Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

The metabolome as a link in the genotype-phenotype map for peroxide resistance in the fruit fly, Drosophila melanogaster

Benjamin R Harrison (ben6@uw.edu)
University of Washington https://orcid.org/0000-0002-4268-7686
Lu Wang
University of Washington Department of Environmental and Occupational Health Sciences
Erika Gajda
University of Washington Department of Pathology
Elise V. Hoffman
University of Washington Department of Pathology
Brian Y. Chung
University of Michigan Department of Molecular and Integrative Physiology
Scott D. Pletcher
Laboratoire de Physiologie Integrative Cellulaire et Moleculaire
Dan Raftery
University of Washington Northwest Metabolomics Research Center
Daniel E.L. Promislow
University of Washington Department of Pathology

Research article

Keywords: genetic variation, complex trait, endophenotype, metabolome, GWAS, quantitative genetics, starvation

Posted Date: November 21st, 2019

DOI: https://doi.org/10.21203/rs.2.17545/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Version of Record: A version of this preprint was published on May 4th, 2020. See the published version at https://doi.org/10.1186/s12864-020-6739-1.

Abstract

Background Genetic association studies that seek to explain the inheritance of complex traits typically fail to explain more than a small fraction of the heritability of the trait under study. Thus we are left with a gap in the map from genotype to phenotype. Several approaches have been used to fill this gap, including those that attempt to map endophenotype such as the transcriptome, proteome or metabolome, that underlie complex traits. Here we used metabolomics to explore the nature of genetic variation for hydrogen peroxide (H2O2) resistance in the sequenced inbred Drosophila Genetic Reference Panel (DGRP).

Results We first studied genetic variation for H2O2 resistance in 180 DGRP lines and identify the insulin signaling modulator u-shaped and several regulators of feeding behavior. We then profiled a portion of the aqueous metabolome in subsets of eight 'high resistance' lines and eight 'low resistance' lines. We used these lines to represent collections of genotypes that were either resistant or sensitive to the stressor, effectively modeling a discrete trait. Across the range of genotypes in both populations, flies exhibited surprising consistency in their metabolomic signature of resistance. Metabolomic profiles were also able to distinguish stress-resistant from stress-sensitive flies with greater accuracy than the genotype of these same lines. Furthermore, we found a metabolic response to H2O2 that was shared among sensitive, but not resistant genotypes. Metabolomic data further implicated at least two pathways, glycogen and folate metabolism, as determinants of sensitivity to H2O2. We also discovered a confounding effect of feeding behavior on assays involving supplemented food.

Conclusions This work suggests that the metabolome can be a point of convergence for genetic variation influencing complex traits, and efficiently elucidate the mechanisms underlying this trait variation.

Background

Phenotypic variation among individuals in a population arises from variation in the genotype, the environment, and the interaction between the two. Genetic variation is a major determinant of many complex traits and, while numerous genetic association studies have failed to explain substantial portions of heritable variation in a given complex trait, use of highly polygenic models have closed this gap considerably [1]. However, highly polygenic models do not easily allow us to identify single gene-level associations with traits, and may not yield mechanistic insight into the pathways that shape trait variation. Genetic variation ultimately affects phenotype though the effect of genes on downstream 'endophenotypes'—the epigenome, transcriptome, proteome, metabolome and microbiome [2, 3]. Several authors have proposed that these endophenotypes, and the metabolome in particular, may serve as a powerful tool in mapping genotype to phenotype, as well as a source of information upon which to construct mechanistic hypotheses [3-7]. In this study, we explore the possibility that genetic effects on phenotypes are filtered through the profile of small molecules, the metabolome, that function downstream of genotype but upstream of phenotype.

The metabolome consists of the small biomolecules, typically less than 2000 Da, that make up the energetic, structural and functional building blocks of all life [8-12]. Given the role of these molecules in cells, researchers have pointed to the metabolome as a key link between genotype and phenotype [3-5, 13-15]. Genetic variation clearly influences the metabolome. For example, genome-wide association studies (GWAS) have identified alleles that potentially explain up to 60 to 80% of the variance in individual features in the human or plant metabolome [16-20]. Some have proposed mapping variation in the abundance of trait-associated metabolites in order to map genotype-to-metabolite-to-phenotype [2, 3, 5, 15, 21].

To explore the potential of metabolomic profiling to bridge the genotype-phenotypic gap and to identify underlying mechanisms of natural variation, here we study resistance to peroxide (H_2O_2) stress in a fruit fly model of genetic variation, the *Drosophila* Genetic Reference Panel (DGRP) [DGRP, 22]. This system provides an ideal model to study the ability of metabolic profiling to bridge the genotype-phenotype gap. First, there is an extensive literature on stress resistance in flies [23-25]. Second, many association studies have examined genetic variation for survival in *Drosophila* [26-28], including in response to oxidative stress [29, 30]. The DGRP, a set of fully sequenced inbred lines, now enables labs around the world to quickly identify loci associated with any trait of interest [26, 31]. Third, numerous studies over the past decade have shown that metabolite profiles in flies are highly sensitive to variation due to genotype and environment [9, 32-37].

We present an analysis of survival time and metabolic profiles in flies from DGRP lines held on H_2O_2 or control food. We identified at least two genes associated with lifespan on H_2O_2 food, including *NPF*, which encodes a feeding-stimulating neuropeptide, and *u-shaped (ush*), a regulator of development and insulin-like signaling (IIS). However, as with most GWAS, naturally occurring variation at these loci are statistically significant but explain only modest levels of trait variation. To better explain phenotypic variation, we turned to metabolome profiles. By comparing highly resistant with highly sensitive lines, we model H_2O_2 resistance as a discrete trait. We found a consistent metabolomic signature of resistance to H_2O_2. Multivariate analysis of metabolome variation across these genotypes allows us to distinguish resistant from sensitive lines, even in samples of flies not exposed to H_2O_2 food. These results suggest that multiple genotypes converge on a similar metabolomic phenotype that is associated with trait variation. Additionally, using univariate analysis of individual metabolites, we found glycogen and folate metabolism are associated with stress resistance and validate the metabolic analysis by showing that metabolite feeding or genetic manipulation of candidate pathways both affect survival on H_2O_2 food.

Finally, we found a strong effect of H_2O_2 on feeding behavior, suggesting that variation in survival of flies on food supplemented with H_2O_2 could be explained in part by variation in starvation resistance. Our results suggest that starvation or nutrient assimilation might be the underlying cause of mortality in historical assays where the stressor has been administered to *Drosophila* in the food.

Results

Resistance to peroxide food within the DGRP

We found substantial variation among 180 DGRP lines for survival on H_2O_2 food. We used a mixed model to study the variation in lifespan and, in addition to significant effects of genotype (random effect, likelihood ratio ^{2} = 36.2, df = 1, $P = 1.7x10^{-9}$, Methods), we found that weight was a significant predictor, with larger flies surviving longer on H_2O_2 (fixed effect, $\beta = 0.49$, $P < 10^{-6}$). After removing the effect of weight, we estimated a broad sense heritability for mean survival on H_2O_2 food of $H^2 = 51.1 \pm 8.9\%$ (mean ±SD).

We compared our measures of H_2O_2 resistance in the DGRP with traits measured in the DGRP in previous studies. These studies have measured either survival or behavioral responses of the DGRP to two other oxidative stressors, paraquat and menadione [29, 30]. We observed Pearson correlations of $r = 0.34$ (n = 180, $P < 10^{-4}$) and $r = 0.35$ (n = 180 $P < 10^{-4}$) between the mean survival time reported here and those measured on food supplemented with paraquat or menadione, respectively (Fig S1 [29]). We found no correlation between our measures of H_2O_2 survival and two behavioral traits, the startle response and climbing, measured by Jordan *et al.* [30] following chronic (13-16 day) exposure to menadione food (data not shown). We also found that the survival times of the DGRP on H_2O_2 food correlate highly with survival measured under starvation in two different labs (Fig S1 [26, 38]) as well as data from our own lab (Fig 1a). It is notable that the correlation between H_2O_2 resistance and starvation was greater than the correlation between H_2O_2 resistance and any other trait published for the DGRP, including survival on food containing paraquat or menadione (Fig S1). These results suggest that H_2O_2 food may affect feeding or nutrient assimilation in *Drosophila*.

To test the possibility that H_2O_2 affects feeding behavior, we measured the amount of food consumed by flies in the CAFE and dye incorporation assays described in the Methods section [39, 40]. Peroxide reduced feeding in each of three different genotypes tested, including strains with both relatively long and short survival time on H_2O_2 food (Fig 1b and c). During the 24h feeding period in the CAFE assay, flies exposed to liquid H_2O_2 food consumed no more than the volume lost due to evaporation in chambers without flies, suggesting that the flies consumed very little H_2O_2 food over 24h (Fig 1c). This finding

suggests that mortality in flies exposed to H_2O_2 is due at least in part to starvation, in addition to any oxidative stress caused by H_2O_2 exposure.

Genetic associations with peroxide resistance.

We used a linear regression model in PLINK [41] to test for associations of H_2O_2 resistance with approximately 1.9 million SNPs with a minor allele frequency (MAF) ≥ 5% and < 30% missing genotypes, while accounting for population structure and a significant effect of the major inversion *In(2L)t* (Methods). We used the *q* value approach [42] to control the false discovery rate (FDR) and at 20% FDR, 14 variants (all were SNPs) were associated with resistance (Fig 2a, Table S1). Pairwise linkage disequilibrium (LD) among the 14 SNPs suggests that they associate with H_2O_2 resistance as seven loci, or groups of SNPs in LD ($r^2 > 0.5$, Fig 2b). With only 180 genotypes, we lack the power to analyze SNP-SNP interactions among these loci. However, our data indicate that H_2O_2 resistance is polygenic within the DGRP.

To investigate these genetic associations further, we performed gene ontology (GO) analysis, looking for biological processes and signaling pathways that are over-represented among the genes associated with survival on H_2O_2 food. To identify gene-level associations, many studies use the minimum P-value of all variants in a gene (P_{min}). One might expect a bias, such that genes with more variants are more likely to have a smaller (more significant) P_{min} by chance alone. Indeed, we found that $-\log_{10}(P_{min})$ was positively associated with the number of variants per gene (Fig S2), potentially biasing gene-trait associations in favor of genes with more variants [43]. To test for this bias, we compared the top 200 genes ranked by P_{min} with the top 200 genes from ten GWAS of randomly permuted phenotypes. Out of 15,322 gene models, the null expectation for such intersection would be 2.6 genes. In contrast, we found an average of 11.6 ± 2.0 genes (mean ± se) in common across the permutations (^{2} = 28.3, df = 3, $P < 1.1\text{x}10^{-7}$), consistent with a bias caused by SNP density. To correct for this bias, we used a permutation approach to derive gene-level *P*-values, P_{gene}, while also accounting for population structure (Methods). Unlike P_{min}, P_{gene} did not associate with the number of variants per gene, and it reduced the number of false-positives when compared to top genes from GWAS of randomized phenotypes (Fig S2, $^2 = 2.2\text{x}10^{-26}$, df = 3, $P = 1$). Thus, P_{gene} increases the accuracy of gene-trait associations. Among the 29 genes at P_{gene} FDR ≤ 0.5 (Table 1) we used gene ontology (GO) enrichment analysis to identify several enriched biological processes and two enriched biological pathways (FDR ≤ 0.05, Table 2). Most of the enriched processes are nested in hierarchical categories and thus are not independent. Also, the enrichment of 7 of the 9 biological processes, and the endothelin signaling pathway, is due entirely to three genes, each encoding adenylyl cyclase (ACXA, ACXB and ACXE, Table 2). Adenylyl cyclase is involved in several signaling pathways, including G-protein coupled and calcium-based signaling. Separately, the platelet-derived

growth factor (PDGF) signaling pathway is enriched (FDR = 0.02) due to three other genes in our dataset, *Rab2*, *Ets21C* and the c-Myc-binding protein homolog CG17202.

[Please place Table 1 here]

Table 2. Process and Pathway Enrichment Among Candidate Genes

PANTHER GO-Slim Biological Process	Gene Content	Input	Expected	Fold Enrichment	P-value	FDR
activation of adenylate cyclase activity (GO:0007190)	32	3*	0.07	46.09	4.69×10^{-5}	0.0353
adenylate cyclase-activating G-protein coupled receptor signaling pathway (GO:0007189)	32	3*	0.07	46.09	4.69×10^{-5}	0.0235
cAMP-mediated signaling (GO:0019933)	41	3*	0.08	35.98	9.38×10^{-5}	0.0282
regulation of cAMP-mediated signaling (GO:0043949)	41	3*	0.08	35.98	9.38×10^{-5}	0.0235
regulation of adenylate cyclase activity (GO:0045761)	42	3*	0.09	35.12	1.00×10^{-4}	0.0216
regulation of lyase activity (GO:0051339)	43	3*	0.09	34.3	1.07×10^{-4}	0.0201
cyclic-nucleotide-mediated signaling (GO:0019935)	50	3*	0.1	29.5	1.64×10^{-4}	0.0274
regulation of biological process (GO:0050789)	1359	11	2.76	3.98	3.87×10^{-5}	0.0582
biological regulation (GO:0065007)	1492	11	3.03	3.62	9.15×10^{-5}	0.0343

PANTHER Pathways	Gene Content	Input	Expected	Fold Enrichment	P-value	FDR
PDGF signaling pathway (P00047)	47	3	0.1	31.38	1.38×10^{-4}	0.0213
Endothelin signaling pathway (P00019)	74	3*	0.15	19.93	0.0495	0.0256

Table 2. Enrichment analysis for the 29 genes associated with peroxide resistance. Gene Content is the number of genes in the respective process/pathway the *Drosophila* genome (n = 13,767 total gene models). Asterisks indicate categories enriched due to the adenylyl cyclase X gene cluster.

We validated these gene-trait associations by using RNAi to manipulate the expression of six of the 29 candidate genes, *nAChRbeta3*, *Ets21C*, *ush*, *Nha1*, *Jon25Bi* and *Marcal1* (P_{gene} < 0.0003 in all cases), and testing the effect of mutation in a seventh candidate *NPF* (P_{gene} < 0.0006). Several of the candidates reside near a cluster of six trait-associated SNPs within a 17.5 kb interval on chromosome 2 (Fig 2a). This interval spans several genes, including *u-shaped (ush,* $P_{gene} = 7.1 \times 10^{-5}$*)*, which contains an intronic C/T SNP associated with H_2O_2 resistance ($P = 2.49 \times 10^{-8}$) and several other SNPs in LD with this SNP (Fig 2b). None of these SNPs are predicted to alter amino acid sequence or splicing, but instead may affect *ush* expression. *ush* has roles in development and growth, including as a negative regulator of PI3K activity within the IIS pathway in the fat body [44]. To validate the effects of *ush* on H_2O_2 resistance, we used the RU486-inducible GAL4 GeneSwitch driver S106 to drive RNAi targeting *ush* in the adult fat body [45]. RU486 treatment of flies carrying both S106 and UAS-RNAi targeting *ush* resulted in shorter survival times on H_2O_2 food than the same genotype without RU486 (Fig 3). We saw the same result with two independent RNAi lines, each targeting a different portion of the *ush* mRNA (Fig 3). We saw no effect of RU486 on the survival of F_1 flies carrying the driver and the empty control P-element in the same genetic background as the UAS-RNAi flies, nor an effect of RNAi on lifespan on food lacking H_2O_2 (see Methods). Knocking down *ush* in the nervous system with the elav-GAL4 gene switch driver did not affect H_2O_2 resistance (data not shown), and ubiquitous knockdown of *ush* appears to be lethal as we failed to recover Act-GAL4/*ush*-RNAi flies in crosses of the *ush*-RNAi construct to the constitutive Act-GAL4 driver. We also detected a strong interaction between RU486 treatment and an Act-GAL GeneSwitch driver in our H_2O_2 food assay and so were unable to test the effect of ubiquitous knockdown of *ush* in adult flies (data not shown). Similar RNAi of five other candidates did not affect H_2O_2 resistance (data not shown).

