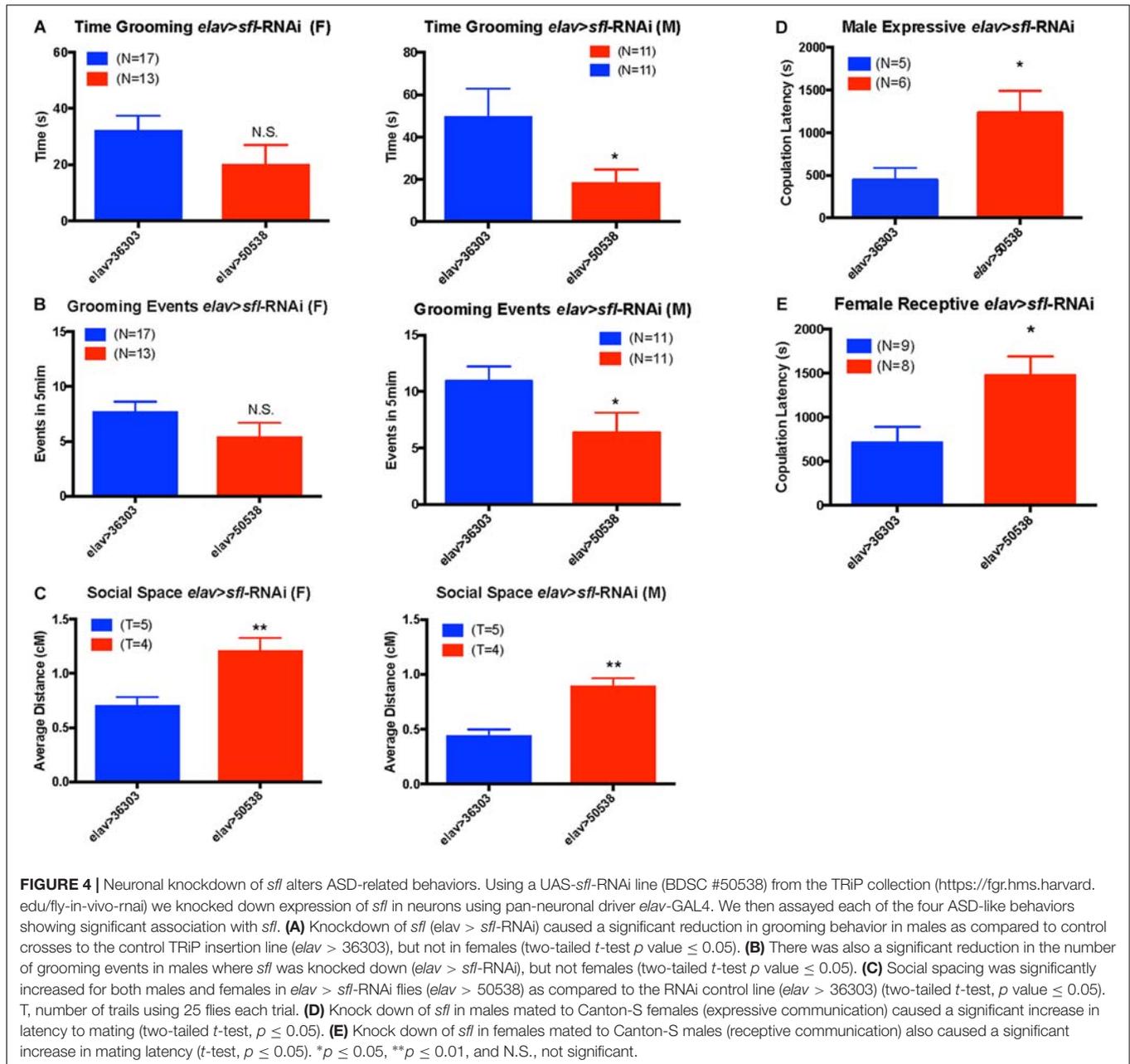
Drosophila as a Model System for Human ASD

