Title: Identification of Microbiota-Induced Gene Expression Changes in the *Drosophila melanogaster* Head

Scott A. Keith*, Rory Eutsey*, Heewook Lee†‡, Brad Solomon†‡, Stacie Oliver*, Carl Kingsford†, N. Luisa Hiller*, Brooke M. McCartney*

* Department of Biological Sciences, Carnegie Mellon University, Pittsburgh, PA 15213

† Computational Biology Department, Carnegie Mellon University, Pittsburgh, PA 15213

‡ Equal contribution

Running title: Microbiota impact *Drosophila* head gene expression

Keywords: *Drosophila,* microbiota, gene expression, immunity, metabolism

Corresponding author:

Brooke M. McCartney

600D Mellon Institute

Department of Biological Sciences

Carnegie Mellon University

4400 Fifth Avenue

Pittsburgh, PA 15213

Phone: (412) 268-5195

Email: bmccartney@cmu.edu

ABSTRACT

Symbiotic microorganisms exert multifaceted impacts on the physiology of their animal hosts. Recent discoveries have shown the gut microbiota influence host brain function and behavior, but the host and microbial molecular factors required to actuate these effects are largely unknown. To uncover molecular mechanisms that underlie the gut-microbiota-brain axis, we used *Drosophila melanogaster* and its bacterial microbiota as a model to identify microbiota-dependent gene expression changes in the host brain and head. Specifically, we employed RNA-seq and nanoString nCounter technology to identify *Drosophila* genes that exhibit altered transcript levels in fly heads upon elimination of the microbiota. The identified genes, some of which exhibited sex-specific differences, have demonstrated or inferred functional roles in the immune response, metabolism, neuronal activity, and stress resistance. Overall, this study reveals microbiota-responsive genes in the fly head, an anatomical structure not previously investigated in this context. Our results serve as a foundation for future investigations of how microbe-driven gene expression changes impact *Drosophila* biology.

INTRODUCTION

Throughout their lifespan, animals engage in complex and dynamic interactions with microbial communities that reside both in and on their bodies, and in their environment. Association with a microbiota impacts many physiological and life history traits of the animal host, and in certain environmental contexts is vitally important to supporting host health and homeostasis (McFall-Ngai *et al.* 2013). The microbial communities occupying the mammalian gut promote host immune development (Atarashi *et al.* 2011; Rosser and Mauri 2016), aid food metabolism (Cockburn and Koropatkin 2016), and alter neural function and behaviors (Hsiao *et al.* 2013; Sampson and Mazmanian 2015; Sharon *et al.* 2016). In these and other ways, symbiotic microorganisms exert their influence on animal health at every hierarchical level of biological organization, ranging from the molecular to the ecological (Kohl and Carey 2016). However, despite the wealth of evidence demonstrating the importance of animal-microbiota interactions, the underlying molecular and mechanistic principles of these interactions are only beginning to be uncovered.

Drosophila melanogaster and its bacterial gut symbionts provide a microbiologically and genetically controllable system with which to interrogate the molecular basis of microbiota-dependent host traits. In the lab and in the wild, *Drosophila* continuously encounter microbe-rich food substrates (Broderick and Lemaitre 2012; Wong *et al.* 2015; Bost *et al.* 2017, 2018; Adair *et al.* 2018). Unlike mammals and other higher order metazoans, the bacterial communities associated with *Drosophila* are low diversity, and are dominated by culturable taxa, predominantly *Lactobacillus* and *Acetobacter* species (Brummel *et al.* 2004; Wong *et al.* 2011, 2013, 2015; Blum *et al.*

2013; Broderick *et al.* 2014; Elya *et al.* 2016; Adair *et al.* 2018). Interactions between *Drosophila* and its external symbionts are also readily studied by generating sterile, germ-free (GF) flies, and gnotobiotic (GNO) flies mono- or polyassociated with defined bacterial isolates (Koyle *et al.* 2016). The microbial tractability of its gut consortium complements the abundant genetic tools that enable thorough investigation of host gene function in *Drosophila*. Thus, the ability to manipulate host genetics and its microbiota makes *Drosophila* an excellent model to study the molecular mechanisms that underlie microbial modulation of host biology.

The fly microbiota strongly impact a variety of physiological traits displayed by laboratory *Drosophila*. Both lactic acid and acetic acid bacteria alter the nutritional availability of particular diet substrates that enables or accelerates larval development, and shapes the metabolic profile of adults (Shin *et al.* 2011; Chaston *et al.* 2014, 2016; Huang and Douglas 2015; Matos *et al.* 2017; Storelli *et al.* 2017; Téfit and Leulier 2017; Sannino *et al.* 2018). Microbes in the gut lumen affect local gut biology by stimulating the proliferation of gut epithelial cells (Buchon *et al.* 2009; Jones *et al.* 2013, 2015; Li *et al.* 2016; Petkau *et al.* 2017), driving reactive oxygen species production (Jones *et al.* 2013, 2015; Guo *et al.* 2014), modulating innate immune activity (Ryu *et al.* 2008; Lee *et al.* 2013; Broderick *et al.* 2014; Combe *et al.* 2014; Sansone *et al.* 2015), and protecting against stressors like pathogens and oxidative injury (Ryu *et al.* 2008; Blum *et al.* 2013; Jones *et al.* 2013, 2015; Sansone *et al.* 2015). Newly emerging evidence is connecting the bacterial microbiota with fly behavior and neural function. Flies demonstrate modified, olfactory-mediated feeding and egg-laying preferences in response to individual bacteria, mixed microbial communities, and microbial fermentation products (Farine *et al.* 2017;

Fischer *et al.* 2017; Kim *et al.* 2017; Leitão-Gonçalves *et al.* 2017; Liu *et al.* 2017; Wong *et al.* 2017). Adult GF *Drosophila* exhibit increased locomotor activity, which can be suppressed by a *Lactobacillus brevis*-derived, secreted metabolic enzyme, xylose isomerase (Schretter *et al.* 2018). The ability of both *Acetobacter pomorum* and *Lactobacillus plantarum* to promote larval development on nutrient-limited diets requires induction of *Drosophila* insulin-like peptides (Shin *et al.* 2011; Storelli *et al.* 2011), which are synthesized by neuroendocrine insulin-producing cells (IPCs) in the brain and released into circulation to induce systemic developmental and metabolic phenotypes (Géminard *et al.* 2009; Nässel and Broeck 2016). Reciprocally, genetic induction of tumors in the optic lobes of the larval brain and eye-antennal discs was sufficient to perturb the structure and diversity of bacterial communities in the larval gut (Jacqueline *et al.* 2017). These discoveries suggest that modulation of neuronal function by *Drosophila*-associated gut microbes can impact a range of physiological, behavioral, and life-history traits in the insect. However, the host molecules and pathways induced by interactions with the microbiota that result in these microbe-dependent host traits remain largely undiscovered.

Several studies have profiled the microbiota's impact on gene expression in the adult *Drosophila* gut (Broderick *et al.* 2014; Guo *et al.* 2014; Elya *et al.* 2016; Petkau *et al.* 2017), embryo (Elgart *et al.* 2016), larvae (Erkosar *et al.* 2017), and whole animals (Combe *et al.* 2014; Dobson *et al.* 2016; Bost *et al.* 2017). These investigations found generally similar functional categories of genes differentially expressed under GF conditions across the indicated tissue- and sample-types, namely genes involved in innate immunity, digestion, metabolism, and cellular homeostatic pathways. Notably, all

but one of these reports (Bost *et al.* 2017) examined the transcriptomic effects of microbiota elimination solely in female flies. The extent to which microbial effects on global gene expression differ between sexes therefore remains minimally explored.

In this study, we sought to identify molecular factors underlying the microbial impacts on *Drosophila* neural function, and behavioral and physiological traits. To accomplish our objective, we screened for microbiota-induced transcriptional changes in the adult *Drosophila* head. The fly head predominantly comprises the brain, eyes, antennae, and head fat body, a major metabolic/endocrine/immune tissue. We hypothesized that the brain and fat body in particular play key, undiscovered roles in mediating microbial impacts on fly behavior and metabolism. Moreover, the head is an anatomical structure that has not been investigated previously in the context of microbiota-dependent global gene expression. We therefore conducted a two-step screen for microbiota-induced transcriptional changes in the head using RNA-seq followed by nanoString technology, and identified both sex-general and sex-specific microbe-sensitive genes. These genes are broadly classifiable by shared functional categories, including innate immunity, neural activity, oxidative stress responses, and metabolism. At the cellular and physiological levels, the microbiota's influence on some of these biological processes in the host insect is well established in the literature. Thus, some of the genes identified in this study are promising candidates that may connect specific known aspects of host function to the microbiota. In other cases, we predict that the genes identified here will lead to the discovery of new microbe-dependent facets of host biology.

MATERIALS AND METHODS

Fly stocks and general rearing

The primary stock used in this study was the Top Banana *Drosophila melanogaster* strain, generously provided by Michael Dickinson's lab at the California Institute of Technology. This population was caught at the Top Banana fruit stand in Seattle, Washington USA (coordinates 47°40'37.0"N 122°22'37.9"W) in September 2013. Stock cultures were reared on yeast-cornmeal-molasses food of the following recipe: vol/vol or wt/vol; 8.5% molasses, 7% cornmeal, 1.1% brewer's yeast, 0.86% agar (MoorAgar), supplemented with 0.27% propionic acid (Sigma) and 13.5 mL/L of 20% methylparaben (Sigma) dissolved in 95-100% ethanol at room temperature, ~25°C. All experimental cultures, derived as described in the following section, were reared in autoclaved bottles and vials of the above diet with ~0.25g autoclaved yeast granules added topically, and maintained in a light-, temperature-, and humidity-controlled incubator on a 12:12 hour light:dark cycle (lights on 8:00am, lights off 8:00pm) at 22-23°C, 70% humidity.

RNA-seq and nanoString experiments were conducted with Top Banana flies harboring a *Wolbachia* infection derived from the wild-caught population; RT-qPCR experiments were conducted with Top Banana cultures cleared of *Wolbachia* by rearing for two generations on indicated diet with 0.25 g/L tetracycline (Sigma), followed by >10 generations on regular diet. The w^{1118} stock used in this study is *Wolbachia*-free. The *Wolbachia* status of stocks was determined by PCR amplification of the *Wolbachia* surface protein (*wsp*) gene from DNA extracts prepared from 5-10 adult flies (Table S2).

Generation of conventional and germ-free *Drosophila* cultures

Synchronous cultures were prepared by collecting embryos laid on apple juice agar plates within a 4-hour time window. Embryos were then collected in embryo wash buffer (1X

177 solution, vol/vol or wt/vol in Milli-Q H_2O: 2% Triton X-100, 7% NaCl). To generate
178 conventional cultures, approximately 150-200 embryos were transferred directly to food
179 bottles by pipette. To generate germ-free and gnotobiotic cultures, embryos were treated
180 for 2 minutes with 50% sodium hypochlorite, rinsed twice in 70% ethanol and twice in
181 sterile Milli-Q H_2O. Approximately 150-200 embryos were then transferred by pipette to
182 either: i) sterile food bottles, for germ-free cultures, or ii) sterile food bottles inoculated
183 with ~10^7 CFU bacterial liquid culture for gnotobiotic fly cultures (see below). All
184 manipulations were performed inside a sterile laminar flow cabinet. Bottle cultures were
185 then incubated under conditions described above and developed to adulthood. Adult flies
186 were collected under light CO_2 anesthesia on a sterilized pad 0-24 hours post-eclosion
187 and transferred to sterile food vials with ~0.05 g autoclaved yeast granules added, 10-20
188 flies of the same sex per vial. For gnotobiotic flies, vials were inoculated with ~10^7 CFU
189 bacterial liquid culture prior to collecting adult flies (see below). Microbial status of
190 gnotobiotic cultures and sterility of germ-free cultures were routinely checked by
191 homogenizing 5-10 flies in phosphate buffer saline (PBS) and plating on bacterial growth
192 media (see below). Flies were aged in vials to 5-6 days post-eclosion for use in all
193 experiments reported here.

194 **Isolation and identification of gut bacteria from conventional Top Banana flies**

195 Bacteria were collected from either surface-sterilized whole flies or dissected guts from
196 conventional Top Banana male and female flies. For surface sterilization, flies were
197 washed once in 1 mL 10% sodium hypochlorite, once in 1 mL 70% ethanol, and three
198 times in PBS. Guts (proventriculus to hindgut, excluding crop and Malpighian tubules)
199 from 10-20 individual animals were dissected in PBS. Surface sterilized animals or

dissected guts were then homogenized in 100 μL PBS and five 10-fold serial dilutions of homogenate were prepared in PBS. Dilutions were plated on MRS (Difco) and Ace (wt/vol or vol/vol: 0.8% yeast extract, 1.5% peptone, 1% dextrose, 1.5% agar, 0.3% acetic acid, 0.5% ethanol) agar plates. MRS plates were sealed with parafilm and incubated at 37° for ~48 hours, and Ace plates were incubated at 30° for ~72 hours. Individual colonies of characteristic morphology were then streaked for isolation on the relevant medium. For taxon identification, the 8F and 1492R primers were used to amplify sequence from the 16S rRNA gene of each isolate (Eden *et al.* 1991; Table S2). Amplicon DNA was purified and Sanger sequenced with both the forward and reverse primers. Sequencing results were then searched for highly similar sequences in both the SILVA database using the SINA alignment service (Pruesse *et al.* 2012) and the NCBI nr/nt database via blastn (Altschul *et al.* 1990; Morgulis *et al.* 2008; Camacho *et al.* 2009). Bacterial taxonomies were assigned based on >97% sequence homology (see Table S1). The 16S rRNA sequence for isolate A22 bore >97% similarity with multiple *Acetobacter* strains of different species.