Another candidate, *NPF* ($P_{gene} = 5.8 \times 10^{-4}$) encodes the ligand neuropeptide F (NPF), which controls feeding, ethanol sensitivity and other behaviors [46]. The npf^{SK1} deletion allele reduced H_2O_2 resistance when compared to wildtype control flies and did not affect survival on control food (Fig 3 and data not shown). These data demonstrate that manipulation of candidate genes from our GWAS affects the H_2O_2 resistance phenotype.

Metabolite profiles associated with peroxide resistance.

To investigate the effect of H_2O_2 on the fly metabolome, and the potential for the metabolome to explain genetic variation in resistance to H_2O_2, we measured untargeted LC-MS profiles in three biological replicates for each of eight resistant (mean survival time = 106.9h, range = 90.4 - 119.7h) and eight sensitive (mean survival time = 58.9h, range = 53.2 - 70.0h) lines, chosen such that the two groups did not differ substantially in size (Fig S3). These lines were subjected to another survival assay and, 24h after

being exposed to H_2O_2 or control food, samples of flies from each line and treatment were flash frozen for aqueous metabolite extraction, while survival measurements were conducted on the remaining flies.

Global metabolite profiling identified a total of 3028 and 2921 features from positive and negative ionization modes, respectively. Profiles from positive and negative modes were analyzed separately. Principal component analysis (PCA) was applied to metabolomic data (see Methods). For the negative mode data, principal component one ($PC1_{neg}$) explained ~12.9% of the variation and separated resistant from sensitive flies but did not separate flies based on treatment alone (Fig 4). PCA using positive mode data gave similar results (data not shown). Analysis along PC1 also separates H_2O_2-treated from control flies among the sensitive genotypes, but not among the resistant genotypes (Fig 4 and data not shown). Principal component analysis thus detected between-group variation in metabolite profiles from sensitive and resistant flies, and further distinguished the effect of treatment on sensitive but not resistant flies.

The separation of resistant and sensitive lines by metabolome profile is striking. However, the genotypes chosen for metabolite profiling were among the extremes of resistance to H_2O_2. This design raises the possibility that resistant and sensitive lines have a genetic signature of resistance. To determine if genotype could also separate lines chosen deliberately with extreme phenotypes, we carried out PCA and hierarchical clustering on the same lines using their genotypes. We analyzed the first ten genotype PCs, which together account for 68.7% of the variance of >2 million genetic variants in the 16 lines used for metabolomics. These PCs failed to clearly separate the resistant and sensitive flies (Fig 5). Similarly, clustering of 6846 LD-pruned variants from these 16 lines also failed to separate the two phenotypes. For comparison, clustering of 2417 negative mode metabolite features separated 14 of the 16 genotypes into two clusters composed of resistant and sensitive lines (Fig 5).

Metabolic pathways associated with peroxide resistance

We next sought to identify individual features and metabolic pathways that predicted the binary trait of resistance or sensitivity to H_2O_2 food. Accordingly, we used logistic regression, identifying a large number of metabolite features from flies on H_2O_2 food or control food whose abundance was associated with trait. Among the features detected in negative mode LC-MS of untreated (control) flies, 252 features associated with trait (FDR < 0.1), 228 of which were also found in the larger set of 637 associated features identified in the metabolome of H_2O_2 treated flies (Fig S4a). In positive mode data, 115 features associated with resistance in H_2O_2 treated flies, but no features did so from flies on control food.

Principal component analysis suggested that the metabolome of sensitive lines is altered by H_2O_2 food, while in resistant lines, the metabolome does not shift in response to H_2O_2 treatment (Fig 4). To identify the specific metabolites whose abundance was affected by treatment in a trait-dependent way we ran a linear regression model, predicting metabolite level in response to trait, treatment and the interaction between the two. Fig S4b shows the clustering of 138 features that show significant interaction term at FDR < 0.1. These data come from analysis of profiles from negative mode only, as no features from positive mode reached our FDR cutoff. Clustering the z-scores for each feature across samples revealed several groups with similar patterns, including one large group of features in which the sensitive flies, but not the resistant flies, showed a decrease in feature abundance on H_2O_2 compared to control treatment (Fig S4). This pattern is similar to the separation of samples across the latent variables revealed in PCA, where the metabolome of sensitive flies, but not resistant flies, was affected by treatment (Fig 4).

We used the *mummichog* software package [47] to identify metabolomic pathways enriched among the features associated with H_2O_2 resistance, or among features with significant treatment by trait interaction effects. This analysis identified metabolites that were enriched in several pathways, including carbohydrate metabolism, amino acids and their biosynthesis, and folate metabolism (see Additional File 1). Many of these pathways share overlapping metabolites and some pathways are nested within other pathways. We chose a subset of the identified pathways for further analysis based on a variety of criteria. One criterion was the significance of each pathway across the different measures of association. Glycogen and folate metabolic pathways were enriched among the features that were significant in at least three of the four analyses (see Additional File 1). Another criterion was the strength of the identification; we gave higher priority to those features that were uniquely assigned to a particular metabolite or pathway, rather than being ambiguously associated with more than one metabolite, or with a metabolite that is present in several pathways.

Glycogen Metabolism

Among the features that showed significant trait by treatment interaction were those with glycogen metabolism, particularly glucose and the maltodextrins containing from two to six glucose moieties (Fig S5). The next maltodextrin in this series, maltoheptaose, has a mass (1153 Da) outside the range of features measured in this experiment (62 to 1000 Da, see Methods). Most of the maltodextrins show the same general relationship with treatment and trait; their abundance is lower in sensitive flies on H_2O_2 food compared to unexposed sensitive flies, and their abundance is not affected by treatment in resistant flies (Fig S4b). In contrast to the unique assignment of features to the maltodextrins, the other metabolites of the glycogen pathway, glucose and the disaccharide maltose, had features that could be assigned to other metabolites of identical mass. Together these data suggest that glycogen metabolism differentiates sensitive verses resistant flies.

Resistant flies appear to maintain a glycogen pool after transfer to H_2O_2 food whereas glycogen is relatively depleted in sensitive flies, suggesting that their survival may be limited by their glycogen reserve (Fig 6b). Metabolite levels were measured 24h after exposure to H_2O_2, and while the glycogen pool of resistant flies appears to persist until this point, we do not know the kinetics of glycogen levels over the course of their typical survival on H_2O_2, which averages 107h. Consistent with a role for glycogen in resistance to H_2O_2 food, flies fed supplemental maltose prior to being exposed to H_2O_2 showed increased resistance to H_2O_2 food in a dose-dependent manner (Fig 6). Maltose could increase survival by providing glucose, or perhaps by some other effect as a disaccharide. Supplementing fly diet with glucose but not lactose extends survival time similar to supplemental maltose, which is consistent with the former hypothesis (Fig 6).

The fat body is a site of glycogen storage in *Drosophila*, and RNAi targeting *glycogenin* (CG44244), the gene encoding the protein core of glycogen, in the fat body increased the sensitivity of flies to H_2O_2 food (Fig 7). We found that knocking down *glycogenin* using the S32 fat-body driver reduced survivorship, while the S106 fat-body driver did not (Fig 7). Glycogen is present in the *Drosophila* brain, and knocking down *glycogenin* with the neuron-specific elavGS driver reduced survival on H_2O_2 substantially (Fig 7) [48].

Folate Metabolism

Metabolites of the folate pathway were overrepresented in features associated with H_2O_2 resistance (see Additional File 1). The folate pathway is central to the synthesis of several amino acids, nucleotides, secondary metabolites and substrates for secondary modifications (e.g. methylation). We detected features corresponding to metabolites both in the folate pathway as well as in peripheral pathways (e.g. *S*-adenosylmethionine), suggesting that the activity of the folate pathway differs between sensitive and resistant flies. To investigate the potential role of the folate pathway in H_2O_2 resistance, we tested the effect of metabolic gene knockdown on survival. Knocking down either CG8665, which encodes 10-formyltetrahydrofolate dehydrogenase (FDH), or *pugilist* (CG4067), which encodes tetrahydrofolate dehydrogenase (THFDH), in the fat body or in neurons reduced the survival of flies on H_2O_2 food (Fig 7). In parallel experiments, RU486 pretreatment failed to affect survival in flies carrying the fat body 32GS driver. These data suggest that folate metabolism in the abdominal fat body and neurons is important for survival on H_2O_2 food.

Supplemental folic acid was shown to reduce the levels of oxidative damage associated with knockdown of *parkin* in *Drosophila* neurons [49]. We also tested whether supplemental folic acid would affect survival on H_2O_2 food and failed to see a significant effect (data not shown).

Discussion

Dissecting Variation for Complex Traits

Next-generation sequencing technology has provided geneticists with unprecedented power to identify single nucleotide variants associated with variation for complex traits in natural populations. But even with extremely large sample size [e.g., 20, 50, 51], the percent of variance explained by SNPs in most studies remains small [52-54], and current models suggest that thousands of SNPs can contribute to any one trait [53, 55]. Our modestly powered GWAS failed to detect SNPs that explain a large amount of the variation in H_2O_2 resistance. Similar to other GWAS studies in this population, our results leave open the possibility that this trait is quite polygenic in the DGRP [28, 56, 57]. The results presented here, however, point to the considerable potential of metabolic profiles to distinguish and predict genetically determined phenotypic variation, and moreover, to identify novel, causal molecular pathways associated with that variation. In light of our findings, we propose a model here whereby a very large number of interacting genetic loci [58] converge through a more limited number of downstream metabolic pathways, which in turn make up the functional and structural building blocks of complex traits (Fig 8).

Metabolomic Analysis

This study illustrates that untargeted metabolomic profiles, even those that include unknown chemical identities, give us tremendous power to 1) predict complex phenotypic variation; 2) identify novel pathways associated with this variation; and 3) in this particular study, suggest a novel hypothesis that resistance to stress might be caused by resistance of the metabolome to environmental perturbation.

The Metabolome as a Predictive Biomarker

First, we find that the untargeted metabolome when compared to genotype predicts phenotype well. While a PCA analysis of latent variation in the metabolome clearly separates sensitive from resistant flies, a similar analysis of allelic variants among genotypes failed to clearly separate genotypes based on their survival on H_2O_2 food (Fig 5). This contrast highlights the ‘proximity’ of the metabolome to trait on the genotype-phenotype map [4, 59].

Numerous other studies have shown metabolomic responses to diet, age, and temperature [33, 60-62], as well as body mass and body composition independent of diet [32], in diverse genotypes. However, these studies were not designed to test the relative power of genotype versus metabolome to predict phenotypic response. Moreover, studies that describe the effect of stress or environment on the metabolome often include only a single genotype and this may fail to resolve the systems-level association between genotype, environment, metabolome and phenotype [63-66]. In light of our results, future studies would benefit from a clearer characterization of the power of the metabolome relative to the genome to distinguish biologically relevant phenotypic variation.

Stress Resistance Pathways Identified by Metabolomics

Second, our metabolomic profiling suggests several possible mechanisms that underlie the phenotypic variation observed here. As we emphasize above in presenting the results, and discuss further here, in *Drosophila* studies, variation in the ability to survive oxidative stress could be confounded with starvation resistance. This is important to keep in mind as we discuss possible mechanisms that underlie the phenotypic variation seen here. Nonetheless, the ability of the metabolome to distinguish resistant and sensitive phenotypes is notable.

Several studies have used metabolomics to study the effects of oxidative stressors in *Drosophila* in a single genetic background. For example, paraquat treatment alters branched-chain amino acid, starch/sucrose, and fatty acid metabolism in the *Drosophila* brain [66]. We also detect significant effects of H_2O_2 food on the first two of these three pathways (see Additional File 1). The lack of evidence for an effect on the third pathway, fatty acid metabolism, is perhaps not surprising, given that our analysis was limited to aqueous metabolites. There are two important caveats to this earlier work as well as to the present study. First is the ambiguous nature of the untargeted metabolite data. For many features in the global metabolome profiles, we know mass/charge ratios but not chemical structure. Second, given the relationship between resistance to H_2O_2 and starvation (Fig 1), we expect that the metabolites associated with exposure to H_2O_2 or paraquat might relate to nutrient intake and/or storage, rather than oxidative stress alone. This confounding influence of H_2O_2 on feeding could explain the results of a variety of studies that use survival on food supplemented with H_2O_2 as a measure of oxidative stress resistance.

One way to disentangle the effects of oxidative stress from a secondary effect of a stressor administered in the diet on nutrient update is to genetically manipulate endogenous levels of reactive oxygen species. Two recent studies compared the metabolome of flies with null mutations in superoxide

dismutase (*sod*), to that of flies either chronically or acutely exposed to paraquat-supplemented diet [63, 64]. Each of these conditions had a distinct effect on the metabolome, with the *sod*$^{-}$ mutant metabolic profile clearly separating from the other conditions in unsupervised clustering [64]. Our analysis suggests that at least some of the differences between the *sod*$^{-}$ and paraquat-treated metabolome could be due to altered feeding or starvation in paraquat-treated flies.

Keeping these caveats in mind, we find strong evidence for two pathways associated with H_2O_2 resistance, glycogen metabolism and folate metabolism. Indeed, the decreased maltodextrin content in flies sensitive to H_2O_2 compared to control food suggests that sensitive flies are exhausting their glycogen pool in response to stress, which is consistent with previous studies of flies on food with paraquat or under starvation [37, 67, 68]. Maltodextrins are intermediates in the glycogen pathway, which is used to store and retrieve glucose for the glycolysis and pentose phosphate pathways (Fig S5). We found that supplemental maltose or glucose enhances survival on H_2O_2 food for all genotypes tested in our study and that RNAi targeting *glycogenin* either in the fat body or in neurons reduces survival on H_2O_2 food (Figs 6 and 7).

Interestingly, knockdown of *glycogenin* affected survival, but only with the S32 driver, which is expressed primarily in the head fat body, and not the S106 driver, which is expressed in the abdominal fat body. This suggests that *glycogenin* functions in the head fat body to mediate resistance to H_2O_2 food [69, 70]. However, we cannot rule out the effect of other differences between the knockdown of *glycogenin* by S32 compared to the S106 driver.

Mummichog also detected enrichment of the folate pathway (see Additional File 1). Several lines of evidence from this work and previous studies indicate a role for the folate pathway in stress resistance in *Drosophila* [29, 49, 71-76]. We show that knocking down either the folate pathway genes *pugilist* or CG8665 in either the fat body or neurons reduces survival on H_2O_2 food while not affecting survival on food without H_2O_2 (Fig 7 and data not shown). Several studies find that transcripts encoding enzymes of the folate pathway are up-regulated following either administration of paraquat or H_2O_2 in food, over-expression of manganese superoxide dismutase, or in mutants with mitochondrial dysfunction [73-78]. Additionally, two missense variants in *pugilist* associate with paraquat resistance in the DGRP [29]. Functional studies show that overexpression of *Drosophila nmdmc*, which encodes methylenetetrahydrofolate dehydrogenase, enhances resistance to paraquat [72], and that folic acid supplementation reduces oxidative damage to lipids and endogenous H_2O_2 levels associated with knockdown of *parkin* in *Drosophila* neurons [49]. While our metabolomic data and previous genetic

studies [71] point to a role for folate metabolism, in our own GWAS analysis, neither *pugilist* nor CG8665 were clearly associated with survival on H_2O_2 (P_{min} > 0.002 in both cases).

