

1
2
3
4 **Full title:** Natural variation in the regulation of neurodevelopmental genes
5 modifies flight performance in *Drosophila*
6

7 Adam N. Spierer^{1*}, Jim A. Mossman^{1,2}, Samuel Pattillo Smith^{1,2}, Lorin Crawford^{2,3,4},
8 Sohini Ramachandran^{1,2}, David M. Rand^{1,2*}
9

10 ¹ Department of Ecology and Evolutionary Biology, Brown University, Providence, RI, United States of
11 America

12 ² Center for Computational Molecular Biology, Brown University, Providence, RI, United States of
13 America

14 ³ Department of Biostatistics, Brown University, Providence, RI, United States of America

15 ⁴ Center for Statistical Sciences, Brown University, Providence, RI, United States of America
16

17 * Corresponding authors

18 Email: Adam_Spierer@alumni.brown.edu (ANS)

19 Email: David_Rand@brown.edu (DMR)

Abstract

The winged insects of the order *Diptera* are colloquially named for their most recognizable phenotype: flight. These insects rely on flight for a number of important life history traits, like dispersal, foraging, and courtship. Despite the importance of flight, relatively little is known about the genetic architecture of variation for flight performance. Accordingly, we sought to uncover the genetic modifiers of flight using a measure of flies' reaction and response to an abrupt drop in a vertical flight column. We conducted an association study using 197 of the *Drosophila* Genetic Reference Panel (DGRP) lines, and identified a combination of additive and marginal variants, epistatic interactions, whole genes, and enrichment across interaction networks. We functionally validated 13 of these candidate genes' (*Adgf-A/Adgf-A2/CG32181*, *bru1*, *CadN*, *CG11073*, *CG15236*, *CG9766*, *CREG*, *Dscam4*, *form3*, *fry*, *Lasp/CG9692*, *Pde6*, *Snoo*) contribution to flight, two of which (*fry* and *Snoo*) also validate a whole gene analysis we introduce for the DGRP: PEGASUS_flies. Overall, our results suggest modifiers of muscle and wing morphology, and peripheral and central nervous system assembly and function are all important for flight performance. Additionally, we identified *ppk23*, an Acid Sensing Ion Channel (ASIC) homolog, as an important hub for epistatic interactions. These results represent a snapshot of the genetic modifiers affecting drop-response flight performance in *Drosophila*, with implications for other insects. It also draws connections between genetic modifiers of performance and BMP signaling and ASICs as targets for treating neurodegeneration and neurodysfunction.

Author summary

Insect flight is a widely recognizable phenotype of winged insects, hence the name: flies. While fruit flies, or *Drosophila melanogaster*, are a genetically tractable model, flight performance is a highly integrative phenotype, making it challenging to comprehensively identify the genetic modifiers that contribute to its genetic architecture. Accordingly, we screened 197 *Drosophila* Genetic Reference Panel lines for their ability to react and respond to an abrupt drop. Using several computational tools, we successfully identified several additive, marginal, and epistatic variants, as well as whole genes and altered sub-networks of gene-gene and protein-protein interaction networks, demonstrating the benefits of using multiple methodologies to elucidate the genetic architecture of complex traits more generally. Many of these significant genes and variants mapped to regions of the genome that affect development of sensory and motor neurons, wing and muscle development, and regulation of transcription factors. We also introduce PEGASUS_flies, a *Drosophila*-adapted version of the PEGASUS platform first used in human studies, to infer gene-level significance of association based on the distribution of individual variant *P*-values. Our results contribute to the debate over the relative importance of individual, additive factors and epistatic, or higher order, interactions, in the mapping of genotype to phenotype.

Introduction

Flight is one of the most distinguishing features of winged insects, especially the taxonomic order *Diptera*. Colloquially named “flies,” these insects rely on their namesake for many facets of their life history: dispersal, foraging, evasion, migration, and mate finding [1]. Because flight is central to flies’ life history, many of the most critical genes for flight are strongly conserved [2, 3].

These “flight-critical” genes are necessary for flight, even as the structures and neural circuits they form are co-opted for other phenotypes, like courtship song and display [4, 5]. For example, *Wingless* is an important developmental patterning gene necessary for wing formation [6] and *Act88F* is one of the main actin isoforms in the indirect flight muscles [7]. We will designate these types of genes that play outsized roles in enabling flight “flight critical” genes, since altering their sequence or expression profile is more likely to result in large flight performance deficits. On the other hand, we will designate “flight-important” genes as those with more modest effects on flight, since they are important but not critical. In an evolutionary context, purifying selection would act on flight-critical genes more strongly than flight-important genes, meaning flight-important genes can harbor natural variants that might otherwise vary the flight phenotype. These genes are found across systems, including metabolism [8], muscle function [9], neuronal function [10, 11], and anatomical development [12, 13]. Genes filling multiple roles across systems are pleiotropic, and those with sufficient natural variation are likely to contribute to complex traits and disease [14, 15]. These traits and diseases’ independent, yet interconnected, genetic architecture make them inherently challenging

to study because they are comprised of several modifiers of small to moderate effect size [16-18].

We can leverage natural variants in flight-important genes to uncover novel associations between genotype and phenotype that otherwise modify flight-critical genes' function, via Genome Wide Association Study (GWAS). The *Drosophila* Genetics Reference Panel [19, 20] (DGRP) is a common resource for performing this type of analysis. The DGRP is a panel of 205 genetically distinct *D. melanogaster* lines represents a snapshot of natural variation. Previous studies on complex and highly polygenic, quantitative traits identify several candidate loci contributing to insect- and *Drosophila*-specific traits [21-23], as well as traits affecting human health and disease [24-27].

Accordingly, this study was designed to identify the genetic modifiers of flight performance and map the underlying genetic architecture. We screened males and females from 197 of the 205 DGRP lines and analyzed both sexes, as well as the average and difference between sexes. Traditional association studies focus on the contribution of additive and dominant variants, however, these fail to identify different types of modifiers with different effect sizes. Accordingly, we took several approaches to identify modifiers at the individual variant, whole gene, and network levels. Accordingly, we identified 180 additive variants, 70 marginal variants, 12161 unique epistatic interactions, and nine interaction sub-networks containing 539 genes contributing to flight performance. We also identified 72 whole genes using `PEGASUS_flies`, a novel modification of the human-based `PEGASUS` program [28] that we modified to work with

105 *Drosophila* and DGRP studies <https://github.com/ramachandran-lab/PEGASUS_flies>
 106 (File S4).

107
 108 Taken together, our results strongly suggest variation in flight performance across
 109 natural populations is affected by cis- and trans-regulatory elements' role in modifying 1)
 110 development of wing morphology, indirect flight musculature, and sensory organs; and
 111 2) the connectivity between the peripheral and central nervous systems. These results
 112 are further supported by functional validations of 13 candidate genes, many with known
 113 roles in altering neurogenesis and development. Overall, our results suggest important
 114 roles for modifiers of BMP signaling in neurodevelopment and pickpocket 23 (ppk23)—a
 115 degenerin/epithelial sodium channels (DEG/ENaC) homologous in humans with Acid
 116 Sensing Ion Channels (ASIC)—in altering affecting flight performance. These findings
 117 address an underexplored body of literature [29-32] calling for genetic and
 118 pharmacological targets of BMP signaling genes and ASIC for treating neuroinjury and
 119 neurodegenerative diseases in humans.

Results

Variation in flight performance across the DGRP

Cohorts of approximately 100 flies from 197 lines of the DGRP (S1 Table) were tested for flight performance using a flight column [33] (Fig 1A). We confirmed the repeatability of our assay by retesting 12 lines of varied ability reared 10 generations apart. We observed very strong agreement between generations ($r = 0.95$; S1 Fig), affirming a role for genetic, rather than environmental or experimental, variation in driving phenotypic variation. We recorded high-speed videos for a weak, intermediate, and strong genotype entering the flight column (Figs 1B-D; S1 File) and concluded this assay is best for studying the reaction and response to an abrupt drop. There was strong agreement between sex-pairs' mean landing height for each genotype ($r = 0.75$; Fig 1E), suggesting the genetic architecture is mostly shared between the sexes. As expected, there was a modest degree of sexual dimorphism in performance, with males outperforming females (male: $0.80\text{m} \pm 0.06 \text{ SD}$; female: $0.73\text{m} \pm 0.07 \text{ SD}$; Fig 1F; S2 Table), though the broad sense heritability (H^2) for each sex was nearly the same ($H^2_{\text{Male}} = 13.5\%$; $H^2_{\text{Female}} = 14.4\%$). In addition to males and females, we also investigated the phenotypic variation in the average (sex-average) and difference (sex-difference) between sexes (S2 Fig).

Fig 1. DGRP lines show differences in flight performance across lines. (A) Flight performance assay measures the average landing height of flies as they fall through a flight column. Vials of flies are sent down the top chute and abruptly stop at the bottom, ejecting flies into a meter-long column. Falling flies will instinctively right themselves and

fly to the periphery, doing so at different times depending on their performance ability. (B-D) Collapsed z-stacks of high-speed video frames from the top quarter of the flight column illustrate these performance differences in (B) weak, (C) intermediate, and (D) strong genotypes. (E) There is sexual dimorphism within genotypes (deviation of red dashed regression line from $y = x$ solid gray line), though sexes are well correlated ($r = 0.75$, $n = 197$). (F) Sexually dimorphic performances are also viewable in the distribution of performances for each male (cyan) and female (red) genotype pair (mean $\pm$ S.E.M.). Sex-genotype pairs are sorted in order of increasing male mean landing height. Genotype performances for genotypes in B-D are indicated on the distribution with the corresponding color-coded asterisk (*) above the respective genotype position.

Before running the association analysis, we tested whether flight performance was a unique phenotype. We compared our phenotype scores for males and female against publically available phenotypes on the DGRP2 webserver. We found no significant regression between flight performance and any of the phenotypes in either sex after correcting for multiple testing ($P \leq 1.85E-3$; S3 Table). This negative result suggests our measure of flight performance is a unique phenotype among those reported.

Association of additive SNPs with flight performance

We conducted a Genome Wide Association Study (GWAS) to identify genetic markers associated with flight performance. We performed an analysis with 1,901,174 common variants ($MAF \geq 0.05$) on the additive genetic effects of four sex-based phenotypes: males, females, sex-average, and sex-difference. Some phenotypes covaried with the

presence of major inversions (S4 Table), so we analyzed association results using a mixed model (Fig 2A) to account for *Wolbachia* infection status, presence of inversions, and polygenic relatedness (S3 and S4 Figs). Annotations and unreferenced descriptors of genes' functions, expression profiles, and orthologs were gathered from autogenerated summaries on FlyBase [34, 35]. These summaries and descriptors were compiled from data supplied by the Gene Ontology Consortium [36, 37], the Berkeley *Drosophila* Genome Project [38], FlyAtlas [39], The Alliance of Genome Resources Consortium [40], modENCODE [34], PAINT[41], the DRSC Integrative Ortholog Prediction Tool (DIOPT) [42], and several transcriptomics and proteomic datasets [9, 12, 39, 43-46].

Fig 2. Variation in flight performance associated with several additive variants, some of which were functionally validated. (A) An additive screen for genetic variants identified several variants that exceeded the traditional DGRP ($P \leq 1E-5$) threshold (gray line). These points (red points) were spread throughout the genome on all but chromosome 4. Sex-average variants pictured, though other sex-based phenotypes had similar profiles. (B) Approximately half of all variants were shared with at least one other sex-based analysis, while the other half of all variants was exclusive to a single analysis. (C) Candidate genes were selected based on the genes that the most significant variants mapped to. Both sexes were tested for flight performance. Validated genes were determined if there was a significant difference between experimental lines homozygous for an insertional mutant in the candidate gene and their background control lines lacking the insertional mutant (red points, Mann-Whitney-

U test, $P \leq 0.05$). Very significant candidate genes (*CadN*, *CG11073/flapper*, and *Dscam4*) each had two independent validation lines.

We filtered additive variants with a strict Bonferroni threshold ($P \leq 2.63E-8$). Taking a MinSNP approach, which identifies significant genes if their lowest (most significant) variant P -value crosses a threshold [28], we identified six unique variants, five of which mapped to six genes (*CG15236*, *CG34215*, *Dscam4*, *Egfr*, *fd96Ca*, *Or85d*) (Table 1). Variants mapping to *Egfr* and *fd96Ca* also contained known embryonic cis-regulatory elements (transcription factor binding sites (TFBS) and a silencer) [47]. Of note, *Dscam4* was deemed “damaged” in 38 of the lines tested [19]; however, the difference between mean landing heights of the damaged vs. undamaged lines was less than 1 cm ($P = 0.32$, Welch’s T-test).

Table 1. Six additive variants surpassed the Bonferroni significance threshold.

Variant	MAF	Annotation		
		Gene (Dmel)	Gene (Hsap)	Regulatory Region
2R_17433667_SNP	0.05128	<i>Egfr</i> (intron)	EGFR	TFBS (<i>bcd</i> , <i>da</i> , <i>dl</i> , <i>gt</i> , <i>hb</i> , <i>kni</i> , <i>Med</i> , <i>prd</i> , <i>sna</i> , <i>tl</i> , <i>twi</i> , <i>disco</i> , <i>Trl</i>)
2R_2718036_DEL	0.05641	<i>CG15236</i> (intron) <i>CG34215</i> (downstream, 764 bp)	- -	-
3L_8237821_SNP	0.0829	<i>Dscam4</i> (intron)	DSCAM	-
3R_20907854_SNP	0.06557	<i>fd96Ca</i> (upstream, 552bp)	FOXB1/ FOXB2	TFBS (<i>dl</i>) Silencer (<i>HDAC</i>)
3R_4379159_SNP	0.05263	<i>Or85d</i> (non-synonymous, C277Y)	-	-
3R_9684126_SNP	0.1514	-	-	-

These variants represented all four sex-based phenotypes and were typically near the minor allele frequency (MAF) > 0.05 limit. All but one mapped to a gene in *Drosophila* (Dmel), and three had human orthologs (Hsap). Additionally, two SNPs mapped to embryonic transcription factor binding sites (TFBS) and a silencer region.

Using the traditional DGRP significance threshold ($P \leq 1E-5$) [48], we identified 180 variants across all four sex-based phenotypes (Fig 2B, S5 Fig, S5 Table). The individual additive variant with the largest effect size contributed 0.045 meters (or 0.97% of the sum of all significant variants) for males and 0.064 meters (1.1% of the sum of all significant variants) for females. For reference, the variant with the smallest significant effect size was 1.7E-4 meters (or 0.0036% of the sum of all significant variants) for males and 5.7E-3 meters (or 0.095% of the sum of all significant variants) for females. All but 19 variants mapped to intergenic or non-coding regions, which are generally indicative of cis-regulatory regions. Of the non-coding variants, 149 mapped to 136 unique genes across the sex-based analyses (Table 2). These included development and function of the nervous system (*aru*, *CadN*, *ChAT*, *chinmo*, *chn*, *CNMaR*, *CSN6*, *DIP-delta*, *Dscam4*, *Egfr*, *fd96Ca*, *form3*, *fry*, *hll*, *htk*, *jeb*, *kek2*, *klg*, *klu*, *Mbs*, *Mmp2*, *nompC*, *Or46a*, *Or85d*, *Pdp1*, *Ptp10D*, *pyd*, *Rbp6*, *rut*, *Sdc*, *SK*, *SKIP*, *Spn*, *Snoo*, *Tmc*), neuromuscular junction (*fend*, *Gad1*, *Gao/Galphao*, *jeb*, *Sdc*, *Syt1*), muscle (*bru1*, *bves*, *CG17839*, *jeb*, *Lasp*, *Pdp1*, *rhea*), cuticle and wing morphogenesis (*CG15236*, *CG34220*, *CG43218*, *Egfr*, *frtz*, *fry*, *Tg*), endoplasmic reticulum (*CG33110*, *CG43783*, *tbc*, *Vti1b*) and Golgi body functions (*Gmap*, *Rab30*, *Vti1b*), and regulation of translation (*mip40*, *mxt*, *Rbm13*, *Wdr37*). Approximately half of all variants were present in two or three sex-based analyses, though the remainder were unique to one (Fig 2B). Several variants mapped to transcription factors (*Asciz*, *Camta*, *CG18011*, *chinmo*, *chn*, *Eip78C*, *fd96Ca*, *Pdp1*, *run*) broadly affecting development and neurogenesis [34, 35]. Despite the enrichment for several annotations, we failed to identify any significant gene

ontology (GO) categories using GOWINDA [49], a GWAS-specific gene set enrichment analysis.

Table 2. Aggregated gene and variant counts by sex-based phenotype for each analysis.

for each analysis:

Additive analysis				
	Male	Female	Sex-Average	Sex-Different
Bonferroni variants ($P \leq 2.63 \cdot 10^{-8}$)	1	4	3	1
Bonferroni MinSNP genes ($P \leq 2.63 \cdot 10^{-8}$)	1	4	3	2
Conventional variants ($P \leq 1.00 \cdot 10^{-5}$)	68	85	85	16
Conventional MinSNP genes ($P \leq 1 \cdot 10^{-5}$)	56	73	69	11
Marginal analysis				
	Male	Female	Sex-Average	Sex-Different
Bonferroni Variants ($P \leq 2.56 \cdot 10^{-8}$)	7	13	62	0
MinSNP Genes ($P \leq 2.56 \cdot 10^{-8}$)	5	7	21	0
Epistatic analysis				
	Male ($P \leq 3.75 \cdot 10^{-9}$)	Female ($P \leq 2.02 \cdot 10^{-9}$)	Sex-Average ($P \leq 4.24 \cdot 10^{-10}$)	Sex-Different
Paired Primary Variants	1	5	18	0
Paired Primary Genes	1	2	6	0
Paired Secondary Variants	42	2188	6139	0
Paired Secondary Genes	28	1061	2419	0
Whole gene analysis				
	Male	Female	Sex-Average	Sex-Different
Bonferroni ($P \leq 3.01 \cdot 10^{-6}$)	23	29	25	23
Network analysis				
	All sex-based phenotypes			
Sub-Networks	9			
Each analysis identified different genetic modifiers (variants, genes, networks). For each analysis, the different variant-, gene-, and network-based analyses identified separate genetic features associated with flight performance.				

General development and neurodevelopmental genes validated to affect flight performance

We performed functional validations on a subset of the genes mapped from variants identified in the Bonferroni and sex-average analysis. We identified 21 unique candidate genes for which a *Minos* enhancer trap *Mi{ET1}* insertional mutation line [50] was publically available [51] (S1 Table; *Adgf-A/Adgf-A2/CG32181*, *bru1*, *CadN*, *CG11073*, *CG15236*, *CG9766*, *CREG*, *Dscam4*, *form3*, *fry*, *Lasp/CG9692*, *Pde6*, *Snoo*). Three additional stocks for *CadN*, *Dscam4*, and *CG11073* were also tested for their strength of association. Finally, an insertion line for *CREG* was also included as a negative control, since it was not significant in the additive or subsequent analyses.

Candidate genes were functionally validated by comparing the distribution in mean landing heights of stocks homozygous for the insertion and their paired control counterpart (S6 Fig) using a Mann-Whitney-U test (Fig 2C; S6 Table). Several were involved in neurodevelopment (*CadN*, *CG9766*, *CG11073*, *CG15236*, *Dscam4*, *form3*, *fry*, and *Snoo*), muscle development (*bru1* and *Lasp*), and transcriptional regulation of gene expression (*Pde6* and *CREG*). Both *CG9766* and *CG11073* are unnamed candidate genes. In validating roles for both these genes, we are naming them *tumbler* (*tumbl*) and *flapper* (*flap*), respectively, based on the tumbling and flapping motions of weaker flies struggling to right themselves in the flight performance assay.

