Stable host gene expression in the gut of adult *Drosophila melanogaster* with different bacterial mono-associations

Carolyn Elya[1], Vivian Zhang[1], Will Ludington[1] & Michael B. Eisen[1,2,3]

[1]Department of Molecular and Cell Biology, [2]Department of Integrative Biology and [3]Howard Hughes Medical Institute, University of California, Berkeley, CA.

Abstract

There is growing evidence that the microbes found in the digestive tracts of animals influence host biology, but we still do not understand how this comes about. Here, we evaluated how different microbial species commonly associated with laboratory-reared *Drosophila melanogaster* impact host biology at the level of gene expression in the dissected adult gut or the entire adult organism. We observed that guts from gnotobiotic animals associated from the embryonic stage with either zero, one or three bacterial species demonstrated indistinguishable transcriptional profiles. Additionally, we found that the gut transcriptional profiles of animals reared in the presence of the yeast *Saccharomyces cerevisiae* alone or in combination with bacteria could recapitulate those of conventionally-reared animals. In contrast, we found whole body transcriptional profiles of conventionally-reared animals were distinct from all of the gnotobiotic treatments tested. Our data suggest that adult flies are insensitive to the ingestion of different bacterial species but that prior to adulthood, different microbes impact the host in ways that lead to global transcriptional differences observable across the whole adult body.

Introduction

The digestive tracts of virtually all animals examined to date are inhabited by microbes and there is increasing evidence that interactions between gut microbes and their animal hosts influence a wide range of host phenotypes [1]. However the molecular basis for these effects is poorly understood.

Drosophila melanogaster has become a model for the study of host-microbe interactions in the gut owing to the relative simplicity of the tissue and its microbial community. The *Drosophila* gut microbe

community is comprised of about 50 bacterial species in wild flies [2–5] and about ten species in lab-reared flies [2,3,6–8]. In lab-reared animals, the majority of these bacteria are lactic acid and acetic acid producers (members of family Lactobacillaceae and Acetobacteraceae, respectively) that can be cultured outside of the host. *Drosophila* embryos can be stripped of their endogenous microflora using household bleach [9] and either maintained under sterile conditions to generate axenic (germ-free) animals or treated with a defined set of microbes to generate gnotobiotic (of known microbial content) animals [10].

A variety of phenotypes have been associated with the presence and/or specific composition of microbes in the *Drosophila melanogaster* gut including nutrition and metabolism [8,10,11], intestinal cell growth [12–14], development [15,16], lifespan [7,17–19] and a variety of different behaviors, including social attraction in larvae and assortative mating in adults of two lab-reared strains [20–22].

The presence of microbes in the environment and gut has a significant effect on physiology, morphology and gene expression in the adult *D. melanogaster* gut [11,19,23], however, the effect of individual microbe species on the gene expression of adult guts has not been examined. Given the known differential phenotypic effects of different bacterial taxa, we hypothesized that the epithelial cells of the digestive tract might respond to different bacteria in different ways, and that these responses might provide clues to the molecular mechanisms that underlie the effects of specific microbes of *Drosophila* physiology and behavior. Here we describe the results of a series of experiments using mRNA sequencing to analyze gene expression in the guts of a laboratory line of *D. melanogaster* reared from embryogenesis to early adulthood with different combinations of bacteria and yeast commonly found in lab populations.

Results

Limited variation in transcription in adult *D. melanogaster* gut in response to mono-association with different bacterial species

To elucidate the effects of specific bacterial species on host gut gene expression we implemented previously published protocols [10] to mono-associate laboratory stocks of *D. melanogaster* with three bacterial species previously shown to associate with *D. melanogaster* in the lab: *Acetobacter*

pasteurianus, *Lactobacillus brevis* and *Lactobacillus plantarum*. *A. pasteurianus* and *L. brevis* were selected based on their high abundance in a survey of wild-type female (CantonS) flies reared on standard media in our lab (Figure S1) as well as their ubiquity in other gut microbe surveys of *D. melanogaster* [2,3,6–8]. *L. plantarum* was chosen because it has been observed in several other gut microbe surveys and has been implicated in mediating specific aspects of fly development and behavior [16,22].

Briefly, sterile embryos of *D. melanogaster* (CantonS) were placed in vials inoculated with 5x106 CFUs of the relevant bacteria and allowed to develop under standard conditions. At five days after eclosion, gut samples from single female adults were prepared for mRNA sequencing and sequenced using standard methods.

After aligning sequencing reads to the *D. melanogaster* genome and estimating transcript abundance (FPKM) for each transcript, we visually examined variability in gene expression within and between experiments and treatments across samples organized by either experimental date (Figure 1A) or bacterial mono-association (Figure 1B). Contrary to our expectation, we did not observe any significant differences in gene expression associated with the different bacterial treatments.

To confirm this visual impression, we used a simple statistical test (ANOVA) to identify individual genes differentially expressed between treatments. After applying a Bonferroni correction for multiple testing, the expression levels of only two genes were significantly associated with treatment: CAH2, a carbonic anhydrase, and CG17574, a gene of unknown function (Figure 2). Even relaxing our significance threshold, and considering the fifty genes that demonstrate the most significant differences in expression within this set, it is clear that there are minimal differences in gene expression between the three different bacterial mono-associations (Figure 2). Instead, it appears that each bacterial species alone affected host gene expression in the same way.
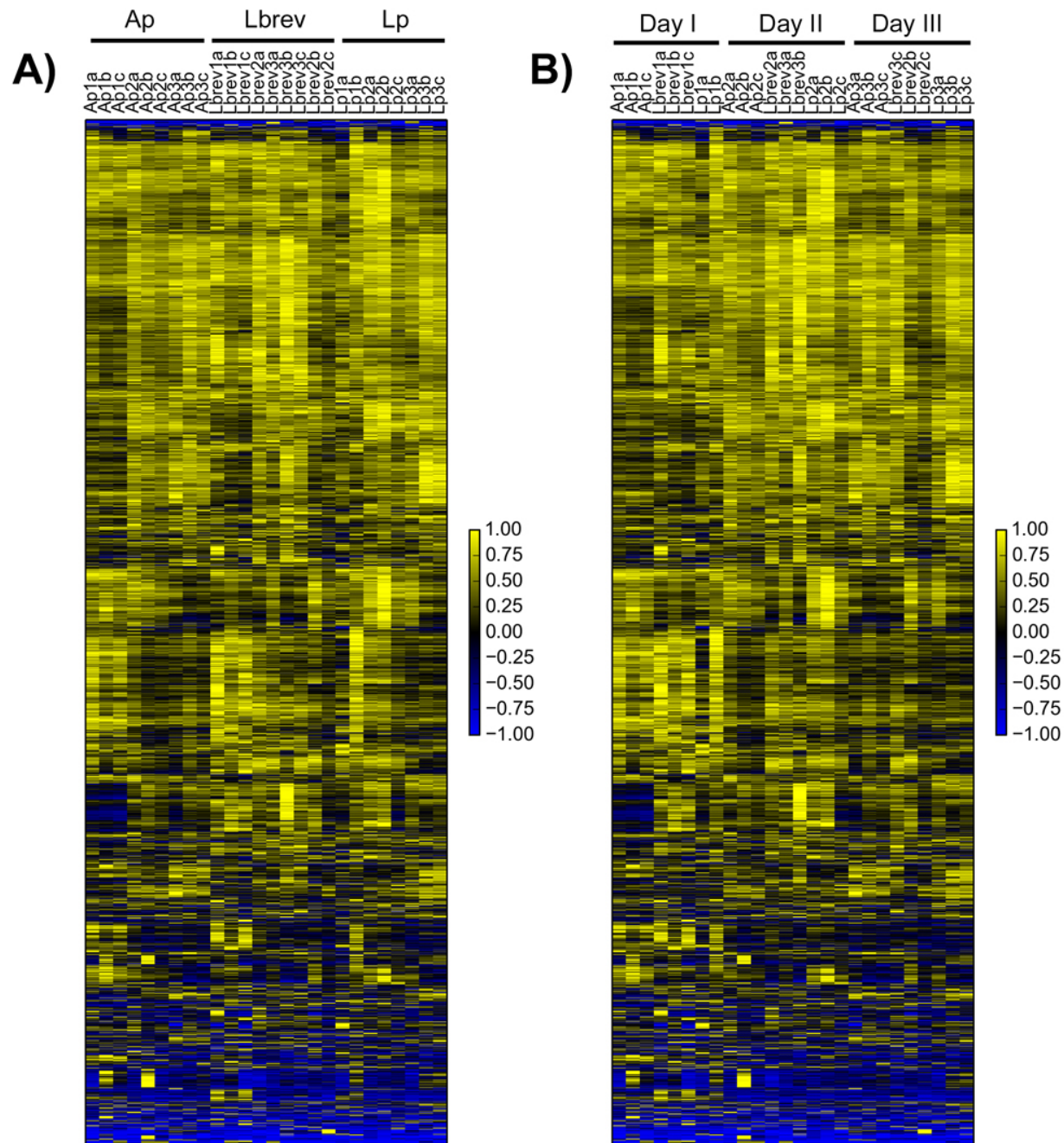
Figure 1) Limited variation in gut gene expression with bacterial mono-association. Expression data from guts dissected from five-day post-eclosion, *Wolbachia*-free, mated female CantonS *D. melanogaster* individuals mono-associated with one of three bacteria (Ap = *A. pastuerianus*, Lbrev = *L. brevis*, Lp = *L. plantarum*) were clustered by gene (average linkage, uncentered correlation) after first filtering out genes that lacked three instances of FPKM greater than two (Gene Cluster 3.0). FPKM values for each gene were normalized to range from -1 to 1 before plotting. A) Samples arranged by bacterial treatment. B) Samples arranged by date of experiment. Scale bars for each heatmap are given to the right of the plot.
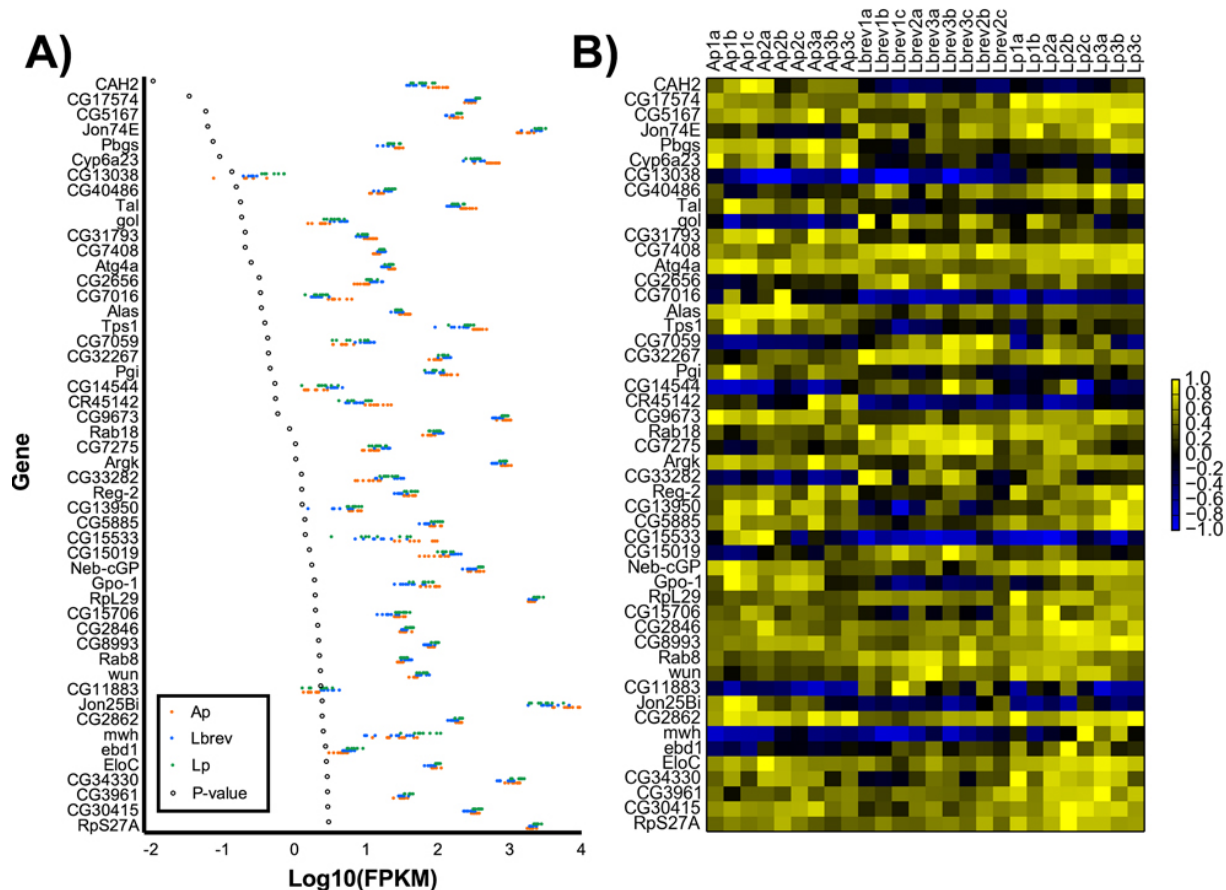
Figure 2) Genes showing greatest difference in expression values from dissected adult guts as determined by one-way ANOVA. A) Scatterplot of log10-transformed FPKM values for each bacteria mono-associated gut replicate (Ap = *A. pastuerianus*, Lbrev = *L. brevis*, Lp = *L. plantarum*). Genes are ordered from lowest ANOVA p-value (top) to highest (bottom). P-values have undergone a Bonferroni correction for multiple testing. B) Data from A presented as a heatmap. FPKM values for each gene are linearly normalized to range from -1 to 1 before plotting.