Adaptive Shifts Versus Robustness of the Metabolome Under Stress

Third, it appears that the metabolome responds to the presence of H_2O_2 more strongly in sensitive than in resistant lines (Figs 4 and S4). One might have expected the opposite pattern, whereby greater resistance is associated with an *adaptive* response of the metabolome to an environmental stressor [e.g., 79, 80, 81]. Instead, it appears that resistance in this study is associated with the ability to maintain the metabolome in its current state. This observation is made possible by having multiple sets of stress-resistant and stress-sensitive genotypes. We interpret these results as suggesting not only that resistance to H_2O_2 food is explained by the metabolome, but also that the metabolome is more *robust* in resistant genotypes, being less likely to change when faced with an external stressor. We do not propose that metabolic robustness is a universal or causal feature of stress resistance, but our data suggest that metabolic robustness and/or resilience might contribute to phenotypic variation in nature.

This observation, combined with the fact that the metabolome predicts phenotype while the genotype does not, suggests the hypothesis that there are genetically diverse ways to achieve resistance to H_2O_2, but that these diverse genetic paths converge at a common metabolome associated with resistance. In future studies it would be worth asking if, across a broad range of phenotypes, there is a correlation between stress resistance and the 'resistance' of the metabolome to stress-induced alteration, or *vice versa* for traits that require metabolic adaptation.

Genomic Analysis

We studied survival on H_2O_2 food in the DGRP, a population representing a sample of natural genetic variation. We estimate broad-sense heritability for mean lifespan under H_2O_2 at H^2 = 51%. This approach likely over-estimates the genetic variance component because it includes some among-line environmental

variance, including variation in the microbiome. Mapping variants associated with mean survival revealed 14 SNPs that were significant and together these variants define 7 loci (Fig 2). We limited our analysis to variants passing MAF and missing genotype filters among 180 DGRP lines. While the DGRP incorporates substantial allelic diversity from the wild, it is a population of inbred lines with different heterozygosity compared to its parent wild population and that has also been purged of deleterious alleles of strong effect [56]. Along with the missing genotypes there are also uncharacterized structural variants within this population [22]. For these reasons, we do not expect this study to identify all variants in the DGRP associated with survival under H_2O_2 stress, nor their mode of gene action. Similar to other studies with the DGRP, we failed to find common alleles with large and highly significant effects, suggesting that variation in survival on H_2O_2 food is influenced by many loci of small effect [56, 57]. It is possible that models incorporating alleles that failed to reach genome-wide significance could explain more of the heritable variation in this study [55]. However, in the relatively small population used for metabolomics, neither clustering nor multivariate analysis of genotypes predicted the discrete resistance trait (Fig 5).

Although we hoped to pinpoint specific genes that influence H_2O_2 resistance, we face the statistical challenge that among the DGRP lines we studied, the number of variants per gene ranged from 1 to 4490 with a mean of 237. To overcome the increased risk of false positives in genes that contain a large number of polymorphisms, we used a permutation approach to measure association between genes and phenotype. Our approach, like other methods, comes with caveats, one being the imperfect annotations of variants and genes. We have limited this analysis to only those variants associated with the FlyBase gene annotation, including 1 kb upstream and 1 kb downstream of the primary transcript [22]. Intergenic variants might affect the expression of trait-associated genes. However, we did not attempt to account for those effects in this study. Additionally, variants that are associated with candidate genes can instead exert their effect on phenotype by modifying the expression of other local genes, and this would result in misattributing the significance to the gene containing the variant rather than the real trait-associated gene [82].

Stress Resistance Pathways Identified by Genetic Association

Our genomic analysis led to several interesting and potentially related genes associated with H_2O_2 resistance, including *ush, NPF* and the pickpocket paralogs *ppk7* and *ppk14*. Each of these genes is implicated in several processes and though we do not rule out causal associations between these processes and resistance to H_2O_2 food, based on published studies, we argue that genetic variation in these genes influences feeding and/or metabolism to explain their effect on H_2O_2 resistance (Fig S7).

As we show, flies on H_2O_2 food substantially reduce food intake, and their lifespans appear to be a function of starvation (Fig 1). Interestingly, feeding behavior is controlled by neuronal signaling involving several of these candidate genes. Larvae require *NPF* to signal the intake of noxious food under starvation conditions [83], and NPF-expressing neurons in adults couple hunger to memory performance [84]. Interestingly, IIS in neurons expressing the NPF receptor repress larval feeding, suggesting that NPF controls feeding through neuronal IIS [83]. We postulate that the candidate *ush* influences survival on H_2O_2 food as a peripheral regulator of IIS [44]. This appears to be a function directly attributed to the USH protein, as its human homolog FOG has been shown to directly bind and inhibit the PI3K complex [44]. Inactivation of PI3K leads to depletion of nutrient stores in the fat body and its constitutive activation reduces both nutrient stores and survival under starvation [85]. PI3K is also a negative regulator of FOXO in the IIS pathway, which has well characterized roles in the response to starvation and to oxidative stressors in food [69, 86-89]. Interestingly, ablation of insulin-like peptide-producing cells in *Drosophila* increases survival both on food containing paraquot and under starvation, alters whole body levels of glycogen, and leads to misregulation of metabolic genes in the glycogen pathway, which could link the effect of NPF and *ush* polymorphism to the variation we see in glycogen metabolism [90, 91].

The effects of two other candidates, *ppk7* and *ppk14*, might also be linked to feeding and metabolism through their potential role in nutrient signaling. *ppk7* and *ppk14* are members of the pickpocket gene family that are expressed in neurons and signal taste cues, modulate feeding, and may influence energy metabolism [92-95]. Interestingly, *Drosophila ppk28* was recently shown to interact with glucagon-like hormone (AKH) signaling, a pathway involved in regulating glycogen metabolism in flies [91, 96]. Together, the genetic analysis of H_2O_2 resistance has revealed pathways whose role in survival could be explored in future studies.

Starvation as a confounding factor in stress assays

Our study used H_2O_2-supplemented food as the stressor. Oxidative stressors such as H_2O_2, paraquat, or menadione are often administered to *Drosophila* by supplementing the diet, and each of these treatments dramatically shortens survival in a dose-dependent manner [29, 87]. After screening the DGRP for survival on H_2O_2 food, we noticed that these survival times correlated closely with survival times under both paraquat and menadione exposure, and even more so with survival times measured under starvation [26].

Feeding is essential to the survival of adult *Drosophila*, and feeding behavior is influenced by a variety of cues, including food acidity or the presence of bitter compounds, hypoxia, and the nutrient content of food [97-100]. While some stressors affect the preference for food of particular composition [98, 101], others may alter feeding behavior by affecting satiety [97]. We show that flies consume very little food

containing 2% H_2O_2. This effect is not limited to H_2O_2, as many supplements have been found to reduce feeding in flies [97], and this may extend to other oxidative stressors including paraquat and menadione. While several studies have detected reduced feeding in response to paraquat-containing food [102, 103], others detected no difference [68]. Contrary to the latter study, our data suggest that feeding is significantly reduced in the three genotypes tested in assays that measured feeding over either 2h or 24h (Fig 1). Several differences in the experimental setting may explain the discrepancy between this study and Riahi *et al*. (2019) [68].

It is possible that the effect we see relates to oxidative stress and not an aversion to food supplementation, as deviation from normoxia alters feeding behavior of *Drosophila* larvae in a manner that appears to rely on H_2O_2-sensitive neurons [104]. Also, the UV light-avoidance of egg-laying females appears to signal through H_2O_2-sensitive taste receptors in *Drosophila* [105], and H_2O_2 also inhibits feeding in *Caenorhabditis elegans* through taste receptors [106]. Also, two recent studies found contradictory roles for the histone methyltransferase G9a in survival in response to oxidative stressors and starvation, indicating that environment and genetic background may affect stress-response in *Drosophila,* but also suggesting that starvation and oxidative stress resistance may share underlying biological pathways [37, 68]. These studies suggest that susceptibility to oxidative stress and starvation are partially separable, however they do not rule out a main effect of H_2O_2 on survival due to starvation. The relationship between lifespan under starvation and survival on H_2O_2 food has implications for studies that draw conclusions about stress resistance in response to agents administered in food. The effects of altered feeding patterns or nutrient deprivation should be accounted for when analyzing the effects of stressors or drugs administered in the diet.

Conclusion

- **Genetic variation in a complex trait converges on the metabolome**

The sample of genotypes in this study show a consistent metabolic signature associated with their phenotype. Thus a potentially wide degree of genotype space may converge on a smaller number of metabolic pathways to shape phenotype.

- **Metabolome robustness associates with stress resistance**

Contrary to the metabolic change that might be expected in animals resisting stress, we find that the metabolome of resistant animals appears robust to stress treatment. This suggests that maintaining metabolism in the presence of certain stressors is a means of survival.

- **Glycogen and folate metabolism and several genes involved in nutrient signaling mediate resistance to peroxide food**

Genetic and metabolomic analysis of peroxide resistance revealed roles for glycogen and folate metabolism and genes with known roles in nutrient signaling. Future studies to understand this network may reveal novel mechanisms of stress resistance.

- **Starvation explains the lifespan response to peroxide food**

Drosophila are sensitive to diet, including additives and food-borne treatments. We show that the response to H_2O_2 in food can be explained by starvation. These effects may dramatically confound assays that examine responses to treatments delivered by supplementing the *Drosophila* diet.

Methods

Genetic Stocks

Drosophila Genome Reference Panel (DGRP) lines were obtained from the Bloomington *Drosophila* Stock Center (BDSC). Genes encoding enzymes involved in the glycogen or folate pathways were identified using the Kyoto Encyclopedia of Genes and Genomes database (http://www.genome.jp/kegg/pathway.html). The expression of GWAS candidates or genes encoding enzymes in the glycogen or folate pathways were manipulated using the GAL4 GeneSwitch/UAS system [45]. The drivers S106 (BDSC #8151), S32 (BDSC #8527) or elavGS (BDSC #43642) were crossed to flies carrying UAS-RNAi transgenes targeting candidate genes: *ush* (CG2762, BDSC #32950 (*ush_44*) and BDSC #44014 (*ush_57*)), *Glycogenin* (CG44244, BDSC #42565), CG8665 (BDSC #62266), and *pugilist* (CG4067, BDSC #42950), or the attP2 background control (BDSC #8622). The NPF^{SK1} allele is a 179bp deletion within the coding region created by CRISPR [107]. NPF^{SK1} was backcrossed at least six times into the Canton-S background prior to testing. Stocks were maintained on standard cornmeal-sugar-yeast food at 25°C on a 12/12h light/dark cycle at 50-70% RH.

Media

Standard food was made by cooking 12g Drosophila agar (type II, Genesee Scientific, El Cajon, CA), 25g brewers yeast (MP Biomedicals, Solon, OH), 55g glucose monohydrate (MP Biomedicals), 30g sucrose, 60g corn meal, 3g methylparaben (Genesee Scientific). 12g 100% ethanol (Decon Labs, King of Prussia, PA), and 3g propionic acid (Fisher Scientific, Pittsburg, PA) per liter of water. A small amount of dry active yeast was sprinkled onto standard food prior to use.

Peroxide food was made in one of two ways, for 2% food, agar was melted into 2% glucose monohydrate and 0.3% propionic acid and, after the food had cooled to less than 60°C, 30% H_2O_2 (Fisher Scientific) was added to reach 2%, or the same volume of water was added for the control food. For the 3% H_2O_2 food, the recipe was the same with the exception that 30% H_2O_2 was added to reach 3% H_2O_2. Approximately 5mL of food was dispensed into 25mm wide x 95mm tall polystyrene vials.

Starvation food was made by melting 2% agar into 0.3% propionic acid and dispensing into vials.

Food supplemented with carbohydrates was made by adding 2% of either D-(+)-maltose, β-lactose (both from Sigma, St. Louis, MO), or additional glucose to the 2% glucose control food.

RU486 food was made by overlaying ~5mL standard food with either 50uL of 25mg mL^{-1} RU486 (mifepristone, Cayman Chemical Company, Ann Arbor, MI) dissolved in 100% ethanol or the same volume of 100% ethanol alone for the -RU control food. Ethanol was allowed to evaporate overnight at 22 to 24°C prior to using the food.

Survival Assays

To measure the variation in resistance to oxidative stress across a lab population, we measured the survival of mated females from 180 DGRP lines in a multi-block design on H_2O_2 food [26]. For each block, flies were raised under low-density conditions by allowing ~50 flies to lay eggs for one day on standard food in bottles. Flies for the assay were then collected over two or three days and then allowed to mate for 24h on standard food. Ten 1 to 3 day-old mated females were then placed in each vial containing H_2O_2 food. Five vials of H_2O_2 food and one or two vials of control food without H_2O_2 were included for each genotype in each trial. Control vials without H_2O_2 were included to confirm that mortality was due to H_2O_2. In knockdown experiments, we included 5 to 8 H_2O_2 vials and 5 to 8 control vials to ensure that any effect of gene knockdown on survival in the absence of H_2O_2 could be measured. For assays involving

the Drosophila Activity Monitor System (DAMS, TriKinetics Inc, Waltham, MA), the activity of 38 to 48 individual flies per genotype was recorded simultaneously every minute over the experiment.

We ran 17 blocks with a mean of 14.5 lines (range = 4 to 35 lines) per block. We used 2% H_2O_2 food for the first 10 blocks and 3% H_2O_2 was used for the last seven blocks. The switch between 2% and 3% H_2O_2 was made accidentally and was realized after the conclusion of the study. We used the following mixed model in the lme4 package to test for an effect of these two food treatments:

$\log_e$ lifespan ~ food + weight + (1 | block) + (1 | genotype) + ϵ.

where lifespan was the $\log_e$ of mean lifespan of a genotype in a block, food and weight were fixed effects, and block and genotype were both random effects along with the error term epsilon. We found no difference between 2% and 3% H_2O_2 doses on lifespan (β = -9.1x10^{-4}, P = 0.988). In all blocks, dead flies in each vial were recorded two to four times per day using D-life software until all H_2O_2-treated flies had died [23]. The survival data were scaled for each block to remove the variance between blocks (see Statistical Analysis). To measure line weights, 24h after beginning a lifespan assay, flies from an extra control food vial were frozen at -80°C and were later collectively weighed on a microbalance (XS105, Mettler Toledo, Columbus, OH).

To compare lifespans on H_2O_2 food to lifespans during starvation, 2 to 5 replicates of twenty 3-to-5 day-old mated females were assayed using D-life on agar food either with or without 2% glucose (see Media). The lifespan of each line on H_2O_2 food was measured twice in separate trials for this comparison, while the lifespan under starvation was measured once. To measure the effect of supplemental carbohydrates on lifespan ten replicates of ten 1-to-3 day-old mated females per genotype were allowed to feed on supplemented food for four days and then transferred without anesthesia to 2% H_2O_2 or control food to assay survival.

Genetic Manipulation

F_1 GAL4/UAS flies were collected over four days (day 0 to 3). These flies were allowed to mate for 24h, at which time they were anesthetized and sexed, and females (ten per vial) were then allowed to feed for two days on RU486 or -RU food. After 48h on RU486 or -RU food, flies were transferred without anesthesia to H_2O_2 or control food to measure survival. Negative genetic controls included F_1 GAL4/attP flies which

were crosses of the GAL4 driver to either the attP2 or attP40 lines from the Transgenic RNAi Project (TRiP) collection, where attP is the empty P-element docking site for the UAS transgenes in the TRiP collection (http://fgr.hms.harvard.edu/fly-in-vivo-rnai). Negative genetic controls were raised, induced and assayed in parallel with experimental flies.

