

ARTICLE

Received 3 Sep 2014 | Accepted 16 Jan 2015 | Published 18 Feb 2015

DOI: 10.1038/ncomms7312

# Host genetic determinants of microbiota-dependent nutrition revealed by genome-wide analysis of *Drosophila melanogaster*

Adam J. Dobson<sup>1,\*†</sup>, John M. Chaston<sup>1,\*†</sup>, Peter D. Newell<sup>1,†</sup>, Leanne Donahue<sup>1</sup>, Sara L. Hermann<sup>1</sup>, David R. Sannino<sup>2</sup>, Stephanie Westmiller<sup>1</sup>, Adam C.-N. Wong<sup>1,†</sup>, Andrew G. Clark<sup>3</sup>, Brian P. Lazzaro<sup>1</sup> & Angela E. Douglas<sup>1,3</sup>

Animals bear communities of gut microorganisms with substantial effects on animal nutrition, but the host genetic basis of these effects is unknown. Here we use *Drosophila* to demonstrate substantial among-genotype variation in the effects of eliminating the gut microbiota on five host nutritional indices (weight, protein, lipid, glucose and glycogen contents); this includes variation in both the magnitude and direction of microbiota-dependent effects. Genome-wide association studies to identify the genetic basis of the microbiota-dependent variation reveal polymorphisms in largely non-overlapping sets of genes associated with variation in the nutritional traits, including strong representation of conserved genes functioning in signalling. Key genes identified by the GWA study are validated by loss-of-function mutations that altered microbiota-dependent nutritional effects. We conclude that the microbiota interacts with the animal at multiple points in the signalling and regulatory networks that determine animal nutrition. These interactions with the microbiota are probably conserved across animals, including humans.

<sup>1</sup>Department of Entomology, Cornell University, Ithaca, New York 14853, USA. <sup>2</sup>Department of Microbiology, Cornell University, Ithaca, New York 14853, USA. <sup>3</sup>Department of Molecular Biology and Genetics, Cornell University, Ithaca, New York 14853, USA. \* These authors contributed equally to this work. † Present addresses: Department of Genetics, Evolution and Environment, Institute of Healthy Ageing, University College London, Gower Street, London WC1E 6BT, UK (A.J.D.); Department of Plant and Wildlife Sciences, Brigham Young University, Provo, Utah 84602, USA (J.M.C.); Biological Sciences, SUNY Oswego, Oswego, New York 13126-3599, USA (P.D.N.); School of Biological Sciences, University of Sydney, Sydney, New South Wales 2006, Australia (A.C.-N.W.). Correspondence and requests for materials should be addressed to A.E.D. (email: aes326@cornell.edu).

Animal phenotypes are strongly influenced by microorganisms that colonize their surfaces (for example, the skin, gut and reproductive tract) and sometimes internal organs<sup>1</sup>. Some microbial effects can be attributed to specific microbial functions, for example, synthesis of specific nutrients or protective toxins<sup>2</sup>. Other microbial effects on the host, including promotion of intestinal homeostasis, immunity and metabolic function, involve complex networks of interactions between the animal host and microbiota<sup>3–7</sup>. These complex interactions have been interpreted as evidence that animal regulatory networks are structured to function in the context of the resident microbiota<sup>1,2</sup>, with the implication that host health and vigour can be prejudiced by mismatch between host function and the composition or activities of the microbiota, a condition known as dysbiosis<sup>8</sup>.

The purpose of this study was to quantify how the effect of the microbiota on host phenotype varies with host genotype, and to elucidate the genetic bases of these microbiota-dependent host traits. This issue has not been addressed directly for any system, even though it has important implications for our understanding of the genetic basis of human diseases linked to microbiota<sup>9</sup> and can potentially make significant contributions to the development of personalized microbial therapies<sup>10–12</sup>. More generally, understanding how the microbiota-dependent phenotype maps onto the host genotype will enrich our understanding of the evolution and function of interactions between animals and their resident microbiota.

Our research was conducted on the fruit fly *Drosophila melanogaster* and its gut microbiota, which is ideally suited for the study of microbiota-dependent effects for three reasons. First, experimental analysis is facilitated by robust methods to eliminate the gut microbiota by egg dechoriation, yielding axenic flies<sup>13,14</sup>. This treatment does not affect the complement of the intracellular bacterium *Wolbachia*, which is present in many *Drosophila* lines and is vertically transmitted via the egg cytoplasm. Second, axenic individuals of *Drosophila* lines studied to date commonly display readily quantified nutritional traits, including elevated levels of indices of triglyceride, glycogen or free glucose<sup>15,16</sup>, and these changes have been linked to altered function of the nutrient-sensing insulin/insulin-like peptide signalling (IIS) and target of rapamycin (TOR) signalling pathways that couple organismal growth to nutrient supply<sup>17,18</sup>. Finally, the superb genetic resources for *Drosophila* can be harnessed to interrogate the genetic architecture of microbiota-dependent effects. In particular, the *Drosophila* Genetic Reference Panel (DGRP) of inbred isofemale lines with sequenced genomes enable genotype–phenotype mapping by genome-wide association (GWA)<sup>19–21</sup> and candidate genes identified from GWA can then be validated experimentally by mutant analysis. The design of this study was also informed by research on the composition of the gut microbiota, which is dominated by bacteria of the Acetobacteraceae ( $\alpha$ -proteobacteria) and Lactobacillales (Firmicutes)<sup>22</sup>. The impact of the gut microbiota on *Drosophila* nutritional indices depends on the composition of the microbiota<sup>23</sup>, which can vary, apparently stochastically, among *Drosophila* stocks maintained under uniform conditions<sup>24,25</sup>. To standardize the microbiota in the test DGRP lines, this study was conducted on flies generated from dechoriated eggs and exposed to isolates of five bacterial species that were isolated from *Drosophila* guts, are found ubiquitously in association with laboratory-cultured and wild-caught *Drosophila* and, in combination, have been shown to restore the nutritional phenotype of *Drosophila* bearing its unmanipulated microbiota<sup>23–25</sup>.

This study focused on the nutritional effects of the *Drosophila* microbiota. Using the DGRP, we demonstrated substantial among-line variation in nutritional response to elimination of

the microbiota and identified host genetic variants (single-nucleotide polymorphisms, SNPs) associated with the microbiota-dependent nutritional traits. Many of the genes identified have fundamental roles in cell signalling and control of gene expression that are conserved in many animals, including humans. Our results show the central role of host genotype in shaping the impact of microbiota on animal nutrition.

## Results

**Genetic variation in microbiota-dependent nutritional traits.** Gnotobiotic *Drosophila* associated with a defined five-species microbiota and axenic (microbe-free) *Drosophila* differed significantly in all of the five nutritional traits tested (Fig. 1). The global mean values of dry weight per fly were depressed in axenic flies and, after controlling for weight, axenic flies also had reduced glycogen and protein contents but elevated glucose and triglyceride (TAG) contents (Fig. 1a and Supplementary Table 1). Nevertheless, the among-line variation in the response to elimination of the microbiota was appreciable for all traits, and genotype accounted for 31%–73% of the total variance (Fig. 1b) and within-line variation was low (Supplementary Fig. 1). For every trait, the variation included lines that differed in response direction, that is, every trait was elevated in axenic flies of some lines, but decreased in other lines (Fig. 1c). Among the lines tested, 51% bore the intracellular bacterium *Wolbachia*, which was not eliminated by the egg dechoriation procedure used to generate axenic flies. The *Wolbachia*-positive lines had significantly elevated glycogen content relative to *Wolbachia*-free lines in both axenic and gnotobiotic flies (Fig. 1), perhaps due to sensitization to insulin signalling<sup>26</sup>.

For each trait, the mean values were significantly correlated across the axenic and gnotobiotic flies (weight  $r = 0.8$ ; protein/weight  $r = 0.61$ ; glucose/weight  $r = 0.46$ ; glycogen/weight  $r = 0.53$ ; TAG/weight  $r = 0.77$ ; all  $P < 0.001$ ), reinforcing the evidence above for a genetic component to these traits. We also tested for correlations among the various phenotypic traits (Supplementary Table 2). No traits were significantly correlated in axenic flies, irrespective of *Wolbachia* status; this result indicates that generalized among-genotype variation in sensitivity to elimination of the microbiota is not a major driver of the microbiota-dependent effects. Among gnotobiotic flies, *Wolbachia* status determined which traits were correlated: TAG was positively correlated with glucose and glycogen contents in *Wolbachia*-positive gnotobiotic flies, but this effect was absent in *Wolbachia*-free flies (Supplementary Table 2). The data did not reveal specific lines whose nutritional status was globally hypersensitive to elimination of the microbiota, evidenced by the absence of positive correlations among response indices for each trait (Supplementary Fig. 2).

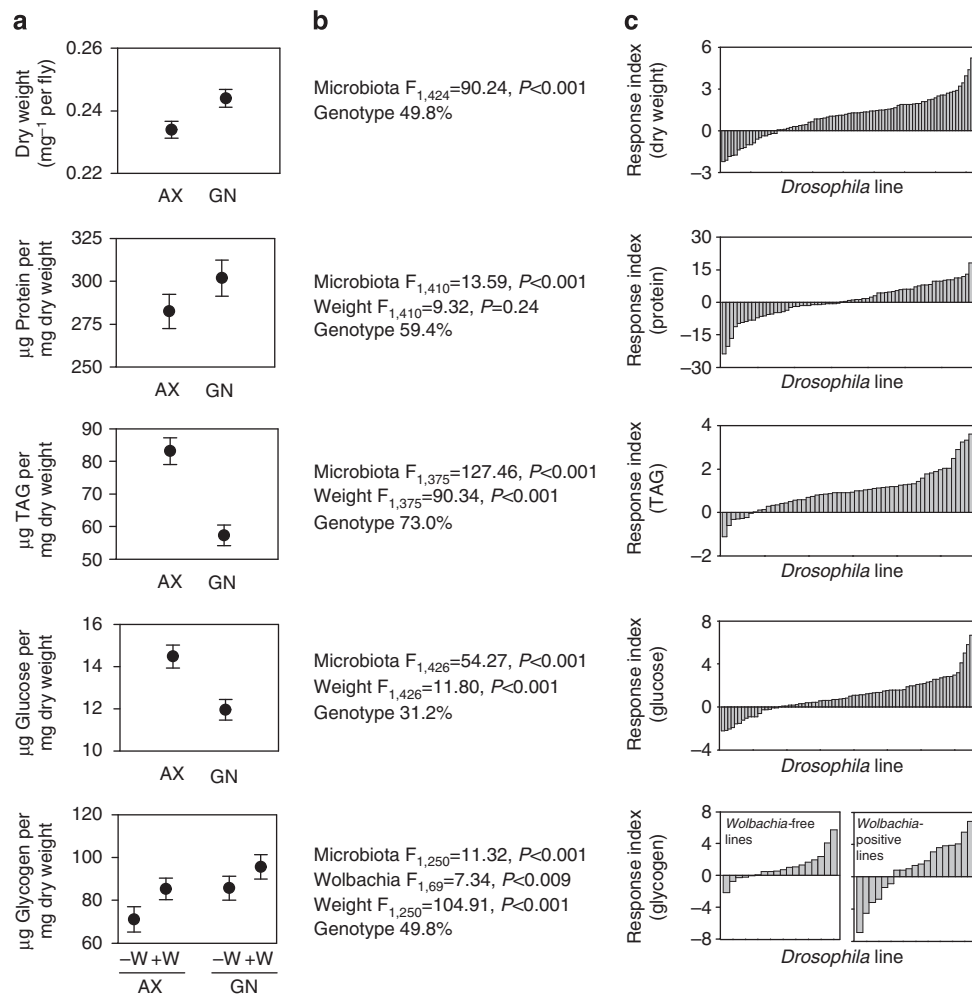
Collectively, these results demonstrate that *Drosophila* nutrition has a strong genetic component, that trait values are influenced by the microbiota and that there is genetic variation in the microbiota-dependence of *Drosophila* nutritional traits. Furthermore, the variation in the response of the various traits to elimination of the microbiota does not appear to be coupled.

**Genotype–phenotype associations.** To determine the association between SNPs, microbiota and the nutritional traits, the interactive effect of microbiota and SNP was tested for biallelic SNP markers in the *Drosophila* genome. We determined SNP significance by  $P$ -value rank, because of  $P$ -value inflation and the ineligibility of standard *post-hoc* corrections for GWA<sup>27</sup>. The top ranking 0.01% of SNPs for each trait (Supplementary Fig. 3) comprised 184–226 SNPs, including SNPs associated with 77–144 genes, with a total of 436 genes mapped across the full data set

(Supplementary Table 3). Linkage disequilibrium between these SNPs was generally low, apart from an island of linked SNPs on chromosome 21 associated with variation in glycogen stores (Supplementary Fig. 4). The mapped SNPs were found more commonly in genic regions than expected, based on genome background (number of SNPs <5,000 bp from known genes: mapped SNPs = 91%, genome background = 57%, proportions test  $P < 0.0001$ ). Significant gene ontology assignments (false discovery rate  $\leq 0.05$ ) and the loci responsible for assignments are presented in Supplementary Table 4.

To gain an overview of the biological functions potentially determining the host phenotypic responses to the microbiota, we interrogated known functions of genes associated with the 20 most significant SNPs found by each GWA (Supplementary Table 3). Associations between microbiota-dependent variation in dry weight and SNPs included a SNP in the transcription factor *Lim3*, a SNP in *dpr10* (a predicted chemosensor enriched in expression in the brain) and two SNPs in the translational modifier *boule* (*bol*). Microbiota-dependent variation in protein content was associated with multiple SNPs in *varicose*, which has

known roles in trachea development<sup>28</sup>, and variation in both dry weight and protein are associated with SNPs in *trol*, a modulator of various signalling pathways, including fibroblast growth factor, hedgehog and Wnt<sup>29,30</sup>. Microbiota-dependent variation in TAG was associated with SNPs in genes coding a number of central regulatory elements including the kinase *happyhour* (*hppy*: negative regulator of epidermal growth factor receptor signalling and a kinase of JNK and IIS/TOR signalling), and the genes coding the tyrosine protein phosphatase *Ptp99A*, *highwire* (*hiw*, a member of the BMP signaling pathway) and *rugose* (*rg*), homologue of the mammalian *Neurobeachin*, which regulates neurotransmitter receptor trafficking implicated in associative learning<sup>31,32</sup>, as well as the transcription factors *klumpfuss* (*klu*) and *scribbler* (*sbb*). The candidate functional significance of IIS/TOR signalling indicated by the SNP in *hppy* is strengthened by lower-ranked associations of TAG with SNPs in two genes *S6K* (ranked 114) and *meltd* (136) coding canonical IIS/TOR proteins downstream of *hppy*, as well as four further *hppy* SNPs (122, 124, 138 and 139). The most significant SNP associated with microbiota-dependent variation in glucose content was in *slit*



**Figure 1 | Effects of elimination of the microbiota on *Drosophila* nutritional traits vary by host genotype.** Data were collected from a total of 108 DGRP lines, in pooled samples of 5 males per line (up to 3 replicate samples per line). **(a)** Population means and s.e. for axenic (AX) and gnotobiotic (GN) flies with s.e. calculated from means per individual DGRP lines. Data are plotted by *Wolbachia* status (–W, *Wolbachia*-free, +W, *Wolbachia*-positive) where effects of *Wolbachia* are significant, and nutritional indices are normalized to line mean dry weight to avoid confounding effects of variation in weight. **(b)** Analysis of variance models of each trait simplified from a full model of microbiota + *Wolbachia* for weight and microbiota + *Wolbachia* + weight for other indices, with genotype nested in experimental block as random effects for all traits (full statistical output in Supplementary Table 1). Percentage variance explained by genotype was calculated as the square of the s.d. around the genotype coefficient. **(c)** Response indices of lines to elimination of the microbiota.

(*sli*), the ligand of *roundabout* signalling that controls intestinal stem cell fate<sup>33</sup>. Other SNPs in this association included *dunce* (*dnc*), a cAMP-phosphodiesterase, which is involved in regulation of multiple behaviours<sup>34</sup>, and the nicotinic acetylcholine receptor *nAChRx2*. Glycogen showed particularly strong associations with the JAK-STAT activated transcription factor *chinmo*<sup>35</sup>, the epidermal growth factor receptor-responsive transcription factor *seven up* (*svp*), as well as the morphogen *nuf*, *numb* (an antagonist of Notch signalling) and *CG5867*, which contains a protein domain for juvenile hormone binding.

In total, SNPs linked to just 26 genes (6% of the 436 genes) were associated with more than one trait (Supplementary Table 5) and no gene mapped to more than two traits. The low incidence of genes associated with multiple traits is consistent with the responses of traits to removal of the microbiota (Supplementary Table 2).

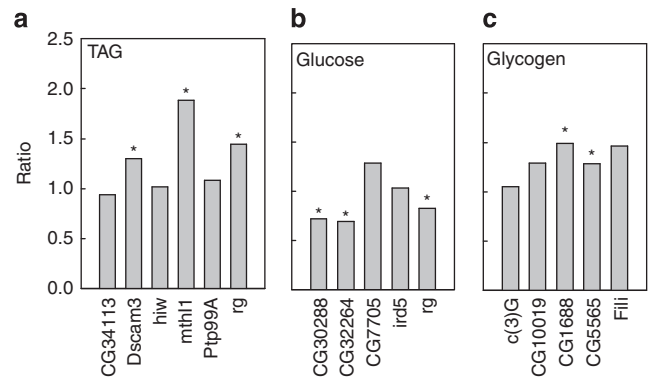
**Functional validation of GWA.** The validation of associations between SNPs and microbiota-dependent phenotypic traits focused on genes linked to mapped SNPs. We reasoned that if a SNP correctly infers the role of a gene, then the trait difference between corresponding flies with a null mutation and wild-type flies could be obtained in either axenic flies or gnotobiotic flies, but not in both treatments. Therefore, our definition of a validated effect of mutation is a statistically significant difference between the mutant and background in one but not both of the axenic and gnotobiotic flies, with a *P*-value  $\leq 0.025$  (correcting for two comparisons) by a mixed-effects model. By this criterion, 9 (56%) of 16 predicted microbiota-dependent effects of the tested genes on TAG, glucose and glycogen were validated (Table 1 and Fig. 2). Some off-target effects were anticipated because mutations are predicted to cause greater disruption of genetic networks than natural polymorphisms. In total, four (27%) of 15 tests for off-target effects of genes validated for nutritional traits predicted in the GWAS study (GWAS) yielded significant effects (Table 1, Supplementary Table 6 and Supplementary Fig. 5). There were two classes of response to microbiota in validated mutants: the microbiota either suppressed the effects of some mutations, such that the ratio of response (Fig. 2) was significantly greater than unity (for example, *Dscam3*, *mthl1* and *rg* for TAG), or the microbiota promoted the effects of

mutations on the host phenotypic trait (for example, glucose levels for *CG32264*, *CG30288* and *rg*).

Taken together, these results provide experimental verification that GWAS is an effective route to identify genes with significant microbiota-dependent effects on nutritional traits of *Drosophila*.

**Discussion**

This study has demonstrated substantial effects of host genetics on the nutritional response of *Drosophila* to elimination of the gut microbiota. For every index tested, both the magnitude and direction of the response varied among the *Drosophila* lines. For example, previous reports of elevated TAG and glucose in axenic flies of the *Drosophila* line Canton S background<sup>15,16,23</sup> are consistent with results obtained for most of the DGRP lines studied here, but a minority displayed the reverse response, and both positive and negative responses to elimination of the



**Figure 2 | GWAS validation.** Effect of microbiota on nutritional indices in fly lines with mutations in GWA-predicted genes for each of (a) TAG, (b) glucose and (c) glycogen. Ratio refers to the ratio of the gnotobiotic:axenic index value in the mutant relative to the ratio of the gnotobiotic:axenic index value in the background *Drosophila* stock. Statistical differences (mixed-effects linear models) between the mutant and its background were calculated under each microbiota treatment (axenic and gnotobiotic) and an effect was assigned where mutant and background stocks were significantly different under only one microbiota treatment (indicated by \*).

**Table 1 | Microbiota-responsive traits among GWA-validated mutants.**

Genes	Wolb.*	SNP rank in GWA	Validated effects of microbiota†			Number of validated effects/total tested	
			TAG	Glucose	Glycogen	Validated‡	Validated off-target effect§
TAG						2/5	1/4
<i>Dscam3</i>	N	13	+				
<i>mthl1</i>	N	17	+	-			
Glucose						2/4	1/4
<i>CG32264</i>	N	3		+			
<i>CG30288</i>	Y	45		+	-		
Glycogen						3/5	1/6
<i>CG1688</i>	Y	1			+		
<i>CG5565</i>	Y	13			+		
<i>Fili</i>	Y	15			+		
TAG and glucose						2/2	1/1
<i>rg</i>	Y	5,6,7 TAG 86 glucose	+	+	-		

GWA, genome-wide association; N, absent; SNP, single-nucleotide polymorphism; TAG, triglyceride; Y, present.

Data sets showing predicted and non-predicted significant effects of mutations on microbiota-dependent nutritional traits are displayed in Supplementary Fig. 3.

\*Incidence of *Wolbachia*, determined by diagnostic PCR of adult flies (Y and N).

†+, predicted effect; -, unpredicted effect.

‡Number validated genes for nutritional trait/total number of genes tested for nutritional trait.

§Number of off-target effects of genes that had GWA-predicted effects/total possible number of off-target effects of genes that had GWA-predicted effects.

||Non-synonymous SNP in coding region in GWA.

microbiota were well represented among the lines for protein content and glycogen content. These results are complemented by the mutant analysis and suggest that interactions between the microbiota and host processes that regulate nutrient allocation are not tightly constrained and are responsive to genetic variation in the host. Host genotypic effects on microbiota-dependent nutritional traits may be general among animals, including the human and mouse, for which some traits of the microbiota are already known to vary with host genotype<sup>36–38</sup>.

A key result of this study is that the effects of eliminating the microbiota on different nutritional indices are not correlated. This suggests that although nutrient allocation patterns are coordinately regulated by integrated signalling circuits, the genetic variation that determines the responsiveness of these networks to the microbiota may have effects specific to individual nutritional indices. Fully consistent with this finding, GWAs revealed that variation in each nutritional trait mapped predominantly to distinct sets of loci and functional processes. Although an experimental design with greater statistical power may have yielded evidence for more pleiotropic genes, these results suggest that many of the genes with microbiota-dependent effects on nutrition have low pleiotropy and raise the possibility that many of the genes, which influence multiple nutritional indices, may generally be microbiota unresponsive.

The functional validations using mutant *Drosophila* provided a critical test for the reliability of the GWAs in predicting genes with microbiota-dependent effects on nutritional traits. These tests confirmed that host nutrition is determined by an interaction of the microbiota and host genotype. The nature of this interaction appears to depend on the gene in question, as the microbiota suppressed nutritional effects of some mutations, but were required to detect nutritional effects of other mutations. The mutants commonly had the effects predicted by GWAs. However, our ability to make statements about trait-specific effects of genes mapped by GWA is tempered by nonspecific effects of mutants of some genes on certain phenotypic traits. Possible technical causes of discrepancies between mutants and GWAs are false negatives in the GWAs and artefacts of mutants' genetic backgrounds. Possible biological causes include off-target effects of mutants, owing to potentially greater disruption of genetic networks by transgenic mutation than by natural polymorphism or the related issue of SNPs having tissue-specific effects on gene function that are not mirrored by mutations. We therefore interpret the results of our functional validations as consistent with the absence of a signal of pleiotropy from GWA mappings, with the caveat that 27% of possible effects were not predicted by GWAs.

Successful validation of mutants is particularly relevant in the context of two limitations of the genome-wide analyses: poor definition of the critical probability for GWAS due to the ineligibility of standard corrections for multiple statistical comparisons<sup>27,39</sup> and inflated *P*-values identified in QQ plots (Supplementary Fig. 6), potentially indicative of mis-specification in GWA models (Supplementary Note 1). We found that the microbiota-dependent effects for each trait were more frequently abrogated in *Drosophila* bearing mutations in genes mapped by specific GWAs than in collective pools for the individual genes of interest, confirming that mapped genes are microbiota-responsive (Fig. 2 and Table 1). We conclude that the validation studies of mutants support the genome-wide mappings as a strategy to generate hypotheses about the genetics underlying variation in impacts of the microbiota on nutrition.

The GWAs identified genes functioning in IIS/TOR signalling, a network identified previously to interact with the gut microbiota<sup>17,18</sup>. IIS/TOR signalling is a nexus of nutrition and growth (for example, see refs 40,41), but with extensive cross-talk with other pathways involved in regulation of growth.

The identification of genes contributing to multiple signalling pathways other than IIS/TOR indicates that the effects of the microbiota on host phenotype cannot be described exclusively in terms of any single defining pathway: instead, our results are congruent with the view that the microbiota interact at multiple points in the signalling and regulatory networks that define animal function. These data provide evidence for the argument made previously<sup>1,2</sup>, but largely without supporting data, that many animal–microbiota interactions are probably conserved across the animal kingdom, including in humans, because the associations with resident microbiota are ubiquitous among animals and important for animal health and vigour.

Immediately relevant to these considerations are genes with no known function, accounting for 43% of mapped genes. The nutritional functions of many of these genes may be microbiota dependent and may have gone undetected by traditional genetic screens, which do not control for microbiota effects. Experimental analyses that compare axenic hosts with hosts that are colonized with different microbial partners could promote understanding of these currently un-annotated genes. The absence of canonical immunity genes or immunity-related gene ontology terms significantly associated with any microbiota-dependent trait is also of interest. Various immune effectors have been implicated in the regulation of interactions with the microbiota in *Drosophila*<sup>42,43</sup> and other animals<sup>44–47</sup>, raising the possibility that genetic variation in immunity may shape the microbiota in a manner that affects traits other than the nutrition-related traits studied here, or this variation is not exposed in associations with a specified microbiota. Our GWA analyses necessarily did not account for variation in the hosts' ability to associate with the microbiota. The replication of host genotypes and properties of the microbiota (for example, total abundance and relative abundance) required to distinguish the effects of these two factors on host phenotype would require thousands of *Drosophila* lines. A key objective of future studies should be to map genetic variation in the predisposition of *Drosophila* for certain microbial communities and to characterize any correspondence between such variation and nutrition.

The GWA mappings elucidated in this study offer a rich inventory of genes with candidate roles in microbiota-dependent traits. Although our experimental design precludes detailed analysis of the evolutionary forces shaping this variation, because the inbred DGRP lines bear unnaturally segregated allelic combinations that expose variants, which may not be subject to selection under field conditions, the host genes identified probably include targets of microbiota-dependent selection. Future research to identify the mode of action and evolutionary significance of genes with microbiota-dependent function will need to account for the complexity of the interactions: that the impacts of the microbiota on host traits are strongly influenced by host genotype and involve variation in multiple genes with largely independent effects on different phenotypic traits.

## Methods

**The flies.** *Drosophila* from the DGRP<sup>19</sup> and transposon (P-element insertion) mutant flies obtained from the Bloomington *Drosophila* Resource Center, IN, USA (Supplementary Table 7) were cultured at 25 °C on a 12 h:12 h light–dark cycle on yeast–glucose medium (11 water, 100 g brewer's yeast (MP Biomedicals 903312), 100 g glucose (Sigma 158968), 1.2% agar (Apex 66–103), 0.84% propionic acid, 0.08% phosphoric acid). Axenic and gnotobiotic flies were generated by allowing flies from stock cultures to oviposit overnight on grape juice agar (1 litre water, 100 g yeast, 100 g glucose, 113 ml Welch's concentrated grape juice). The eggs were picked from the agar with a paint brush and dechorionated by 5 min of washing in 0.6% hypochlorite (10% Clorox brand bleach), to remove all microorganisms associated with the egg surface (but not the cytoplasmic *Wolbachia*), followed by three rinses with sterile water. To produce axenic flies, the eggs were added to 7.5 ml sterile yeast–glucose food medium in 50 ml Falcon tubes (Fisher Scientific) using aseptic technique in a laminar flow cabinet. To produce gnotobiotic flies,

50 µl bacterial suspension at  $10^8$  cells per ml was added to tubes containing the axenic eggs. The bacterial suspension comprised equal proportions of *Acetobacter pomorum* DmelCS\_004, *Acetobacter tropicalis* DmelCS\_006, *Lactobacillus brevis* DmelCS\_003, *Lactobacillus fructivorans* DmelCS\_002 and *Lactobacillus plantarum* DmelCS\_001. Bacteria were streaked from clonal glycerol stocks kept at  $-80^\circ\text{C}$  onto modified Man, Rogosa and Sharpe (mMRS) 1.2% agar and grown at  $30^\circ\text{C}$ , either aerobically (*Acetobacter* spp.) or under  $\text{CO}_2$  (*Lactobacillus* spp.). Individual colonies were then inoculated into 10 ml mMRS broth. *Acetobacter* spp. were grown overnight at  $30^\circ\text{C}$  with shaking. *L. brevis* and *L. plantarum* were grown overnight at  $30^\circ\text{C}$  without shaking. *L. fructivorans* was grown for 7 days without shaking. For each *Drosophila* line, flies were reared from eggs in up to six replicate vials per treatment and phenotyped for GWA at 6–8 days post eclosion. The microbiological status of axenic and gnotobiotic flies was checked by plating a sample of fly homogenate (generated as below) onto 1.2% mMRS agar and incubating for 48 h at  $30^\circ\text{C}$  under aerobic conditions for *Acetobacter* and under  $\text{CO}_2$  for *Lactobacillus*. As low bacterial counts can be generated by contamination during fly sorting, our criteria for discarding axenic fly samples was  $>600$  colony-forming units per fly, equivalent to  $<1\%$  of colony-forming units per gnotobiotic fly.

For axenic and gnotobiotic flies of each line, the day on which  $\geq 50\%$  flies eclosed was recorded and the flies were harvested 5–7 days later (both treatments in any line were assayed on the same day). At 4–7 h after the beginning of the daily light cycle, the flies in each vial were lightly anaesthetized with  $\text{CO}_2$  and males were selected for analysis.