Generation of gnotobiotic *Drosophila* cultures

Overnight cultures of *L. plantarum* L32 and *L. brevis* L28 were grown in MRS broth statically at 37°C, and *A. pasteurianus* A40, and *Acetobacter sp.* A22 were grown in MRS broth at 30°C with shaking. Cells were pelleted, washed twice with PBS, and resuspended in 500μL PBS. Cells were then diluted to a suspension of OD_{600}=1 and five serial 10-fold dilutions were prepared. Defined volumes of each dilution were then plated on MRS or Ace agar and incubated at 37°C or 30°C for 48 hours. Colonies were then counted

manually and CFU/mL constants for OD_{600}=1 cell suspensions were calculated for each bacterial isolate (Table S1).

To inoculate sterile food bottles with bacterial cell suspensions for gnotobiotic fly culture generation, overnight cultures were grown and washed in PBS as described. Cell suspensions were then diluted to a concentration of ~10^7 cells in 150 μL volume (calculations based on obtained CFU/mL constants). The entire 150 μL was then pipetted directly onto the surface of sterile food bottles ~4-5 hours prior to addition of dechorionated, sterilized fly embryos (as described above). A cell suspension volume of 50 μL was used to inoculate vials with ~10^7 CFU of each bacteria prior to transfer of newly-eclosed adult flies, as described above.

RNA extraction and sequencing

For each biological replicate, heads from 50 adult male flies were dissected in PBS and transferred to tubes containing 400 μL Trizol Reagent (Invitrogen) and ~50 μL zirconia beads (1.0 mm). Head tissues were then immediately homogenized with a mini-beadbeater (Biospec Products) for three 30 second pulses, with 10-15 second pauses in between. RNA was then extracted using the PureLink RNA Mini Kit (Life Technologies) exactly following the kit protocol. Paired-end sequencing was conducted at the University of Southern California Molecular Genomics Core facility on an Illumina NextSeq instrument. The following read-pair counts for each sample were obtained: CV.1 53638762 read pairs; CV.2 44963314 read pairs; CV.3 45583437 read pairs; GF.1 42930013 read pairs; GF.2 46242287 read pairs; GF.3 51418597 read pairs. Raw read data were provided to C. Kingsford's group at Carnegie Mellon for analyses described below.

RNA-Seq data analysis

Transcript expression quantification estimates were generated using Salmon (Patro *et al.* 2017) version 0.6.0 using Release-84 of Ensembl's *Drosophila melanogaster* cDNA library as a reference index. Each paired end file set was processed individually using the variational Bayesian EM algorithm using 30 bootstrapped samples to compute abundance estimates. Transcript-level expression values were converted to gene-level expression using the biomaRt package of the Bioconductor software project (Durinck *et al.* 2009). Differential expression was then processed on the gene level using DESeq2 (Love *et al.* 2014) on both the full dataset and for each leave-one-out analysis. A gene was defined as "differentially expressed" if it exhibited nonzero expression in both sample sets and had a Benjamini-Hochberg adjusted p-value of <0.1 (see "Results"). Tissue enrichment and Gene Ontology (GO) term enrichment analyses were conducted using the FlyMine integrated database (Lyne *et al.* 2007).

NanoString nCounter analysis

Custom barcoded nanoString probes for genes of interest and housekeeping genes were designed by nanoString Technologies (File S6, sheet "Codeset"). For each biological replicate, heads from 5-10 adult flies (males and virgin females) were removed in PBS and immediately homogenized in Trizol by bead beating, as described in "RNA extraction and sequencing". RNA was then extracted using the Direct-zol RNA MiniPrep Kit (Zymo Research) exactly following the kit protocol. Hybridization with probe set on nCounter chips was performed following manufacturer's protocol with 70-100ng RNA per sample. Quality assessment of raw data and normalization were performed using nSolver Analysis Software 3.0. Of note, one sample each of the following conditions were flagged by

nSolver QC analysis due to positive control failure: CV males, GF females, GNO females (File S6, sheet: "Raw mRNA counts"). Data from these samples were therefore excluded from subsequent analysis. For genes of interest, mRNA counts were normalized to the geometric mean of count values from each sample for the following four housekeeping genes: *14-3-3ε*, *Su(Tpl)*, *cyp1*, *cyp33*. As our aim was to identify strongly microbiota-responsive genes in either male or female *Drosophila* heads, and not to examine the impact of sex on expression of our genes of interest, we examined expression data across CV, GF, and GNO conditions independently for each sex. Additionally, we refined this analysis by examining the genes with the highest magnitude average expression difference, arbitrarily defined as genes with l2fc>0.4 or l2fc<-0.4 in either the GF *vs.* CV or GF *vs.* GNO comparisons. For each gene examined, we first conducted Levene's test for equal variance and the Shapiro-Wilk normality test. Data that met the parametric assumptions according to results of these tests ($p<0.05$) were analyzed via one-way analysis of variance (ANOVA), followed by the Tukey HSD *post hoc* test to compare means between groups. Data that did not meet assumptions of equal variance and a normal distribution were analyzed via the Kruskal-Wallis test with Dunn's multiple comparisons test.

RT-qPCR analysis

For RT-qPCR experiments, heads were dissected and RNA extracted exactly as described in "NanoString nCounter Analysis". Each biological replicate represented RNA extracted from heads of 10 individual animals. Pure quality RNA ($A_{260nm/280nm}$ ~2.1, 400-500ng) was used as template for cDNA synthesis using the qScript cDNA synthesis kit (QuantaBio). Product from cDNA synthesis reactions was used for qPCR with the

PerfeCTa SYBR Green Supermix (QuantaBio) in an Applied Biosystems 7300 Real Time PCR System instrument. Primer efficiencies were calculated using LinregPCR software (Ramakers *et al.* 2003; Table S2). Data were normalized to either *Rpl32* or the average of *Rpl32* and *14-3-3ε* housekeeping genes, as indicated, and expression fold changes relative to GF were calculated using the $2^{-\Delta\Delta Ct}$ method. Normalized data were analyzed via one-way analysis of variance (ANOVA), followed by the Tukey HSD *post hoc* test to compare means between groups. Sequences for all primers used are listed in Table S2.

Statistical analysis

The statistical methods and tests employed for each set of experiments are indicated in the relevant Methods sections and are reported in the figure legends. Analyses and figure generation were performed with R version 3.4.0 and GraphPad Prism 7 software.

Data and reagent availability

All fly stocks and bacterial isolates used in this study are available upon request. All supplemental figures and files have been deposited in figshare. Raw RNA-seq data have been deposited in the National Center for Biotechnology Information Sequence Read Archive, BioProject accession number PRJNA514099. File S1 contains gene-mapped RNA-Seq read counts (sheet “Counts”); weighted average of transcript length for each gene based on the number of reads for each transcript (sheet “Transcript length”); abundance representing transcripts per million (TPM; sheet “TPM”); and initial statistical analysis of results, as described in corresponding Results section, with genes ordered by ascending Benjamini-Hochberg adjusted p-values (sheet “Statistical comparison”). File S2 summarizes descriptive results of each leave-out analysis conducted. File S3 lists genes identified as significantly differentially expressed (p-adj<0.1) in ≥2 leave-out

314 analyses, i.e. genes with highest degree of statistical support. File S4 lists the 343 genes
315 identified as significantly differentially expressed (p-adj<0.1) in ≥1 leave-out analysis; p-
316 adj values for each analysis conducted are provided for each gene. File S5 contains
317 results from tissue and GO-term enrichment analyses conducted on the genes listed in
318 File S4. File S6 contains the codeset used in nanoString experiments, raw and normalized
319 nCounter mRNA count values, and differential expression calculated as $\log_2$ fold change
320 values. Table S1 lists the four bacteria isolated from CV Top Banana flies employed for
321 GNO microbial conditions, their closest taxonomic assignment based on 16S rRNA gene
322 sequence data, and the empirically determined CFU/mL constants utilized to equilibrate
323 planktonic bacterial cultures for inoculation of GNO fly cultures (see Materials and
324 Methods). Table S2 lists and provides pertinent information for all primers used in this
325 study.

326 **RESULTS**

327 **Transcriptomic comparison of the heads of microbiota-associated and germ-free**
328 ***Drosophila***

329 To investigate whether and how microbial symbionts that primarily occupy the gut
330 lumen alter gene expression in head tissues, we compared the head transcriptomes of
331 CV adult male *Drosophila* to their sterile GF siblings via RNA-seq. Specifically, we
332 collected and sequenced RNA from the heads of fifty animals per replicate, and examined
333 three, independently reared biological replicates of each microbial condition. We initially
334 analyzed the data by read normalization to transcripts per million (TPM) values, and
335 comparison via the Wald Chi-squared test, with a Benjamini-Hochberg false discovery
336 rate (FDR)-adjusted p-value cutoff of 0.1 serving as our preliminary statistical significance

criterion. In this primary analysis, out of 13180 genes with nonzero total read counts, we identified 50 genes as being significantly differentially expressed in the heads of GF *vs.* CV male *Drosophila* (File S1; Sheet "Statistical comparison", highlighted cells). These 50 genes included 14 genes that showed elevated transcript levels in GF fly heads [$\log_2$ fold change (l2fc)>0], and 36 genes that showed decreased transcript levels relative to CV (l2fc<0). Overall, the magnitude of expression differences we observed were modest: 6 of the 14 elevated genes had l2fc>0.5, and 19 of the 36 decreased genes had l2fc<-0.5. Notably, we did not observe any gene expression changes greater than two-fold (l2fc>1 or l2fc<-1; Figure 1A, File S1).

Closer, visual inspection of the data revealed a considerable degree of gene-to-gene variability in normalized expression levels across replicates for the CV condition, with each replicate displaying a distinct expression pattern (Figure S1) as revealed by principal component analyses (PCA; Figure S2). Moreover, one replicate of the GF condition (designated GF replicate 3; Figure 1B, Figure S1) showed a global expression pattern markedly distinct from the other two replicates, and our PCA supported this replicate as an extreme outlier (Figure S2). The GF outlier immediately suggested the possibility that this this sample was derived from microbe-contaminated fly cultures. However, matched-sample 16S rRNA gene sequence profiling of the dissected guts of all flies used for our RNA-seq samples does not suggest contamination of any GF samples (data available upon request). Similarly, we speculated that the considerable transcriptomic variability among CV flies could be attributable to diversity in the abundance and composition of their un-manipulated microbial communities. This possibility is consistent with the known features of the laboratory *Drosophila* gut

360 microbiota, namely its inconstancy, transiency under certain rearing conditions,
361 dependence on the diet substrate, and high inter-generational and inter-individual
362 compositional variability (Wong *et al.* 2011, 2013, 2015; Blum *et al.* 2013; Broderick *et al.*
363 2014; Chaston *et al.* 2016; Elya *et al.* 2016; Early *et al.* 2017; Jacqueline *et al.* 2017). Our
364 16S rDNA profiling of the gut microbiota of the CV samples did not reveal any substantive
365 differences, but we cannot rule out the possibility that small populations of distinct
366 microbial cohorts influenced the head transcriptional profiles of these samples. Moreover,
367 the Top Banana fly stock utilized in our study harbors a *Wolbachia* infection presumably
368 derived from the original wild-caught population (see Materials and Methods). Variable
369 titers and cohorts of *Wolbachia* endosymbionts may constitute additional sources of
370 variability in the transcriptomic profiles of GF and CV samples.

371 To address the potential for type II error (i.e. false-negatives) in our initial analysis,
372 resulting from the substantial inter-replicate variability of our CV samples and the GF
373 outlier replicate, we conducted a leave-one-out analysis for each replicate in the
374 experiment, resulting in seven total result outputs. Importantly, the number of genes that
375 met our significance criterion (FDR adjusted p-value<0.1) varied dramatically depending
376 on which replicate was omitted from the analysis: for example, omission of GF replicate
377 2 from analysis of the dataset yielded 8 total differentially expressed genes, while
378 omission of CV replicate 1 yielded 177 significantly different genes (File S2, File S3, File
379 S4). The number of significant genes identified in each of the six analyses were generally
380 consistent with the grouping and distribution patterns observed in our PCA results; for
381 example, omission of GF replicate 3, the extreme outlier GF sample, increased the
382 number of significantly different genes to 162, compared to the other analyses.

383 From this approach, we obtained a list of 343 unique candidate genes (Figure 1B;
384 File S4 for FDR-adjusted p-value outputs from each analysis) that were significantly
385 altered in ≥one of our seven analyses (the analysis of the full dataset and each of the six
386 leave-one-out analyses). Notably, most of these genes were significantly differentially
387 expressed in only one out of seven of the analyses; 66 of the 343 candidates were
388 significant in ≥two analyses, and only six genes were differentially expressed in all seven
389 analyses (File S3, File S4). To begin to prioritize individual genes for subsequent analysis,
390 we characterized the functional biological properties of these 343 candidate genes.
391 Examination of tissue-specific expression, based on Affy Calls ("Up" vs. "Down" relative
392 to whole-body expression levels) derived from the FlyAtlas expression database
393 (Chintapalli *et al.* 2007), revealed that of the 343 genes, 176 genes are highly expressed
394 in the head, 107 genes are highly expressed in the fat body (importantly, the FlyAtlas
395 dataset reflects abdominal fat body expression, and not expression in the head-localized
396 fat body), 99 genes are highly expressed in the eye, and 70 genes are highly expressed
397 in the brain (Figure S3, File S5). Gene Ontology (GO) terms enriched (Holm-Bonferroni
398 adjusted p-value<0.05) in the 248/343 genes down-regulated in GF fly heads primarily
399 encompassed functional categories related to immune function and antimicrobial
400 responses (Figure 1C, File S5). GO terms related to amino acid and fatty acid biosynthetic
401 processes were also enriched within the down-regulated genes (File S5). Enriched GO
402 terms in the 95 genes transcriptionally elevated in GF, compared to CV, heads were
403 dominated by responses to organic chemical stimuli (Figure 1D, File S5).