Association of gene-level significance and interaction networks with flight performance

The minSNP approach on the additive variants prioritizes the identification of genes containing variants with larger effects [28]. However, this approach ignores linkage

blocks and gene length, which can bias results. It is important to account for gene length because many neurodevelopmental genes can be lengthy and exceed 100kb (*CadN*, 131kb). One alternative approach is Precise, Efficient Gene Association Score Using SNPs (PEGASUS), which assesses whole gene significance scores based on the distribution of a gene's variant *P*-value distributions with respect to a null chi-squared distribution [28]. This approach enriches for whole genes of moderate effect and enables the identification of genes that might go undetected in a minSNP approach.

While PEGASUS is configured for human populations, we developed PEGASUS_flies, a modified version for *Drosophila* <https://github.com/ramachandran-lab/PEGASUS_flies>. This platform is configured to work with DGRP data sets, and can be customized to accept other *Drosophila*-based or model screening panels. From our additive variants, PEGASUS_flies identified 72 unique genes across the all sex-based phenotypes, whose gene scores passed a Bonferroni threshold ($P \leq 3.03E-6$; S7 Table). These genes were present on five of the six chromosome arms tested (Fig 3A). They were generally different from those identified in the additive approach's minSNP analyses (Fig 3B, S7 Fig), though 15 overlapped (*CG17839*, *CG32506*, *CG33110*, *Gmap*, *Mbs*, *Pdp1*, *Rab30*, *VAcHT*, *aru*, *bves*, *fry*, *mip40*, *mxt*, *oys*, *sdK*). The relatively low overlap between these two gene sets is to be expected, since they prioritize variants of large effect vs. whole genes of moderate effect. Overall, genes annotations were enriched for neural development and function (*aru*, *bchs*, *CG13506*, *ChAT*, *Ccn*, *daw*, *dsf*, *Dip-δ*, *dpr6*, *fry*, *fz2*, *Mbs*, *Pdp1*, *sdK*), wing and development (*CycE*, *daw*, *dsx*, *egr*, *fry*, *fz2*, *Gart*, *HnRNP-K*, *Mbs*, *sno*), Rab GTPase activity (*ca*, *CG32506*, *Gmap*, *plx*,

Rab30), and regulators of transcription (*dsf*, *fry*, *HBS1*, *luna*, *MED23*, *mip40*, *Pdp1*, *Rab30*, *SAP130*, *Tgi*). Different sex-based phenotypes varied in how unique certain whole genes were to a given phenotype (Fig 3C). Genes identified in the sex-average analysis were generally shared with the male and female phenotypes, while genes in the sex-difference analysis were generally unique. Interestingly, *Ccn* was present in both the male and sex-difference, and *dsf* and *sdk* were both present in the sex-average and sex-difference.

Fig 3. PEGASUS_flies identifies different genetic modifiers than the additive screen. (A) PEGASUS_flies results plotted as a Manhattan plot. For the sex-average phenotype, several genes (red points, labeled with gene symbol) exceed a strict Bonferroni significance threshold (gray dashed line, $P \leq 3.43E-6$) identified several genes. (B) PEGASUS_flies prioritizes genetic modifiers of moderate effect, taking into account linkage blocks and gene length. Significant PEGASUS_flies (red) compared against genes significant under a minSNP approach for additive variants (blue) have very little overlap between the two sets (purple). (C) Many of the genes PEGASUS_flies identifies are unique to a sex-based phenotype, though the sex-average genes were generally found in other analyses.

Taking advantage of the gene-level significance scores, we leveraged publicly available gene-gene and protein-protein interaction networks to identify altered sub-networks of genes that connect to the flight performance phenotype. A local False Discovery Rate (lFDR) was calculated for each sex-based phenotype (S8 Table), for which gene-scores

were either $-\log_{10}$ transformed if they passed or set to 0 if they did not. Transformed scores for each sex-based phenotype were analyzed together in Hierarchical HotNet [52], which returned a consensus network consisting of nine sub-networks of genes (S9 Table). The largest network identified 512 genes and was significantly enriched for several GO terms, including transcription factor binding, histone and chromatin modification, regulation of nervous system development, and regulation of apoptosis (S10 Table). The other eight networks were comprised of 27 genes, which together had several significant GO terms, including regulation of gene expression through alternative splicing, maintenance of the intestinal epithelium, and the Atg1/ULK1 kinase complex (S11 Table).

Association of epistatic interactions with flight performance

Epistatic interactions account for a substantial fraction of genetic variation in complex traits [53] but they are statistically and computationally challenging to identify. To circumvent the barriers associated with performing an exhaustive, pairwise search across all possible combinations ($n = 1.81E12$), we turned to MArginal ePIstasis Test (MAPIT) to focus the search area. MAPIT is a linear mixed modeling approach that identifies variants more likely to have an effect on other variants. These putative hub variants represent more central and interconnected genes in a larger genetic network proposed by the Omnigenic Inheritance model [54, 55]. Accordingly, we identified 70 unique significant marginal variants exceeding a Bonferroni threshold ($P \leq 2.56E-8$) across male, female, and sex-average phenotypes. The sex-difference analysis yielded no significant variants (S8 Fig; S12 Table). From these, only 14 had significant epistatic

interactions with other variants in the genome (S13 Table), which we will discuss in order of the male, female, and sex-average results and contextualized with their epistatic interactions.

In males, there were seven significant marginal variants that mapped to five genes (*CG5645*, *CG18507*, *cv-c*, *sog*, *Ten-a*). Of the variants, only one (*X_15527230_SNP*) that mapped to a novel transcription start site in the BMP antagonist of *short gastrulation* (*sog*; human ortholog of *CHRD*) had significant interactions. This marginal SNP interacted with 42 other variants across 28 unique genes (S13 Table). Several of these genes are important in neuron development, signaling, and function (*CG13579*, *Dh31*, *nAChRalpha4*, *Sdc simj*, *sqz*, and *trio*), supporting accumulating evidence of a neurodevelopmental basis for variation in flight performance.

In females, there were 14 significant marginal variants that mapped to six genes (*CG6123*, *CG7573*, *CG42741*, *ppk23*, *Src64B*, *twi*). Of these variants, five mapped to two genes (*CG42671* and *ppk23*) with epistatic interactions. One intronic SNP (*3L_11217593_SNP*) mapped to *CG42671*. Little is known about this gene and there are no human orthologs, but we can gain insights into its function based on the 51 epistatic variants that mapped to 37 genes with annotations for regulation of gene expression (*arx*, *bi*, *CG6843*, *Ches-1-like*, *dve*, *HDAC1*, *Moe*, and *RpL26*, *Sdc*, *Tgi*), and neural development, signaling, and function (*cact*, *CG13579*, *HDAC1*, *ed*, *ngl3*, *nrm*, *numb*, *Sdc*). The other four variants (*X_17459818_SNP*, *X_17459830_SNP*, *X_17460743_DEL*, *X_17460820_SNP*) mapped to a 1002 bp region downstream of

pickpocket 23 (*ppk23*; human homologs in ASIC gene family). *ppk23* is a member of the degenerin (DEG)/epithelial Na⁺ channel (ENaC) gene family that functions as subunits of non-voltage gated, amiloride-sensitive cation channels. It is involved in chemo- and mechanosensation, typically in the context of foraging, pheromone detection, and courtship behaviors [56, 57]. These marginal variants significantly interacted with 2162 variants, which mapped to 1042 genes that were also largely found in the sex-average analysis.

The sex-average phenotype had 62 significant marginal variants (11 also found in females) mapping to 21 genes (*Art2*, *CG10936*, *CG15630*, *CG15651*, *CG18507*, *CG3921*, *CG42671*, *CG42741*, *CG5645*, *CG6123*, *CG9313*, *CR44176*, *cv-c*, *Fad2*, *natalisin*, *ppk23*, *Rbfox1*, *Rgk1*, *Src64B*, *twi*). Of the 62 marginal variants, 18 had significant epistatic interactions: nine were intergenic, seven mapped to *ppk23*, and the remaining four mapped to single genes: *CG42671*, *CG10936*, *CG9313*, and *CG15651* (S13 Table). Previously identified in the female analysis, *ppk23* had the greatest number of interactions, placing it close to the center of a highly interconnected genetic landscape (Fig 4A). The seven marginal variants interacted with 4895 variants across 2010 unique genes, 11 of which mapped to genes that also contained significant marginal variants (*A2bp1*, *cv-c*, *Fad2*, *CG9313*, *CG10936*, *CG42741*, *Rgk1*, *sog*, *Src64B*, *twi*, *Ten-a*). The 2010 unique genes had significant GO term enrichment for neuronal growth, organization and differentiation (S14 Table). One of *ppk23*'s interactors was *CG42671*, itself a gene with a significant marginal variant in the sex-average epistasis screen and previously mentioned in the female epistasis screen. For

the sex-average epistasis screen, *CG42671* interacted with 1013 variants across 616 genes. These genes were significantly enriched in a gene set enrichment analysis for genes involved in neurodevelopment, particularly neuron growth and movement (S15 Table). While this gene is understudied and lacks substantive annotations, but based on its interactors' significant GO categories, it is very likely *CG42671* is involved in growth and neuronal target finding. *CG10936* has few annotations, though it was identified in a screen for nociception [58]. It paired with 29 genes annotated for neurogenesis and function (*CG42788*, *Dh31*, *fru*, *hiw*, *lilli*, *nAChRalpha4*), as well as regulation of gene expression through chromatin modification (*Etl1* and *lilli*) and alternative splicing (*Srp54* and *U2af38*). One SNP (2R_16871314_SNP) was mapped to both the 3' UTR of *CG9313* and 29 bp downstream of *CG15651*. *CG9313* (orthologous to human DNAI1) is an ATP-dependent microtubule motor and is involved in the sensory perception of sound in *Drosophila* and proprioception, as well as sperm development [59]. *CG15651* is predicted to localize to the rough endoplasmic reticulum and Golgi body during embryogenesis, early larval, and late pupation stages where it is expressed in the central nervous system. Its human ortholog, FKRP (fukutin related protein), is implicated in intellectual disability and it is a candidate gene therapy target for muscular dystrophy [60-62]. These genes' shared function in nervous system development is reflected in the variants that map to 87 genes with annotations for neuron development, patterning, and function (*5-HT2B*, *cwo*, *dally*, *dx*, *Dysb*, *enok*, *erm*, *mbi*, *Ngf1*, *nmo*, *Sdc*, *Sema1a*, *sNPF*, *tup*,). Several genes were also annotated for endoplasmic reticulum function (*bark*, *CG5885*, *CG15651*, *Fatp3*, *PAPLA1*, *Trc8*, *Uggt*); chromatin remodeling (*CG43902*, *enok*, *erm*, *lncRNA:roX1*, *tim*); transcription and alternative splicing (*cwo*,

bru3, *CG6841*, *CG9650*, *CG15710*, *enok*, *luna*, *mbi*, *tim*, *tup*); and gene product regulation (*bru3*, *cwo*, *CG5885*, *CG9650*, *CG15710*, *luna*, *tRNA:Arg-TCT-2-1*, *tup*). Finally, there was a 669 bp region with six intergenic variants (chr3L:6890373 - 6891042). This region lacked regulatory annotations, yet collectively interacted with 513 variants mapping to 309 genes, many of which were shared with *ppk23*, *CG42671*, and *CG10936*. Similarly, these genes had significant GO term enrichment for neurodevelopment and neuron function (S16 Table).

Fig 4. Flight performance is a larger complex trait comprised of several smaller traits. (A) The genetic architecture of epistatically interacting genes generally coordinated through *ppk23*. A few other genes mapped to from marginal variants had epistatic interactions with marginal variants in *ppk23*. (B) Genes or genes mapped to from variants across different analyses were not identified in more than three analyses. Roughly half or more genes were unique to each analysis. (C) Flight performance has a complex genetic architecture, with the key developmental gene *Egfr* and BMP signaling pathway contributing to wing and neurodevelopment. These processes are both important for structuring the sensory organs that enable the fly to use mechanosensory channels for proprioception. Signals from the sensory organs on the wing, head, and body travel to the brain and thoracic ganglion, which sends signals through the motor neurons to the direct and indirect flight musculature that is also differentially assembled and innervated to generate power and control the wing angle during flight.

There were epistatic interactions between several of the genes identified from marginal variants (Fig 4A). Since marginal variants represent those more likely to interact with other variants, their interaction with one another suggests a highly interconnected genetic architecture underlying flight performance. Additionally, the breadth of epistatic interactions from a small, focused subset of marginal variants supports an important role for epistasis in the genetic architecture of flight performance. There are likely many more variants that interact with one another. But based on strong enrichment for neurodevelopmental genes from the very limited subset of marginal variants we tested, we hypothesize that flight performance in wildtype *Drosophila* is modulated by neural function and circuitry.

No evidence for adult transcriptome variation affecting flight performance

Since many variants mapped to cis-regulatory elements and trans-regulatory genes, we sought to test whether regulatory variation was affecting developmental or adult homeostasis. Accordingly, performed a Weighted Gene Co-expression Network Analysis (WGCNA)[63] using 177 publically available DGRP transcriptomic profiles for young adults of both sexes [64]. We clustered genes by similarity in expression profile, then correlated those clusters' eigenvalues with the mean and standard deviation of flight performance, as well as the proportion of flies that fell through the column over the total assayed. No clusters across sex or phenotype had a significant correlation. This result squares well with our previous observation that many of the significant variants map to genes involved in pre-adult development, rather than genes that maintain adult homeostasis (S9 Fig).

438

439 **Flight performance is modulated by an interconnected genetic architecture**

440 The genetic architecture of flight performance is comprised of many different types of
 441 genetic modifiers. Many of the variants map to genes that are found across analytic
 442 platforms (Fig 4B). Most variants were unique to a single analysis, suggesting that
 443 association studies should consider using multiple different analyses to enhance the
 444 power to detect variants and genes in their study. However, many genes and genes
 445 identified from variants were identified in two (148) or three (23) analyses. Those
 446 involved in three analyses include: *aru*, *CG2964*, *CG13506*, *CG15651*, *CG17839*,
 447 *CG42671*, *CycE*, *daw*, *Diap1*, *Egfr*, *fz2*, *Gart*, *Gmap*, *Mbs*, *MED23*, *mip40*, *mxt*, *Pdp1*,
 448 *Rab30*, *rhea*, *sog*, *sona*, *Tgi*) analyses. This suggests that individual genes can contain
 449 variants with different types of effects or have differential contributions to the overall
 450 genetic architecture. A complete lookup table of all genes and genes identified from
 451 variants is available (S17 Table).

Discussion

We tested flight performance of 197 DGRP lines, identifying several additive and marginal variants, epistatic interactions, whole genes, and a consensus network of altered sub-networks that associated with variation our phenotype. We identified many cis-regulatory variants mapped to genes with annotations for wing morphology, indirect flight muscle performance, and neurodevelopment of sensory and neuromuscular junctions.

Variation in gene regulation drives variation in flight performance

Variation in gene expression is a major contributor to phenotypic variation [65, 66]. Association studies with the DGRP lines often map variants to intergenic and non-coding regions of genes [22, 48, 67]. These regulatory elements can be cis-regulators, like transcription factor binding sites (TFBS), enhancers, or silencers; or they can be trans-regulatory, like transcription factors, splicosomes, or chromatin modifiers. In the present study, the vast majority of variants in the additive, marginal, and epistatic analyses mapped to introns or within 1kb of a gene, suggesting a cis-regulatory role.

When cis-regulatory elements lie in important developmental genes, their effects can be magnified as the organism continues through development. The most significant additive variant we identified mapped to an *epidermal growth factor receptor* (*Egfr*, human homolog *EGFR*) intron. Encoding a key transmembrane tyrosine kinase receptor, *Egfr* is a pleiotropic gene affecting developmental and homeostatic processes throughout the life and anatomy of the fly. It is well known for its role in embryonic

patterning and implications in human cancers [68, 69]. The variant also mapped to several overlapping TFBS for transcription factors known to affect embryonic development in a highly dose-dependent manner (*bcd*, *da*, *dl*, *gt*, *hb*, *kni*, *Med*, *prd*, *sna*, *tll*, *twi*, *disco*, *Trl*). Variation in patterning cells fated to become tissues and organs can be magnified through the adult stage, especially when that receptor is also known to affect other developmental processes [70]. Other intronic variants were identified in *Egfr* through the epistatic interactions with *ppk23*, illustrating how different types of genetic modifiers can exist within the same gene.

The role of cis- and trans-regulatory elements goes even further when there is variation in cis-regulatory elements of trans-regulatory genes. One of the Bonferroni additive variants mapped to an intronic region of *Forkhead domain 96Ca* (*fd96Ca*; human homologs *FOXB1* and *FOXB2*), a TFBS for *dorsal* (*dl*), and a silencer for *histone deacetylase 1* (*HDAC1*). *fd96Ca* is a fork head box transcription factor expressed in neuroblasts along the longitudinal axis of the embryo and in some sensory neurons in the embryonic head [71]. Trans-regulators, like *fd96Ca*, are proposed to have a large impact on phenotypic variation under the Omnigenic Inheritance model [54, 55]. Similar to *Egfr*, regulatory variation in a gene that helps determine cell fates can have larger effects if not enough cells are allocated for differentiation later in life. This can begin a cascade that amplifies downstream [72] and may hint at why trans-regulators were significant Gene Ontology (GO) terms in the consensus network.

There are likely non-coding regions of the genome that correspond with more cryptic regulatory regions. Six intergenic, marginal variants in a 669bp stretch (chr3L:6890373 - 6891042) had a number of significant epistatic interactions with developmental and neurodevelopmental genes. These variants lacked regulatory annotations in the DGRP2 annotation file, however these annotations were collected during embryogenesis [47] so it is possible these sites are activated by trans-regulators during different times in development. Nonetheless, based on its epistatic interactions, it is likely an important cis-regulatory region that affects general development from an early stage in the fly life cycle.

Our results suggest genetic variation in regulatory (non-coding) regions has a greater affect on variation of flight performance than variation in protein coding regions. While non-synonymous variants can have large effects on flight performance [73-75], they were uncommon in our screen compared with variation in non-coding regions. This may be a result of strong purifying selection acting against them in a natural setting. Many of the candidate modifiers of flight are more commonly expressed during development [39, 45, 46]. This observation is supported by our lack of evidence for adult transcriptomic variation correlating with flight performance. Additionally, the flight phenotype was highly heritable, suggesting our phenotype was not an artifact of environmental or experimental variation. Finally, the constructs we used to validate candidate genes created genetic variation in intronic regions, rather than post-transcriptionally modifying gene expression with an RNAⁱ construct. Our successful validation of several candidate genes suggests variation in the non-coding regions of the candidate genes is sufficient

for observing phenotypic differences. Further, insertion of the constructs into intronic regions both positively and negatively affected performance, even when done at independent sites in the same gene, suggesting a more nuanced impact of genetic variation in cis-regulatory regions. We conclude that modifiers of cis- and trans-regulation in pre-adult stages are more likely to modify flight performance in wild populations than variation in coding sequence.