Simple behaviors in *D. melanogaster* have been used to model human conditions for over a century (Ugur et al., 2016). Early studies of epilepsy, neurodegeneration and circadian rhythm may be the most familiar examples. However, after we discovered a strong molecular connection between human disease associated

genes and the fly genome [namely that 75% of the genes known to cause human disorders have homologs in flies (Reiter et al., 2001)] there has been renewed interest in utilizing this easily tractable genetic model organism specifically for the study of human disorders, even among researchers unfamiliar with the abundant genetic tools available in flies (Wangler et al., 2015). In terms of this study, an interest in Fragile X syndrome modeling in flies by many labs somewhat parallels our use of behaviors in flies to approximate ASD-like behaviors (Drozd et al., 2018). In particular, one group utilized grooming behavior, as we did, as a read out for ASD-like repetitive behavior (Tauber et al., 2011). In fact, the fly social spacing assay we use here (Simon et al., 2012) has been used to assess a curious phenomenon in humans with ASD, specifically that the age of the father increases the risk of ASD in offspring (Brenman-Suttner et al., 2018). What is unique about our study is that we have combined these previously published ASD assays as a screening tool for the DGRP in order to evaluate the influence of multiple ASD related genes at once, which may be difficult in other systems, either mice or human GWAS, because they are not single genes with strong effects, but rather have more subtle genetic influence on the individual ASD-like behaviors and the combined Z-score across these behaviors.

Here we used an un-biased GWA approach in *Drosophila* to detect genes involved in regulating behaviors relevant to ASD. The DGRP lines used in this study allow for the analysis of naturally occurring behaviors in animals that represent the genetic variation of individuals from a wild caught population with little to no background effects and complete genome sequences for all 205 DGRP lines (Mackay et al., 2012). Using only 40 of the 205 DGRP lines available, we identified associations between three ASD-related behaviors: social communication during mating, social reciprocity during the social space test, and repetitive behavior during grooming. Importantly, our analysis implicated several known ASD genes in regulating these behaviors, providing validity to the approach. Moreover, this analysis revealed a single novel gene, *sfl*, that influences all three behavioral domains.

While there are now many single genes, especially many new *de novo* mutations, known to affect ASD behavior in humans, we were not looking for a single gene at the onset of this study. Our hypothesis was that multiple genes would have subtle effects on ASD behavior and would be difficult, if not impossible, to identify in mice or humans using GWAS, but could be potentially be identified using the power of the DGRP system vs. the more complex genomes of mice or humans. We were, in fact, surprised to find a single gene affecting all three domains tested since we expected a more cumulative effect of multiple genes across all of the domains. The Z-score approach, which also revealed the influence of *sfl* across all the behaviors tested, was the original intention of the project. However, we soon realized that finding individual genes affecting specific ASD-like domains like grooming behavior or social communication were also interesting results. Although the finding of a single gene that can influence all three behavioral domains was unexpected, the mechanism by which a human heparan sulfate (HS) N-sulfotransferase may be involved in ASD phenotypes in



humans is not. HS proteoglycans are involved in many aspects of neural development including neurogenesis, axon guidance, and synaptogenesis (Yamaguchi, 2001). Furthermore, loss of HS has been firmly linked to ASD-like phenotypes in murine models. The mouse knock-out for *exostosin-1* (*Ext1*), a gene encoding an HS-modifying enzyme essential for the production of HS in the brain, also shows strong ASD-like behaviors including impaired social interaction and increased repetitive behaviors as well as impaired ultrasonic vocalization (Yamaguchi, 2001). Furthermore, a SNP in *EXT1* was the most significant genic association (2.94×10^{-7}) in a recent GWAS from the ASD working group of the PGC (Autism Spectrum Disorders Working Group of the Psychiatric Genomics Consortium, 2017).

In addition to *sfl*, we identified an association between combined Z-score and the non-synonymous coding variant M330L in the gene *Fps85D*, also known as *FER tyrosine kinase*. DGRP lines carrying the lysine variant tended to have higher Z-scores (Supplementary Data 3), and thus increases in the behavioral measures, compared to those carrying the methionine variant. These data suggest that the expression or function of *Fps85D* also may impact the behavioral domains we tested. Additional studies will be required to confirm this hypothesis.

The canonical BTBR mouse model (*BTBR T + tf/J*) has been used extensively to study ASD behaviors. BTBR mice have been shown to lack fractal-like structures (fractones) in the sub-ventricular zone of the brain (Meyza et al., 2012). These

fractones carry HS cargo to various locations in the brain, thus depriving the developing brain of a key component required for proper neural development and synaptogenesis (Irie et al., 2012). Finally, studies on postmortem human ASD brain samples using immunofluorescence to detect changes in HS in the sub-ventricular zone recapitulate the findings from the BTBR mouse model, showing decreased levels of HS in ASD brain vs. controls (Pearson et al., 2013). While there is no direct evidence to suggest that the human ortholog of *sfl*, *NDST2*, is itself an ASD gene, the highly conserved HS functions between flies and humans point, in this case, to a shared process associated with ASD, if not a specific gene. Our study now connects these changes in HS levels in neurons directly to the behaviors of social communication, repetitive behavior, and social reciprocity through an HS modifying enzyme encoded by *sfl*.