404 In summary, our RNA-Seq examination of the microbiota's impact on the head
405 transcriptome of adult male *Drosophila* revealed a relatively small number of low-

406 magnitude gene expression changes resulting from elimination of microbes. The limited
407 number of robustly differentially expressed genes could in part be attributed to a high
408 degree of inter-replicate variability in the expression profiles of each CV fly sample. This
409 explanation is supported by the substantial increase in the number of candidate genes
410 that met our statistical cutoff depending on the exclusion of certain replicates from our
411 analysis. Nevertheless, this initial experiment provided evidence of genes expressed in
412 the adult *Drosophila* head that were transcriptionally responsive to sterile rearing
413 conditions. This list of candidates formed the basis for our subsequent analyses.

414 **Microbial impact on immune, metabolic, and oxidative stress response gene**
415 **expression in *Drosophila* heads identified by secondary nanoString analysis**

416 Our overall goal was to identify specific genes exhibiting microbiota-dependent
417 expression changes in the *Drosophila* head. We reasoned that these gene identities
418 would inform hypotheses about the molecular bases of known and novel microbiota-
419 regulated host physiological traits, and would thereby guide subsequent mechanistic
420 investigations into the roles of these genes within a microbial context. Given this objective,
421 and the modest expression changes and high variance observed in our RNA-Seq results,
422 we conducted a second round of screening on a subset of candidate genes utilizing
423 nanoString nCounter technology. The nCounter system employs a probe-capture
424 barcode platform for direct, digital measurement of mRNA transcripts. These features of
425 the platform enabled us to validate putative gene expression changes via an alternative
426 methodology that obviates the potential bias introduced by the enzymatic processing
427 steps of library preparation. Expression differences consistently observed via both
428 techniques (indirect, relative, genome-wide *vs.* direct, target-based) would therefore have

429 considerable empirical support as microbiota-affected genes. Additionally, the high-
430 sensitivity of the automated nCounter technology afforded the potential to reveal evidence
431 of type I ("false positive") errors from our RNA-Seq analyses.

432 For our nCounter study we generated probes for 92 genes chosen from the 343
433 identified candidates (File S5). Our selection criteria included statistical robustness (i.e.
434 the genes with comparatively high support shared by multiple analyses of RNA-seq
435 results, described above) and magnitude of putative differential expression. Further, given
436 our aim of identifying molecular factors involved in microbiota-modulated neural function
437 and behavioral traits, we also prioritized and selected genes implicated in fly behaviors
438 and with known roles in the brain.

439 The multiplexed nature of the nCounter technology also allowed us to test the
440 additional variables of a GNO microbial condition and host sex. Which bacterial taxa or
441 cohorts of taxa, and what functional attributes of microbial populations contribute to a
442 microbiota-dependent host trait, are crucial questions inherent to all studies of the
443 mechanistic basis of host-microbe symbioses. Moreover, as indicated above, the
444 microbiota of CV-reared laboratory *Drosophila* is inconstant and compositionally variable
445 within and across fly populations. We generated flies with a standardized, GNO
446 microbiota for examination in our nanoString experiment, to: (i) begin to identify the
447 specific bacteria responsible for observed gene expression changes, facilitating future
448 explorations of microbial mechanisms responsible for identified host traits, and (ii) reduce
449 inter-sample variability in microbiota composition as a potential factor contributing to
450 variable gene expression. Our GNO flies were reared from embryo to 5-6-day adulthood
451 in continuous polyassociation with four bacterial strains isolated from CV Top Banana

cultures: *Lactobacillus brevis*, *Lactobacillus plantarum*, *Acetobacter sp.* (unresolved species/strain-level taxonomic assignment; see Materials and Methods), and *Acetobacter pasteurianus* (Table S1). These strains represent the bacterial taxa most commonly associated with laboratory *Drosophila* cultures (Broderick and Lemaitre 2012; Douglas 2018). In addition, as noted above, reports surveying microbiota effects on *Drosophila* gene expression are dominated by studies focused solely on female flies. Our RNA-seq experiment used only male heads, and we speculated that some putative expression changes might be male-specific, or exhibit a different expression pattern in females. We therefore included both male and female flies in our nanoString study to determine whether any tested candidate genes responded differently to host microbial condition as a function of the animal's sex. In all, our final experimental matrix for nanoString analysis included heads from male and female CV, GF, and GNO flies.

As with our RNA-seq results, most gene expression changes observed in the nanoString study were modest. More specifically, we observed relatively few genes with high-magnitude expression differences in both GF *vs.* CV and GF *vs.* GNO comparisons (Figure 2A, 2B). For both males and females, more genes were strongly down-regulated in the heads of GF compared to microbe-associated flies than were upregulated in GF heads (Figure 2A, 2B, File S6). Notably, genes down-regulated in GF male heads were generally more strongly repressed relative to CV flies than relative to GNO flies (Figure 2A; the magnitude of the blue bars exceeds that of the gray bars). Conversely, in females down-regulated genes were suppressed to a greater magnitude in relation to the GNO condition than *vs.* CV animals (Figure 2B; the magnitude of the gray bars exceeds that of the blue bars). These trends, in addition to specific gene expression differences (see

below), suggest that host sex impacts head-localized transcriptional responses to the microbial environment, consistent with previous observations in the gut (Bost *et al.* 2017).

The fundamental objective of our secondary nanoString screen was to identify individual genes with robust, microbiota-affected expression in either male or female *Drosophila* heads. We therefore examined the genes with the greatest magnitude of change in either the GF/CV or GF/GNO comparison (arbitrarily defined as l2fc>0.4 or l2fc<-0.4; 32 genes for males, 40 genes for females: File S6, sheet “Log2FoldChange”), and statistically compared the effect of the three microbial conditions on these genes separately for each sex (see Materials and Methods; File S6, sheet “Log2FoldChange”). Using this approach, we identified 19 genes that were differentially expressed ($p<0.05$) in ≥one of the four comparisons: (i) male GF *vs.* CV, (ii) male GF *vs.* GNO, (iii) female GF *vs.* CV, (iv) female GF *vs.* GNO. For both males and females, a greater number of significant expression changes were observed in the GF to GNO comparison (11 genes for males, 13 genes for females) than in the GF to CV comparison (4 genes for males, 1 gene for females). Six of the 19 genes were differentially expressed in both males and females, and for all six the directions of change between microbial conditions were the same in both sexes. Notably, while all 92 of the genes in our codeset were selected from our leave-out analyses-derived 343 RNA-seq candidates, most of the genes (11/19) with significant expression differences in our nanoString results were significantly different in only one leave-out analysis out of our seven total comparisons (File S6, sheet “RNAseq comparison”).

The genes down-regulated in the heads of both male and female GF flies relative to microbe-associated flies were dominated by immune genes (Figure 3), specifically, two

498 extracellular peptidoglycan recognition proteins (PGRPs), *PGRP-SB1* and *PGRP-SD*,
499 and eight antimicrobial peptides (AMPS). As previously indicated, more significant
500 expression differences were observed in the GF *vs.* GNO comparison than in the GF *vs.*
501 CV comparison. The PGRP genes were significantly reduced exclusively in GF female
502 compared to GNO female heads (Figure 3A, B). Four AMPs were significantly reduced
503 only in male GF *vs.* male GNO heads (*DptA*, *DptB*, *Dro*, and *edin*; Figure 3F, G, H, I), and
504 the remaining four were reduced in both male and female GF *vs.* GNO and/or CV heads
505 (*AttA*, *AttB*, *AttC*, and *Mtk*; Figure 3C, D, E, J). While only these ten immune genes
506 achieved statistical significance in ≥one comparison, the observable trend of elevated
507 transcript levels in microbe-associated flies compared to sterile flies was consistent
508 between sexes for all immune genes assayed (Figure 3A-J; File S6), including those with
509 no statistically significant differences in any comparison (Figure S4). As noted in our RNA-
510 seq results, the failure to achieve statistical significance likely reflects substantial
511 transcript count variability among the microbe-associated conditions, as compared to the
512 more consistent, low transcript levels detected in GF samples.

513 Seven additional genes that exhibited statistically significant responses to host
514 microbial condition in ≥one comparison had known or predicted functions in aging,
515 oxidative stress resistance and general detoxification, and metabolism. Interestingly,
516 each of these genes responded to host microbial condition in a sex specific manner. The
517 cytochrome P450 gene *cyp6a17* was elevated in GF male, but not female heads (Figure
518 4A), while the small mitochondrial heat shock protein *Hsp22* and the serine
519 hydroxymethyl transferase *Shmt*/*CG3011* were upregulated exclusively in GF females
520 (Figure 4B, C). Conversely, four experimentally uncharacterized genes with predicted

metabolic functions (inferred from sequence homology and predicted domain architecture) were all down-regulated in GF *vs.* microbe-associated heads, again with gene-by-gene sex differences. *CG10960* (functionally annotated as involved in sugar transport), *CG4757* (inferred carboxylesterase activity), and *CG3036* (a putative anion transporter) were all suppressed in GF females, but not males (Figure 4D, E, F). Another uncharacterized gene, *CG10512*, with predicted malate dehydrogenase activity, was more highly expressed in both male and female GNO, but not CV, heads relative to GF heads (Figure 4G). An uncharacterized gene with no domain-based annotation and no experimentally determined function, *CG7296*, was also elevated in male GNO *vs.* GF heads (Figure 4H).

Individual bacteria alter *Arc1* expression in the adult *Drosophila* head

Drosophila activity-regulated cytoskeleton associated protein 1 (*Arc1*) also responded strongly to microbial condition in Top Banana heads. Specifically, we found *Arc1* was elevated approximately twofold in GF male heads compared to both CV and GNO males, and more moderately elevated in GF female compared to GNO female heads (Figure 5A). *Arc1* encodes an ~29 kDa protein evolutionarily related to retrotransposon Group-specific antigen (Gag)-like proteins. Arc1 proteins multimerize to form capsid-like structures that facilitate intercellular mRNA transfer via exosomes at the larval neuromuscular junction (Ashley *et al.* 2018), and this mechanism is conserved in mammalian Arc/Arg3.1 (Pastuzyn *et al.* 2018). Moreover, *Drosophila Arc1* function in a particular subset of neurons in the larval brain promotes systemic metabolic homeostasis and prevents hyperlipidemia (Mosher *et al.* 2015). These metabolic consequences of *Arc1* genetic perturbation resemble those observed in flies grown under sterile conditions

544 (Shin *et al.* 2011; Storelli *et al.* 2011; Ridley *et al.* 2012; Wong *et al.* 2014; Newell and
545 Douglas 2014; Dobson *et al.* 2015; Huang and Douglas 2015; Chaston *et al.* 2016; Kim
546 *et al.* 2017; Judd *et al.* 2018; Kamareddine *et al.* 2018).

547 Numerous published reports have demonstrated that association with individual
548 bacterial species and strains is sufficient to recapitulate CV- and/or polyassociated GNO-
549 like host molecular and physiological traits. In some examples, a mixed microbial
550 community's impacts on the host are attributable to the activities of a specific, single
551 community member, while in other examples monoassociation with any bacterial
552 commensal is sufficient to recapitulate the effects of the polymicrobial community (Shin
553 *et al.* 2011; Storelli *et al.* 2011; Newell and Douglas 2014; Elya *et al.* 2016; Daisley *et al.*
554 2017; Téfit and Leulier 2017; Fischer *et al.* 2017; Kim *et al.* 2017; Leitão-Gonçalves *et al.*
555 2017; Matos *et al.* 2017; Obadia *et al.* 2017; Wong *et al.* 2017; Judd *et al.* 2018; Sannino
556 *et al.* 2018). We asked whether suppression of *Arc1* expression in the heads of CV and
557 GNO relative to GF flies is due to association with specific bacterial taxa, or represents a
558 generalized response to the presence of microbes. Given that our standardized,
559 simplified, four bacteria GNO microbial community was sufficient to restore CV level
560 expression to male flies (Figure 5A), we tested this hypothesis by focusing on these four
561 species. Specifically, we reared Top Banana *Drosophila* cultures in monoassociation with
562 each of the four bacteria comprising our GNO condition, and measured relative *Arc1*
563 expression in the heads of adult male flies via RT-qPCR. Consistent with our nanoString
564 results (Figure 5A), polyassociation with the four-species GNO community resulted in
565 significantly decreased *Arc1* expression relative to GF heads, and monoassociation with
566 *Acetobacter sp.*, *A. pasteurianus*, or *L. brevis* was sufficient to downregulate *Arc1* to the

same extent (Figure 5B). *Arc1* expression was also reduced in *L. plantarum* monoassociated heads relative to GF, but the difference was not significant (Figure 5B). These results suggest that the transcriptional downregulation of *Arc1* in the heads of GNO flies is primarily attributable to interactions with the two *Acetobacter* isolates and *L. brevis,* but not to the presence of *L. plantarum*. As each of these bacteria alone recapitulates the transcriptional difference induced by the polymicrobial community, this microbiota-dependent *Arc1* suppression may reflect a common host response to the presence of certain prokaryotic organisms.

Microbiota-dependent gene expression changes in the *Drosophila* head are sensitive to host genetic background

Host genotype substantially affects microbial impacts on certain host traits in *Drosophila* (Brummel *et al.* 2004; Broderick *et al.* 2014; Dobson *et al.* 2015; Chaston *et al.* 2016; Obata *et al.* 2018). Because our gene expression screen employed the recently established wild-type *Drosophila* stock Top Banana, we asked whether select microbiota-dependent gene expression changes are unique to this genetic background, or represent more common microbial effects. To address this, we used RT-qPCR to measure relative transcript levels of three strongly microbiota-responsive genes from our screen (*Arc1*, *Hsp22*, and *DptA*) in the heads of CV-, GF-, and GNO-reared w^{1118} flies, a widely-utilized laboratory *Drosophila* wild-type stock.