Host transcription in the gut is markedly different between conventional and yeast mono-association compared to axenic and bacteria mono-association treatments

After observing a lack of differences in our initial set of samples, we decided to expand our dataset to include other treatments that might shed light on this unexpected result. To this end, we measured gene expression in guts from conventionally reared flies, axenic flies and flies mono-associated with the yeast *Saccharomyces cerevisiae* (which is known to be associated with flies in the wild [24] and can alone provide complete nutrition for the developing fly [25]) rather than bacteria. Additionally, we tested if providing a simplified multi-species microbial community (either all three of *A. pasteurianus*, *L.*

Elya et al, 2016   5

brevis and *L. plantarum* with or without *S. cerevisiae*) would yield a transcriptional program more similar to that of conventional than to mono-associated samples.

Based on studies demonstrating that microbes play critical roles in animal development, we anticipated that the expression pattern of guts from axenic flies would be distinct from those of mono-associated, poly-associated (bearing a simplified microbial community) or conventional animals [11,23]. We also expected expression in mono-associated and poly-associated flies to either closely resemble that of conventionally-reared flies or lie somewhere in between the axenic and conventional samples. Finally, we did not expect that providing yeast, either alone or with all three of our selected bacteria, would have a significant impact on gut gene expression, as previous studies suggest that most vegetative *S. cerevisiae* ingested by *D. melanogaster* do not survive passage through the digestive tract [26,27]. If *S. cerevisiae* is only metabolized and does not actively survive the GI tract, we predicted that providing live *S. cerevisiae* would have no significant impact on the flies, since they already were provided dead *S. cerevisiae* in their diet.

To test these predictions, we dissected and sequenced mRNA from wild-type, 5-day-old, female CantonS flies that were either axenic, conventional, mono-associated with *S. cerevisiae*, poly-associated with an equal amounts of *A. pasteurianus*, *L. brevis* and *L. plantarum* or poly-associated with an equal amounts of *A. pasteurianus*, *L. brevis*, *L. plantarum* with or without *S. cerevisiae*.

As above, we first examined the aggregate gene expression data visually (Figure 3A). In contrast to our expectations, the expression data clearly showed that guts from gnotobiotic animals whose treatment included yeast gave transcription patterns most similar to those of guts from conventional animals. We were further surprised to observe that rather than being distinct from all other samples as we had expected, transcriptional patterns of guts from axenic animals most closely resembled that of animals either mono- or poly-associated with bacteria. Taken together, samples appeared to fall within one of two transcriptional regimes: "conventional-like" (conventional and yeast-mono and poly-associations) and "bacteria-like" (axenic, bacteria mono- and poly-associations).

To understand which genes most distinguished these two transcriptional groups, we again performed a gene-by-gene ANOVA. This analysis revealed more than 2,000 genes with significant expression differences between the two groups (p-value under 0.05 after Bonferroni correction). Genes that were

more highly expressed in "conventional-like" samples were enriched for annotations for a variety of metabolic processes, most notably involving lipids and amino acids (Figure 3B). Genes more highly expressed in "bacterial-like" samples were also enriched in metabolism annotations, although this enrichment was less pronounced and more general than for the former set (Figure 3C).
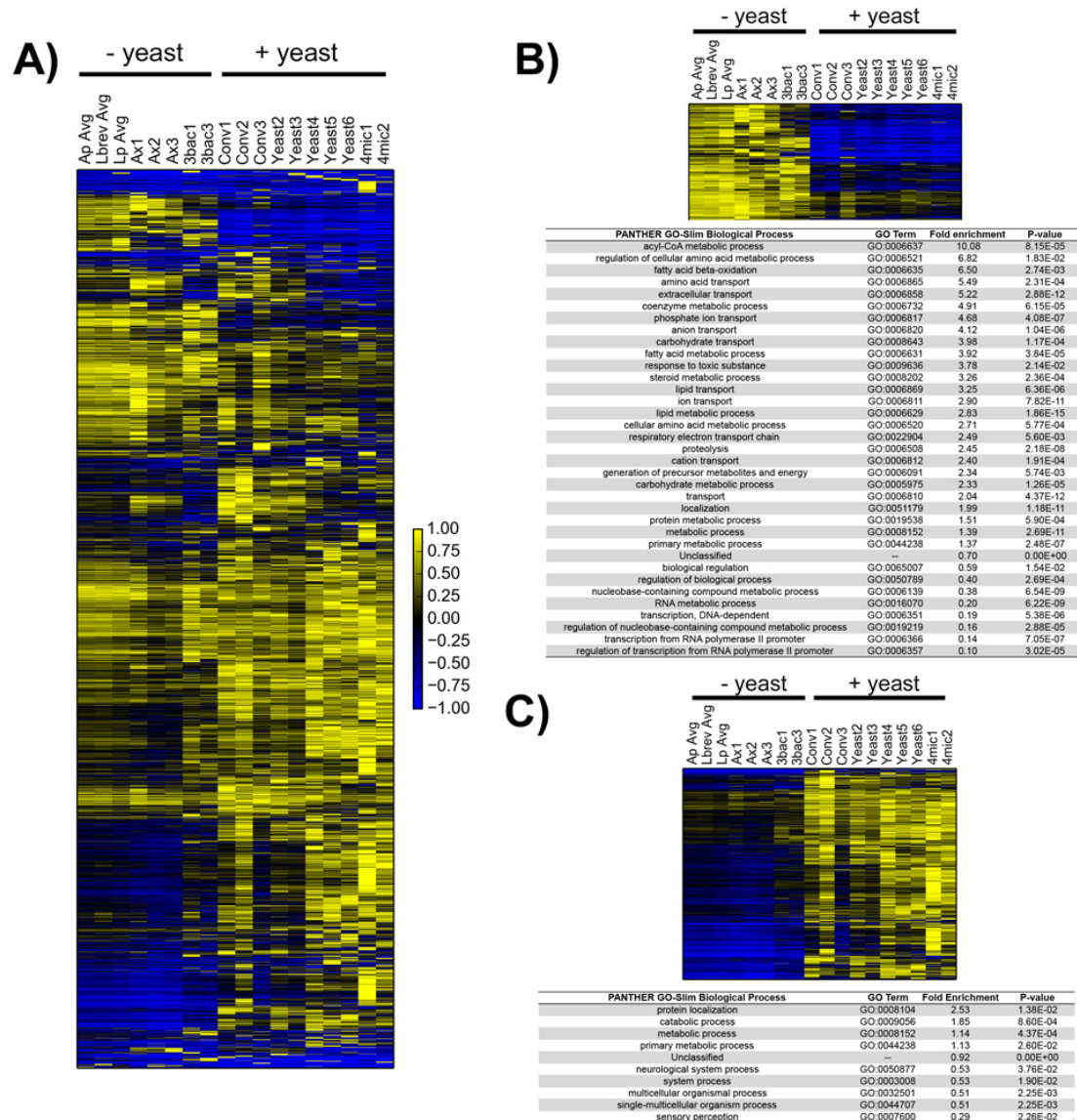
Overall, these gut expression data are consistent with a model in which yeast shapes the transcriptional program of the gut in mated, young adult females. Specifically, these results argue that yeast induces metabolic changes that are not evoked by association with any one species of bacteria associated with lab-reared *D. melanogaster* or a simplified community comprised of the all three bacterial species considered herein.

Global host gene expression in conventional animals is distinct among treatments

As our expectation that host transcription within the gut (i.e. proximal transcription) would change in response to mono-association with different bacterial taxa did not hold true, we were curious if the same would hold true for whole animals. We therefore extracted and sequenced mRNA from individual whole, 5-day post-eclosion, mated female CantonS flies that were prepared with the same microbial treatments as our gut samples (i.e. mono-associated with either *A. pasteurianus*, *L. brevis*, *L. plantarum* or *S. cerevisiae*, poly-associated with *A. pasteurianus*, *L. brevis* and *L. plantarum*, poly-associated with *A. pasteurianus*, *L. brevis*, *L. plantarum* and *S. cerevisiae*, axenic or conventional).

From our observation that gene expression in the gut tracked with exposure to yeast, we hypothesized that global transcription would also be dependent on the presence of *S. cerevisiae*. Based on other gene-expression studies examining axenic and conventional animals, we anticipated that the majority of the affected genes still would be those expressed in the gut [11,23].

Data from whole animals (Figure 4A) show that conventionally-reared flies have a transcription profile distinct from all of the samples tested. ANOVA comparing conventional samples to all other samples showed that ~1700 genes display different patterns of expression between these two groups after applying a strict Bonferroni correction for multiple testing. Many biological processes are enriched among the genes expressed at higher levels in the conventionally reared samples (Figure 4B), while

Figure 3) Table B) PANTHER GO-Slim Biological Process:

PANTHER GO-Slim Biological Process	GO Term	Fold enrichment	P-value
acyl-CoA metabolic process	GO:0006637	10.08	8.15E-05
regulation of cellular amino acid metabolic process	GO:0006521	6.82	1.83E-02
fatty acid beta-oxidation	GO:0006635	6.50	2.74E-03
amino acid transport	GO:0006865	5.49	2.31E-04
extracellular transport	GO:0006858	5.22	2.88E-12
coenzyme metabolic process	GO:0006732	4.91	6.15E-05
phosphate ion transport	GO:0006817	4.68	4.08E-07
anion transport	GO:0006820	4.12	1.04E-06
carbohydrate transport	GO:0008643	3.98	1.17E-04
fatty acid metabolic process	GO:0006631	3.92	3.84E-05
response to toxic substance	GO:0009636	3.78	2.14E-02
steroid metabolic process	GO:0008202	3.26	2.36E-04
lipid transport	GO:0006869	3.25	6.36E-06
ion transport	GO:0006811	2.90	7.82E-11
lipid metabolic process	GO:0006629	2.83	1.86E-15
cellular amino acid metabolic process	GO:0006520	2.71	5.77E-04
respiratory electron transport chain	GO:0022904	2.49	5.60E-03
proteolysis	GO:0006508	2.45	2.18E-08
cation transport	GO:0006812	2.40	1.91E-04
generation of precursor metabolites and energy	GO:0006091	2.34	5.74E-03
carbohydrate metabolic process	GO:0005975	2.33	1.26E-05
transport	GO:0006810	2.04	4.37E-12
localization	GO:0051179	1.99	1.18E-11
protein metabolic process	GO:0019538	1.51	5.90E-04
metabolic process	GO:0008152	1.39	2.69E-11
primary metabolic process	GO:0044238	1.37	2.48E-07
Unclassified	--	0.70	0.00E+00
biological regulation	GO:0065007	0.59	1.54E-02
regulation of biological process	GO:0050789	0.40	2.69E-04
nucleobase-containing compound metabolic process	GO:0006139	0.38	6.54E-09
RNA metabolic process	GO:0016070	0.20	6.22E-09
transcription, DNA-dependent	GO:0006351	0.19	5.38E-06
regulation of nucleobase-containing compound metabolic process	GO:0019219	0.16	2.88E-05
transcription from RNA polymerase II promoter	GO:0006366	0.14	7.05E-07
regulation of transcription from RNA polymerase II promoter	GO:0006357	0.10	3.02E-05