**Dry-weight quantification and nutritional assays.** For GWA, three pools of 5  $\text{CO}_2$ -anaesthetized flies per treatment per line were flash frozen on dry ice, desiccated over 7 days at  $55^\circ\text{C}$ , then weighed on a Mettler Toledo (MX5) microbalance to the nearest microgram. For each microbiota treatment per line, a second set of three pools of 5  $\text{CO}_2$ -anaesthetized flies were homogenized in 125 µl TE extraction buffer (35 mM Tris, 25 mM KCl, 10 mM  $\text{MgCl}_2$ , pH 7.5, 0.1% (v/v) Triton X-100) with ceramic beads (MP Biomedicals 6540–434) for 30 s in a tissue homogenizer (MP Biomedicals FastPrep-24). Twenty microlitres of homogenate was immediately flash frozen for subsequent protein quantification, the remainder was heat treated at  $72^\circ\text{C}$  for 15 min, then flash frozen. Samples were subsequently thawed and assayed for soluble protein, glucose, glycogen and TAG content, according to established protocols<sup>15</sup>. Protein was quantified by the Lowry assay using the BioRad DC kit (BioRad 500-0111). Glucose and glycogen were quantified with the Sigma Glucose assay kit (Sigma-Aldrich GAGO20-1KT): glucose homogenates were assayed directly from TE-buffered homogenates, whereas glycogen was digested to glucose in a 30-min incubation at  $30^\circ\text{C}$  with 5 µl amyloglucosidase solution (Sigma A7420 at  $1\text{ U ml}^{-1}$ ), before subtracting the previous glucose reading. TAG was quantified using the Sigma Triglyceride Assay kit (TR0100-1KT), in which glycerol content is determined before and after digestion with supplied Triglyceride Reagent. TAG content was determined as difference in glycerol content between the TAG-digested and -undigested measurements of glycerol.

**Wolbachia status of the flies.** DNA was isolated from adult flies by the salting out method<sup>48</sup> for most DGRP lines used in this study and for all P-element insertion mutants and backgrounds. Pools of five males were homogenized in 180 µl lysis buffer (20 mM Tris-HCl, 2 mM EDTA, 1.2% Triton X-100, with 20 mg lysozyme per ml) with a sterile pestle, incubated at  $37^\circ\text{C}$  for 60 min and vortexed with 100-µl 0.1-mm glass beads (Scientific Industries SI-BG01) for 5 min. Twenty microlitres of  $10\times$  extraction buffer (2 M Tris-HCl, pH 8.5, 2.5 M NaCl, 250 mM EDTA, 5% (w/v) SDS) was added to the samples, which were mixed with 10 µl 20 mg proteinase K per ml (Qiagen) and incubated at  $55^\circ\text{C}$  for 60 min. One hundred microlitres of 3 M sodium acetate (pH 5.2) were added to each sample before incubation at  $-20^\circ\text{C}$  for 10 min. Samples were centrifuged for 5 min at  $5,000g$  and  $4^\circ\text{C}$ . Three hundred microlitres of isopropanol (cooled to  $-20^\circ\text{C}$ ) were then added to the supernatant, mixed and then incubated at room temperature for 30 min. Following 30 min centrifugation at  $5,000g$  at room temperature, the pellets were rinsed in 500 µl 70% ethanol (cooled to  $-20^\circ\text{C}$ ), vortexed, centrifuged for 5 min at  $5,000g$  at room temperature and resuspended in sterile Milli-Q water. *Wolbachia*-specific primers Wsp81F and Wsp691R (5'-AAA AAT TAA ACG CTC CA-3' and 5'-TGG TCC AAT AAG TGA TGA AGA AAC-3', respectively)<sup>49</sup> were used for specific detection of *Wolbachia* in duplicate samples. PCR reactions (25 µl) contained 20.575 µl ddH<sub>2</sub>O, 2.5 µl buffer, 0.25 µl 20 µM Wsp81F primer, 0.25 µl 20 µM Wsp691R primer, 0.3 µl 20 mM dNTPs, 1 µl 50 mM  $\text{MgCl}_2$ , 0.125 µl Platinum Taq DNA polymerase ( $5\text{ U } \mu\text{l}^{-1}$ ) and 0.5 µl template DNA ( $2\text{--}100\text{ ng } \mu\text{l}^{-1}$ ). Cycle parameters were as follows:  $95^\circ\text{C}$  for 5 min, 30 cycles of  $60\text{ s } 95^\circ\text{C}$ ,  $60\text{ s } 54^\circ\text{C}$ ,  $60\text{ s } 72^\circ\text{C}$  and a final extension  $60\text{ s } 95^\circ\text{C}$ , 5 min at  $72^\circ\text{C}$ . There was high congruence between our detection of *Wolbachia* and previous descriptions of DGRP *Wolbachia* status<sup>19</sup>.

**Quantitative and GWA analyses.** All data were analysed in R (v3.0.2). The phenotypic traits were analysed with mixed models using the lme function from the nlme library<sup>50,51</sup>, excluding any row of the data matrix with missing data points. The interactive effect of microbiota and *Wolbachia* as binary factors was analysed with genotype nested within experimental block as a random factor.

Mean line dry weight was an additional covariate for the other indices. The interaction of microbiota and *Wolbachia* was simplified where the *Wolbachia* term was not significant. Full statistical analysis is presented in Supplementary Table 1.

Response indices were calculated for each trait as the difference between mean axenic and gnotobiotic values for the trait in each line, divided by the difference in mean of axenic and gnotobiotic trait value across all lines<sup>27</sup>. To calculate correlations between traits, nutritional traits were normalized to line mean dry weight within each microbiota condition, hence, weight was excluded from these analyses to avoid autocorrelative artefacts. Significance of pairwise Pearson's correlations were corrected for ten multiple comparisons by Bonferroni correction.

To associate phenotypic traits with genetic variants, GWAs were conducted, using the data for all replicates. R scripts were constructed specifically for this analysis so that all covariates could be fully accounted for. We associated traits with any marker that had a minor allele frequency greater  $>3$  in the pool of tested lines, giving a different subset of markers for each trait tested (number SNPs tested per trait: dry weight 1,881,661 SNPs; protein 1,805,384 SNPs; TAG 1,702,250 SNPs; glucose 1,977,756 SNPs; glycogen 1,447,725 SNPs). Models and numbers of DGRP lines used in each GWA are presented in Supplementary Table 8, and degrees of freedom in Supplementary Table 1. QQ plots of P-values from these associations were inflated above the null expectation (Supplementary Fig. 5) and this inflation could not be removed by model respecification or additional covariates (Supplementary Note 1 and Supplementary Fig. 6). Inferences about function and expression of genes bearing significant SNPs were based on Flybase FB2013\_06 (ref. 52).

**Empirical validation of GWA.** The validation experiments used flies with P-element insertion mutations in genes that included the 20 most significant SNPs and the five most significant non-synonymous SNPs in coding regions of genes associated with each trait. To exclude mutants with generally deleterious effects on fitness and condition, we excluded from study genes in which lethal mutations are known and balanced homozygous-lethal mutants. A total of 15 genes were tested after applying these criteria: 5 for TAG, 4 for glucose, 5 for glycogen and 1 (rg) with predicted effects on both TAG and glycogen. Predictions were not validated for protein, because only a single mutant was available according to these guidelines. The validation experiments were conducted on three replicate sets of five male flies in each of three separate experiments, by exactly the same procedures as in the GWAS, except that dry weights of five individual flies per replicate were determined.

The effect of the microbiota on each mutant was analysed, relative to a *Wolbachia*-status-matched background stock, by a linear mixed model (lme in R v3.0.2) accounting for genotype (mutant versus background) as a fixed effect and experiment ( $n = 3$ , 3 replicates per experiment) as a random effect. The effect was calculated separately for the mutant/background pair raised under each of axenic or gnotobiotic conditions: a significant effect of genotype of ( $P < 0.025$ , that is, correcting for two comparisons) under only one of axenic or gnotobiotic conditions indicated the mutant had a different response to the microbiota than the wild-type or background host. The ratio of the gnotobiotic:axenic index value in the mutant, relative to the ratio of the gnotobiotic:axenic index value of the background was calculated to display interactions among host genotype (mutant versus background) and treatment (gnotobiotic or axenic).

## References

- McFall-Ngai, M. *et al.* Animals in a bacterial world, a new imperative for the life sciences. *Proc. Natl Acad. Sci. USA* **110**, 3229–3236 (2013).
- Douglas, A. E. Symbiosis as a general principle in eukaryotic evolution. *Cold Spring Harb. Perspect. Biol.* **6**, a016113 (2014).
- Sommer, F. & Backhed, F. The gut microbiota—masters of host development and physiology. *Nat. Rev. Microbiol.* **11**, 227–238 (2013).
- Maynard, C. L., Elson, C. O., Hatton, R. D. & Weaver, C. T. Reciprocal interactions of the intestinal microbiota and immune system. *Nature* **489**, 231–241 (2012).
- Tremaroli, V. & Backhed, F. Functional interactions between the gut microbiota and host metabolism. *Nature* **489**, 242–249 (2012).
- Mortha, A. *et al.* Microbiota-dependent crosstalk between macrophages and ILC3 promotes intestinal homeostasis. *Science* **343**, 1249288 (2014).
- Hooper, L. V. & Macpherson, A. J. Immune adaptations that maintain homeostasis with the intestinal microbiota. *Nat. Rev. Immunol.* **10**, 159–169 (2010).
- Stecher, B., Maier, L. & Hardt, W. D. 'Blooming' in the gut: how dysbiosis might contribute to pathogen evolution. *Nat. Rev. Microbiol.* **11**, 277–284 (2013).
- Qin, J. *et al.* A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature* **490**, 55–60 (2012).
- van Nood, E., Dijkgraaf, M. G. & Kellner, J. J. Duodenal infusion of feces for recurrent *Clostridium difficile*. *N. Eng. J. Med.* **368**, 2145 (2013).
- Hungin, A. P. *et al.* Systematic review: probiotics in the management of lower gastrointestinal symptoms in clinical practice -- an evidence-based international guide. *Aliment. Pharmacol. Ther.* **38**, 864–886 (2013).
- Hardy, H., Harris, J., Lyon, E., Beal, J. & Foeys, A. D. Probiotics, prebiotics and immunomodulation of gut mucosal defences: homeostasis and immunopathology. *Nutrients* **5**, 1869–1912 (2013).

13. Bakula, M. The persistence of a microbial flora during postembryogenesis of *Drosophila melanogaster*. *J. Invert. Pathol.* **14**, 365–374 (1969).
14. Ridley, E. V., Wong, A. C. & Douglas, A. E. Microbe-dependent and nonspecific effects of procedures to eliminate the resident microbiota from *Drosophila melanogaster*. *Appl. Environ. Microbiol.* **79**, 3209–3214 (2013).
15. Ridley, E. V., Wong, A. C., Westmiller, S. & Douglas, A. E. Impact of the resident microbiota on the nutritional phenotype of *Drosophila melanogaster*. *PLoS ONE* **7**, e36765 (2012).
16. Wong, A. C., Dobson, A. J. & Douglas, A. E. Gut microbiota dictates the metabolic response of *Drosophila* to diet. *J. Exp. Biol.* **217**, 1894–1990 (2014).
17. Shin, S. C. *et al.* *Drosophila* microbiome modulates host developmental and metabolic homeostasis via insulin signaling. *Science* **334**, 670–674 (2011).
18. Storelli, G. *et al.* *Lactobacillus plantarum* promotes *Drosophila* systemic growth by modulating hormonal signals through TOR-dependent nutrient sensing. *Cell Metab.* **14**, 403–414 (2011).
19. Mackay, T. F. *et al.* The *Drosophila melanogaster* Genetic Reference Panel. *Nature* **482**, 173–178 (2012).
20. Jemc, J. C., Milutinovich, A. B., Weyers, J. J., Takeda, Y. & Van Doren, M. raw functions through JNK signaling and cadherin-based adhesion to regulate *Drosophila* gonad morphogenesis. *Dev. Biol.* **367**, 114–125 (2012).
21. Huang, W. *et al.* Natural variation in genome architecture among 205 *Drosophila melanogaster* Genetic Reference Panel lines. *Genome Res.* **24**, 1193–1208 (2014).
22. Broderick, N. A. & Lemaitre, B. Gut-associated microbes of *Drosophila melanogaster*. *Gut Microbes* **3**, 307–321 (2012).
23. Newell, P. D. & Douglas, A. E. Interspecies interactions determine the impact of the gut microbiota on nutrient allocation in *Drosophila melanogaster*. *Appl. Environ. Microbiol.* **80**, 788–796 (2014).
24. Wong, A. C., Chaston, J. M. & Douglas, A. E. The inconstant gut microbiota of *Drosophila* species revealed by 16S rRNA gene analysis. *ISME J.* **7**, 1922–1932 (2013).
25. Chandler, J. A., Lang, J. M., Bhatnagar, S., Eisen, J. A. & Kopp, A. Bacterial communities of diverse *Drosophila* species: ecological context of a host-microbe model system. *PLoS Genet.* **7**, e1002272 (2011).
26. Ikeya, T., Broughton, S., Alic, N., Grandison, R. & Partridge, L. The endosymbiont *Wolbachia* increases insulin/IGF-like signalling in *Drosophila*. *Proc. R. Soc. B* **276**, 3799–3807 (2009).
27. Jordan, K. W. *et al.* Genome-wide association for sensitivity to chronic oxidative stress in *Drosophila melanogaster*. *PLoS ONE* **7**, e38722 (2012).
28. Wu, V. M. *et al.* *Drosophila* Varicose, a member of a new subgroup of basolateral MAGUKs, is required for septate junctions and tracheal morphogenesis. *Development* **134**, 999–1009 (2007).
29. Lindner, J. R. *et al.* The *Drosophila* Perlecan gene *trol* regulates multiple signaling pathways in different developmental contexts. *BMC Dev. Biol.* **7**, 121 (2007).
30. Kamimura, K. *et al.* Perlecan regulates bidirectional Wnt signaling at the *Drosophila* neuromuscular junction. *J. Cell Biol.* **200**, 219–233 (2013).
31. Volders, K. *et al.* *Drosophila* rugose is a functional homolog of mammalian Neurobeachin and affects synaptic architecture, brain morphology, and associative learning. *J. NeuroSci.* **32**, 15193–15204 (2012).
32. Nair, R. *et al.* Neurobeachin regulates neurotransmitter receptor trafficking to synapses. *J. Cell Biol.* **200**, 61–80 (2013).
33. Biteau, B. & Jasper, H. Slit/Robo signaling regulates cell fate decisions in the intestinal stem cell lineage of *Drosophila*. *Cell Rep.* **7**, 1867–1875 (2014).
34. Davis, R. L. Olfactory memory formation in *Drosophila*: from molecular to systems neuroscience. *Ann. Rev. Neurosci.* **28**, 275–302 (2005).
35. Flaherty, M. S. *et al.* *chinmo* is a functional effector of the JAK/STAT pathway that regulates eye development, tumor formation, and stem cell self-renewal in *Drosophila*. *Dev. Cell* **18**, 556–568 (2010).
36. Benson, A. K. *et al.* Individuality in gut microbiota composition is a complex polygenic trait shaped by multiple environmental and host genetic factors. *Proc. Natl Acad. Sci. USA* **107**, 18933–18938 (2010).
37. Parks, B. W. *et al.* Genetic control of obesity and gut microbiota composition in response to high-fat, high-sucrose diet in mice. *Cell Metab.* **17**, 141–152 (2013).
38. Tims, S. *et al.* Microbiota conservation and BMI signatures in adult monozygotic twins. *ISME J.* **7**, 707–717 (2013).
39. Brown, E. B., Layne, J. E., Zhu, C., Jegga, A. G. & Rollmann, S. M. Genome-wide association mapping of natural variation in odour-guided behaviour in *Drosophila*. *Genes Brain Behav.* **12**, 503–515 (2013).
40. Böhni, R. *et al.* Autonomous control of cell and organ size by CHICO, a *Drosophila* homolog of vertebrate IRS1–4. *Cell* **97**, 865–875 (1999).
41. Luong, N. *et al.* Activated FOXO-mediated insulin resistance is blocked by reduction of TOR activity. *Cell. Metab.* **4**, 133–142 (2006).
42. Ryu, J. H. *et al.* Innate immune homeostasis by the homeobox gene *caudal* and commensal-gut mutualism in *Drosophila*. *Science* **319**, 777–782 (2008).
43. Broderick, N. A., Buchon, N. & Lemaitre, B. Microbiota-induced changes in *Drosophila melanogaster* host gene expression and gut morphology. *mBio* **5**, e01117–14 (2014).
44. Login, F. H. *et al.* Antimicrobial peptides keep insect endosymbionts under control. *Science* **334**, 362–365 (2011).
45. Franzenburg, S. *et al.* Distinct antimicrobial peptide expression determines host species-specific bacterial associations. *Proc. Natl Acad. Sci. USA* **110**, E3730–E3738 (2013).
46. Chu, H. & Mazmanian, S. K. Innate immune recognition of the microbiota promotes host-microbial symbiosis. *Nat. Immunol.* **14**, 668–675 (2013).
47. McFall-Ngai, M., Nyholm, S. V. & Castillo, M. G. The role of the immune system in the initiation and persistence of the *Euprymna scolopes-Vibrio fischeri* symbiosis. *Semin. Immunol.* **22**, 48–53 (2010).
48. Cenis, J. L., Perez, P. & Fereres, A. Identification of aphid (Homoptera, Aphididae) species and clones by random amplified polymorphic DNA. *Ann. Ent. Soc. Am.* **86**, 545–550 (1993).
49. Zhou, W., Rousset, F. & O’Neil, S. Phylogeny and PCR-based classification of *Wolbachia* strains using *wsp* gene sequences. *Proc. Biol. Sci.* **265**, 509–515 (1998).
50. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org> (2013).
51. Pinheiro, J., Bates, D., DebRoy, S. & Sarkar, D. Linear and nonlinear mixed effects models. *R package version 3.1-117*. <http://CRAN.R-project.org/package=nlme> (2014).
52. Marygold, S. J. *et al.* FlyBase: improvements to the bibliography. *Nucleic Acids Res.* **41**, D751–D757 (2013).

## Acknowledgements

We thank David Duneau, Søren Franzenburg, Brenda Gustin, Ginny Howick, Xiangfeng Jing, Susan Rottschaefer, Calum Russell, Robin Schwenke and Rob Unckless for their assistance in experiments. Nicolas Buchon provided advice in the design of mutant studies, and Julien Ayroles, Keegan Kelsey, Jason Mezey, Cornelia Scheitz and Rob Unckless provided valuable analytical support. Financial support was provided by NIH grant 1R01GM095372 and The Sarkaria Institute for Insect Physiology and Toxicology.

## Author contributions

Conceived and designed experiments: A.E.D., J.M.C., A.J.D., B.P.L. and A.G.C. Conducted experiments: A.J.D., J.M.C., P.D.N., S.L.H., S.W., D.R.S., L.D., A.C.-N.W., B.P.L. Analysed data: A.J.D., J.M.C. and A.E.D.

## Additional information

**Supplementary Information** accompanies this paper at <http://www.nature.com/naturecommunications>

**Competing financial interests:** The authors declare no competing financial interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**How to cite this article:** Dobson, A. J. *et al.* Host genetic determinants of microbiota-dependent nutrition revealed by genome-wide analysis of *Drosophila melanogaster*. *Nat. Commun.* **6**:6312 doi: 10.1038/ncomms7312 (2015).

## **Supplementary Materials**

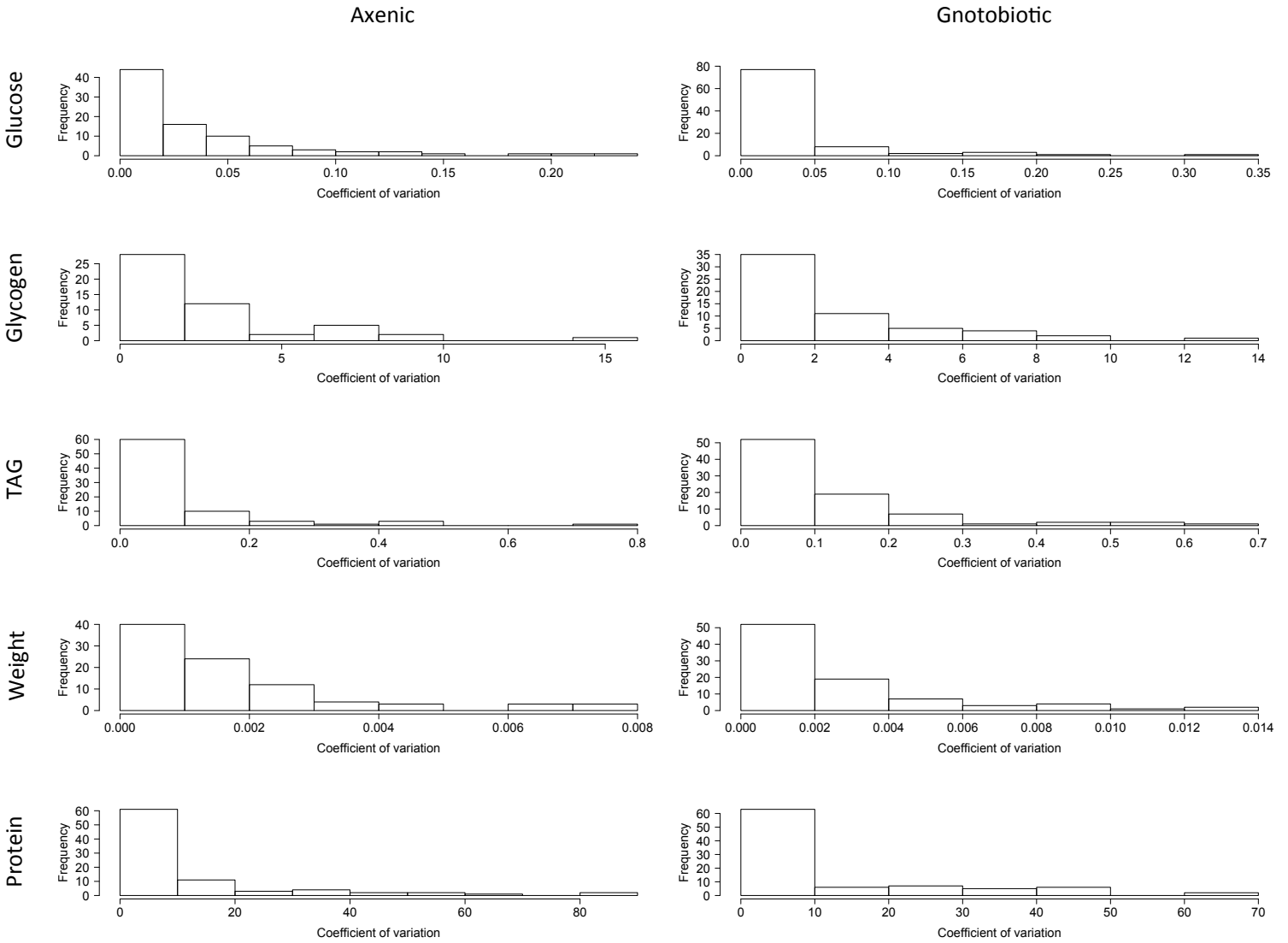
Supplementary Figures 1-7

Supplementary Tables 1-7

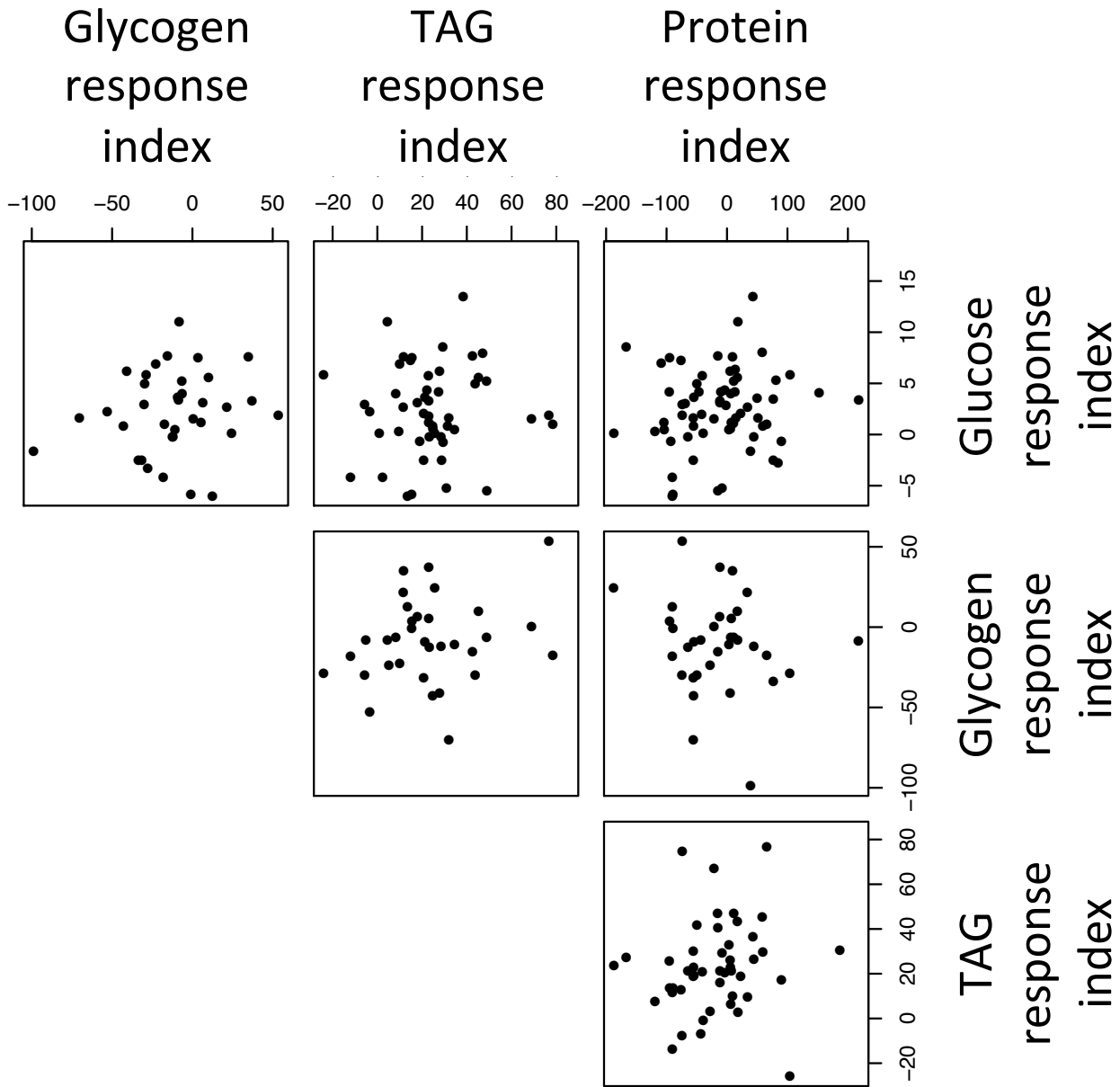
Supplementary Note 1



# Supplementary Figure 1

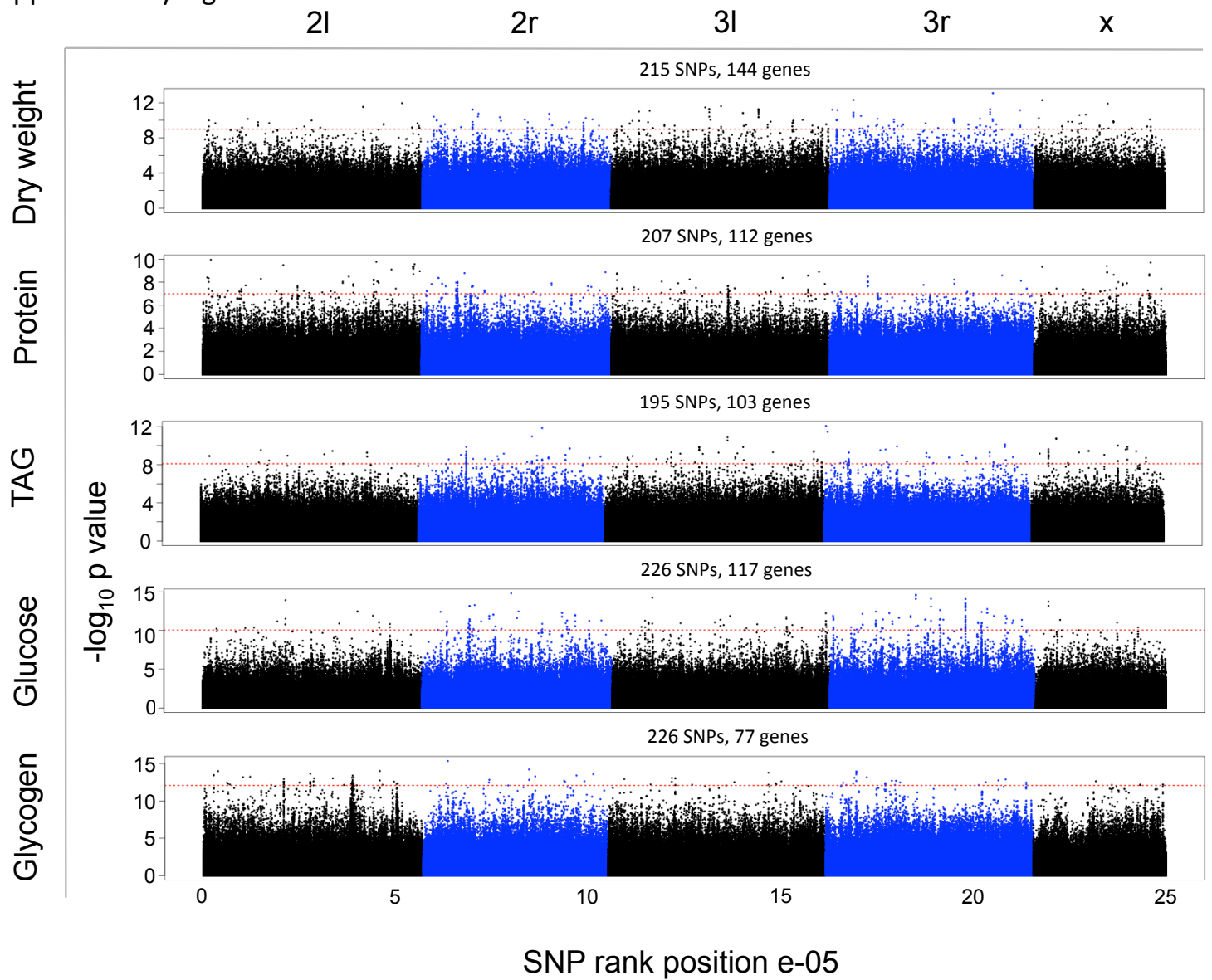


Supplementary Figure 1 . Coefficients of variation (variance/mean) for each DGRP line in each condition. Histograms show that within-line variance was generally low, and equivalent between gnotobiotic and axenic flies.