As in Top Banana flies (Figure 5A), *Arc1* was expressed at lower levels in the heads of bacteria-associated CV and GNO male w^{1118} *Drosophila* relative to GF males (Figure 6A), however there was no difference in expression across conditions in w^{1118} female heads (Figure 6B). Similarly, *Hsp22*, which exhibited female-specific elevation in

590 the heads of GF Top Banana flies compared to GNO females (Figure 4B) was unaffected
591 in w^{1118}, showing roughly equivalent expression across all conditions (Figure 6C, D). The
592 representative AMP gene *DptA*, showed a high fold expression change in the heads of
593 CV and GNO w^{1118} males and females compared to their sterile siblings like in Top
594 Banana (Figure 3F, 6E,F). However, the fold elevation relative to GF was highest and
595 statistically significant only in GNO females (Figure 6E, F). Nonetheless, this observation
596 was consistent with the robust AMP expression changes observed in our preceding
597 screen results. These data provide additional support for the hypotheses that host sex
598 and genotype are important factors governing *Drosophila*'s transcriptional response to
599 association with the microbiota. More specifically, these data suggest that some gene
600 expression changes identified in our screen may be sensitive to host genomic
601 composition.

602 **DISCUSSION**

603 To date a variety of microbe-dependent physiological and behavioral traits have
604 been described in *Drosophila* (Broderick and Lemaitre 2012; Martino *et al.* 2017; Douglas
605 2018), however the host molecular processes underlying these effects are largely
606 unknown. In this study, we utilized RNA-seq and nanoString technology to uncover
607 microbiota-responsive genes in the adult *Drosophila* head. Our RNA-seq investigation of
608 gene expression changes induced by GF rearing revealed a high degree of inter-replicate
609 variability in global expression profiles, particularly among CV samples. Additionally, all
610 putative microbiota-dependent gene expression changes we identified, in both RNA-seq
611 and nanoString analyses, were of relatively low magnitude (Figure 1A, Figure 2). In
612 contrast, previous studies examining microbiota-induced gene expression identified

613 stronger transcriptional differences (i.e. ≥twofold) in the adult *Drosophila* gut (Broderick
614 *et al.* 2014; Guo *et al.* 2014; Elya *et al.* 2016; Petkau *et al.* 2017), in the embryos of sterile
615 parents (Elgart *et al.* 2016), in third instar larvae (Erkosar *et al.* 2017), and in whole flies
616 (Combe *et al.* 2014; Dobson *et al.* 2016; Bost *et al.* 2017). Excepting the proboscis
617 mouthparts and anterior-most foregut epithelium, the fly head largely consists of tissues
618 that are unlikely to directly interact with microbes. It may therefore follow that expression
619 changes in the head, while potentially no less functionally important, would be modest
620 compared to microbial impacts on gene expression in the gut. Moreover, the head
621 contains multiple highly distinct cell- and tissue-types (i.e. eyes, ocelli, brain, antennae,
622 tracheal tissue, fat body, etc). Large transcriptional changes in small cell numbers could
623 therefore be masked by sampling whole heads. Nevertheless, our study is the first
624 transcriptome-scale assay of microbiota-dependent gene expression in an adult fly
625 anatomical structure distinct from the gut. This study is also noteworthy as only the
626 second to examine microbial effects on gene expression in both male and female flies
627 (Bost *et al.* 2017), and our findings are congruent with that report. Specifically, in both our
628 RNA-seq results and those of Bost *et al.*, chemical response genes, including cytochrome
629 p450-encoding genes, were the dominant functional categories upregulated in GF male
630 animals, while immune genes constituted the major downregulated functional categories
631 in sterile males. Moreover, our nanoString experiment revealed multiple genes that were
632 microbiota-responsive only in one sex, consistent with the sex-specific transcriptome
633 profiles observed by Bost *et al.* These findings emphasize host sex as a critical variable
634 that affects molecular outcomes of host-microbe association in *Drosophila*.

Most of the genes revealed by our screen have known or inferred roles in immune function, metabolism, aging, and oxidative stress responses. Intriguingly, numerous published studies have provided evidence that commensal microbes also modulate each of these physiological processes; examples of these connections are described in the following sections. We propose that the genes identified here are promising candidates for future study of the functional molecular basis for known microbe-dependent *Drosophila* traits.

Microbiota-induced immune gene expression in the adult *Drosophila* head

Innate immune genes, namely AMPs and PGRPs, were suppressed in the heads of GF compared to microbiota-associated flies in both our RNA-seq and nanoString experiments (Figure 1A, 1C, Figure 3, Figure S4). Prior transcriptomic studies have also demonstrated microbiota-driven transcriptional induction of IMD pathway genes in the adult gut and in whole animals (Ryu *et al.* 2008; Buchon *et al.* 2009; Broderick *et al.* 2014; Guo *et al.* 2014; Clark *et al.* 2015; Sansone *et al.* 2015; Dobson *et al.* 2016; Petkau *et al.* 2017; Bost *et al.* 2017). Some reports suggest that this priming of intestinal immune function by the commensal microbiota is important to maintenance of gut homeostasis and protection against injury in young flies; as the animal ages, over-induction of immune responses contributes to immunosenescence, inflammation, deterioration of gut homeostasis and integrity, and ultimate mortality (Ryu *et al.* 2008; Buchon *et al.* 2009; Blum *et al.* 2013; Guo *et al.* 2014; Clark *et al.* 2015; Sansone *et al.* 2015; Li *et al.* 2016; Lindberg *et al.* 2018). Whether microbiota-induced IMD activity in adult organs and tissues other than the gut also contributes to these phenomena is not known, and our data suggest this question warrants investigation. Importantly, the differences in immune

gene activation we observed may reflect transcriptional changes in head cell types that most directly contact microbes, such as the labellum and anterior-most foregut, or the head fat body. While less is known about immune activity in the head fat body, the abdominal fat body is the principle site of humoral immune induction during pathogenic infection (Ferrandon *et al.* 2007). Interestingly, *DptB* (one of the AMPs that emerged from our screen) was recently shown to be highly upregulated in adult male heads following long-term-memory-inducing behavioral training paradigms, and reciprocally its expression in the head fat body was required for the formation of long-term memories (Barajas-azpeleta *et al.* 2018). Further examples in the literature also point to connections between immune pathway activity in head tissues and complex physiological and behavioral phenotypes. AMP transcription in the brain increases with age, and overexpression of IMD-regulated AMPs in the adult brain is sufficient to reduce longevity and geotactic proficiency; conversely, genetic dampening of immune activity in both the CNS and glia is sufficient to extend lifespan and climbing ability (Kounatidis *et al.* 2017). In *Drosophila* models of neurological pathology, such as Fragile X syndrome, glial activity and activation of innate immune function in the brain contributes to neurodegenerative phenotypes and disease symptoms (Cao *et al.* 2013; Petersen *et al.* 2013; O'Connor *et al.* 2017). Enteric infection by the pathogen *Pectinobacterium carotovora carotovora 15* also exacerbates the neurological and physiological deterioration of a Drosophila Alzheimer's disease model, in a manner involving hemocyte recruitment to the adult brain (Wu *et al.* 2017). Potential contributions of the commensal microbiota to these neuroimmunological processes and their impacts on animal health have not been thoroughly examined. Our observation of microbiota-induced innate immune responses

in the heads of adult flies may therefore have important implications for the similar physiological and life-history traits of flies.

Modulation of aging and stress response gene expression by the microbiota

Our screen also revealed microbiota-responsive genes with known roles in the oxidative stress response and in pro-longevity functions. Organisms must respond to reactive oxygen species (ROS)-induced oxidative injury throughout their lifespans. ROS are constantly generated as metabolic byproducts, and their accumulation and the consequent macromolecular damage over time is a major contributor to cellular senescence and normal aging (Johnson *et al.* 1996; Lin and Beal 2003; Ewald 2018). Accordingly, in flies and other model organisms, overexpression of genes that promote proteostasis–such as chaperones, proteasomal subunits, and reducing enzymes–curtails age-related ROS elevation and extends longevity (Lin and Beal 2003; Back *et al.* 2012; Wang *et al.* 2013). Three genes we found to have increased expression in GF fly heads in one or both screen steps, *Hsp22*, *NAD-dependent methylenetetrahydrofolate dehydrogenase-methenyltetrahydrofolate cyclohydrolase* (*Nmdmc*), and *Ecdysone-induced protein 71CD/methionine-S-sulfoxide reductase A* (*Eip71CD*/*MsrA*), all have been shown to enhance oxidative stress resistance and lifespan upon overexpression (Morrow *et al.* 2004; Roesijadi *et al.* 2007; Chung *et al.* 2010; Yu *et al.* 2015). *Hsp22* encodes a small mitochondrial chaperone that functions in the unfolded protein response, and was upregulated specifically in female GF flies in our nanoString study (Figure 4B). While it did not meet our criterion of statistical significance, *Nmdmc* was also highly upregulated, on average, in the heads of GF *vs.* GNO female flies in our nanoString results (File S6). This gene encodes a mitochondria-localized folate-dependent enzyme

704 involved in purine biosynthesis, and may extend lifespan by increasing mitochondrial DNA
705 copy number through an unknown mechanism (Yu *et al.* 2015). *Eip71CD*/*MsrA* was
706 among the most statistically robust genes elevated in GF *vs.* CV male heads in our RNA-
707 seq screen (Figure 1A, File S1, File S3), and was the most highly upregulated (though
708 not significantly different) gene in GF *vs*. GNO male heads assayed by nanoString (File
709 S6). *MsrA* is a repair enzyme that reduces methionine-S-sulfoxide (generated by
710 oxidative damage to proteins) to methionine, and its pan-neuronal overexpression is
711 sufficient for enhanced longevity (Chung *et al.* 2010). Another gene of interest with
712 putative detoxification and stress response function was *Cyp6a17*, which increased in GF
713 male, but not female, heads compared to both CV and GNO animals (Figure 4A).
714 *Cyp6a17* belongs to the cytochrome P450 gene family, which encodes a broad range of
715 enzymes that oxidize toxic compounds (Bergé *et al.* 1998). In flies, *Cyp6a17* is enriched
716 in the mushroom body of the adult brain where its activity downstream of cAMP-PKA
717 signaling modulates temperature preference behavior via an unknown mechanism (Kang
718 *et al.* 2011).

719 Higher basal expression of these pro-longevity genes in GF *Drosophila* may
720 contribute to the extended lifespan that has been observed for GF animals (Petkau *et al.*
721 2014; Clark *et al.* 2015; Fast *et al.* 2018; Iatsenko *et al.* 2018; Sannino *et al.* 2018).
722 Interestingly, the *Drosophila* microbiota also induce secreted ROS production by
723 enterocytes that limits the gut bacterial population size, preventing dysbiosis and
724 premature aging (Guo *et al.* 2014; Iatsenko *et al.* 2018). Thus, the complete absence of
725 microbes should result in less oxidative damage, and therefore less upregulation of
726 oxidative damage counteracting genes, at least in the gut. Induction of *Nmdmc*, *Eip71CD*,

and *Hsp22* in the heads of GF flies must therefore occur via signals independent of microbe-induced ROS.

Increased expression of these stress response genes under non-stress conditions may reflect a basal susceptibility to oxidative injury in adults lacking microbial symbionts. Consistent with this, upregulation of *Cyp6a17* (and other cytochrome P450 genes observed in our RNA-seq analysis; File S5) may suggest a greater burden on the animal to detoxify ingested compounds and metabolic byproducts in the absence of microbes that normally serve this function. The baseline stressed condition may be exacerbated upon encountering environmental stressors, suggesting an explanation for the greater susceptibility of GF flies to chemical oxidative challenge (Jones *et al.* 2015; Naudin *et al.* 2019). Microbiota-dependent ROS production has not been examined in tissues other than the gut, nor has there been past indication of the sex-specific effects that we observed. This suggests that there is much more to learn about the molecular mechanisms connecting the microbiota to host oxidative stress resistance and lifespan.

Novel microbiota-regulated metabolic genes

Many of the most profound effects of the microbiota on animal physiology occur via their metabolic activities. Bacterial commensals in the mammalian gut utilize host dietary polysaccharides as carbon sources, and many of the resultant secondary metabolites are absorbed by host tissues where they can have systemic effects on multiple organ systems (Cockburn and Koropatkin 2016; Daïen *et al.* 2017). Mice and rats raised GF or subjected to aggressive antibiotic regimens exhibit reduced body fat levels and a decreased basal metabolic rate (Smith *et al.* 2007). Removal of the *Drosophila* microbiota also drastically impacts host metabolic function. On nutrient-rich

diets, bacteria modulate the glucose content of the diet substrate influencing the metabolic and nutritional profiles of adults (Ridley *et al.* 2012; Newell and Douglas 2014; Newell *et al.* 2014; Wong *et al.* 2014; Dobson *et al.* 2015; Huang and Douglas 2015). Microbial enhancement of dietary nutritional availability, particularly B-vitamins, promotes the fitness of the host–assayed by parameters including developmental rate, adult mass, and fecundity–particularly on protein-limited diets (Storelli *et al.* 2011, 2017; Wong *et al.* 2014; Leitão-Gonçalves *et al.* 2017; Matos *et al.* 2017; Bing *et al.* 2018; Sannino *et al.* 2018). *Acetobacter*-derived acetic acid induces both insulin signaling and the IMD pathway to reduce hyperlipidemic phenotypes and promote fly development (Shin *et al.* 2011; Hang *et al.* 2014; Kamareddine *et al.* 2018). However, host mechanisms activated by microbial depletion that yield these metabolic perturbations are unknown.