The implications here are that even subtle changes in HS modifying enzymes or HS proteoglycan synthesis may have a large effect on behavior. One possible reason for this is the diversity of proteoglycan chains that can be created in order to specify appropriate interactions between cells in the brain. In fact, mutations in the cell membrane bound glypican 4 (GCP4) protein have been found in ASD (Doan et al., 2016) and it would not be surprising to find other HS proteoglycans associated with human psychiatric conditions based on their roles in synaptogenesis (Condomitti and de Wit, 2018).

The DGRP collection was primarily designed as a tool to investigate naturally occurring variation in a single population. Our study is the first to utilize the DGRP for the direct study of behaviors related to autism, but not the first to use this powerful tool for the study of human genetic disease. This collection has been used to successfully identify modifiers of neurodegeneration in a fly Parkinson's disease model (Lavoy et al., 2018) and also candidate modifiers of retinitis pigmentosa using an F1 crossing strategy in combination with the GAL4/UAS system (Chow et al., 2016).

Caveats and Future Directions

In this study, we used a GWA approach in conjunction with the DGRP to identify candidate genes for regulating ASD-related behaviors in the fly. One major caveat to the study is that we selected naturally occurring behaviors in flies that may or may not be analogous to ASD behaviors in humans, although the detection of associations with known ASD associated genes lends confidence in the use of these behavioral traits. We also are aware that there are several ways in which the individual behaviors tested here could yield false positive or false negative results. For example, we used mating behavior as a proxy for social communication and assume that the defects in latency to mating are due to defects in male song or female reception to the mating song. Another possibility, however, is that the flies being tested for a given DGRP line could have locomotor defects that cause them to take longer to mate, or cause them to not mate at all. It is unlikely that these animals do not copulate, however, since each DGRP line can be maintained as an isogenic stock. Similarly, in the grooming behavior, flies with severe motor defects may want to groom excessively, but simply cannot do so due to physical limitations, although we are confident that we would

have easily observed such global motor defects had they been present. In the end, such movement defects could eliminate a few individual DGRP lines from the analysis, but it should not alter the positive results we found for associations between a given behavior and particular genes, especially those where multiple SNPs were found in or near a given gene.

Although RNAi knock down experiments in neurons confirmed a role for *sfl* in all three behaviors, it is likely that a larger GWA screen using all 205 sequenced lines would reveal additional genes associated with all three behavior domains (and even more affecting two behavioral domains only). Our secondary analysis of all three behavior domains at once using normalized Z-scores may be a more useful direct link to ASD than the genes found in the individual association studies. While *sfl* was the only gene found in both the three behavioral domains and the combined Z-score analysis, which is why we focused on *sfl* for secondary validation studies, other genes were identified in the Z-score analysis that may similarly, contribute to relevant behaviors that require further testing. Future studies should include more detailed analysis of HS residues in postmortem human brain as well as additional candidate gene approaches focused on HS synthesis and modification genes to identify risk variants in human ASD cohorts. Finally, additional validation studies using RNAi knock-down or over-expression in neurons to verify the involvement of many of the candidate genes from this study could be performed to expand the list of ASD associated candidate genes identified using this approach.

MATERIALS AND METHODS

Fly Stocks

All stocks used in this study came from the Bloomington Stock Center (BDSC) including the Genetic Reference Panel (DGRP) screening kit, the control Canton S line (64349), TRiP UAS-RNAi lines for *sfl* (50538), and the control TRiP insertion site (36303). The pan neuronal *elav*-GAL4 stock came from Dr. Hugo Bellen, Baylor College of Medicine, Houston, TX, United States. All flies were grown on standard corn meal agar media (BDSC) and kept at 25°C in a humidified environmental chamber on a 12 h day/night cycle.