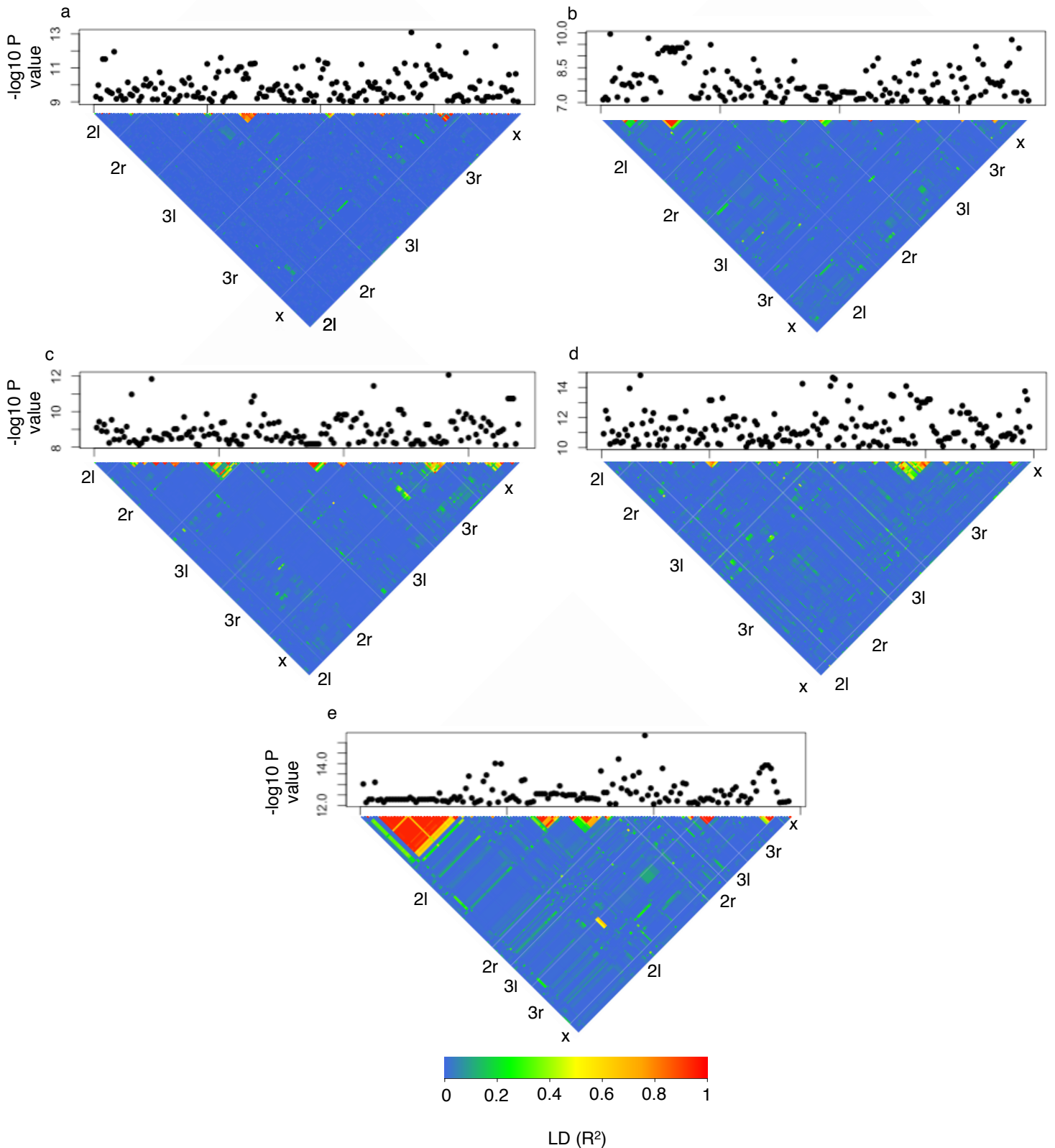
Supplementary Figure 2. Absence of elevated global sensitivity of *Drosophila* nutrition to microbiota. Response indices - calculated as the difference in trait value for each line in axenic or gnotobiotic conditions, normalized to mean population differences between axenic and gnotobiotic flies - are uncorrelated.

Supplementary Figure 3



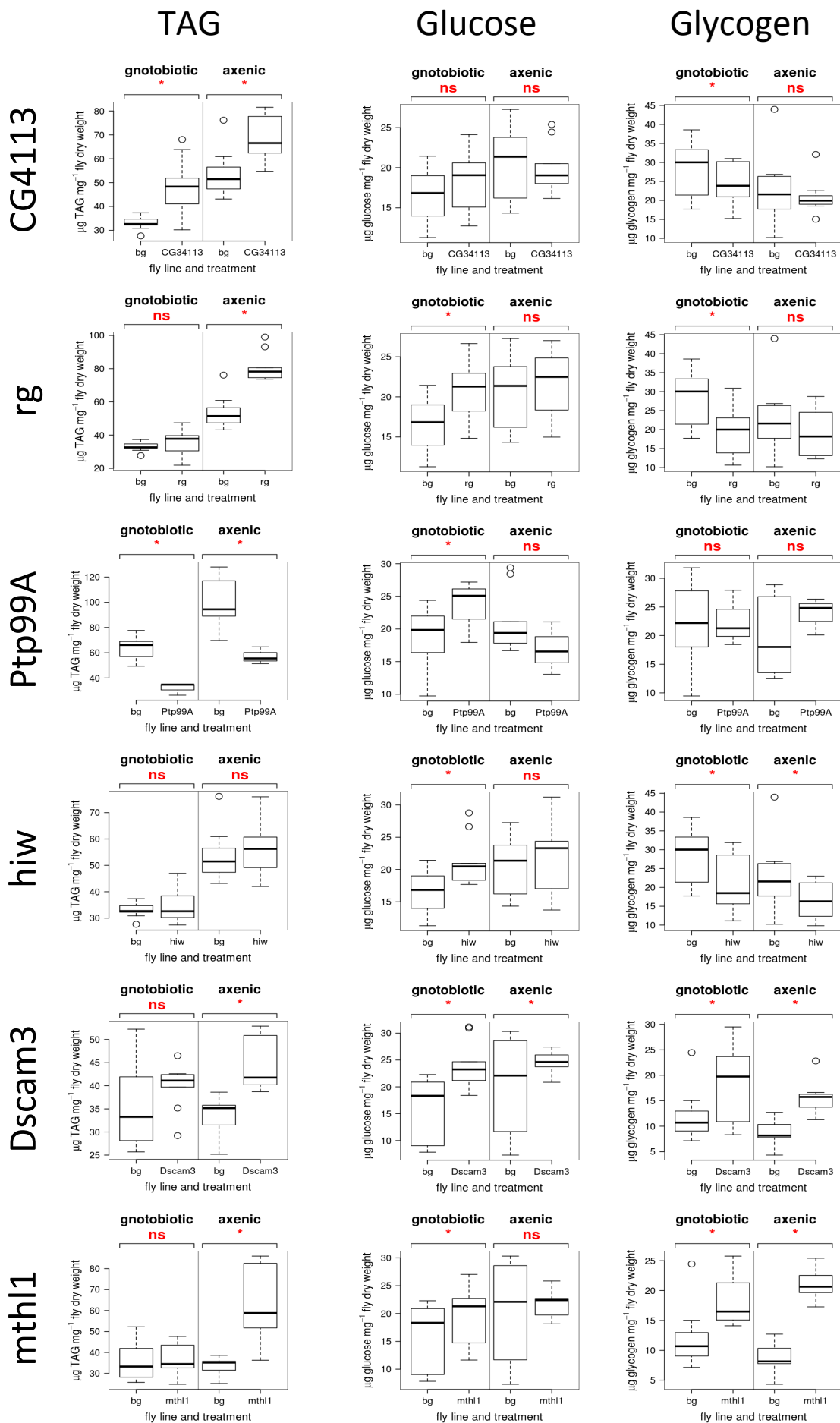
Supplementary Figure 3. Variation in microbiota-responsiveness of phenotypic traits maps to numerous SNPs throughout the *Drosophila* genome. Manhattan plots show  $-\log_{10}$  p-values of the interaction of biallelic SNP markers and microbiota treatment for each phenotypic trait.

Supplementary Figure 4



Supplementary Figure 4. Pairwise linkage disequilibrium (LD) among significant SNPs identified in GWAs. Pairwise LD was determined for all of the top 0.01% SNPs in GWAs for (a) dry weight; (b) protein; (c) TAG; (d) glucose content; (e) glycogen content. The distribution of SNPs (top panel); the heat map (bottom panel) is based on the LD measure  $R^2$  between SNPs, with the scale shown.

# Supplementary Figure 5a



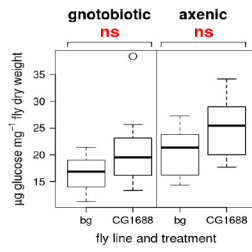
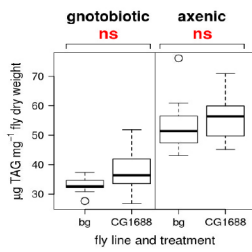
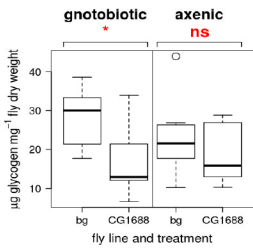
# Supplementary Figure 5b

## Glycogen

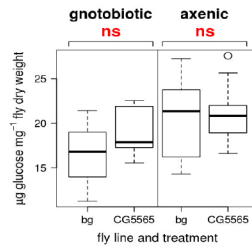
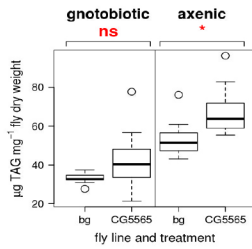
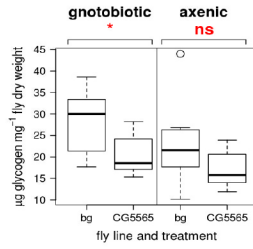
## TAG

## Glucose

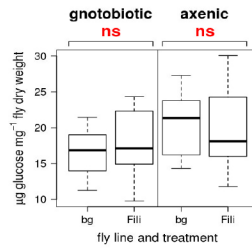
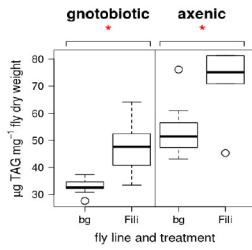
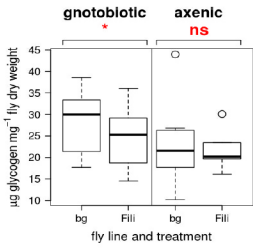
CG1688



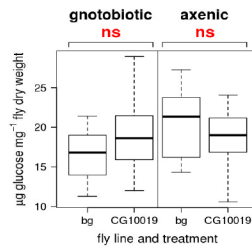
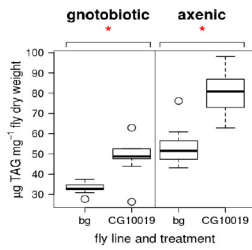
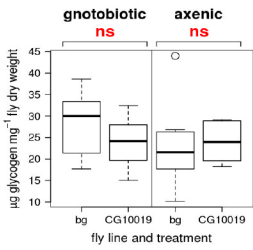
CG5565



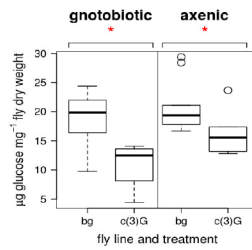
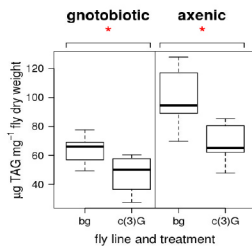
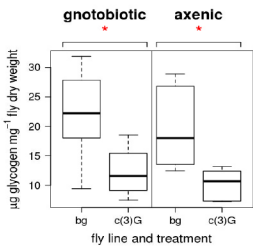
Fili



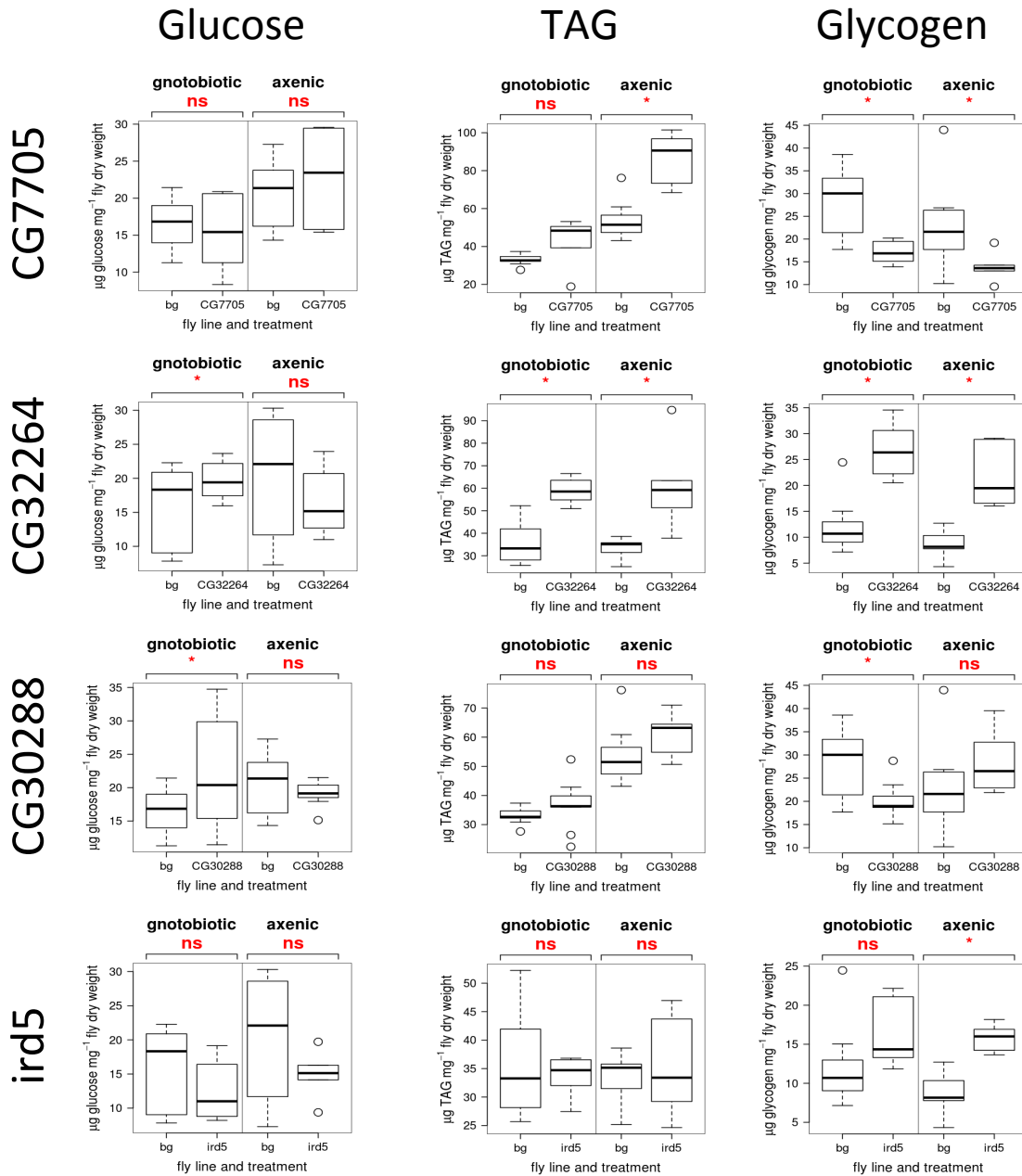
CG10019



c(3)G

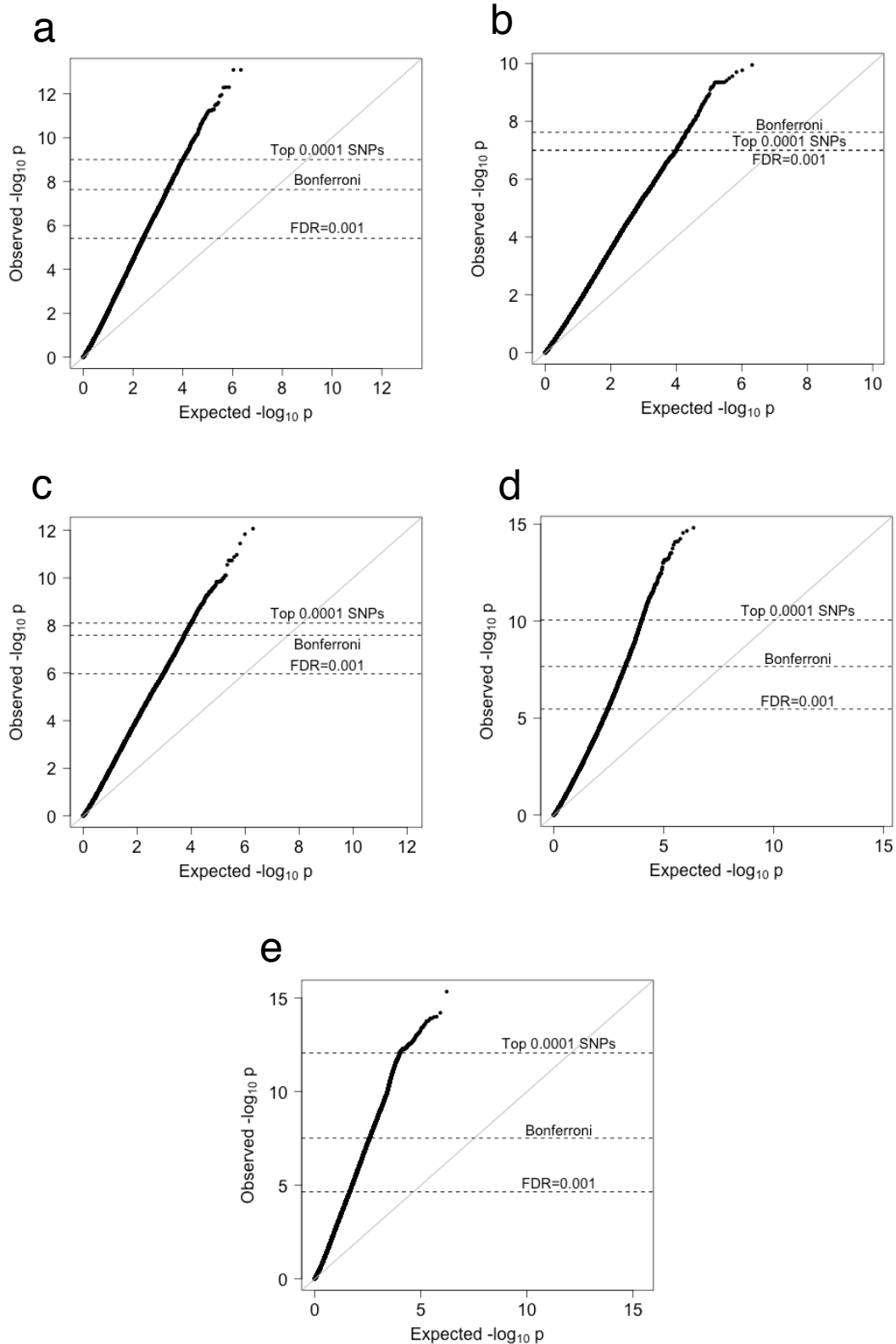


# Supplementary Figure 5c



Supplementary Figure 5 . Effect of gene mutations on response of *Drosophila* to microbiota treatments. Candidate genes identified by the GWA study were tested for (a) TAG, (b) glycogen content and (c) glucose content. Differences between the mutant and background (bg) stocks between axenic and gnotobiotic flies were tested by a linear mixed model with critical probability below 0.025 (\*, N=9, ns=no significant difference) to correct for two comparisons. A microbiota effect was accepted where there was a significant difference between the mutant and a background fly stock in only one microbial condition. ns denotes absence of a significant difference. To avoid confounding *Wolbachia* effects, background stocks and mutant stocks were matched for *Wolbachia*-status.

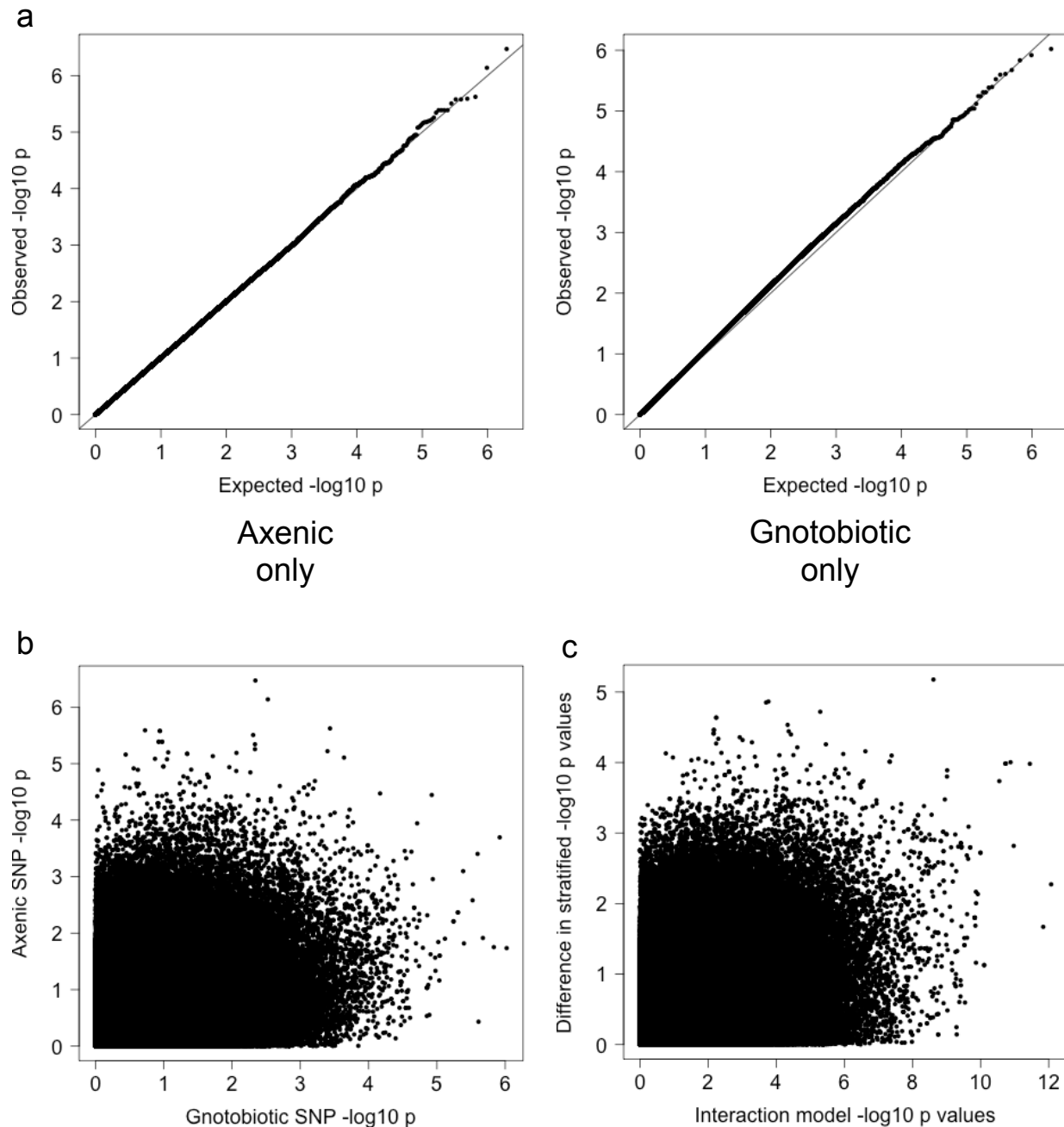
Supplementary Figure 6



Supplementary Figure 6. Quantile-quantile plots of genome-wide p-values ( $-\log_{10}$ ). P-values for all traits were inflated: observed p-values are inflated above null expectation for all traits. Horizontal dashed lines represent critical thresholds determined by 0.001% False Discovery Rate, Bonferroni correction, and the top 0.0001 proportion of SNPs (used in this study). Top 0.0001 proportion SNPs is the most conservative criterion for all traits under study except protein ( $-\log_{10}$  False Discovery Rate for protein=6.997, Bonferroni critical threshold=7.62, top 0.0001 proportion=7.000).



## Supplementary Figure 7



Supplementary Figure 7. Comparison of p-values ( $-\log_{10}$ ) from genome-wide association and effects of SNP in genome-wide treatment-stratified analysis. Example based on TAG mapping. P-values of SNP\*microbiota interaction were inflated (Supplementary Figure 5), but p-values for each SNP are not inflated when axenic and gnotobiotic samples were mapped separately in a treatment-stratified analysis (a). Treatment-stratified analysis can identify SNPs with treatment-specific effects (b), and some SNPs are identified by both modeling approaches (c). Both stratified and interaction-based analyses suggest some shared SNPs ( $r=0.32$ ,  $p<2.2e-16$ ), including SNPs in the gene *Rugose*, which was validated by mutant analysis (see Supplementary Note, Fig. 2, Table 1)

## Supplementary Tables 1-7

### Supplementary Table 1 Statistical analysis of phenotypic traits

Two approaches were adopted to model the data prior to GWA. The first approach (a) and (b) is statistically the most correct, but interpretation is difficult and power is low, particularly when considering the additional interaction of genetic polymorphisms by GWA. In this approach, an interaction of microbiota and *Wolbachia* status was fit as a fixed effect for dry weight, and a three-way interaction of dry weight, microbiota and *Wolbachia* status was fit as a fixed effect for the other traits. (These interactions were fitted because some traits showed relations to these predictors whilst others did not, and we wished to fit equivalent models to all traits.) For all models, genotype was nested in experimental block as a random effect. Full models are presented in (a). These models were simplified by stepwise removal of non-significant higher interactions, and the minimal models produced are presented in Table S1B. In the second approach (c), simpler models were fitted, in which microbiota, weight and *Wolbachia* status were additive, with genotype nested in experimental block as a random effect. The statistical analysis presented in Fig. 1b uses these models simplified by removal of non-significant factors.

#### (a) Full statistical models (all higher interactions)

	Term	DF	F	P	% variance explained*
Glucose	Microbiota	1,422	67.82	<0.0001	
	Wolbachia	1,101	0.74	0.39	
	Weight	1,422	12.62	<0.001	
	Microbiota * Wolbachia	1,422	0.76	0.38	
	Microbiota * weight	1,422	4.74	0.03	
	Wolbachia * weight	1,422	10.33	<0.002	
	Microbiota * Wolbachia *	1,422	0	0.9	
	Weight				
	Genotype				31.16
Glycogen	Microbiota	1,246	11.76	0.0007	
	Wolbachia	1,69	6.83	0.01	
	Weight	1,246	102.44	<0.0001	
	Microbiota * Wolbachia	1,246	6.22	0.01	
	Microbiota * weight	1,246	1.82	0.18	
	Wolbachia * weight	1,246	2.91	0.09	
	Microbiota * Wolbachia *	1,246	1.10	0.3	
	Weight				
	Genotype				53.85
Weight	Microbiota	1,423	90.03	<0.0001	
	Wolbachia	1,102	0.01	0.92	
	Microbiota * Wolbachia	1,423	0.006	0.94	
	Genotype				
TAG	Microbiota	1,371	229.94	<0.0001	
	Wolbachia	1,101	0	0.98	

	Weight	1,371	8.81	<0.005	
	Microbiota * Wolbachia	1,371	0.01	0.92	
	Microbiota * weight	1,371	0.18	0.67	
	Wolbachia * weight	1,371	0.72	0.40	
	Microbiota * Wolbachia *	1,371	1.71	0.19	
	Weight				
	Genotype				71.05
Protein	Microbiota	1,406	13.63	0.0003	
	Wolbachia	1,106	1.32	0.25	
	Weight	1,406	9.44	0.002	
	Microbiota * Wolbachia	1,406	0.30	0.6	
	Microbiota * weight	1,406	5.57	0.02	
	Wolbachia * weight	1,406	2.31	0.1	
	Microbiota * Wolbachia *	1,406	1.13	0.3	
	Weight				
	Genotype				72.41

(b) Minimal models from simplification of higher-level interactive models

Trait	Df	Effect	F	P	% variance explained*
Glucose	1,422	Microbiota	67.98	<0.0001	
	1,101	Wolbachia	0.75	0.39	
	1,422	Weight	12.67	<0.0005	
	1,422	Microbiota * weight	4.75	<0.03	
	1,422	Wolbachia * weight	10.35	<0.005	
	1,422	Genotype			
Glycogen	1,250	Weight	104.91	<0.0001	
	1,250	Microbiota	11.32	0.0009	
	1,69	Wolbachia	7.25	0.009	
		Genotype (random)			
Weight	1,424	Microbiota	90.27	<0.0001	
		Genotype (random)			
TAG	1,375	Weight	8.69	<0.005	
		Microbiota	231.46	<0.0001	
		Genotype (random)			
Protein	1,409	Microbiota	13.68	0.0002	
	1,409	Weight	9.20	0.003	
	1,409	Weight * Microbiota	5.53	0.02	
		Genotype (random)			

(c) Reduced trait models

Trait	Df	Effect	F	P	% variance explained*
Glucose	1,426	Microbiota	66.34	<0.0001	

	1,101	Wolbachia	0.71	0.4	
	1,426	Weight	12.10	0.0006	
		Genotype			30.75
Glycogen	1,250	Microbiota	11.32	0.0009	
	1,69	Wolbachia	7.34	0.009	
	1,250	Weight	104.91	<0.0001	
		Genotype			49.77
Weight	1,424	Microbiota	90.24	<0.0001	
	1,102	Wolbachia	0.01	0.9	
		Genotype			77.47
TAG	1,375	Microbiota	231.31	<0.0001	
	1,101	Wolbachia	0	0.98	
	1,375	Weight	8.56	<0.005	
		Genotype			71.63
Protein	1,410	Microbiota	13.59	0.0003	
	1,102	Wolbachia	1.41	0.2	
	1,410	Weight	9.39	0.24	
		Genotype			59.41

\* square of line (random effect) standard deviation

**Supplementary Table 2 Pairwise correlations among phenotypic traits in gnotobiotic and axenic flies.**

Correlations less significant than Bonferroni critical threshold ( $p < 0.008$ ) for six tests. Significant correlations are indicated in bold.