Second to immune activity, the GO-terms enriched among the genes downregulated in GF flies in our RNA-seq study were dominated by metabolic functions, specifically amino acid and fatty acid biosynthetic processes (File S5). This observation raises the intriguing possibility that microbe-induced expression of certain genes found here enables flies to derive the optimal nutritional benefit from their diet substrate, and potentially from dead microbial cells themselves. We also identified, via nanoString, the serine hydroxyl-methyl transferase *Shmt* and five unstudied genes with domain-based predicted functions in metabolic activity, including sugar transport and enzymatic organic chemical modifiers, which responded in both directions to elimination of the microbiota, more frequently in females than in males (Figure 4). Substantial additional work is required to investigate the uncharacterized, putative metabolic *Drosophila* genes

identified here as potential nodes in the complex interplay between laboratory *Drosophila*, its bacterial symbionts, sex, diet, and nutrition.

Arc1 is the *Drosophila* homolog of mammalian activity-regulated cytoskeleton associated protein (ARC), a neuronal protein that is required for strengthening synaptic connections, dendrite maturation, and learning and memory formation in mice (Tzingounis and Nicoll 2006; Shepherd and Bear 2011). Genetic defects in *Arc*/*Arg3.1* have been connected to neurological disorders, including Alzheimer's disease and Fragile X syndrome (Park *et al.* 2008; Rudinskiy *et al.* 2012). Like its vertebrate homologues, Drosophila *Arc1* is expressed in the larval and adult brain and neuroendocrine cells (Mattaliano *et al.* 2007; Mosher *et al.* 2015). Interestingly, flies null for *Arc1* exhibit metabolic defects, specifically increased fat stores and dysregulation of central carbon metabolism (Mosher *et al.* 2015), as well as enhanced starvation resistance (Mattaliano *et al.* 2007). *Drosophila Arc1* encodes a retroviral GAG-like protein which multimerizes to form capsid-like structures. These structures mediate trans-synaptic, vesicular transfer of *Arc1* mRNA and other mRNAs at the larval neuromuscular junction (Ashley *et al.* 2018), a mechanism conserved in the vertebrate protein (Pastuzyn *et al.* 2018). Connections between the molecular mechanism of Arc1 activity and its functions in neuronal activity and metabolism are unknown. Interestingly, in addition to our identification of *Arc1* as a microbe-responsive host gene, *Arc1* appears in several published transcriptomic datasets as among the most significantly microbiota-responsive genes. However, while we found a significant elevation of *Arc1* in the heads of GF males and females (Fig. 5A), these studies all observed a decrease in *Arc1* expression in gut or whole GF females (Guo *et al.* 2014; Dobson *et al.* 2016; Petkau *et al.* 2017). This discrepancy may further

795 underscore the importance of tissue-specificity of microbial impacts on gene expression.
796 Additionally, these studies were predominantly conducted in the Canton-S wild-type fly
797 stock, and our RT-qPCR experiments conducted on w^{1118} *Drosophila* suggest that the
798 *Arc1* expression change may be sensitive to host genotype, sex, and/or an interaction
799 between the two (Figure S5). Nevertheless, the potential connections between the
800 bacterial microbiota, *Arc1* expression and function, and *Drosophila* metabolism represent
801 a promising avenue for future investigation.

802 **Summary and concluding remarks**

803 Together this study has revealed genes that exhibit altered expression in the head
804 of young adult *Drosophila* upon elimination of the microbiota. We hypothesize that some
805 of these genes contribute to the host molecular mechanisms underlying known
806 microbiota-impacted traits, including metabolic function, stress resistance, and aging. In
807 addition, many genes not prioritized in our screening process nevertheless trended with
808 differential expression patterns in our RNA-seq study, and may still represent bona fide,
809 biologically relevant microbiota-regulated genes. Examples include genes with roles in
810 circadian rhythms, visual and odor perception, and cellular responses to hypoxia (File
811 S5). We predict that analysis of these gene expression changes at the tissue and cell
812 type specific level will reveal connections between the gut microbiota and many novel
813 aspects of host physiology and behavior.

814 **ACKNOWLEDGMENTS**

815 We thank members of the McCartney, Hiller, Minden, and Mitchell labs for helpful
816 discussions during the performance of the study and preparation of this manuscript. The
817 Top Banana fly stock was a generous gift from Dr. Michael Dickinson's lab (CalTech). Dr.

Carol Woolford (Mitchell lab, Carnegie Mellon University) performed the nanoString hybridizations. Dr. Anagha Kadam assisted with use of nSolver analysis software and LinRegPCR software. We would like to thank the Woolford, Mitchell, and Hinman labs and the Molecular Biosensor and Imaging Center at Carnegie Mellon University for reagents and equipment. Funding for this work was provided by a Carnegie Mellon University ProSEED/BrainHub seed grant to B.M.M., N.L.H., and C.K..

LITERATURE CITED

Adair, K. L., M. Wilson, A. Bost, and A. E. Douglas, 2018 Microbial community assembly in wild populations of the fruit fly Drosophila melanogaster. ISME J. 12: 959-972

Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman, 1990 Basic local alignment search tool. J. Mol. Biol. 215: 403–10.

Ashley, J., B. Cordy, D. Lucia, L. G. Fradkin, V. Budnik *et al.*, 2018 Retrovirus-like Gag Protein Arc1 Binds RNA and Traffics across Synaptic Boutons. Cell 172: 262–270.e11.

Atarashi, K., T. Tanoue, T. Shima, A. Imaoka, T. Kuwahara *et al.*, 2011 Induction of Colonic Regulatory T Cells. Science 331: 337–342.

Back, P., B. P. Braeckman, and F. Matthijssens, 2012 ROS in Aging Caenorhabditis elegans : Damage or Signaling? Oxid. Med. Cell. Longev. 2012: 1–14.

Barajas-azpeleta, R., J. Wu, J. Gill, and R. Welte, 2018 Antimicrobial peptides modulate long-term memory. PLoS Genet. 1–26.

Bergé, J. B., R. Feyereisen, and M. Amichot, 1998 Cytochrome P450 monooxygenases and insecticide resistance in insects. Philos. Trans. R. Soc. B Biol. Sci. 353: 1701–1705.

841 Bing, X. L., J. Gerlach, G. Loeb, and N. Buchon, 2018 Nutrient-dependent impact of
842 microbes on Drosophila suzukii development. MBio 9: e02199-17.
843 Blum, J. E., C. N. Fischer, J. Miles, and J. Handelsman, 2013 Frequent Replenishment
844 Sustains the Beneficial Microbiome of Drosophila melanogaster. MBio 4: 1–8.
845 Bost, A., S. Franzenburg, K. L. Adair, V. G. Martinson, G. Loeb *et al.*, 2017 How gut
846 transcriptional function of *Drosophila melanogaster* varies with the presence and
847 composition of the gut microbiota. Mol. Ecol. 27: 1–12.
848 Bost, A., V. G. Martinson, S. Franzenburg, K. L. Adair, A. Albasi *et al.*, 2018 Functional
849 variation in the gut microbiome of wild *Drosophila* populations. Mol. Ecol. 27: 2834–
850 2845.
851 Broderick, N. A., N. Buchon, and B. Lemaitre, 2014 Microbiota-Induced Changes in
852 Drosophila melanogaster Host Gene Expression and Gut Morphology MBio 5: 1–
853 13.
854 Broderick, N. A., and B. Lemaitre, 2012 Gut-associated microbes of Drosophila
855 melanogaster. Gut Microbes 3: 307–321.
856 Brummel, T., A. Ching, L. Seroude, A. F. Simon, and S. Benzer, 2004 Drosophila
857 lifespan enhancement by exogenous bacteria. Proc. Natl. Acad. Sci. 101: 12974–
858 12979.
859 Buchon, N., N. a. Broderick, S. Chakrabarti, and B. Lemaitre, 2009 Invasive and
860 indigenous microbiota impact intestinal stem cell activity through multiple pathways
861 in Drosophila. Genes Dev. 23: 2333–2344.
862 Camacho, C., G. Coulouris, V. Avagyan, N. Ma, J. Papadopoulos *et al.*, 2009 BLAST+:
863 Architecture and applications. BMC Bioinformatics 10: 1–9.

Cao, Y., S. Chtarbanova, A. J. Petersen, and B. Ganetzky, 2013 Dnr1 mutations cause neurodegeneration in Drosophila by activating the innate immune response in the brain. Proc. Natl. Acad. Sci. 110: E1752–E1760.

Chaston, J. M., A. J. Dobson, P. D. Newell, and A. E. Douglas, 2016 Host genetic control of the microbiota mediates the Drosophila nutritional phenotype. Appl. Environ. Microbiol. 82: 671–679.

Chaston, J., P. Newell, and A. Douglas, 2014 Metagenome-wide association of microbial determinants of host phenotype in Drosophila melanogaster. MBio 5: 1–12.

Chintapalli, V. R., J. Wang, and J. A. T. Dow, 2007 Using FlyAtlas to identify better Drosophila melanogaster models of human disease. Nat. Genet. 39: 715–720.

Chung, H., A. kyeong Kim, S. A. Jung, S. W. Kim, K. Yu *et al.*, 2010 The Drosophila homolog of methionine sulfoxide reductase A extends lifespan and increases nuclear localization of FOXO. FEBS Lett. 584: 3609–3614.

Clark, R. I., A. Salazar, R. Yamada, S. Fitz-Gibbon, M. Morselli *et al.*, 2015 Distinct Shifts in Microbiota Composition during Drosophila Aging Impair Intestinal Function and Drive Mortality. Cell Rep. 12: 1656–1667.

Cockburn, D. W., and N. M. Koropatkin, 2016 Polysaccharide Degradation by the Intestinal Microbiota and Its Influence on Human Health and Disease. J. Mol. Biol. 428: 3230–3252.

Combe, B. E., A. Defaye, N. Bozonnet, D. Puthier, J. Royet *et al.*, 2014 Drosophila microbiota modulates host metabolic gene expression via IMD/NF-KB signaling. PLoS One 9: e94729.

887 Daïen, C. I., G. V. Pinget, J. K. Tan, and L. Macia, 2017 Detrimental impact of
888 microbiota-accessible carbohydrate-deprived diet on gut and immune homeostasis:
889 An overview. Front. Immunol. 8: 1-7.
890 Daisley, B. A., M. Trinder, T. W. McDowell, H. Welle, J. S. Dube *et al.*, 2017
891 Neonicotinoid-induced pathogen susceptibility is mitigated by Lactobacillus
892 plantarum immune stimulation in a Drosophila melanogaster model. Sci. Rep. 7:
893 2703.
894 Dobson, A. J., J. M. Chaston, and A. E. Douglas, 2016 The Drosophila transcriptional
895 network is structured by microbiota. BMC Genomics 17: 1–9.
896 Dobson, A. J., J. M. Chaston, P. D. Newell, L. Donahue, S. L. Hermann *et al.*, 2015
897 Host genetic determinants of microbiota-dependent nutrition revealed by genome-
898 wide analysis of Drosophila melanogaster. Nat Commun 6: 6312.
899 Douglas, A. E., 2018 The Drosophila model for microbiome research. Lab Anim. (NY).
900 47: 157–164.
901 Durinck, S., P. T. Spellman, E. Birney, and W. Huber, 2009 Mapping identifiers for the
902 integration of genomic datasets with the R/Bioconductor package biomaRt. Nat.
903 Protoc. 4: 1184.
904 Early, A. M., N. Shanmugarajah, N. Buchon, and A. G. Clark, 2017 Drosophila
905 Genotype Influences Commensal Bacterial Levels. PLoS One 12: e0170332.
906 Eden, P. a, T. M. Schmidt, R. P. Blakemore, and N. R. Pace, 1991 Phylogenetic
907 analysis of Aquaspirillum magnetotacticum using polymerase chain reaction-
908 amplified 16S rRNA-specific DNA. Int. J. Syst. Bacteriol. 41: 324–325.
909 Elgart, M., S. Stern, O. Salton, Y. Gnainsky, Y. Heifetz *et al.*, 2016 Impact of gut

910 microbiots on the fly's germ line. Nat. Commun. 7: 11280
911 Elya, C., V. Zhang, W. B. Ludington, and M. B. Eisen, 2016 Stable host gene
912 expression in the gut of adult Drosophila melanogaster with different bacterial
913 mono-associations. PLoS One 11: E0167357.
914 Erkosar, B., S. Kolly, J. R. van der Meer, and T. J. Kawecki, 2017 Adaptation to
915 Chronic Nutritional Stress Leads to Reduced Dependence on Microbiota in
916 Drosophila melanogaster . MBio 8: 1–18.
917 Ewald, C., 2018 Redox Signaling of NADPH Oxidases Regulates Oxidative Stress
918 Responses, Immunity and Aging. Antioxidants 7: 130.
919 Farine, J.-P., W. Habbachi, J. Cortot, S. Roche, and J.-F. Ferveur, 2017 Maternally-
920 transmitted microbiota affects odor emission and preference in Drosophila larva.
921 Sci. Rep. 7: 6062.
922 Fast, D., A. Duggal, and E. Foley, 2018 Monoassociation with Lactobacillus plantarum
923 Disrupts Intestinal Homeostasis in Adult Drosophila melanogaster. MBio 9: e01114-
924 18.
925 Ferrandon, D., J. L. Imler, C. Hetru, and J. A. Hoffmann, 2007 The Drosophila systemic
926 immune response: Sensing and signalling during bacterial and fungal infections.
927 Nat. Rev. Immunol. 7: 862–874.
928 Fischer, C., E. P. Trautman, J. M. Crawford, E. V Stabb, J. Handelsman *et al.*, 2017
929 Metabolite exchange between microbiome members produces compounds that
930 influence *Drosophila* behavior. Elife 6: 1–25.
931 Géminard, C., E. J. Rulifson, and P. Léopold, 2009 Remote Control of Insulin Secretion
932 by Fat Cells in Drosophila. Cell Metab. 10: 199–207.