Table C) PANTHER GO-Slim Biological Process:

PANTHER GO-Slim Biological Process	GO Term	Fold Enrichment	P-value
protein localization	GO:0008104	2.53	1.38E-02
catabolic process	GO:0009056	1.85	8.60E-04
metabolic process	GO:0008152	1.14	4.37E-04
primary metabolic process	GO:0044238	1.13	2.60E-02
Unclassified	--	0.92	0.00E+00
neurological system process	GO:0050877	0.53	3.76E-02
system process	GO:0003008	0.53	1.90E-02
multicellular organismal process	GO:0032501	0.51	2.25E-03
single-multicellular organism process	GO:0044707	0.51	2.25E-03
sensory perception	GO:0007600	0.29	2.26E-02

Figure 3) Yeast drive genome-wide difference in gut gene expression. A) Average linkage hierarchical clustering was performed in Gene Cluster 3.0 across all genes that are expressed at least at two FPKM in at least two out of 11 samples. Bacteria mono-association data has been averaged across each treatment to collapse down into a single column. FPKM values for each gene are normalized to range from -1 to 1 before plotting. Abbreviations: Ap avg = average for *A. pasteurianus*-mono-associated samples; Lbrev avg = average for *L. brevis*-mono-associated samples, Lp avg = average for *L plantarum*-mono-associated samples, 3bac = poly-associated (without yeast), Ax = axenic, Conv = conventional, Yeast = *S. cerevisiae*-mono-associated, 4mic = poly-associated (with yeast). Scale bar is shown at bottom right. B) Top) heatmap of 579 genes that are overexpressed in axenic, bacteria-mono-associated and poly-associated (without yeast) guts compared to other gut samples (Bonferroni p-value>0.05, ANOVA). Bottom) Results from Panther GO-SLIM biological function enrichment test [28] for gene set above compared to reference set of all genes observed across all gut datasets (556 were identified by Panther and used for analysis out of 579) C) Top) Heatmap of 1728 genes that are overexpressed in conventional, yeast mono-associated and poly-associated (with yeast) compared to other gut samples (Bonferroni p-value>0.05, ANOVA). Results from Panther GO-Slim biological processes enrichment test with 1728 (1663 identified) genes compared to reference set of all genes observed across all gut datasets. Note for B) and C): all individual sample values were used for ANOVA analysis, not the average value as plotted in A).

genes expressed at lower levels show a marked enrichment for processes involving protein folding and biogenesis (Figure 4C).
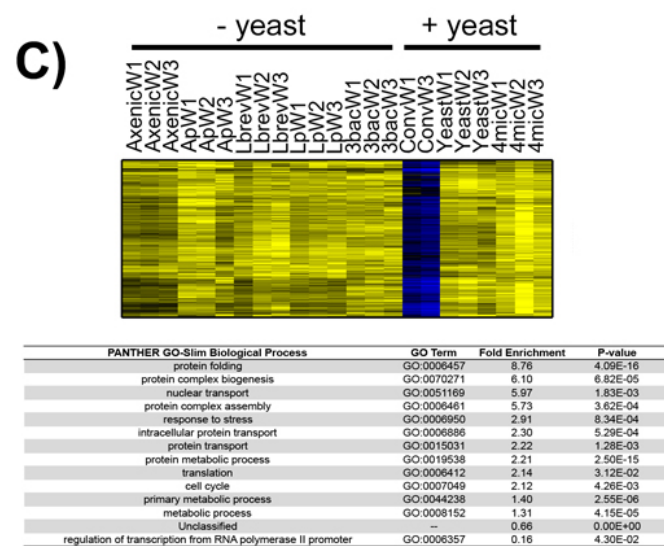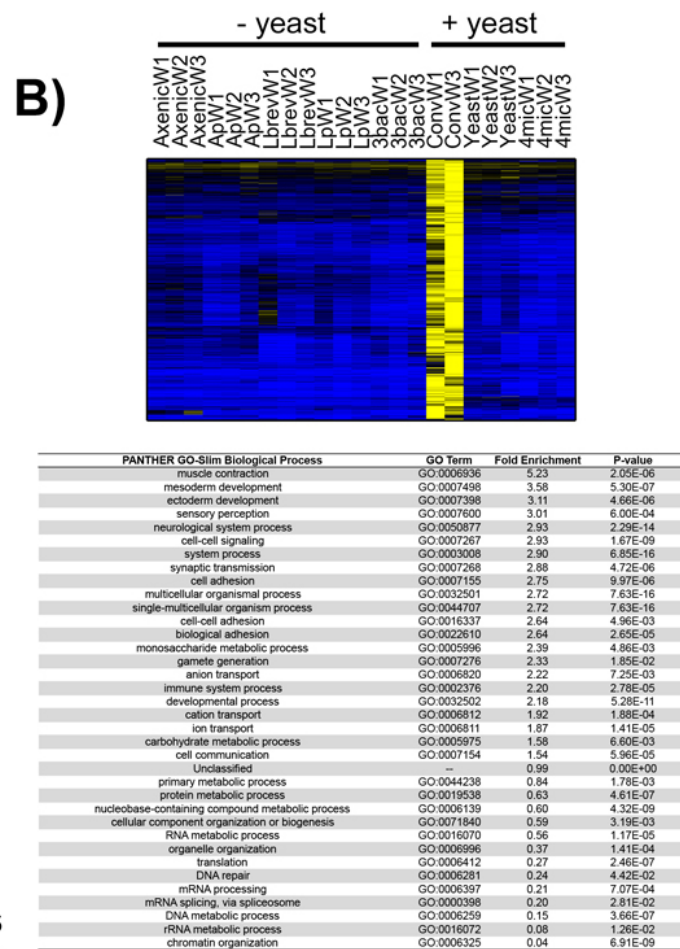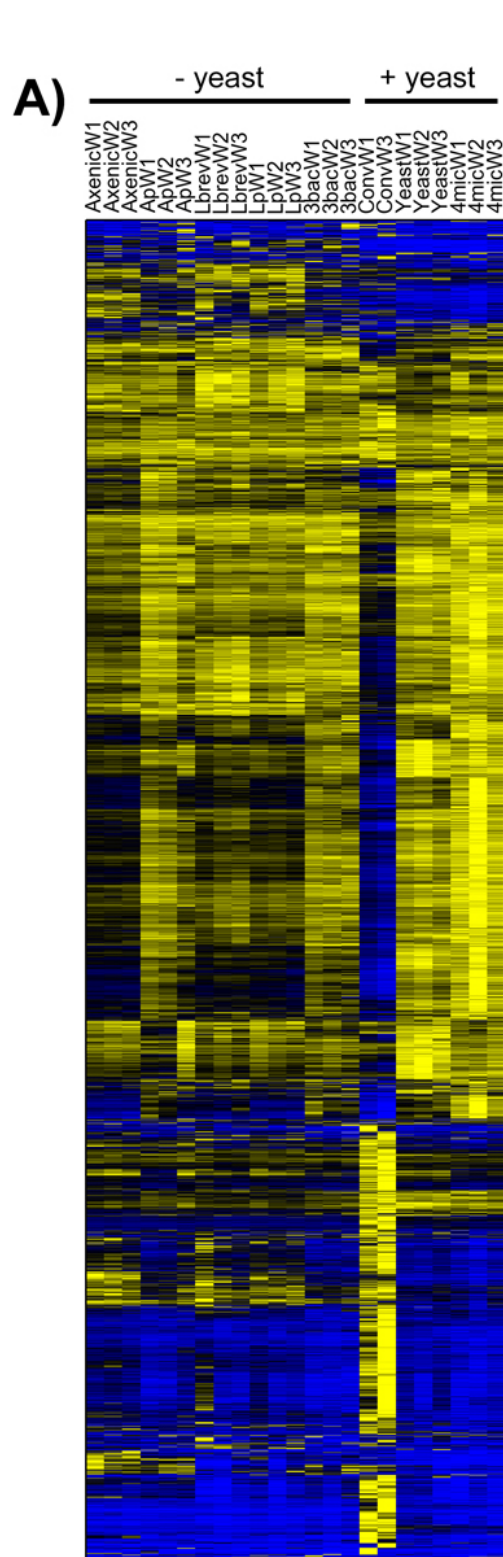
Though the global pattern clearly showed that conventional samples exhibit a unique transcription pattern, we reasoned that there could still be a subset of genes that maintain a high similarity between yeast-containing and conventional samples. Gene-by-gene analysis demonstrated that there was only a small overlap in transcription between conventional and yeast-containing samples (~170 genes showed similar patterns of expression between samples, Figure S2) and so better supports a model in which none of the gnotobiotic treatments tested in this study could recapitulate conventional levels of gene expression at the level of the whole animal.

To determine if any differences manifested between different bacterial mono-associated animals at the level of the entire organism that were not apparent in dissected guts, we completed a gene-by-gene analysis as per Figure 2. As in our guts, there are few significant differences between whole flies that are mono-associated with different bacteria once a Bonferroni correction is applied (Figure S3), though the data do demonstrate more variance than what was observed in dissected guts alone.

Overall, the whole animal data suggest that, while yeast alone may suffice to recapitulate conventional host gene expression in the gut, yeast alone or in a simplified mock community are not enough to generate an animal that demonstrates an overall conventional-like transcriptional program.

Figure 4) Analysis of gene expression trends from gnotobiotic whole flies (next page). A) Transcriptome-wide heatmap from axenic, conventional, yeast-mono-associated, bacteria-mono-associated and poly-associated whole flies clustered by gene expression Average linkage hierarchical clustering using an uncentered correlation similarity metric was performed in Gene Cluster 3.0 across all genes that are expressed at least at two FPKM across two out of eleven samples. Abbreviations: Ap = A. *pasteurianus*-mono-associated; Lbrev = *L. brevis*-mono-associated, Lp = *L plantarum*-mono-associated, 3bac = poly-associated without yeast, Ax = axenic, Conv = conventional, Yeast = *S. cerevisiae*-mono-associated, 4mic = poly-associated with yeast. Scale bar is shown at bottom right. B) Top) heatmap of 1159 of 1385 genes that are overexpressed in conventional whole flies compared to other whole fly samples (Bonferroni p-value>0.05, ANOVA). Genes absent in heatmap did not pass filtering criteria. Bottom) Results from Panther GO-SLIM biological function enrichment test [28] for gene set above (1278 genes were identified of 1385) compared to reference set of all genes observed across all whole fly datasets. C) Top) Heatmap 351 that are overexpressed in all non-conventional whole-fly samples compared to conventional whole flies (Bonferroni p-value>0.05, ANOVA). Results from Panther GO-Slim biological processes enrichment test with gene set above (348 genes were identified out of 351) compared to reference set of all genes observed across all whole fly datasets.