Trait <sup>1</sup>	Correlation coefficient (Pearson's r) and p value						
	Gnotobiotic flies		Axenic flies				
	Glycogen	TAG	Protein	Glycogen	TAG	Protein	
<i>Wolbachia</i> -positive flies	Glucose	0.356, p=0.013	<b>0.625,</b> <b>p&lt;0.001</b>	0.003, p=0.981	-0.119, p=0.448	0.047, p=0.753	-0.010, p=0.943
	Glycogen		<b>0.411,</b> <b>p=0.006</b>	-0.150, p=0.316		0.328, p=0.048	0.189, p=0.231
	TAG			0.196, p=0.155			0.141, p=0.366
<i>Wolbachia</i> -negative flies	Glucose	0.174, p=0.241	0.016, p=0.910	-0.201, p=0.137	0.260, p=0.074	-0.065, p=0.637	0.131, p=0.310
	Glycogen		0.346, p=0.017	-0.256, p=0.094		0.289, p=0.052	0.324, p=0.022
	TAG			0.317, p=0.026			0.283, p=0.033

<sup>1</sup> Nutritional traits expressed as µg per mg dry weight.

**Supplementary Table 3a.** SNPs significantly associated with dry weight.

Position (Chromosome arm, base)	Gene symbol	FlyBase ID	Site class	Bases from gene	p-value (-log10)
chrM_2L_1185490	Tgt	FBgn0031321	SYNONYMOUS_CODING	NA	9.95
chrM_2L_16084738	beat-1a	FBgn0013433	DOWNSTREAM	3050	9.76
chrM_x_19685987			SNP is more than 5000 bp away from known genes	NA	9.70
chrM_2L_20953314			SNP is more than 5000 bp away from known genes	NA	9.56
chrM_2L_7429548	CG5261	FBgn0031912	INTRON	NA	9.49
chrM_x_12965665	CG42629	FBgn0261388	INTRON	NA	9.41
chrM_2L_20795209	vari	FBgn0250785	INTRON	NA	9.35
chrM_2L_20797248	CG9328	FBgn0032886	UPSTREAM	469	9.35
chrM_2L_20797248	vari	FBgn0250785	UPSTREAM	208	9.35
chrM_2L_20798645	CG9328	FBgn0032886	INTRON	NA	9.35
chrM_2L_20800885	CG9328	FBgn0032886	INTRON	NA	9.35
chrM_2L_20802914	CG9328	FBgn0032886	INTRON	NA	9.35
chrM_2L_20803879	CG9328	FBgn0032886	INTRON	NA	9.35
chrM_x_2423941	trol	FBgn0261451	INTRON	NA	9.33
chrM_2L_20792942	vari	FBgn0250785	INTRON	NA	9.24
chrM_2L_20797373	vari	FBgn0250785	UPSTREAM	333	9.18
chrM_2L_20797373	CG9328	FBgn0032886	UPSTREAM	344	9.18
chrM_2L_20799189	CG9328	FBgn0032886	INTRON	NA	9.17
chrM_2L_17783645	CadN2	FBgn0262018	DOWNSTREAM	3633	9.10
chrM_2L_22327628	CG17018	FBgn0039972	INTRON	NA	8.95
chrM_3L_20360659	in	FBgn0001259	NON_SYNONYMOUS_CODING	NA	8.90
chrM_2R_20115159	CG30419	FBgn0050419	INTRON	NA	8.86
chrM_x_13007135	CG42237	FBgn0250862	UPSTREAM	2035	8.85
chrM_x_13007135	CG43313	FBgn0263005	UPSTREAM	660	8.85
chrM_2R_7328465	CG7777	FBgn0033635	INTRON	NA	8.79
chrM_3L_829321	CG13900	FBgn0035162	SYNONYMOUS_CODING	NA	8.77
chrM_2L_20804507	CG9328	FBgn0032886	INTRON	NA	8.70
chrM_x_19500414	CG14207	FBgn0031037	INTRON	NA	8.69
chrM_3L_829204	CG13900	FBgn0035162	SYNONYMOUS_CODING	NA	8.68
chrM_x_14071021	CG11581	FBgn0030540	SYNONYMOUS_CODING	NA	8.64
chrM_3R_24599729			SNP is more than 5000 bp away from known genes	NA	8.59
chrM_x_19497297	CG14220	FBgn0031036	UPSTREAM	1414	8.58
chrM_x_19497297	CG14207	FBgn0031037	UPSTREAM	1894	8.58
chrM_3L_18488333	AICR2	FBgn0036789	INTRON	NA	8.57
chrM_3R_9161426	CG8784	FBgn0038140	UPSTREAM	4259	8.48
chrM_3R_9161426	CG8795	FBgn0038139	UPSTREAM	1541	8.48
chrM_2L_13104688	CG5142	FBgn0032470	UPSTREAM	2183	8.48

chrM_2l_863637	aru	FBgn0029095	SYNONYMOUS_CODING	NA	8.41
chrM_3l_15654441	Eig71Ek	FBgn0014851	INTRON	NA	8.37
chrM_2r_4114655	LRP1	FBgn0053087	DOWNSTREAM	712	8.36
chrM_2r_4114655	CG34217	FBgn0085246	DOWNSTREAM	840	8.36
chrM_2l_970579			SNP is more than 5000 bp away from known genes	NA	8.35
chrM_2l_5505633			SNP is more than 5000 bp away from known genes	NA	8.29
chrM_3l_7321096	CG8602	FBgn0035763	SYNONYMOUS_CODING	NA	8.25
chrM_3r_18966517	hh	FBgn0004644	INTRON	NA	8.23
chrM_2l_15766186	CG3793	FBgn0028507	UTR_3_PRIME	NA	8.19
chrM_2l_15766186	wek	FBgn0001990	UTR_3_PRIME	NA	8.19
chrM_2l_15766316	wek	FBgn0001990	UTR_3_PRIME	NA	8.19
chrM_3l_829318	CG13900	FBgn0035162	SYNONYMOUS_CODING	NA	8.18
chrM_2l_15766217	CG3793	FBgn0028507	UTR_3_PRIME	NA	8.17
chrM_2l_15766217	wek	FBgn0001990	UTR_3_PRIME	NA	8.17
chrM_3r_26256962	CG15533	FBgn0039768	NON_SYNONYMOUS_CODING	NA	8.12
chrM_2l_16194853	Ca-alpha1D	FBgn0001991	DOWNSTREAM	4774	8.07
chrM_3l_4630926	axo	FBgn0262870	INTRON	NA	8.06
chrM_3r_9161779	CG8784	FBgn0038140	UPSTREAM	3906	8.06
chrM_3r_9161779	CG8795	FBgn0038139	UPSTREAM	1894	8.06
chrM_2l_16258204	EndoGI	FBgn0028515	INTRON	NA	8.05
chrM_2l_12468829	bun	FBgn0259176	INTRON	NA	8.05
chrM_2l_976907			SNP is more than 5000 bp away from known genes	NA	8.02
chrM_2r_6734164	CG30015	FBgn0050015	INTRON	NA	8.02
chrM_3r_9161655	CG8784	FBgn0038140	UPSTREAM	4030	8.01
chrM_3r_9161655	CG8795	FBgn0038139	UPSTREAM	1770	8.01
chrM_x_14709844	rut	FBgn0003301	INTRON	NA	8.01
chrM_2r_6578432	stan	FBgn0024836	INTRON	NA	7.96
chrM_2l_16362492	jhamt	FBgn0028841	UPSTREAM	3427	7.94
chrM_2r_6734151	CG30015	FBgn0050015	INTRON	NA	7.93
chrM_2l_12468777	bun	FBgn0259176	INTRON	NA	7.92
chrM_x_16074471			SNP is more than 5000 bp away from known genes	NA	7.91
chrM_3l_6254829	ImpL3	FBgn0001258	INTRON	NA	7.90
chrM_2r_15450499	sm	FBgn0003435	INTRON	NA	7.90
chrM_3r_18966534	hh	FBgn0004644	INTRON	NA	7.87
chrM_x_15275362	CG9164	FBgn0030634	UTR_3_PRIME	NA	7.86
chrM_x_15275367	CG9164	FBgn0030634	UTR_3_PRIME	NA	7.86
chrM_2r_4681981	sns	FBgn0024189	UPSTREAM	4041	7.86
chrM_2r_4681981	CG8746	FBgn0033330	DOWNSTREAM	699	7.86
chrM_3l_22163688	olf413	FBgn0037153	INTRON	NA	7.84
chrM_2r_2796520	CG30158	FBgn0050158	INTRON	NA	7.84

chrX_15275732	CG9164	FBgn0030634	INTRON	NA	7.83
chr2L_15766306	wek	FBgn0001990	UTR_3_PRIME	NA	7.83
chr2L_13608642	kuz	FBgn0259984	INTRON	NA	7.82
chr2L_13608590	kuz	FBgn0259984	INTRON	NA	7.78
chrX_15275419	CG9164	FBgn0030634	SYNONYMOUS_CODING	NA	7.78
chr3R_15320159	Dys	FBgn0260003	INTRON	NA	7.76
chr2L_13608604	kuz	FBgn0259984	INTRON	NA	7.75
chr2R_15460277	sm	FBgn0003435	INTRON	NA	7.74
chr2L_4051281	ed	FBgn0000547	INTRON	NA	7.72
chrX_15276037	CG9164	FBgn0030634	SYNONYMOUS_CODING	NA	7.71
chr2R_6728017	CG30015	FBgn0050015	SYNONYMOUS_CODING	NA	7.70
chr2L_13608651	kuz	FBgn0259984	INTRON	NA	7.68
chr3L_10615226	CG43127	FBgn0262592	DOWNSTREAM	2602	7.67
chr3L_10615226	CG42521	FBgn0260396	DOWNSTREAM	940	7.67
chr3L_10615235	CG43127	FBgn0262592	DOWNSTREAM	2593	7.67
chr3L_10615235	CG42521	FBgn0260396	DOWNSTREAM	931	7.67
chr3L_10615239	CG43127	FBgn0262592	DOWNSTREAM	2589	7.67
chr3L_10615239	CG42521	FBgn0260396	DOWNSTREAM	927	7.67
chr3R_9162157	CG8784	FBgn0038140	UPSTREAM	3528	7.65
chr3R_9162157	CG8795	FBgn0038139	UPSTREAM	2272	7.65
chr2R_18510065	CG33143	FBgn0053143	INTRON	NA	7.65
chr2R_8238733	ana3	FBgn0033718	DOWNSTREAM	150	7.65
chr2R_8238733	CG30047	FBgn0050047	UPSTREAM	565	7.65
chr2L_8638991	Sema-1a	FBgn0011259	INTRON	NA	7.65
chr2R_5298432	wun	FBgn0016078	INTRON	NA	7.64
chr2R_8130048	CG43316	FBgn0263021	UPSTREAM	3	7.61
chr2R_8130048	CG43315	FBgn0263020	DOWNSTREAM	784	7.61
chr2R_8133105	CG43316	FBgn0263021	DOWNSTREAM	2327	7.60
chr2R_8133105	CG43244	FBgn0262889	UPSTREAM	938	7.60
chr2R_8133171	CG43244	FBgn0262889	UPSTREAM	1004	7.59
chr2R_8133171	CG43316	FBgn0263021	DOWNSTREAM	2393	7.59
chr2R_18785975	nahoda	FBgn0034797	INTRON	NA	7.59
chr3L_2220998			SNP is more than 5000 bp away from known genes	NA	7.55
chr2L_8639036	Sema-1a	FBgn0011259	INTRON	NA	7.53
chr3L_7849597	Pdp1	FBgn0016694	INTRON	NA	7.51
chr3L_11101797	CG12522	FBgn0036131	UTR_3_PRIME	NA	7.47
chr3L_829351	CG13900	FBgn0035162	SYNONYMOUS_CODING	NA	7.47
chr3L_829360	CG13900	FBgn0035162	SYNONYMOUS_CODING	NA	7.47
chrX_14404284			SNP is more than 5000 bp away from known genes	NA	7.46
chr3L_11101725	CG12522	FBgn0036131	NON_SYNONYMOUS_CODING	NA	7.45
chr2R_6730132	CG30015	FBgn0050015	SYNONYMOUS_CODING	NA	7.44
chrX_4618689	Proc-R	FBgn0029723	INTRON	NA	7.44
chr2L_18381797	Fas3	FBgn0000636	INTRON	NA	7.44



chr3r_26729403	Cyp4c3	FBgn0015032	DOWNSTREAM	2993	7.43
chr3r_26729403	CG33483	FBgn0053483	UPSTREAM	1520	7.43
chr2l_970899			SNP is more than 5000 bp away from known genes	NA	7.43
chr3l_1277218	CG32333	FBgn0052333	INTRON	NA	7.43
chr3l_1277225	CG32333	FBgn0052333	INTRON	NA	7.43
chr3l_1277226	CG32333	FBgn0052333	INTRON	NA	7.43
chr3l_1277227	CG32333	FBgn0052333	INTRON	NA	7.43
chrx_21315361	CG42343	FBgn0259245	INTRON	NA	7.43
chrx_21315362	CG42343	FBgn0259245	INTRON	NA	7.43
chr2r_6570464	stan	FBgn0024836	INTRON	NA	7.42
chr2l_3711191	cutlet	FBgn0015376	NON_SYNONYMOUS_CODING	NA	7.41
chr3l_10614430	CG43127	FBgn0262592	DOWNSTREAM	3398	7.40
chr3l_10614430	CG42521	FBgn0260396	DOWNSTREAM	1736	7.40
chr3l_2574451	msn	FBgn0010909	INTRON	NA	7.40
chr3l_10616787	CG43127	FBgn0262592	DOWNSTREAM	1041	7.39
chr3l_10616787	CG42521	FBgn0260396	UPSTREAM	111	7.39
chr3l_5479057	CG4835	FBgn0035607	DOWNSTREAM	2294	7.37
chr3l_18488284	AICR2	FBgn0036789	INTRON	NA	7.36
chrx_8916007	rdgA	FBgn0261549	INTRON	NA	7.36
chrx_8916007	CR43836	FBgn0264384	INTRON	NA	7.36
chrx_13408215	Neto	FBgn0052635	INTRON	NA	7.32
chr3l_10631766	CG8009	FBgn0036090	SYNONYMOUS_CODING	NA	7.31
chr3l_2220902			SNP is more than 5000 bp away from known genes	NA	7.31
chr2l_8639111	Sema-1a	FBgn0011259	INTRON	NA	7.30
chr2r_18509918	CG33143	FBgn0053143	INTRON	NA	7.30
chr2r_6570419	stan	FBgn0024836	INTRON	NA	7.29
chr2r_19462480	apt	FBgn0015903	INTRON	NA	7.29
chr2r_9279637	CG43691	FBgn0263774	NON_SYNONYMOUS_CODING	NA	7.28
chrx_19293197	kek5	FBgn0031016	INTRON	NA	7.27
chr3l_18488329	AICR2	FBgn0036789	INTRON	NA	7.27
chr2l_2585503	Drp1	FBgn0026479	UPSTREAM	55	7.25
chr2l_2585503	CG15394	FBgn0250835	UPSTREAM	3351	7.25
chr2r_6570465	stan	FBgn0024836	INTRON	NA	7.25
chrx_12656071	CG12717	FBgn0030420	SYNONYMOUS_CODING	NA	7.25
chr2l_9135470	CG32982	FBgn0052982	INTRON	NA	7.24
chr2r_9279659	CG43691	FBgn0263774	SYNONYMOUS_CODING	NA	7.24
chr3l_10707096			SNP is more than 5000 bp away from known genes	NA	7.23
chr3l_10707097			SNP is more than 5000 bp away from known genes	NA	7.23
chr2l_1296034	robo3	FBgn0041097	DOWNSTREAM	1347	7.23
chr3l_6903932	Prat2	FBgn0041194	DOWNSTREAM	1167	7.23
chr2l_11397558			SNP is more than 5000 bp away	NA	7.22

			from known genes		
chrM_2r_1563527	scaf	FBgn0033033	DOWNSTREAM	727	7.22
chrM_2r_1563527	kune	FBgn0033032	DOWNSTREAM	3225	7.22
chrM_2r_6575211	stan	FBgn0024836	INTRON	NA	7.22
chrM_2l_6843244	sens-2	FBgn0051632	DOWNSTREAM	4239	7.21
chrM_x_12001735	CR43960	FBgn0264675	INTRON	NA	7.19
chrM_x_12001735	CG2750	FBgn0030376	INTRON	NA	7.19
chrM_2l_3711118	cutlet	FBgn0015376	SYNONYMOUS_CODING	NA	7.19
chrM_2l_3711130	cutlet	FBgn0015376	SYNONYMOUS_CODING	NA	7.19
chrM_2r_6729331	CG30015	FBgn0050015	SYNONYMOUS_CODING	NA	7.19
chrM_2l_3575889	CG43815	FBgn0264363	UPSTREAM	321	7.18
chrM_2l_3575889	CR43822	FBgn0264370	DOWNSTREAM	83	7.18
chrM_3r_20665770	CG13636	FBgn0039232	UTR_3_PRIME	NA	7.17
chrM_3l_855884	dpr20	FBgn0035170	INTRON	NA	7.16
chrM_3l_14205832	nuf	FBgn0013718	INTRON	NA	7.16
chrM_3l_14205829	nuf	FBgn0013718	INTRON	NA	7.16
chrM_3l_14205831	nuf	FBgn0013718	INTRON	NA	7.16
chrM_3r_3726146	wtrw	FBgn0260005	SYNONYMOUS_CODING	NA	7.15
chrM_2l_16054783	beat-1a	FBgn0013433	INTRON	NA	7.15
chrM_3l_10632370	CG8009	FBgn0036090	SYNONYMOUS_CODING	NA	7.14
chrM_2r_13812178	elk	FBgn0011589	SYNONYMOUS_CODING	NA	7.13
chrM_2l_15805377			SNP is more than 5000 bp away from known genes	NA	7.13
chrM_2r_9279672	CG43691	FBgn0263774	NON_SYNONYMOUS_CODING	NA	7.12
chrM_2l_10586312	Trim9	FBgn0051721	INTRON	NA	7.12
chrM_2l_11770148			SNP is more than 5000 bp away from known genes	NA	7.12
chrM_3l_13766739	bru-3	FBgn0264001	INTRON	NA	7.12
chrM_3l_10004092	dpr6	FBgn0040823	INTRON	NA	7.12
chrM_3l_10004108	dpr6	FBgn0040823	INTRON	NA	7.12
chrM_3r_1270839	CG14669	FBgn0037326	INTRON	NA	7.12
chrM_3l_18488341	AICR2	FBgn0036789	INTRON	NA	7.12
chrM_2l_13608629	kuz	FBgn0259984	INTRON	NA	7.10
chrM_3l_6903985	Prat2	FBgn0041194	DOWNSTREAM	1114	7.09
chrM_3l_4621795	CG32246	FBgn0052246	SYNONYMOUS_CODING	NA	7.09
chrM_x_9214582	c12.2	FBgn0040234	INTRON	NA	7.08
chrM_2l_9513712	CG33298	FBgn0032120	INTRON	NA	7.08
chrM_x_3538650	AlstR	FBgn0028961	INTRON	NA	7.07
chrM_3r_20559590	tok	FBgn0004885	INTRON	NA	7.06
chrM_3l_6574746	jv	FBgn0263973	INTRON	NA	7.05
chrM_3l_6574746	CG18769	FBgn0042185	INTRON	NA	7.05
chrM_3r_21380698	msi	FBgn0011666	INTRON	NA	7.03
chrM_3r_10187030	CG31320	FBgn0051320	UTR_5_PRIME	NA	7.02
chrM_2r_6729208	CG30015	FBgn0050015	SYNONYMOUS_CODING	NA	7.01
chrM_3r_23781753	CG5527	FBgn0039564	SYNONYMOUS_CODING	NA	7.01

chr <sub>m</sub> _2r_5298746	wun	FBgn0016078	INTRON	NA	7.00
chr <sub>m</sub> _3l_10631736	CG8009	FBgn0036090	SYNONYMOUS_CODING	NA	7.00
chr <sub>m</sub> _x_10073599	CG43902	FBgn0264503	INTRON	NA	7.00
chr <sub>m</sub> _3l_22916473	CG12768	FBgn0037206	UTR_3_PRIME	NA	7.00

**Supplementary Table 3b.** SNPs significantly associated with protein.

Position (Chromosome arm, base)	Gene symbol	FlyBase ID	Site class	Bases from gene	p-value (-log10)
chrM_2L_1185490	Tgt	FBgn0031321	SYNONYMOUS_CODING	NA	9.95
chrM_2L_16084738	beat-1a	FBgn0013433	DOWNSTREAM	3050	9.76
chrM_x_19685987			SNP is more than 5000 bp away from known genes	NA	9.70
chrM_2L_20953314			SNP is more than 5000 bp away from known genes	NA	9.56
chrM_2L_7429548	CG5261	FBgn0031912	INTRON	NA	9.49
chrM_x_12965665	CG42629	FBgn0261388	INTRON	NA	9.41
chrM_2L_20795209	vari	FBgn0250785	INTRON	NA	9.35
chrM_2L_20797248	CG9328	FBgn0032886	UPSTREAM	469	9.35
chrM_2L_20797248	vari	FBgn0250785	UPSTREAM	208	9.35
chrM_2L_20798645	CG9328	FBgn0032886	INTRON	NA	9.35
chrM_2L_20800885	CG9328	FBgn0032886	INTRON	NA	9.35
chrM_2L_20802914	CG9328	FBgn0032886	INTRON	NA	9.35
chrM_2L_20803879	CG9328	FBgn0032886	INTRON	NA	9.35
chrM_x_2423941	trol	FBgn0261451	INTRON	NA	9.33
chrM_2L_20792942	vari	FBgn0250785	INTRON	NA	9.24
chrM_2L_20797373	vari	FBgn0250785	UPSTREAM	333	9.18
chrM_2L_20797373	CG9328	FBgn0032886	UPSTREAM	344	9.18
chrM_2L_20799189	CG9328	FBgn0032886	INTRON	NA	9.17
chrM_2L_17783645	CadN2	FBgn0262018	DOWNSTREAM	3633	9.10
chrM_2L_22327628	CG17018	FBgn0039972	INTRON	NA	8.95
chrM_3L_20360659	in	FBgn0001259	NON_SYNONYMOUS_CODING	NA	8.90
chrM_2R_20115159	CG30419	FBgn0050419	INTRON	NA	8.86
chrM_x_13007135	CG42237	FBgn0250862	UPSTREAM	2035	8.85
chrM_x_13007135	CG43313	FBgn0263005	UPSTREAM	660	8.85
chrM_2R_7328465	CG7777	FBgn0033635	INTRON	NA	8.79
chrM_3L_829321	CG13900	FBgn0035162	SYNONYMOUS_CODING	NA	8.77
chrM_2L_20804507	CG9328	FBgn0032886	INTRON	NA	8.70
chrM_x_19500414	CG14207	FBgn0031037	INTRON	NA	8.69
chrM_3L_829204	CG13900	FBgn0035162	SYNONYMOUS_CODING	NA	8.68
chrM_x_14071021	CG11581	FBgn0030540	SYNONYMOUS_CODING	NA	8.64
chrM_3R_24599729			SNP is more than 5000 bp away from known genes	NA	8.59
chrM_x_19497297	CG14220	FBgn0031036	UPSTREAM	1414	8.58
chrM_x_19497297	CG14207	FBgn0031037	UPSTREAM	1894	8.58
chrM_3L_18488333	AICR2	FBgn0036789	INTRON	NA	8.57
chrM_3R_9161426	CG8784	FBgn0038140	UPSTREAM	4259	8.48
chrM_3R_9161426	CG8795	FBgn0038139	UPSTREAM	1541	8.48
chrM_2L_13104688	CG5142	FBgn0032470	UPSTREAM	2183	8.48

chrn_2l_863637	aru	FBgn0029095	SYNONYMOUS_CODING	NA	8.41
chrn_3l_15654441	Eig71Ek	FBgn0014851	INTRON	NA	8.37
chrn_2r_4114655	LRP1	FBgn0053087	DOWNSTREAM	712	8.36
chrn_2r_4114655	CG34217	FBgn0085246	DOWNSTREAM	840	8.36
chrn_2l_970579			SNP is more than 5000 bp away from known genes	NA	8.35
chrn_2l_5505633			SNP is more than 5000 bp away from known genes	NA	8.29
chrn_3l_7321096	CG8602	FBgn0035763	SYNONYMOUS_CODING	NA	8.25
chrn_3r_18966517	hh	FBgn0004644	INTRON	NA	8.23
chrn_2l_15766186	CG3793	FBgn0028507	UTR_3_PRIME	NA	8.19
chrn_2l_15766186	wek	FBgn0001990	UTR_3_PRIME	NA	8.19
chrn_2l_15766316	wek	FBgn0001990	UTR_3_PRIME	NA	8.19
chrn_3l_829318	CG13900	FBgn0035162	SYNONYMOUS_CODING	NA	8.18
chrn_2l_15766217	CG3793	FBgn0028507	UTR_3_PRIME	NA	8.17
chrn_2l_15766217	wek	FBgn0001990	UTR_3_PRIME	NA	8.17
chrn_3r_26256962	CG15533	FBgn0039768	NON_SYNONYMOUS_CODING	NA	8.12
chrn_2l_16194853	Ca-alpha1D	FBgn0001991	DOWNSTREAM	4774	8.07
chrn_3l_4630926	axo	FBgn0262870	INTRON	NA	8.06
chrn_3r_9161779	CG8784	FBgn0038140	UPSTREAM	3906	8.06
chrn_3r_9161779	CG8795	FBgn0038139	UPSTREAM	1894	8.06
chrn_2l_16258204	EndoGI	FBgn0028515	INTRON	NA	8.05
chrn_2l_12468829	bun	FBgn0259176	INTRON	NA	8.05
chrn_2l_976907			SNP is more than 5000 bp away from known genes	NA	8.02
chrn_2r_6734164	CG30015	FBgn0050015	INTRON	NA	8.02
chrn_3r_9161655	CG8784	FBgn0038140	UPSTREAM	4030	8.01
chrn_3r_9161655	CG8795	FBgn0038139	UPSTREAM	1770	8.01
chrn_x_14709844	rut	FBgn0003301	INTRON	NA	8.01
chrn_2r_6578432	stan	FBgn0024836	INTRON	NA	7.96
chrn_2l_16362492	jhamt	FBgn0028841	UPSTREAM	3427	7.94
chrn_2r_6734151	CG30015	FBgn0050015	INTRON	NA	7.93
chrn_2l_12468777	bun	FBgn0259176	INTRON	NA	7.92
chrn_x_16074471			SNP is more than 5000 bp away from known genes	NA	7.91
chrn_3l_6254829	ImpL3	FBgn0001258	INTRON	NA	7.90
chrn_2r_15450499	sm	FBgn0003435	INTRON	NA	7.90
chrn_3r_18966534	hh	FBgn0004644	INTRON	NA	7.87
chrn_x_15275362	CG9164	FBgn0030634	UTR_3_PRIME	NA	7.86
chrn_x_15275367	CG9164	FBgn0030634	UTR_3_PRIME	NA	7.86
chrn_2r_4681981	sns	FBgn0024189	UPSTREAM	4041	7.86
chrn_2r_4681981	CG8746	FBgn0033330	DOWNSTREAM	699	7.86
chrn_3l_22163688	olf413	FBgn0037153	INTRON	NA	7.84
chrn_2r_2796520	CG30158	FBgn0050158	INTRON	NA	7.84

chrX_15275732	CG9164	FBgn0030634	INTRON	NA	7.83
chr2L_15766306	wek	FBgn0001990	UTR_3_PRIME	NA	7.83
chr2L_13608642	kuz	FBgn0259984	INTRON	NA	7.82
chr2L_13608590	kuz	FBgn0259984	INTRON	NA	7.78
chrX_15275419	CG9164	FBgn0030634	SYNONYMOUS_CODING	NA	7.78
chr3R_15320159	Dys	FBgn0260003	INTRON	NA	7.76
chr2L_13608604	kuz	FBgn0259984	INTRON	NA	7.75
chr2R_15460277	sm	FBgn0003435	INTRON	NA	7.74
chr2L_4051281	ed	FBgn0000547	INTRON	NA	7.72
chrX_15276037	CG9164	FBgn0030634	SYNONYMOUS_CODING	NA	7.71
chr2R_6728017	CG30015	FBgn0050015	SYNONYMOUS_CODING	NA	7.70
chr2L_13608651	kuz	FBgn0259984	INTRON	NA	7.68
chr3L_10615226	CG43127	FBgn0262592	DOWNSTREAM	2602	7.67
chr3L_10615226	CG42521	FBgn0260396	DOWNSTREAM	940	7.67
chr3L_10615235	CG43127	FBgn0262592	DOWNSTREAM	2593	7.67
chr3L_10615235	CG42521	FBgn0260396	DOWNSTREAM	931	7.67
chr3L_10615239	CG43127	FBgn0262592	DOWNSTREAM	2589	7.67
chr3L_10615239	CG42521	FBgn0260396	DOWNSTREAM	927	7.67
chr3R_9162157	CG8784	FBgn0038140	UPSTREAM	3528	7.65
chr3R_9162157	CG8795	FBgn0038139	UPSTREAM	2272	7.65
chr2R_18510065	CG33143	FBgn0053143	INTRON	NA	7.65
chr2R_8238733	ana3	FBgn0033718	DOWNSTREAM	150	7.65
chr2R_8238733	CG30047	FBgn0050047	UPSTREAM	565	7.65
chr2L_8638991	Sema-1a	FBgn0011259	INTRON	NA	7.65
chr2R_5298432	wun	FBgn0016078	INTRON	NA	7.64
chr2R_8130048	CG43316	FBgn0263021	UPSTREAM	3	7.61
chr2R_8130048	CG43315	FBgn0263020	DOWNSTREAM	784	7.61
chr2R_8133105	CG43316	FBgn0263021	DOWNSTREAM	2327	7.60
chr2R_8133105	CG43244	FBgn0262889	UPSTREAM	938	7.60
chr2R_8133171	CG43244	FBgn0262889	UPSTREAM	1004	7.59
chr2R_8133171	CG43316	FBgn0263021	DOWNSTREAM	2393	7.59
chr2R_18785975	nahoda	FBgn0034797	INTRON	NA	7.59
chr3L_2220998			SNP is more than 5000 bp away from known genes	NA	7.55
chr2L_8639036	Sema-1a	FBgn0011259	INTRON	NA	7.53
chr3L_7849597	Pdp1	FBgn0016694	INTRON	NA	7.51
chr3L_11101797	CG12522	FBgn0036131	UTR_3_PRIME	NA	7.47
chr3L_829351	CG13900	FBgn0035162	SYNONYMOUS_CODING	NA	7.47
chr3L_829360	CG13900	FBgn0035162	SYNONYMOUS_CODING	NA	7.47
chrX_14404284			SNP is more than 5000 bp away from known genes	NA	7.46
chr3L_11101725	CG12522	FBgn0036131	NON_SYNONYMOUS_CODING	NA	7.45
chr2R_6730132	CG30015	FBgn0050015	SYNONYMOUS_CODING	NA	7.44
chrX_4618689	Proc-R	FBgn0029723	INTRON	NA	7.44
chr2L_18381797	Fas3	FBgn0000636	INTRON	NA	7.44