Behavioral Testing

Mating, grooming, and social spacing behaviors have all been previously described (Greenspan and Ferveur, 2000; Simon et al., 2012). For mating and grooming assays, single flies or pairs of flies (one male and one female) were video recorded using an Apple iPhone 5 for later analysis. Average latency to mating (seconds) was scored as the time until the male fly successfully mounted the female fly and was recorded for each line for both expressive and receptive social communication. Total average time grooming in a 5 min interval was used for repetitive behavior. For social spacing, 25 flies were used in each experiment (all of the same genotype and sex) and the flies were allowed to equilibrate in the social space triangle for 1 h prior to photographing with an Apple iPhone 5. ImageJ was used to calculate the nearest neighbor

distance for each individual fly and nearest neighbor distances were averaged within each experiment.

A total of 40 DGRP lines were selected at random from the collection of 205 lines and used for all three behavior tests. Each individual test on each DGRP line was performed three independent times to obtain and average performance value on that test (latency to mating, time grooming, or social spacing). **Supplementary Data 1** contains all of the average input values for each DGRP line used in this study. These average values were input into the GWAS tool (males and females listed separately). The output of the GWAS tool² provides a large amount of data including the heatmaps illustrated in **Figures 1, 2** as well as the table output data (Allele frequencies, sex effects, location of associated SNPs and other information) available in **Supplementary Data 2** (each Excel Tab represents GWAS for a single behavior domain). A complete set of all raw data used to calculate the average values for input into GWAS (individual animals tested or individual trails for social spacing) is compiled in **Supplementary Data 4**.

Z-Score Analysis

A summed Z-score was calculated for each DGRP line to combine measures from all behavioral domains into one integrated score. Z-scores were calculated by finding the mean and standard deviation for each behavioral assay (separating male and females for grooming and social space), and then calculating how many standard deviations from the mean each DGRP line fell. Z-scores were summed across all behavioral assays to give each DGRP line a summed Z-score. The summed Z-score was used for GWAS analysis².

sfl Knockdown Experiments

Using the GAL4/UAS system (Duffy, 2002) and a UAS-RNAi line (BDSC #50538) available from the Harvard TRiP collection, expression levels of *sfl* were knocked down in all neurons using the *elav*-GAL4 driver. The control line, BDSC# 36303, contains an attP40 insertion site without the UAS-RNAi insert and is used as a control for experiments utilizing these TRiP lines. Female *elav*-GAL4 virgins were crossed to males for the TRiP collection and the offspring allowed to mature at least 3 days prior to

² <http://dgrp2.gnets.ncsu.edu/>

REFERENCES

- American Psychiatric Association (2013). *DSM-5 Task Force: Diagnostic and Statistical Manual of Mental Disorders*. Washington, DC: American Psychiatric Association.
- Aranha, M. M., and Vasconcelos, M. L. (2018). Deciphering *Drosophila* female innate behaviors. *Curr. Opin. Neurobiol* 52, 139–148. doi: 10.1016/j.conb.2018.06.005
- Autism Spectrum Disorders Working Group of the Psychiatric Genomics Consortium (2017). Meta-analysis of GWAS of over 16,000 individuals with autism spectrum disorder highlights a novel locus at 10q24.32 and a significant overlap with schizophrenia. *Mol. Autism* 8:21.
- Brenman-Suttner, D. B., Long, S. Q., Kamesan, V., de Belle, J. N., Yost, R. T., Kanippayoor, R. L., et al. (2018). Progeny of old parents have increased social space in *Drosophila melanogaster*. *Sci. Rep.* 8:3673. doi: 10.1038/s41598-018-21731-0

behavior testing. All animals were reared in an incubator on a 12 h day/night cycle. Behavior tests were performed as described above using at least 10 animals per genotype for grooming and mating behaviors and at least 4 independent runs (25 flies each run) for the social spacing behaviors. Statistical analysis was performed using Prism 7.0 to calculate the standard error of the mean, significance by Student's *t*-test (*p* value ≤ 0.05) and graph the results.

AUTHOR CONTRIBUTIONS

KH designed and executed the experiments, interpreted the data, and prepared the manuscript. DF, PC, and BM executed the experiments and interpreted the data. LR and JO'D developed the concept, managed the project, interpreted the data, and prepared the manuscript. JS interpreted the data and prepared the manuscript.

FUNDING

This work was supported in part by the National Institutes of Health (R01NS059902 and R21HD091541 to LR) and a Duplication 15q Alliance fellowship award to KH.