Feeding Assays

To measure feeding rate, we used both dye incorporation and CAFE assays. For both assays, flies were allowed to mate for 24h and then separated sexes over light CO_2 anesthesia and transferred to agar-only food for 24h of starvation. For dye incorporation, after starvation, flies were immediately transferred without anesthesia into vials that contained either H_2O_2 or control food with 2.5% FD&C Blue Dye #1 (Spectrum Chemicals, Gardena, CA). After 2h on dye-containing food, flies were flash frozen in liquid N_2, homogenized in water, centrifuged at 16,000rcf for 1 minute, and the absorbance of the supernatant was measured at 630nm. The absorbance of each sample was normalized by dividing by the number of flies in the sample (n = 7 to 11 flies per sample).

For the CAFE assay, ten replicates of 10 mated females were starved for 24h and transferred without anesthesia to assay chambers. Assay chambers were 15mL conical bottom polystyrene tubes (Corning Inc, Corning, NY) containing water under a foam partition to maintain humidity but not allow flies to drink, and fitted with a 0.75mm ID glass capillary (World Precision Instruments, Sarasota, FL) which had been filled with 2% glucose, 0.3% propionic acid supplemented with either 2% H_2O_2 or water. Flies were housed in the assay chambers at 25°C on a 12/12h D/L cycle in an incubator at 60-70% RH for 24h before the volume of food consumed was assayed by measuring the difference in height of the top of the liquid food in the capillary and multiplying by $\pi \cdot 0.375mm^2$.

Metabolomic Analysis

Eight of the resistant and eight of the sensitive DGRP lines were selected based on their lifespans and line weights to reduce the effect of fly size on resistance. Lifespan for these lines on H_2O_2 food was again measured in a single block, and 24h after exposure to the H_2O_2 or control food, 5 flies per replicate were collected, flash frozen in liquid nitrogen, and then stored at -80°C. Each *Drosophila* sample was weighed and then homogenized in 200µL water with PBS in a microfuge tube immersed in an ice bath. Methanol (800 µL) was then added, followed by vortexing for 2 min and incubation at -20°C for 30 min to precipitate proteins. Samples were sonicated in an ice bath for 10 min and then centrifuged at 17,000rcf for 5 min at 4°C. From each tube, 900µL supernatant was transferred to a new microfuge tube for drying under vacuum at 30°C (~3h). The completely dried samples were reconstituted in 100µL 40% water/60%

HPLC-grade acetonitrile (ACN, Fisher Scientific) for liquid chromatography-mass spectroscopy (LC-MS) analysis. A pooled quality control (QC) sample was made by combining ~5μL aliquots from each reconstituted sample. The QC was analyzed once for every ten study samples to serve as a technical replicate throughout the data set to assess process reproducibility and allow for data normalization to account for any instrument drift.

LC-MS analysis was performed using an LC-QTOF-MS system (Agilent Technologies, Santa Clara, CA) consisting of an Agilent 1200 SL liquid chromatography system coupled online with an Agilent 6520 time-of-flight mass spectrometer. A 5μL aliquot of reconstituted sample was injected onto a 2.1×150mm Waters BEH-Amide 2.5μm particle column at 35°C. The metabolites were gradient-eluted at 0.3mL/min using mobile phase A, 5mM ammonium formate (Sigma) and 0.0125% formic acid (Sigma) in 97% water/3% ACN, and mobile phase B, 5mM ammonium formate and 0.0125% formic acid in 3% water/97% ACN (98% B for 1 min, 98% to 77% B in 6.5 min, 77% to 39% B in 4.5 min and 39% B for 7 min). The MS interface capillary was maintained at 325°C with a nebulizing gas pressure of 45psig, and a drying gas flow of 9L/min. The capillary voltage for positive ion injection was 3.5kV. Agilent MassHunter Workstation Data Acquisition software B.02.01 (B2116.30) was used to acquire all data from 60 to 1000m/Z using centroid mode with a threshold of 200 or 0.01%. LC-MS data was processed using XCMS online (version 2.2.5) and a list of ion intensities for each detected peak was generated using a retention time index and m/z data as the identifiers for each ion.

Statistical Analysis

Demography

Data were collected in D-life and all calculations were performed in R [23, 108]. Mean lifespan was estimated from H_2O_2 assays using the restricted mean (default settings) in the Kaplan-Meier model with the survival package [109]. To correct for block effects, lifespans were centered (mean = 0) and scaled to unit variance (SD = 1). For lines measured in more than one block, we calculated the mean block-centered lifespans across blocks. We then used residuals from linear regression of lifespan versus weight as our measure of H_2O_2 resistance.

Genetic Analysis

We calculated broad-sense heritability estimates within each of the nine blocks in which at least 18 randomly chosen DGRP lines were included. For each block, heritability was estimated by: V_L / V_L + V_E,

where V_L is the among-line variance and V_E is the average within-line variance [110]. Mean heritability and its variance were calculated from these nine estimates.

We use stepwise regression in the MASS package to identify significant covariates among the chromosomal inversions and Wolbachia status [22]. *In(2L)t* was the only significant covariate (P = 0.0277, ANOVA) and residuals from the linear regression of H_2O_2 resistance on *In(2L)t* genotype were used in a linear model of SNP-phenotype associations in PLINK [41]. Approximately 1.93 million SNPs with a MAF ≥0.05 and < 30% missing genotypes were tested for association with H_2O_2 resistance from 180 DGRP lines using a linear model. To account for population structure, we used the Tracy-Windom test in the AssocTests package to evaluate eigenvalues from 20 PCs of genotype, and retained the first four PCs as covariates in the model (α = 0.05, [111]). Genome-wide significance was determined by controlling for FDR at 0.2 using the *q* value method [42].

Gene-level associations with H_2O_2 resistance were derived using permutation to evaluate the significance of the 4810 genes that had at least one variant associated with phenotype with $P \leq 0.01$. We refer to SNPs within 1kb upstream and 1kb downstream of the gene model in FB release 5.49 as associated with that gene [22, 112]. In the initial analysis, 10,000 permutations of phenotype were performed, and the association between phenotype and each SNP in a gene was tested in the same linear model used for SNP-phenotype associations described above, including genotype PCs as covariates. The maximum test statistic (T_{max}) among the SNPs in each gene from the real GWAS was compared to the T_{max} for each of 10,000 permutations to derive a one-tailed empirical P-value for each gene (P_{gene}). The sets of SNPs in genes with $P_{gene} \leq 0.01$ were then subjected to 1 million more permutations and the resulting P_{gene} values were used as our measure of gene-trait association. We use *q* values to estimate the FDR for each P_{gene}. Over-representation by biological process and pathways was tested using Fisher's exact test in PANTHER version 13.1 and the GO-slim subset of biological processes and PANTHER pathways among the 13,788 gene models in *Drosophila melanogaster* [113].

Metabolomic Analysis

Global LC-MS provided measures of 3028 features from positive mode and 2921 features from negative mode in a total of 93 samples. Principal component analysis (PCA) was performed on 2270 and 2417 features with no missing values in positive and negative mode, separately. The $\log_2$-transformed feature abundance was centered and scaled prior to PC using the prcomp R function. The separation of samples according to treatment, trait or both was examined across the first ten principal components (PCs) and those components with the clearest separation are shown.

For univariate analysis, we fit logistic models for each metabolomic feature, with the log odds of being a resistant genotype as the outcome variable, and the $\log_2$ feature abundance as the independent variable. This was done for control and H_2O_2-treated groups separately, controlling for the potentially confounding effect of line weight:

$$Log(\frac{P}{1-P})$$

~ weight + $\log_2$ (feature).

Weight was also included in linear modeling of both main and interaction effects of treatment and trait to look for metabolites that associate with trait, treatment, or their interaction:

Log_2 (feature) ~ treatment + trait + treatment*trait + weight + ϵ.

To identify specific metabolic pathways associated with resistance to H_2O_2 food, we carried out a metabolomic equivalent of gene set enrichment analysis with *mummichog* version 1.05 (Subramanian *et al*., 2005, Li *et al*. 2013). The BioCyc *D. melanogaster* (version 16.5) metabolic model was used to test the enrichment of pathways and networks. Separate analyses were performed for the features fitted in the logistic or linear regression models at an FDR cutoff of < 0.1 [114]. One hundred permutations of the data were performed by *mummichog* to estimate the null distribution.

Declarations

Abbreviations

H_2O_2: hydrogen peroxide

GWAS: genome-wide association studies

DGRP: *Drosophila* genetic reference panel

IIS: insulin-like signaling

MAF: minor allele frequency

FDR: false discovery rate

LD: linkage disequilibrium

SNP: single nucleotide polymorphism

GO: gene ontology

RNAi: RNA interference

PCA: principal component analysis

PC: principal component

LC-MS: liquid chromatography mass spectrometry

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

LC-MS data collected in this study are available at MetaboLights [115]:

www.ebi.ac.uk/metabolights/MTBLS1190. All other data and R code are available on GitHub: https://github.com/ben6uw/Harrison-et-al-2019-Data-Repository

Competing interests

1. Promislow receives consultancy fees as a member of the External Advisory Board for the WALTHAM Center for Pet Nutrition, and is a co-founder of Fido-Gen, Inc. None of the other authors have competing interests.

Funding

This work was supported by grants NSF DMS1561814 to DP, and NIH grants AG049494 to DP and AG013280 to SP and DP. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author's contributions

BH designed the experiments, acquired and interpreted data, and wrote the manuscript

LW interpreted the data, and wrote the manuscript

EG acquired the H_2O_2 resistance data across the DGRP

EH acquired the feeding rate and supplemental carbohydrate data

BC acquired and interpreted the data for the *NPF*^{SK1} mutant

SP interpreted the data for the *NPF*^{SK1} mutant

DR designed the LC-MS experiments

DP designed the experiments, interpreted the data and wrote the manuscript

Acknowledgments

We thank: Nick Force, William Gordon, Kelly Jiang, Ijay Okeke, Sarah Promislow, Allie Shiff, Erin Tudor, and Deborah Xi for their assistance in measuring resistance to H_2O_2 food across the DGRP. Ming Yang wrote the script to run gene-level permutations. Xiaqing Zhao measured resistance to H_2O_2 food and starvation. Shu Kondo generously provided the *NPF*^{SK1} allele. Elizabeth King, Clement Chow, James MacDonald and Theo Bammler reviewed the manuscript.

Aaron Rosen, Gabriel Endicott and Haley Dixon diligently maintained fly stocks.

References

1. Boyle EA, Li YI, Pritchard JK: **An Expanded View of Complex Traits: From Polygenic to Omnigenic.** *Cell* 2017, **169:**1177-1186.
2. Patti GJ, Yanes O, Siuzdak G: **Innovation: Metabolomics: the apogee of the omics trilogy.** *Nat Rev Mol Cell Biol* 2012, **13:**263-269.
3. Johnson CH, Ivanisevic J, Siuzdak G: **Metabolomics: beyond biomarkers and towards mechanisms.** *Nat Rev Mol Cell Biol* 2016, **17:**451-459.
4. Zampieri M, Sauer U: **Metabolomics-driven understanding of genotype-phenotype relations in model organisms.** *Current Opinion in Systems Biology* 2017, **6:**28-36.
5. Handakumbura PP, Stanfill B, Rivas-Ubach A, Fortin D, Vogel JP, Jansson C: **Metabotyping as a Stopover in Genome-to-Phenome Mapping.** *Sci Rep* 2019, **9:**1858.
6. Fu J, Keurentjes JJ, Bouwmeester H, America T, Verstappen FW, Ward JL, Beale MH, de Vos RC, Dijkstra M, Scheltema RA, et al: **System-wide molecular evidence for phenotypic buffering in Arabidopsis.** *Nat Genet* 2009, **41:**166-167.
7. Parker BL, Calkin AC, Seldin MM, Keating MF, Tarling EJ, Yang P, Moody SC, Liu Y, Zerenturk EJ, Needham EJ, et al: **An integrative systems genetic analysis of mammalian lipid metabolism.** *Nature* 2019, **567:**187-193.
8. Wang Z, Klipfell E, Bennett BJ, Koeth R, Levison BS, Dugar B, Feldstein AE, Britt EB, Fu X, Chung YM, et al: **Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease.** *Nature* 2011, **472:**57-63.
9. Hoffman JM, Soltow QA, Li S, Sidik A, Jones DP, Promislow DE: **Effects of age, sex, and genotype on high-sensitivity metabolomic profiles in the fruit fly, Drosophila melanogaster.** *Aging Cell* 2014, **13:**596-604.
10. Mahieu NG, Patti GJ: **Systems-Level Annotation of a Metabolomics Data Set Reduces 25000 Features to Fewer than 1000 Unique Metabolites.** *Anal Chem* 2017, **89:**10397-10406.
11. Fuhrer T, Zamboni N: **High-throughput discovery metabolomics.** *Curr Opin Biotechnol* 2015, **31:**73-78.
12. Sevin DC, Sauer U: **Ubiquinone accumulation improves osmotic-stress tolerance in Escherichia coli.** *Nat Chem Biol* 2014, **10:**266-272.
13. Cirulli ET, Guo L, Leon Swisher C, Shah N, Huang L, Napier LA, Kirkness EF, Spector TD, Caskey CT, Thorens B, et al: **Profound Perturbation of the Metabolome in Obesity Is Associated with Health Risk.** *Cell Metab* 2019, **29:**488-500.e482.