PANTHER GO-Slim Biological Process	GO Term	Fold Enrichment	P-value
muscle contraction	GO:0006936	5.23	2.05E-06
mesoderm development	GO:0007498	3.58	5.30E-07
ectoderm development	GO:0007398	3.11	4.66E-06
sensory perception	GO:0007600	3.01	6.00E-04
neurological system process	GO:0050877	2.93	2.29E-14
cell-cell signaling	GO:0007267	2.93	1.67E-09
system process	GO:0003008	2.90	6.85E-16
synaptic transmission	GO:0007268	2.88	4.72E-06
cell adhesion	GO:0007155	2.75	9.97E-06
multicellular organismal process	GO:0032501	2.72	7.63E-16
single-multicellular organism process	GO:0044707	2.72	7.63E-16
cell-cell adhesion	GO:0016337	2.64	4.96E-03
biological adhesion	GO:0022610	2.64	2.65E-05
monosaccharide metabolic process	GO:0005996	2.39	4.86E-03
gamete generation	GO:0007276	2.33	1.85E-02
anion transport	GO:0006820	2.22	7.25E-03
immune system process	GO:0002376	2.20	2.78E-05
developmental process	GO:0032502	2.18	5.28E-11
cation transport	GO:0006812	1.92	1.88E-04
ion transport	GO:0006811	1.87	1.41E-05
carbohydrate metabolic process	GO:0005975	1.58	6.60E-03
cell communication	GO:0007154	1.54	5.96E-05
Unclassified	--	0.99	0.00E+00
primary metabolic process	GO:0044238	0.84	1.78E-03
protein metabolic process	GO:0019538	0.63	4.61E-07
nucleobase-containing compound metabolic process	GO:0006139	0.60	4.32E-09
cellular component organization or biogenesis	GO:0071840	0.59	3.19E-03
RNA metabolic process	GO:0016070	0.56	1.17E-05
organelle organization	GO:0006996	0.37	1.41E-04
translation	GO:0006412	0.27	2.46E-07
DNA repair	GO:0006281	0.24	4.42E-02
mRNA processing	GO:0006397	0.21	7.07E-04
mRNA splicing, via spliceosome	GO:0000398	0.20	2.81E-02
DNA metabolic process	GO:0006259	0.15	3.66E-07
rRNA metabolic process	GO:0016072	0.08	1.26E-02
chromatin organization	GO:0006325	0.04	6.91E-09

PANTHER GO-Slim Biological Process	GO Term	Fold Enrichment	P-value
protein folding	GO:0006457	8.76	4.09E-16
protein complex biogenesis	GO:0070271	6.10	6.82E-05
nuclear transport	GO:0051169	5.97	1.83E-03
protein complex assembly	GO:0006461	5.73	3.62E-04
response to stress	GO:0006950	2.91	8.34E-04
intracellular protein transport	GO:0006886	2.30	5.29E-04
protein transport	GO:0015031	2.22	1.28E-03
protein metabolic process	GO:0019538	2.21	2.50E-15
translation	GO:0006412	2.14	3.12E-02
cell cycle	GO:0007049	2.12	4.26E-03
primary metabolic process	GO:0044238	1.40	2.55E-06
metabolic process	GO:0008152	1.31	4.15E-05
Unclassified	--	0.66	0.00E+00
regulation of transcription from RNA polymerase II promoter	GO:0006357	0.16	4.30E-02

Microbial load is highly variable between lab-reared flies raised under identical gnotobiotic conditions

Prior to conducting the experiments described above, we had been concerned about inter-animal variation in microbial load producing highly variable gene expression patterns. However, the relative lack of variation in gene expression within treatments led us to wonder if our gnotobiotic protocol generated flies with less variability in microbial load than we had anticipated.

We attempted to determine the total bacterial load of the animals we sequenced by counting 16S rRNAs in our sequencing reads. Although our sequencing library preparation protocol involves a polyA selection step to enrich for eukaryotic mRNAs, we nonetheless sequence many *Drosophila* rRNAs, which are not polyadenylated. We therefore expected to have sequenced some microbial rRNAs. However, we did not detect any and, in retrospect, believe this is due to a failure to lyse bacterial cells. Instead, to determine the likely range of microbial loads that our sequenced animals possessed we repeated the preparation of mono-associated animals five times, and determined the microbial loads of multiple animals per treatment per replicate for each of these preparations. Briefly, animals were surface-sterilized, individually homogenized and plated in two dilutions on appropriate media to determine the number of colony forming units (CFUs) they contained (Figure 5A).

Consistent with previous work [23], we observed large variability in the microbial load of gnotobiotic flies, despite being reared under nominally identical conditions within a given treatment. The average load among mono-associated animals varies with bacterial species, with *L. plantarum* colonizing more densely than *L. brevis*, which in turn colonizes more densely than *A. pastuerianus*. Conventional animals are on average about 10 - 100x more densely colonized than mono-associated animals. Notably, animals mono-associated with *S. cerevisiae* on average contain only about ten viable yeast cells.

Given these observations we think it is almost certain that the animals we sequenced were effectively colonized with the relevant bacterial species, but that the amount of bacteria they contained was likely highly variable. Despite this presumed variability in the animals we sampled, we did not observe appreciable variation within and between treatments in our transcriptional data. This suggests that variability in microbial load does not play as much of a role in defining transcriptional response as we had anticipated. Still, we cannot completely exclude the possibility that the expression of a few genes

did vary in response to microbial load and that we were unable to distinguish this true biological variation from experimental variation.

Wild-caught flies have comparable microbial loads to lab-reared gnotobiotic flies

We recognized throughout this experiment that our method of preparing gnotobiotic animals is one of many possible methods. Other studies have reported higher inocula for preparing gnotobiotic animals [15,16,29], or inoculating axenic flies upon eclosion rather than associating from embryo [12,19]. We wanted to understand how closely our experimentally-manipulated animals reflected their wild counterparts in terms of gut microbial density. We therefore sampled wild flies from Berkeley, CA and compared their microbial load to that of animals mono-associated in the laboratory. In doing so we had to address several potential complications. It has been well established that microbial load in laboratory *D. melanogaster* positively correlates with age [6,7,16], and thus we wanted to sample animals that would be comparable in age to those used in our sequencing experiments. However, we did not want to capture newly eclosed flies and rear them to the appropriate age in non-natural conditions, as this would deprive them of the substrate needed to replenish their natural gut microbiome and influence their microbial load [30].

We therefore established a stable food source for wild flies using a clean dishpan baited with organic watermelon (referred to as "the fendel") and then began leaving small pieces of organic fruit (grapes, apple or banana) near, but not touching, the watermelon in the fendel for 24 hours. After 24 hours, the fruit was transferred to a clean vial and any embryos that were deposited were grown up to approximately 5 days PE before surface-sterilization and plating, as done with the gnotobiotic flies. Since our rearing method necessarily involved using a finite number of baits, and we did not know how these baits could impact the final microbial load, we also collected wild flies of unknown age that came to feed in the fendel and subjected them to same plating and DNA extraction protocol. For each fly, DNA was extracted from half of the same homogenate to assay using a non-culture based method (qPCR) to determine microbial load, since some of the taxa within the microbial communities of wild flies might not necessarily grow well under lab conditions, and to confirm species identification, as we wanted only to sample wild *D. melanogaster* (*D. simulans* and *D. immigrans* are morphologically similar to *D. melanogaster*; all three species are found in Berkeley, CA). Culture-dependent and -independent
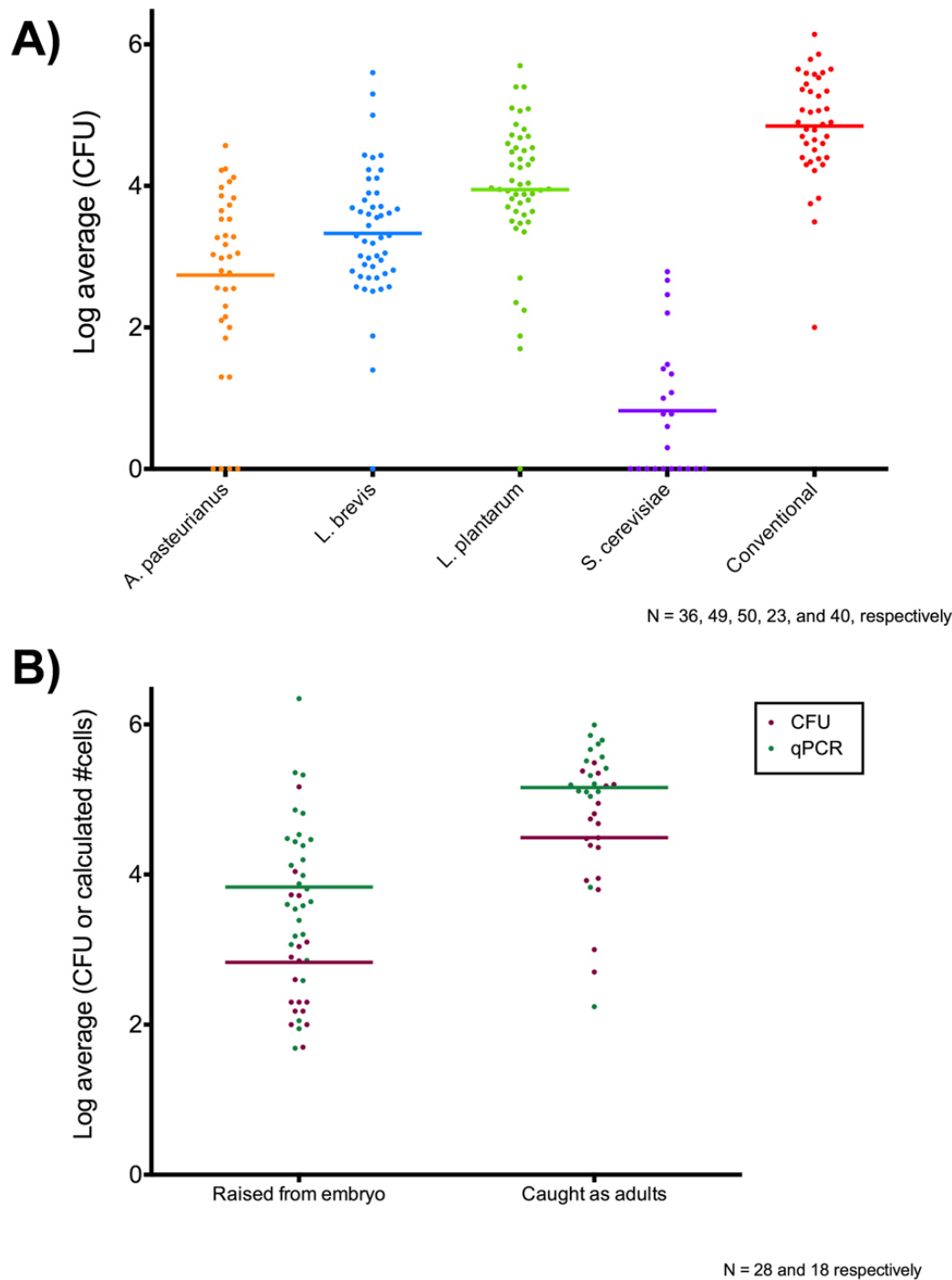
Figure 5) Microbial load of female *D. melanogaster* individuals. A) Log10 transformed average number of colony forming units (CFU) from plating individual gnotobiotic and conventional, lab-reared CantonS, *Wolbachia*-free, mated, 5-day post-eclosion females on two separate plates. B) Log10 transformed average number of CFU and estimated microbial cells (yeast and bacteria combined) by qPCR for individual, female, wild *D. melanogaster* raised from embryos (ranging from three to ten days post-eclosion) or caught as adults (of unknown age). The mean for each group is plotted as a horizontal line.

Elya et al, 2016   13

estimations of total microbial loads from flies ranging from 3-10 days and flies of unknown age are shown in Figure 5B.

Our microbial load data showed that five-day old flies are associated with fewer bacteria than wild-caught adults of indeterminate age (presumably older than our young flies and exposed to more substrates as an adult than our reared flies). The qPCR estimates of microbial load was generally higher than our culturing estimates, which is not surprising given the greater diversity of microbial species associated with wild *D. melanogaster* and the expectation that these different species will not necessarily thrive under our chosen culture conditions [2,3]. The number of yeast detected by qPCR was two to three orders of magnitude lower than that for bacteria, consistent with our observation for gnotobiotic animals (Figure S4). As with the gnotobiotic, lab-reared flies, wild fly microbial loads were highly variable even within flies of relatively controlled age. Nonetheless, the association estimates of gnotobiotic and wild flies showed that, with the exception of yeast mono-associated flies, our lab-reared flies bear microbial loads similar to wild flies.