ACKNOWLEDGMENTS

We thank several undergraduate researchers who assisted in aspects of the study and the technical support of Ken Gentner. We also thank Drs. Anholt and MacKay for development of the DGRP resource which made this work possible. Stocks obtained from the Bloomington *Drosophila* Stock Center (NIH P4OD018537) were used in this study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2019.00574/full#supplementary-material>

- Chow, C. Y., Kelsey, K. J., Wolfner, M. F., and Clark, A. G. (2016). Candidate genetic modifiers of retinitis pigmentosa identified by exploiting natural variation in *Drosophila*. *Hum. Mol. Genet.* 25, 651–659. doi: 10.1093/hmg/ddv502
- Condomitti, G., and de Wit, J. (2018). Heparan sulfate proteoglycans as emerging players in synaptic specificity. *Front. Mol. Neurosci.* 11:14. doi: 10.3389/fnmol.2018.00014
- De Rubeis, S., He, X., Goldberg, A. P., Poultney, C. S., Samocha, K., Cicek, A. E., et al. (2014). Synaptic, transcriptional and chromatin genes disrupted in autism. *Nature* 515, 209–215. doi: 10.1038/nature13772
- Devlin, B., Melhem, N., and Roeder, K. (2011). Do common variants play a role in risk for autism? Evidence and theoretical musings. *Brain Res.* 1380, 78–84. doi: 10.1016/j.brainres.2010.11.026
- Doan, R. N., Bae, B. I., Cubelos, B., Chang, C., Hossain, A. A., Al-Saad, S., et al. (2016). Mutations in human accelerated regions disrupt cognition and social behavior. *Cell* 167, 341–354.e312. doi: 10.1016/j.cell.2016.08.071