14. Williams S, Dew-Budd K, Davis K, Anderson J, Bishop R, Freeman K, Davis D, Bray K, Perkins L, Hubickey J, Reed LK: **Metabolomic and Gene Expression Profiles Exhibit Modular Genetic and Dietary Structure Linking Metabolic Syndrome Phenotypes in Drosophila.** *G3 (Bethesda)* 2015, **5:**2817-2829.
15. Shin SY, Fauman EB, Petersen AK, Krumsiek J, Santos R, Huang J, Arnold M, Erte I, Forgetta V, Yang TP, et al: **An atlas of genetic influences on human blood metabolites.** *Nat Genet* 2014, **46:**543-550.
16. Adamski J, Suhre K: **Metabolomics platforms for genome wide association studies–linking the genome to the metabolome.** *Curr Opin Biotechnol* 2013, **24:**39-47.
17. Wen W, Li K, Alseekh S, Omranian N, Zhao L, Zhou Y, Xiao Y, Jin M, Yang N, Liu H, et al: **Genetic Determinants of the Network of Primary Metabolism and Their Relationships to Plant Performance in a Maize Recombinant Inbred Line Population.** *Plant Cell* 2015, **27:**1839-1856.
18. Suhre K, Shin SY, Petersen AK, Mohney RP, Meredith D, Wagele B, Altmaier E, Deloukas P, Erdmann J, Grundberg E, et al: **Human metabolic individuality in biomedical and pharmaceutical research.** *Nature* 2011, **477:**54-60.
19. Wu S, Tohge T, Cuadros-Inostroza A, Tong H, Tenenboim H, Kooke R, Meret M, Keurentjes JB, Nikoloski Z, Fernie AR, et al: **Mapping the Arabidopsis Metabolic Landscape by Untargeted Metabolomics at Different Environmental Conditions.** *Mol Plant* 2018, **11:**118-134.
20. Klarin D, Damrauer SM, Cho K, Sun YV, Teslovich TM, Honerlaw J, Gagnon DR, DuVall SL, Li J, Peloso GM, et al: **Genetics of blood lipids among ~300,000 multi-ethnic participants of the Million Veteran Program.** *Nat Genet* 2018, **50:**1514-1523.
21. Rinschen MM, Ivanisevic J, Giera M, Siuzdak G: **Identification of bioactive metabolites using activity metabolomics.** *Nat Rev Mol Cell Biol* 2019, **20:**353-367.
22. Huang W, Massouras A, Inoue Y, Peiffer J, Ramia M, Tarone AM, Turlapati L, Zichner T, Zhu D, Lyman RF, et al: **Natural variation in genome architecture among 205 Drosophila melanogaster Genetic Reference Panel lines.** *Genome Res* 2014, **24:**1193-1208.
23. Linford NJ, Bilgir C, Ro J, Pletcher SD: **Measurement of lifespan in Drosophila melanogaster.** *J Vis Exp* 2013.
24. Sohal RS, Orr WC: **The redox stress hypothesis of aging.** *Free Radic Biol Med* 2012, **52:**539-555.
25. Pickering AM, Vojtovich L, Tower J, KJ AD: **Oxidative stress adaptation with acute, chronic, and repeated stress.** *Free Radic Biol Med* 2013, **55:**109-118.
26. Mackay TF, Richards S, Stone EA, Barbadilla A, Ayroles JF, Zhu D, Casillas S, Han Y, Magwire MM, Cridland JM, et al: **The Drosophila melanogaster Genetic Reference Panel.** *Nature* 2012, **482:**173-178.
27. Mackay TF: **The nature of quantitative genetic variation for Drosophila longevity.** *Mech Ageing Dev* 2002, **123:**95-104.
28. Ivanov DK, Escott-Price V, Ziehm M, Magwire MM, Mackay TF, Partridge L, Thornton JM: **Longevity GWAS Using the Drosophila Genetic Reference Panel.** *J Gerontol A Biol Sci Med Sci* 2015, **70:**1470-1478.

29. Weber AL, Khan GF, Magwire MM, Tabor CL, Mackay TF, Anholt RR: **Genome-wide association analysis of oxidative stress resistance in Drosophila melanogaster.** *PLoS One* 2012, **7:**e34745.
30. Jordan KW, Craver KL, Magwire MM, Cubilla CE, Mackay TF, Anholt RR: **Genome-wide association for sensitivity to chronic oxidative stress in Drosophila melanogaster.** *PLoS One* 2012, **7:**e38722.
31. King EG, Macdonald SJ, Long AD: **Properties and Power of the Drosophila Synthetic Population Resource for the Routine Dissection of Complex Traits.** *Genetics* 2012, **191:**935.
32. Reed LK, Lee K, Zhang Z, Rashid L, Poe A, Hsieh B, Deighton N, Glassbrook N, Bodmer R, Gibson G: **Systems genomics of metabolic phenotypes in wild-type Drosophila melanogaster.** *Genetics* 2014, **197:**781-793.
33. Hariharan R, Hoffman JM, Thomas AS, Soltow QA, Jones DP, Promislow DE: **Invariance and plasticity in the Drosophila melanogaster metabolomic network in response to temperature.** *BMC Syst Biol* 2014, **8:**139.
34. Malmendal A, Overgaard J, Bundy JG, Sorensen JG, Nielsen NC, Loeschcke V, Holmstrup M: **Metabolomic profiling of heat stress: hardening and recovery of homeostasis in Drosophila.** *Am J Physiol Regul Integr Comp Physiol* 2006, **291:**R205-212.
35. Overgaard J, Malmendal A, Sorensen JG, Bundy JG, Loeschcke V, Nielsen NC, Holmstrup M: **Metabolomic profiling of rapid cold hardening and cold shock in Drosophila melanogaster.** *J Insect Physiol* 2007, **53:**1218-1232.
36. Coquin L, Feala JD, McCulloch AD, Paternostro G: **Metabolomic and flux-balance analysis of age-related decline of hypoxia tolerance in Drosophila muscle tissue.** *Mol Syst Biol* 2008, **4:**233.
37. An PNT, Shimaji K, Tanaka R, Yoshida H, Kimura H, Fukusaki E, Yamaguchi M: **Epigenetic regulation of starvation-induced autophagy in Drosophila by histone methyltransferase G9a.** *Sci Rep* 2017, **7:**7343.
38. Everman ER, Morgan TJ: **Antagonistic pleiotropy and mutation accumulation contribute to age-related decline in stress response.** *Evolution* 2018, **72:**303-317.
39. Ja WW, Carvalho GB, Mak EM, de la Rosa NN, Fang AY, Liong JC, Brummel T, Benzer S: **Prandiology of Drosophila and the CAFE assay.** *Proc Natl Acad Sci U S A* 2007, **104:**8253-8256.
40. Wong R, Piper MD, Wertheim B, Partridge L: **Quantification of food intake in Drosophila.** *PLoS One* 2009, **4:**e6063.
41. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC: **PLINK: a tool set for whole-genome association and population-based linkage analyses.** *Am J Hum Genet* 2007, **81:**559-575.
42. Storey JD, Tibshirani R: **Statistical significance for genomewide studies.** *Proc Natl Acad Sci U S A* 2003, **100:**9440-9445.
43. Mirina A, Atzmon G, Ye K, Bergman A: **Gene size matters.** *PLoS One* 2012, **7:**e49093.
44. Hyun S, Lee JH, Jin H, Nam J, Namkoong B, Lee G, Chung J, Kim VN: **Conserved MicroRNA miR-8/miR-200 and its target USH/FOG2 control growth by regulating PI3K.** *Cell* 2009, **139:**1096-1108.

45. Roman G, Endo K, Zong L, Davis RL: **P[Switch], a system for spatial and temporal control of gene expression in Drosophila melanogaster.** *Proc Natl Acad Sci U S A* 2001, **98:**12602-12607.
46. Fadda M, Hasakiogullari I, Temmerman L, Beets I, Zels S, Schoofs L: **Regulation of Feeding and Metabolism by Neuropeptide F and Short Neuropeptide F in Invertebrates.** *Front Endocrinol (Lausanne)* 2019, **10:**64.
47. Li S, Park Y, Duraisingham S, Strobel FH, Khan N, Soltow QA, Jones DP, Pulendran B: **Predicting network activity from high throughput metabolomics.** *PLoS Comput Biol* 2013, **9:**e1003123.
48. Yamada T, Habara O, Kubo H, Nishimura T: **Fat body glycogen serves as a metabolic safeguard for the maintenance of sugar levels in Drosophila.** *Development* 2018, **145**.
49. Srivastav S, Singh SK, Yadav AK, Srikrishna S: **Folic acid supplementation rescues anomalies associated with knockdown of parkin in dopaminergic and serotonergic neurons in Drosophila model of Parkinson's disease.** *Biochem Biophys Res Commun* 2015, **460:**780-785.
50. Jansen PR, Watanabe K, Stringer S, Skene N, Bryois J, Hammerschlag AR, de Leeuw CA, Benjamins JS, Munoz-Manchado AB, Nagel M, et al: **Genome-wide analysis of insomnia in 1,331,010 individuals identifies new risk loci and functional pathways.** *Nat Genet* 2019, **51:**394-403.
51. Timmers PR, Mounier N, Lall K, Fischer K, Ning Z, Feng X, Bretherick AD, Clark DW, Agbessi M, Ahsan H, et al: **Genomics of 1 million parent lifespans implicates novel pathways and common diseases and distinguishes survival chances.** *Elife* 2019, **8**.
52. Park JH, Wacholder S, Gail MH, Peters U, Jacobs KB, Chanock SJ, Chatterjee N: **Estimation of effect size distribution from genome-wide association studies and implications for future discoveries.** *Nat Genet* 2010, **42:**570-575.
53. Zhang Y, Qi G, Park JH, Chatterjee N: **Estimation of complex effect-size distributions using summary-level statistics from genome-wide association studies across 32 complex traits.** *Nat Genet* 2018, **50:**1318-1326.
54. Pilling LC, Kuo CL, Sicinski K, Tamosauskaite J, Kuchel GA, Harries LW, Herd P, Wallace R, Ferrucci L, Melzer D: **Human longevity: 25 genetic loci associated in 389,166 UK biobank participants.** *Aging (Albany NY)* 2017, **9:**2504-2520.
55. Yang J, Zeng J, Goddard ME, Wray NR, Visscher PM: **Concepts, estimation and interpretation of SNP-based heritability.** *Nat Genet* 2017, **49:**1304-1310.
56. Mackay TFC, Huang W: **Charting the genotype-phenotype map: lessons from the Drosophila melanogaster Genetic Reference Panel.** *Wiley Interdiscip Rev Dev Biol* 2018, **7**.
57. Ober U, Ayroles JF, Stone EA, Richards S, Zhu D, Gibbs RA, Stricker C, Gianola D, Schlather M, Mackay TF, Simianer H: **Using whole-genome sequence data to predict quantitative trait phenotypes in Drosophila melanogaster.** *PLoS Genet* 2012, **8:**e1002685.
58. Huang W, Richards S, Carbone MA, Zhu D, Anholt RR, Ayroles JF, Duncan L, Jordan KW, Lawrence F, Magwire MM, et al: **Epistasis dominates the genetic architecture of Drosophila quantitative traits.** *Proc Natl Acad Sci U S A* 2012, **109:**15553-15559.

59. Fiehn O: **Metabolomics--the link between genotypes and phenotypes.** *Plant Mol Biol* 2002, **48:**155-171.
60. Avanesov AS, Ma S, Pierce KA, Yim SH, Lee BC, Clish CB, Gladyshev VN: **Age- and diet-associated metabolome remodeling characterizes the aging process driven by damage accumulation.** *Elife* 2014, **3:**e02077.
61. Laye MJ, Tran V, Jones DP, Kapahi P, Promislow DE: **The effects of age and dietary restriction on the tissue-specific metabolome of Drosophila.** *Aging Cell* 2015, **14:**797-808.
62. MacMillan HA, Knee JM, Dennis AB, Udaka H, Marshall KE, Merritt TJ, Sinclair BJ: **Cold acclimation wholly reorganizes the Drosophila melanogaster transcriptome and metabolome.** *Sci Rep* 2016, **6:**28999.
63. Knee JM, Rzezniczak TZ, Barsch A, Guo KZ, Merritt TJ: **A novel ion pairing LC/MS metabolomics protocol for study of a variety of biologically relevant polar metabolites.** *J Chromatogr B Analyt Technol Biomed Life Sci* 2013, **936:**63-73.
64. Doran ML, Knee JM, Wang N, Rzezniczak TZ, Parkes TL, Li L, Merritt TJS: **Metabolomic analysis of oxidative stress: Superoxide dismutase mutation and paraquat induced stress in Drosophila melanogaster.** *Free Radic Biol Med* 2017.
65. Zampieri M, Zimmermann M, Claassen M, Sauer U: **Nontargeted Metabolomics Reveals the Multilevel Response to Antibiotic Perturbations.** *Cell Rep* 2017, **19:**1214-1228.
66. Shukla AK, Ratnasekhar C, Pragya P, Chaouhan HS, Patel DK, Chowdhuri DK, Mudiam MKR: **Metabolomic Analysis Provides Insights on Paraquat-Induced Parkinson-Like Symptoms in Drosophila melanogaster.** *Mol Neurobiol* 2016, **53:**254-269.
67. Rzezniczak TZ, Merritt TJ: **Interactions of NADP-reducing enzymes across varying environmental conditions: a model of biological complexity.** *G3 (Bethesda)* 2012, **2:**1613-1623.
68. Riahi H, Brekelmans C, Foriel S, Merkling SH, Lyons TA, Itskov PM, Kleefstra T, Ribeiro C, van Rij RP, Kramer JM, Schenck A: **The histone methyltransferase G9a regulates tolerance to oxidative stress-induced energy consumption.** *PLoS Biol* 2019, **17:**e2006146.
69. Hwangbo DS, Gershman B, Tu MP, Palmer M, Tatar M: **Drosophila dFOXO controls lifespan and regulates insulin signalling in brain and fat body.** *Nature* 2004, **429:**562-566.
70. Poirier L, Shane A, Zheng J, Seroude L: **Characterization of the Drosophila gene-switch system in aging studies: a cautionary tale.** *Aging Cell* 2008, **7:**758-770.
71. Bauer M, Katzenberger JD, Hamm AC, Bonaus M, Zinke I, Jaekel J, Pankratz MJ: **Purine and folate metabolism as a potential target of sex-specific nutrient allocation in Drosophila and its implication for lifespan-reproduction tradeoff.** *Physiol Genomics* 2006, **25:**393-404.
72. Yu S, Jang Y, Paik D, Lee E, Park JJ: **Nmdmc overexpression extends Drosophila lifespan and reduces levels of mitochondrial reactive oxygen species.** *Biochem Biophys Res Commun* 2015, **465:**845-850.
73. Celardo I, Lehmann S, Costa AC, Loh SH, Miguel Martins L: **dATF4 regulation of mitochondrial folate-mediated one-carbon metabolism is neuroprotective.** *Cell Death Differ* 2017, **24:**638-648.

74. Landis GN, Abdueva D, Skvortsov D, Yang J, Rabin BE, Carrick J, Tavaré S, Tower J: **Similar gene expression patterns characterize aging and oxidative stress in Drosophila melanogaster.** *Proc Natl Acad Sci U S A* 2004, **101:**7663-7668.

75. Girardot F, Monnier V, Tricoire H: **Genome wide analysis of common and specific stress responses in adult drosophila melanogaster.** *BMC Genomics* 2004, **5:**74.

76. Tufi R, Gandhi S, de Castro IP, Lehmann S, Angelova PR, Dinsdale D, Deas E, Plun-Favreau H, Nicotera P, Abramov AY, et al: **Enhancing nucleotide metabolism protects against mitochondrial dysfunction and neurodegeneration in a PINK1 model of Parkinson's disease.** *Nat Cell Biol* 2014, **16:**157-166.

77. Curtis C, Landis GN, Folk D, Wehr NB, Hoe N, Waskar M, Abdueva D, Skvortsov D, Ford D, Luu A, et al: **Transcriptional profiling of MnSOD-mediated lifespan extension in Drosophila reveals a species-general network of aging and metabolic genes.** *Genome Biol* 2007, **8:**R262.

78. Landis G, Shen J, Tower J: **Gene expression changes in response to aging compared to heat stress, oxidative stress and ionizing radiation in Drosophila melanogaster.** *Aging (Albany NY)* 2012, **4:**768-789.

79. Colinet H, Renault D: **Similar post-stress metabolic trajectories in young and old flies.** *Exp Gerontol* 2018, **102:**43-50.

80. Ding MZ, Li BZ, Cheng JS, Yuan YJ: **Metabolome analysis of differential responses of diploid and haploid yeast to ethanol stress.** *Omics* 2010, **14:**553-561.

81. Nikiforova VJ, Kopka J, Tolstikov V, Fiehn O, Hopkins L, Hawkesford MJ, Hesse H, Hoefgen R: **Systems rebalancing of metabolism in response to sulfur deprivation, as revealed by metabolome analysis of Arabidopsis plants.** *Plant Physiol* 2005, **138:**304-318.

82. Enattah NS, Sahi T, Savilahti E, Terwilliger JD, Peltonen L, Jarvela I: **Identification of a variant associated with adult-type hypolactasia.** *Nat Genet* 2002, **30:**233-237.