Variation in wing and indirect flight muscle development contributes to variance in flight performance

Flight performance is a complex trait comprised of coordination across several smaller developmental and functional, complex traits[13, 76, 77]. The central role of *Egfr* in development means it can have wide range of functional effects on adult morphology. Natural variants in *Egfr* are known to cause developmental differences in wing morphology that can significantly alter flight performance [70, 77], in part through interactions with the Bone Morphogenetic Protein (BMP) signaling pathway [13, 70, 78]. BMP signaling is also an established modifier of wing development, as it forms dose-dependent gradients that pattern the wing size and shape [79, 80], as well as sensory and neuromuscular circuits [6, 81]. We identified several modifiers of BMP signaling (*cmpy*, *Cul2*, *cv-2*, *cv-c*, *dpp*, *dally*, *daw*, *egr*, *gbb*, *hiw*, *kek5*, *Lis-1*, *Lpt*, *lqf*, *ltl*, *Mad*, *nmo*, *scw*, *srw*, *Snoo*, *tkv*, *trio*) across all analyses and functionally validated *Snoo*—discussed below. Among the modifiers of BMP signaling, *short gastrulation* (*sog*; human homolog *Chordin*) stood out as a known source of natural variants that modifies flight

performance in natural populations [13]. *sog* affects wing morphology through its role as a *dpp* antagonist in patterning the dorsoventral axis of the wings [80, 82, 83]. *sog* is also noteworthy for its interconnectedness to other genes containing both a significant marginal variant and variants that had epistatic interactions with other significant marginal variants: *ppk23* and *CG42671* (formerly *CG18490* and *CG34240*)—discussed below. Marginal variants represent a class of variants that are statistically more likely to interact with other variants [84], via epistasis. Their identification hints at a more interconnected role in the genetic architecture. In this case, identification of *sog* suggests a more interconnected role for this antagonist of BMP signaling in modifying flight performance.

In addition to wing morphology, we identified several modifiers known to affect flight muscle function. The indirect flight muscles (IFM) power flight through the alternating dorsoventral and dorsolongitudinal muscle contraction to deform the cuticle and move the wings [85, 86], while the direct flight muscles control flight through precise adjust of the wing angle [87]. We identified two genes with known roles in flight [88, 89] from the additive screen that we successfully validated: *Lasp* and *bru1*. *Lasp* (human ortholog *LASP1*), is the only nebulin family gene in *Drosophila*, and shown to modify sarcomere and thin filament length, and myofibril diameter [88]. We also identified *bruno 1* (*bru1* or *aret*; human homolog *CLEF1* and *CLEF2*), a transcription factor that controls alternative splicing of myofibrils in the IFM [9, 89], among other developmental processes. *bru1* had two intronic variants, one of which mapped to a TFBS for *twi*—one of the genes identified from a significant marginal variant.

Using our newly developed platform `PEGASUS_flies` to find significant whole genes, we also identified *tropomodulin* (*tmod*; human homolog *TMOD1*) and *Glycerol-3-phosphate dehydrogenase 1* (*Gpdh1*; human homolog *GPD1*). These two genes were previously validated for their roles in flight performance [8, 90] and are responsible for muscle function and metabolism within muscles, respectively. The identification and previous validation of *tmod* and *Gpdh1* is noteworthy because neither had a significant variant exceed the additive screen's significance threshold ($P \leq 1E-5$). This finding demonstrates a successful proof-of-principle for `PEGASUS_flies`' ability to identify genetic modifiers that would otherwise be overlooked in a traditional minSNP approach in an additive screen. Additionally, we successfully validated *fry*, identified in both the additive and whole gene screens. Taken together, the prior and current validation of these genes establishes `PEGASUS_flies` as a verified platform for identifying modifiers of complex traits.

Neurodevelopmental genes play an important role in modifying flight performance

Many neurodevelopmental genes with diverse functions were identified across analyses. Because neurodevelopmental genes can play several roles, many of which are unannotated in GO databases, GO term enrichment analyses can be underpowered. This may explain why we failed to identify any GO terms for additive variants in the GOwinda analysis [22]. However, their identification through other GO analyses on the epistatic and network-based analyses is encouraging.

Several neurodevelopmental genes overlapped between the additive minSNP and PEGASUS_flies whole gene approach. These genes (*aru*, *ChAT*, *Ccn*, *DIP-δ*, *dsf*, *dsx*, *fry*, *Mbs*, *sdk*, *VACHT*), lend additional support to the likelihood these genes were not false positives. For example, *fry* and *Sidekick* (*sdk*) both coordinate dendritic target finding functions with DSCAM family genes [91, 92]. This is in agreement with several significant GO terms for axon guidance and neuronal targeting in the consensus network's largest sub-network (S11 Table) and for the genes identified from epistatic interactions with *ppk23*, *CG42671*, and an intergenic region (chr3L:6890373 - 6891042)(S14-16 Tables). Accordingly neurodevelopmental genes are present throughout our study, and represent a highly interconnected group of genes that likely plays an important role in flight performance.

Underscoring this interconnectedness is the identification of several neurodevelopmental genes that mapped to epistatic interactions with a common, significant marginal variant in *sog*. This variant was significant in males and mapped to a new transcription start site. In addition to affecting wing morphology, *sog* also plays a role in neurodevelopment (*CG13579*, *dib*, *Hk*, *lncRNA:rox1*, *nAChRa4*, *Sdc*, *simj*, *sqz*, *Toll-4*, *trio*) [36, 37, 81, 83]. Several of these genes were involved in neuromuscular growth and function (*CG13579*, *Hyperkinetic* (*Hk*), *nicotinic acetylcholine receptor α 4* (*nAChRa4*), *Syndecan* (*Sdc*), *squeeze* (*sqz*), *trio*) [81, 93-97], suggesting an important connection between neurodevelopmental phenotypes and their role in activating direct and indirect flight muscles. However, some of the genes interacting with *sog* can affect

sensory neurons as well. For example, *trio* is also present in sensory neurons and is capable of modifying chemosensation [21]. Other *sog* variants that had epistatic interactions with marginal variants in *CG42671* (formerly *CG18490* and *CG34240*) and *ppk23*—discussed below, two genes with known or putative roles in developing the peripheral nervous system (PNS).

In addition to neuromuscular genes, we validated genes involved in patterning the PNS. One of the Bonferroni variants from the additive screen mapped to *Down Syndrome Cell Adhesion Molecule 4* (*Dscam4*; human ortholog *DSCAM*). DSCAMs are a conserved family of extracellular, immunoglobulin proteins that promote cell-cell adhesion. They are found in complex (type IV) dendrite arborization neurons that promote dendritic target recognition and dendrite self-avoidance in the developing PNS [34] and in the brain and central nervous system (CNS) [98, 99]. Type IV dendritic arborization neurons transduce signals from sensory neurons (e.g. photoreceptors, chemosensors, and mechanosensors), to the CNS [99-102]. Dscams are expressed differentially and combinatorially in different neurons, which allows them to create highly interconnected neural circuits [99]. They also work with other cell-cell adhesion proteins, like cadherins, in patterning the nervous system. *Cadherin-N* (*CadN* or *N-cad*) interacts with *Dscam2* and *Dscam4* in patterning olfactory receptor neurons (ORN), like *Or46a* (significant additive hit) and *Or59c* (significant epistatic hit with *ppk23*) [102-105]. Given their importance in patterning sensory neuron circuits and strong significance in the additive screen, we independently validated *Dscam4* and *CadN* using two separate insertional mutants for each. Both pairs of insertional mutants in both genes were significant,

though the direction of effect was reversed, reiterating how cis-regulatory regions can differentially affect genes' expression levels. Our double validation for each supports a greater level of confidence in *Dscam4* and *CadN* as modifiers of the peripheral nervous system important for flight performance.

We validated two other dendrite patterning genes that also help to form sensory organs on the wing and body that contribute to proprioception: *furry* (*fry*; human homolog FRYL) and *Sno oncogene* (*Snoo* or *dSno*; human homolog SKI). These two conserved proteins are expressed along the same types of sensory neurons as Dscams and cadherins that promote dendrite field patterning, dendrite self-avoidance, and development of sensory organs [106]. *fry* assists Dscams and cadherins in dendritic tiling of chemosensors (olfaction or gustation) and mechanosensors (proprioception) [105-107] that directly connect to sensory microchaete (hairs or bristles) organized along the wing and body in specific patterns [108]. Meanwhile, *Snoo* interacts with the wingless pathway [6, 109], and is an important antagonist of *Medea* (*Med* or *dSmad4*; human homolog *Smad4*)—an important regulatory of the BMP-to-activin- β pathway [110]. *Snoo* is known to modify wing shape [110], dendritic tiling, and the development of sensory organs (microchate and campaniform sensilla) on the wing [6, 111]. These sensory organs play different roles; wing chaete can function as chemosensors (olfaction and gustation) and mechanosensors [100, 112], while campaniform sensilla measure strain on the deformed wing blade [113-116]. Together, these sensory organs aid in proprioception of flight [11] and delineate a direct connection between the role of

proper development of the wings' sensory organs and the proper development of the neural circuitry connecting them to the CNS in modifying flight performance.

We functionally validated two candidate genes with only tangential evidence of their function that we are naming *flapper* (*flap*, formerly *CG11073*) and *flippy* (*flip*, formerly *CG9766*). *flapper* is expressed in the peripodial epithelium cells of the eye, leg, and wing imaginal discs [117]. It is expressed at very high levels during 16-18 hours of embryogenesis, pupariation [45] and in the head, eyes, and carcass in the adult stage [118]. It was previously identified as a candidate gene in a screen for modifiers of circadian rhythm [119] and was significantly upregulated in flies bred for aggressive behavior [120], but both studies failed to functionally validate the gene. *flapper* was also implicated in the downregulation of amyloid- β peptides [121] and in late life fecundity [122] suggesting it may play a basic role in development that affects several phenotypes. Accordingly, we hypothesize it plays some role in patterning neural circuitry of sensory neurons on the cuticle and eyes, and facilitates neural circuit assembly in the brain. The other gene, *flippy* (human homolog *FANK1*), is pleiotropic with important roles in neuroanatomical development [43, 123] and sperm development [46]. It is important in the development of trichogen cells, which are precursors to the chaete flies use for mechanosensation. In humans, *FANK1* plays roles in spermatogenesis and apoptosis, and is a putative evolutionary target of balancing selection [124, 125]. Given *flippy*'s pleiotropic role in neurodevelopment and gametogenesis, it may also be under stabilizing selection brought about by contrasting selective pressures for neural function and fitness.

Finally, qualitative observations of differentially performing DGRP lines support a role for proprioception as a modifier of flight performance. High-speed videos of strong, intermediate, and weak lines show strong lines react quicker to an abrupt free fall and are better at controlling their descent than the intermediate fliers, and much more than weak fliers. This direct evidence corroborates with the validation screen and inferential association analyses to support a role for natural variants in genes that affect 1) sensory neural circuit connectivity, 2) development and function of neuromuscular junctions, and 3) the integration of these two onto wings of varying morphologies for modifying flight performance in a natural population.

Important implications for acid sensing ion channels in flight performance and neural function flight

Pickpocket genes encode a conserved group of degenerin/epithelial sodium channels (DEG/ENaC) that function as non-voltage gated, amiloride-sensitive cation channels [56]. They are found in the brain, thoracic ganglion [57, 126], neuromuscular junctions[126, 127], and trachea [128], though pickpocket family genes are most commonly found along type IV dendrite arborization sensory neurons that connect chemo- or mechanosensory organs to the CNS [105-107, 126, 129-132] on the head, legs, and wings [57, 127, 133-136]. Chemosensing microchaete can contain olfactory receptor neurons (ORN), gustatory receptor neurons (GRN), and ionotropic receptors (IR), which are useful for foraging and pheromone detection [11, 57, 127, 133-139]. In this study, we identified six pickpocket genes (*ppk1*, *ppk8*, *ppk9*, *ppk10*, *ppk12*, *ppk23*),

10 gustatory receptors (*Gr10a*, *Gr10b*, *Gr28b*, *Gr36b*, *Gr36c*, *Gr39a*, *Gr59a*, *Gr59d*,
Gr61a, *Gr64a*), 12 olfactory receptors and binding proteins (*Or24a*, *Or45a*, *Or46a*,
Or49a, *Or59b*, *Or59c*, *Or67d*, *Or71a*, *Or85d*, *Obp8a*, *Obp28a*, *Obp47a*), and 13
ionotropic receptors (*Ir41a*, *Ir47a*, *Ir47b*, *Ir51a*, *Ir56b*, *Ir56c*, *Ir56d*, *Ir60d*, *Ir60f*, *Ir62a*,
Ir64a, *Ir67b*, *Ir75d*) from the additive, marginal, epistatic, and network approaches.
Or85d was identified from the 2nd most significant additive variant and only non-
synonymous SNP that passed a Bonferroni threshold in the additive search. And yet,
despite a combined 41 pickpocket, gustatory receptor, olfactory receptor, and ionotropic
receptor genes, only six (*ppk10*, *ppk12*, *Gr59d*, *Or24a*, *Ir41a*, and *Ir60d*) overlapped
with an olfactory screen testing for genetic associations across 14 odors [21].
Accordingly, we hypothesize a more nuanced role for these chemosensors in aiding
proprioception during flight.

The magnitude of significant marginal variants and epistatic interactions that mapped to
ppk23 suggests this ion transporter has a much more interconnected role in the genetic
architecture of flight performance than previously thought. *ppk23* is a modifier of flies'
ability to track odors during free flight, but not a modifier of odorless flight [140]. Our
results support a role for *ppk23* in modifying flight, along with all but eight (*Or46a*,
Or49a, *Or85d*, *Gr36b*, *Gr36c*, *Ir60d*, *Ir60f*, *ppk10*) of the 41 previously listed pickpocket
and chemoreceptor genes that *ppk23* interacted with. Like *sog*, *ppk23* is likely a central
modifier of performance based on the number of epistatic interactions with variants
mapping to genes identified in the marginal variant screen (*A2bp1/Rbfox1*, *cv-c*, *Fad2*,
CG9313, *CG10936*, *CG42741*, *Rgk1*, *sog*, *Src64B*, *twi*, *Ten-a*). Some of these play

roles in sensory signal processing (*A2bp1/Rbfox1*, *CG9313*, *CG10936*, *Fad2*, *Rgk1*), neuron growth (*sog* and *Src64b*), neuromuscular junction development (*cv-c*, *Src64b*, *Ten-a*), and transcription factors (*A2bp1/Rbfox1*, *CG42741*, *twi*) [141, 142], several of which had significant epistatic interactions of their own. Of these, *CG10936* is proposed to be involved in sensory perception [142], but has limited annotations otherwise. Our work supports this hypothesized function. *ppk23*, in addition to these interactions, is known to modulate physiology and lifespan [143], broadening its canonical roles in chemo- and mechanosensation. Taken together, *ppk23* likely has strong connections to many systems beyond detection of stimulation that have deeper connections to organismal biology.

The interconnectedness of *ppk23* also provides clues about the sexual dimorphism observed in flight. While males generally outperform females, likely due to differences in weight, sex failed to explain ~25% of the variation between the two groups. Like most pickpocket family genes, *ppk23* is well established as an important factor in chemosensation, pheromone detection, and courtship [57, 126, 130]—highly sex-specific phenotypes. One of *ppk23*'s epistatic interactions mapped to *fruitless* (*fru*; human homolog *ZBTB24*), a transcription factor responsible for sex-specific neural phenotypes involved in courtship and pheromone detection [144] that co-localizes with *ppk23* differentially between sexes, on the leg and wing microchaete [4, 57, 126, 127, 143]. In addition to the PNS, *ppk23* and *fru* have sex-specific co-localization patterns in the thoracic ganglion. This cluster of neurons central to the “escape” response, allowing for ultra-fast processing of and response to flight-associated cues [11, 145]. Males show

more connections between *ppk23* and *fru* in the thoracic ganglion, and co-localization in neurons crossing the midline between the two sides of the anterior-most, pro-thoracic ganglion [57, 126]. *fru* is also expressed in vMS2 motor neurons connecting the thoracic ganglion to the flight musculature, likely involved in courtship song generation and aggression behaviors [146, 147]. The connection between sensory neurons, *ppk23*, *fru*, and motor neurons involved in wing motion draw a clear connection between a potential mechanism delineating the sex-difference phenotype we observed. Given the prior connections between *ppk23*, *sog*, and the epistatic interactions between them that annotate to sensory neurons and motor neuron neuromuscular junctions, there are likely other important connections underlying the ability of flies to process proprioceptive signals that are relayed directly to the flight musculature during our assay that have yet to be uncovered. Some of these connections may lie in the genes identified using PEGASUS_flies' for the sex-difference analysis, like *doublesex* (*dsx*), an interactor of *fru* and *ppk1* in patterning sex-specific neural networks for courtship; dissatisfaction (*dsf*), a modifier of courtship behavior [146, 148-150]; and several other genes: *blue cheese* (*bchs*), *Ccn*, *CG13506*, *defective proboscis extension response 6* (*dpr6*), *pollux* (*plx*), *sidekick* (*sdk*), *eiger* (*egr*) [4, 34, 151, 152]. Further study of these genes may yield promising insights into the sex-differences we observed in flight performance, as well as sex-specific behavioral traits.

A proposed model for understanding the genetic architecture of flight performance

Flight performance is likely an epiphenomenon of several interconnected complex traits. While we are unable to identify every modifier, we likely identified the main components of the genetic architecture. Accordingly, we propose the following model to synthesize our findings (Fig 4C).

Epidermal growth factor receptor is a key gene in a canonical developmental pathway. It can affect wing morphology, sensory organ development, and neurodevelopment, on its own and through the BMP signaling pathway. Proper development of these structures and circuits enables well-connected sensory neurons to receive external stimuli regarding proprioception. These signals are transduced through the thoracic ganglion, with sex-specific differences potentially modulated through *ppk23*, *fru*, and *dsx*. The thoracic ganglion processes these signals and activates motor neurons, which innervate the direct (control) and indirect (power) flight musculature at neuron muscular junctions. Activating these muscles allows the properly developed wings to flap and generate lift.

Implications for BMP signaling and pickpocket genes in neuroinjury and neurodegeneration

The complexity of congenital, neurodegenerative diseases lies in the mix of genetic elements with very modest effect size. Association screens with *Drosophila* present a compelling model for identifying these sources of variation, especially in neuron-centric traits [25, 48, 153, 154]. Our results present a strong link between flight performance and BMP signaling—a proposed candidate pathway for therapeutic interventions in

several neurodegenerative diseases [30, 32, 155]. Mutations in *thickveins* (*tkv*) human homologs *BMPR1A* and *BMPR1B* are linked to familial Alzheimer's Disease [156], while mutants of *Superoxide dismutase 1* (*dSOD1*; human homolog *SOD1*) associated with Amyotrophic Lateral Sclerosis (ALS) can be rescued by activators of BMP signaling expressed in proprioceptive and motor neurons [157]. Our validation of the BMP antagonist *Snoo* confirms BMP signaling plays a role in flight performance. Given the number of epistatic interactions between *ppk23* and BMP signaling genes, it is very likely our data uncovers important modifiers of the BMP pathway that affect neurodysfunction in humans.

In addition to BMP signaling, we propose an expanded role for *ppk23*, and pickpocket family genes more generally, in neurobiology and neurodysfunction therapeutics. Acid Sensing Ion Channel (ASIC) family genes, the human homolog of the pickpocket family, can function as neuronal damage sensors. They detect drops in pH around neurons, often caused injury, damage, and dysfunction, which can elicit an inflammation response [31, 158]. These channels are found all over the brain and spinal column, supporting a functional and protective role following traumatic brain injury (concussion) and cerebral ischemia (stroke) [29, 31]. They are also identified as a potential target for genetic and/or pharmacological interventions of neurodegeneration and neuroinflammation [158]. Accordingly, our results break ground in identifying candidate genetic interactions that might be useful for such interventions.