Discussion

Our results are broadly consistent with previous studies showing that the presence of microbes in the environment and food has a significant effect on gene expression in the adult *D. melanogaster* gut, and we add several new observations to our understanding of the effects of microbes on host gut gene expression. Raising flies in the presence of three bacterial species associated with laboratory stocks of *D. melanogaster* (two of which are the highly abundant in lab-reared flies), either individually or in combination, has minimal effects on adult gut gene expression compared to flies raised axenically. There are, correspondingly, few differences in gut gene expression between flies raised on monocultures of different bacteria. In contrast, the presence of the yeast *S. cerevisiae*, either alone or in combination with bacteria, has a large effect on adult gut gene expression, suggesting that yeast plays a more important role than bacteria in shaping physiology of the adult gut.

Insensitivity of adult female gut to bacterial species

Contrary to our expectations, we did not find significant differences in gene expression between the guts of *D. melanogaster* individuals mono-associated with three different bacterial species (*A.*

pasteurianus, *L. brevis* or *L. plantarum*) commonly found in lab-reared *D. melanogaster*. This was surprising not only because of previous reports that these mono-associations had phenotypic consequences, but also because we expected, and confirmed, high variability in bacterial load between individuals within the same treatment group. We believe the most parsimonious explanation for our data is that the adult *D. melanogaster* gut is largely indifferent to the identity of bacteria species that occupy it under laboratory colonization levels that reflect those found in wild flies. But several important caveats warrant discussion.

First, that there is no difference in gut gene expression in response to these three strains does not mean that the adult gut is insensitive to bacterial identity. Indeed, while the species found in association with laboratory *D. melanogaster* are also found in wild flies, it is possible that the species that bacterial species that thrive amongst lab-reared flies are not representative of natural populations, and that the *Drosophila* gut would respond in different ways to different bacterial species. We also note that while the *A. pasteurianus* strains used in these experiments was collected from flies, the *Lactobacillus* strains came from non-fly sources (we used the sequenced strains of *L. plantarum* from the ATCC and *L. brevis* from the NBRC), and there may be important inter-strain differences that this choice obscured. Second, although the guts of five day old adult females show no differences between different bacterial mono-associations does not indicate that their guts did not respond differently at earlier points in development; we may simply have missed a critical period for these interactions. In fact, compelling work from several groups has demonstrated that certain fly-associated bacterial strains make significant contributions to larval development [15,16,29] and gut physiology [12,13]. Third, it is possible that the adult female *D. melanogaster* gut is capable of a differential response to these bacteria, but does so only under different conditions (e.g. different diet). Finally, it is possible that the guts in the flies we sampled were responding differently to these bacteria, but that these responses were not reflected at the level of gene expression. For example it could be that the common metabolic demands of the gut dominate its gene expression, and that the gut transduces bacteria-specific signals to other tissues.

Nonetheless, we think it is interesting that gene expression in the adult female *D. melanogaster* gut is so consistently insensitive to varying composition and levels of these bacteria under these conditions. There is a certain logic in insulating the gut from the effects of the constantly varying mix of diverse microbes that wild *D. melanogaster* are exposed to on their preferred substrate of rotting fruits, and our gene expression data may reflect precisely such an insulation. This is not to argue in any way that

bacteria in the gut have no effect on the fly, just simply that the epithelial cells of the adult gut do not appear to be the locus for these effects.

Effect of yeast on gut gene expression

One of the more striking results from this study is the dramatic effect of living *S. cerevisiae* in the environment on gene expression in the adult gut. The strong effect of yeast on *Drosophila* physiology is unsurprising, given the important role that fungi play in the *Drosophila* life cycle, especially as food for larvae. However, the fact that yeast have a large impact on gut gene expression despite the near complete absence of living yeast cells in the gut suggests an indirect effect, either due to differences in development or to the effects of yeast on the nutritional value or other properties of the media. One obvious possible factor is ethanol. We did not perform a chemical analysis of the media, but it is all but certain that the presence of *S. cerevisiae* led to quantifiable ethanol production. *D. melanogaster* preferentially oviposit in substrates emitting volatile fermentation products, and they have evolved a tolerance for relatively high ethanol levels as both larvae and adults. The direct effect of ethanol of gene expression in the *D. melanogaster* gut has not been studied, but dietary ethanol has been shown to have significant effects on gut morphology and physiology [31].

Interestingly, the gut expression data revealed that axenic or bacteria mono-associated guts overexpressed genes involved in fatty acid production and transport relative to yeast mono-associated or conventional guts. This observation is consistent with the finding that axenic animals contain more lipids than conventional animals [10], and suggests that sensation by the gut underlies the amount of fat produced by different gnotobiotic animals.

Whole animal gene expression paradox

Intriguingly, our experiments showed that, while yeast-containing treatments were sufficient to drive conventional-like transcription in the adjacent host tissue (i.e. guts), yeast alone or as part of a simplified microbial community (*A. pasteurianus*, *L. brevis*, *L. plantarum* and *S. cerevisiae*) was not sufficient to produce an animal-wide transcriptional program comparable to that of conventional animals. At first glance, these two findings seem to present a paradox. That is, one might think that if the difference

between animals is what microbes are contained in their gut and if the gut is not responding to this change then one should also not observe changes outside of the gut tissue.

We can imagine several ways in which to reconcile these seemingly contradictory results. First, given the repeated observation that different microbial taxa have profound influences on host biology at the larval stages [12,13,15,16,29], we believe that differences in host gut gene expression would likely be observed at earlier timepoints. If this were the case, differences in response to the conventional samples (either due to species, fungal or bacterial, that were not introduced in our simplified community and/or strain-specific effects that could not be recapitulated by our chosen isolates) could lead to differences that accumulate over developmental time and result in animals with markedly different global transcriptional programs. In turn, these accumulated differences could manifest in several ways. The simplest case is that developmental differences accumulated to result in transcriptional differences in an otherwise physically indistinguishable animal. More likely, it could be that certain tissues are over- or underrepresented in conventional animals versus other treatments (which would appear in our data as either higher or lower contributions of those tissues to the overall transcriptional program). Alternatively, the metabolic state of the adult conventional animal could be distinct from those of other treatments (feeding back on transcription to cause big changes).

Lastly, we observed excessive microbial growth on the food in the conventional vials and the conventional animals were much more likely to die between eclosion and collection than any other treatment. It is possible that this sort of stressful environment selected for animals that were acutely adapted to deal with excessive microbial growth, which manifested at the level of global transcription.

Challenges in utilizing *Drosophila* as a model system to study gut microbe-host biology

Lab-reared flies and their associated bacteria offer a convenient and powerful model for studying microbe-host interactions. However, the diversity and composition of the microbiomes of laboratory flies is limited compared to wild flies [2], likely because we have selected for microbes that thrive on the substrates that we use to culture flies in the lab. A consequence of this selection is that the effects of microbes associated with *D. melanogaster* in the lab may not reflect those of microbes they associate with in the wild.

Elya et al, 2016   17

It is also unclear exactly what it means for microbes to be associated with a fly. Though we generally refer to the microbes found within the fly gut as having colonized the gut, it has yet to be firmly established whether stable colonization of the gut occurs and under what conditions. Analogous to the mucus layer in the mammalian gut, the *Drosophila* gut possesses a barrier, the peritrophic matrix (PM), that impedes the direct contact between objects greater than ~250 KDa (including food particles and microbial cells) and host cells [32,33]. Similar to the mammalian mucosal barrier, recent evidence suggests that the PM is a dynamic structure capable of responding to pathogenic microbes [33]. However, while it has been established that some microbes in the vertebrate gut dine on mucus protecting the gut epithelium [34], it has not yet been firmly established whether *Drosophila*-associated microbes actually adhere to or subsist on the molecules comprising or associated with the PM. Additionally, flies appear to lack the crypt structures, present in vertebrate guts, that can provide hideout for microbes allowing them to persist during hard times [35]. At present, we know that lab-reared flies can lose their microbiota by continually transferring onto sterile food [30,36]. The turnover of fly flora under these conditions suggests that stable colonization of the gut tissue is not occurring under these circumstances, although it is possible that there are different conditions under which stable colonization could be observed.

A recent study of wild flies illustrates the importance of microbes to fly biology and highlights the distinct manner in which flies associate with microbial partners. Yamada et al. (2015) discovered that a yeast species, *Issatchenki orientalis*, increases amino acid availability and extends lifespan in flies reared on a protein-deficient diet [36]. They found that lifespan could be extended not just by providing flies with live *I. orientalis*, but also by providing heat-killed *I. orientalis.* Of note, the authors observed that daily transfer of flies onto sterile medium resulted in the loss of association with *I. orientalis.* Thus, there is evidence for flies benefiting from a non-stable association with a microbial species, and the benefit is conferred even when the microbes are dead. This might not be how we typically conceptualize important microbial interactions in animals, but it's possible that these are the most critical microbial interactions for flies.

Materials and Methods

Fly and microbial stocks

Wolbachia-free CantonS fly stocks were reared on medium from UC Berkeley's Koshland fly kitchen (Koshland diet; 0.68% agar, 6.68% cornmeal, 2.7% yeast, 1.6% sucrose, 0.75% sodium tartrate tetrahydrate, 5.6 mM CaCl2, 8.2% molasses, 0.09% tegosept, 0.77% ethanol, 0.46% propionic acid) at 25C. *Acetobacter pasteurianus* (CNE7) was isolated from lab-reared WT *Drosophila melanogaster* and grown on de Man, Rogosa and Sharpe (MRS) agar (Research Products International Corp) at 30C. *Lactobacillus brevis* (NBRC 107147) and *Lactobacillus plantarum* (ATCC 8014) were acquired from NBRC and ATCC culture collections, respectively, and grown on MRS agar at 30C. *Saccharomyces cerevisiae* (ASQ HI) was isolated from wild Hawaiian *Drosophila* by Alli Quan and grown on YPD agar at 30C (A. Quan, personal communication).

Shotgun DNA sequencing of WT fly guts

Female CantonS flies age five days post-eclosion (PE) were collected via cold anesthesia, surface-sterilized with 10% bleach for ten minutes and rinsed with sterile 1x PBS before dissecting out guts. Forceps used for dissection were treated with 3.5% H2O2 between animals to remove DNA contaminants. Dissections included the proventriculus to the rectal ampulla, leaving the Malpighian tubules attached. Each individual dissected guts was flash frozen in sterile-filtered Buffer ATL (QIAamp Micro Kit, QIAGEN) and stored at -80C. DNA was extracted according to QIAGEN's QIAamp Micro Kit tissue protocol, with modifications. After the overnight digestion with proteinase K, 0.1 mm Zirconium bead and 1 volume buffer WJL (2M Guanidinium thiocyanate, 0.5 M EDTA, 1.8% Tris base, 8% NaCl, pH 8.5) were added to each sample, Samples were then bead beat twice for one minute at 4C with a 30 second break in between, spun 5 minutes at ~14,000xg and the supernatant was transferred to a new tube. Beads were resuspended in two volumes buffer WJL and beat again an additional minute before spinning down and pooling supernatant. Beads were washed once more with two volumes buffer WJL before spinning down and pooling supernatant a final time. An additional spin was performed to pellet any carried-over beads; supernatant was transferred to a new tube. Each sample then received 1 ug of carrier RNA dissolved in buffer AE before proceeding with ethanol precipitation and elution per the manufacturer's protocol. DNA samples were quantified (Qubit dsDNA HS assay kit, ThermoFisher

Scientific) then treated with 120 ng/uL RNase A (QIAGEN) for 1 hour at 37C. Each RNase-treated sample was used to generate an indexed next-generation sequencing library using the TruSeq Nano DNA kit (Illumina). Samples were pooled and sequenced at 100 paired-end reads at the UC Davis Genome Center on a HiSeq 2500. Reads were first filtered by aligning to the *Drosophila melanogaster* genome (version 6.01) using bowtie2, then aligned to the Green Genes 16S database (release 13-5). Species that were identified by two reads or fewer were removed before rarefaction in R (http://www.jennajacobs.org/R/rarefaction.html). Samples that did not rarefy to saturation were discarded.