83. Wu Q, Zhao Z, Shen P: **Regulation of aversion to noxious food by Drosophila neuropeptide Y- and insulin-like systems.** *Nat Neurosci* 2005, **8:**1350-1355.

84. Krashes MJ, DasGupta S, Vreede A, White B, Armstrong JD, Waddell S: **A neural circuit mechanism integrating motivational state with memory expression in Drosophila.** *Cell* 2009, **139:**416-427.

85. Britton JS, Lockwood WK, Li L, Cohen SM, Edgar BA: **Drosophila's insulin/PI3-kinase pathway coordinates cellular metabolism with nutritional conditions.** *Dev Cell* 2002, **2:**239-249.

86. Junger MA, Rintelen F, Stocker H, Wasserman JD, Vegh M, Radimerski T, Greenberg ME, Hafen E: **The Drosophila forkhead transcription factor FOXO mediates the reduction in cell number associated with reduced insulin signaling.** *J Biol* 2003, **2:**20.

87. Tettweiler G, Miron M, Jenkins M, Sonenberg N, Lasko PF: **Starvation and oxidative stress resistance in Drosophila are mediated through the eIF4E-binding protein, d4E-BP.** *Genes Dev* 2005, **19:**1840-1843.

88. Teleman AA: **Molecular mechanisms of metabolic regulation by insulin in Drosophila.** *Biochem J* 2009, **425:**13-26.

89. Bjedov I, Toivonen JM, Kerr F, Slack C, Jacobson J, Foley A, Partridge L: **Mechanisms of life span extension by rapamycin in the fruit fly Drosophila melanogaster.** *Cell Metab* 2010, **11:**35-46.
90. Broughton SJ, Piper MDW, Ikeya T, Bass TM, Jacobson J, Driege Y, Martinez P, Hafen E, Withers DJ, Leevers SJ, Partridge L: **Longer lifespan, altered metabolism, and stress resistance in Drosophila from ablation of cells making insulin-like ligands.** *Proceedings of the National Academy of Sciences of the United States of America* 2005, **102:**3105.
91. Buch S, Melcher C, Bauer M, Katzenberger J, Pankratz MJ: **Opposing effects of dietary protein and sugar regulate a transcriptional target of Drosophila insulin-like peptide signaling.** *Cell Metab* 2008, **7:**321-332.
92. Ainsley JA, Kim MJ, Wegman LJ, Pettus JM, Johnson WA: **Sensory mechanisms controlling the timing of larval developmental and behavioral transitions require the Drosophila DEG/ENaC subunit, Pickpocket1.** *Dev Biol* 2008, **322:**46-55.
93. Freeman EG, Dahanukar A: **Molecular neurobiology of Drosophila taste.** *Curr Opin Neurobiol* 2015, **34:**140-148.
94. Olds WH, Xu T: **Regulation of food intake by mechanosensory ion channels in enteric neurons.** *Elife* 2014, **3**.
95. Wegman LJ, Ainsley JA, Johnson WA: **Developmental timing of a sensory-mediated larval surfacing behavior correlates with cessation of feeding and determination of final adult size.** *Dev Biol* 2010, **345:**170-179.
96. Waterson MJ, Chung BY, Harvanek ZM, Ostojic I, Alcedo J, Pletcher SD: **Water sensor ppk28 modulates Drosophila lifespan and physiology through AKH signaling.** *Proc Natl Acad Sci U S A* 2014, **111:**8137-8142.
97. Branch A, Shen P: **Central and Peripheral Regulation of Appetite and Food Intake in Drosophila.** In *Appetite and Food Intake: Central Control.* Edited by nd, Harris RBS. Boca Raton (FL): CRC Press/Taylor & Francis

(c) 2017 by Taylor & Francis Group, LLC.; 2017: 17-38

98. Vigne P, Frelin C: **Hypoxia modifies the feeding preferences of Drosophila. Consequences for diet dependent hypoxic survival.** *BMC Physiol* 2010, **10:**8.
99. Edgecomb RS, Harth CE, Schneiderman AM: **Regulation of feeding behavior in adult Drosophila melanogaster varies with feeding regime and nutritional state.** *J Exp Biol* 1994, **197:**215-235.
100. Min KJ, Tatar M: **Drosophila diet restriction in practice: do flies consume fewer nutrients?** *Mech Ageing Dev* 2006, **127:**93-96.
101. Deshpande SA, Yamada R, Mak CM, Hunter B, Soto Obando A, Hoxha S, Ja WW: **Acidic Food pH Increases Palatability and Consumption and Extends Drosophila Lifespan.** *J Nutr* 2015, **145:**2789-2796.
102. Zeng C, Du Y, Alberico T, Seeberger J, Sun X, Zou S: **Gender-specific prandial response to dietary restriction and oxidative stress in Drosophila melanogaster.** *Fly (Austin)* 2011, **5:**174-180.

103. Galikova M, Diesner M, Klepsatel P, Hehlert P, Xu Y, Bickmeyer I, Predel R, Kuhnlein RP: **Energy Homeostasis Control in Drosophila Adipokinetic Hormone Mutants.** *Genetics* 2015, **201:**665-683.

104. Kim MJ, Ainsley JA, Carder JW, Johnson WA: **Hyperoxia-triggered aversion behavior in Drosophila foraging larvae is mediated by sensory detection of hydrogen peroxide.** *J Neurogenet* 2013, **27:**151-162.

105. Guntur AR, Gou B, Gu P, He R, Stern U, Xiang Y, Yang CH: **H2O2-Sensitive Isoforms of Drosophila melanogaster TRPA1 Act in Bitter-Sensing Gustatory Neurons to Promote Avoidance of UV During Egg-Laying.** *Genetics* 2017, **205:**749-759.

106. Bhatla N, Horvitz HR: **Light and hydrogen peroxide inhibit C. elegans Feeding through gustatory receptor orthologs and pharyngeal neurons.** *Neuron* 2015, **85:**804-818.

107. Ameku T, Yoshinari Y, Texada MJ, Kondo S, Amezawa K, Yoshizaki G, Shimada-Niwa Y, Niwa R: **Midgut-derived neuropeptide F controls germline stem cell proliferation in a mating-dependent manner.** *PLoS Biol* 2018, **16:**e2005004.

108. R Core Team: **R: A Language and Environment for Statistical Computing.** Vienna, Austria: R Foundation for Statistical Computing; 2018.

109. Terry T: **A Package for Survival Analysis*in S*. version 2.38.** https://CRAN.R-project.org/package=survival . 2015.

110. Garlapow ME, Huang W, Yarboro MT, Peterson KR, Mackay TF: **Quantitative Genetics of Food Intake in Drosophila melanogaster.** *PLoS One* 2015, **10:**e0138129.

111. Patterson N, Price AL, Reich D: **Population structure and eigenanalysis.** *PLoS Genet* 2006, **2:**e190.

112. Gramates LS, Marygold SJ, Santos GD, Urbano JM, Antonazzo G, Matthews BB, Rey AJ, Tabone CJ, Crosby MA, Emmert DB, et al: **FlyBase at 25: looking to the future.** *Nucleic Acids Res* 2017, **45:**D663-d671.

113. Mi HaHXaMAaTHaMCaKDaTPD: **PANTHER version 11: expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements.** *Nucleic Acids Research* 2017, **45:**D183-D189.

114. Benjamini Y, Hochberg Y: **Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing.** *Journal of the Royal Statistical Society Series B (Methodological)* 1995, **57:**289-300.

115. Haug K, Salek RM, Conesa P, Hastings J, de Matos P, Rijnbeek M, Mahendraker T, Williams M, Neumann S, Rocca-Serra P, et al: **MetaboLights–an open-access general-purpose repository for metabolomics studies and associated meta-data.** *Nucleic Acids Res* 2013, **41:**D781-786.

Tables

Table 1. Genes Associated with Peroxide Resistance.

FlyBase ID	Gene	Name	Lead Variant	Pmin	Pgene	FDR
FBgn0031803	ppk14	Pickpocket 14	2L_6353972_SNP	2.12 x 10-6	9.00 x10-6	0.004797
FBgn0031261	nAChRbeta3	Nicotinic acetylcholine receptor beta 3 (Dbeta3) subunit	2L_544231_SNP	9.73 x10-7	2.20 x10-5	0.011704
FBgn0031802	ppk7	Pickpocket 7	2L_6351358_SNP	2.12 x10-6	2.70 x10-5	0.014337
FBgn0034098	krimp	FI20010p1	2R_12124705_SNP	1.68 x10-6	3.10 x10-5	0.01643
FBgn0034099	CG15708	AT21920p	2R_12127442_SNP	1.68 x10-6	3.10 x10-5	0.01643
FBgn0005660	Ets21C	DNA-binding protein D-ETS-6	2L_546065_SNP	9.73 x10-7	3.30 x10-5	0.017424
FBgn0002031	l(2)37Cc	Protein l(2)37Cc	2L_19121067_SNP	7.57 x10-6	4.50 x10-5	0.023715
FBgn0003963	ush	Zinc finger protein ush	2L_475238_SNP	4.51 x10-8	7.10 x10-5	0.037346
FBgn0031865	Nha1	Na[+]/H[+] hydrogen antiporter 1, isoform A	2L_6860010_SNP	5.32 x 10-8	0.000123	0.064575
FBgn0020906	Jon25Bi	Jonah 25Bi	2L_4953449_SNP	4.23 x10-5	0.000214	0.112136
FBgn0031655	Marcal1	SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A-like protein 1	2L_4954910_SNP	1.27 x10-5	0.000253	0.132319
FBgn0263831	Gen	Flap endonuclease GEN	3L_5141974_SNP	7.43 x10-6	0.000262	0.136764
FBgn0037730	DmelCG9444	CG9444, isoform A	3R_5506170_SNP	9.71 x10-5	0.000272	0.141712
FBgn0005616	msl-2	E3 ubiquitin-protein ligase msl-2	2L_3461213_SNP	3.65 x10-5	0.000367	0.19084
FBgn0263102	psq	Pipsqueak, isoform M	2R_6444567_SNP	3.29 x10-5	0.000373	0.193587
FBgn0004103	Pp1-87B	Serine/threonine-protein phosphatase alpha-2 isoform	3R_8248858_SNP	0.0001966	0.000437	0.226366
FBgn0083960	DmelCG34124	FI23230p1	2L_4957415_SNP	1.27 x10-5	0.000461	0.238337
FBgn0040510	ACXA	Adenylyl cyclase X, isoform A	2L_12920145_SNP	5.17 x10-5	0.000472	0.243552
FBgn0085227	CG34198	HDC07368	2R_15569059_SNP	0.0001068	0.000489	0.251835
FBgn0040506	ACXE	Adenylyl cyclase X, isoform E	2L_12924675_SNP	8.53 x10-5	0.000517	0.265738
FBgn0027109	NPF	Neuropeptide F	3R_12434359_SNP	9.19 x10-5	0.000578	0.296514
FBgn0031538	DmelCG3246	CG3246	2L_3459329_SNP	3.65 x10-5	0.000611	0.312832
FBgn0052391	DmelCG32391	Uncharacterized protein	3L_7060319_SNP	1.15 x10-5	0.000723	0.369453
FBgn0038043	CG17202	c-Myc-binding protein homolog	3R_8247496_SNP	0.0001966	0.000762	0.38862
FBgn0005640	Eip63E	Ecdysone-induced protein	3L_3514849_SNP	1.80 x10-5	0.00091	0.46319

		63E, isoform N				
FBgn0051641	stai	stathmin	2L_6099379_SNP	6.01 x10-6	0.000922	0.468376
FBgn0014009	Rab2	GH01619p	2R_2583312_SNP	5.43 x10-5	0.000932	0.472524
FBgn0262467	Scox	AT19154p	2L_4967292_SNP	0.0002591	0.000943	0.477158
FBgn0040509	ACXB	Adenylyl cyclase X, isoform B	2L_12916421_SNP	0.0003891	0.00096	0.4848

Table 1. The gene identifications from Fly Base (release 5.49) for all genes with P_{gene} FDR < 0.5. The variant with the smallest P value for each gene (P_{min}) is listed as the lead variant.

Figures

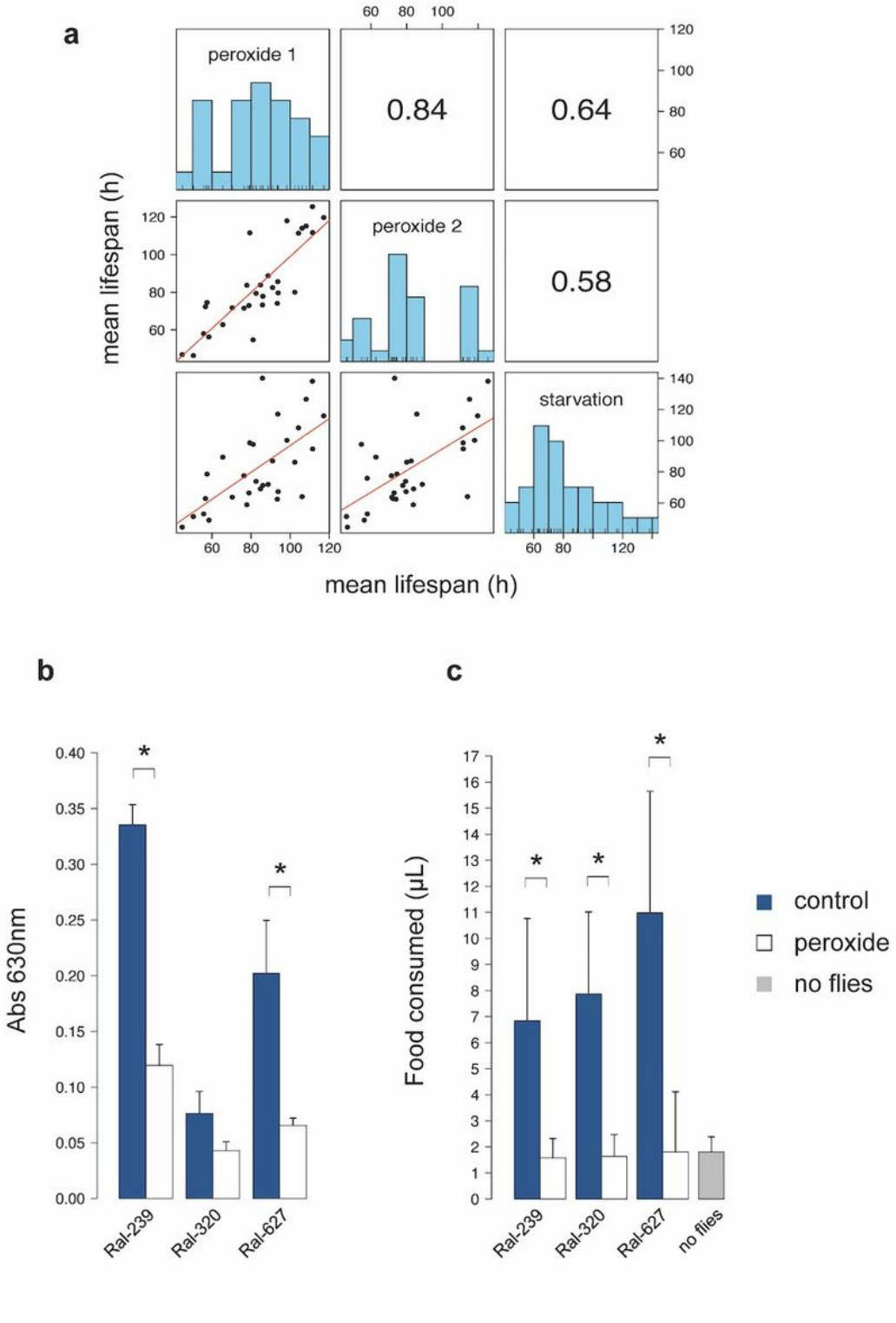

Figure 1

Starvation explains the lifespan effect of H2O2 food (a) The correlations between mean lifespan of 31 DGRP lines on food containing either 2% glucose and 2% hydrogen peroxide in replicate trials (peroxide 1 and peroxide 2) or no glucose (starvation). Below the diagonal are plots of trait values (mean survival (h)). Least-squares linear regression lines are shown in red. Above the diagonal are Spearman correlation coefficients for each pair of traits. (b and c) Feeding assays, the mean (+SD) absorbance at 630nm of

extracts from three replicate vials of seven to eleven 1 to 5 day old mated females from the indicated DGRP line (b). Flies were exposed to peroxide (open bars) or control food (colored bars) for two hours of feeding prior to dye extraction. (c) The mean (+SD) volume of liquid food consumed by ten replicates of ten 2 to 4 day-old mated females was measured using the CAFE assay. CAFE food contained 2% glucose and, either water (colored bars) or 2% peroxide (open bars). Additional apparatus were set up without flies (no flies) to measure volume loss due to evaporation. Asterisks indicate $p < 0.05$ (Welch's t-test).