Materials and Methods

Drosophila Stocks and Husbandry

All stocks were obtained from Bloomington *Drosophila* Stock Center (<https://bdsc.indiana.edu/>), including 197 *Drosophila* Genetic Reference Panel (DGRP) lines [20], 23 *Drosophila* Gene Disruption Project lines using the Mi{ET1} construct [159, 160], and two genetic background lines (w^{1118} and y^1w^{67c23} ; S1 Table).

Flies were reared at 25° under a 12-h light-dark cycle. Stocks were density controlled and grown on a standard cornmeal media [161]. Two to three days post-eclosion, flies were sorted by sex under light CO₂ anesthesia and given five days to recover before phenotyping.

Flight performance assay

Flight performance was measured following the protocol refined by Babcock and Ganetzky [33]. Briefly, each sex-genotype combination consisted of approximately 100 flies, divided into groups of 20 flies across five glass vials. These vials were gently tapped to draw flies down, and unplugged before a rapid inversion down a 25 cm chute. Vials stopped at the bottom, ejecting the flies into a 100 cm long x 13.5 cm diameter cylinder lined with a removable acrylic sheet coated in TangleTrap adhesive. Free falling flies instinctively right themselves before finding a place to land, which ended up immobilizing them at their respective landing height. Flies that passed through the column were caught in a pan of mineral oil and were excluded from the analysis.

After all vials in a run were released, the acrylic sheet was removed and pinned to a white poster board. A digital image was recorded on a fixed Raspberry PiCamera (V2) and the x,y coordinates of all flies were located with the ImageJ/FIJI Find Maxima function with a noise tolerance of 30 [162]. For each sex-genotype combination, the mean landing height was calculated for only the flies that landed on the acrylic sheet.

High-speed video capture of flight column

High-speed videos of flies leaving the flight column were recorded at 1540 frames per second using a Phantom Miro m340 camera recording at a resolution of 1920 x 1080 with an exposure of 150 μ s (Data available in File S1). The camera was equipped with a Nikon Micro NIKKOR (105 mm, 1:2.8D) lens and Veritas Constellation 120 light source.

Estimating heritability

Individual fly landing heights were adjusted for covariate status by adding the difference between the DGRP webserver's adjusted and raw line means for each sex, and added them back to the individual landing height of the respective sex and genotype. Using these adjusted landing heights by sex, we performed a random effects analysis of variance using the R (v.3.5.2) package lme4 (v.1.1.23): $Y \sim \mu + L + \varepsilon$. Here, Y is the adjusted flight score, μ is the combined mean, L is the line mean, and ε is the residual. From this, sex-specific broad sense heritability (H^2) estimates were calculated from the among line (σ_L^2) and error (σ_E^2) variance components: $H^2 = \sigma_L^2 / (\sigma_L^2 + \sigma_E^2)$.

Genome wide association mapping

Flight performance scores for males and females were submitted to the DGRP2 GWAS pipeline (<http://dgrp2.gnets.ncsu.edu/>) [19, 20] and results for each sex, and the average (sex-average) and difference (sex-difference) between them were all considered (S3 Table). In total, 1,901,174 variants with a minor allele frequency (MAF) ≥ 0.05 were analyzed (Data available in File S2). All reported additive variant P -values result from a linear mixed model analysis, including *Wolbachia* infection and presence of five major inversions as covariates. Variants were filtered for significance using the conventional $P \leq 1\text{E-}5$ threshold [48]. Effect size estimates were calculated as one-half the difference between the mean landing heights for lines homozygous for the major vs. minor allele. The contribution of individual variants to the overall effects was estimated as the absolute value of an individual variant's effect size divided by the sum of the absolute values for all conventionally significant ($P < 1\text{e-}5$) variants' effect sizes.

Candidate gene disruption screen

Candidate genes were validated using insertional mutant stocks generated from Gene Disruption Project [51]. These stocks contain a *Minos* enhancer trap construct Mi{ET1}[50] and were built on either w^{1118} or $y^1 w^{67c23}$ backgrounds (BDSC_6326 and BDSC_6599, respectively).

Control and experiment line genetic backgrounds were isogenized with five successive rounds of backcrossing the insertional mutant line to its respective control. Validation of flight phenotypes was done using offspring of single-pair (1M x 1F) crosses between the control and insert lines. Heterozygous flies from these crosses were mated in pairs and

the homozygous offspring lacking the insertion were collected as the control. Candidate heterozygous/homozygous positive lines were mated as pairs once more and lines producing only homozygous positive offspring were used as experimental lines (S1 Fig). Experimental lines were checked for a GFP reporter three generations later to confirm their genotype. The finalized recombinant backcrossed control and experimental lines for each sex-genotype combination were assayed for flight performance, and tested for significance, via Mann-Whitney U-tests.

Calculating gene-score significance

Gene-scores were calculated using Precise, Efficient Gene Association Score Using SNPs (PEGASUS) [28]. Originally implemented with human datasets, we modified the program to work with *Drosophila* datasets, which we call PEGASUS_flies. It also contains default values adjusted for *Drosophila*, a linkage disequilibrium file, and gene annotations drawn from the FB5.57 annotation file, available on the DGRP webserver. PEGASUS_flies is available at: https://github.com/ramachandran-lab/PEGASUS_flies, and as File S4.

Identifying altered sub-networks of gene-gene and protein-protein interaction networks

Returned gene-scores were filtered for genes of high confidence using the Twilight package (v.1.60.0) in R (Scheid and Spang 2005). Here, we estimated the local False Discovery Rate (lFDR) of all previously output gene scores using the twilight function. Taking the inflection point of the (1 – lFDR) curve, our high-confidence gene scores

ranged from 0.65 – 0.73 for the four, sex-based phenotypes (S8 Table). High confidence genes were $-\log_{10}$ transformed, while the remaining were set to 0.

Hierarchical HotNet was used to identify altered sub-networks of interacting genes or proteins [52] based on network topology generated from several gene-gene or protein-protein interaction networks. The four adjusted, sex-based gene-score vectors were mapped in the program to fifteen interaction networks obtained from High-quality INTERactomes (HINT)[163], the *Drosophila* Interactions Database (Droidb)[164, 165], and the *Drosophila* RNAⁱ Screening Center (DRSC) Integrative Ortholog Prediction Tool (DIOPT)[166]. Consensus networks were calculated from 100 permutations of all four gene-score vectors on each of the fifteen interaction networks and filtered to include at least three members. The largest sub-network and the remaining eight sub-networks were passed to Gene Ontology enRichment anaLysis and visualizAtion tool (GORilla) to identify enrichment for gene ontology (GO) categories [167, 168].

Screening for epistatic interactions

Epistatic hub genes were identified using MArginal ePlstasis Test (MAPIT), a linear mixed modeling approach that tests the significance of each SNP's marginal effect on a chosen phenotype. MAPIT requires a complete genotype matrix, without missing data. SNPs were imputed using BEAGLE 4.1 [169, 170] and then filtered for $MAF \geq 0.05$ using VCFtools (v.0.1.16) [171]. MAPIT was run using the Davies method on the imputed genome (File S2), DGRP2 webserver-adjusted phenotype scores for each sex-

based phenotype (S2 Table), DGRP2 relatedness matrix, and covariate file containing *Wolbachia* infection and the presence of five major inversions.

Resulting marginal effect P -values (data available File S3) were filtered to a Bonferroni threshold ($P \leq 2.56e-8$) and tested for pairwise epistatic interactions in a set-by-all framework against the initial 1,901,174 SNPs (unimputed; $MAF \geq 0.05$) using the `PLINK` `-epistasis` flag (v.1.90)[172]. Results were filtered for all P -values that exceeded a Bonferroni threshold, calculated as $0.05 / (\text{the number of Bonferroni marginal effect } P\text{-values} \times 1,901,174 \text{ SNPs})$.

Annotating FBgn and orthologs

Flybase gene (FBgn) identifiers were converted to their respective *D. melanogaster* (Dmel) or *H. sapiens* (Hsap) gene symbols using the *Drosophila* RNAⁱ Stock Center (DRSC) Integrative Ortholog Prediction Tool (DIOPT)[42]. FBgn were filtered for all high to moderate confidence genes, or low confidence genes if they contained the best forward and reverse score.

Calculating an empirically simulated significance threshold

We sought to simulate an empirically derived significance threshold that was unique to our data set and separate from the traditional DGRP and Bonferroni thresholds used in other studies. Using the genotype-phenotype matrix, two separate datasets were simulated ($n = 1000$) for each sex-based phenotype. The first randomized the genotype-

phenotype matrix using all available line means, while the second randomized subsets of 150 genotype-phenotype pairs.

Simulated associations were run with `PLINK` [172](v.1.90) on each dataset type for each sex-based phenotype. The 5th percentile most-significant *P*-value across all permutations in a simulation type was deemed the “empirically simulated significance threshold.”

GO term analysis

GOWINDA [49] was implemented to perform a Gene Ontology (GO) analysis that corrects for gene size in GWA studies. We conducted this analysis for male (n=418), female (n=473), sex-average (n=527), and sex-difference (n=214) candidate SNPs exceeding a relaxed $P < 1E-4$ significance threshold, against the 1,901,174 SNPs with $MAF \geq 0.05$. We ran 100,000 simulations of GOWINDA using the gene mode and including all SNPs within 2000 bp.

Gene Ontology enRichment anaLysis and visualiZAtion tool (GORilla)[167, 168] was run on `PEGASUS_flies` gene-scores and `Hierarchical Hotnet` sub-networks using the default commands and a gene list compiled from all genes available in the FB5.57 annotation file.

Weighted Gene Co-expression Network Analysis

To test whether ambient adult transcriptomes could explain the observed phenotypic variation, we turned to the publically available DGRP2 microarray data, downloaded from the DGRP2 webserver [20]. These data represent the transcriptomes for untreated young adult flies of each sex. We performed Weighted Gene Co-expression Network Analysis (WGCNA) analyses using the available R package [63] to cluster and correlate the expression profiles of genes from 177 shared, DGRP lines. This analysis was run using the following parameters: power = 16 (from soft threshold analysis ≥ 0.9), merging threshold = 0.0, signed network type, maximum blocksize = 1000, minimum module size = 30.

Data availability

All data required to rerun the outlined analyses either publically available through FlyBase (<http://flybase.org/>) [34, 35, 39], the DGRP2 webserver (<http://dgrp2.gnets.ncsu.edu/>), or in the Supplement: S1-S18 tables, <https://doi.org/10.26300/v4rm-sa82>; S1 file, <https://doi.org/10.26300/dwvm-vt70>; S2 file, <https://doi.org/10.26300/317y-p682>; S3 file, <https://doi.org/10.26300/xcrh-c744>; S4 file, <https://doi.org/10.26300/qhc7-dp70>.

Acknowledgements

We thank Priya Nakka and Matthew A Reyna for their guidance through PEGASUS and Hierarchical HotNet, respectively. David Boerma for his assistance recording high-speed footage and F. Lemieux for maintaining *Drosophila* stocks. We are grateful for critical feedback from M. Tatar and T. Roberts and our anonymous reviewers.

Conflict of Interest statement

The authors declare no conflict of interest.

Funding

This research was funded by National Institutes of Health grant GM067862 (to DMR).

Author contributions

ANS, JAM, and DMR conceived and designed the study. ANS performed validation crosses, while ANS and JAM collected data. ANS performed the statistical analyses guidance from SR and LC. SPS and ANS designed and implemented PEGASUS_flies. ANS and DMR wrote the manuscript.

Supplemental Results

Establishing an empirically defined significance threshold

While the Bonferroni significance threshold is conservative, the conventional $P = 1\text{E-}5$ threshold might be considered lax. Accordingly, we simulated two sets of genotype-phenotype matrices; one set “shuffled” the genotype-phenotype matrix while the other set randomly “subsampled” 150 of the 197 lines.

The significance threshold for each sex-based phenotype in each simulation was determined by taking the 5th percentile of the most significant P -value across 1000 permutations [173]. Despite these efforts, the resulting significance thresholds were even more stringent than the Bonferroni (S18 Table) and resulted in only one variant (2R_2718036_DEL) mapping to *CG15236* and *CG34215* in the shuffled sex-difference set. *CG15236*’s function is not well known, but it is expressed during embryogenesis and pupariation in the developing brain and central nervous system and putatively affects the wing veins [174, 175]. *CG34215* is less understood, though it is expressed at varying levels throughout developmental and adult stages [34] and contains a single domain Von Willebrand factor type C domain—thought to play a role in anti-viral capabilities [176].

References

1. Brodsky AK. The evolution of insect flight: Oxford University Press; 1994.
2. Edwards JS. The evolution of insect flight: Implications for the evolution of the nervous system. *Brain Behavior and Evolution*. 1997;50(1):8-12. doi: 10.1159/000113317. PubMed PMID: WOS:A1997XE45100002.
3. Ugur B, Chen KC, Bellen HJ. Drosophila tools and assays for the study of human diseases. *Disease Models & Mechanisms*. 2016;9(3):235-44. doi: 10.1242/dmm.023762. PubMed PMID: WOS:000371439600003.
4. Pavlou HJ, Goodwin SF. Courtship behavior in *Drosophila melanogaster*: towards a 'courtship connectome'. *Current Opinion in Neurobiology*. 2013;23(1):76-83. doi: 10.1016/j.conb.2012.09.002. PubMed PMID: WOS:000314562900013.
5. Weitkunat M, Schnorrer F. A guide to study *Drosophila* muscle biology. *Methods*. 2014;68(1):2-14. doi: <https://doi.org/10.1016/j.ymeth.2014.02.037>.
6. Quijano JC, Stinchfield MJ, Zerlanko B, Gibbens YY, Takaesu NT, Hyman-Walsh C, et al. The Sno Oncogene Antagonizes Wingless Signaling during Wing Development in *Drosophila*. *Plos One*. 2010;5(7). doi: 10.1371/journal.pone.0011619. PubMed PMID: WOS:000279980800008.
7. Nongthomba U, Pasalodos-Sanchez S, Clark S, Clayton JD, Sparrow JC. Expression and function of the *Drosophila* ACT88F actin isoform is not restricted to the indirect flight muscles. *Journal of muscle research and cell motility*. 2001;22(2):111-9.
8. Montooth KL, Marden JH, Clark AG. Mapping determinants of variation in energy metabolism, respiration and flight in *Drosophila*. *Genetics*. 2003;165(2):623-35.
9. Kao SY, Nikonova E, Ravichandran K, Spletter ML. Dissection of *Drosophila melanogaster* Flight Muscles for Omics Approaches. *Jove-Journal of Visualized Experiments*. 2019;(152). doi: 10.3791/60309. PubMed PMID: WOS:000493379500080.
10. Frye MA, Dickinson MH. Closing the loop between neurobiology and flight behavior in *Drosophila*. *Current Opinion in Neurobiology*. 2004;14(6):729-36. doi: 10.1016/j.conb.2004.10.004. PubMed PMID: WOS:000225988600011.
11. Lehmann FO, Bartussek J. Neural control and precision of flight muscle activation in *Drosophila*. *Journal of Comparative Physiology a-Neuroethology Sensory Neural and Behavioral Physiology*. 2017;203(1):1-14. doi: 10.1007/s00359-016-1133-9. PubMed PMID: WOS:000393670200001.
12. Okada H, Ebhardt HA, Vonesch SC, Aebersold R, Hafen E. Proteome-wide association studies identify biochemical modules associated with a wing-size phenotype in *Drosophila melanogaster*. *Nature Communications*. 2016;7. doi: 10.1038/ncomms12649. PubMed PMID: WOS:000384969200001.

- 1067 13. Marcus JM. The development and evolution of crossveins in insect wings.
1068 Journal of Anatomy. 2001;199:211-6. doi: 10.1017/s0021878201008226.
1069 PubMed PMID: WOS:000170738500022.
- 1070 14. Lobell AS, Kaspari RR, Negron YLS, Harbison ST. The Genetic Architecture of
1071 Ovariole Number in *Drosophila melanogaster*: Genes with Major, Quantitative,
1072 and Pleiotropic Effects. *G3-Genes Genomes Genetics*. 2017;7(7):2391-403. doi:
1073 10.1534/g3.117.042390. PubMed PMID: WOS:000404991600035.
- 1074 15. Watanabe K, Stringer S, Frei O, Mirkov MU, de Leeuw C, Polderman TJC, et al.
1075 A global overview of pleiotropy and genetic architecture in complex traits. *Nature*
1076 *Genetics*. 2019;51(9):1339-+. doi: 10.1038/s41588-019-0481-0. PubMed PMID:
1077 WOS:000484010800010.
- 1078 16. McClellan J, King MC. Genetic Heterogeneity in Human Disease. *Cell*.
1079 2010;141(2):210-7. doi: 10.1016/j.cell.2010.03.032. PubMed PMID:
1080 WOS:000276738400008.
- 1081 17. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, et al.
1082 Finding the missing heritability of complex diseases. *Nature*.
1083 2009;461(7265):747-53. doi: 10.1038/nature08494. PubMed PMID:
1084 WOS:000270547500027.
- 1085 18. McCarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little J, Ioannidis JPA, et
1086 al. Genome-wide association studies for complex traits: consensus, uncertainty
1087 and challenges. *Nature Reviews Genetics*. 2008;9(5):356-69. doi:
1088 10.1038/nrg2344. PubMed PMID: WOS:000255057300012.
- 1089 19. Mackay TFC, Richards S, Stone EA, Barbadilla A, Ayroles JF, Zhu DH, et al. The
1090 *Drosophila melanogaster* Genetic Reference Panel. *Nature*.
1091 2012;482(7384):173-8. doi: 10.1038/nature10811. PubMed PMID:
1092 WOS:000299994100029.
- 1093 20. Huang W, Massouras A, Inoue Y, Peiffer J, Ramia M, Tarone AM, et al. Natural
1094 variation in genome architecture among 205 *Drosophila melanogaster* Genetic
1095 Reference Panel lines. *Genome Research*. 2014;24(7):1193-208. doi:
1096 10.1101/gr.171546.113. PubMed PMID: WOS:000338185000013.
- 1097 21. Arya GH, Magwire MM, Huang W, Serrano-Negron YL, Mackay TFC, Anholt
1098 RRH. The Genetic Basis for Variation in Olfactory Behavior in *Drosophila*
1099 *melanogaster*. *Chemical Senses*. 2015;40(4):233-43. doi:
1100 10.1093/chemse/bjv001. PubMed PMID: WOS:000355703500004.
- 1101 22. Chow CY, Wolfner MF, Clark AG. Large Neurological Component to Genetic
1102 Differences Underlying Biased Sperm Use in *Drosophila*. *Genetics*.
1103 2013;193(1):177-85. doi: 10.1534/genetics.112.146357. PubMed PMID:
1104 WOS:000312894700011.
- 1105 23. Battlay P, Leblanc PB, Green L, Garud NR, Schmidt JM, Fournier-Level A, et al.
1106 Structural Variants and Selective Sweep Foci Contribute to Insecticide