933 Guo, L., J. Karpac, S. L. Tran, and H. Jasper, 2014 PGRP-SC2 promotes gut immune
934 homeostasis to limit commensal dysbiosis and extend lifespan. Cell 156: 109–122.
935 Hang, S., A. E. Purdy, W. P. Robins, Z. Wang, M. Mandal *et al.*, 2014 The acetate
936 switch of an intestinal pathogen disrupts host insulin signaling and lipid metabolism.
937 Cell Host Microbe 16: 592–604.
938 Hsiao, E. Y., S. W. McBride, S. Hsien, G. Sharon, E. R. Hyde *et al.*, 2013 Microbiota
939 Modulate Behavioral and Physiological Abnormalities Associated with
940 Neurodevelopmental Disorders. Cell 155: 1451–1.
941 Huang, J.-H., and A. E. Douglas, 2015 Consumption of dietary sugar by gut bacteria
942 determines *Drosophila* lipid content. Biol. Lett. 11: 20150469.
943 Iatsenko, I., J. P. Boquete, and B. Lemaitre, 2018 Microbiota-Derived Lactate Activates
944 Production of Reactive Oxygen Species by the Intestinal NADPH Oxidase Nox and
945 Shortens Drosophila Lifespan. Immunity 49: 929–942.e5.
946 Jacqueline, C., L. Brazier, D. Faugère, F. Renaud, F. Thomas *et al.*, 2017 Can intestinal
947 microbiota be associated with non-intestinal cancers? Sci. Rep. 7: 12722.
948 Johnson, T. E., G. J. Lithgow, and S. Murakami, 1996 Interventions that increase the
949 response to stress offer the potential for effective life prolongation and increased
950 health. Journals Gerontol. - Ser. A Biol. Sci. Med. Sci. 51: B392-B395.
951 Jones, R. M. M., C. Desai, T. M. M. Darby, L. Luo, A. A. A. Wolfarth *et al.*, 2015
952 Lactobacilli Modulate Epithelial Cytoprotection through the Nrf2 Pathway. Cell Rep.
953 12: 1217–1225.
954 Jones, R. M., L. Luo, C. S. Ardita, A. N. Richardson, Y. M. Kwon *et al.*, 2013 Symbiotic
955 lactobacilli stimulate gut epithelial proliferation via Nox-mediated generation of

reactive oxygen species. EMBO J. 32: 3017–28.

Judd, A. M., M. K. Matthews, R. Hughes, M. Veloz, C. E. Sexton *et al.*, 2018 Bacterial methionine metabolism genes influence Drosophila melanogaster starvation resistance 1 2. 1–35.

Kamareddine, L., W. P. Robins, C. D. Berkey, J. J. Mekalanos, and P. I. Watnick, 2018 The Drosophila Immune Deficiency Pathway Modulates Enteroendocrine Function and Host Metabolism. Cell Metab. 28: 449-462.

Kang, J., J. Kim, and K. W. Choi, 2011 Novel cytochrome P450, cyp6a17, is required for temperature preference behavior in Drosophila. PLoS One 6: E29800.

Kim, G., J. H. Huang, J. G. McMullen, P. D. Newell, and A. E. Douglas, 2017 Physiological responses of insects to microbial fermentation products: insights from the interactions between Drosophila and acetic acid. J. Insect Physiol. 106: 13-19.

Kohl, K. D., and H. V. Carey, 2016 A place for host–microbe symbiosis in the comparative physiologist's toolbox. J. Exp. Biol. 219: 3496–3504.

Kounatidis, I., S. Chtarbanova, Y. Cao, M. Hayne, D. Jayanth *et al.*, 2017 NF-κB Immunity in the Brain Determines Fly Lifespan in Healthy Aging and Age-Related Neurodegeneration. Cell Rep. 19: 836–848.

Koyle, M. L., M. Veloz, A. M. Judd, A. C.-N. Wong, P. D. Newell *et al.*, 2016 Rearing the Fruit Fly Drosophila melanogaster Under Axenic and Gnotobiotic Conditions. J. Vis. Exp. 1–8.

Lee, K. A., S. H. Kim, E. K. Kim, E. M. Ha, H. You *et al.*, 2013 Bacterial-derived uracil as a modulator of mucosal immunity and gut-microbe homeostasis in drosophila. Cell 153: 797–811.

979 Leitão-Gonçalves, R., Z. Carvalho-Santos, A. P. Francisco, G. T. Fioreze, M. Anjos *et*
980 *al.*, 2017 Commensal bacteria and essential amino acids control food choice
981 behavior and reproduction. PLOS Biol. 15: e2000862.
982 Li, H., Y. Qi, and H. Jasper, 2016 Preventing Age-Related Decline of Gut
983 Compartmentalization Limits Microbiota Dysbiosis and Extends Lifespan. Cell Host
984 Microbe 19: 240–253.
985 Lin, M. T., and M. F. Beal, 2003 The oxidative damage theory of aging, Clin. Neurosci.
986 Res. 2: 305-315.
987 Lindberg, B. G., X. Tang, W. Dantoft, P. Gohel, S. Seyedoleslami Esfahani *et al.*, 2018
988 Nubbin isoform antagonism governs Drosophila intestinal immune homeostasis.
989 PLoS Pathog. 14: E1006936.
990 Liu, W., K. Zhang, Y. Li, W. Su, K. Hu *et al.*, 2017 Enterococci Mediate the Oviposition
991 Preference of Drosophila melanogaster through Sucrose Catabolism. Sci. Rep. 7:
992 1–14.
993 Love, M. I., W. Huber, and S. Anders, 2014 Moderated estimation of fold change and
994 dispersion for RNA-seq data with DESeq2. Genome Biol. 15: 550.
995 Lyne, R., R. Smith, K. Rutherford, M. Wakeling, A. Varley *et al.*, 2007 FlyMine: an
996 integrated database for Drosophila and Anopheles genomics. Genome Biol. 8:
997 R129.
998 Martino, M., D. Ma, and F. Leulier, 2017 Microbial influence on Drosophila biology. Curr.
999 Opin. Microbiol. 38: 165–170.
1000 Matos, R. C., M. Schwarzer, H. Gervais, P. Courtin, P. Joncour *et al.*, 2017 D-
1001 Alanylation of teichoic acids contributes to Lactobacillus plantarum-mediated

1002 Drosophila growth during chronic undernutrition. Nat. Microbiol. 2: 1635-1647.
1003 Mattaliano, M. D., E. S. Montana, K. M. Parisky, J. T. Littleton, and L. C. Griffith, 2007
1004 The Drosophila ARC homolog regulates behavioral responses to starvation. Mol.
1005 Cell. Neurosci. 36: 211–221.
1006 McFall-Ngai, M., M. G. Hadfield, T. C. G. Bosch, H. V Carey, T. Domazet-Lošo *et al.*,
1007 2013 Animals in a bacterial world, a new imperative for the life sciences. Proc. Natl.
1008 Acad. Sci. 110: 3229–3236.
1009 Morgulis, A., G. Coulouris, Y. Raytselis, T. L. Madden, R. Agarwala *et al.*, 2008
1010 Database indexing for production MegaBLAST searches. Bioinformatics 24: 1757-
1011 1764.
1012 Morrow, G., M. Samson, S. Michaud, and R. M. Tanguay, 2004 Overexpression of the
1013 small mitochondrial Hsp22 extends Drosophila life span and increases resistance
1014 to oxidative stress. FASEB J. 18: 598–599.
1015 Mosher, J., W. Zhang, R. Z. Blumhagen, A. D'Alessandro, T. Nemkov *et al.*, 2015
1016 Coordination between Drosophila Arc1 and a specific population of brain neurons
1017 regulates organismal fat. Dev. Biol. 405: 280–290.
1018 Nässel, D. R., and J. Vanden Broeck, 2016 Insulin/IGF signaling in Drosophila and
1019 other insects: Factors that regulate production, release and post-release action of
1020 the insulin-like peptides. Cell. Mol. Life Sci. 73: 271–290.
1021 Naudin, C. R., R. M. Jones, B. J. Saeedi, L. Luo, T. M. Darby *et al.*, 2019 Lactococcus
1022 lactis subsp. cremoris is an efficacious beneficial bacterium that limits tissue injury
1023 in the intestine. iScience 12: 356–367.
1024 Newell, P. D., J. M. Chaston, Y. Wang, N. J. Winans, D. R. Sannino *et al.*, 2014 In vivo

function and comparative genomic analyses of the drosophila gut microbiota identify candidate symbiosis factors. Front. Microbiol. 5: 576.

Newell, P. D., and A. E. Douglas, 2014 Interspecies Interactions Determine the Impact of the Gut Microbiota on Nutrient Allocation in Drosophila melanogaster. Appl. Environ. Microbiol. 80: 788–796.

O'Connor, R. M., E. F. Stone, C. R. Wayne, E. V. Marcinkevicius, M. Ulgherait *et al.*, 2017 A Drosophila model of Fragile X syndrome exhibits defects in phagocytosis by innate immune cells. J. Cell Biol. 216: 595–605.

Obadia, B., Z. T. Güvener, V. Zhang, J. A. Ceja-Navarro, E. L. Brodie *et al.*, 2017 Probabilistic Invasion Underlies Natural Gut Microbiome Stability. Curr. Biol. 27: 1999-2006.

Obata, F., C. O. Fons, and A. P. Gould, 2018 Early-life exposure to low-dose oxidants can increase longevity via microbiome remodelling in Drosophila. Nat. Commun. 9: 975.

Park, S., J. M. Park, S. Kim, J. A. Kim, J. D. Shepherd *et al.*, 2008 Elongation Factor 2 and Fragile X Mental Retardation Protein Control the Dynamic Translation of Arc/Arg3.1 Essential for mGluR-LTD. Neuron 59: 70–83.

Pastuzyn, E. D., C. E. Day, R. B. Kearns, M. Kyrke-Smith, A. V. Taibi *et al.*, 2018 The Neuronal Gene Arc Encodes a Repurposed Retrotransposon Gag Protein that Mediates Intercellular RNA Transfer. Cell 173: 275.

Patro, R., G. Duggal, M. I. Love, R. A. Irizarry, and C. Kingsford, 2017 Salmon provides fast and bias-aware quantification of transcript expression. Nat. Methods 14: 417-419.

1048 Petersen, A. J., R. J. Katzenberger, and D. A. Wassarman, 2013 The innate immune
1049 response transcription factor relish is necessary for neurodegeneration in a
1050 Drosophila model of ataxia-telangiectasia. Genetics 194: 133–142.
1051 Petkau, K., M. Ferguson, S. Guntermann, and E. Foley, 2017 Constitutive Immune
1052 Activity Promotes Tumorigenesis in Drosophila Intestinal Progenitor Cells. Cell
1053 Rep. 20: 1784–1793.
1054 Petkau, K., B. D. Parsons, A. Duggal, and E. Foley, 2014 A deregulated intestinal cell
1055 cycle program disrupts tissue homeostasis without affecting longevity in drosophila.
1056 J. Biol. Chem. 289: 28719–28729.
1057 Pruesse, E., J. Peplies, and F. O. Glöckner, 2012 SINA: Accurate high-throughput
1058 multiple sequence alignment of ribosomal RNA genes. Bioinformatics 28: 1823–
1059 1829.
1060 Ramakers, C., J. M. Ruijter, R. H. Lekanne Deprez, and A. F. M. Moorman, 2003
1061 Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR)
1062 data. Neurosci. Lett. 339: 62–66.
1063 Ridley, E. V., A. C. N. Wong, S. Westmiller, and A. E. Douglas, 2012 Impact of the
1064 resident microbiota on the nutritional phenotype of drosophila melanogaster. PLoS
1065 One 7: E36765.
1066 Roesijadi, G., S. Rezvankhah, D. M. Binninger, and H. Weissbach, 2007 Ecdysone
1067 induction of MsrA protects against oxidative stress in Drosophila. Biochem.
1068 Biophys. Res. Commun. 354: 511–516.
1069 Rosser, E. C., and C. Mauri, 2016 A clinical update on the significance of the gut
1070 microbiota in systemic autoimmunity. J. Autoimmun. 74: 85–93.

1071 Rudinskiy, N., J. M. Hawkes, R. A. Betensky, M. Eguchi, S. Yamaguchi *et al.*, 2012
1072 Orchestrated experience-driven Arc responses are disrupted in a mouse model of
1073 Alzheimer's disease. Nat. Neurosci. 15: 1422–1429.
1074 Ryu, J.-H., S.-H. Kim, H.-Y. Lee, J. Y. Bai, Y.-D. Nam *et al.*, 2008 Innate Immune
1075 Homeostasis by the Homeobox Gene Caudal and Commensal-Gut Mutualism in
1076 Drosophila. Science 319: 777–782.
1077 Sampson, T. R., and S. K. Mazmanian, 2015 Control of brain development, function,
1078 and behavior by the microbiome. Cell Host Microbe 17: 565–576.
1079 Sannino, D. R., A. J. Dobson, K. Edwards, E. R. Angert, and N. Buchon, 2018 The
1080 *Drosophila melanogaster* Gut Microbiota Provisions Thiamine to Its Host. MBio 9:
1081 e00155-18.
1082 Sansone, C. L., J. Cohen, A. Yasunaga, J. Xu, G. Osborn *et al.*, 2015 Microbiota-
1083 Dependent Priming of Antiviral Intestinal Immunity in Drosophila. Cell Host Microbe
1084 18: 571–81.
1085 Schretter, C. E., J. Vielmetter, I. Bartos, Z. Marka, S. Marka *et al.*, 2018 A gut microbial
1086 factor modulates locomotor behaviour in Drosophila. Nature 563: 402-406.
1087 Sharon, G., T. R. Sampson, D. H. Geschwind, and S. K. Mazmanian, 2016 The Central
1088 Nervous System and the Gut Microbiome. Cell 167: 915–932.
1089 Shepherd, J. D., and M. F. Bear, 2011 New views of Arc, a master regulator of synaptic
1090 plasticity. Nat. Neurosci. 14: 279–284.
1091 Shin, S. C., S. H. Kim, H. You, B. Kim, A. C. Kim *et al.*, 2011 Drosophila microbiome
1092 modulates host developmental and metabolic homeostasis via insulin signaling.
1093 Science 334: 670–674.