Preparation of axenic and gnotobiotic flies

Freshly-laid embryos from CantonS reared on Koshland diet were used to generate axenic, gnotobiotic and conventional animals. Axenic and gnotobiotic animals were prepared using a previously published protocol [10] with some modifications. Briefly, for all treatments, embryos were first collected on molasses plates and rinsed with distilled water. Conventional animals were generated by transferring water-rinsed embryos to YG diet (10% yeast, 10% glucose, 1.2% agar, 0.42% propionic acid, sterilized by autoclaving [37]. To generate axenic and gnotobiotic flies, water-rinsed embryos were subjected to 5 minutes in 10% household bleach, changing bleach solution once halfway through, then briefly rinsed in 70% ethanol before being rinsed with copious distilled water. For axenic animals, cleaned embryos were transferred to sterile YG diet. For gnotobiotic animals, YG vials were inoculated with 50 uL of 1 x 10^8 cell/mL suspension of one species (mono-associated) or an equal mixture of *A. pasteurianus* (CNE7), *L. brevis* (NBRC 107147) and *L. plantarum* with or without *S. cerevisiae* (HI ASQ) (poly-associated). To prepare cell suspensions, overnight liquid monocultures of microbes were pelleted, washed and resuspended to 1 x 10^{8}cell/mL in 1x PBS following OD600 conversions (given in [10]) for *A. pasteurianus*, *L. brevis* and *L. plantarum* or 3 x 10^8 cell/mL per OD600 of 1.0 for *S. cerevisiae* [38].

Fly collection and validation of gnotobiotic treatment

Gnotobiotic and conventional CantonS females reared on YG diet were collected at age five days PE via cold anesthesia using aseptic technique. For gut samples, individual guts were dissected from each of three female flies (from proventriculus through rectal ampulla including Malpighian tubules) in sterile 1x PBS using Dumont 55 forceps. Cleanly dissected guts were immediately deposited in Trizol

(ThermoFisher Scientific) and flash frozen in liquid nitrogen. Between samples, forceps were treated with 70% ethanol, flamed then dipped in 3.5% hydrogen peroxide and sterile water to prevent nucleic acid carryover between samples. For whole fly samples, individual cold-anesthetized animals were transferred to Trizol and frozen as above. *A. pastuerianus*, *L. brevis* and *L. plantarum* mono-associated animals for gut samples were collected for each of three independent gnotobiotic preparations. All other samples were collected from a single experiment, with preparations for different treatments occurring over the course of several months. Each time animals were prepared and sampled for sequencing three individuals (females when possible, males when not) were independently homogenized and plated on suitable agar medium to verify they possessed (or in the case of axenic animals, lacked) the expected microbial species. For mono-associated animals species verification was achieved by first confirming that plates showed only one morphology, performing 16S or ITS colony PCR on five representative colonies for each treatment and Sanger sequencing of the resultant products. For conventional and poly-associated animals treatment was first verified by checking morphology of resultant colonies performing 16S or ITS colony PCR on five representative colonies for each treatment and Sanger sequencing of the resultant products.

RNA preparation and sequencing

RNA was prepared from each thawed sample by homogenizing with an RNase-free pestle (Kimble Chase), washing the pestle with 750 uL Trizol, then proceeding using the manufacturer's protocol with 10 ug glycogen carrier per sample. RNA quality was checked by running on a RNA 6000 Pico chip on a Bioanalyzer 2100 (Agilent Technologies) and quantified using a Qubit fluorometer (Qubit RNA HS assay kit, ThermoFisher Scientific). High quality RNA was then treated with Turbo DNase (ThermoScientific) per the manufacturer's protocol. For gut samples, RNAseq libraries were prepared using the TruSeq RNA v2 kit (Illumina) starting with 100-200 ng of DNase-treated total RNA for each sample. For whole fly samples, RNAseq libraries were prepared using 400-500 ng of RNA per sample. Samples were multiplexed and sequenced using 100-150 bp paired-end reads on a HiSeq 2500 at the QB3 Vincent J. Coates Genomic Sequencing Facility at UC Berkeley. Reads were aligned to the *D. melanogaster* genome (version 6.01) using Tophat using the "--no-mixed" option (Table 1). Transcript abundance was calculated from aligned reads using Cufflinks based on a reference transcriptome lacking RNA genes, as we find these to be wildly variable in any RNAseq experiment (M. Eisen, personal communication). Data were analyzed using hierarchical clustering by gene (Cluster 3.0), ANOVA between grouped treatments

(scipy.stats) and GO term analysis (Panther, [28]). Data were plotted using matplotlib (Python) and Prism 6 (GraphPad). Heatmap color scales were defined with Chris Slocum's custom_cmap.py (http://schubert.atmos.colostate.edu/~cslocum/custom_cmap.html).

Table 1. Alignment statistics for all mRNAseq libraries reported in the present manuscript.

Lane	Read type^	Sample type	Sample ID	Total input read pairs*	Concordantly-aligned pairs*	Concordant alignment rate*
1	100 PE	Gut	Ap1a	4.53E+07	2.84E+07	62.6
1	100 PE	Gut	Ap1b	2.74E+07	1.70E+07	61.8
1	100 PE	Gut	Ap1c	4.96E+07	3.10E+07	62.5
1	100 PE	Gut	Lbrev1a	1.39E+07	8.61E+06	61.9
1	100 PE	Gut	Lbrev1b	2.08E+07	1.30E+07	62.2
1	100 PE	Gut	Lbrev1c	2.06E+07	1.22E+07	59.1
1	100 PE	Gut	Lp1a	1.94E+07	1.17E+07	60.6
1	100 PE	Gut	Lp1b	2.46E+07	1.51E+07	61.3
1	100 PE	Gut	Lp1c	5.41E+07	3.53E+07	65.2
2	100 PE	Gut	Ap2a	1.14E+07	8.56E+06	74.9
2	100 PE	Gut	Ap2b	1.71E+07	1.28E+07	75.2
2	100 PE	Gut	Ap2c	1.84E+07	1.39E+07	75.7
2	100 PE	Gut	Lbrev2a	1.34E+07	1.05E+07	78.0
2	100 PE	Gut	Lbrev2b	2.37E+07	1.80E+07	76.0
2	100 PE	Gut	Lbrev2c	4.64E+07	3.55E+07	76.5
2	100 PE	Gut	Lp2a	3.92E+07	3.01E+07	76.8
2	100 PE	Gut	Lp2b	2.12E+07	1.63E+07	76.9
3	100 PE	Gut	Ap3a	2.19E+07	1.70E+07	77.4
3	100 PE	Gut	Ap3b	1.54E+07	1.20E+07	77.9
3	100 PE	Gut	Ap3c	2.07E+07	1.60E+07	77.4
3	100 PE	Gut	Lbrev3a	2.21E+07	1.71E+07	77.4
3	100 PE	Gut	Lbrev3b	1.70E+07	1.34E+07	78.4
3	100 PE	Gut	Lbrev3c	2.61E+07	2.00E+07	76.8
3	100 PE	Gut	Lp2c	2.05E+07	1.57E+07	76.5
3	100 PE	Gut	Lp3a	2.19E+07	1.68E+07	76.7
3	100 PE	Gut	Lp3b	3.88E+07	3.06E+07	78.9
3	100 PE	Gut	Lp3c	1.57E+07	1.22E+07	77.5

Elya et al, 2016

4	100 PE	Gut	Ax1	1.58E+07	1.19E+07	75.7
4	100 PE	Gut	Ax2	3.34E+07	2.56E+07	76.6
4	100 PE	Gut	Ax3	2.96E+07	2.28E+07	77.2
4	100 PE	Gut	Conv1	3.86E+07	2.98E+07	77.3
4	100 PE	Gut	Conv2	2.50E+07	1.92E+07	77.0
4	100 PE	Gut	Conv3	3.33E+07	2.54E+07	76.4
4	100 PE	Gut	Yeast1	3.26E+07	2.67E+07	81.8
4	100 PE	Gut	Yeast2	1.65E+07	1.25E+07	75.7
4	100 PE	Gut	Yeast3	2.10E+07	1.59E+07	75.5
5	100 PE	Whole	ApW1	6.75E+06	5.73E+06	84.9
5	100 PE	Whole	ApW2	7.51E+06	6.30E+06	83.9
5	100 PE	Whole	ApW3	3.77E+06	3.23E+06	85.6
5	100 PE	Whole	AxenicW1	1.05E+07	8.91E+06	85.0
5	100 PE	Whole	AxenicW2	7.64E+06	6.45E+06	84.4
5	100 PE	Whole	AxenicW3	7.97E+06	6.71E+06	84.3
5	100 PE	Whole	ConvW1	2.17E+06	1.84E+06	84.6
5	100 PE	Whole	ConvW2	1.04E+07	8.81E+06	85.1
5	100 PE	Whole	ConvW3	7.14E+06	6.06E+06	84.9
5	100 PE	Whole	LbrevW1	7.77E+06	6.69E+06	86.1
5	100 PE	Whole	LbrevW2	8.81E+06	7.60E+06	86.3
5	100 PE	Whole	LbrevW3	8.03E+06	6.88E+06	85.7
5	100 PE	Whole	LpW1	7.72E+06	6.55E+06	84.8
5	100 PE	Whole	LpW2	1.09E+07	9.40E+06	86.1
5	100 PE	Whole	LpW3	1.22E+07	1.05E+07	86.4
5	100 PE	Whole	YeastW1	1.07E+07	9.21E+06	86.0
5	100 PE	Whole	YeastW2	1.37E+07	1.18E+07	85.6
5	100 PE	Whole	YeastW3	1.37E+07	1.16E+07	84.8
6	150 PE	Gut	3bac1	4.38E+06	2.39E+06	54.6
6	150 PE	Gut	3bac3	5.02E+06	2.83E+06	56.4
6	150 PE	Whole	3bacW1	9.11E+06	5.03E+06	55.2
6	150 PE	Whole	3bacW2	6.13E+06	3.41E+06	55.6
6	150 PE	Whole	3bacW3	2.04E+06	1.15E+06	56.3
6	150 PE	Gut	4mic1	3.53E+06	1.97E+06	55.8

6	150 PE	Gut	4mic2	4.70E+06	2.71E+06	57.7
6	150 PE	Gut	4mic3	3.78E+06	2.20E+06	58.1
6	150 PE	Whole	4micW1	5.86E+06	3.21E+06	54.7
6	150 PE	Whole	4micW2	6.97E+06	3.51E+06	50.4
6	150 PE	Whole	4micW3	1.09E+07	5.90E+06	54.2
6	150 PE	Gut	Yeast4	2.80E+06	1.60E+06	57.0
6	150 PE	Gut	Yeast5	3.76E+06	2.08E+06	55.4
6	150 PE	Gut	Yeast6	3.68E+06	1.97E+06	53.5

*Value obtained from the "aligment_summary.txt" output following Tophat alignment of reads to v6.01 *D. melanogaster* genome.

^All libraries were sequenced on the Illumina HiSEq 2500 platform

Determining gut microbe colonization of gnotobiotic flies

Gnotobiotic (mono-associated), axenic and conventional flies were prepared and reared on YG diet five separate times as described above. CantonS females were collected at age five days PE via cold anesthesia using aseptic technique. Animals were surface sterilized using a one minute incubation in 95-100% ethanol then rinsed in sterile water. For conventional and gnotobiotic treatments, a ten flies or as many as available were transferred to individual tubes containing 1x PBS, homogenized with sterile pestles and plated at the dilutions and on the media as shown in Table 2.

Table 2. Plating media and dilutions for determining gut microbial load of lab-reared flies.

Treatment	Plate A	Plate B
Conventional	1/200, MRS	1/200, YPD
A. pasteurianus	1/10, MRS	1/20, MRS
L. brevis	1/50, MRS	1/100, MRS
L. plantarum	1/50, MRS	1/100, MRS
S. cerevisiae	1/4, YPD	1/4, YPD

Elya et al, 2016   24

For axenic preparations, three females were homogenized in each of two tubes of 1x PBS without prior surface sterilization. One homogenate was plated on MRS and one on YPD. A negative control was performed for each set of flies sampled by homogenizing an equal volume of 1x PBS and plating the half on each MRS and YPD. All plates were incubated at 30C for three days or until plates showed visible colonies. Colonies were counted by hand with the aid of a light board and a handheld counter. The count from each plate was corrected for dilution and averaged over two plates for each fly.