- Resistance in the Drosophila Genetic Reference Panel. *G3-Genes Genomes Genetics*. 2018;8(11):3489-97. doi: 10.1534/g3.118.200619. PubMed PMID: WOS:000449381500009.
24. Chow CY, Kelsey KJP, Wolfner MF, Clark AG. Candidate genetic modifiers of retinitis pigmentosa identified by exploiting natural variation in *Drosophila*. *Human Molecular Genetics*. 2016;25(4):651-9. doi: 10.1093/hmg/ddv502. PubMed PMID: WOS:000372151000003.
25. Carbone MA, Yamamoto A, Huang W, Lyman RA, Meadors TB, Yamamoto R, et al. Genetic architecture of natural variation in visual senescence in *Drosophila*. *Proceedings of the National Academy of Sciences of the United States of America*. 2016;113(43):E6620-E9. doi: 10.1073/pnas.1613833113. PubMed PMID: WOS:000386087100012.
26. Zhou SS, Morozova TV, Hussain YN, Luoma SE, McCoy L, Yamamoto A, et al. The Genetic Basis for Variation in Sensitivity to Lead Toxicity in *Drosophila melanogaster*. *Environmental Health Perspectives*. 2016;124(7):1062-70. doi: 10.1289/ehp.1510513. PubMed PMID: WOS:000380749300029.
27. Montgomery SL, Vorojeikina D, Huang W, Mackay TFC, Anholt RRH, Rand MD. Genome-Wide Association Analysis of Tolerance to Methylmercury Toxicity in *Drosophila* Implicates Myogenic and Neuromuscular Developmental Pathways. *Plos One*. 2014;9(10). doi: 10.1371/journal.pone.0110375. PubMed PMID: WOS:000343943700033.
28. Nakka P, Raphael BJ, Ramachandran S. Gene and Network Analysis of Common Variants Reveals Novel Associations in Multiple Complex Diseases. *Genetics*. 2016;204(2):783-+. doi: 10.1534/genetics.116.188391. PubMed PMID: WOS:000385871400031.
29. Xiong ZG, Xu TL. The role of ASICs in cerebral ischemia. *Wiley Interdisciplinary Reviews: Membrane Transport and Signaling*. 2012;1(5):655-62.
30. Deshpande M, Feiger Z, Shilton AK, Luo CC, Silverman E, Rodal AA. Role of BMP receptor traffic in synaptic growth defects in an ALS model. *Molecular Biology of the Cell*. 2016;27(19):2898-910. doi: 10.1091/mbc.E16-07-0519. PubMed PMID: WOS:000385183000003.
31. Huang Y, Jiang N, Li J, Ji YH, Xiong ZG, Zha XM. Two aspects of ASIC function: Synaptic plasticity and neuronal injury. *Neuropharmacology*. 2015;94:42-8. doi: 10.1016/j.neuropharm.2014.12.010. PubMed PMID: WOS:000356735000007.
32. Pinto C, Cardenas P, Osses N, Henriquez JP. Characterization of Wnt/beta-catenin and BMP/Smad signaling pathways in an in vitro model of amyotrophic lateral sclerosis. *Frontiers in Cellular Neuroscience*. 2013;7. doi: 10.3389/fncel.2013.00239. PubMed PMID: WOS:000327856200001.
33. Babcock DT, Ganetzky B. An Improved Method for Accurate and Rapid Measurement of Flight Performance in *Drosophila*. *Jove-Journal of Visualized*

1147 Experiments. 2014;(84). doi: 10.3791/51223. PubMed PMID:
1148 WOS:000348604100055.

1149 34. dos Santos G, Schroeder AJ, Goodman JL, Strelets VB, Crosby MA, Thurmond
1150 J, et al. FlyBase: introduction of the *Drosophila melanogaster* Release 6
1151 reference genome assembly and large-scale migration of genome annotations.
1152 Nucleic Acids Research. 2015;43(D1):D690-D7. doi: 10.1093/nar/gku1099.
1153 PubMed PMID: WOS:000350210400101.

1154 35. Grumblin G, Strelets V, FlyBase C. FlyBase: anatomical data, images and
1155 queries. Nucleic Acids Research. 2006;34:D484-D8. doi: 10.1093/nar/gkj068.
1156 PubMed PMID: WOS:000239307700106.

1157 36. Carbon S, Douglass E, Dunn N, Good B, Harris NL, Lewis SE, et al. The Gene
1158 Ontology Resource: 20 years and still GOing strong. Nucleic Acids Research.
1159 2019;47(D1):D330-D8. doi: 10.1093/nar/gky1055. PubMed PMID:
1160 WOS:000462587400049.

1161 37. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene
1162 Ontology: tool for the unification of biology. Nature Genetics. 2000;25(1):25-9.
1163 doi: 10.1038/75556. PubMed PMID: WOS:000086884000011.

1164 38. Frise E, Hammonds AS, Celniker SE. Systematic image-driven analysis of the
1165 spatial *Drosophila* embryonic expression landscape. Molecular Systems Biology.
1166 2010;6. doi: 10.1038/msb.2009.102. PubMed PMID: WOS:000274193500002.

1167 39. Chintapalli VR, Wang J, Dow JAT. Using FlyAtlas to identify better *Drosophila*
1168 *melanogaster* models of human disease. Nature Genetics. 2007;39(6):715-20.
1169 doi: 10.1038/ng2049. PubMed PMID: WOS:000246859100012.

1170 40. Consortium TAoGR. Alliance of Genome Resources Portal: unified model
1171 organism research platform. Nucleic acids research. 2020;48(D1):D650-D8.

1172 41. Gaudet P, Livstone MS, Lewis SE, Thomas PD. Phylogenetic-based propagation
1173 of functional annotations within the Gene Ontology consortium. Briefings in
1174 Bioinformatics. 2011;12(5):449-62. doi: 10.1093/bib/bbr042. PubMed PMID:
1175 WOS:000295171700010.

1176 42. Hu YH, Flockhart I, Vinayagam A, Bergwitz C, Berger B, Perrimon N, et al. An
1177 integrative approach to ortholog prediction for disease-focused and other
1178 functional studies. BMC Bioinformatics. 2011;12. doi: 10.1186/1471-2105-12-357.
1179 PubMed PMID: WOS:000295225400001.

1180 43. Mummery-Widmer JL, Yamazaki M, Stoeger T, Novatchkova M, Bhalerao S,
1181 Chen D, et al. Genome-wide analysis of Notch signalling in *Drosophila* by
1182 transgenic RNAi. Nature. 2009;458(7241):987-U59. doi: 10.1038/nature07936.
1183 PubMed PMID: WOS:000265412900033.

1184 44. Karr TL. Fruit flies and the sperm proteome. Human Molecular Genetics.
1185 2007;16:R124-R33. doi: 10.1093/hmg/ddm252. PubMed PMID:
1186 WOS:000251165500002.

45. Casas-Vila N, Bluhm A, Sayols S, Dinges N, Dejung M, Altenhein T, et al. The developmental proteome of *Drosophila melanogaster*. *Genome Research*. 2017;27(7):1273-85. doi: 10.1101/gr.213694.116. PubMed PMID: WOS:000404735500016.
46. Brown JB, Boley N, Eisman R, May GE, Stoiber MH, Duff MO, et al. Diversity and dynamics of the *Drosophila* transcriptome. *Nature*. 2014;512(7515):393-9. doi: 10.1038/nature12962. PubMed PMID: WOS:000340840600025.
47. Negre N, Brown CD, Ma LJ, Bristow CA, Miller SW, Wagner U, et al. A cis-regulatory map of the *Drosophila* genome. *Nature*. 2011;471(7339):527-31. doi: 10.1038/nature09990. PubMed PMID: WOS:000288702200065.
48. Mackay TF, Huang W. Charting the genotype–phenotype map: lessons from the *Drosophila melanogaster* Genetic Reference Panel. *Wiley Interdisciplinary Reviews: Developmental Biology*. 2018;7(1).
49. Kofler R, Schlötterer C. Gowinda: unbiased analysis of gene set enrichment for genome-wide association studies. *Bioinformatics*. 2012;28(15):2084-5.
50. Metaxakis A, Oehler S, Klinakis A, Savakis C. Minos as a genetic and genomic tool in *Drosophila melanogaster*. *Genetics*. 2005;171(2):571-81. doi: 10.1534/genetics.105.041848. PubMed PMID: WOS:000233194500014.
51. Bellen HJ, Levis RW, He YC, Carlson JW, Evans-Holm M, Bae E, et al. The *Drosophila* Gene Disruption Project: Progress Using Transposons With Distinctive Site Specificities. *Genetics*. 2011;188(3):731-U341. doi: 10.1534/genetics.111.126995. PubMed PMID: WOS:000292538900022.
52. Reyna MA, Leiserson MDM, Raphael BJ. Hierarchical HotNet: identifying hierarchies of altered subnetworks. *Bioinformatics*. 2018;34(17):972-80. doi: 10.1093/bioinformatics/bty613. PubMed PMID: WOS:000444317200045.
53. Huang W, Richards S, Carbone MA, Zhu DH, Anholt RRH, Ayroles JF, et al. Epistasis dominates the genetic architecture of *Drosophila* quantitative traits. *Proceedings of the National Academy of Sciences of the United States of America*. 2012;109(39):15553-9. doi: 10.1073/pnas.1213423109. PubMed PMID: WOS:000309604500014.
54. Liu XY, Li YI, Pritchard JK. Trans Effects on Gene Expression Can Drive Omnigenic Inheritance. *Cell*. 2019;177(4):1022-+. doi: 10.1016/j.cell.2019.04.014. PubMed PMID: WOS:000466843000020.
55. Boyle EA, Li YI, Pritchard JK. An Expanded View of Complex Traits: From Polygenic to Omnigenic. *Cell*. 2017;169(7):1177-86. doi: 10.1016/j.cell.2017.05.038. PubMed PMID: WOS:000403332400008.
56. Adams CM, Anderson MG, Motto DG, Price MP, Johnson WA, Welsh MJ. Ripped pocket and pickpocket, novel *Drosophila* DEG/ENaC subunits expressed in early development and in mechanosensory neurons. *Journal of Cell Biology*.

1226 1998;140(1):143-52. doi: 10.1083/jcb.140.1.143. PubMed PMID:
1227 WOS:000071500600014.

1228 57. Lu BK, LaMora A, Sun YS, Welsh MJ, Ben-Shahar Y. ppk23-Dependent
1229 Chemosensory Functions Contribute to Courtship Behavior in *Drosophila*
1230 *melanogaster*. *Plos Genetics*. 2012;8(3). doi: 10.1371/journal.pgen.1002587.
1231 PubMed PMID: WOS:000302254800062.

1232 58. Neely GG, Hess A, Costigan M, Keene AC, Goulas S, Langeslag M, et al. A
1233 Genome-wide *Drosophila* Screen for Heat Nociception Identifies alpha 2 delta 3
1234 as an Evolutionarily Conserved Pain Gene. *Cell*. 2010;143(4):628-38. doi:
1235 10.1016/j.cell.2010.09.047. PubMed PMID: WOS:000284149100019.

1236 59. zur Lage P, Newton FG, Jarman AP. Survey of the Ciliary Motility Machinery of
1237 *Drosophila* Sperm and Ciliated Mechanosensory Neurons Reveals Unexpected
1238 Cell-Type Specific Variations: A Model for Motile Ciliopathies. *Frontiers in*
1239 *Genetics*. 2019;10. doi: 10.3389/fgene.2019.00024. PubMed PMID:
1240 WOS:000457405100001.

1241 60. Brockington M, Yuva Y, Prandini P, Brown SC, Torelli S, Benson MA, et al.
1242 Mutations in the fukutin-related protein gene (FKRP) identify limb girdle muscular
1243 dystrophy 2I as a milder allelic variant of congenital muscular dystrophy MDC1C.
1244 *Human Molecular Genetics*. 2001;10(25):2851-9. doi: 10.1093/hmg/10.25.2851.
1245 PubMed PMID: WOS:000172868200001.

1246 61. Inlow JK, Restifo LL. Molecular and comparative genetics of mental retardation.
1247 *Genetics*. 2004;166(2):835-81. doi: 10.1534/genetics.166.2.835. PubMed PMID:
1248 WOS:000220390600020.

1249 62. Vannoy CH, Xiao W, Lu PJ, Xiao X, Lu QL. Efficacy of Gene Therapy Is
1250 Dependent on Disease Progression in Dystrophic Mice with Mutations in the
1251 FKRP Gene. *Molecular Therapy-Methods & Clinical Development*. 2017;5:31-42.
1252 doi: 10.1016/j.omtm.2017.02.002. PubMed PMID: WOS:000406299600004.

1253 63. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation
1254 network analysis. *BMC Bioinformatics*. 2008;9(1):559. doi: 10.1186/1471-2105-9-
1255 559.

1256 64. Huang W, Carbone MA, Magwire MM, Peiffer JA, Lyman RF, Stone EA, et al.
1257 Genetic basis of transcriptome diversity in *Drosophila melanogaster*.
1258 *Proceedings of the National Academy of Sciences*. 2015;112(44):E6010-E9.

1259 65. Cheatle Jarvela AM, Hinman VF. Evolution of transcription factor function as a
1260 mechanism for changing metazoan developmental gene regulatory networks.
1261 *EvoDevo*. 2015;6(1):3. doi: 10.1186/2041-9139-6-3.

1262 66. Oleksiak MF, Churchill GA, Crawford DL. Variation in gene expression within and
1263 among natural populations. *Nature genetics*. 2002;32(2):261.

67. Wittkopp PJ, Haerum BK, Clark AG. Evolutionary changes in cis and trans gene regulation. *Nature*. 2004;430(6995):85-8. doi: 10.1038/nature02698. PubMed PMID: WOS:000222356800048.
68. Crossman SH, Streichan SJ, Vincent JP. EGFR signaling coordinates patterning with cell survival during *Drosophila* epidermal development. *Plos Biology*. 2018;16(10). doi: 10.1371/journal.pbio.3000027. PubMed PMID: WOS:000449322300040.
69. Sibilila M, Kroismayr R, Lichtenberger BM, Natarajan A, Hecking M, Holcman M. The epidermal growth factor receptor: from development to tumorigenesis. *Differentiation*. 2007;75(9):770-87. doi: 10.1111/j.1432-0436.2007.00238.x. PubMed PMID: WOS:000250816400002.
70. Paul L, Wang SH, Manivannan SN, Bonanno L, Lewis S, Austin CL, et al. Dpp-induced Egfr signaling triggers postembryonic wing development in *Drosophila*. *Proceedings of the National Academy of Sciences of the United States of America*. 2013;110(13):5058-63. doi: 10.1073/pnas.1217538110. PubMed PMID: WOS:000318031900048.
71. Hacker U, Grossniklaus U, Gehring WJ, Jackle H. DEVELOPMENTALLY REGULATED DROSOPHILA GENE FAMILY ENCODING THE FORK HEAD DOMAIN. *Proceedings of the National Academy of Sciences of the United States of America*. 1992;89(18):8754-8. doi: 10.1073/pnas.89.18.8754. PubMed PMID: WOS:A1992JN50300071.
72. Albert FW, Kruglyak L. The role of regulatory variation in complex traits and disease. *Nature Reviews Genetics*. 2015;16(4):197-212. doi: 10.1038/nrg3891. PubMed PMID: WOS:000352121800007.
73. Haigh SE, Salvi SS, Sevdali M, Stark M, Goulding D, Clayton JD, et al. *Drosophila* indirect flight muscle specific Act88F actin mutants as a model system for studying congenital myopathies of the human ACTA1 skeletal muscle actin gene. *Neuromuscular Disorders*. 2010;20(6):363-74. doi: 10.1016/j.nmd.2010.03.008. PubMed PMID: WOS:000279099100001.
74. Drummond DR, Hennessey ES, Sparrow JC. CHARACTERIZATION OF MISSENSE MUTATIONS IN THE ACT88F GENE OF DROSOPHILA-MELANOGASTER. *Molecular & General Genetics*. 1991;226(1-2):70-80. doi: 10.1007/bf00273589. PubMed PMID: WOS:A1991FK53800010.
75. Maughan DW, Vigoreaux IO. An integrated view of insect flight muscle: Genes, motor molecules, and motion. *News in Physiological Sciences*. 1999;14:87-92. PubMed PMID: WOS:000081066000001.
76. Ennos AR. COMPARATIVE FUNCTIONAL-MORPHOLOGY OF THE WINGS OF DIPTERA. *Zoological Journal of the Linnean Society*. 1989;96(1):27-47. doi: 10.1111/j.1096-3642.1989.tb01820.x. PubMed PMID: WOS:A1989U946600003.

77. Pitchers W, Nye J, Marquez EJ, Kowalski A, Dworkin I, Houle D. A Multivariate Genome-Wide Association Study of Wing Shape in *Drosophila melanogaster*. *Genetics*. 2019;211(4):1429-47. doi: 10.1534/genetics.118.301342. PubMed PMID: WOS:000463935700021.
78. Hevia CF, Lopez-Varea A, Esteban N, de Celis JF. A Search for Genes Mediating the Growth-Promoting Function of TGF beta in the *Drosophila melanogaster* Wing Disc. *Genetics*. 2017;206(1):231-49. doi: 10.1534/genetics.116.197228. PubMed PMID: WOS:000401127800016.
79. Cruz C, Glavic A, Casado M, de Celis JF. A Gain-of-Function Screen Identifying Genes Required for Growth and Pattern Formation of the *Drosophila melanogaster* Wing. *Genetics*. 2009;183(3):1005-26. doi: 10.1534/genetics.109.107748. PubMed PMID: WOS:000272295800021.
80. Yu K, Sturtevant MA, Biehs B, Francois V, Padgett RW, Blackman RK, et al. The *Drosophila* decapentaplegic and short gastrulation genes function antagonistically during adult wing vein development. *Development*. 1996;122(12):4033-44. PubMed PMID: WOS:A1996WC55400034.
81. Ball RW, Warren-Paquin M, Tsurudome K, Liao EH, Elazzouzi F, Cavanagh C, et al. Retrograde BMP Signaling Controls Synaptic Growth at the NMJ by Regulating Trio Expression in Motor Neurons. *Neuron*. 2010;66(4):536-49. doi: 10.1016/j.neuron.2010.04.011. PubMed PMID: WOS:000278771100009.
82. O'Connor MB, Umulis D, Othmer HG, Blair SS. Shaping BMP morphogen gradients in the *Drosophila* embryo and pupal wing. *Development*. 2006;133(2):183-93. PubMed PMID: WOS:000235345700001.
83. Wharton KA, Serpe M. Fine-tuned shuttles for bone morphogenetic proteins. *Current Opinion in Genetics & Development*. 2013;23(4):374-84. doi: 10.1016/j.gde.2013.04.012. PubMed PMID: WOS:000324362900003.
84. Crawford L, Zeng P, Mukherjee S, Zhou X. Detecting epistasis with the marginal epistasis test in genetic mapping studies of quantitative traits. *Plos Genetics*. 2017;13(7). doi: 10.1371/journal.pgen.1006869. PubMed PMID: WOS:000406615300011.
85. Lehmann FO, Dickinson MH. The changes in power requirements and muscle efficiency during elevated force production in the fruit fly *Drosophila melanogaster*. *Journal of Experimental Biology*. 1997;200(7):1133-43. PubMed PMID: WOS:A1997WW57600007.
86. Dickinson MH, Tu MS. The function of dipteran flight muscle. *Comparative Biochemistry and Physiology a-Physiology*. 1997;116(3):223-38. doi: 10.1016/s0300-9629(96)00162-4. PubMed PMID: WOS:A1997WJ67300005.
87. Kozopas KM, Nusse R. Direct flight muscles in *Drosophila* develop from cells with characteristics of founders and depend on DWnt-2 for their correct