1094 Smith, K., K. D. McCoy, and A. J. Macpherson, 2007 Use of axenic animals in studying
1095 the adaptation of mammals to their commensal intestinal microbiota. Semin.
1096 Immunol. 19: 59–69.
1097 Storelli, G., A. Defaye, B. Erkosar, P. Hols, J. Royet *et al.*, 2011 Lactobacillus plantarum
1098 promotes drosophila systemic growth by modulating hormonal signals through
1099 TOR-dependent nutrient sensing. Cell Metab. 14: 403–414.
1100 Storelli, G., M. Strigini, T. Grenier, L. Bozonnet, M. Schwarzer *et al.*, 2017 Drosophila
1101 Perpetuates Nutritional Mutualism by Promoting the Fitness of Its Intestinal
1102 Symbiont Lactobacillus plantarum. Cell Metab. 1–16.
1103 Téfit, M. A., and F. Leulier, 2017 *Lactobacillus plantarum* favors the early emergence of
1104 fit and fertile adult Drosophila upon chronic undernutrition. J. Exp. Biol. 220: 900-
1105 907.
1106 Tzingounis, A. V., and R. A. Nicoll, 2006 Arc/Arg3.1: Linking Gene Expression to
1107 Synaptic Plasticity and Memory. Neuron 52: 403–407.
1108 Wang, L., J. Karpac, and H. Jasper, 2013 Promoting longevity by maintaining metabolic
1109 and proliferative homeostasis. J. Exp. Biol. 217: 109–118.
1110 Wong, A. C., J. M. Chaston, and A. E. Douglas, 2013 The inconstant gut microbiota of
1111 Drosophila species revealed by 16S rRNA gene analysis. ISME J 7: 1922–1932.
1112 Wong, A. C., A. J. Dobson, and A. E. Douglas, 2014 Gut microbiota dictates the
1113 metabolic response of Drosophila to diet. J Exp Biol 217: 1894–1901.
1114 Wong, A. C. N., Y. Luo, X. Jing, S. Franzenburg, A. Bost *et al.*, 2015 The host as the
1115 driver of the microbiota in the gut and external environment of Drosophila
1116 melanogaster. Appl. Environ. Microbiol. 81: 6232–6240.

Wong, C. N. A., P. Ng, and A. E. Douglas, 2011 Low-diversity bacterial community in the gut of the fruitfly Drosophila melanogaster. Environ. Microbiol. 13: 1889–1900.

Wong, A. C.-N., Q.-P. Wang, J. Morimoto, A. M. Senior, M. Lihoreau *et al.*, 2017 Gut Microbiota Modifies Olfactory-Guided Microbial Preferences and Foraging Decisions in Drosophila. Curr. Biol. 13: 163–167.

Wu, S.-C., Z.-S. Cao, K.-M. Chang, and J.-L. Juang, 2017 Intestinal microbial dysbiosis aggravates the progression of Alzheimer's disease in Drosophila. Nat. Commun. 8: 24.

Yu, S., Y. Jang, D. Paik, E. Lee, and J. J. Park, 2015 Nmdmc overexpression extends Drosophila lifespan and reduces levels of mitochondrial reactive oxygen species. Biochem. Biophys. Res. Commun. 465: 845–850.

FIGURE LEGENDS

Figure 1
Overview of the head transcriptomic response microbiota elimination in adult male Top Banana *Drosophila*. (A) Volcano plot representation of RNA-seq results plotting all genes with non-zero read counts according to statistical significance (Benjamini-Hochberg adjusted p-value) as a function of fold change in GF *vs.* CV male heads. Genes with $\log_2$ fold change >0.5 and FDR-adjusted p-value <0.05 are highlighted. (B) Heatmap representation of 343 candidate genes identified as statistically significant (p-adj<0.1) in at least one leave-out analysis conducted on the entire RNA-seq dataset (File S4). Each column represents a biological replicate of the indicated condition (RNA from 50 individual heads/replicate); colors represent row Z-scores derived from transcripts per million (TPM) values for each gene. (C) GO-terms enriched (Holm-Bonferroni adjusted p-value<0.001, ≥20 genes) in the 248/343 candidate genes downregulated (l2fc<0) in the heads of male GF compared to CV flies. (D) GO-terms enriched (Holm-Bonferroni adjusted p-value<0.05) in the 95/343 candidate genes upregulated (l2fc>0) in the heads of male GF compared to CV flies.

Figure 2
Microbiota-dependent gene expression changes in adult male (A) and female (B) Top Banana *Drosophila* heads revealed by secondary screening with the nanoString nCounter platform. Blue bars represent $\log_2$ fold change values comparing GF and CV heads; gray bars represent $\log_2$ fold change values comparing GF heads to GNO flies reared in polyassociation with a four-species microbial community consisting of *Acetobacter* and *Lactobacillus* bacteria isolated from CV Top Banana cultures. Data are independently ordered for each sex based on $\log_2$ fold change values for the GF *vs.* GNO comparison, with GF *vs.* CV value adjacent for each gene. Fold change values are calculated based on means of normalized mRNA counts for each condition (n=5 replicates each for CV males, GF females, GNO females, and n=6 replicates each for CV females, GF males, GNO males; 5-10 heads/replicate). The values plotted here are also found in File S6 "Log2FoldChange".

Figure 3
The microbiota induce innate immune gene expression in the heads of adult flies. nanoString data are represented as individual gene plots for each immunity-related gene for which microbial condition yielded a statistically significant expression difference in at least one sex. n=5 replicates each for CV males, GF females, GNO females, and n=6 replicates each for CV females, GF males, GNO males; 5-10 heads/replicate. Data were analyzed independently for each sex via one-way ANOVA with Tukey HSD *post-hoc* tests or Kruskal-Wallis test with Dunn's multiple comparisons test when assumptions of equal variance and normal distribution were broken. p<0.05 *, p<0.01 **, ns=not significant.

Figure 4
Sex-specific microbial modulation of expression of genes involved in detoxification, oxidative stress resistance, and predicted metabolic function in the heads of adult Top Banana *Drosophila*. nanoString data are represented as plots of individual genes for

which microbial condition yielded a statistically significant expression difference in at least one sex. n=5 replicates each for CV males, GF females, GNO females, and n=6 replicates each for CV females, GF males, GNO males; 5-10 heads/replicate. Data were analyzed independently for each sex via one-way ANOVA with Tukey HSD *post-hoc* tests or Kruskal-Wallis test with Dunn's multiple comparisons test when assumptions of equal variance and normal distribution were broken. $p<0.05$ *, $p<0.01$ **, ns=not significant.

Figure 5

Microbial suppression of *Arc1* expression in adult male Top Banana *Drosophila* heads can be mediated by interaction with specific, individual bacteria. (A) nanoString data demonstrating elevation of *Arc1* mRNA in the heads of male and, more moderately, female flies. n=5 replicates each for CV males, GF females, GNO females, and n=6 replicates each for CV females, GF males, GNO males; 5-10 heads/replicate. (B) Monoassociation with only three members of the four-species GNO bacterial community–*Acetobacter sp.*, *L. brevis*, and *A. pasteurianus*, but not *L. plantarum*–is sufficient to reduce *Arc1* expression in the heads of male Top Banana flies, as measured by RT-qPCR. n=4 biological replicates of each condition, 10 heads/replicate. Figure represents fold change relative to GF flies following normalization to the average of housekeeping genes *Rpl32* and *14-3-3ε* as calculated via $\Delta\Delta C_t$ values. Error bars represent standard error of the mean. Data from both experiments were analyzed via one-way ANOVA with Tukey HSD *post-hoc* tests. $p<0.05$ *, $p<0.01$ **, $p<0.0005$ ***, ns=not significant.

Figure 6

Candidate microbiota-dependent gene expression effects identified in Top Banana *Drosophila* are impacted by host genotype. (A) Male, but not (B) female, w^{1118} flies exhibit microbiota-induced reduction of *Arc1* expression in the head. (C) & (D) *Hsp22* expression in male and female w^{1118} heads is unaffected by microbial condition, unlike the female-specific elevation observed in GF Top Banana flies via nanoString. (E) & (F) *DptA* expression is elevated in the heads of GNO w^{1118} females compared to GF females; a similar trend that does not achieve statistical significance is observed in male heads and CV female heads. n=3 biological replicates of each condition, 10 heads/replicate. Figure represents fold change relative to GF flies following normalization to *Rpl32* $\Delta\Delta C_t$ values. Error bars represent standard error of the mean. Data were analyzed via one-way ANOVA with Tukey HSD *post-hoc* tests. $p<0.0005$ ***, ns=not significant.