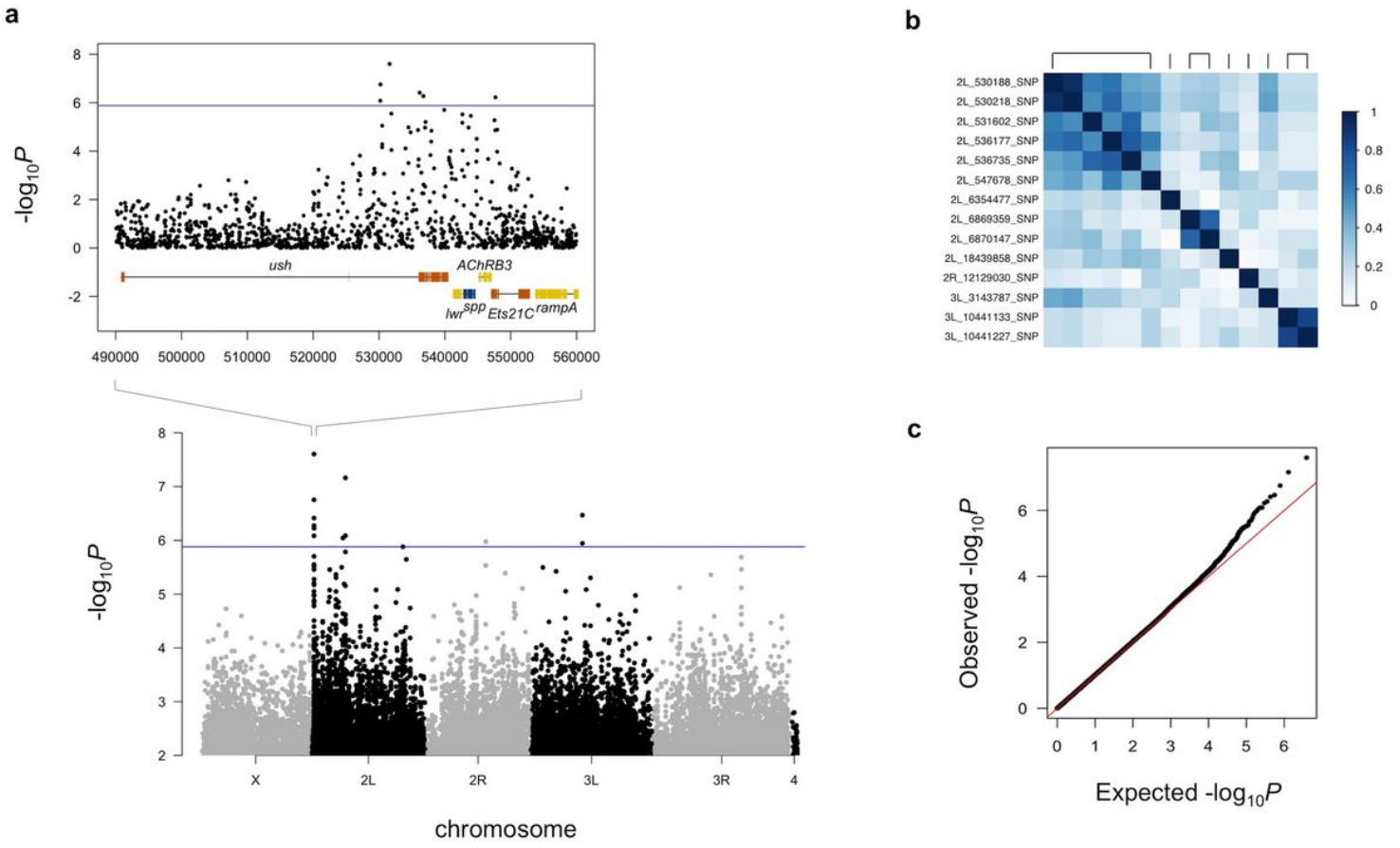

Figure 2

Genome-Wide Association. (a) −log10 P-values for each polymorphism (MAF ≥ 0.05, < 30% missing) in association with lifespan on H2O2 food, P-values above 10-2 not shown, the FDR < 0.2 level of significance is shown by a blue line. Inset shows a locus on chromosome 2L that includes 6 resistance-associated SNPs in LD ($r \geq 0.5$) and several candidate genes, including ush. (b) A pairwise LD plot for all 14 resistance-associated SNPs (FDR < 0.2), groups of SNPs in LD are indicated with bars across the top, and a color scale for r is shown. (c) A Q-Q plot of −log10 P-values for the variants tested in the GWAS showing little to no inflation after adjusting for population structure (Methods).

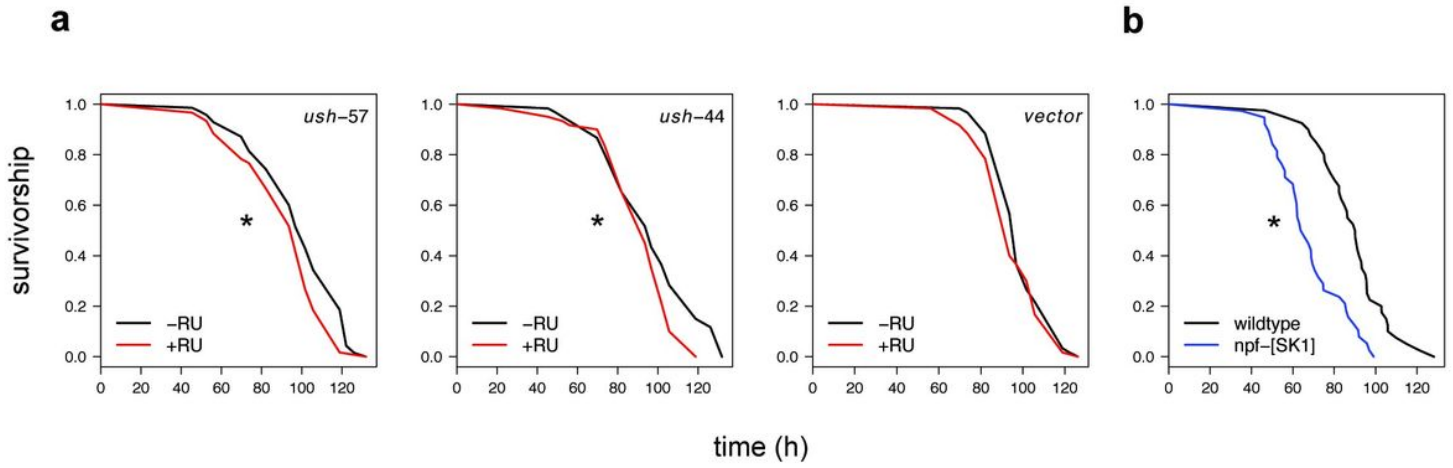

Figure 3

Candidate Gene Validation. (a) Mated F1 females from crosses of the GeneSwitch106 driver to either UAS-RNAi lines targeting u-shaped (ush-57 and ush-44) or a UAS-control line (attP2, vector), that were maintained on RU486 food (red line) or –RU food (black line) for 48h prior to being transferred to H2O2 food or control food. Survivorship was recorded for six replicate vials of ten flies per vial on H2O2 food (n = 60 flies) and two replicate vials on control food (n = 20 flies). In another experiment (b), survivorship of wild type females (n = 40) were compared to NPFSK1 females (n = 38) on H2O2 food or control food in an activity monitoring system (Methods). Asterisks indicate significant effect of RU486, or of the NPFSK1 mutation ($P < 0.05$) from the log-rank test using the survival package in R. Results are representative of at least two independent experiments, and there was no mortality observed on control food during these experiments.

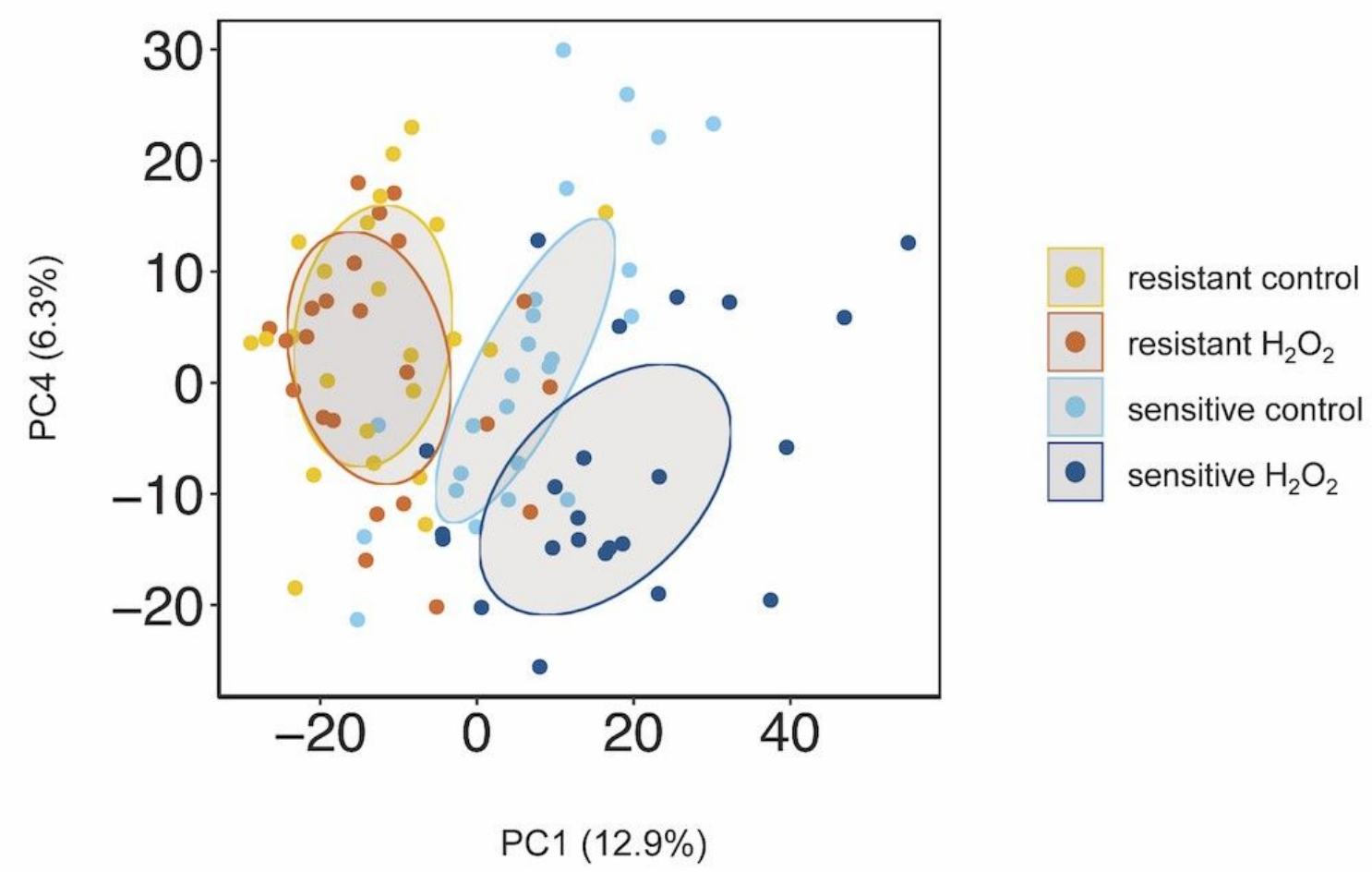

Figure 4

Projections of Metabolomic Principal Components Principal components that demonstrate the separation of resistant and sensitive genotypes, as well as the effect of treatment (Control vs. H2O2) from and negative mode. Trait and treatment groups are indicated by colored points and ellipses (50% CI). The percentage of the variance explained by each PC is shown in parentheses.

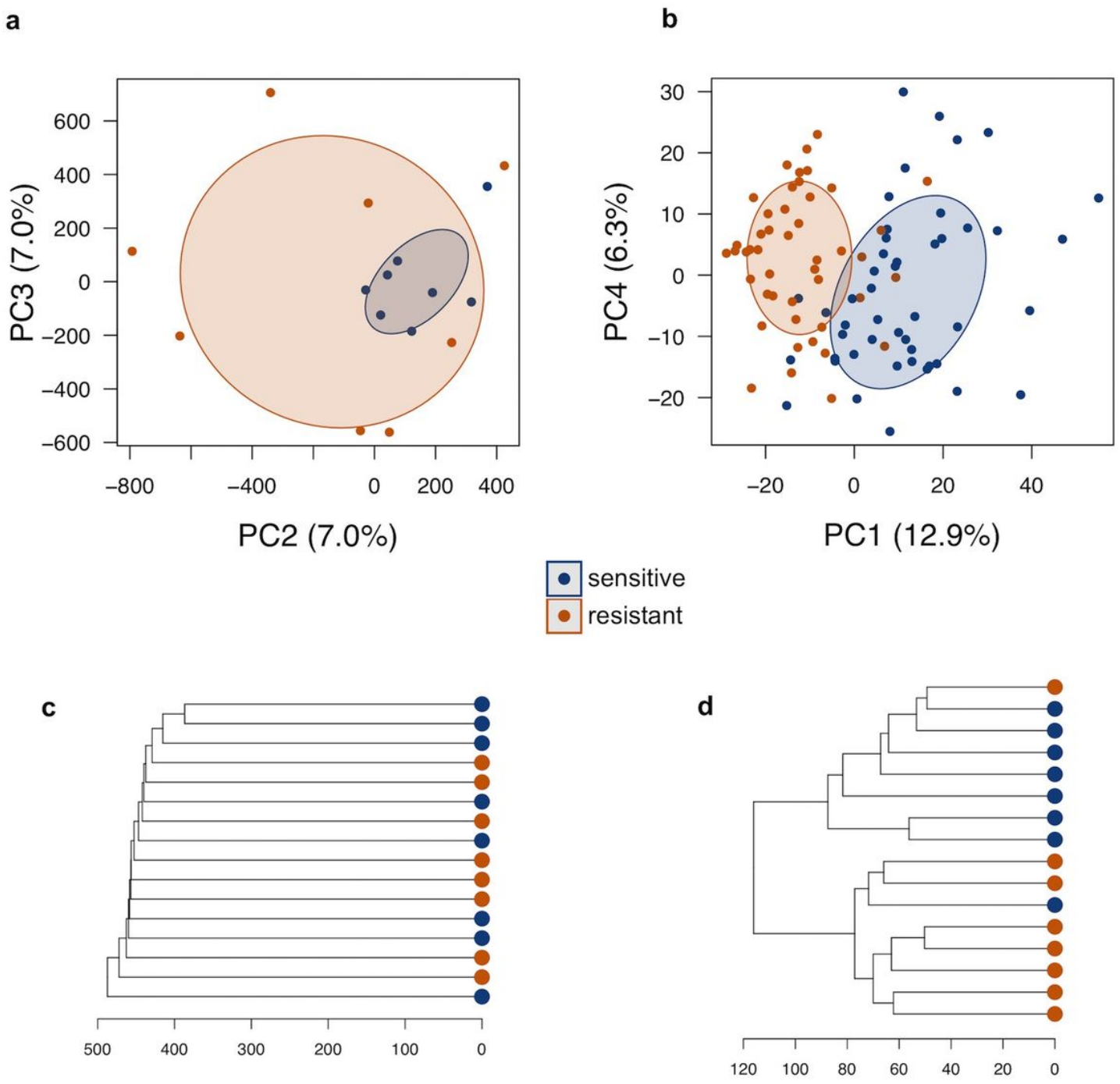

Figure 5

Metabolome Outperforms Genotype as a Predictor of Phenotype (a) PCA of >2.8x106 variants (0% missing calls) from the 16 genotypes in the metabolomics analysis. (b) PCs 1 and 4 of 2417 metabolite features from the same genotypes as in (a). In all plots, sensitive genotypes are colored in blue and resistant genotypes are colored in orange. The two PCs that gave the most separation between phenotype are plotted and the variance explained by each PC is shown in parentheses. Ellipses are the 50% confidence intervals. (c) Representative clustering of 6846 LD-pruned variants (0% missing data, Methods) from the 16 genotypes used for metabolomics. (d) Clustering of mean metabolite levels for

2417 metabolite features in untreated flies. Data were clustered using the minimal variance method (Ward's D2) in the hclust package.