- patterning. *Developmental Biology*. 2002;243(2):312-25. doi: 10.1006/dbio.2002.0572. PubMed PMID: WOS:000174465500010.
88. Fernandes I, Schock F. The nebulin repeat protein Lasp regulates I-band architecture and filament spacing in myofibrils. *Journal of Cell Biology*. 2014;206(4):559-72. doi: 10.1083/jcb.201401094. PubMed PMID: WOS:000340739500009.
89. Spletter ML, Barz C, Yeroslaviz A, Schonbauer C, Ferreira IRS, Sarov M, et al. The RNA-binding protein Arrest (Bruno) regulates alternative splicing to enable myofibril maturation in *Drosophila* flight muscle. *Embo Reports*. 2015;16(2):178-91. doi: 10.15252/embr.201439791. PubMed PMID: WOS:000350695600009.
90. Mardahl-Dumesnil M, Fowler VM. Thin filaments elongate from their pointed ends during myofibril assembly in *Drosophila* indirect flight muscle. *Journal of Cell Biology*. 2001;155(6):1043-53. doi: 10.1083/jcb.200108026. PubMed PMID: WOS:000172730200017.
91. Fuerst PG, Burgess RW. Adhesion molecules in establishing retinal circuitry. *Current Opinion in Neurobiology*. 2009;19(4):389-94. doi: 10.1016/j.conb.2009.07.013. PubMed PMID: WOS:000270107600007.
92. Yamagata M, Sanes JR. Dscam and Sidekick proteins direct lamina-specific synaptic connections in vertebrate retina. *Nature*. 2008;451(7177):465-U6. doi: 10.1038/nature06469. PubMed PMID: WOS:000252554100045.
93. Hewes RS, Taghert PH. Neuropeptides and neuropeptide receptors in the *Drosophila melanogaster* genome. *Genome research*. 2001;11(6):1126-42.
94. Ueda A, Wu CF. Effects of Social Isolation on Neuromuscular Excitability and Aggressive Behaviors in *Drosophila*: Altered Responses by Hk and gsts1, Two Mutations Implicated in Redox Regulation. *Journal of Neurogenetics*. 2009;23(4):378-94. doi: 10.3109/01677060903063026. PubMed PMID: WOS:000271470500003.
95. Allan DW, St Pierre SE, Miguel-Aliaga I, Thor S. Specification of neuropeptide cell identity by the integration of retrograde BMP signaling and a combinatorial transcription factor code. *Cell*. 2003;113(1):73-86. doi: 10.1016/s0092-8674(03)00204-6. PubMed PMID: WOS:000182282900008.
96. Fontaine B, Sassoon D, Buckingham M, Changeux J-P. Detection of the nicotinic acetylcholine receptor alpha - subunit mRNA by in situ hybridization at neuromuscular junctions of 15 - day - old chick striated muscles. *The EMBO Journal*. 1988;7(3):603-9.
97. Nguyen MU, Kwong J, Chang J, Gillet VG, Lee RM, Johnson KG. The extracellular and cytoplasmic domains of syndecan cooperate postsynaptically to promote synapse growth at the *Drosophila* neuromuscular junction. *PloS one*. 2016;11(3).

98. Zhan XL, Clemens JC, Neves G, Hattori D, Flanagan JJ, Hummel T, et al. Analysis of Dscam diversity in regulating axon guidance in *Drosophila* mushroom bodies. *Neuron*. 2004;43(5):673-86. doi: 10.1016/j.neuron.2004.07.020. PubMed PMID: WOS:000223691900013.
99. Neves G, Zucker J, Daly M, Chess A. Stochastic yet biased expression of multiple Dscam splice variants by individual cells. *Nature Genetics*. 2004;36(3):240-6. doi: 10.1038/ng1299. PubMed PMID: WOS:000189250400012.
100. Stocker RF. THE ORGANIZATION OF THE CHEMOSENSORY SYSTEM IN *DROSOPHILA-MELANOGASTER* - A REVIEW. *Cell and Tissue Research*. 1994;275(1):3-26. doi: 10.1007/bf00305372. PubMed PMID: WOS:A1994MM22200001.
101. Smith SA, Shepherd D. Central afferent projections of proprioceptive sensory neurons in *Drosophila* revealed with the enhancer-trap technique. *Journal of Comparative Neurology*. 1996;364(2):311-23. PubMed PMID: WOS:A1996TQ97200009.
102. Tadros W, Xu SW, Akin O, Yi CH, Shin GJE, Millard SS, et al. Dscam Proteins Direct Dendritic Targeting through Adhesion. *Neuron*. 2016;89(3):480-93. doi: 10.1016/j.neuron.2015.12.026. PubMed PMID: WOS:000373564900009.
103. Hummel T, Zipursky L. Afferent induction of olfactory glomeruli requires N-cadherin. *Neuron*. 2004;42(1):77-88. doi: 10.1016/s0896-6273(04)00158-8. PubMed PMID: WOS:000221458400009.
104. Hummel T, Vasconcelos ML, Clemens JC, Fishilevich Y, Vosshall LB, Zipursky SL. Axonal targeting of olfactory receptor neurons in *Drosophila* is controlled by Dscam. *Neuron*. 2003;37(2):221-31. doi: 10.1016/s0896-6273(02)01183-2. PubMed PMID: WOS:000181054900008.
105. Soba P, Zhu S, Emoto K, Younger S, Yang SJ, Yu HH, et al. *Drosophila* sensory neurons require Dscam for dendritic self-avoidance and proper dendritic field organization. *Neuron*. 2007;54(3):403-16. doi: 10.1016/j.neuron.2007.03.029. PubMed PMID: WOS:000246855700011.
106. Emoto K, He Y, Ye B, Grueber WB, Adler PN, Jan LY, et al. Control of dendritic branching and tiling by the tricorned-kinase/furry signaling pathway in *Drosophila* sensory neurons. *Cell*. 2004;119(2):245-56. doi: 10.1016/j.cell.2004.09.036. PubMed PMID: WOS:000224577200012.
107. Matsubara D, Horiuchi SY, Shimono K, Usui T, Uemura T. The seven-pass transmembrane cadherin Flamingo controls dendritic self-avoidance via its binding to a LIM domain protein, Espinas, in *Drosophila* sensory neurons. *Genes & Development*. 2011;25(18):1982-96. doi: 10.1101/gad.16531611. PubMed PMID: WOS:000295082400009.

- 1420 108. Cong JL, Geng W, He B, Liu JC, Charlton J, Adler PN. The furry gene of
1421 Drosophila is important for maintaining the integrity of cellular extensions during
1422 morphogenesis. *Development*. 2001;128(14):2793-802. PubMed PMID:
1423 WOS:000170209900015.
- 1424 109. Fischer S, Bayersdorfer F, Harant E, Reng R, Arndt S, Bosserhoff AK, et al.
1425 fussel (fuss) - A Negative Regulator of BMP Signaling in Drosophila
1426 melanogaster. *Plos One*. 2012;7(8). doi: 10.1371/journal.pone.0042349. PubMed
1427 PMID: WOS:000307437900031.
- 1428 110. Takaesu NT, Hyman-Walsh C, Ye Y, Wisotzkey RG, Stinchfield MJ, O'Connor
1429 MB, et al. dSno facilitates Baboon signaling in the Drosophila brain by switching
1430 the affinity of Medea away from Mad and toward dSmad2. *Genetics*.
1431 2006;174(3):1299-313.
- 1432 111. Luo K. Signaling cross talk between TGF- β /Smad and other signaling pathways.
1433 *Cold Spring Harbor perspectives in biology*. 2017;9(1):a022137.
- 1434 112. Furman D, Bukharina T. How Drosophila melanogaster forms its
1435 mechanoreceptors. *Current genomics*. 2008;9(5):312-23.
- 1436 113. Ainsley JA, Pettus JM, Bosenko D, Gerstein CE, Zinkevich N, Anderson MG, et
1437 al. Enhanced locomotion caused by loss of the Drosophila DEG/ENaC protein
1438 pickpocket1. *Current Biology*. 2003;13(17):1557-63. doi: 10.1016/s0960-
1439 9822(03)00596-7. PubMed PMID: WOS:000185171300028.
- 1440 114. Yamashita S, Takigahira T, Takahashi KH. Genome-wide association analysis of
1441 host genotype and plastic wing morphological variation of an endoparasitoid
1442 wasp *Asobara japonica* (Hymenoptera: Braconidae). *Genetica*. 2018;146(3):313-
1443 21. doi: 10.1007/s10709-018-0022-2. PubMed PMID: WOS:000434148400006.
- 1444 115. Dickinson MH. COMPARISON OF ENCODING PROPERTIES OF
1445 CAMPANIFORM SENSILLA ON THE FLY WING. *Journal of Experimental*
1446 *Biology*. 1990;151:245-61. PubMed PMID: WOS:A1990DN79800014.
- 1447 116. Dickinson MH, Hannaford S, Palka J. The evolution of insect wings and their
1448 sensory apparatus. *Brain Behavior and Evolution*. 1997;50(1):13-24. doi:
1449 10.1159/000113318. PubMed PMID: WOS:A1997XE45100003.
- 1450 117. Firth LC, Baker NE. Spitz from the retina regulates genes transcribed in the
1451 second mitotic wave, peripodial epithelium, glia and plasmatocytes of the
1452 Drosophila eye imaginal disc. *Developmental Biology*. 2007;307(2):521-38. doi:
1453 10.1016/j.ydbio.2007.04.037. PubMed PMID: WOS:000248019100027.
- 1454 118. Chintapalli VR, Wang J, Dow JA. Using FlyAtlas to identify better Drosophila
1455 melanogaster models of human disease. *Nature genetics*. 2007;39(6):715-20.
- 1456 119. Harbison ST, Kumar S, Huang W, McCoy LJ, Smith KR, Mackay TFC. Genome-
1457 Wide Association Study of Circadian Behavior in Drosophila melanogaster.
1458 *Behavior Genetics*. 2019;49(1):60-82. doi: 10.1007/s10519-018-9932-0. PubMed
1459 PMID: WOS:000455331800005.

120. Dierick HA, Greenspan RJ. Molecular analysis of flies selected for aggressive behavior. *Nature Genetics*. 2006;38(9):1023-31. doi: 10.1038/ng1864. PubMed PMID: WOS:000240112100019.
121. Page RM, Munch A, Horn T, Kuhn PH, Colombo A, Reiner O, et al. Loss of PAFAH1B2 Reduces Amyloid-beta Generation by Promoting the Degradation of Amyloid Precursor Protein C-Terminal Fragments. *Journal of Neuroscience*. 2012;32(50):18204-14. doi: 10.1523/jneurosci.2681-12.2012. PubMed PMID: WOS:000312404700026.
122. Durham MF, Magwire MM, Stone EA, Leips J. Genome-wide analysis in *Drosophila* reveals age-specific effects of SNPs on fitness traits. *Nature Communications*. 2014;5. doi: 10.1038/ncomms5338. PubMed PMID: WOS:000340615500029.
123. Neumuller RA, Richter C, Fischer A, Novatchkova M, Neumuller KG, Knoblich JA. Genome-Wide Analysis of Self-Renewal in *Drosophila* Neural Stem Cells by Transgenic RNAi. *Cell Stem Cell*. 2011;8(5):580-93. doi: 10.1016/j.stem.2011.02.022. PubMed PMID: WOS:000290927600017.
124. Zheng Z, Zheng H, Yan W. Fank1 is a testis-specific gene encoding a nuclear protein exclusively expressed during the transition from the meiotic to the haploid phase of spermatogenesis. *Gene Expression Patterns*. 2007;7(7):777-83. doi: 10.1016/j.modgep.2007.05.005. PubMed PMID: WOS:000249331000008.
125. DeGiorgio M, Lohmueller KE, Nielsen R. A Model-Based Approach for Identifying Signatures of Ancient Balancing Selection in Genetic Data. *Plos Genetics*. 2014;10(8). doi: 10.1371/journal.pgen.1004561. PubMed PMID: WOS:000341577800047.
126. Thistle R, Cameron P, Ghorayshi A, Dennison L, Scott K. Contact chemoreceptors mediate male-male repulsion and male-female attraction during *Drosophila* courtship. *Cell*. 2012;149(5):1140-51.
127. Ben-Shahar Y. Sensory Functions for Degenerin/Epithelial Sodium Channels (DEG/ENaC). In: Friedmann T, Dunlap JC, Goodwin SF, editors. *Advances in Genetics*, Vol 76. *Advances in Genetics*. 762011. p. 1-26.
128. Liu L, Johnson WA, Welsh MJ. *Drosophila* DEG/ENaC pickpocket genes are expressed in the tracheal system, where they may be involved in liquid clearance. *Proceedings of the National Academy of Sciences of the United States of America*. 2003;100(4):2128-33. doi: 10.1073/pnas.252785099. PubMed PMID: WOS:000181073000122.
129. Ng R, Salem SS, Wu ST, Wu ML, Lin HH, Shepherd AK, et al. Amplification of *Drosophila* Olfactory Responses by a DEG/ENaC Channel. *Neuron*. 2019;104(5):947-+. doi: 10.1016/j.neuron.2019.08.041. PubMed PMID: WOS:000500919100012.

- 1499 130. Gorczyca DA, Younger S, Meltzer S, Kim SE, Cheng L, Song W, et al.
1500 Identification of Ppk26, a DEG/ENaC Channel Functioning with Ppk1 in a
1501 Mutually Dependent Manner to Guide Locomotion Behavior in *Drosophila*. *Cell*
1502 *Reports*. 2014;9(4):1446-58. doi: 10.1016/j.celrep.2014.10.034. PubMed PMID:
1503 WOS:000345529600025.
- 1504 131. Kuo CT, Jan LY, Jan YN. Dendrite-specific remodeling of *Drosophila* sensory
1505 neurons requires matrix metalloproteases, ubiquitin-proteasome, and ecdysone
1506 signaling. *Proceedings of the National Academy of Sciences of the United States*
1507 *of America*. 2005;102(42):15230-5. doi: 10.1073/pnas.0507393102. PubMed
1508 PMID: WOS:000232811800053.
- 1509 132. Grueber WB, Ye B, Moore AW, Jan LY, Jan YN. Dendrites of distinct classes of
1510 *Drosophila* sensory neurons show different capacities for homotypic repulsion.
1511 *Current Biology*. 2003;13(8):618-26. doi: 10.1016/s0960-9822(03)00207-0.
1512 PubMed PMID: WOS:000182490000017.
- 1513 133. Mauthner SE, Hwang RY, Lewis AH, Xiao Q, Tsubouchi A, Wang Y, et al. Balboa
1514 Binds to Pickpocket In Vivo and Is Required for Mechanical Nociception in
1515 *Drosophila* Larvae. *Current Biology*. 2014;24(24). doi:
1516 10.1016/j.cub.2014.10.038. PubMed PMID: WOS:000346580700022.
- 1517 134. Zelle KM, Lu BK, Pyfrom SC, Ben-Shahar Y. The Genetic Architecture of
1518 Degenerin/Epithelial Sodium Channels in *Drosophila*. *G3-Genes Genomes*
1519 *Genetics*. 2013;3(3):441-50. doi: 10.1534/g3.112.005272. PubMed PMID:
1520 WOS:000315950000006.
- 1521 135. Paukert M, Sidi S, Russell C, Siba M, Wilson SW, Nicolson T, et al. A family of
1522 acid-sensing ion channels from the zebrafish - Widespread expression in the
1523 central nervous system suggests a conserved role in neuronal communication.
1524 *Journal of Biological Chemistry*. 2004;279(18):18783-91. doi:
1525 10.1074/jbc.M401477200. PubMed PMID: WOS:000221041500086.
- 1526 136. Jeong YT, Oh SM, Shim J, Seo JT, Kwon JY, Moon SJ. Mechanosensory
1527 neurons control sweet sensing in *Drosophila*. *Nature Communications*. 2016;7.
1528 doi: 10.1038/ncomms12872. PubMed PMID: WOS:000385384100006.
- 1529 137. Frye MA, Dickinson MH. Motor output reflects the linear superposition of visual
1530 and olfactory inputs in *Drosophila*. *Journal of Experimental Biology*.
1531 2004;207(1):123-31. doi: 10.1242/jeb.00725. PubMed PMID:
1532 WOS:000189011000022.
- 1533 138. Taylor GK, Krapp HG. Sensory systems and flight stability: what do insects
1534 measure and why? *Advances in insect physiology*. 2007;34:231-316.
- 1535 139. Sherman A, Dickinson MH. Summation of visual and mechanosensory feedback
1536 in *Drosophila* flight control. *Journal of Experimental Biology*. 2004;207(1):133-42.
1537 doi: 10.1242/jeb.00731. PubMed PMID: WOS:000189011000023.

140. Houot B, Gigot V, Robichon A, Ferveur JF. Free flight odor tracking in *Drosophila*: Effect of wing chemosensors, sex and pheromonal gene regulation. *Scientific Reports*. 2017;7. doi: 10.1038/srep40221. PubMed PMID: WOS:000391403200001.
141. Shukla JP, Deshpande G, Shashidhara LS. Ataxin 2-binding protein 1 is a context- specific positive regulator of Notch signaling during neurogenesis in *Drosophila melanogaster*. *Development*. 2017;144(5):905-15. doi: 10.1242/dev.140657. PubMed PMID: WOS:000395650100018.
142. Jin M, Aibar S, Ge ZQ, Chen R, Aerts S, Mardon G. Identification of novel direct targets of *Drosophila* *Sine oculis* and *Eyes absent* by integration of genome-wide data sets. *Developmental Biology*. 2016;415(1):157-67. doi: 10.1016/j.ydbio.2016.05.007. PubMed PMID: WOS:000378187700014.
143. Gendron CM, Kuo TH, Harvanek ZM, Chung BY, Yew JY, Dierick HA, et al. *Drosophila* Life Span and Physiology Are Modulated by Sexual Perception and Reward. *Science*. 2014;343(6170):544-8. doi: 10.1126/science.1243339. PubMed PMID: WOS:000330343700048.
144. Kimura KI, Ote M, Tazawa T, Yamamoto D. Fruitless specifies sexually dimorphic neural circuitry in the *Drosophila* brain. *Nature*. 2005;438(7065):229-33. doi: 10.1038/nature04229. PubMed PMID: WOS:000233133500050.
145. Strausfeld NJ. Brain and optic lobes. *Encyclopedia of Insects*: Elsevier; 2009. p. 121-30.
146. Yu JY, Kanai MI, Demir E, Jefferis G, Dickson BJ. Cellular Organization of the Neural Circuit that Drives *Drosophila* Courtship Behavior. *Current Biology*. 2010;20(18):1602-14. doi: 10.1016/j.cub.2010.08.025. PubMed PMID: WOS:000282385600020.
147. Ewing AW. NEUROMUSCULAR BASIS OF COURTSHIP SONG IN *DROSOPHILA* - ROLE OF THE DIRECT AND AXILLARY WING MUSCLES. *Journal of Comparative Physiology*. 1979;130(1):87-93. doi: 10.1007/bf02582977. PubMed PMID: WOS:A1979GS00900010.
148. Shirangi TR, Wong AM, Truman JW, Stern DL. Doublesex Regulates the Connectivity of a Neural Circuit Controlling *Drosophila* Male Courtship Song. *Developmental Cell*. 2016;37(6):533-44. doi: 10.1016/j.devcel.2016.05.012. PubMed PMID: WOS:000378204200008.
149. Rezaval C, Pavlou HJ, Dornan AJ, Chan YB, Kravitz EA, Goodwin SF. Neural Circuitry Underlying *Drosophila* Female Postmating Behavioral Responses. *Current Biology*. 2012;22(13):1155-65. doi: 10.1016/j.cub.2012.04.062. PubMed PMID: WOS:000306379600018.
150. Finley KD, Taylor BJ, Milstein M, McKeown M. dissatisfaction, a gene involved in sex-specific behavior and neural development of *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences*. 1997;94(3):913-8.