Determining gut microbe colonization of wild flies

To collect wild flies of unknown age, flies were directly caught either by baiting a closed bottle trap with banana or by directly aspirating from an uncovered plastic dishwashing pan (heretofore referred to as a "fendel") that was baited with organic watermelon and an assortment of other organic fruits. All baiting and capture was performed in the spring of 2015 at a personal residence in Berkeley, CA. Flies were captured in the morning and transferred into sterile, empty vials (to avoid providing a diet that might manipulate their gut flora) before sampling shortly thereafter. Flies were recovered from vials via cold anesthesia using aseptic technique. Males were discarded.

To collect wild flies of roughly five days PE, a piece of tap-water rinsed organic fruit (grape, banana or fuji apple) was placed in the open-bait fendel and left for 24 hours for provide an oviposition substrate for wild females. After 24 hours, the fruit pieces were placed in sterile vials and the deposited embryos were allowed to develop at room temperature. Vials were monitored daily for newly-eclosed flies. All flies from a vial were collected when the majority of the adults present in the vial were 5 days PE via cold anesthesia using aseptic technique. Males were discarded.

All wild flies were then surface sterilized using a one minute incubation in 95-100% ethanol and rinsed in sterile water before homogenization with a sterile pestle in 1x PBS. Half of the homogenate was diluted to plate 1/200 across MRS agar and YPD agar for each fly and plates were diluted at 30C for three days or until colonies emerged. The remaining half was then processed for DNA extraction as described for shotgun DNA sequencing or frozen at -30C for later processing. A negative control was performed for each set of flies collected by homogenizing an equal volume of 1x PBS with a sterile pestle and splitting

in half as for fly samples. Colonies were counted by hand with the aid of a light board and a handheld counter. The count from each plate was corrected for dilution and averaged over two plates for each fly.

All sampled wild females were genotyped by PCR and Sanger sequencing cytochrome c oxidase II primers tLYS (GTTTAAGAGACCAGTACTTG) and tLEU (ATGGCAGATTAGTGCAATGG)[39]. Colonization data from flies that were not identified as *D. melanogaster* (e.g. *D. simulans*, *D. immigrans*, *D. persimilis*) were discarded.

To provide a culture-independent method for estimating gut colonization in wild animals, all DNA samples from confirmed wild *D. melanogaster* females were subjected to two sets of triplicate qPCR reactions, one using universal bacteria 16S primer pair P1 (CCTACGGGAGGCAGCAG) and P2 (ATTACCGCGGCTGCTGG) at 100 nM per reaction and the other using fungal ITS primer pair Y1 (GCGGTAATTCCAGCTCCAATAG) and Y2 (GCCACAAGGACTCAAGGTTAG) at 800 nM per reaction [40]. Reactions were performed on a Roche480 LightCycler using Sybr Green mastermix (Roche) templated with 10-20 ng of total fly DNA for each reaction. Four-point standard curves were run for each experiment for *A. pasteurianus* and *L. plantarum* (P1, P2 primer pair) and *S. cerevisiae* (Y1, Y2 primer pair). The amplification program used for qPCR is as follows: 95C for 5 min then 45 cycles of 95 C for 10 seconds, 65C for 10 seconds (decreasing 1 C for the first 10 cycles until reaching 55C), and 72C for 10 seconds. Primers were validated using DNA extracted from monocultures of *A. pasteurianus, L. brevis, L. plantarum* and *S. cerevisiae.* The number of bacterial and yeast cells present in the whole animal was estimated assuming an average genome size of 3 Mb for bacteria and 25 Mb for yeast.

References

1. McFall-Ngai M, Hadfield MG, Bosch TCG, Carey HV, Domazet-Lošo T, Douglas AE, et al. Animals in a bacterial world, a new imperative for the life sciences. Proc Natl Acad Sci U S A. 2013;110: 3229–3236.

2. Chandler JA, Lang JM, Bhatnagar S, Eisen JA, Kopp A. Bacterial communities of diverse *Drosophila* species: ecological context of a host-microbe model system. PLoS Genet. 2011;7: e1002272.

3. Staubach F, Baines JF, Künzel S, Bik EM, Petrov DA. Host species and environmental effects on bacterial communities associated with *Drosophila* in the laboratory and in the natural environment. PLoS One. 2013;8: e70749.

4.   Corby-Harris, V, Pontaroli AC, Shimkets LJ, Bennetzen JL, Habel KE, Promislow DEL. Geographical Distribution and Diversity of Bacteria Associated with Natural Populations of *Drosophila melanogaster*. Appl Environ Microbiol. 2007;73: 3470–3479.

5.   Cox CR, Gilmore MS. Native Microbial Colonization of *Drosophila melanogaster* and Its Use as a Model of Enterococcus faecalis Pathogenesis. Infect Immun. 2007;75: 1565–1576.

6.   Wong CNA, Ng P, Douglas AE. Low-diversity bacterial community in the gut of the fruitfly *Drosophila melanogaster*. Environ Microbiol. 2011;13: 1889–1900.

7.   Ren C, Webster P, Finkel SE, Tower J. Increased Internal and External Bacterial Load during *Drosophila* Aging without Life-Span Trade-Off. Cell Metab. Elsevier Inc.; 2007;6: 144–152.

8.   Ridley EV, Wong AC-N, Westmiller S, Douglas AE. Impact of the resident microbiota on the nutritional phenotype of *Drosophila melanogaster*. PLoS One. 2012;7: e36765.

9.   Bakula M. The persistence of a microbial flora during postembryogenesis of *Drosophila melanogaster*. J Invertebr Pathol. 1969;14: 365–374.

10. Newell PD, Douglas AE. Interspecies interactions determine the impact of the gut microbiota on nutrient allocation in *Drosophila melanogaster*. Appl Environ Microbiol. 2014;80: 788–796.

11. Erkosar Combe B, Defaye A, Bozonnet N, Puthier D, Royet J, Leulier F. <italic>Drosophila</italic> Microbiota Modulates Host Metabolic Gene Expression via IMD/NF-κB Signaling. PLoS One. Public Library of Science; 2014;9: e94729.

12. Jones RM, Desai C, Darby TM, Luo L, Wolfarth AA, Scharer CD, et al. Lactobacilli Modulate Epithelial Cytoprotection through the Nrf2 Pathway. Cell Rep. 2015;12: 1217–1225.

13. Jones RM, Luo L, Ardita CS, Richardson AN, Kwon YM, Mercante JW, et al. Symbiotic lactobacilli stimulate gut epithelial proliferation via Nox-mediated generation of reactive oxygen species. EMBO J. 2013;32: 3017–3028.

14. Ryu JH, Kim SH, Lee HY, Bai JY, Nam YD, Bae JW, et al. Innate Immune Homeostasis by the Homeobox Gene Caudal and Commensal-Gut Mutualism in *Drosophila*. Science. 2008;319: 777–782.

15. Shin SC, Kim S-H, You H, Kim B, Kim AC, Lee K-A, et al. *Drosophila* microbiome modulates host developmental and metabolic homeostasis via insulin signaling. Science. 2011;334: 670–674.

16. Storelli G, Defaye A, Erkosar B, Hols P, Royet J, Leulier F. Lactobacillus plantarum Promotes *Drosophila* Systemic Growth by Modulating Hormonal Signals through TOR-Dependent Nutrient Sensing. Cell Metab. Elsevier Inc.; 2011;14: 403–414.

17. Li H, Qi Y, Jasper H. Preventing Age-Related Decline of Gut Compartmentalization Limits Microbiota Dysbiosis and Extends Lifespan. Cell Host Microbe. 2016;19: 240–253.

18. Guo L, Karpac J, Tran SL, Jasper H. PGRP-SC2 Promotes Gut Immune Homeostasis to Limit Commensal Dysbiosis and Extend Lifespan. Cell. Elsevier Inc.; 2014;156: 109–122.

19. Fast D, Duggal A, Foley E. Lactobacillus plantarum is a pathobiont for adult *Drosophila* [Internet]. bioRxiv. 2016. p. 049981. doi:10.1101/049981

20. Venu I, Durisko Z, Xu J, Dukas R. Social attraction mediated by fruit flies' microbiome. J Exp Biol. 2014;217: 1346–1352.

21. Najarro MA, Sumethasorn M, Lamoureux A, Turner TL. Choosing mates based on the diet of your ancestors: replication of non-genetic assortative mating in *Drosophila melanogaster*. PeerJ. 2015;3: e1173.

22. Sharon G, Segal D, Ringo JM, Hefetz A, Zilber-Rosenberg I, Rosenberg E. Commensal bacteria play a role in mating preference of *Drosophila melanogaster*. Proceedings of the National Academy of Sciences. National Acad Sciences; 2010;107: 20051–20056.

23. Broderick NA, Buchon N, Lemaitre B. Microbiota-induced changes in *Drosophila melanogaster* host gene expression and gut morphology. MBio. 2014;5: e01117–14.

24. Chandler JA, Eisen JA, Kopp A. Yeast communities of diverse *Drosophila* species: comparison of two symbiont groups in the same hosts. Appl Environ Microbiol. American Society for Microbiology; 2012;78: 7327–7336.

25. Sang JH, King RC. Nutritional Requirements of Axenically Cultured *Drosophila Melanogaster* Adults. J Exp Biol. 1961;38: 793–809.

26. Coluccio AE, Rodriguez RK, Kernan MJ, Neiman AM. The yeast spore wall enables spores to survive passage through the digestive tract of *Drosophila*. PLoS One. 2008;3: e2873.

27. Shihata AME-TA, Mrak EM. The Fate of Yeast in the Digestive Tract of *Drosophila*. Am Nat. The University of Chicago Press for The American Society of Naturalists; 1951;85: 381–383.

28. Mi H, Poudel S, Muruganujan A, Casagrande JT, Thomas PD. PANTHER version 10: expanded protein families and functions, and analysis tools. Nucleic Acids Res. 2016;44: D336–42.

29. Erkosar B, Storelli G, Mitchell M, Bozonnet L, Bozonnet N, Leulier F. Pathogen Virulence Impedes Mutualist-Mediated Enhancement of Host Juvenile Growth via Inhibition of Protein Digestion. Cell Host Microbe. 2015;18: 445–455.

30. Blum JE, Fischer CN, Miles J, Handelsman J. Frequent Replenishment Sustains the Beneficial Microbiome of *Drosophila melanogaster*. MBio. 2013;4: e00860–13–e00860–13.

31. Hartman JR, Dybas LK, Geer BW. At high dietary levels ethanol alters the structure of mid- and hindgut epithelial cells of *Drosophila melanogaster* larvae. J Exp Zool. 1993;267: 365–376.

32. Hegedus D, Erlandson M, Gillott C, Toprak U. New Insights into Peritrophic Matrix Synthesis, Architecture, and Function. Annu Rev Entomol. 2009;54: 285–302.

33. Kuraishi T, Binggeli O, Opota O, Buchon N, Lemaitre B. Genetic evidence for a protective role of the peritrophic matrix against intestinal bacterial infection in *Drosophila melanogaster*. Proc Natl Acad Sci U S A. 2011;108: 15966–15971.

34. Johansson MEV, Larsson JMH, Hansson GC. The two mucus layers of colon are organized by the MUC2 mucin, whereas the outer layer is a legislator of host–microbial interactions. Proceedings of the National Academy of Sciences. 2011;108: 4659–4665.

35. Lemaitre B, Miguel-Aliaga I. The Digestive Tract of *Drosophila melanogaster*. Annu Rev Genet. 2013;47: 377–404.

36. Yamada R, Deshpande SA, Bruce KD, Mak EM, Ja WW. Microbes Promote Amino Acid Harvest to Rescue Undernutrition in *Drosophila*. Cell Rep. 2015; doi:10.1016/j.celrep.2015.01.018

37. Ridley EV, Wong ACN, Douglas AE. Microbe-dependent and nonspecific effects of procedures to eliminate the resident microbiota from *Drosophila melanogaster*. Appl Environ Microbiol. 2013;79: 3209–3214.