chr3r_26729403	Cyp4c3	FBgn0015032	DOWNSTREAM	2993	7.43
chr3r_26729403	CG33483	FBgn0053483	UPSTREAM	1520	7.43
chr2l_970899			SNP is more than 5000 bp away from known genes	NA	7.43
chr3l_1277218	CG32333	FBgn0052333	INTRON	NA	7.43
chr3l_1277225	CG32333	FBgn0052333	INTRON	NA	7.43
chr3l_1277226	CG32333	FBgn0052333	INTRON	NA	7.43
chr3l_1277227	CG32333	FBgn0052333	INTRON	NA	7.43
chrx_21315361	CG42343	FBgn0259245	INTRON	NA	7.43
chrx_21315362	CG42343	FBgn0259245	INTRON	NA	7.43
chr2r_6570464	stan	FBgn0024836	INTRON	NA	7.42
chr2l_3711191	cutlet	FBgn0015376	NON_SYNONYMOUS_CODING	NA	7.41
chr3l_10614430	CG43127	FBgn0262592	DOWNSTREAM	3398	7.40
chr3l_10614430	CG42521	FBgn0260396	DOWNSTREAM	1736	7.40
chr3l_2574451	msn	FBgn0010909	INTRON	NA	7.40
chr3l_10616787	CG43127	FBgn0262592	DOWNSTREAM	1041	7.39
chr3l_10616787	CG42521	FBgn0260396	UPSTREAM	111	7.39
chr3l_5479057	CG4835	FBgn0035607	DOWNSTREAM	2294	7.37
chr3l_18488284	AICR2	FBgn0036789	INTRON	NA	7.36
chrx_8916007	rdgA	FBgn0261549	INTRON	NA	7.36
chrx_8916007	CR43836	FBgn0264384	INTRON	NA	7.36
chrx_13408215	Neto	FBgn0052635	INTRON	NA	7.32
chr3l_10631766	CG8009	FBgn0036090	SYNONYMOUS_CODING	NA	7.31
chr3l_2220902			SNP is more than 5000 bp away from known genes	NA	7.31
chr2l_8639111	Sema-1a	FBgn0011259	INTRON	NA	7.30
chr2r_18509918	CG33143	FBgn0053143	INTRON	NA	7.30
chr2r_6570419	stan	FBgn0024836	INTRON	NA	7.29
chr2r_19462480	apt	FBgn0015903	INTRON	NA	7.29
chr2r_9279637	CG43691	FBgn0263774	NON_SYNONYMOUS_CODING	NA	7.28
chrx_19293197	kek5	FBgn0031016	INTRON	NA	7.27
chr3l_18488329	AICR2	FBgn0036789	INTRON	NA	7.27
chr2l_2585503	Drp1	FBgn0026479	UPSTREAM	55	7.25
chr2l_2585503	CG15394	FBgn0250835	UPSTREAM	3351	7.25
chr2r_6570465	stan	FBgn0024836	INTRON	NA	7.25
chrx_12656071	CG12717	FBgn0030420	SYNONYMOUS_CODING	NA	7.25
chr2l_9135470	CG32982	FBgn0052982	INTRON	NA	7.24
chr2r_9279659	CG43691	FBgn0263774	SYNONYMOUS_CODING	NA	7.24
chr3l_10707096			SNP is more than 5000 bp away from known genes	NA	7.23
chr3l_10707097			SNP is more than 5000 bp away from known genes	NA	7.23
chr2l_1296034	robo3	FBgn0041097	DOWNSTREAM	1347	7.23
chr3l_6903932	Prat2	FBgn0041194	DOWNSTREAM	1167	7.23
chr2l_11397558			SNP is more than 5000 bp away	NA	7.22

			from known genes		
chrM_2r_1563527	scaf	FBgn0033033	DOWNSTREAM	727	7.22
chrM_2r_1563527	kune	FBgn0033032	DOWNSTREAM	3225	7.22
chrM_2r_6575211	stan	FBgn0024836	INTRON	NA	7.22
chrM_2l_6843244	sens-2	FBgn0051632	DOWNSTREAM	4239	7.21
chrM_x_12001735	CR43960	FBgn0264675	INTRON	NA	7.19
chrM_x_12001735	CG2750	FBgn0030376	INTRON	NA	7.19
chrM_2l_3711118	cutlet	FBgn0015376	SYNONYMOUS_CODING	NA	7.19
chrM_2l_3711130	cutlet	FBgn0015376	SYNONYMOUS_CODING	NA	7.19
chrM_2r_6729331	CG30015	FBgn0050015	SYNONYMOUS_CODING	NA	7.19
chrM_2l_3575889	CG43815	FBgn0264363	UPSTREAM	321	7.18
chrM_2l_3575889	CR43822	FBgn0264370	DOWNSTREAM	83	7.18
chrM_3r_20665770	CG13636	FBgn0039232	UTR_3_PRIME	NA	7.17
chrM_3l_855884	dpr20	FBgn0035170	INTRON	NA	7.16
chrM_3l_14205832	nuf	FBgn0013718	INTRON	NA	7.16
chrM_3l_14205829	nuf	FBgn0013718	INTRON	NA	7.16
chrM_3l_14205831	nuf	FBgn0013718	INTRON	NA	7.16
chrM_3r_3726146	wtrw	FBgn0260005	SYNONYMOUS_CODING	NA	7.15
chrM_2l_16054783	beat-1a	FBgn0013433	INTRON	NA	7.15
chrM_3l_10632370	CG8009	FBgn0036090	SYNONYMOUS_CODING	NA	7.14
chrM_2r_13812178	elk	FBgn0011589	SYNONYMOUS_CODING	NA	7.13
chrM_2l_15805377			SNP is more than 5000 bp away from known genes	NA	7.13
chrM_2r_9279672	CG43691	FBgn0263774	NON_SYNONYMOUS_CODING	NA	7.12
chrM_2l_10586312	Trim9	FBgn0051721	INTRON	NA	7.12
chrM_2l_11770148			SNP is more than 5000 bp away from known genes	NA	7.12
chrM_3l_13766739	bru-3	FBgn0264001	INTRON	NA	7.12
chrM_3l_10004092	dpr6	FBgn0040823	INTRON	NA	7.12
chrM_3l_10004108	dpr6	FBgn0040823	INTRON	NA	7.12
chrM_3r_1270839	CG14669	FBgn0037326	INTRON	NA	7.12
chrM_3l_18488341	AICR2	FBgn0036789	INTRON	NA	7.12
chrM_2l_13608629	kuz	FBgn0259984	INTRON	NA	7.10
chrM_3l_6903985	Prat2	FBgn0041194	DOWNSTREAM	1114	7.09
chrM_3l_4621795	CG32246	FBgn0052246	SYNONYMOUS_CODING	NA	7.09
chrM_x_9214582	c12.2	FBgn0040234	INTRON	NA	7.08
chrM_2l_9513712	CG33298	FBgn0032120	INTRON	NA	7.08
chrM_x_3538650	AlstR	FBgn0028961	INTRON	NA	7.07
chrM_3r_20559590	tok	FBgn0004885	INTRON	NA	7.06
chrM_3l_6574746	jv	FBgn0263973	INTRON	NA	7.05
chrM_3l_6574746	CG18769	FBgn0042185	INTRON	NA	7.05
chrM_3r_21380698	msi	FBgn0011666	INTRON	NA	7.03
chrM_3r_10187030	CG31320	FBgn0051320	UTR_5_PRIME	NA	7.02
chrM_2r_6729208	CG30015	FBgn0050015	SYNONYMOUS_CODING	NA	7.01
chrM_3r_23781753	CG5527	FBgn0039564	SYNONYMOUS_CODING	NA	7.01



chr <sub>m</sub> _2r_5298746	wun	FBgn0016078	INTRON	NA	7.00
chr <sub>m</sub> _3l_10631736	CG8009	FBgn0036090	SYNONYMOUS_CODING	NA	7.00
chr <sub>m</sub> _x_10073599	CG43902	FBgn0264503	INTRON	NA	7.00
chr <sub>m</sub> _3l_22916473	CG12768	FBgn0037206	UTR_3_PRIME	NA	7.00

**Supplementary Table 3c.** SNPs significantly associated with TAG.

Position (Chromosome arm, base)	Gene symbol	FlyBase ID	Site class	Bases from gene	p-value (-log10)
chrM_3r_962497	CG12589	FBgn0040684	DOWNSTREAM	282	12.07
chrM_2r_15067444	hppy	FBgn0263395	INTRON	NA	11.84
chrM_3r_1698167	CG34113	FBgn0083949	INTRON	NA	11.45
chrM_2r_14167711	sbb	FBgn0010575	INTRON	NA	10.97
chrM_3l_11006124	klu	FBgn0013469	UPSTREAM	4191	10.88
chrM_x_5022378	rg	FBgn0086911	INTRON	NA	10.73
chrM_x_5022421	rg	FBgn0086911	INTRON	NA	10.73
chrM_x_5022446	rg	FBgn0086911	INTRON	NA	10.73
chrM_3l_11006052	klu	FBgn0013469	UPSTREAM	4119	10.55
chrM_3r_25235093	Ptp99A	FBgn0004369	INTRON	NA	10.11
chrM_3r_25235095	Ptp99A	FBgn0004369	INTRON	NA	10.11
chrM_x_14919621	be	FBgn0052594	INTRON	NA	10.00
chrM_x_14919621	hiw	FBgn0030600	INTRON	NA	10.00
chrM_3r_13295471	Dscam3	FBgn0261046	DOWNSTREAM	4186	9.92
chrM_2r_7927904	RpIII128	FBgn0004463	SYNONYMOUS_CODING	NA	9.87
chrM_3r_25235132	Ptp99A	FBgn0004369	INTRON	NA	9.86
chrM_x_16536018	mth11	FBgn0030766	NON_SYNONYMOUS_CODING	NA	9.86
chrM_3l_12203385	app	FBgn0260941	UTR_3_PRIME	NA	9.85
chrM_3l_8705647			SNP is more than 5000 bp away from known genes	NA	9.83
chrM_3l_8710596	SrpRbeta	FBgn0011509	DOWNSTREAM	1453	9.83
chrM_3l_8710596	CG32022	FBgn0052022	DOWNSTREAM	2784	9.83
chrM_3l_8710971	SrpRbeta	FBgn0011509	DOWNSTREAM	1078	9.83
chrM_3l_8710971	CG32022	FBgn0052022	DOWNSTREAM	2409	9.83
chrM_2r_17584889	CG17922	FBgn0034656	UPSTREAM	232	9.70
chrM_2r_17584889	HmgZ	FBgn0010228	DOWNSTREAM	167	9.70
chrM_x_16536055	mth11	FBgn0030766	SYNONYMOUS_CODING	NA	9.65
chrM_x_3859812	CG6428	FBgn0029689	UTR_3_PRIME	NA	9.64
chrM_3l_8705580			SNP is more than 5000 bp away from known genes	NA	9.60
chrM_3r_11772759	Sap47	FBgn0013334	INTRON	NA	9.60
chrM_x_16234678	Dsp1	FBgn0011764	DOWNSTREAM	1874	9.60
chrM_2l_5694680			SNP is more than 5000 bp away from known genes	NA	9.54
chrM_3l_8710322	SrpRbeta	FBgn0011509	DOWNSTREAM	1727	9.53
chrM_3l_8710322	CG32022	FBgn0052022	DOWNSTREAM	3058	9.53
chrM_3r_11022714	CG42542	FBgn0260659	UTR_5_PRIME	NA	9.52
chrM_x_10141989	Hk	FBgn0263220	INTRON	NA	9.45
chrM_x_10142017	Hk	FBgn0263220	INTRON	NA	9.44
chrM_2l_11798122	crol	FBgn0020309	INTRON	NA	9.44

chrM_3l_20499430	CG5078	FBgn0037005	NON_SYNONYMOUS_CODING	NA	9.42
chrM_x_3862905	CG6414	FBgn0029690	DOWNSTREAM	126	9.42
chrM_x_3862905	CG6428	FBgn0029689	UPSTREAM	308	9.42
chrM_x_17777078	mnb	FBgn0259168	INTRON	NA	9.41
chrM_2r_7928303	RpIII128	FBgn0004463	SYNONYMOUS_CODING	NA	9.41
chrM_2r_7928294	RpIII128	FBgn0004463	SYNONYMOUS_CODING	NA	9.38
chrM_3l_20464996	CG32428	FBgn0052428	INTRON	NA	9.34
chrM_3l_15154324	CG7011	FBgn0036489	SYNONYMOUS_CODING	NA	9.31
chrM_3l_6151699	Cpr65Ay	FBgn0085300	UTR_5_PRIME	NA	9.30
chrM_x_8724841			SNP is more than 5000 bp away from known genes	NA	9.29
chrM_3r_23465789	CG34294	FBgn0085323	NON_SYNONYMOUS_CODING	NA	9.28
chrM_3r_6858920	CG34114	FBgn0083950	INTRON	NA	9.28
chrM_3r_6858921	CG34114	FBgn0083950	INTRON	NA	9.28
chrM_2l_15289419	CG15260	FBgn0028850	DOWNSTREAM	2007	9.27
chrM_2l_15289419	ms(2)35Ci	FBgn0011239	UPSTREAM	3049	9.27
chrM_3l_10354912	mir-276a	FBgn0262439	UPSTREAM	3416	9.26
chrM_x_3864288	CG6414	FBgn0029690	INTRON	NA	9.26
chrM_3r_16254063	mun	FBgn0262869	INTRON	NA	9.23
chrM_3l_11770117	Pi3K68D	FBgn0015278	INTRON	NA	9.19
chrM_x_2453446	Pdfr	FBgn0260753	INTRON	NA	9.16
chrM_3l_7055771	CG8641	FBgn0035733	SYNONYMOUS_CODING	NA	9.16
chrM_2r_7928168	RpIII128	FBgn0004463	SYNONYMOUS_CODING	NA	9.16
chrM_2r_11473329			SNP is more than 5000 bp away from known genes	NA	9.16
chrM_3l_9009365	Doc2	FBgn0035956	INTRON	NA	9.13
chrM_2l_10970127	CG33129	FBgn0053129	UTR_5_PRIME	NA	9.10
chrM_x_12949796	CG42629	FBgn0261388	INTRON	NA	9.03
chrM_2r_7369979	inv	FBgn0001269	INTRON	NA	9.03
chrM_2r_17237640	CG17974	FBgn0034624	INTRON	NA	9.02
chrM_2r_17237620	CG17974	FBgn0034624	INTRON	NA	9.01
chrM_x_3867078	CG6414	FBgn0029690	UPSTREAM	1697	9.00
chrM_x_3867078	CG6428	FBgn0029689	UPSTREAM	4481	9.00
chrM_2r_7886493	otk	FBgn0004839	DOWNSTREAM	2490	9.00
chrM_3r_20449285	mld	FBgn0263490	UTR_3_PRIME	NA	8.99
chrM_2l_7927199	Snoo	FBgn0085450	INTRON	NA	8.95
chrM_2r_14596876	CG15087	FBgn0034380	SYNONYMOUS_CODING	NA	8.94
chrM_3l_2627277	Pxn	FBgn0011828	INTRON	NA	8.91
chrM_2l_1215397	CR43263	FBgn0262944	DOWNSTREAM	1457	8.91
chrM_2l_1215397	CG42329	FBgn0259229	UPSTREAM	3921	8.91
chrM_3r_23465893	CG34294	FBgn0085323	NON_SYNONYMOUS_CODING	NA	8.91
chrM_x_3862818	CG6414	FBgn0029690	DOWNSTREAM	213	8.91
chrM_x_3862818	CG6428	FBgn0029689	UPSTREAM	221	8.91
chrM_x_3866252	CG6428	FBgn0029689	UPSTREAM	3655	8.90
chrM_x_3866252	CG6414	FBgn0029690	UPSTREAM	871	8.90

chrM_2r_10804295	CG10249	FBgn0027596	INTRON	NA	8.90
chrM_2l_15300432	CG15262	FBgn0028852	DOWNSTREAM	1065	8.87
chrM_2l_15300432	CR43764	FBgn0264264	UPSTREAM	3019	8.87
chrM_2r_20255301	bs	FBgn0004101	INTRON	NA	8.86
chrM_3r_16254471	mun	FBgn0262869	INTRON	NA	8.85
chrM_x_3866193	CG6414	FBgn0029690	UPSTREAM	812	8.83
chrM_x_3866193	CG6428	FBgn0029689	UPSTREAM	3596	8.83
chrM_3l_11615069	CG6084	FBgn0086254	INTRON	NA	8.83
chrM_3r_25894792	sima	FBgn0015542	INTRON	NA	8.79
chrM_2r_7928169	RpIII128	FBgn0004463	SYNONYMOUS_CODING	NA	8.77
chrM_3r_18805524	Nha2	FBgn0263390	DOWNSTREAM	1657	8.76
chrM_3r_18805524	CG17244	FBgn0039031	DOWNSTREAM	1900	8.76
chrM_3r_16638943	Fancd2	FBgn0038827	INTRON	NA	8.75
chrM_3l_6151698	Cpr65Ay	FBgn0085300	UTR_5_PRIME	NA	8.75
chrM_3l_6151719	Cpr65Ay	FBgn0085300	UTR_5_PRIME	NA	8.75
chrM_3l_2202045	CG8960	FBgn0035315	DOWNSTREAM	3956	8.75
chrM_2r_15827307	CR43421	FBgn0263340	DOWNSTREAM	80	8.75
chrM_2r_15827307	CG43195	FBgn0262824	DOWNSTREAM	4461	8.75
chrM_2r_15827415	CG43195	FBgn0262824	DOWNSTREAM	4353	8.75
chrM_2r_7928187	RpIII128	FBgn0004463	SYNONYMOUS_CODING	NA	8.73
chrM_2r_7928060	RpIII128	FBgn0004463	SYNONYMOUS_CODING	NA	8.72
chrM_x_19403091	Pfrx	FBgn0027621	UTR_3_PRIME	NA	8.71
chrM_2r_19147755	CG12782	FBgn0034838	DOWNSTREAM	4037	8.71
chrM_3r_22011079	CG31086	FBgn0051086	UTR_5_PRIME	NA	8.68
chrM_3r_22011079	CG31323	FBgn0051323	UTR_3_PRIME	NA	8.68
chrM_2r_7928600	RpIII128	FBgn0004463	SYNONYMOUS_CODING	NA	8.67
chrM_2r_15827697	CG42690	FBgn0261580	DOWNSTREAM	4766	8.67
chrM_2r_15827697	CG43195	FBgn0262824	DOWNSTREAM	4071	8.67
chrM_3l_11615116	CG6084	FBgn0086254	INTRON	NA	8.66
chrM_3l_2202046	CG8960	FBgn0035315	DOWNSTREAM	3957	8.65
chrM_x_3865307	CG6414	FBgn0029690	SYNONYMOUS_CODING	NA	8.64
chrM_2r_7927883	RpIII128	FBgn0004463	SYNONYMOUS_CODING	NA	8.62
chrM_2r_8612003	CG42663	FBgn0261545	INTRON	NA	8.61
chrM_3l_5802659	S6k	FBgn0015806	INTRON	NA	8.61
chrM_2r_9611603	fas	FBgn0000633	INTRON	NA	8.60
chrM_3r_6863343	CG34114	FBgn0083950	INTRON	NA	8.59
chrM_3l_20467038	CG5969	FBgn0036998	NON_SYNONYMOUS_CODING	NA	8.59
chrM_3l_4695034	Gef64C	FBgn0035574	INTRON	NA	8.59
chrM_2r_10879323	CG7639	FBgn0033989	INTRON	NA	8.58
chrM_2r_6222131	CG42732	FBgn0261698	INTRON	NA	8.56
chrM_3l_21597451	AP-2	FBgn0261953	INTRON	NA	8.54
chrM_2r_15067594	hppy	FBgn0263395	INTRON	NA	8.54
chrM_x_16536043	mthl1	FBgn0030766	SYNONYMOUS_CODING	NA	8.54
chrM_2r_15067615	hppy	FBgn0263395	INTRON	NA	8.52

chrM_3l_2318672	DmsR-1	FBgn0035331	UPSTREAM	4021	8.52
chrM_2r_17237960	CG17974	FBgn0034624	START_GAINED	NA	8.52
chrM_2r_17237736	CG17974	FBgn0034624	INTRON	NA	8.50
chrM_3l_12253473	app	FBgn0260941	INTRON	NA	8.48
chrM_2r_14596572	CG15087	FBgn0034380	SYNONYMOUS_CODING	NA	8.47
chrM_2r_14596681	CG15087	FBgn0034380	SYNONYMOUS_CODING	NA	8.47
chrM_3r_6862427	CG34114	FBgn0083950	INTRON	NA	8.47
chrM_3r_25316854	CG2006	FBgn0039664	UTR_3_PRIME	NA	8.46
chrM_3r_25316854	Spase12	FBgn0040623	UTR_5_PRIME	NA	8.46
chrM_2r_17237408	CG17974	FBgn0034624	INTRON	NA	8.46
chrM_2r_7928213	RpIII128	FBgn0004463	SYNONYMOUS_CODING	NA	8.46
chrM_3l_7133507	melt	FBgn0023001	SYNONYMOUS_CODING	NA	8.44
chrM_2l_6373214	slam	FBgn0043854	INTRON	NA	8.43
chrM_2r_15067697	hppy	FBgn0263395	INTRON	NA	8.43
chrM_2r_15067698	hppy	FBgn0263395	INTRON	NA	8.43
chrM_2r_10323860	Oaz	FBgn0261613	INTRON	NA	8.42
chrM_2r_7928552	RpIII128	FBgn0004463	SYNONYMOUS_CODING	NA	8.42
chrM_2r_7928570	RpIII128	FBgn0004463	SYNONYMOUS_CODING	NA	8.42
chrM_3l_18355537			SNP is more than 5000 bp away from known genes	NA	8.40
chrM_2r_14170724	sbb	FBgn0010575	SYNONYMOUS_CODING	NA	8.38
chrM_x_14919650	be	FBgn0052594	INTRON	NA	8.38
chrM_x_14919650	hiw	FBgn0030600	INTRON	NA	8.38
chrM_3r_6296733			SNP is more than 5000 bp away from known genes	NA	8.36
chrM_3r_6296833			SNP is more than 5000 bp away from known genes	NA	8.36
chrM_3l_21397522	rgn	FBgn0261258	INTRON	NA	8.36
chrM_2r_7928696	RpIII128	FBgn0004463	SYNONYMOUS_CODING	NA	8.34
chrM_3r_14198670	gl	FBgn0004618	UPSTREAM	210	8.34
chrM_3r_21068896	bam	FBgn0000158	UTR_3_PRIME	NA	8.33
chrM_3r_21068896	CG11854	FBgn0039299	UTR_3_PRIME	NA	8.33
chrM_3r_6871565	CG34114	FBgn0083950	UTR_3_PRIME	NA	8.33
chrM_3r_6862432	CG34114	FBgn0083950	INTRON	NA	8.32
chrM_x_19403120	Pfx	FBgn0027621	UTR_3_PRIME	NA	8.30
chrM_3l_5702364	sif	FBgn0085447	INTRON	NA	8.30
chrM_3r_6863302	CG34114	FBgn0083950	INTRON	NA	8.29
chrM_3r_6871879	CG34114	FBgn0083950	INTRON	NA	8.29
chrM_2r_13721761	grh	FBgn0259211	INTRON	NA	8.27
chrM_3r_11996375	CG6006	FBgn0063649	UTR_3_PRIME	NA	8.27
chrM_2r_14172109	sbb	FBgn0010575	SYNONYMOUS_CODING	NA	8.26
chrM_3r_23931978			SNP is more than 5000 bp away from known genes	NA	8.25
chrM_2r_14176935	sbb	FBgn0010575	INTRON	NA	8.24
chrM_2r_14176935	CG14502	FBgn0034321	INTRON	NA	8.24

chrM_2l_5521711	CG7371	FBgn0031710	SYNONYMOUS_CODING	NA	8.24
chrM_3l_8323849	CG34461	FBgn0250833	DOWNSTREAM	2442	8.23
chrM_3l_8323849	CG34462	FBgn0085491	DOWNSTREAM	512	8.23
chrM_3l_22948868	Mes2	FBgn0037207	DOWNSTREAM	3520	8.22
chrM_3l_22948868	CG32461	FBgn0052461	UPSTREAM	466	8.22
chrM_3r_6871372	CG34114	FBgn0083950	UTR_3_PRIME	NA	8.21
chrM_3r_23933504			SNP is more than 5000 bp away from known genes	NA	8.21
chrM_2r_7928567	RpIII128	FBgn0004463	SYNONYMOUS_CODING	NA	8.19
chrM_x_6769215	CG43736	FBgn0263993	INTRON	NA	8.19
chrM_3r_5738033	Glut4EF	FBgn0263097	INTRON	NA	8.18
chrM_3l_6151439	Cpr65Ay	FBgn0085300	SYNONYMOUS_CODING	NA	8.18
chrM_3l_6151458	Cpr65Ay	FBgn0085300	INTRON	NA	8.18
chrM_3l_6151478	Cpr65Ay	FBgn0085300	INTRON	NA	8.18
chrM_3l_6151484	Cpr65Ay	FBgn0085300	INTRON	NA	8.18
chrM_3l_6151486	Cpr65Ay	FBgn0085300	INTRON	NA	8.18
chrM_3r_17035860	SNF4Agamma	FBgn0264357	INTRON	NA	8.18
chrM_3r_6858781	CG34114	FBgn0083950	INTRON	NA	8.17
chrM_3r_6691275	SelR	FBgn0037847	SYNONYMOUS_CODING	NA	8.17
chrM_2r_7929617	RpIII128	FBgn0004463	SYNONYMOUS_CODING	NA	8.17
chrM_2r_7886473	otk	FBgn0004839	DOWNSTREAM	2510	8.17
chrM_3l_8941723	CG43783	FBgn0264305	INTRON	NA	8.16
chrM_3l_8941723	orb2	FBgn0264307	INTRON	NA	8.16
chrM_3r_4899495	pum	FBgn0003165	INTRON	NA	8.16
chrM_3l_6151331	Cpr65Ay	FBgn0085300	SYNONYMOUS_CODING	NA	8.16
chrM_2r_7935363	Drep-1	FBgn0024732	INTRON	NA	8.16
chrM_x_3866186	CG6414	FBgn0029690	UPSTREAM	805	8.14
chrM_x_3866186	CG6428	FBgn0029689	UPSTREAM	3589	8.14
chrM_2r_14172542	sbb	FBgn0010575	INTRON	NA	8.13
chrM_2r_7886483	otk	FBgn0004839	DOWNSTREAM	2500	8.12
chrM_x_4460281	CG32773	FBgn0052773	INTRON	NA	8.10