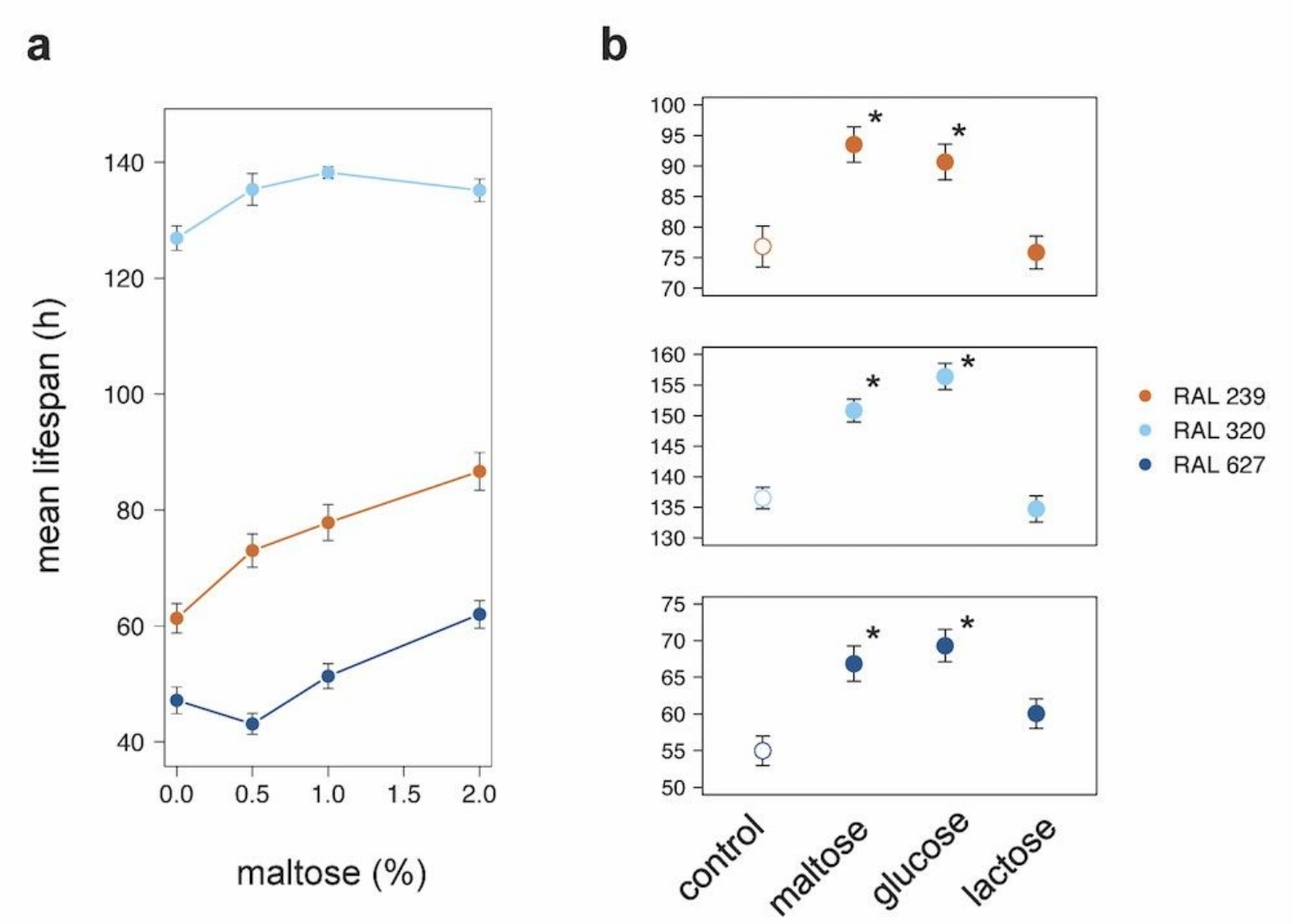

Figure 6

Supplemental Maltose or Glucose Increases Resistance to Peroxide. Females of three different DGRP lines were allowed to feed on media containing 2% glucose with the indicated amount of supplemental maltose (a). (b) Mated females fed on glucose food containing 2% of either additional maltose, additional glucose, or additional lactose, or control without additional carbohydrate for four days prior to being transferred to 2% glucose food containing 2% H2O2 for survival analysis. Points are the mean lifespans (± 1 s.e.) of 49 to 60 flies. ANOVA was used to compare the effect of maltose (P =0.0035) after removing the effect of line in (a). The log-rank test was used in (b) to test the difference between lifespans of carbohydrate-fed verses control flies (asterisk, Bonferroni-corrected $P \leq 0.05$).

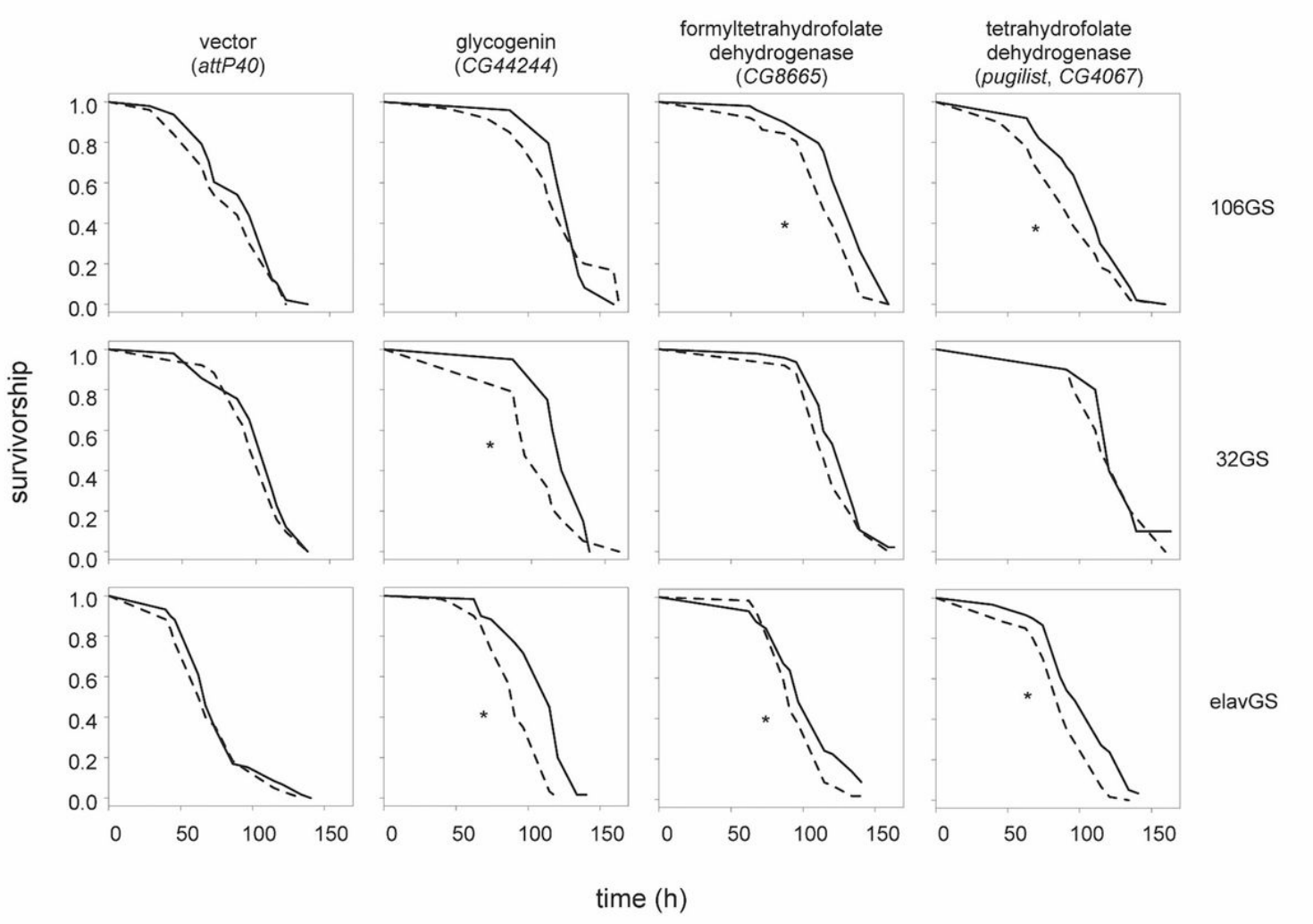

Figure 7

Knockdown of Genes in Glycogen or Folate Metabolism. F1 females from crosses of UAS-RNAi lines, targeting the indicated gene with either the GeneSwitch106 (106GS, top row), GeneSwitch32 (32GS, middle row), or elavGeneSwitch (elavGS, bottom row) drivers, were maintained on RU486 food (dashed line) or –RU control food (solid line) for 48h prior to being transferred to H2O2 or control food. Survivorship was recorded for five replicate vials of ten flies per vial on H2O2 food and two to five replicates on control food (n = 20 or 5 flies). Asterisks indicate significant effect of RU486 ($P < 0.05$, log-rank test). Data are representative of one or two experiments, and the mortality of flies on control food (not shown) did not exceed 15% in any condition and did not depend on RU486 exposure.

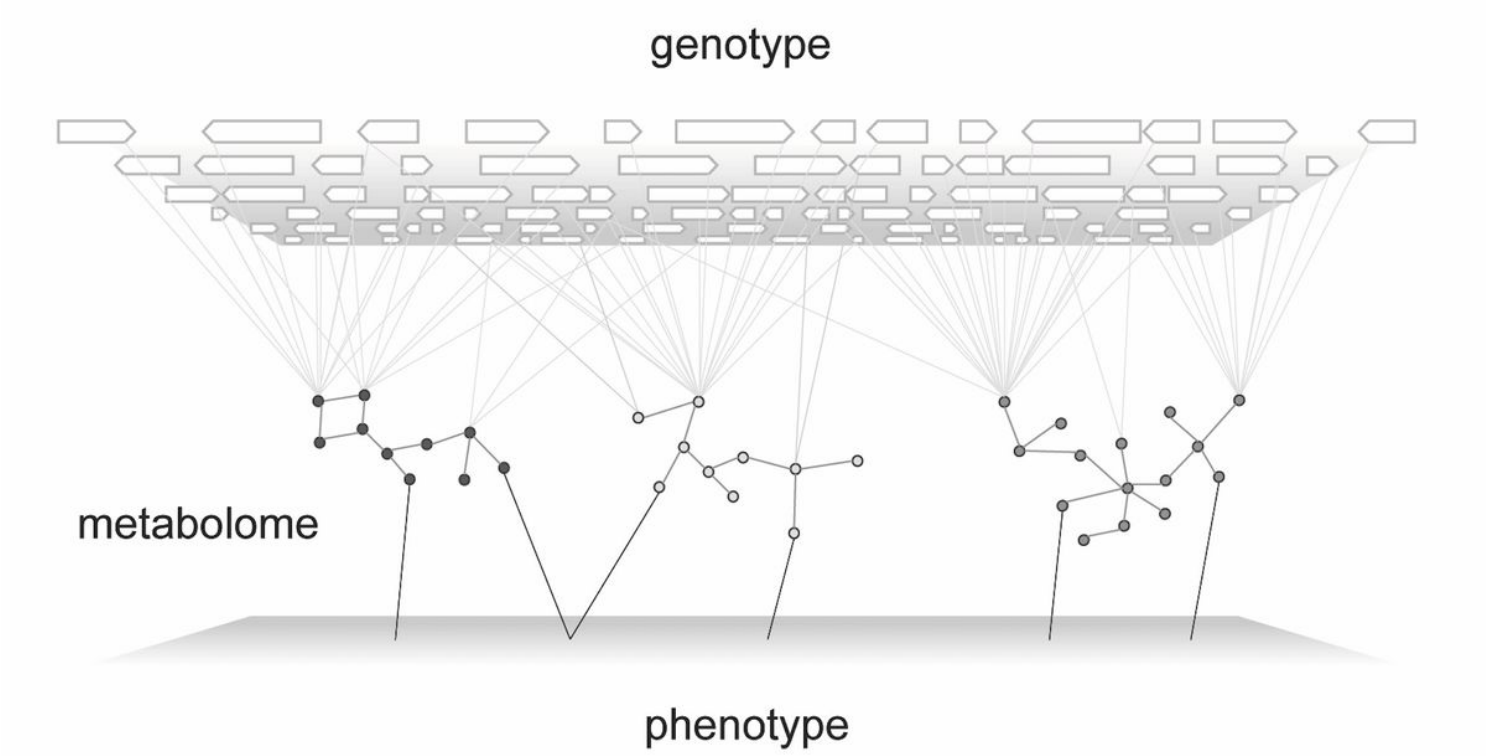

Figure 8

The Metabolome as a Bridge in the Genotype to Phenotype Map. Phenotypic variation, represented by the bottom plain, is influenced by the action of a large number of genes and genetic variants (upper plain) on a smaller number of metabolites or metabolic pathways. Direct influences on phenotype are shown as dark lines whereas indirect connections are shown with grey lines. Genetic variation can influence the levels of individual metabolites or the relationship between metabolites and, as a result, affect metabolic pathways that share direct connections to phenotype; or, genetic variation can affect phenotype more directly (not depicted). Two or more genes might act epistatically to influence the level of a single metabolite, and a single gene might act pleiotropically on multiple metabolites.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- FigS2.tif
- TableS2.xlsx
- FigS4.tif
- FigS5.tif
- FigS3.tif
- FigS6.tif
- FigS1.tif