FIGURE 1

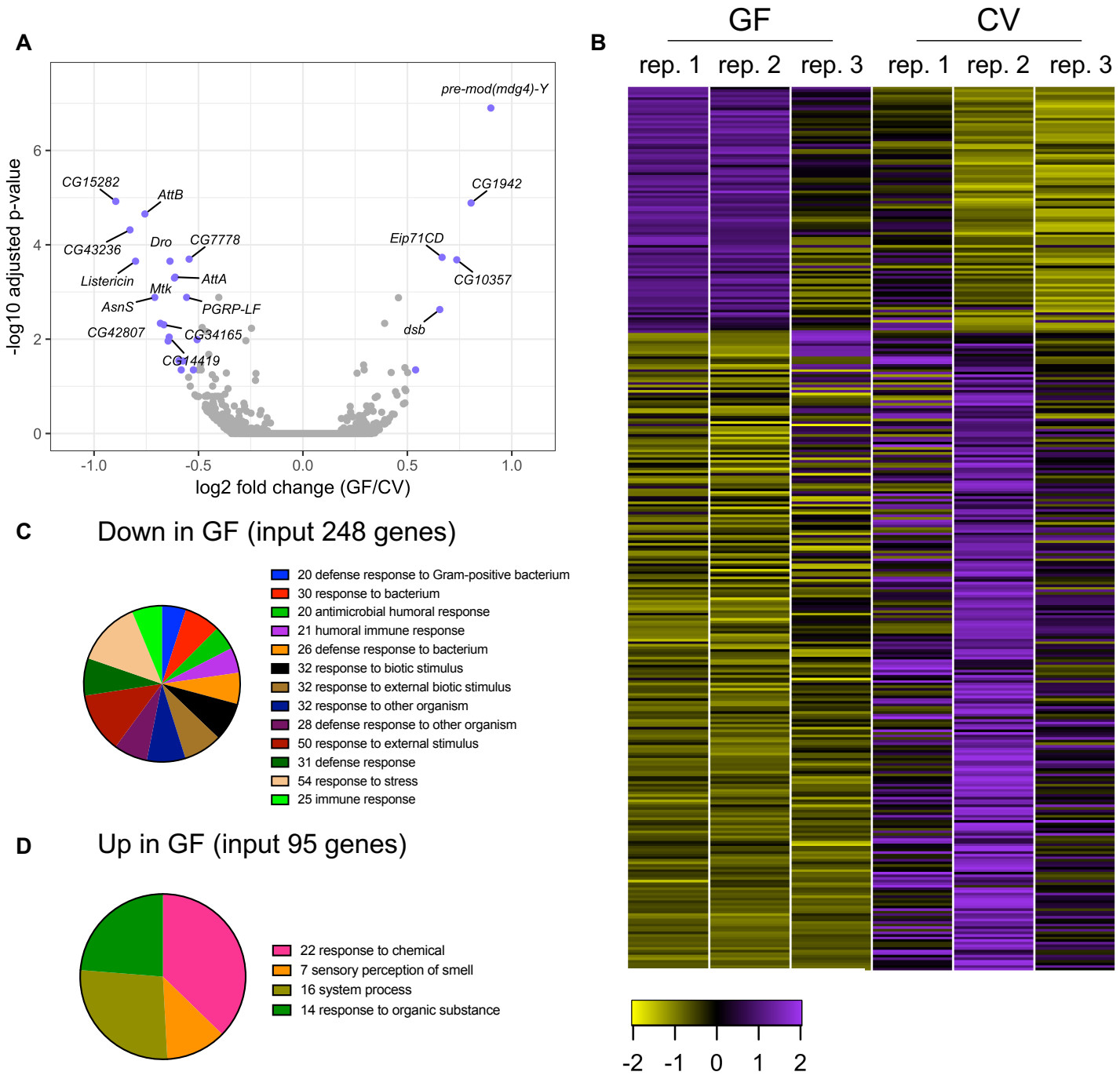

FIGURE 2

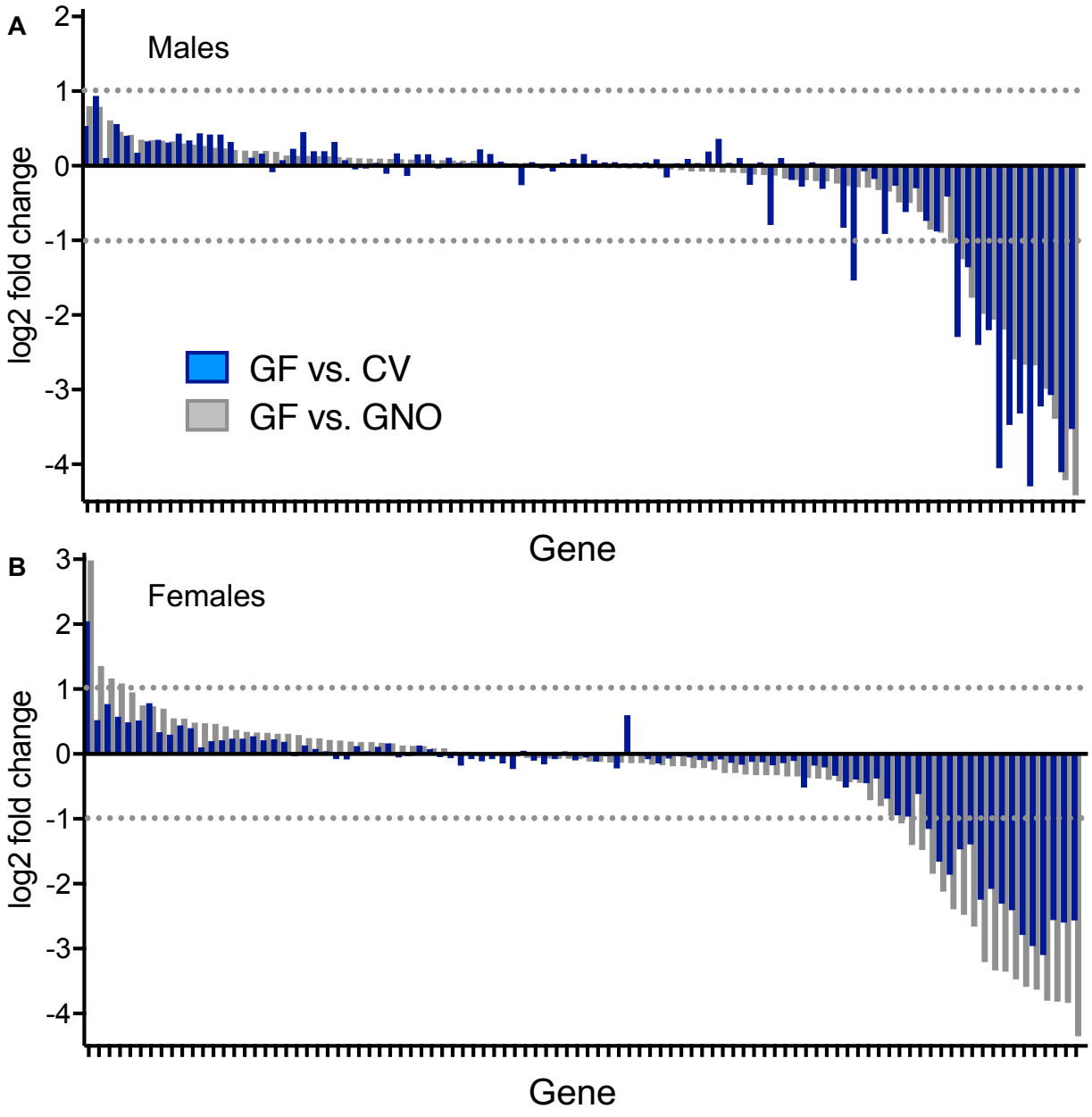

FIGURE 3

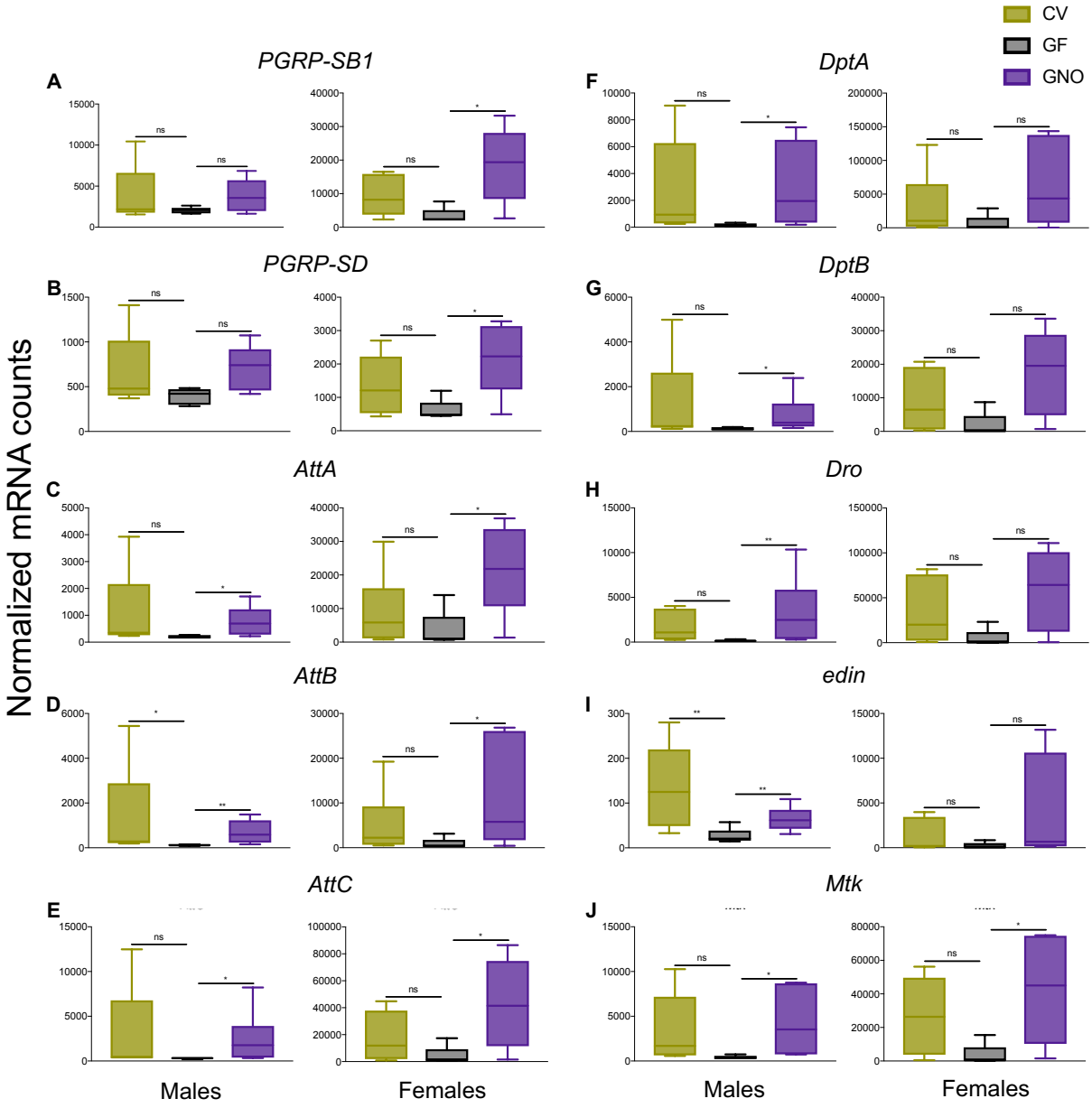

FIGURE 4

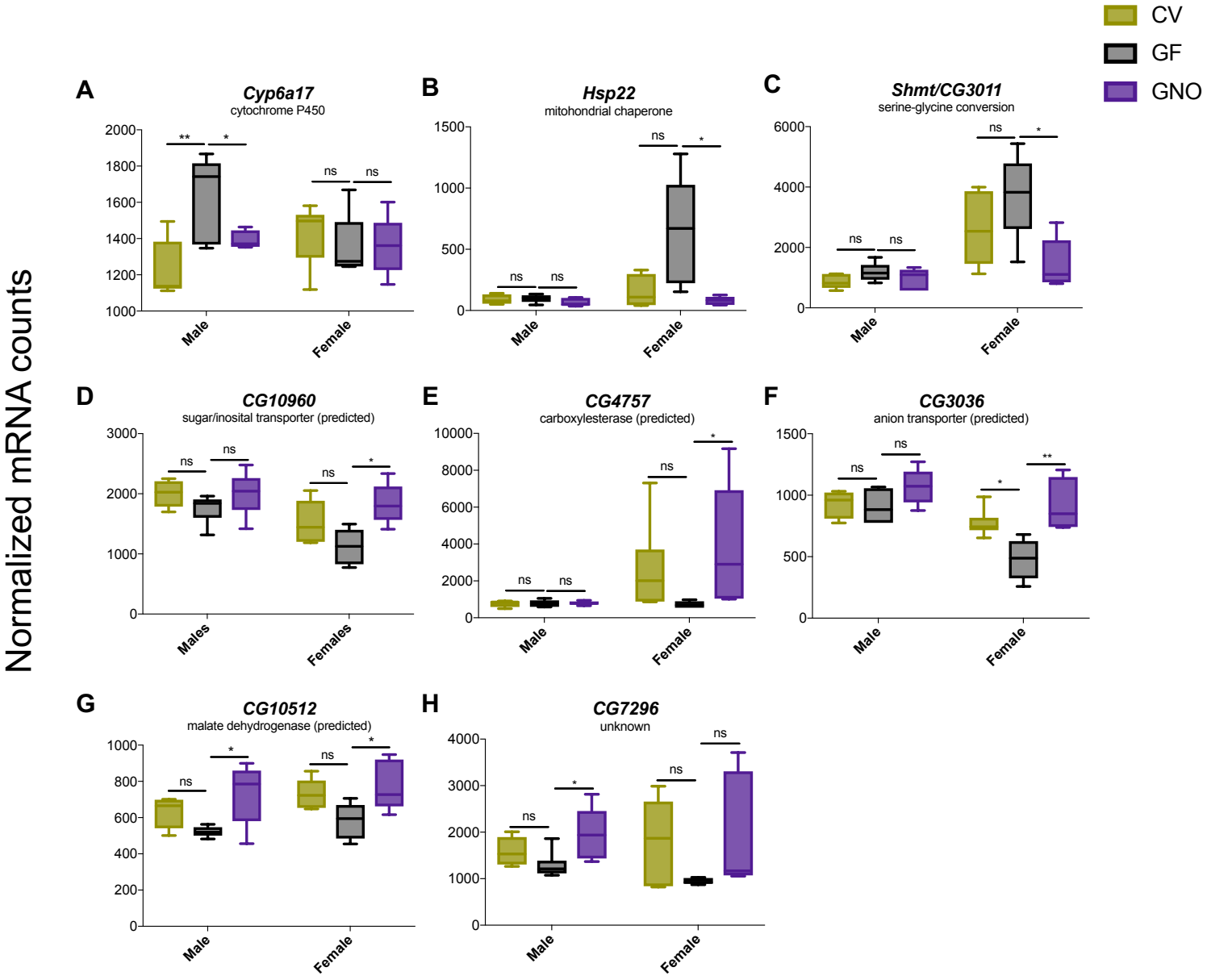

FIGURE 5

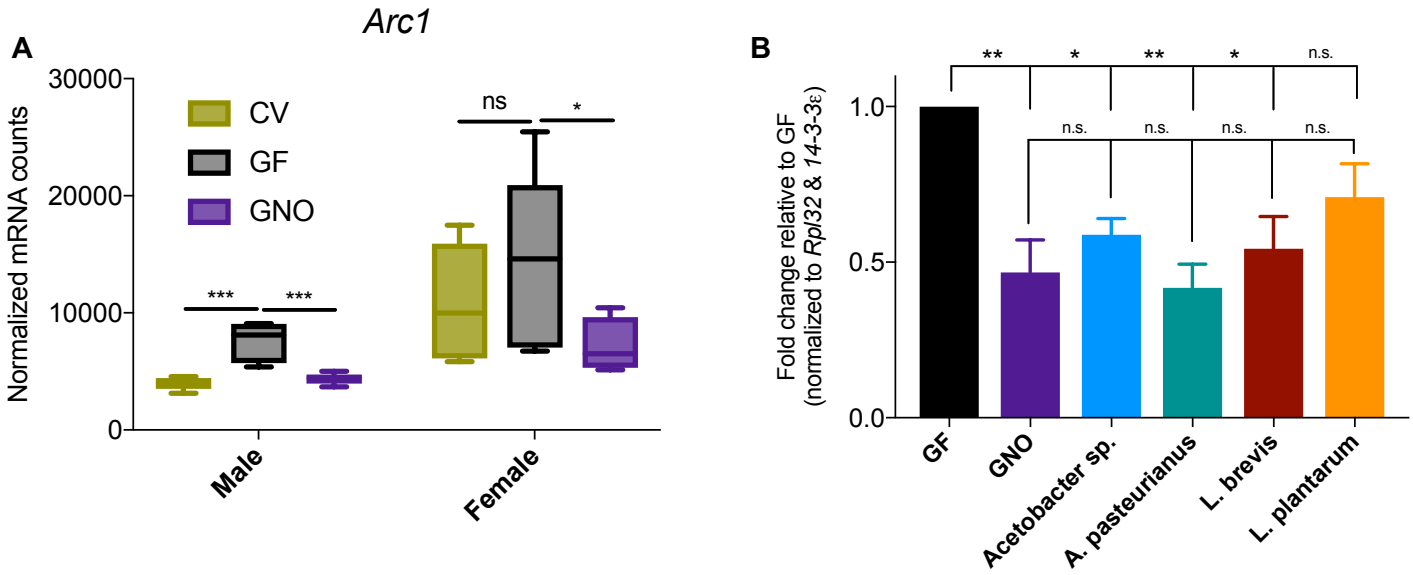

FIGURE 6