151. Billeter JC, Vilella A, Allendorfer JB, Dornan AJ, Richardson M, Gailey DA, et al. Isoform-specific control of male neuronal differentiation and behavior in *Drosophila* by the fruitless gene. *Current Biology*. 2006;16(11):1063-76. doi: 10.1016/j.cub.2006.04.039. PubMed PMID: WOS:000238245900019.
152. Finley KD, Taylor BJ, Milstein M, McKeown M. dissatisfaction, a gene involved in sex-specific behavior and neural development of *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences of the United States of America*. 1997;94(3):913-8. doi: 10.1073/pnas.94.3.913. PubMed PMID: WOS:A1997WG23400026.
153. Chow CY, Reiter LT. Etiology of Human Genetic Disease on the Fly. *Trends in Genetics*. 2017.
154. Lavoy S, Chittoor-Vinod VG, Chow CY, Martin I. Genetic Modifiers of Neurodegeneration in a *Drosophila* Model of Parkinson's Disease. *Genetics*. 2018;209(4):1345-56. doi: 10.1534/genetics.118.301119. PubMed PMID: WOS:000440014100024.
155. Bayat V, Jaiswal M, Bellen HJ. The BMP signaling pathway at the *Drosophila* neuromuscular junction and its links to neurodegenerative diseases. *Current Opinion in Neurobiology*. 2011;21(1):182-8. doi: 10.1016/j.conb.2010.08.014. PubMed PMID: WOS:000288876100024.
156. Kang MJ, Hansen TJ, Mickiewicz M, Kaczynski TJ, Fye S, Gunawardena S. Disruption of Axonal Transport Perturbs Bone Morphogenetic Protein (BMP) - Signaling and Contributes to Synaptic Abnormalities in Two Neurodegenerative Diseases. *Plos One*. 2014;9(8). doi: 10.1371/journal.pone.0104617. PubMed PMID: WOS:000340879300038.
157. Held A, Major P, Sabin A, Reenan RA, Lipscombe D, Wharton KA. Circuit Dysfunction in SOD1-ALS Model First Detected in Sensory Feedback Prior to Motor Neuron Degeneration Is Alleviated by BMP Signaling. *Journal of Neuroscience*. 2019;39(12):2347-64. doi: 10.1523/jneurosci.1771-18.2019. PubMed PMID: WOS:000461756300015.
158. Ortega-Ramirez A, Vega R, Soto E. Acid-Sensing Ion Channels as Potential Therapeutic Targets in Neurodegeneration and Neuroinflammation. *Mediators of Inflammation*. 2017. doi: 10.1155/2017/3728096. PubMed PMID: WOS:000411707000001.
159. Bellen HJ, Levis RW, He Y, Carlson JW, Evans-Holm M, Bae E, et al. The *drosophila* gene disruption project: progress using transposons with distinctive site-specificities. *Genetics*. 2011;genetics. 111.126995.
160. Metaxakis A, Oehler S, Klinakis A, Savakis C. Minos as a genetic and genomic tool in *Drosophila melanogaster*. *Genetics*. 2005.

161. Elgin S, Miller D. Special techniques for tissue isolation and injection. II. Mass rearing of flies and mass production and harvesting of embryos. *Genetics and biology of Drosophila*. 1978.
162. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an open-source platform for biological-image analysis. *Nature Methods*. 2012;9(7):676-82. doi: 10.1038/nmeth.2019. PubMed PMID: WOS:000305942200021.
163. Das J, Yu HY. HINT: High-quality protein interactomes and their applications in understanding human disease. *Bmc Systems Biology*. 2012;6. doi: 10.1186/1752-0509-6-92. PubMed PMID: WOS:000310442400001.
164. Yu JK, Pacifico S, Liu GZ, Finley RL. DroiD: the Drosophila Interactions Database, a comprehensive resource for annotated gene and protein interactions. *Bmc Genomics*. 2008;9. doi: 10.1186/1471-2164-9-461. PubMed PMID: WOS:000260678600001.
165. Murali T, Pacifico S, Yu JK, Guest S, Roberts GG, Finley RL. DroiD 2011: a comprehensive, integrated resource for protein, transcription factor, RNA and gene interactions for Drosophila. *Nucleic Acids Research*. 2011;39:D736-D43. doi: 10.1093/nar/gkq1092. PubMed PMID: WOS:000285831700117.
166. Hu Y, Flockhart I, Vinayagam A, Bergwitz C, Berger B, Perrimon N, et al. An integrative approach to ortholog prediction for disease-focused and other functional studies. *BMC bioinformatics*. 2011;12(1):357.
167. Eden E, Navon R, Steinfeld I, Lipson D, Yakhini Z. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *Bmc Bioinformatics*. 2009;10. doi: 10.1186/1471-2105-10-48. PubMed PMID: WOS:000264007400001.
168. Eden E, Lipson D, Yogev S, Yakhini Z. Discovering motifs in ranked lists of DNA sequences. *Plos Computational Biology*. 2007;3(3):508-22. doi: 10.1371/journal.pcbi.0030039. PubMed PMID: WOS:000246191000017.
169. Browning SR, Browning BL. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *American Journal of Human Genetics*. 2007;81(5):1084-97. doi: 10.1086/521987. PubMed PMID: WOS:000250480900018.
170. Browning BL, Browning SR. Genotype Imputation with Millions of Reference Samples. *American Journal of Human Genetics*. 2016;98(1):116-26. doi: 10.1016/j.ajhg.2015.11.020. PubMed PMID: WOS:000368050800009.
171. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al. The variant call format and VCFtools. *Bioinformatics*. 2011;27(15):2156-8. doi: 10.1093/bioinformatics/btr330. PubMed PMID: WOS:000292778700023.
172. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: A tool set for whole-genome association and population-based linkage

- analyses. American Journal of Human Genetics. 2007;81(3):559-75. doi: 10.1086/519795. PubMed PMID: WOS:000249128200012.
173. Doerge RW, Churchill GA. Permutation Tests for Multiple Loci Affecting a Quantitative Character. Genetics. 1996;142(1):285-94.
174. Krupp JJ, Yaich LE, Wessells RJ, Bodmer R. Identification of genetic loci that interact with cut during Drosophila wing-margin development. Genetics. 2005;170(4):1775-95. doi: 10.1534/genetics.105.043125. PubMed PMID: WOS:000232033300028.
175. Vonesch SC, Lamparter D, Mackay TFC, Bergmann S, Hafen E. Genome-Wide Analysis Reveals Novel Regulators of Growth in Drosophila melanogaster. PLOS Genetics. 2016;12(1):e1005616. doi: 10.1371/journal.pgen.1005616.
176. Chen YH, Jia XT, Zhao L, Li CZ, Zhang SA, Chen YG, et al. Identification and functional characterization of Dicer2 and five single VWC domain proteins of Litopenaeus vannamei. Developmental and Comparative Immunology. 2011;35(6):661-71. doi: 10.1016/j.dci.2011.01.010. PubMed PMID: WOS:000289333600005.

Supporting information

S1 Fig. DGRP lines' mean flight performance is highly repeatable across

generations. Set of genotypes ($n = 12$) reared 10 generations apart show very strong agreement ($r = 0.95$) in mean flight performance scores. The regression line (red line) through the point pairs (black points) has nearly the same slope and y-intercept as the $x = y$ line (gray dashed line).

S2 Fig. Sex-average and sex-difference phenotypic distributions are amenable to

an association study. Distribution in mean landing height (m) for (A) sex-average and (B) sex-difference phenotypes suggest ample phenotypic variation exists to run an association study. Each plot is sorted in order of increasing phenotype score, independent of one another.

S3 Fig. QQ-plots show enrichment for some additive variants across each of the

sex-based phenotypes. Plots comparing the theoretical vs. observed P -value distribution across (A) males, (B) females, (C) sex-average, and (D) sex-difference phenotypes. Red line denotes $y = x$.

S4 Fig. Top additive associations are spaced throughout the genome. Top additive

variants, those reported in DGRP2 webserver file with the `top.annot` suffix, are largely free of linkage blocks. There is a larger block on X, corresponding with 10 variants that map to intronic and one synonymous coding site in *CG32506*. The heat component corresponds with likelihood of being in a linkage block from less (0 - blue) to more likely (1 - red).

S5 Fig. Additional sex-based phenotype Manhattan plots for additive analysis. (A)

Males, (B) females, and (C) sex-difference phenotypes all have significant

additive variants pass a traditional DGRP threshold ($P \leq 1E-5$, gray solid line, red points), and at least one variant pass a Bonferroni threshold ($P \leq 2.63E-8$, gray dashed line, red dot with black outline). Variants are arranged in order of relative genomic position by chromosome and plotted by the $-\log_{10}$ of the P -value. The sex-average is displayed in text.

S6 Fig. Genetic crosses performed for deriving experimental and control stocks

used to validate candidate genes. All crosses are represented with females on the left and males on the right. Ten single pair crosses of a female genetic control, either w^{1118} (pictured) or $y[1] w[67c23]$, in white boxes were crossed with the respective $M\{ET1\}$ insertional mutant line in green boxes. After the initial cross, heterozygous flies were backcrossed to the respective genetic control for five generations. In the sixth generation, single pairs of heterozygous flies were crossed. Progeny without the $Avic\backslash GFPE.3xP3$ marker were collected as homozygous nulls, while several vials of putatively homozygous mutants (no progeny without marker) were crossed again to confirm genotype. Stocks were monitored for two additional generations to confirm mutant carrier status before a homozygous mutant stock was selected as an experimental line.

S7 Fig. Significant whole genes are distributed throughout the genome and sex-

based phenotypes. Whole gene analyses conducted with `PEGASUS_flies` for (A) males, (B) females, and (C) sex-difference phenotypes showed enrichment for significant whole genes across these three, and the sex-average (displayed in text). Each dot represents a whole gene, ordered by position across the

chromosomes and plotted as the $-\log_{10}$ of the gene-score. Points above the Bonferroni threshold ($P \leq 3.03E-6$, gray line) are colored in red.

S8 Fig. Significant marginal variants are unevenly distributed across sex-based

phenotypes. (A) Males had very few significant variants pass a Bonferroni threshold ($P \leq 2.56E-8$, gray solid line, red points), while (B) females had more and (C) sex-average had the most. (D) Sex-difference had no significant marginal variants. Variants are arranged in order of relative genomic position by chromosome and significance scores $-\log_{10}$ transformed.

S9 Fig. Trait-relationship correlation matrix shows no correlation between

measured phenotypes and young adult transcriptome. Neither sexes' mean landing height, standard deviation in landing height, or proportion of flies that fell through the column (fallen) were significant with a cluster of similarly expressed genes in a Weighted Gene Co-expression Network Analysis (WGCNA). Colored modules on the left represent WGCNA-generated clusters of genes and the color of each table cell corresponds with the magnitude of correlation coefficient (top number in cell). The bottom number in each cell is the significance of the correlation. No clusters were significantly correlated with any sex-phenotype combination.

S1 Table. *Drosophila* stocks used in this study.

S2 Table. Raw and adjusted flight performance metrics.

1743 **S3 Table.** No significant correlations were observed between flight performance and
1744 other DGRP phenotypes.

1745 **S4 Table.** Up to two inversions were significant covariates in three of the sex-based
1746 analyses.

1747 **S5 Table.** Several additive variants associated with flight performance.

1748 **S6 Table.** Several candidate genes were validated for flight performance.

1749 **S7 Table.** Several gene-scores pass a Bonferroni threshold across all four sex-based
1750 phenotypes.

1751 **S8 Table.** Twilight-estimated local False Discovery Rate (lFDR) cutoff thresholds for
1752 PEGASUS_fly gene-scores.

1753 **S9 Table.** Hierarchical HotNet sub-network gene annotations.

1754 **S10 Table.** Large sub-network from Hierarchical HotNet is enriched for trans-regulatory
1755 factors and neurodevelopmental Gene Ontology terms.

1756 **S11 Table.** Collection of smaller sub-networks from Hierarchical HotNet are collectively
1757 enriched for mRNA splicing and autophagy Gene Ontology terms.

1758 **S12 Table.** Significant marginal variants identified from MAginal ePIstasis Test
1759 (MAPIT).

1760 **S13 Table.** Epistatic interactions play a large role in shaping the genetic architecture of
1761 flight performance.

1762 **S14 Table.** Epistatic interactions with *pickpocket 23* (*ppk23*) are enriched in a Gene
1763 Ontology (GO) term analysis from neurodevelopmental genes.

1764 **S15 Table.** Genes mapped to from epistatic interactions with *CG42671* are significantly
1765 enriched for neurodevelopment in a Gene Ontology (GO) analysis.

1766 **S16 Table.** Gene set enrichment analysis for significant epistatic interactors within a
 1767 669 bp intergenic region between chr3L:6890373 - 6891042 suggests
 1768 enrichment for neurodevelopmental Gene Ontology categories.

1769 **S17 Table.** Master lookup table for all genes identified.

1770 **S18 Table.** Empirically derived P-values from simulated permutations of the genotype-
 1771 phenotype matrix.

1772

1773 **S1 File.** High-speed gifs of strong, intermediate, and weak genotypes in flight
 1774 performance assay; <https://doi.org/10.26300/dwvm-vt70>

1775 **S2 File.** BEAGLE-imputed DGRP2 genome; <https://doi.org/10.26300/317y-p682>

1776 **S3 File.** All marginal *P*-value significance scores calculated using MAPIT;
 1777 <https://doi.org/10.26300/xcrh-c744>

1778 **S4 File.** Source code for PEGASUS_flies; <https://doi.org/10.26300/qhc7-dp70> -- Also
 1779 available at: https://github.com/ramachandran-lab/PEGASUS_flies

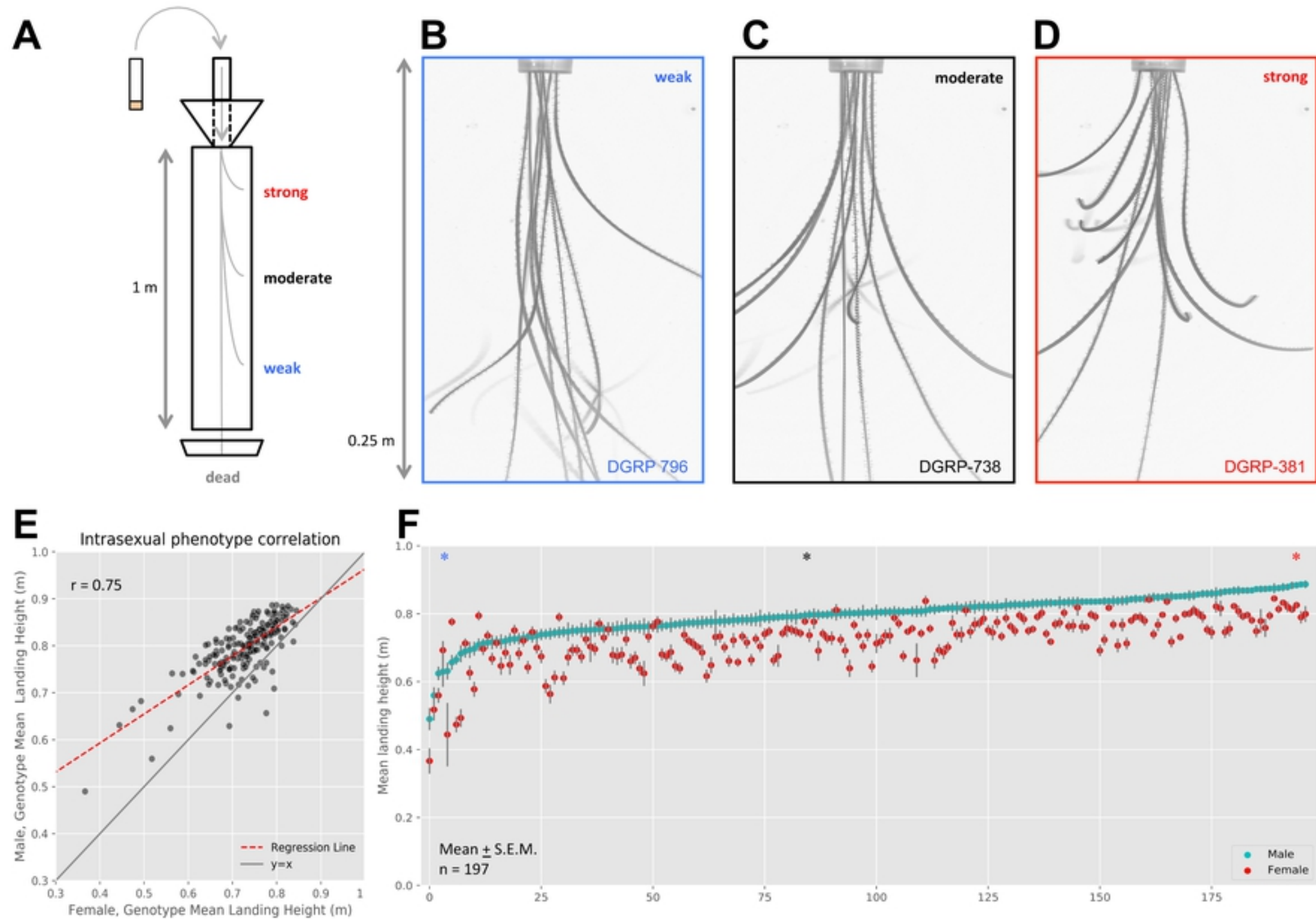
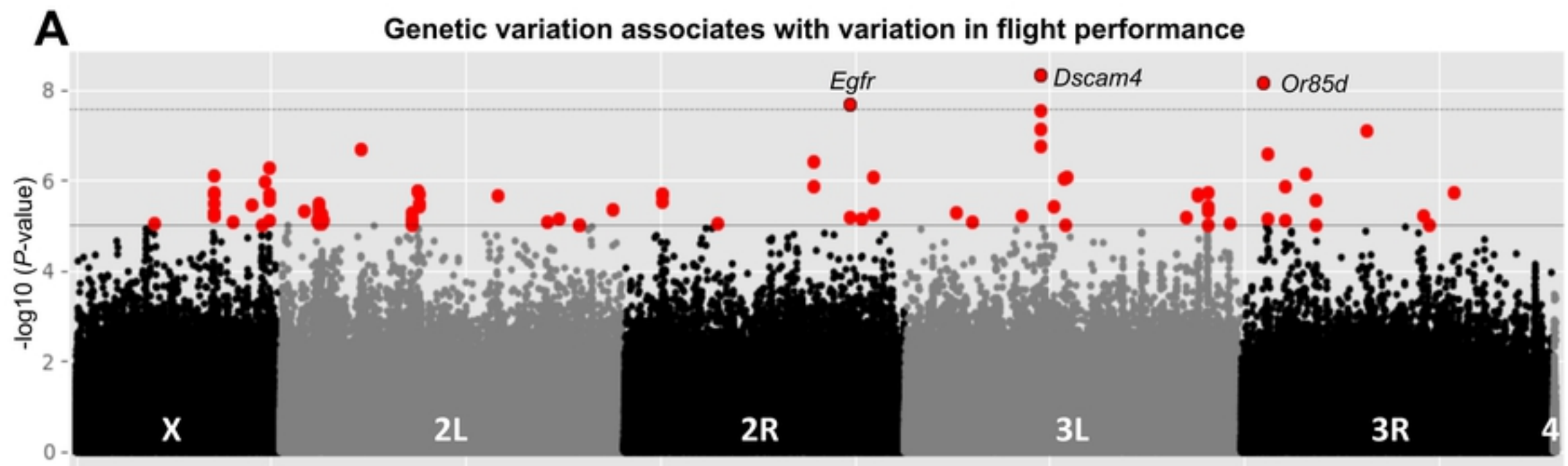