38. Day A, Schneider C, Schneider BL. Yeast Cell Synchronization. Cell Cycle Checkpoint Control Protocols. 2003. pp. 55–76.

39. Liu H, Beckenbach AT. Evolution of the mitochondrial cytochrome oxidase II gene among 10 orders of insects. Mol Phylogenet Evol. 1992;1: 41–52.

40. Xu W, Huang Z, Zhang X, Li Q, Lu Z, Shi J, et al. Monitoring the microbial community during solid-state acetic acid fermentation of Zhenjiang aromatic vinegar. Food Microbiol. Elsevier Ltd; 2011;28: 1175–1181.

Supplementary Figures
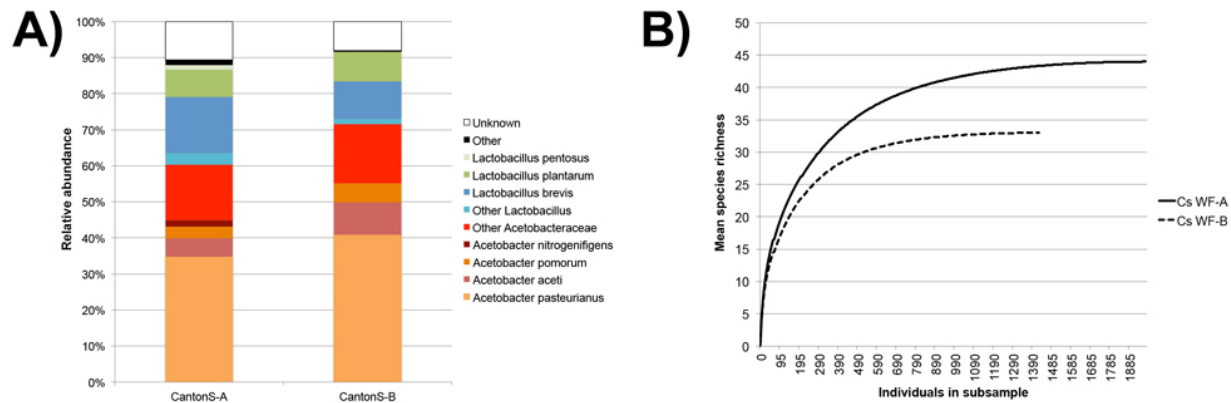
Figure S1. Gut bacterial community analysis of *Wolbachia*-free CantonS five-day mated females reared on Koshland diet. A) Relative species abundance of gut bacteria as determined by shotgun sequencing reads to the Green Genes 16S rRNA database release 13-5. "Other" category includes species not listed in key. "Unknown" category includes reads that aligned to 16S rRNA sequences included in the Green Genes database annotated as "unknown" (e.g. unknown compost). B) Rarefaction curve using data shown in A [0].
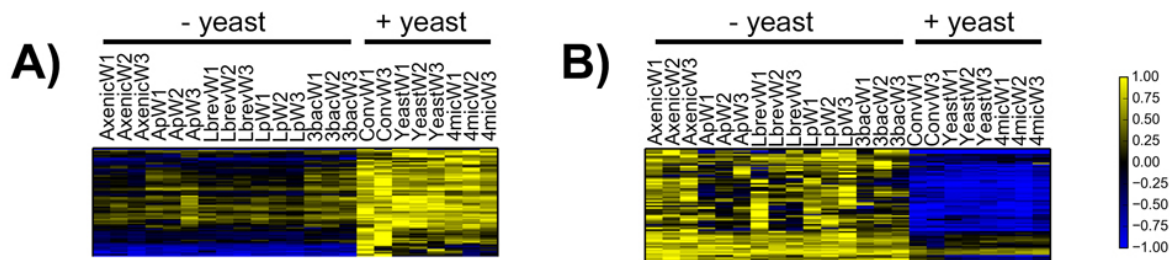
Figure S2. Genes expressed more similarly within yeast-containing and conventional whole flies versus all other samples. A) heatmap of 72 genes that are overexpressed in conventional, yeast- mono- and poly-associated whole flies compared to other whole fly samples (Bonferroni p-value>0.05, ANOVA). B) Heatmap of 67 genes that are overexpressed in axenic and bacteria mono- and poly-associated whole flies compare to other whole fly samples (Bonferroni p-value>0.05, ANOVA). Abbreviations: Ap = *A. pasteurianus*-mono-associated; Lbrev = *L. brevis*-mono-associated, Lp = *L plantarum*-mono-associated, 3bac = poly-associated without yeast, Ax = axenic, Conv = conventional, Yeast = *S. cerevisiae*-mono-associated, 4mic = poly-associated with yeast.
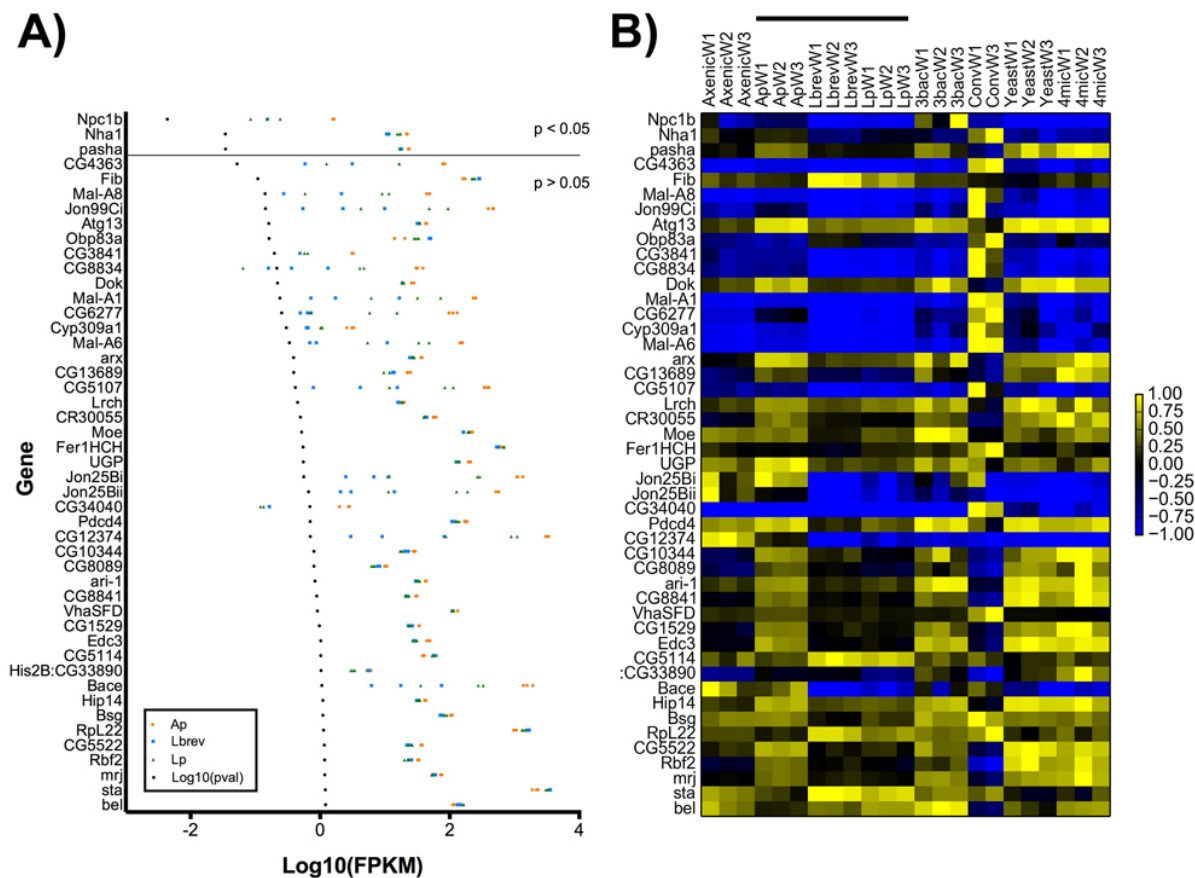
Figure S3. Genes showing greatest difference in expression values between different bacteria mono-associations in whole adults as determined by one-way ANOVA. A) Scatterplot of log10-transformed FPKM values for each bacteria mono-associated whole fly replicate. Genes are ordered from lowest ANOVA p-value (top) to highest (bottom). P-values have undergone a Bonferroni correction for multiple testing. B) Data from A presented as a heatmap. FPKM values for each gene are normalized to range from -1 to 1 before plotting. Black line above heatmap denotes bacteria mono-association samples. Abbreviations: Ap = *A. pasteurianus*-mono-associated; Lbrev = *L. brevis*-mono-associated, Lp = *L plantarum*-mono-associated, 3bac = poly-associated without yeast, Ax = axenic, Conv = conventional, Yeast = *S. cerevisiae*-mono-associated, 4mic = poly-associated with yeast.
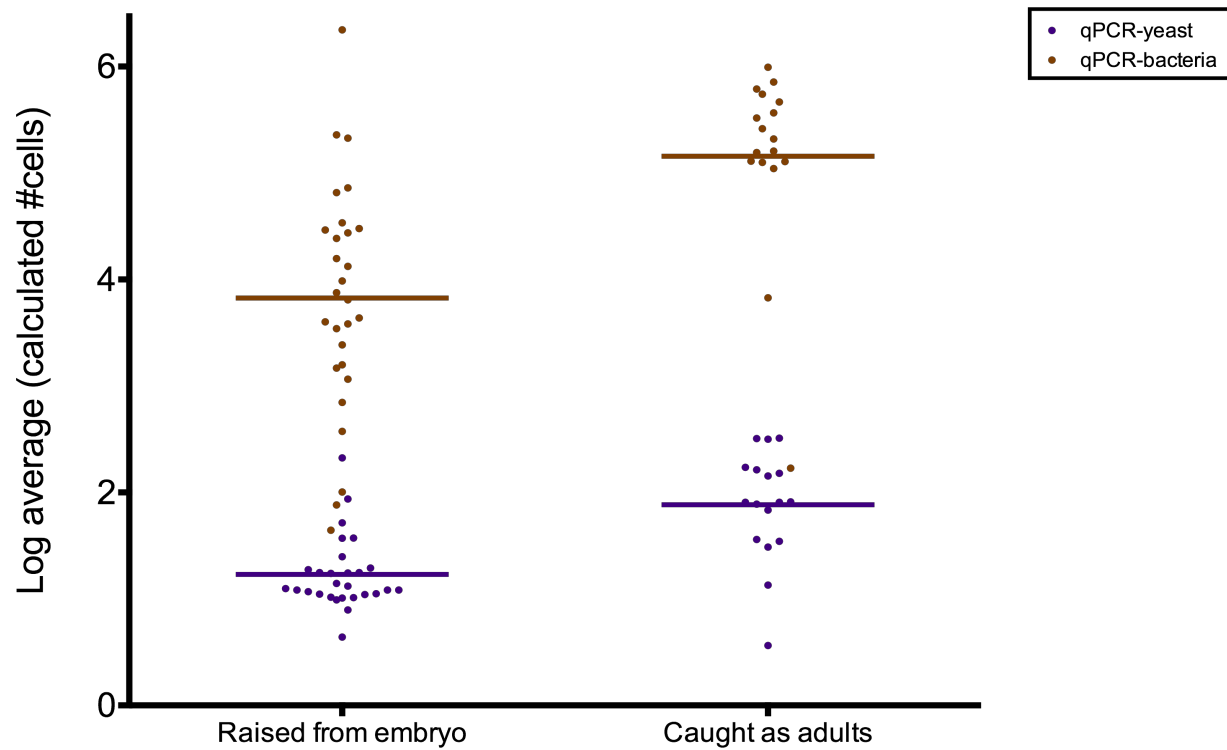
N = 28 and 18 respectively

Figure S4. Loads of bacteria and yeast in wild flies estimated from qPCR. Log10 transformed average number of estimated bacteria cells or yeast cells by qPCR for individual, female, wild *D. melanogaster* raised from embryos (ranging from 3-10 days post-eclosion) or caught as adults (of unknown age). The mean for each group is plotted as a horizontal line.