**Supplementary Table 3d.** SNPs significantly associated with glucose.

Position (Chromosome arm, base)	Gene symbol	FlyBase ID	Site class	Bases from gene	p-value (-log10)
chrM_2r_11773694	sli	FBgn0264089	INTRON	NA	14.81
chrM_3r_14502104	CG7705	FBgn0038639	UTR_3_PRIME	NA	14.65
chrM_3r_14502107	CG7705	FBgn0038639	UTR_3_PRIME	NA	14.55
chrM_3l_3798748	CG32264	FBgn0052264	INTRON	NA	14.24
chrM_3r_16174735			SNP is more than 5000 bp away from known genes	NA	14.11
chrM_3r_14502098	CG7705	FBgn0038639	UTR_3_PRIME	NA	14.09
chrM_3r_20323256			SNP is more than 5000 bp away from known genes	NA	14.08
chrM_2l_7586215	Spn7	FBgn0083141	DOWNSTREAM	2881	13.93
chrM_x_3202872	dnc	FBgn0000479	INTRON	NA	13.74
chrM_3r_20320641	nAcRalpha- 96Ab	FBgn0000039	DOWNSTREAM	3766	13.51
chrM_3r_20323341			SNP is more than 5000 bp away from known genes	NA	13.51
chrM_3r_20320706	nAcRalpha- 96Ab	FBgn0000039	DOWNSTREAM	3831	13.43
chrM_2r_8327699	Dyb	FBgn0033739	INTRON	NA	13.29
chrM_3r_20325371			SNP is more than 5000 bp away from known genes	NA	13.21
chrM_3r_20325398			SNP is more than 5000 bp away from known genes	NA	13.21
chrM_x_3202877	dnc	FBgn0000479	INTRON	NA	13.19
chrM_3r_20324464			SNP is more than 5000 bp away from known genes	NA	13.16
chrM_2r_7755308	CG13185	FBgn0033661	DOWNSTREAM	73	13.15
chrM_2r_7755308	CR42532	FBgn0260437	DOWNSTREAM	224	13.15
chrM_2r_7755310	CR42532	FBgn0260437	DOWNSTREAM	226	13.15
chrM_2r_7755310	CG13185	FBgn0033661	DOWNSTREAM	71	13.15
chrM_3r_16174621	Ir92a	FBgn0038789	DOWNSTREAM	4997	13.12
chrM_3r_20325232			SNP is more than 5000 bp away from known genes	NA	13.01
chrM_3r_20325263			SNP is more than 5000 bp away from known genes	NA	13.01
chrM_3r_20325166			SNP is more than 5000 bp away from known genes	NA	13.01
chrM_3r_22761063	sda	FBgn0015541	INTRON	NA	12.78
chrM_3r_20325098			SNP is more than 5000 bp away from known genes	NA	12.66
chrM_3r_17216223	bap	FBgn0004862	DOWNSTREAM	52	12.65
chrM_2r_12389852	Sema-2a	FBgn0011260	INTRON	NA	12.46
chrM_2l_14013222	bgm	FBgn0027348	UPSTREAM	525	12.46

chrM_2l_14013222	tRNA:Q:34E	FBgn0028909	UPSTREAM	2586	12.46
chrM_3r_9889686	foxo	FBgn0038197	INTRON	NA	12.45
chrM_2r_4270801	pdm3	FBgn0261588	INTRON	NA	12.42
chrM_3r_22150509	CG42765	FBgn0261833	UTR_3_PRIME	NA	12.40
chrM_3r_22762222	sda	FBgn0015541	INTRON	NA	12.32
chrM_3r_22762340	sda	FBgn0015541	INTRON	NA	12.32
chrM_2r_16437575	Obp57d	FBgn0043536	UTR_5_PRIME	NA	12.30
chrM_2r_16437575	Cpr57A	FBgn0034517	UTR_5_PRIME	NA	12.30
chrM_3r_22759527	sda	FBgn0015541	INTRON	NA	12.28
chrM_3r_11453171			SNP is more than 5000 bp away from known genes	NA	12.25
chrM_3l_22652231			SNP is more than 5000 bp away from known genes	NA	12.24
chrM_3r_8126925	svp	FBgn0003651	UTR_3_PRIME	NA	12.16
chrM_2r_9977354	Prosap	FBgn0040752	INTRON	NA	12.06
chrM_2r_9977388	Prosap	FBgn0040752	INTRON	NA	12.06
chrM_2r_17406887	CG30288	FBgn0050288	NON_SYNONYMOUS_CODING	NA	11.98
chrM_2r_17406888	CG30288	FBgn0050288	SYNONYMOUS_CODING	NA	11.98
chrM_2r_17406891	CG30288	FBgn0050288	SYNONYMOUS_CODING	NA	11.98
chrM_3r_20321322	nAcRalpha-96Ab	FBgn0000039	DOWNSTREAM	4447	11.92
chrM_2l_15607491	CG15255	FBgn0028950	DOWNSTREAM	3241	11.92
chrM_3r_24833255	CG1443	FBgn0039620	INTRON	NA	11.90
chrM_3r_1655708	CG1208	FBgn0037386	UPSTREAM	1642	11.89
chrM_3r_1655708	glob3	FBgn0037385	UPSTREAM	76	11.89
chrM_3r_23456407	Tusp	FBgn0039530	SYNONYMOUS_CODING	NA	11.89
chrM_2r_9580801	CG6220	FBgn0033865	UTR_5_PRIME	NA	11.89
chrM_2r_9580819	CG6220	FBgn0033865	UTR_5_PRIME	NA	11.89
chrM_3l_10798454	CG12523	FBgn0036102	DOWNSTREAM	3253	11.88
chrM_3r_14037670	CG7985	FBgn0028499	INTRON	NA	11.88
chrM_2r_16437573	Obp57d	FBgn0043536	UTR_5_PRIME	NA	11.79
chrM_2r_16437573	Cpr57A	FBgn0034517	UTR_5_PRIME	NA	11.79
chrM_3l_16206905	CG13073	FBgn0036577	NON_SYNONYMOUS_CODING	NA	11.75
chrM_3r_14037792	CG7985	FBgn0028499	INTRON	NA	11.72
chrM_3r_20317482	nAcRalpha-96Ab	FBgn0000039	DOWNSTREAM	607	11.69
chrM_3r_20317482	mir-1017	FBgn0262389	DOWNSTREAM	3078	11.69
chrM_3r_12896218	ns1	FBgn0038473	DOWNSTREAM	27	11.66
chrM_3r_12896218	Dad	FBgn0020493	DOWNSTREAM	83	11.66
chrM_2r_12282635	Sema-2b	FBgn0264273	INTRON	NA	11.61
chrM_3r_1784495	CG11373	FBgn0040679	DOWNSTREAM	1550	11.59
chrM_2l_7586341	Spn7	FBgn0083141	DOWNSTREAM	3007	11.54
chrM_3l_16207076	CG13073	FBgn0036577	NON_SYNONYMOUS_CODING	NA	11.53
chrM_3r_24974551	Pglym78	FBgn0014869	INTRON	NA	11.53
chrM_2r_9580820	CG6220	FBgn0033865	UTR_5_PRIME	NA	11.53



chr3_17234756	CheB93b	FBgn0051438	UPSTREAM	89	11.48
chr3_17234756	CheB93a	FBgn0038888	DOWNSTREAM	710	11.48
chr2_7755346	CR42532	FBgn0260437	DOWNSTREAM	262	11.45
chr2_7755346	CG13185	FBgn0033661	DOWNSTREAM	35	11.45
chr3_10604164	jvl	FBgn0263929	INTRON	NA	11.44
chr3_17199596	mod(mdg4)	FBgn0002781	INTRON	NA	11.43
chr3_15422589	Dys	FBgn0260003	DOWNSTREAM	1472	11.42
chr3_15422589	Cpr92A	FBgn0038714	DOWNSTREAM	635	11.42
chr3_20331821	nAcRbeta-96A	FBgn0004118	DOWNSTREAM	3116	11.41
chr3_17634951	CASK	FBgn0013759	INTRON	NA	11.41
chr3_10045451	dpr6	FBgn0040823	INTRON	NA	11.41
chr3_1780242			SNP is more than 5000 bp away from known genes	NA	11.40
chr3_26182438	hdc	FBgn0010113	INTRON	NA	11.39
chrX_5028928	rg	FBgn0086911	INTRON	NA	11.38
chr3_9880633	foxo	FBgn0038197	UPSTREAM	2057	11.32
chr3_25526492	dmrt99B	FBgn0039683	DOWNSTREAM	3745	11.32
chr2_19526189	retn	FBgn0004795	INTRON	NA	11.31
chr2_16802118	Rx	FBgn0020617	UPSTREAM	2167	11.30
chr2_16802118	CG9235	FBgn0034560	DOWNSTREAM	3805	11.30
chr3_16174327	Ir92a	FBgn0038789	DOWNSTREAM	4703	11.28
chr3_3112460	CG14955	FBgn0035399	SYNONYMOUS_CODING	NA	11.28
chr2_18533385	CG2852	FBgn0034753	INTRON	NA	11.26
chr3_22678550	CG14455	FBgn0037175	DOWNSTREAM	910	11.23
chr3_22678550	CG14456	FBgn0037176	UPSTREAM	1772	11.23
chr2_7754966	CG13185	FBgn0033661	DOWNSTREAM	415	11.23
chr2_7754966	Damm	FBgn0033659	UPSTREAM	1078	11.23
chr2_17406122	CG10494	FBgn0034634	SYNONYMOUS_CODING	NA	11.22
chr3_11875006	ird5	FBgn0024222	NON_SYNONYMOUS_CODING	NA	11.22
chr2_8926920			SNP is more than 5000 bp away from known genes	NA	11.20
chr2_6905242	neuroligin	FBgn0031866	SYNONYMOUS_CODING	NA	11.20
chr3_20317663	mir-1017	FBgn0262389	DOWNSTREAM	3259	11.20
chr3_20317663	nAcRalpha-96Ab	FBgn0000039	DOWNSTREAM	788	11.20
chr2_7755499	CG13185	FBgn0033661	UTR_3_PRIME	NA	11.19
chr2_5309507	CG13955	FBgn0033412	UTR_3_PRIME	NA	11.18
chr2_5309558	CG13955	FBgn0033412	UTR_3_PRIME	NA	11.12
chr2_5309516	CG13955	FBgn0033412	UTR_3_PRIME	NA	11.11
chr3_21760732	CcapR	FBgn0039396	INTRON	NA	11.09
chr3_20317278	mir-1017	FBgn0262389	DOWNSTREAM	2874	11.08
chr3_20317278	nAcRalpha-96Ab	FBgn0000039	DOWNSTREAM	403	11.08
chr2_16357459	CG4455	FBgn0028506	DOWNSTREAM	2920	11.08
chr2_16357459	CaBP1	FBgn0025678	UPSTREAM	725	11.08

chr3l_9582091	CG42673	FBgn0261555	INTRON	NA	11.07
chrX_14434961			SNP is more than 5000 bp away from known genes	NA	11.04
chr3l_3480897	CG42324	FBgn0259224	INTRON	NA	11.03
chr3r_22170308	CG6036	FBgn0039421	SYNONYMOUS_CODING	NA	10.99
chr3r_25469389	DopR2	FBgn0015129	INTRON	NA	10.97
chr3r_24838734	CG1443	FBgn0039620	INTRON	NA	10.97
chr2r_16715238	sktl	FBgn0016984	UTR_3_PRIME	NA	10.96
chr3l_6526677	sfl	FBgn0020251	INTRON	NA	10.96
chr3r_22150945	CG42765	FBgn0261833	SYNONYMOUS_CODING	NA	10.96
chr3r_22153499	CG33970	FBgn0053970	INTRON	NA	10.96
chr3r_21773493	CcapR	FBgn0039396	INTRON	NA	10.95
chr3l_3480886	CG42324	FBgn0259224	INTRON	NA	10.93
chr2l_12157628	CG31760	FBgn0051760	INTRON	NA	10.92
chr3r_9876056			SNP is more than 5000 bp away from known genes	NA	10.92
chr3r_9877316			SNP is more than 5000 bp away from known genes	NA	10.92
chr3l_3798591	CG32264	FBgn0052264	INTRON	NA	10.91
chr2r_7659759	ths	FBgn0033652	INTRON	NA	10.90
chr2r_14511905	Atg7	FBgn0034366	SYNONYMOUS_CODING	NA	10.87
chr3l_3480920	CG42324	FBgn0259224	INTRON	NA	10.86
chr3r_11913419	Mhcl	FBgn0026059	INTRON	NA	10.86
chr2l_17541695	CR43304	FBgn0262996	UPSTREAM	159	10.85
chr3l_22669025	CG14457	FBgn0037174	INTRON	NA	10.84
chr3l_12975803			SNP is more than 5000 bp away from known genes	NA	10.82
chr3r_9875626			SNP is more than 5000 bp away from known genes	NA	10.82
chr2l_7581515	CG13791	FBgn0031923	UPSTREAM	2956	10.81
chr2l_7581515	Spn7	FBgn0083141	UPSTREAM	417	10.81
chr3r_20324384			SNP is more than 5000 bp away from known genes	NA	10.77
chr3r_8152616	CG10013	FBgn0038012	DOWNSTREAM	1680	10.75
chr3r_8152616	tRNA:CR31432	FBgn0051432	UPSTREAM	4260	10.75
chr3r_9875628			SNP is more than 5000 bp away from known genes	NA	10.75
chr3r_9880410	foxo	FBgn0038197	UPSTREAM	2280	10.74
chr3l_22668986	CG14457	FBgn0037174	INTRON	NA	10.72
chr3l_16755248	Nrt	FBgn0004108	UPSTREAM	3924	10.72
chr3r_17161541	CG16791	FBgn0038881	INTRON	NA	10.72
chr3l_10318672			SNP is more than 5000 bp away from known genes	NA	10.71
chr2r_7754671	Damm	FBgn0033659	UPSTREAM	783	10.69
chr2r_7754671	CG13185	FBgn0033661	DOWNSTREAM	710	10.69
chr3r_22106946	CG6154	FBgn0039420	DOWNSTREAM	4174	10.66

chrn_3r_14692992			SNP is more than 5000 bp away from known genes	NA	10.64
chrn_3r_22106984	CG6154	FBgn0039420	DOWNSTREAM	4212	10.63
chrn_3l_10045569	dpr6	FBgn0040823	INTRON	NA	10.61
chrn_3r_22105138	CG6154	FBgn0039420	DOWNSTREAM	2366	10.60
chrn_2r_5255470	CG13954	FBgn0033405	INTRON	NA	10.59
chrn_2r_17405453	CG10494	FBgn0034634	INTRON	NA	10.58
chrn_2r_10210389	Shroom	FBgn0085408	INTRON	NA	10.54
chrn_3r_26182554	hdc	FBgn0010113	INTRON	NA	10.54
chrn_3r_14502909	CG7705	FBgn0038639	NON_SYNONYMOUS_CODING	NA	10.54
chrn_3r_9879517	foxo	FBgn0038197	UPSTREAM	3173	10.54
chrn_2r_16800044	CG9235	FBgn0034560	DOWNSTREAM	1731	10.52
chrn_2r_16800044	Rx	FBgn0020617	UPSTREAM	4241	10.52
chrn_3l_16241481	CG13055	FBgn0036583	DOWNSTREAM	1179	10.51
chrn_3r_20320839	nAcRalpha-96Ab	FBgn0000039	DOWNSTREAM	3964	10.49
chrn_3r_20320840	nAcRalpha-96Ab	FBgn0000039	DOWNSTREAM	3965	10.49
chrn_3r_20321267	nAcRalpha-96Ab	FBgn0000039	DOWNSTREAM	4392	10.49
chrn_2r_7649214			SNP is more than 5000 bp away from known genes	NA	10.48
chrn_3l_2793655	CG43444	FBgn0263392	SYNONYMOUS_CODING	NA	10.48
chrn_3r_20321790	nAcRalpha-96Ab	FBgn0000039	DOWNSTREAM	4915	10.48
chrn_3l_22885113	slif	FBgn0037203	NON_SYNONYMOUS_CODING	NA	10.47
chrn_3r_22761381	sda	FBgn0015541	INTRON	NA	10.44
chrn_3l_3122546	prominin-like	FBgn0026189	INTRON	NA	10.44
chrn_2l_17538006	CR43304	FBgn0262996	UPSTREAM	3848	10.44
chrn_3r_6122742	CG43143	FBgn0262617	SYNONYMOUS_CODING	NA	10.43
chrn_3r_22078656	CR42745	FBgn0261708	DOWNSTREAM	2	10.43
chrn_3r_22078656	Ald	FBgn0000064	DOWNSTREAM	3032	10.43
chrn_3r_22078711	Ald	FBgn0000064	DOWNSTREAM	2977	10.43
chrn_3r_22078711	CR42745	FBgn0261708	DOWNSTREAM	57	10.43
chrn_3l_14838510	CG17839	FBgn0036454	INTRON	NA	10.43
chrn_x_17577242	chas	FBgn0263258	INTRON	NA	10.42
chrn_3r_1783135	CG11373	FBgn0040679	DOWNSTREAM	2910	10.41
chrn_2l_4796273	CG3294	FBgn0031628	INTRON	NA	10.40
chrn_3l_16243525	CG13055	FBgn0036583	UTR_3_PRIME	NA	10.39
chrn_2l_16371899	jhamt	FBgn0028841	DOWNSTREAM	4922	10.38
chrn_3l_18067744	CG34252	FBgn0085281	UPSTREAM	1262	10.37
chrn_3r_10489240	CG3509	FBgn0038252	UTR_5_PRIME	NA	10.34
chrn_3r_1785474	CG11373	FBgn0040679	DOWNSTREAM	571	10.32
chrn_2r_7602270	pyr	FBgn0033649	INTRON	NA	10.32
chrn_2l_4025345			SNP is more than 5000 bp away from known genes	NA	10.32

chr3_7928167	dpr17	FBgn0051361	INTRON	NA	10.31
chr3_12075501	Sema-5c	FBgn0250876	UPSTREAM	616	10.31
chr3_20324392			SNP is more than 5000 bp away from known genes	NA	10.29
chr3_10603802	jvl	FBgn0263929	INTRON	NA	10.28
chr3_17217074	bap	FBgn0004862	DOWNSTREAM	903	10.28
chr2_7755638	CG13185	FBgn0033661	UTR_3_PRIME	NA	10.27
chr3_20323340			SNP is more than 5000 bp away from known genes	NA	10.26
chr2_1630242			SNP is more than 5000 bp away from known genes	NA	10.24
chr2_10726987	NaPi-T	FBgn0016684	DOWNSTREAM	3238	10.22
chr3_21779541			SNP is more than 5000 bp away from known genes	NA	10.21
chr2_8199330	CG30203	FBgn0050203	SYNONYMOUS_CODING	NA	10.21
chr3_16242135	CG13055	FBgn0036583	DOWNSTREAM	525	10.20
chr2_7755021	Damm	FBgn0033659	UPSTREAM	1133	10.20
chr2_7755021	CG13185	FBgn0033661	DOWNSTREAM	360	10.20
chr3_16069675	CG34139	FBgn0083975	INTRON	NA	10.18
chr2_14514424	sec6	FBgn0034367	SYNONYMOUS_CODING	NA	10.18
chr2_16799987	Rx	FBgn0020617	UPSTREAM	4298	10.17
chr2_16799987	CG9235	FBgn0034560	DOWNSTREAM	1674	10.17
chr3_19697594	Pli	FBgn0025574	INTRON	NA	10.16
chr3_7159978	pros	FBgn0004595	INTRON	NA	10.13
chr3_1776208			SNP is more than 5000 bp away from known genes	NA	10.13
chr2_16955049	Cht4	FBgn0022700	UPSTREAM	457	10.12
chr2_16955049	Cht9	FBgn0034582	DOWNSTREAM	438	10.12
chr3_3511193	ImpE2	FBgn0001254	UTR_3_PRIME	NA	10.11
chr3_22088943	CG6154	FBgn0039420	INTRON	NA	10.10
chr3_24783133	CG14521	FBgn0039617	INTRON	NA	10.09
chr2_3800321	CG30377	FBgn0050377	INTRON	NA	10.09
chr3_6790523			SNP is more than 5000 bp away from known genes	NA	10.08
chr3_20317618	mir-1017	FBgn0262389	DOWNSTREAM	3214	10.08
chr3_20317618	nAcRalpha-96Ab	FBgn0000039	DOWNSTREAM	743	10.08
chr3_14038537	CG7985	FBgn0028499	NON_SYNONYMOUS_CODING	NA	10.08
chr3_20857037	CG32432	FBgn0052432	INTRON	NA	10.07
chr3_11912960	CG5718	FBgn0036222	DOWNSTREAM	2750	10.07
chr3_11912960	CG11588	FBgn0036221	UPSTREAM	1653	10.07
chr3_17213559	bap	FBgn0004862	UPSTREAM	1046	10.06
chr2_18435053	px	FBgn0003175	INTRON	NA	10.06
chr2_10019514	Prosap	FBgn0040752	INTRON	NA	10.06

**Supplementary Table 3e.** SNPs significantly associated with glycogen.

Position (Chromosome arm, base)	Gene symbol	FlyBase ID	Site class	Bases from gene	p-value (-log10)
chrM_2r_5672241	CG1688	FBgn0027589	INTRON	NA	15.35
chrM_2r_13547241			SNP is more than 5000 bp away from known genes	NA	14.22
chrM_2l_16179928	Ca-alpha1D	FBgn0001991	INTRON	NA	14.01
chrM_2l_1701762	chinmo	FBgn0086758	UTR_3_PRIME	NA	13.99
chrM_3r_8090545	svp	FBgn0003651	INTRON	NA	13.92
chrM_3r_8090560	svp	FBgn0003651	INTRON	NA	13.92
chrM_3r_8090059	svp	FBgn0003651	INTRON	NA	13.80
chrM_3l_14221223	nuf	FBgn0013718	INTRON	NA	13.77
chrM_3r_8090911	svp	FBgn0003651	INTRON	NA	13.77
chrM_2l_9442461	numb	FBgn0002973	INTRON	NA	13.64
chrM_2r_19178990			SNP is more than 5000 bp away from known genes	NA	13.57
chrM_3r_8087463	svp	FBgn0003651	INTRON	NA	13.55
chrM_2l_1357798	CG5565	FBgn0031335	UPSTREAM	1874	13.45
chrM_2l_1357798	CG31659	FBgn0051659	DOWNSTREAM	1127	13.45
chrM_2r_17814413	Fili	FBgn0085397	INTRON	NA	13.41
chrM_2l_13238126	CG5867	FBgn0027586	INTRON	NA	13.40
chrM_2r_14131654			SNP is more than 5000 bp away from known genes	NA	13.28
chrM_2l_4285760	tutl	FBgn0010473	INTRON	NA	13.23
chrM_2l_3750295	CG10019	FBgn0031568	SYNONYMOUS_CODING	NA	13.18
chrM_3r_9599438	CG9297	FBgn0038181	INTRON	NA	13.15
chrM_2l_13331663	CG43778	FBgn0264308	UPSTREAM	542	13.15
chrM_2l_13151655			SNP is more than 5000 bp away from known genes	NA	13.11
chrM_3r_7834991	KLHL18	FBgn0037978	INTRON	NA	13.09
chrM_3l_5902828	CG13288	FBgn0035648	SYNONYMOUS_CODING	NA	13.07
chrM_3l_6255918	ImpL3	FBgn0001258	UPSTREAM	124	13.03
chrM_3l_6255918	CG10163	FBgn0035697	UPSTREAM	4058	13.03
chrM_2l_11353767			SNP is more than 5000 bp away from known genes	NA	13.02
chrM_2l_9719890	Nckx30C	FBgn0028704	INTRON	NA	13.01
chrM_2l_7341907	Wnt6	FBgn0031902	INTRON	NA	12.93
chrM_3l_1672550	CG13930	FBgn0035256	SYNONYMOUS_CODING	NA	12.92
chrM_3r_25141415	Cnx99A	FBgn0015622	UTR_3_PRIME	NA	12.88
chrM_2r_9657492	Cpr50Ca	FBgn0033867	INTRON	NA	12.82
chrM_2r_13571056	CG42561	FBgn0260763	UPSTREAM	1167	12.82
chrM_2r_13571056	Sema-1b	FBgn0016059	UPSTREAM	56	12.82
chrM_2l_13237948	CG5867	FBgn0027586	INTRON	NA	12.81
chrM_3r_24344229	Dhc98D	FBgn0013813	INTRON	NA	12.75

chrM_2l_16179908	Ca-alpha1D	FBgn0001991	INTRON	NA	12.75
chrM_3r_12687758	msa	FBgn0263610	INTRON	NA	12.71
chrM_2r_16870322	CG9313	FBgn0034566	SYNONYMOUS_CODING	NA	12.71
chrM_3r_13125078			SNP is more than 5000 bp away from known genes	NA	12.70
chrM_3r_8084450	CG3942	FBgn0038008	DOWNSTREAM	2136	12.69
chrM_3r_8084450	svp	FBgn0003651	UPSTREAM	21	12.69
chrM_x_11068133	CG1394	FBgn0030277	UPSTREAM	3452	12.64
chrM_2r_18458956	dnr1	FBgn0260866	INTRON	NA	12.64
chrM_3l_15770767			SNP is more than 5000 bp away from known genes	NA	12.62
chrM_2l_9671873			SNP is more than 5000 bp away from known genes	NA	12.62
chrM_2l_9671890			SNP is more than 5000 bp away from known genes	NA	12.62
chrM_2l_13228678	CG16815	FBgn0032491	UPSTREAM	964	12.59
chrM_2l_13228678	CG16813	FBgn0032490	DOWNSTREAM	271	12.59
chrM_2l_17814576	CadN2	FBgn0262018	INTRON	NA	12.58
chrM_2l_7341286	Wnt6	FBgn0031902	INTRON	NA	12.58
chrM_3l_5896481	CG32413	FBgn0052413	INTRON	NA	12.57
chrM_2r_16939478	shg	FBgn0003391	SYNONYMOUS_CODING	NA	12.56
chrM_2l_9440580	CG33723	FBgn0053723	SYNONYMOUS_CODING	NA	12.56
chrM_2l_7339170	Wnt6	FBgn0031902	INTRON	NA	12.55
chrM_2l_7339406	Wnt6	FBgn0031902	INTRON	NA	12.55
chrM_2l_7339683	Wnt6	FBgn0031902	INTRON	NA	12.55
chrM_2l_7339702	Wnt6	FBgn0031902	INTRON	NA	12.55
chrM_2l_7341804	Wnt6	FBgn0031902	INTRON	NA	12.55
chrM_2l_7341240	Wnt6	FBgn0031902	INTRON	NA	12.54
chrM_2l_7356644	Wnt6	FBgn0031902	DOWNSTREAM	4102	12.53
chrM_3l_11266076			SNP is more than 5000 bp away from known genes	NA	12.52
chrM_3r_22878347	NepYr	FBgn0004842	INTRON	NA	12.51
chrM_2l_9436881	RpS28-like	FBgn0085211	INTRON	NA	12.50
chrM_2l_9437018	RpS28-like	FBgn0085211	SYNONYMOUS_CODING	NA	12.50
chrM_2l_9437175	RpS28-like	FBgn0085211	NON_SYNONYMOUS_CODING	NA	12.50
chrM_2r_9630118	fas	FBgn0000633	INTRON	NA	12.48
chrM_3r_13491003	CG14322	FBgn0038532	INTRON	NA	12.47
chrM_2l_8604840	Sema-1a	FBgn0011259	INTRON	NA	12.46
chrM_3r_26750427	Ptx1	FBgn0020912	INTRON	NA	12.45
chrM_2l_2435535	dpp	FBgn0000490	INTRON	NA	12.44
chrM_3r_5439629	CG8312	FBgn0037720	INTRON	NA	12.42
chrM_2l_13230915	CG16815	FBgn0032491	DOWNSTREAM	195	12.41
chrM_2l_13230915	Prosalph6T	FBgn0032492	UPSTREAM	1124	12.41
chrM_2l_13230969	CG16815	FBgn0032491	DOWNSTREAM	249	12.41
chrM_2l_13230969	Prosalph6T	FBgn0032492	UPSTREAM	1070	12.41

chrM_2l_13224805	CG15480	FBgn0032489	SYNONYMOUS_CODING	NA	12.38
chrM_2l_13300620	CG31729	FBgn0051729	SYNONYMOUS_CODING	NA	12.35
chrM_2l_18335794	Fas3	FBgn0000636	INTRON	NA	12.35
chrM_2l_9440489	CG33723	FBgn0053723	NON_SYNONYMOUS_CODING	NA	12.34
chrM_2l_9441322	numb	FBgn0002973	INTRON	NA	12.34
chrM_2l_9442020	numb	FBgn0002973	INTRON	NA	12.34
chrM_2l_9442205	numb	FBgn0002973	INTRON	NA	12.34
chrM_3l_15066795	cp309	FBgn0086690	INTRON	NA	12.34
chrM_3r_11617903	c(3)G	FBgn0000246	NON_SYNONYMOUS_CODING	NA	12.33
chrM_3r_11619250	Acyp2	FBgn0038363	DOWNSTREAM	134	12.33
chrM_3r_11619250	wah	FBgn0262527	UPSTREAM	275	12.33
chrM_3r_11619668	wah	FBgn0262527	UTR_5_PRIME	NA	12.33
chrM_2l_7340121	Wnt6	FBgn0031902	INTRON	NA	12.33
chrM_2l_9442387	numb	FBgn0002973	INTRON	NA	12.32
chrM_3r_7651045	CG6959	FBgn0037956	INTRON	NA	12.31
chrM_2l_2546748	CG3515	FBgn0031431	DOWNSTREAM	4396	12.30
chrM_2l_13228575	CG16813	FBgn0032490	DOWNSTREAM	168	12.30
chrM_2l_13228575	CG16815	FBgn0032491	UPSTREAM	1067	12.30
chrM_2l_13228583	CG16815	FBgn0032491	UPSTREAM	1059	12.30
chrM_2l_13228583	CG16813	FBgn0032490	DOWNSTREAM	176	12.30
chrM_2l_13106310	CG5142	FBgn0032470	UPSTREAM	561	12.29
chrM_2l_13106330	CG5142	FBgn0032470	UPSTREAM	541	12.29
chrM_2r_5353777	Camta	FBgn0259234	INTRON	NA	12.29
chrM_2l_9442402	numb	FBgn0002973	INTRON	NA	12.29
chrM_2l_13207481	Pect	FBgn0032482	INTRON	NA	12.29
chrM_2l_13208062	Pect	FBgn0032482	INTRON	NA	12.29
chrM_2l_13208293	Pect	FBgn0032482	SYNONYMOUS_CODING	NA	12.29
chrM_2l_13210238	Pect	FBgn0032482	UTR_3_PRIME	NA	12.29
chrM_2l_13211137	CG15482	FBgn0032483	SYNONYMOUS_CODING	NA	12.29
chrM_2l_13211524	CG15482	FBgn0032483	SYNONYMOUS_CODING	NA	12.29
chrM_2l_13213641	kek4	FBgn0032484	NON_SYNONYMOUS_CODING	NA	12.29
chrM_2l_13213862	kek4	FBgn0032484	SYNONYMOUS_CODING	NA	12.29
chrM_2l_13213996	kek4	FBgn0032484	NON_SYNONYMOUS_CODING	NA	12.29
chrM_2l_13224922	CG15480	FBgn0032489	SYNONYMOUS_CODING	NA	12.29
chrM_2l_13225156	CG15480	FBgn0032489	NON_SYNONYMOUS_CODING	NA	12.29
chrM_2l_13225357	CG15480	FBgn0032489	UPSTREAM	151	12.29
chrM_2l_13225357	CG16812	FBgn0032488	DOWNSTREAM	807	12.29
chrM_2l_13228419	CG16815	FBgn0032491	UPSTREAM	1223	12.29
chrM_2l_13228419	CG16813	FBgn0032490	DOWNSTREAM	12	12.29
chrM_2l_13228557	CG16815	FBgn0032491	UPSTREAM	1085	12.29
chrM_2l_13228557	CG16813	FBgn0032490	DOWNSTREAM	150	12.29
chrM_2l_13229662	CG16815	FBgn0032491	START_GAINED	NA	12.29
chrM_2l_13230959	CG16815	FBgn0032491	DOWNSTREAM	239	12.26
chrM_2l_13230959	Prosalph6T	FBgn0032492	UPSTREAM	1080	12.26

chrM_2l_9437201	RpS28-like	FBgn0085211	UTR_3_PRIME	NA	12.26
chrM_2l_18203425	CG42750	FBgn0261804	INTRON	NA	12.26
chrM_3r_11621027	wah	FBgn0262527	SYNONYMOUS_CODING	NA	12.26
chrM_2l_13195793	Edem2	FBgn0032480	NON_SYNONYMOUS_CODING	NA	12.25
chrM_2l_13314585	CG43778	FBgn0264308	INTRON	NA	12.24
chrM_3r_11618661	Acyp2	FBgn0038363	INTRON	NA	12.23
chrM_3l_15170441			SNP is more than 5000 bp away from known genes	NA	12.22
chrM_3r_26745576	Ptx1	FBgn0020912	INTRON	NA	12.22
chrM_2l_13228742	CG16813	FBgn0032490	DOWNSTREAM	335	12.22
chrM_2l_13228742	CG16815	FBgn0032491	UPSTREAM	900	12.22
chrM_2l_13228654	CG16815	FBgn0032491	UPSTREAM	988	12.22
chrM_2l_13228654	CG16813	FBgn0032490	DOWNSTREAM	247	12.22
chrM_2l_13228904	CG16813	FBgn0032490	DOWNSTREAM	497	12.22
chrM_2l_13228904	CG16815	FBgn0032491	UPSTREAM	738	12.22
chrM_3l_3951603	CG14984	FBgn0035480	INTRON	NA	12.21
chrM_x_21896622	l(1)G0196	FBgn0027279	INTRON	NA	12.20
chrM_2l_2690903	CG31690	FBgn0051690	INTRON	NA	12.19
chrM_2l_7338067	Wnt6	FBgn0031902	INTRON	NA	12.19
chrM_2l_13311901	CG43778	FBgn0264308	INTRON	NA	12.18
chrM_2l_13235256	CG15479	FBgn0032493	UPSTREAM	623	12.17
chrM_2l_13235256	CG5867	FBgn0027586	UPSTREAM	1237	12.17
chrM_3l_14220524	nuf	FBgn0013718	INTRON	NA	12.16
chrM_3l_7751640	CG32373	FBgn0052373	NON_SYNONYMOUS_CODING	NA	12.16
chrM_x_21895406	l(1)G0196	FBgn0027279	SYNONYMOUS_CODING	NA	12.16
chrM_x_18070453	Andorra	FBgn0030898	DOWNSTREAM	2106	12.16
chrM_x_18070453	Frq2	FBgn0083228	UPSTREAM	1192	12.16
chrM_2l_16457967	Idgf3	FBgn0020414	DOWNSTREAM	4791	12.16
chrM_2l_13228978	CG16815	FBgn0032491	UPSTREAM	664	12.15
chrM_2l_13228978	CG16813	FBgn0032490	DOWNSTREAM	571	12.15
chrM_2l_13224925	CG15480	FBgn0032489	SYNONYMOUS_CODING	NA	12.15
chrM_2l_13207803	Pect	FBgn0032482	INTRON	NA	12.15
chrM_2l_7337340	Wnt6	FBgn0031902	INTRON	NA	12.14
chrM_2l_12371256	bru-2	FBgn0262475	INTRON	NA	12.14
chrM_x_13772755	mamo	FBgn0263115	INTRON	NA	12.14
chrM_3r_13032388			SNP is more than 5000 bp away from known genes	NA	12.14
chrM_3r_6320696			SNP is more than 5000 bp away from known genes	NA	12.12
chrM_3l_6262733			SNP is more than 5000 bp away from known genes	NA	12.12
chrM_3l_6262764			SNP is more than 5000 bp away from known genes	NA	12.12
chrM_2l_13301067	CG31729	FBgn0051729	UTR_3_PRIME	NA	12.12
chrM_3r_22133524			SNP is more than 5000 bp away	NA	12.11



			from known genes		
chr <sub>m</sub> _2l_6230925	Ugt37b1	FBgn0026755	DOWNSTREAM	4083	12.11
chr <sub>m</sub> _2l_13728140	CG9014	FBgn0028847	UPSTREAM	4498	12.09
chr <sub>m</sub> _2l_9672002			SNP is more than 5000 bp away from known genes	NA	12.08
chr <sub>m</sub> _3r_11615466	c(3)G	FBgn0000246	UTR_3_PRIME	NA	12.07
chr <sub>m</sub> _2r_11580458	CG30087	FBgn0050087	SYNONYMOUS_CODING	NA	12.07
chr <sub>m</sub> _3l_10657107	CG32068	FBgn0052068	INTRON	NA	12.07