- Drozd, M., Bardoni, B., and Capovilla, M. (2018). Modeling fragile X syndrome in *Drosophila*. *Front. Mol. Neurosci.* 11:124. doi: 10.3389/fnmol.2018.00124
- Duffy, J. B. (2002). GAL4 system in *Drosophila*: a fly geneticist's Swiss army knife. *Genesis* 34, 1–15. doi: 10.1002/gene.10150
- Greenspan, R. J., and Ferveur, J. F. (2000). Courtship in *Drosophila*. *Annu. Rev. Genet.* 34, 205–232.
- Hu, Y., Comjean, A., Roesel, C., Vinayagam, A., Flockhart, I., Zirin, J., et al. (2017). FlyRNAi.org-the database of the *Drosophila* RNAi screening center and transgenic RNAi project: 2017 update. *Nucleic Acids Res.* 45, D672–D678. doi: 10.1093/nar/gkw977
- Iossifov, I., Ronemus, M., Levy, D., Wang, Z., Hakker, I., Rosenbaum, J., et al. (2012). De novo gene disruptions in children on the autistic spectrum. *Neuron* 74, 285–299. doi: 10.1016/j.neuron.2012.04.009
- Irie, F., Badie-Mahdavi, H., and Yamaguchi, Y. (2012). Autism-like socio-communicative deficits and stereotypies in mice lacking heparan sulfate. *Proc. Natl. Acad. Sci. U.S.A.* 109, 5052–5056. doi: 10.1073/pnas.1117881109
- Kalsner, L., Twachtman-Bassett, J., Tokarski, K., Stanley, C., Dumont-Mathieu, T., Cotney, J., et al. (2018). Genetic testing including targeted gene panel in a diverse clinical population of children with autism spectrum disorder: findings and implications. *Mol. Genet. Genomic Med.* 6, 171–185. doi: 10.1002/mgg3.354
- Kamikouchi, A. (2013). Auditory neuroscience in fruit flies. *Neurosci. Res.* 76, 113–118. doi: 10.1016/j.neures.2013.04.003
- Lavoy, S., Chittoor-Vinod, V. G., Chow, C. Y., and Martin, I. (2018). Genetic modifiers of neurodegeneration in a *Drosophila* model of Parkinson's disease. *Genetics* 209, 1345–1356. doi: 10.1534/genetics.118.301119
- Li, Z., Chen, J., Yu, H., He, L., Xu, Y., Zhang, D., et al. (2017). Genome-wide association analysis identifies 30 new susceptibility loci for schizophrenia. *Nat. Genet.* 49, 1576–1583. doi: 10.1038/ng.3973
- Liao, W., Gandal, M. J., Ehrlichman, R. S., Siegel, S. J., and Carlson, G. C. (2012). MeCP2^{+/−} mouse model of RTT reproduces auditory phenotypes associated with Rett syndrome and replicate select EEG endophenotypes of autism spectrum disorder. *Neurobiol. Dis.* 46, 88–92. doi: 10.1016/j.nbd.2011.12.048
- Mackay, T. F., Richards, S., Stone, E. A., Barbadilla, A., Ayroles, J. F., Zhu, D., et al. (2012). The *Drosophila melanogaster* genetic reference panel. *Nature* 482, 173–178. doi: 10.1038/nature10811
- Mackay, T. F. C., and Huang, W. (2018). Charting the genotype-phenotype map: lessons from the *Drosophila melanogaster* genetic reference panel. *Wiley Interdiscip. Rev. Dev. Biol.* 7:e289. doi: 10.1002/wdev.289
- Meyza, K. Z., Blanchard, D. C., Pearson, B. L., Pobbe, R. L., and Blanchard, R. J. (2012). Fractone-associated N-sulfated heparan sulfate shows reduced quantity in BTBR T+tf/J mice: a strong model of autism. *Behav. Brain Res.* 228, 247–253. doi: 10.1016/j.bbr.2011.11.004
- Mol, C., Chen, A., Kager, R. W. J., and Ter Haar, S. M. (2017). Prosody in birdsong: a review and perspective. *Neurosci. Biobehav. Rev.* 81, 167–180. doi: 10.1016/j.neubiorev.2017.02.016
- Nakatani, J., Tamada, K., Hatanaka, F., Ise, S., Ohta, H., Inoue, K., et al. (2009). Abnormal behavior in a chromosome-engineered mouse model for human 15q11-13 duplication seen in autism. *Cell* 137, 1235–1246. doi: 10.1016/j.cell.2009.04.024
- Pardinas, A. F., Holmans, P., Pocklington, A. J., Escott-Price, V., Ripke, S., Carrera, N., et al. (2018). Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nat. Genet.* 50, 381–389. doi: 10.1038/s41588-018-0059-2
- Pearson, B. L., Corley, M. J., Vasconcellos, A., Blanchard, D. C., and Blanchard, R. J. (2013). Heparan sulfate deficiency in autistic postmortem brain tissue from the subventricular zone of the lateral ventricles. *Behav. Brain Res.* 243, 138–145. doi: 10.1016/j.bbr.2012.12.062
- Reiter, L. T., Potocki, L., Chien, S., Gribskov, M., and Bier, E. (2001). A systematic analysis of human disease-associated gene sequences in *Drosophila melanogaster*. *Genome Res.* 11, 1114–1125. doi: 10.1101/gr.169101
- Sanders, S. J., He, X., Willsey, A. J., Ercan-Sencicek, A. G., Samocha, K. E., Cicek, A. E., et al. (2015). Insights into autism spectrum disorder genomic architecture and biology from 71 risk loci. *Neuron* 87, 1215–1233. doi: 10.1016/j.neuron.2015.09.016
- Schizophrenia Working Group of the Psychiatric Genomics Consortium (2014). Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 511, 421–427. doi: 10.1038/nature13595
- Simon, A. F., Chou, M. T., Salazar, E. D., Nicholson, T., Saini, N., Metchev, S., et al. (2012). A simple assay to study social behavior in *Drosophila*: measurement of social space within a group. *Genes Brain Behav.* 11, 243–252. doi: 10.1111/j.1601-183X.2011.00740.x
- Smith, S. E., Zhou, Y. D., Zhang, G., Jin, Z., Stoppel, D. C., Anderson, M. P., et al. (2011). Increased gene dosage of Ube3a results in autism traits and decreased glutamate synaptic transmission in mice. *Sci. transl. med.* 3:103ra197.
- Tauber, J. M., Vanlandingham, P. A., and Zhang, B. (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. doi: 10.1371/journal.pone.0027100
- Ugur, B., Chen, K., and Bellen, H. J. (2016). *Drosophila* tools and assays for the study of human diseases. *Dis. Model Mech.* 9, 235–244. doi: 10.1242/dmm.023762
- Wangler, M. F., Yamamoto, S., and Bellen, H. J. (2015). Fruit flies in biomedical research. *Genetics* 199, 639–653. doi: 10.1534/genetics.114.171785
- Yamaguchi, Y. (2001). Heparan sulfate proteoglycans in the nervous system: their diverse roles in neurogenesis, axon guidance, and synaptogenesis. *Semin. Cell Dev. Biol.* 12, 99–106. doi: 10.1006/scdb.2000.0238

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Hope, Flatten, Cavitch, May, Sutcliffe, O'Donnell and Reiter. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.