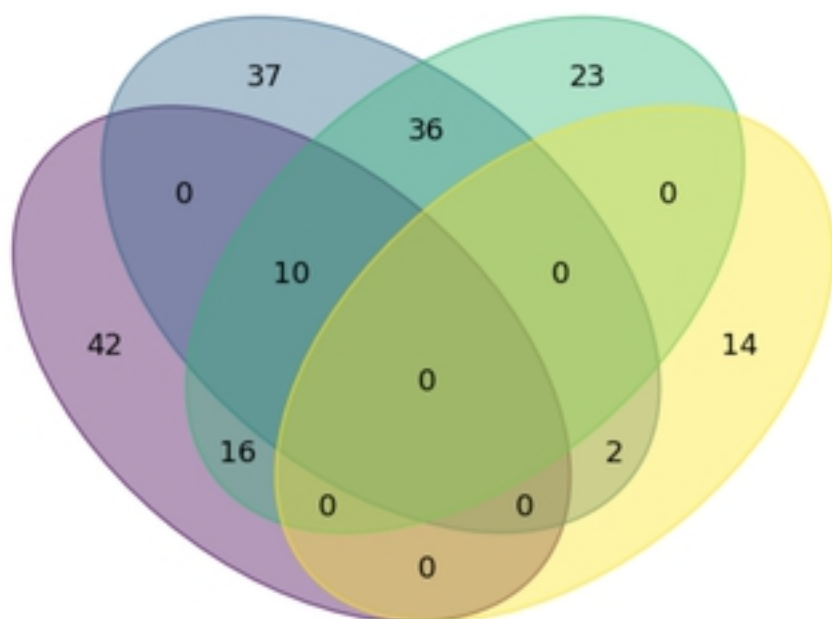
Figure 1



B

Four-way Venn Diagram of Additive genes

Male
Female
Average
Difference



C

Confirmation of several candidate genes

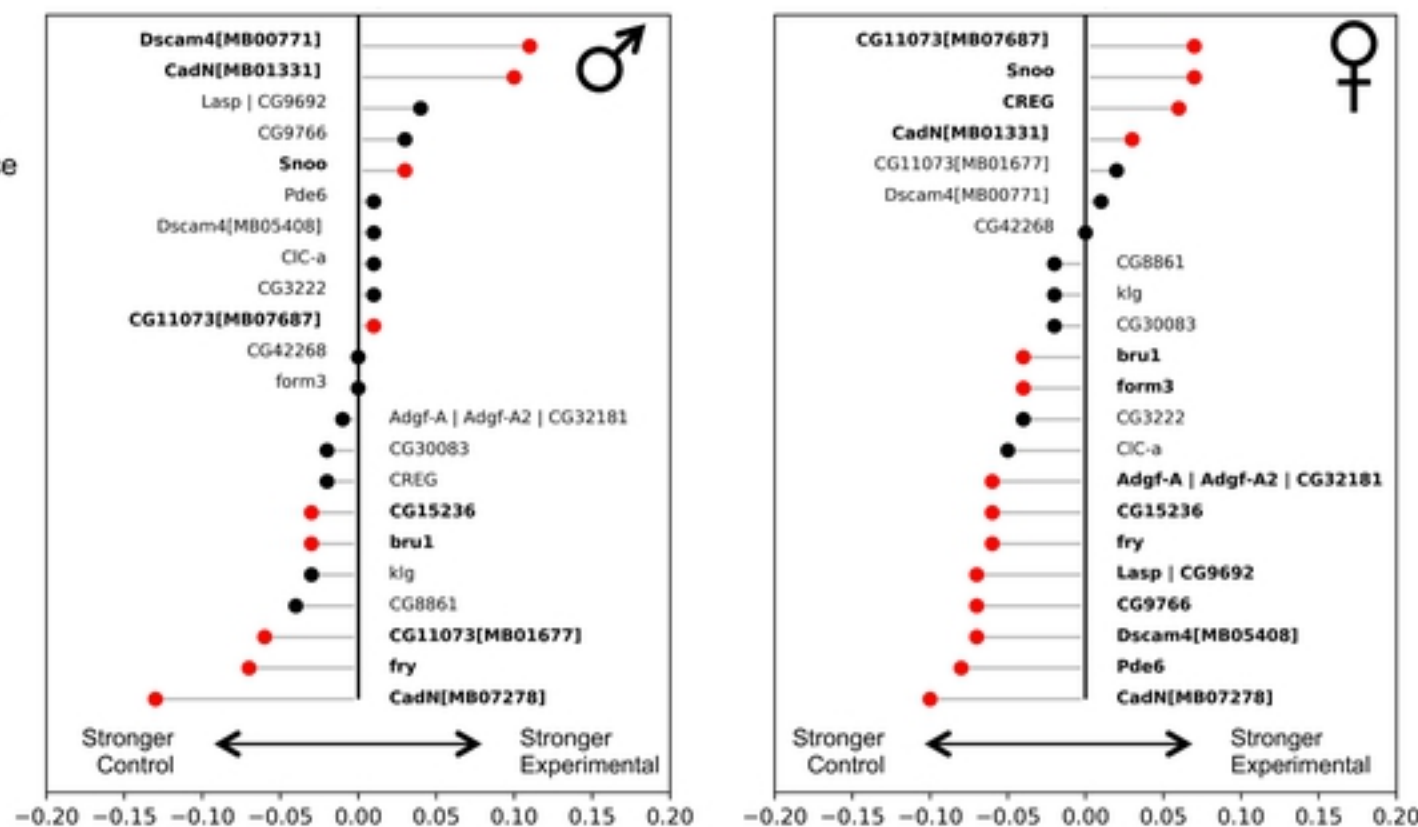


Figure 2

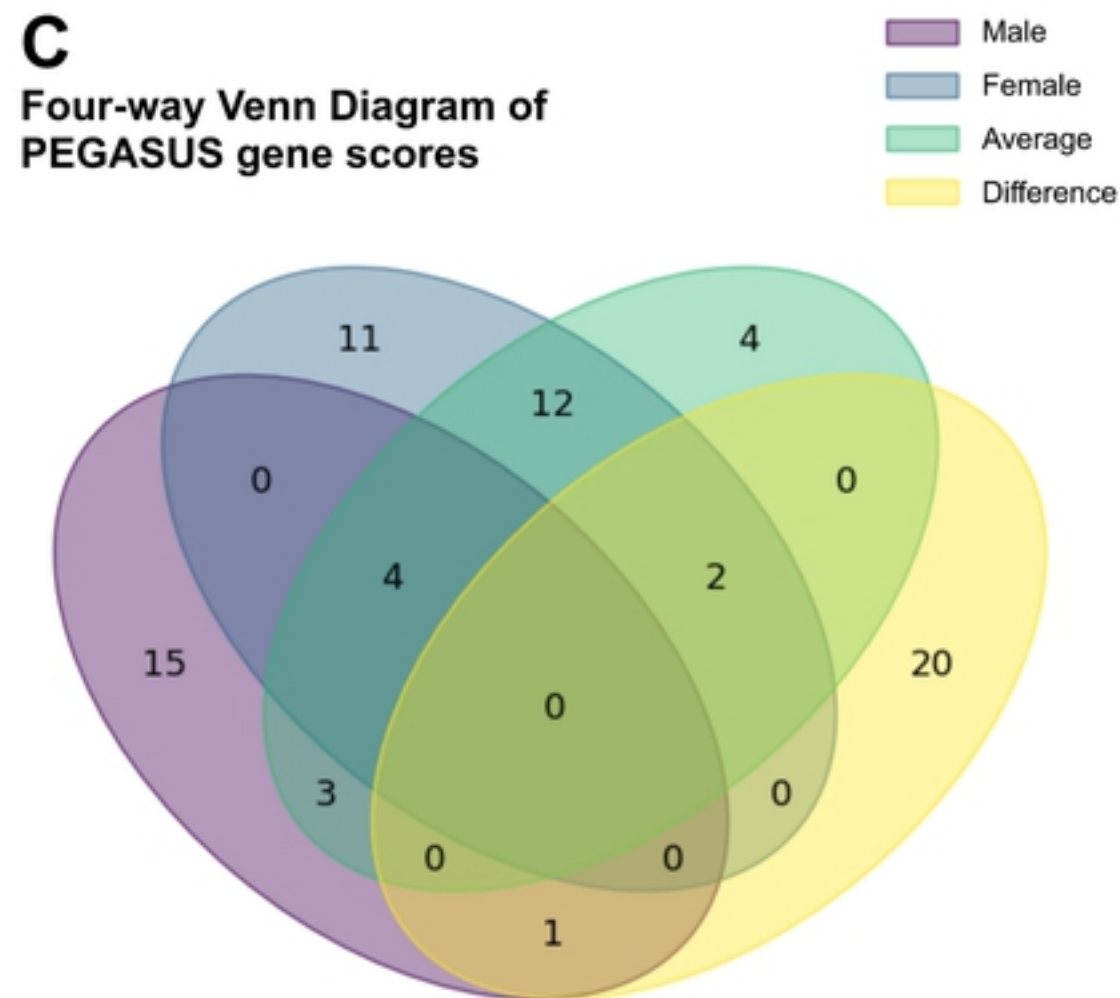
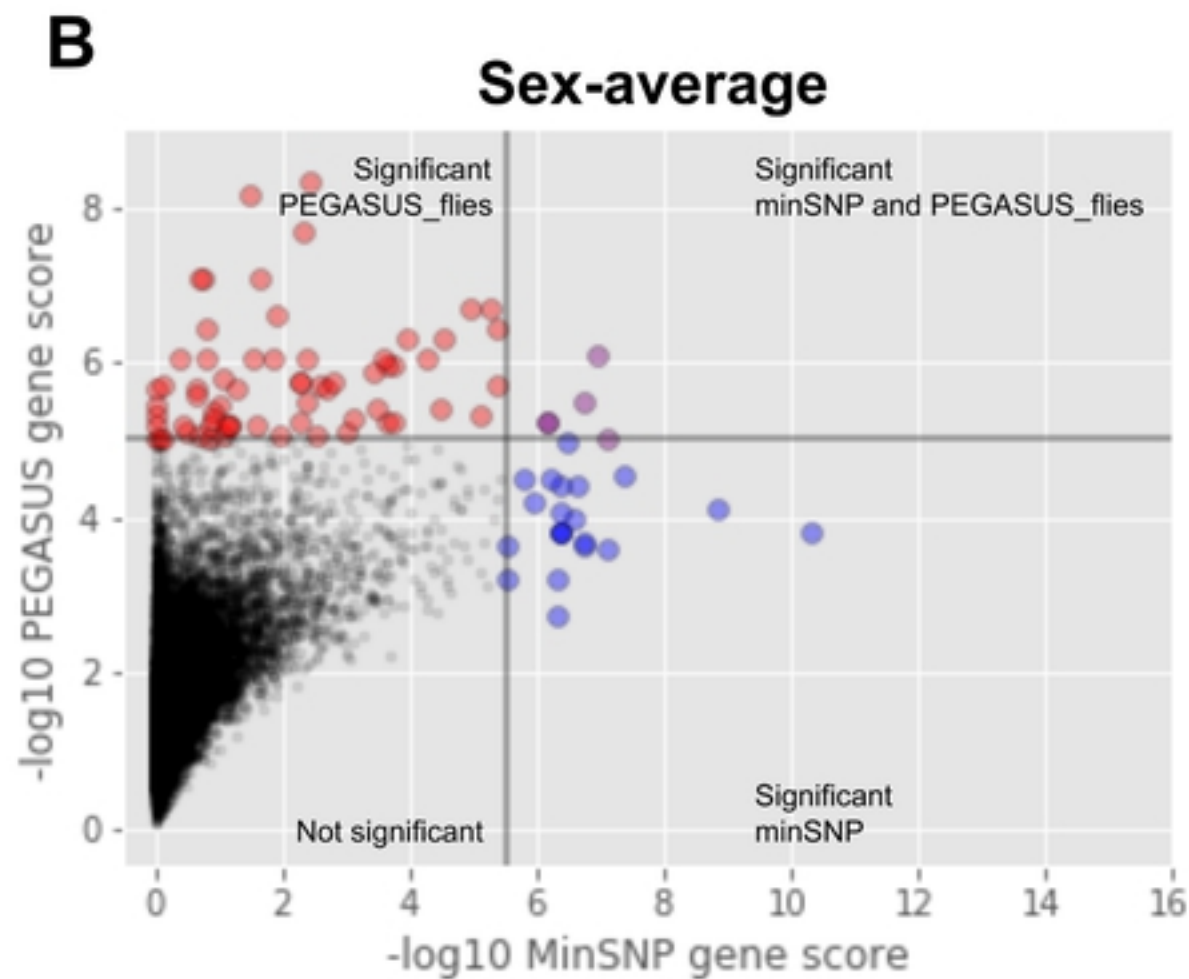
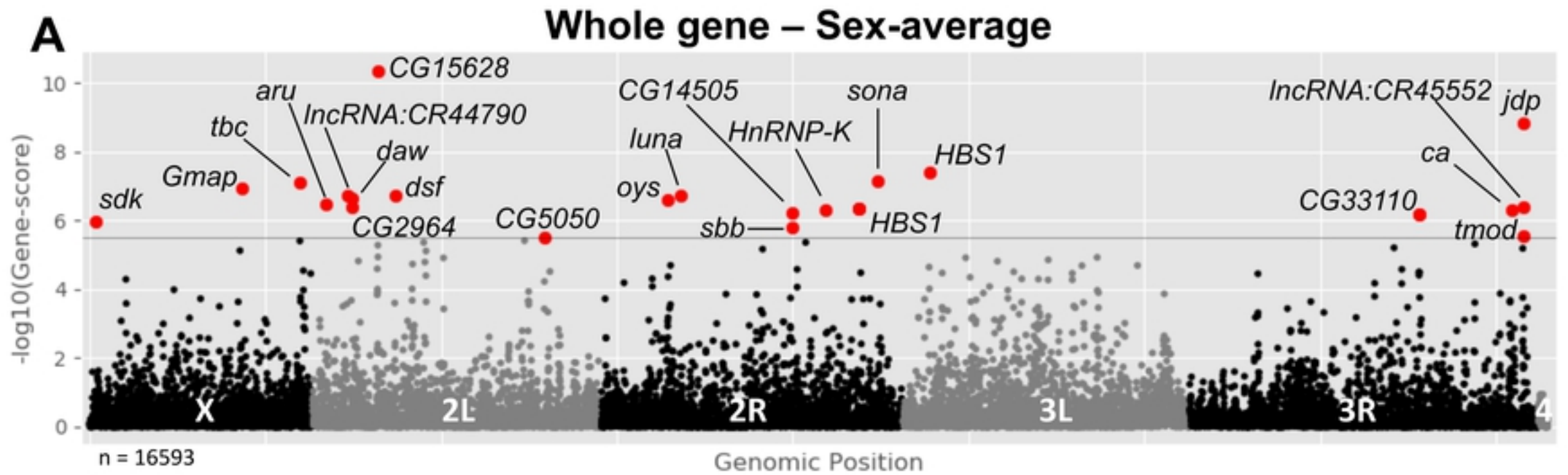


Figure 3

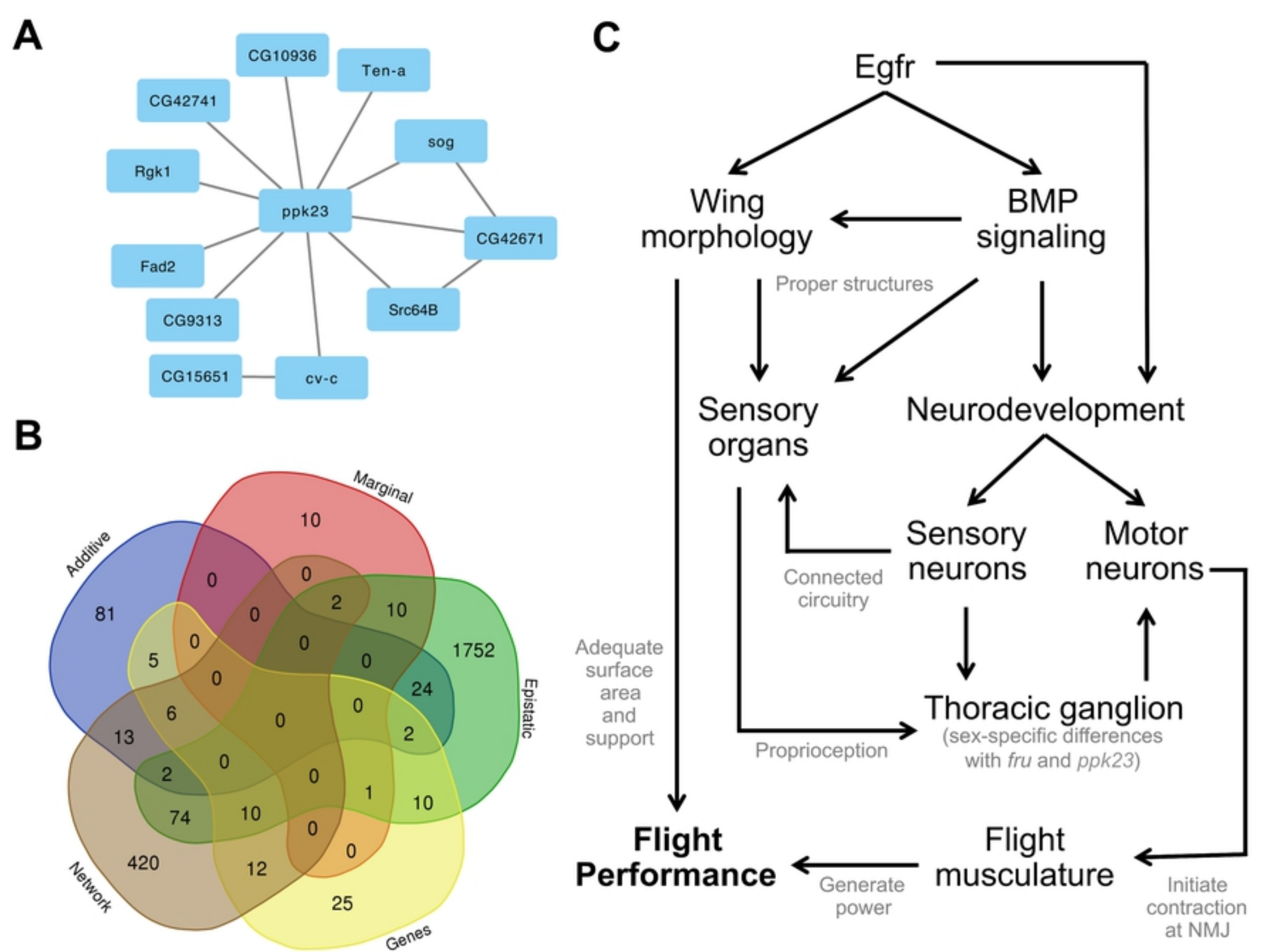


Figure 4