**Supplementary Table 4** Gene ontology (GO) assignments for genes significantly associated with microbiota-dependent phenotypic traits. The number of GO categories assigned to a trait depended on the SNPs in question: for example, there were many GO assignments to the TAG-associated list, driven in particular by SNPs in IIS/TOR genes (principally *Happyhour*) and regulators of transcription. By contrast, only two significant GO assignments were made to glucose content, driven by one gene (*slamdance*), indicating that the strong GO enrichment associated with other traits is not an artefact of the relatively large number of input SNPs

GO category	Adjusted P-value	GO description	Associated genes
<b>(a) Dry weight</b>			
GO:0090254	0.0355	cell elongation involved in imaginal disc-derived wing morphogenesis	tkv
GO:0072663	0.0230	establishment of protein localization to peroxisome	Pex1
GO:0003840	0.0433	gamma-glutamyltransferase activity	CG1492
GO:0004466	0.0028	long-chain-acyl-CoA dehydrogenase activity	CG7461
GO:0007140	0.0028	male meiosis	bol
GO:0007127	0.0328	meiosis I	bol
GO:0044771	0.0010	meiotic cell cycle phase transition	bol
GO:0008315	0.0010	meiotic G2/M1 transition	bol
GO:0005771	0.0318	multivesicular body	tkv
GO:0045705	0.0461	negative regulation of salivary gland boundary specification	Abd-B tkv
GO:0016667	0.0010	oxidoreductase activity, acting on a sulfur group of donors	SelR
GO:0016671	0.0003	oxidoreductase activity, acting on a sulfur group of donors, disulfide as acceptor	SelR
GO:0033743	0.0003	peptide-methionine (R)-S-oxide reductase activity	SelR
GO:0018206	0.0003	peptidyl-methionine modification	SelR
GO:0043204	0.0003	perikaryon	bol
GO:0043574	0.0230	peroxisomal transport	Pex1
GO:0090068	0.0390	positive regulation of cell cycle process	bol
GO:0045836	0.0003	positive regulation of meiosis	bol
GO:0072662	0.0230	protein localization to peroxisome	Pex1
GO:0006625	0.0230	protein targeting to peroxisome	Pex1
GO:0040020	0.0003	regulation of meiosis	bol
GO:0051445	0.0003	regulation of meiotic cell cycle	bol
GO:0045704	0.0461	regulation of salivary gland boundary specification	Abd-B tkv
GO:0016246	0.0461	RNA interference	hdc Tis11
GO:0048137	0.0003	spermatocyte division	Pex1 bol
GO:0050431	0.0230	transforming growth factor beta binding	tkv
GO:0005025	0.0250	transforming growth factor beta receptor activity, type I	tkv
GO:0007181	0.0230	transforming growth factor beta receptor complex assembly	tkv
GO:0005024	0.0318	transforming growth factor beta-activated receptor activity	tkv
GO:0004675	0.0318	transmembrane receptor protein serine/threonine kinase activity	tkv
GO:0009826	0.0308	unidimensional cell growth	tkv Dok

GO:0017099	0.0008	very-long-chain-acyl-CoA dehydrogenase activity	CG7461
<b>(b) protein</b>			
GO:0031532	0.0361	actin cytoskeleton reorganization	nuf
GO:0008261	0.0003	allatostatin receptor activity	AICR2
GO:0021960	0.0139	anterior commissure morphogenesis	hh
GO:0048099	0.0344	anterior/posterior lineage restriction, imaginal disc	hh
GO:0061564	0.0037	axon development	trol rut tok beat-la hh EndoGl msn sm stan CG12717 Fas3 Sema-1a Trim9 kuz
GO:0007411	0.0049	axon guidance	trol tok beat-la EndoGl msn sm stan Fas3 Sema-1a Trim9 kuz
GO:0007409	0.0025	axonogenesis	trol rut tok beat-la hh EndoGl msn sm stan CG12717 Fas3 Sema-1a Trim9 kuz
GO:0055034	0.0418	Bolwig's organ development	beat-la hh
GO:0001746	0.0418	Bolwig's organ morphogenesis	beat-la hh
GO:0007420	0.0013	brain development	hh bun ed stan Sema-1a
GO:0016339	0.0115	calcium-dependent cell-cell adhesion	stan Fas3
GO:0071013	0.0092	catalytic step 2 spliceosome	CG13900
GO:0060326	0.0037	cell chemotaxis	hh wun
GO:0000904	0.0324	cell morphogenesis involved in differentiation	trol rut tok beat-la hh EndoGl msn

			sm stan CG12717 Fas3 Sema-1a Trim9 kuz
GO:0048667	0.0169	cell morphogenesis involved in neuron differentiation	trol rut tok beat-1a hh EndoGI msn sm stan CG12717 Fas3 Sema-1a Trim9 kuz
GO:0032990	0.0486	cell part morphogenesis	trol rut tok beat-1a hh EndoGI msn sm stan CG12717 Fas3 Sema-1a Trim9 kuz
GO:0048858	0.0486	cell projection morphogenesis	trol rut tok beat-1a hh EndoGI msn sm stan CG12717 Fas3 Sema-1a Trim9 kuz
GO:0008037	0.0493	cell recognition	trol tok beat-1a axo stan Fas3 Sema-1a Trim9
GO:0007166	0.0008	cell surface receptor signaling pathway	Proc-R trol rut wek hh kek5 rdgA bun ed EndoGI aru stan Sema-1a

			Trim9 AICR2 kuz
GO:0005911	0.0055	cell-cell junction	vari ed wun stan Fas3
GO:0006928	0.0198	cellular component movement	trol tok beat-1a apt hh bun CG13636 EndoGI msn sm wun stan Fas3 Sema-1a Trim9 kuz
GO:0022412	0.0404	cellular process involved in reproduction in multicellular organism	wek apt NA hh bun CG13636 ed wun msi Fas3 kuz
GO:0051716	0.0301	cellular response to stimulus	Proc-R elk trol rut rut wek CG14669 apt hh kek5 CG30158 rdgA bun ed EndoGI msn aru wun msi stan Sema-1a Trim9 AICR2 kuz
GO:0007417	0.0002	central nervous system development	apt hh bun ed stan CG12717 Sema-1a kuz
GO:0021955	0.0018	central nervous system neuron axonogenesis	hh CG12717

GO:0021954	0.0024	central nervous system neuron development	hh CG12717
GO:0021953	0.0024	central nervous system neuron differentiation	hh CG12717
GO:0021952	0.0015	central nervous system projection neuron axonogenesis	hh CG12717
GO:0006935	0.0008	chemotaxis	trol tok beat-1a hh EndoG1 msn sm wun stan Fas3 Sema-1a Trim9 kuz
GO:0008234	0.0325	cysteine-type peptidase activity	hh CG12717
GO:0003684	0.0002	damaged DNA binding	CG13900
GO:0070593	0.0027	dendrite self-avoidance	stan
GO:0004175	0.0161	endopeptidase activity	tok CG5527 hh kuz
GO:0030900	0.0139	forebrain development	hh
GO:0004930	0.0316	G-protein coupled receptor activity	Proc-R stan AICR2
GO:0035232	0.0404	germ cell attraction	hh
GO:0035233	0.0117	germ cell repulsion	wun
GO:0008347	0.0090	glial cell migration	hh kuz
GO:0060914	0.0288	heart formation	hh
GO:0003007	0.0008	heart morphogenesis	hh kuz
GO:0042802	0.0117	identical protein binding	bru-3 wek bun Pdp1 wun
GO:0016539	0.0114	intein-mediated protein splicing	hh
GO:0040011	0.0347	locomotion	trol tok beat-1a apt hh rdgA bun CG13636 EndoG1 msn sm wun stan Fas3 Sema-1a Trim9 kuz
GO:0007311	0.0002	maternal specification of dorsal/ventral axis, oocyte, germ-line encoded	wek
GO:0006509	0.0002	membrane protein ectodomain proteolysis	kuz
GO:0033619	0.0002	membrane protein proteolysis	kuz
GO:0004222	0.0018	metalloendopeptidase activity	tok CG5527

			kuz
GO:0008237	0.0417	metallopeptidase activity	tok CG5527 kuz
GO:0016319	0.0060	mushroom body development	bun ed stan
GO:0030178	0.0300	negative regulation of Wnt receptor signaling pathway	stan
GO:0014017	0.0417	neuroblast fate commitment	kuz
GO:0007400	0.0271	neuroblast fate determination	kuz
GO:0097485	0.0055	neuron projection guidance	trol tok beat-1a EndoGI msn sm stan Fas3 Sema-1a Trim9 kuz
GO:0048812	0.0404	neuron projection morphogenesis	trol rut tok beat-1a hh EndoGI msn sm stan CG12717 Fas3 Sema-1a Trim9 kuz
GO:0008038	0.0491	neuron recognition	trol tok beat-1a axo stan Fas3 Sema-1a Trim9
GO:0007218	0.0002	neuropeptide signaling pathway	Proc-R stan AICR2
GO:0007220	0.0007	Notch receptor processing	kuz
GO:0007219	0.0329	Notch signaling pathway	bun EndoGI kuz
GO:0070160	0.0218	occluding junction	vari wun Fas3
GO:0004984	0.0071	olfactory receptor activity	AICR2
GO:0030707	0.0440	ovarian follicle cell development	apt hh bun CG13636 ed Fas3 kuz
GO:0005113	0.0125	patched binding	hh
GO:0061320	0.0066	pericardial nephrocyte differentiation	kuz
GO:0007365	0.0018	periodic partitioning	hh stan
GO:0010647	0.0364	positive regulation of cell communication	trol hh

			EndoG stan
GO:0007228	0.0449	positive regulation of hh target transcription factor activity	hh
GO:0009967	0.0300	positive regulation of signal transduction	trol hh EndoG stan
GO:0023056	0.0320	positive regulation of signaling	trol hh EndoG stan
GO:0030177	0.0023	positive regulation of Wnt receptor signaling pathway	stan
GO:0071011	0.0172	precatalytic spliceosome	CG13900
GO:0016540	0.0273	protein autoprocessing	hh
GO:0042803	0.0065	protein homodimerization activity	bru-3 wek bun Pdp1 wun
GO:0030908	0.0114	protein splicing	hh
GO:0048057	0.0002	R3/R4 development	stan
GO:0005057	0.0316	receptor signaling protein activity	msn stan
GO:0090066	0.0006	regulation of anatomical structure size	vari stan Sema-1a kuz
GO:0030516	0.0320	regulation of axon extension	kuz
GO:0050770	0.0002	regulation of axonogenesis	stan Sema-1a kuz
GO:0060284	0.0440	regulation of cell development	bun ed stan Sema-1a kuz
GO:0001558	0.0198	regulation of cell growth	bun kuz
GO:0022604	0.0025	regulation of cell morphogenesis	ed msn stan Sema-1a kuz
GO:0010769	0.0003	regulation of cell morphogenesis involved in differentiation	stan Sema-1a kuz
GO:0031344	0.0007	regulation of cell projection organization	stan Sema-1a kuz
GO:0008361	0.0209	regulation of cell size	Sema-1a kuz
GO:0032535	0.0347	regulation of cellular component size	Sema-1a kuz
GO:0090175	0.0092	regulation of establishment of planar polarity	stan
GO:0061387	0.0320	regulation of extent of cell growth	kuz
GO:0051960	0.0092	regulation of nervous system development	bun ed stan Sema-1a kuz
GO:0050767	0.0007	regulation of neurogenesis	bun ed stan Sema-1a kuz
GO:0045664	0.0028	regulation of neuron differentiation	stan



			Sema-1a kuz
GO:0010975	0.0003	regulation of neuron projection development	stan Sema-1a kuz
GO:0035386	0.0002	regulation of Roundabout signaling pathway	kuz
GO:0035159	0.0023	regulation of tube length, open tracheal system	vari stan
GO:0035150	0.0117	regulation of tube size	vari stan
GO:0035151	0.0057	regulation of tube size, open tracheal system	vari stan
GO:0042221	0.0347	response to chemical stimulus	trol rut tok beat-1a apt hh Pdp1 EndoGI msn wtrw sm aru wun msi stan Fas3 Sema-1a Trim9 kuz
GO:0009605	0.0169	response to external stimulus	trol tok beat-1a hh rdgA Pdp1 EndoGI msn sm wun stan Fas3 Sema-1a Trim9 kuz
GO:0030529	0.0323	ribonucleoprotein complex	CG13900 sm aru
GO:0035385	0.0018	Roundabout signaling pathway	kuz
GO:0007367	0.0004	segment polarity determination	hh stan
GO:0071526	0.0418	semaphorin-plexin signaling pathway	trol Sema-1a
GO:0005918	0.0139	septate junction	vari wun Fas3
GO:0019991	0.0345	septate junction assembly	vari wun
GO:0007165	0.0347	signal transduction	Proc-R elk trol rut wek CG14669 apt hh kek5

			CG30158 rdgA bun ed EndoGI msn aru stan Sema-1a Trim9 AICR2 kuz
GO:0007227	0.0449	signal transduction downstream of smoothened	hh
GO:0030532	0.0002	small nuclear ribonucleoprotein complex	CG13900
GO:0004994	0.0002	somatostatin receptor activity	AICR2
GO:0005681	0.0316	spliceosomal complex	CG13900
GO:0048867	0.0440	stem cell fate determination	kuz
GO:0008146	0.0094	sulfotransferase activity	CG9164
GO:0042330	0.0018	taxis	trol tok beat-1a hh rdgA EndoGI msn sm wun stan Fas3 Sema-1a Trim9 kuz
GO:0021537	0.0139	telencephalon development	hh
GO:0016782	0.0109	transferase activity, transferring sulfur- containing groups	CG9164
GO:0005686	0.0002	U2 snRNP	CG13900
GO:0031532	0.0361	actin cytoskeleton reorganization	nuf
GO:0008261	0.0003	allatostatin receptor activity	AICR2
<b>(c) TAG</b>			
GO:0003873	0.0040	6-phosphofructo-2-kinase activity	Pfrx
GO:0055090	0.0001	acylglycerol homeostasis	hppy
GO:0030534	0.0353	adult behavior	Hk hiw pum hppy
GO:0008106	0.0071	alcohol dehydrogenase (NADP+) activity	CG6084
GO:0004032	0.0064	alditol:NADP+ 1-oxidoreductase activity	CG6084
GO:0004033	0.0071	aldo-keto reductase (NADP) activity	CG6084
GO:0019438	0.0029	aromatic compound biosynthetic process	grh crol Snoo bs AP-2 Doc2 hiw pum Pdf inv sbb RpIII128
GO:0007412	0.0296	axon target recognition	sbb
GO:0007610	0.0216	behavior	Hk mnb Gef64C S6k

			Sap47 hiw mld pum Pdfr sbb hppy be orb2
GO:0009058	0.0287	biosynthetic process	grh crol Snoo bs AP-2 Doc2 hiw mld pum app Pdfr inv sbb RpIII128 orb2
GO:0097367	0.0016	carbohydrate derivative binding	mnb CG8641 S6k RpIII128 hppy Pfrx
GO:0019200	0.0305	carbohydrate kinase activity	Pfrx
GO:0019203	0.0278	carbohydrate phosphatase activity	Pfrx
GO:0006725	0.0429	cellular aromatic compound metabolic process	sif Fancd2 grh crol CG8641 Snoo bs AP-2 Doc2 hiw Drep-1 pum Pdfr inv sbb RpIII128
GO:0044249	0.0216	cellular biosynthetic process	grh crol Snoo bs AP-2 Doc2 hiw mld pum app Pdfr inv sbb RpIII128 orb2
GO:0034645	0.0053	cellular macromolecule biosynthetic process	grh crol Snoo bs AP-2

			Doc2 hiw pum app inv sbb RpIII128 orb2
GO:0044260	0.0027	cellular macromolecule metabolic process	Fancd2 grh crol mnb SNF4Agamma S6k Snoo bs AP-2 Doc2 hiw Drep-1 pum SelR app inv sbb Ptp99A RpIII128 hppy Spase12 orb2
GO:0044237	0.0069	cellular metabolic process	sif Fancd2 grh crol mnb SNF4Agamma CG8641 S6k Pi3K68D Snoo CG6428 bs AP-2 Doc2 hiw mld Drep-1 pum SelR app Pdfr inv sbb Ptp99A RpIII128 hppy Spase12 Pfrx orb2
GO:0044271	0.0040	cellular nitrogen compound biosynthetic process	grh crol Snoo bs AP-2 Doc2 hiw pum

			Pdfr inv sbb RpIII128
GO:0034641	0.0251	cellular nitrogen compound metabolic process	sif Fancd2 grh crol CG8641 Snoo bs AP-2 Doc2 hiw Drep-1 pum Pdfr inv sbb RpIII128 Spase12
GO:0032869	0.0040	cellular response to insulin stimulus	melt hppy
GO:1901699	0.0059	cellular response to nitrogen compound	melt hppy
GO:0071417	0.0048	cellular response to organonitrogen compound	melt hppy
GO:1901701	0.0445	cellular response to oxygen-containing compound	melt hppy
GO:1901653	0.0040	cellular response to peptide	melt hppy
GO:0071375	0.0040	cellular response to peptide hormone stimulus	melt hppy
GO:0048878	0.0445	chemical homeostasis	SNF4Agamma hppy
GO:0003677	0.0345	DNA binding	grh bs AP-2 inv RpIII128
GO:0003899	0.0001	DNA-directed RNA polymerase activity	RpIII128
GO:0000428	0.0001	DNA-directed RNA polymerase complex	RpIII128
GO:0005666	0.0001	DNA-directed RNA polymerase III complex	RpIII128
GO:0006003	0.0040	fructose 2,6-bisphosphate metabolic process	Pfrx
GO:0006000	0.0040	fructose metabolic process	Pfrx
GO:0004331	0.0040	fructose-2,6-bisphosphate 2-phosphatase activity	Pfrx
GO:0010467	0.0364	gene expression	grh crol Snoo bs AP-2 Doc2 hiw pum inv sbb RpIII128 Spase12 orb2
GO:0007030	0.0031	Golgi organization	hppy
GO:0018130	0.0029	heterocycle biosynthetic process	grh crol Snoobs AP-2 Doc2 hiw pum

			Pdfr inv sbb RpIII128
GO:0046483	0.0296	heterocycle metabolic process	sif Fancd2 grh crol CG8641 Snoo bs AP-2 Doc2 hiw Drep-1 pum Pdfr inv sbb RpIII128
GO:0008286	0.0040	insulin receptor signaling pathway	melt hppy
GO:0070013	0.0001	intracellular organelle lumen	RpIII128
GO:0030537	0.0296	larval behavior	S6k sbb
GO:0055088	0.0021	lipid homeostasis	SNF4Agamma hppy
GO:0009059	0.0054	macromolecule biosynthetic process	grh crol Snoo bs AP-2 Doc2 hiw pum app inv sbb RpIII128 orb2
GO:0043170	0.0392	macromolecule metabolic process	Fancd2 grh crol mnb SNF4Agamma S6k Snoo bs AP-2 Doc2 hiw Drep-1 pum SelR app inv sbb Ptp99A RpIII128 hppy Spase12 orb2
GO:0031974	0.0001	membrane-enclosed lumen	RpIII128
GO:0010648	0.0148	negative regulation of cell communication	crol Snoo hiw pum

			sbb hppy
GO:0042059	0.0149	negative regulation of epidermal growth factor receptor signaling pathway	pum hppy
GO:1901185	0.0149	negative regulation of ERBB signaling pathway	pum hppy
GO:0048585	0.0136	negative regulation of response to stimulus	crol Snoo hiw pum sbb hppy
GO:0009968	0.0095	negative regulation of signal transduction	crol Snoo hiw pum sbb hppy
GO:0023057	0.0148	negative regulation of signaling	crol Snoo hiw pum sbb hppy
GO:0045879	0.0080	negative regulation of smoothened signaling pathway	sbb
GO:0006807	0.0475	nitrogen compound metabolic process	sif Fancd2 grh crol CG8641 Snoo CG6428 bs AP-2 Doc2 hiw Drep-1 pum Pdfr inv sbb RpIII128 Spase12
GO:0055029	0.0001	nuclear DNA-directed RNA polymerase complex	RpIII128
GO:0031981	0.0001	nuclear lumen	RpIII128
GO:0044428	0.0050	nuclear part	RpIII128
GO:0090304	0.0097	nucleic acid metabolic process	Fancd2 grh crol Snoo bs AP-2 Doc2 hiw Drep-1 pum inv sbb RpIII128
GO:0034654	0.0021	nucleobase-containing compound biosynthetic process	grh crol Snoo bs AP-2 Doc2 hiw

			pum Pdfr inv sbb RpIII128
GO:0006139	0.0222	nucleobase-containing compound metabolic process	sif Fancd2 grh crol CG8641 Snoo bs AP-2 Doc2 hiw Drep-1 pum Pdfr inv sbb RpIII128
GO:0005654	0.0001	nucleoplasm	RpIII128
GO:0044451	0.0001	nucleoplasm part	RpIII128
GO:0001882	0.0001	nucleoside binding	mnb CG8641 S6k RpIII128 hppy Pfrx
GO:0016779	0.0001	nucleotidyltransferase activity	RpIII128
GO:0043233	0.0001	organelle lumen	RpIII128
GO:1901362	0.0021	organic cyclic compound biosynthetic process	grh crol Snoo bs AP-2 Doc2 hiw mld pum Pdfr inv sbb RpIII128
GO:1901360	0.0378	organic cyclic compound metabolic process	sif Fancd2 grh crol CG8641 Snoo bs AP-2 Doc2 hiw mld Drep-1 pum Pdfr inv sbb RpIII128
GO:1901576	0.0254	organic substance biosynthetic process	grh crol Snoo bs AP-2 Doc2 hiw



			mld pum app Pdfr inv sbb RpIII128 orb2
GO:0008443	0.0084	phosphofructokinase activity	Pfrx
GO:0043065	0.0144	positive regulation of apoptotic process	hppy
GO:0010740	0.0378	positive regulation of intracellular protein kinase cascade	hppy
GO:0046330	0.0022	positive regulation of JNK cascade	hppy
GO:0043410	0.0140	positive regulation of MAPK cascade	hppy
GO:0032874	0.0022	positive regulation of stress-activated MAPK cascade	hppy
GO:0070304	0.0022	positive regulation of stress-activated protein kinase signaling cascade	hppy
GO:0032008	0.0001	positive regulation of TOR signaling cascade	hppy
GO:0015459	0.0158	potassium channel regulator activity	Hk
GO:0044238	0.0205	primary metabolic process	sif Fancd2 grh crol mnb SNF4Agamma CG8641 S6k Pi3K68D Snoo CG6428 bs AP-2 Doc2 hiw mld Drep-1 pum SelR app Pdfr inv sbb Ptp99A RpIII128 hppy Spase12 Pfrx orb2
GO:0005057	0.0217	receptor signaling protein activity	sif hppy
GO:0004702	0.0063	receptor signaling protein serine/threonine kinase activity	hppy
GO:1900076	0.0097	regulation of cellular response to insulin stimulus	hppy
GO:0080135	0.0086	regulation of cellular response to stress	melt SNF4Agamma S6k hppy
GO:0042058	0.0305	regulation of epidermal growth factor receptor signaling pathway	pum hppy
GO:1901184	0.0305	regulation of ERBB signaling pathway	pum hppy
GO:0040008	0.0305	regulation of growth	S6k hiw pum sbb hppy

GO:0040009	0.0001	regulation of growth rate	hppy
GO:0046626	0.0097	regulation of insulin receptor signaling pathway	hppy
GO:0046328	0.0296	regulation of JNK cascade	hppy
GO:0080134	0.0339	regulation of response to stress	melt grh SNF4Agamma S6k hppy
GO:0008589	0.0205	regulation of smoothed signaling pathway	sbb
GO:0032872	0.0297	regulation of stress-activated MAPK cascade	hppy
GO:0070302	0.0297	regulation of stress-activated protein kinase signaling cascade	hppy
GO:0032006	0.0003	regulation of TOR signaling cascade	hppy
GO:0032868	0.0040	response to insulin stimulus	melt hppy
GO:1901698	0.0297	response to nitrogen compound	melt hppy
GO:0010243	0.0246	response to organonitrogen compound	melt hppy
GO:1901652	0.0040	response to peptide	melt hppy
GO:0043434	0.0040	response to peptide hormone stimulus	melt hppy
GO:0032549	0.0001	ribonucleoside binding	mnb CG8641 S6k RpIII128 hppy Pfrx
GO:0032774	0.0009	RNA biosynthetic process	grh crol Snoo bs AP-2 Doc2 hiw pum inv sbb RpIII128
GO:0016070	0.0133	RNA metabolic process	grh crol Snoo bs AP-2 Doc2 hiw pum inv sbb RpIII128
GO:0034062	0.0001	RNA polymerase activity	RpIII128
GO:0030880	0.0001	RNA polymerase complex	RpIII128
GO:0044708	0.0020	single-organism behavior	Hk mnb S6k Sap47 hiw mld pum Pdfr sbb hppy be orb2
GO:0036094	0.0053	small molecule binding	mnb

			SNF4Agamma CG8641 S6k Snoo RpIII128 hppy Pfrx orb2
GO:0007224	0.0392	smoothened signaling pathway	sbb
GO:0005214	0.0001	structural constituent of chitin-based cuticle	Cpr65Ay
GO:0042302	0.0001	structural constituent of cuticle	Cpr65Ay
GO:0050308	0.0070	sugar-phosphatase activity	Pfrx
GO:0031929	0.0012	TOR signaling cascade	hppy
GO:0031931	0.0003	TORC1 complex	hppy
GO:0003714	0.0020	transcription corepressor activity	sbb
GO:0006383	0.0001	transcription from RNA polymerase III promoter	RpIII128
GO:0006351	0.0009	transcription, DNA-dependent	grh crol Snoo bs AP-2 Doc2 hiw pum inv sbb RpIII128
GO:0016740	0.0001	transferase activity	mnb SNF4Agamma S6k Pi3K68D app RpIII128 hppy Pfrx
GO:0016772	0.0001	transferase activity, transferring phosphorus-containing groups	mnb SNF4Agamma S6k Pi3K68D RpIII128 hppy Pfrx
GO:0070328	0.0001	triglyceride homeostasis	hppy
GO:0045479	0.0347	vesicle targeting to fusome	bam
GO:0003873	0.0040	6-phosphofructo-2-kinase activity	Pfrx
<b>(d) glucose</b>			
GO:0004177	0.0369	aminopeptidase activity	sda
GO:0007638	0.0192	mechanosensory behavior	sda
<b>(e) glycogen</b>			
GO:0010309	0.0239	acireductone dioxygenase [iron(II)-requiring] activity	CG32068
GO:0006066	0.0016	alcohol metabolic process	Pect
GO:0009308	0.0001	amine metabolic process	Pect
GO:0009952	0.0002	anterior/posterior pattern specification	dpp shg Wnt6
GO:0008356	0.0007	asymmetric cell division	numb dpp shg
GO:0055059	0.0000	asymmetric neuroblast division	numb
GO:0009798	0.0000	axis specification	dpp shg Wnt6
GO:0045180	0.0000	basal cortex	numb

GO:0045178	0.0000	basal part of cell	numb
GO:0009058	0.0057	biosynthetic process	svp RpS28-like dpp bru-2 Pect Ptx1 Cnx99A Wnt6 Camta dnr1 CG32068
GO:0006816	0.0117	calcium ion transport	Nckx30C Ca-alpha1D Cnx99A CG9297
GO:0060070	0.0000	canonical Wnt receptor signaling pathway	Wnt6
GO:0060911	0.0034	cardiac cell fate commitment	svp fas
GO:0060913	0.0000	cardiac cell fate determination	svp fas
GO:0042684	0.0009	cardioblast cell fate commitment	svp fas
GO:0007510	0.0000	cardioblast cell fate determination	svp fas
GO:0010002	0.0003	cardioblast differentiation	svp dpp fas
GO:0035051	0.0000	cardiocyte differentiation	numb svp dpp fas
GO:0072358	0.0000	cardiovascular system development	numb svp dpp shg fas
GO:0005938	0.0243	cell cortex	numb
GO:0044448	0.0003	cell cortex part	numb
GO:0051301	0.0029	cell division	numb dpp shg nuf
GO:0045165	0.0313	cell fate commitment	numb svp dpp tutl fas nuf
GO:0060581	0.0000	cell fate commitment involved in pattern specification	numb nuf
GO:0001709	0.0000	cell fate determination	numb svp dpp fas nuf
GO:0060582	0.0000	cell fate determination involved in pattern specification	numb nuf
GO:0001708	0.0113	cell fate specification	numb dpp
GO:0071944	0.0197	cell periphery	numb NepYr Ca-alpha1D shg kek4 Cnx99A Wnt6

			CadN2 Fas3 Sema-1a
GO:0008283	0.0005	cell proliferation	numb svp dpp wah
GO:0007166	0.0475	cell surface receptor signaling pathway	numb dpp NepYr Wnt6 Camta Sema-1a
GO:0044106	0.0001	cellular amine metabolic process	Pect
GO:0006576	0.0000	cellular biogenic amine metabolic process	Pect
GO:0071396	0.0198	cellular response to lipid	svp
GO:0071407	0.0234	cellular response to organic cyclic compound	svp
GO:0071383	0.0050	cellular response to steroid hormone stimulus	svp
GO:0007417	0.0007	central nervous system development	numb svp shg chinmo Sema-1a
GO:0072359	0.0000	circulatory system development	numb svp dpp shg fas
GO:0070567	0.0000	cytidyltransferase activity	Pect
GO:0002181	0.0000	cytoplasmic translation	RpS28-like
GO:0044445	0.0029	cytosolic part	RpS28-like
GO:0022626	0.0001	cytosolic ribosome	RpS28-like
GO:0022627	0.0000	cytosolic small ribosomal subunit	RpS28-like
GO:0055123	0.0348	digestive system development	svp dpp shg fas
GO:0048565	0.0348	digestive tract development	svp dpp shg fas
GO:0048546	0.0155	digestive tract morphogenesis	svp dpp fas
GO:0072511	0.0119	divalent inorganic cation transport	Nckx30C Ca-alpha1D Cnx99A CG9297
GO:0070838	0.0119	divalent metal ion transport	Nckx30C Ca-alpha1D Cnx99A CG9297
GO:0035050	0.0000	embryonic heart tube development	numb
GO:0048619	0.0075	embryonic hindgut morphogenesis	svp dpp fas
GO:0048568	0.0000	embryonic organ development	numb shg
GO:0006897	0.0258	endocytosis	numb wah
GO:0009453	0.0001	energy taxis	svp
GO:0050673	0.0012	epithelial cell proliferation	svp
GO:0061331	0.0002	epithelial cell proliferation involved in Malpighian tubule morphogenesis	svp
GO:2001013	0.0002	epithelial cell proliferation involved in renal tubule morphogenesis	svp
GO:0042439	0.0000	ethanolamine-containing compound	Pect

		metabolic process	
GO:0004306	0.0000	ethanolamine-phosphate cytidyltransferase activity	Pect
GO:0044421	0.0003	extracellular region part	dpp CG5867 Wnt6
GO:0005615	0.0000	extracellular space	dpp CG5867 Wnt6
GO:0007503	0.0000	fat body development	svp
GO:0048859	0.0001	formation of anatomical boundary	numb dpp nuf
GO:0010160	0.0000	formation of organ boundary	numb dpp nuf
GO:0001664	0.0000	G-protein coupled receptor binding	Wnt6
GO:0008347	0.0003	glial cell migration	numb
GO:0042063	0.0034	gliogenesis	numb
GO:0007507	0.0000	heart development	numb svp dpp shg fas
GO:0061525	0.0117	hindgut development	svp dpp fas
GO:0007442	0.0117	hindgut morphogenesis	svp dpp fas
GO:0035035	0.0006	histone acetyltransferase binding	wah
GO:0009755	0.0155	hormone-mediated signaling pathway	svp
GO:0042802	0.0219	identical protein binding	svp dpp shg Camta
GO:0004879	0.0097	ligand-activated sequence-specific DNA binding RNA polymerase II transcription factor activity	svp
GO:0072002	0.0000	Malpighian tubule development	numb svp dpp fas
GO:0007443	0.0053	Malpighian tubule morphogenesis	svp dpp fas
GO:0061382	0.0000	Malpighian tubule tip cell differentiation	numb
GO:0042693	0.0000	muscle cell fate commitment	numb
GO:0042694	0.0000	muscle cell fate specification	numb
GO:0048519	0.0335	negative regulation of biological process	numb dpp shg bru-2 wah Wnt6 dnr1
GO:0009890	0.0197	negative regulation of biosynthetic process	bru-2 Wnt6 dnr1
GO:0031327	0.0197	negative regulation of cellular biosynthetic process	bru-2 Wnt6 dnr1
GO:0048523	0.0082	negative regulation of cellular process	numb dpp shg bru-2 wah

			Wnt6 dnr1
GO:0010629	0.0372	negative regulation of gene expression	dpp Wnt6
GO:0051172	0.0054	negative regulation of nitrogen compound metabolic process	Wnt6 dnr1
GO:0045746	0.0000	negative regulation of Notch signaling pathway	numb
GO:0045934	0.0148	negative regulation of nucleobase-containing compound metabolic process	Wnt6
GO:0051253	0.0115	negative regulation of RNA metabolic process	Wnt6
GO:0000122	0.0005	negative regulation of transcription from RNA polymerase II promoter	Wnt6
GO:0045892	0.0107	negative regulation of transcription, DNA-dependent	Wnt6
GO:0061319	0.0000	nephrocyte differentiation	numb dpp
GO:0061351	0.0001	neural precursor cell proliferation	numb
GO:0014016	0.0002	neuroblast differentiation	numb
GO:0055057	0.0000	neuroblast division	numb
GO:0014017	0.0000	neuroblast fate commitment	numb
GO:0007400	0.0000	neuroblast fate determination	numb
GO:0007405	0.0001	neuroblast proliferation	numb
GO:0007270	0.0000	neuron-neuron synaptic transmission	svp
GO:0005112	0.0000	Notch binding	numb
GO:0007219	0.0055	Notch signaling pathway	numb
GO:0044545	0.0031	NSL complex	wah
GO:0016779	0.0000	nucleotidyltransferase activity	Pect
GO:0048645	0.0000	organ formation	numb dpp nuf
GO:0007389	0.0007	pattern specification process	numb dpp shg Wnt6 CadN2 nuf
GO:0061320	0.0000	pericardial nephrocyte differentiation	numb dpp
GO:0007422	0.0042	peripheral nervous system development	numb tutl
GO:0042331	0.0001	phototaxis	svp
GO:0009891	0.0232	positive regulation of biosynthetic process	Wnt6 Camta
GO:0031328	0.0232	positive regulation of cellular biosynthetic process	Wnt6 Camta
GO:0051130	0.0085	positive regulation of cellular component organization	numb
GO:0045807	0.0000	positive regulation of endocytosis	numb
GO:0010628	0.0054	positive regulation of gene expression	Wnt6 Camta
GO:0010557	0.0096	positive regulation of macromolecule biosynthetic process	Wnt6 Camta
GO:0010604	0.0377	positive regulation of macromolecule metabolic process	Wnt6 Camta
GO:0051173	0.0195	positive regulation of nitrogen compound metabolic process	Wnt6 Camta
GO:0045935	0.0117	positive regulation of nucleobase-containing compound metabolic process	Wnt6 Camta
GO:0051254	0.0075	positive regulation of RNA metabolic process	Wnt6 Camta
GO:0045893	0.0047	positive regulation of transcription, DNA-dependent	Wnt6 Camta
GO:0051050	0.0000	positive regulation of transport	numb
GO:0005515	0.0093	protein binding	numb svp dpp shg

			wah Cnx99A Edem2 Wnt6 Camta KLHL18 CadN2 CG16815 Sema-1a
GO:0046982	0.0070	protein heterodimerization activity	nuf svp dpp
GO:0042803	0.0114	protein homodimerization activity	svp dpp shg Camta
GO:0048052	0.0000	R1/R6 cell differentiation	svp
GO:0007462	0.0000	R1/R6 cell fate commitment	svp
GO:0048056	0.0157	R3/R4 cell differentiation	svp
GO:0007464	0.0156	R3/R4 cell fate commitment	svp
GO:0045466	0.0468	R7 cell differentiation	svp CadN2
GO:0007465	0.0011	R7 cell fate commitment	svp
GO:0005102	0.0000	receptor binding	numb dpp Wnt6 Sema-1a
GO:0003002	0.0002	regionalization	numb dpp shg Wnt6 nuf
GO:0009786	0.0000	regulation of asymmetric cell division	numb
GO:0051302	0.0000	regulation of cell division	numb
GO:0030100	0.0000	regulation of endocytosis	numb wah
GO:0032879	0.0138	regulation of localization	numb wah Cnx99A Sema-1a
GO:0008593	0.0004	regulation of Notch signaling pathway	numb
GO:0006357	0.0187	regulation of transcription from RNA polymerase II promoter	dpp Ptx1 Wnt6 Camta
GO:0051049	0.0004	regulation of transport	numb wah Cnx99A
GO:0060627	0.0000	regulation of vesicle-mediated transport	numb wah
GO:0061318	0.0000	renal filtration cell differentiation	numb dpp
GO:0072001	0.0000	renal system development	numb svp dpp fas
GO:0061326	0.0000	renal tubule development	numb svp dpp fas
GO:0061333	0.0053	renal tubule morphogenesis	svp dpp fas
GO:0033993	0.0377	response to lipid	svp
GO:0048545	0.0050	response to steroid hormone stimulus	svp
GO:0007622	0.0027	rhythmic behavior	numb
GO:0048511	0.0218	rhythmic process	numb
GO:0044391	0.0009	ribosomal subunit	RpS28-like



GO:0005840	0.0024	ribosome	RpS28-like
GO:0000982	0.0156	RNA polymerase II core promoter proximal region sequence-specific DNA binding transcription factor activity	svp Camta
GO:0001078	0.0005	RNA polymerase II core promoter proximal region sequence-specific DNA binding transcription factor activity involved in negative regulation of transcription	svp
GO:0001227	0.0005	RNA polymerase II transcription regulatory region sequence-specific DNA binding transcription factor activity involved in negative regulation of transcription	svp
GO:0008052	0.0000	sensory organ boundary specification	numb nuf
GO:0045035	0.0000	sensory organ precursor cell division	numb
GO:0016360	0.0000	sensory organ precursor cell fate determination	numb nuf
GO:0015935	0.0001	small ribosomal subunit	RpS28-like
GO:0048103	0.0000	somatic stem cell division	numb shg
GO:0048863	0.0000	stem cell differentiation	numb svp dpp shg wah fas
GO:0017145	0.0003	stem cell division	numb dpp shg
GO:0048865	0.0001	stem cell fate commitment	numb
GO:0048867	0.0000	stem cell fate determination	numb
GO:0072089	0.0002	stem cell proliferation	numb
GO:0043401	0.0050	steroid hormone mediated signaling pathway	svp
GO:0003707	0.0055	steroid hormone receptor activity	svp
GO:0003735	0.0017	structural constituent of ribosome	RpS28-like
GO:0000795	0.0011	synaptonemal complex	c(3)G
GO:0006366	0.0168	transcription from RNA polymerase II promoter	dpp Ptx1 Wnt6 Camta
GO:0035295	0.0001	tube development	numb svp dpp shg fas
GO:0001655	0.0000	urogenital system development	numb svp dpp fas
GO:0007419	0.0000	ventral cord development	numb svp
GO:0016055	0.0002	Wnt receptor signaling pathway	Wnt6
GO:0010309	0.0239	acireductone dioxygenase [iron(II)-requiring] activity	CG32068
GO:0006066	0.0016	alcohol metabolic process	Pect
GO:0009308	0.0001	amine metabolic process	Pect
GO:0009952	0.0002	anterior/posterior pattern specification	dpp shg Wnt6

**Supplementary Table 5 Genes with SNPs significantly associated with more than one microbiota-dependent trait**

	Protein	TAG	Glucose	Glycogen
Dry weight	<i>CG32982</i> <i>CG43244</i> <i>CG43316</i> <i>hh</i> <i>kek5</i> <i>rdgA</i> <i>trol</i>	<i>CG42663</i> <i>grh</i> <i>SelR</i>	<i>Ald</i> <i>CG42324</i> <i>CR42745</i> <i>hdc</i> <i>sda</i> <i>Sema-5c</i> <i>Tusp</i>	<i>KLHL18</i>
Protein		<i>CG42629</i>	<i>dpr6</i> <i>Dys</i> <i>jhamt</i>	<i>Ca-alpha1D</i> <i>CadN2</i> <i>CG5142</i> <i>Fas3</i> <i>ImpL3</i> <i>nuf</i>
TAG			<i>rg</i>	<i>fas</i>
Glucose				<i>svp</i>

**Supplementary Table 6 *Drosophila* stocks.** The incidence of *Wolbachia*, determined experimentally by diagnostic PCR assay conducted on adult flies, is indicated (Y, present; n, absent).

A. DGRP lines

DGRP line	Wolbachia presence	DGRP line	Wolbachia presence	DGRP line	Wolbachia presence	DGRP line	Wolbachia presence
RAL_26	n	RAL_158	n	RAL_321	Y	RAL_385	n
RAL_28	n	RAL_161	n	RAL_332	n	RAL_386	n
RAL_38	n	RAL_176	Y	RAL_335	Y	RAL_392	n
RAL_41	n	RAL_181	Y	RAL_336	n	RAL_393	Y
RAL_42	n	RAL_195	n	RAL_338	Y	RAL_398	Y
RAL_45	n	RAL_208	n	RAL_340	Y	RAL_399	n
RAL_57	n	RAL_217	n	RAL_350	n	RAL_405	Y
RAL_59	n	RAL_223	Y	RAL_352	Y	RAL_409	Y
RAL_69	Y	RAL_227	Y	RAL_357	n	RAL_426	n
RAL_73	Y	RAL_228	n	RAL_358	n	RAL_427	n
RAL_80	Y	RAL_235	Y	RAL_360	Y	RAL_439	n
RAL_83	n	RAL_237	Y	RAL_361	Y	RAL_440	Y
RAL_85	n	RAL_239	n	RAL_362	Y	RAL_441	Y
RAL_88	n	RAL_272	Y	RAL_365	n	RAL_443	n
RAL_91	n	RAL_280	Y	RAL_367	n	RAL_461	Y
RAL_93	n	RAL_303	n	RAL_371	n	RAL_486	Y
RAL_105	n	RAL_304	Y	RAL_373	n	RAL_492	n
RAL_109	n	RAL_306	Y	RAL_374	Y	RAL_502	n
RAL_129	n	RAL_309	n	RAL_377	n	RAL_513	Y
RAL_136	Y	RAL_315	n	RAL_379	n	RAL_514	n
RAL_149	Y	RAL_317	Y	RAL_380	Y	RAL_531	Y
RAL_153	Y	RAL_318	Y	RAL_381	n	RAL_535	Y

RAL_554	Y	RAL_750	Y	RAL_821	Y	RAL_882	Y
RAL_563	n	RAL_765	n	RAL_822	Y	RAL_884	Y
RAL_584	Y	RAL_776	Y	RAL_837	Y	RAL_887	Y
RAL_642	n	RAL_783	Y	RAL_842	Y	RAL_890	Y
RAL_703	n	RAL_786	Y	RAL_843	n	RAL_897	Y
RAL_712	Y	RAL_787	Y	RAL_849	n	RAL_900	n
RAL_714	n	RAL_796	Y	RAL_850	Y	RAL_907	n
RAL_716	Y	RAL_801	Y	RAL_852	Y	RAL_908	n
RAL_721	Y	RAL_805	Y	RAL_855	Y	RAL_911	n
RAL_737	Y	RAL_808	n	RAL_857	Y	RAL_913	Y
RAL_738	Y	RAL_810	n	RAL_861	Y		
RAL_748	Y	RAL_819	Y	RAL_879	Y		

B. Mutant stocks

Symbol	Wolbachia	BDSC#	Genotype	Chr(s)	Breakpoints/Insertions
<i>mth1</i>	n	15318	<i>y</i> [1] <i>w</i> [67 <i>c</i> 23]; <i>r</i> { <i>w</i> [+ <i>m</i> C] <i>y</i> [+ <i>r</i> <i>m</i> Dint2]}=EPgy2} <i>mth1</i> [EY00861]	1	14F5, X:1.655397770..1655397770 (R5)
<i>ird5</i>	n	19825	<i>y</i> [1] <i>w</i> [67 <i>c</i> 23]; <i>P</i> { <i>w</i> [+ <i>m</i> C] <i>y</i> [+ <i>m</i> Dint2]}=EPgy2} <i>ird5</i> [EY02434]	1,3	89B1, 3R:1.1875112..1.1875112 (R5 flank)
HDAC4	Y	19942	<i>y</i> [1] <i>w</i> [67 <i>c</i> 23]; <i>P</i> { <i>w</i> [+ <i>m</i> C] <i>y</i> [+ <i>m</i> Dint2]}=EPgy2} <i>HDAC4</i> [EY08910]	1	11E8, X:1.3172806..1.3172806 (R5 flank)
Edem2	Y	21030	<i>y</i> [1] <i>w</i> [67 <i>c</i> 23]; <i>P</i> { <i>y</i> [+ <i>t</i> 7.7] <i>w</i> [+ <i>m</i> C]= <i>why</i> } <i>Edem2</i> [DG03809]	1,2	34A9, 2L:1.3193696..1.3194225 (R5 flank)
CG32373/ <i>Ank2</i>	n	22665	<i>y</i> [1] <i>w</i> [67 <i>c</i> 23]; <i>Mi</i> { <i>ET1</i> } <i>CG32373</i> [ <i>MB00005</i> ] <i>Ank2</i> [ <i>MB00005</i> ]	1,3	66A10, 3L:7705072..7705072 (R5 flank)
Dscam3	n	22728	<i>y</i> [1] <i>w</i> [67 <i>c</i> 23]; <i>Mi</i> { <i>ET1</i> } <i>Dscam3</i> [ <i>MB00214</i> ]	1,3	90B1, 3R:1.3319443..1.3319443 (R5 flank)
CG32264	n	22806	<i>y</i> [1] <i>w</i> [67 <i>c</i> 23]; <i>Mi</i> { <i>ET1</i> } <i>CG32264</i> [ <i>MB00461</i> ]	1,3	63F1, 3L:3.793458..3.793458 (R5 flank)
CG7341	n	22937	<i>y</i> [1] <i>w</i> [67 <i>c</i> 23]; <i>Mi</i> { <i>ET1</i> } <i>CG7341</i> [ <i>MB00927</i> ]	1,3	75B10, 3L:1.8113031..1.8113030 (R5 flank)
<i>rg</i>	Y	23371	<i>w</i> [1.118]; <i>Mi</i> { <i>ET1</i> } <i>rg</i> [ <i>MB01845</i> ]	1	4F2, X:5.101540 (R5 flank)
CG1688	Y	23791	<i>w</i> [1.118]; <i>Mi</i> { <i>ET1</i> } <i>CG1688</i> [ <i>MB02215</i> ]	1,2	46B3, 2R:5.688310 (R5 flank)
<i>hiw</i>	Y	23830	<i>w</i> [1.118]; <i>Mi</i> { <i>ET1</i> } <i>hiw</i> [ <i>MB02880</i> ] <i>CG5544</i> [ <i>MB02880</i> ]	1	13A5, X:1.4944473 (R5 flank)
<i>Fili</i>	Y	23968	<i>w</i> [1.118]; <i>Mi</i> { <i>ET1</i> } <i>Fili</i> [ <i>MB02223</i> ]	1,2	58A3, 2R:1.7810972 (R5 flank)
CG34113	Y	24681	<i>w</i> [1.118]; <i>Mi</i> { <i>ET1</i> } <i>CG34113</i> [ <i>MB04218</i> ]	1,3	83C8, 3R:1.755288 (R5 flank)
Ptp99A	n	24843	<i>w</i> [1.118]; <i>Mi</i> { <i>ET1</i> } <i>Ptp99A</i> [ <i>MB04947</i> ]	1,3	99A9, 3R:2.5242468 (R5 flank)
CG5565	Y	25492	<i>w</i> [1.118]; <i>Mi</i> { <i>ET1</i> } <i>CG5565</i> [ <i>MB06279</i> ]	1,2	22A1, 2L:1.355053 (R5 flank)
CG15533	Y	26036	<i>w</i> [1.118]; <i>Mi</i> { <i>ET1</i> } <i>CG15533</i> [ <i>MB05761</i> ]	1,3	99F4, 3R:2.6255260 (R5 flank)

CG18437	y	26099	w[1118]; M{ET1}CG18437[MB08151]	1;3	98A15, 3R:23494101 (R5 flank)
CG7461	n	26409	w[1118]; M{ET1}CG7461[MB08844]	1;2	56C9, 2R:15193421 (R5 flank)
CG30288	y	26457	w[1118]; M{ET1}CG30288[MB09262]	1;2	57E9, 2R:17407616 (R5 flank)
kek4	y	29204	w[1118]; M{ET1}kek4[MB11415]	1;2	34A10, 2L:13213226..13213226 (R5 flank)
CG10019	y	29248	w[1118]; M{ET1}CG10019[MB11741]	1;2	24C3, 2L:3751143..3751143 (R5 flank)
c(3)G	n	30110	w[1118]; P{w[+mC]=EP}c(3)G[G5001]	1;3	89A5, 3R:11618273 (R5 flank)
CG7705	y	31820	w[1118]; P{w[+mC]=EP}CG7705[G18394]	1;3	91C1, 3R:14503232 (R5 flank)

**Supplementary Table 7 Statistical models and sample sizes for genome-wide associations**

Trait	Model fitted at each SNP	Number of DGRP lines tested <sup>†</sup>
Dry weight	$Dry\ weight \sim SNP_i + microbiota + (SNP_i * microbiota) + (day / DGRR\ line)$	108
Glucose	$Glucose \sim Mean\ line\ dry\ weight + SNP_i + microbiota + (SNP_i * microbiota) + (day / DGRR\ line)$	107
Triglyceride	$TAG \sim Mean\ line\ dry\ weight + SNP_i + microbiota + (SNP_i * microbiota) + (day / DGRR\ line)$	107
Protein	$Protein \sim Mean\ line\ dry\ weight + SNP_i + microbiota + (SNP_i * microbiota) + (day / DGRR\ line)$	108
Glycogen	$Glycogen \sim Mean\ line\ dry\ weight + Wolbachia\ status + SNP + microbiota + (SNP * microbiota) + (day / DGRR\ line)$	74

<sup>†</sup> up to 3 replicates per line

## 1 **Supplementary Note 1**

2 This section addresses two potential caveats in our GWAs that are rarely considered in DGRP  
3 studies: inflation of p-values in quantile-quantile (QQ) plots that is indicative of potential mis-  
4 specification of the statistical model; and rationalized assignment of critical probabilities for  
5 identification of SNPs significantly associated with phenotypic traits.

6

7 **Inflation of p-values.** In this study, the QQ plots of p-values for the interaction of SNP and  
8 presence/absence of microbiota were inflated above the null expectation (Supplementary Fig. 5).  
9 From our QQ plots we inferred that the cause of inflation was likely to be universal to all traits  
10 and chromosomes, and therefore we tested ways to ameliorate QQ plot inflation focusing on the  
11 TAG phenotypic data, using two complementary approaches.

12 In the first approach, we conducted a stratified analysis, in which the effect of SNP was  
13 modeled in separate GWAs for axenic and gnotobiotic flies. QQ plots of p-values for SNPs  
14 conformed well to a null distribution (Supplementary Fig. 6a). Certain SNPs were positively  
15 correlated in this analysis whilst others were much more significant in one microbiota condition  
16 than the other (Supplementary Fig. 5b), indicating that this analysis was able to distinguish  
17 putative genotypic effects on the traits under study, and for SNPs with effects specific to the  
18 presence/absence of microbiota. Although this stratified approach is an unconventional means to  
19 test genotype-environment interactions, we sought to test whether it provided independent  
20 corroboration of our full GWA analysis. We reasoned that a SNP with a strong microbiota-  
21 interactive effect would have a low p-value in one microbiota condition and a high p-value in the  
22 other. We calculated the difference in p-values for each SNP in each treatment, and multiplied  
23 negative differences by -1 on the basis that the magnitude of difference matters for determining

24 effects, but not direction of difference. We compared the full and stratified mapping approaches  
25 by correlating p-values of each SNP from the full GWA to the difference in p-value of each SNP  
26 in the two parts of the stratified analysis. These statistics were somewhat correlated, providing  
27 some independent validation of our mapping approach (Supplementary Fig.5c). Of the 11 SNPs  
28 with a GWA p-value ( $-\log_{10}$ )  $\geq 10$ , five showed a difference of  $\geq 3.75$  in p-values ( $-\log_{10}$ ) in the  
29 stratified analysis.

30 As a second approach to reduce p-value inflation, alternative models were tested on the TAG  
31 data and SNPs in the X chromosome. We randomized our data either within or across microbiota  
32 treatments, i.e. generating data that were randomly distributed with respect to genotype or both  
33 genotype and treatment. Inflation was removed by permutation between but not within the two  
34 treatments, confirming that inflation is caused by an unidentified aspect of modeling the  
35 phenotypic data across these treatments. We interpret this result to indicate that additional  
36 microbiota-dependent covariates, the identity of which remains to be established, are required to  
37 fully explain variance in the *Drosophila* traits under study. We tested for effects of variation in  
38 CFU counts associated with gnotobiotic flies, which would not be accounted for by statistically  
39 modeling microbiota as a binary factor in GWAs. We accounted for variation CFU of  
40 *Acetobacter*, *Lactobacilli*, and total CFU in separate analyses, by replacing the binary microbiota  
41 term in our GWAs with mean line CFU, such that gnotobiotic samples were modeled with mean  
42 line CFU ( $\log_{10}N$ ) as a covariate for each bacterial taxon and total CFU (in three separate  
43 analyses), and axenic samples were given a CFU count of zero. GWAs of the interaction of SNP  
44 and these continuous variables remained inflated. We also considered the possibility that  
45 removal of microbiota may have had systematic effects on the water content of flies used in this  
46 study, based on the observation that the microbiota affects dry weight of *Drosophila* from the



47 DGRP (this study) but not the wet weight of wild-type male *Drosophila* Canton S reared on the  
48 same diet used in this study<sup>1,2</sup>. In principle, such effects could cause additional phenotypic  
49 variance when dry weight is used as a covariate. However, QQ plot inflation was not removed by  
50 normalizing dry weights of axenic flies to the dry weights of gnotobiotic flies [this was achieved  
51 by adding the difference in grand means (all lines) of axenic and gnotobiotic dry weight to the  
52 dry weight of axenic flies of each DGRP line]. We also tested whether inflation in the QQ plots  
53 of genome-wide SNP-microbiota interactions could be driven by inflation in the genotype QQ  
54 plots, caused by cryptic population structure in the DGRP. We compressed the genotype matrix  
55 by principal components analysis (PCA) and included the first three principal components as  
56 covariates in the GWA. This approach did not remove inflation from the QQ plots. Other  
57 strategies that did not bring p-values to a null distribution included running GWA of mean trait  
58 values per line without line as a random effect, re-running our GWA in SAS rather than R, and  
59 removal of experimental block as a random effect. These analyses suggest that p-value  
60 distributions, which are not routinely reported in DGRP GWA studies, may commonly deviate  
61 from a null expectation due to an intrinsic property of the lines or the models used to analyse  
62 them, and therefore that inclusion of QQ plots should be a standard for the field.

63

64 **Assignment of critical probabilities.** The critical probability for identifying significant genetic  
65 associations with phenotypic traits is poorly defined because standard corrections for multiple  
66 comparisons are widely recognized as ineligible for GWA studies in the DGRP<sup>3,4</sup>. The widely-  
67 adopted convention is to assign a single arbitrary probability cutoff to determine the statistical  
68 significance of SNPs<sup>3,4</sup>. This was not followed here because our multiple phenotypes have  
69 different distributions of p-values (Supplementary Fig. 1) and a single arbitrary cutoff would,

70 consequently, bias the number of SNPs reported as significantly associated with each trait.  
71 Instead, we focused on the top ranking 0.01% of SNPs (Supplementary Fig. 1), yielding 184-226  
72 SNPs, including SNPs associated with 77-144 genes for each trait, and a total of 436 genes  
73 mapped across the full dataset (Supplementary Table 3). These numbers are similar to those  
74 presented in other recent *Drosophila* GWAs<sup>4,5</sup>.

75

## 76 **References**

- 77 1. Wong, A.C., Dobson, A.J. & Douglas, A.E. Gut microbiota dictates the metabolic response of  
78 *Drosophila* to diet. *J. Exp. Biol.* **217**, 1894-1901 (2014).
- 79 2. Newell, P.D. & Douglas, A.E. Interspecies interactions determine the impact of the gut  
80 microbiota on nutrient allocation in *Drosophila melanogaster*. *Appl. Environ. Microbiol.* **80**,  
81 788-796 (2014).
- 82 3. Jordan, K.W., Craver, K.L., Magwire, M.M., Cubilla, C.E., Mackay, T.F. & Anholt, R.R.  
83 Genome-wide association for sensitivity to chronic oxidative stress in *Drosophila*  
84 *melanogaster*. *PLoS One* **7**, e38722 (2012).
- 85 4. Brown, E.B., Layne, J.E., Zhu, C., Jegga, A.G. & Rollmann, S.M. Genome-wide association  
86 mapping of natural variation in odour-guided behaviour in *Drosophila*. *Genes, Brain Behav.*  
87 **12**, 503-515 (2013).
- 88 5. Burke, M.K., King, E.G., Shahrestani, P., Rose, M.R. & Long, A.D. Genome-wide association  
89 study of extreme longevity in *Drosophila melanogaster*. *Genome Biol. Evol.* **6**, 1-11 (2014).

90

91