Modular Evolution of the *Drosophila* Metabolome

Benjamin R. Harrison,[†,1] Jessica M. Hoffman,[†,2] Ariana Samuelson,[3] Daniel Raftery [iD],[4] and Daniel E.L. Promislow [iD][*,1,3]

[1]Department of Lab Medicine & Pathology, University of Washington School of Medicine, Seattle, WA, USA
[2]Department of Biology, University of Alabama at Birmingham, Birmingham, AL, USA
[3]Department of Biology, University of Washington, Seattle, WA, USA
[4]Department of Anesthesiology & Pain Medicine, University of Washington School of Medicine, Seattle, WA, USA
[†]These authors contributed equally to this work.
*Corresponding author: E-mail: promislo@uw.edu.
Associate editor: Patricia Wittkopp

Abstract

Comparative phylogenetic studies offer a powerful approach to study the evolution of complex traits. Although much effort has been devoted to the evolution of the genome and to organismal phenotypes, until now relatively little work has been done on the evolution of the metabolome, despite the fact that it is composed of the basic structural and functional building blocks of all organisms. Here we explore variation in metabolite levels across 50 My of evolution in the genus *Drosophila*, employing a common garden design to measure the metabolome within and among 11 species of *Drosophila*. We find that both sex and age have dramatic and evolutionarily conserved effects on the metabolome. We also find substantial evidence that many metabolite pairs covary after phylogenetic correction, and that such metabolome coevolution is modular. Some of these modules are enriched for specific biochemical pathways and show different evolutionary trajectories, with some showing signs of stabilizing selection. Both observations suggest that functional relationships may ultimately cause such modularity. These coevolutionary patterns also differ between sexes and are affected by age. We explore the relevance of modular evolution to fitness by associating modules with lifespan variation measured in the same common garden. We find several modules associated with lifespan, particularly in the metabolome of older flies. Oxaloacetate levels in older females appear to coevolve with lifespan, and a lifespan-associated module in older females suggests that metabolic associations could underlie 50 My of lifespan evolution.

Key words: metabolome, evolution, *Drosophila*, coevolution, modular.

Introduction

The development of high-dimensional "omics" methods has had a dramatic impact on the nature of comparative studies. First, genome data enable researchers to derive accurate, high-resolution phylogenies to probe the evolution of gene families and the genomic signatures of selection (e.g., Whitkus et al. 1992; Clark et al. 2007). Second, transcriptomics, proteomics, and metabolomics allow for the study of adaptation at the molecular level with a breadth not previously possible (e.g., von Mering et al. 2003; Spirin et al. 2006; Bedford and Hartl 2009; Brawand et al. 2011; Gordon and Ruvinsky 2012; Martin and Fraser 2018; Cope et al. 2020). Metabolomic methods are able to measure the levels of hundreds to thousands of metabolite features with a high level of accuracy (Jones et al. 2012). Given the central role that the metabolome plays in organismal structure and function, and the fact that it integrates upstream genetic and environmental variation, it is surprising how little is known about the evolution of metabolite abundance (Flowers et al. 2007; Noda-Garcia et al. 2018).

There are just a handful of comparative studies of the animal metabolome (Khaitovich et al. 2008; Fu et al. 2011; Park et al. 2012; Blekhman et al. 2014; Bozek et al. 2014, 2017; Ma et al. 2015). Comparative studies are often confounded by lineage-specific environments, phylogenetic nonindependence, and measurement error in estimating species-level phenotypes. Lineage-specific environments are particularly relevant in metabolomics studies as environment, genotype, sex, age, and tissue/organ can each have dramatic effects on metabolome profiles (Steuer 2006; Hoffman et al. 2014; Li et al. 2017; Khrameeva et al. 2018; Wilinski et al. 2019). Comparative analysis is also affected by phylogenetic nonindependence, although methods to handle phylogenetic confounding, including within systems-level data, are well established (Felsenstein 2008; Dunn et al. 2018). We address these issues using a common garden design, with multiple strains per species, and phylogenetic correction using a standard approach (Felsenstein 2008).

There is abundant evidence that gene expression within species is modular, where groups of genes covary in their expression, sharing many more such covarying partners

Open Access

within the group (module), than with genes in other groups (Hartwell et al. 1999; Wagner et al. 2007). Patterns of covariation are of great interest as they are widley assumed to reveal functional associations (Ge et al. 2001). In contrast to within species coexpression, phylogenetic patterns of coexpression and modularity are less explored. There is a substantial literature describing comparative analyses of modularity in morphology (Klingenberg 2014), but relatively limited comparative analyses of covariation and modularity in gene content (von Mering et al. 2003) or in gene expression (von Mering et al. 2003; Fraser et al. 2004; Innocenti and Chenoweth 2013; Martin and Fraser 2018; Cope et al. 2020).

The covariation of cellular traits across species may indicate evolution in pathway activity, functional interactions, or of development. Theories of biochemical pathway evolution are focused in large part on the evolution of metabolic enzymes, perhaps owing to the relative scarcity of systems-level data on metabolite abundance (Noda-Garcia et al. 2018). Measuring the divergence and covariation among the metabolome can provide such insight. For instance, comparisons of the tissue-level metabolome in just four mammal species highlighted the divergence of the human brain metabolome as a putative hallmark of human evolution (Fu et al. 2011; Bozek et al. 2014, 2015). Thus, the metabolome may be reshaped in coordination with evolutionary change in organ systems and other organismal phenotypes. All of these possibilities provide a strong rationale for taking a systems-level approach to understand the evolutionary context of natural variation.

The fruit fly genus Drosophila offers a powerful model to examine evolutionary dynamics of metabolic pathways. Although studies of the classic model species Drosophila melanogaster predominate, evolutionary biologists have long appreciated the potential of this diverse clade, with both classic and contemporary studies in Drosophila pseudoobscura (e.g., Dobzhansky 1946), the Hawaiian Drosophila (Carson and Kaneshiro 1976), and cactophilic Drosophila species (Markow et al. 1983), among many others (Kambysel and Heed 1971; Schnebel and Grossfield 1983; Partridge et al. 1987; Coyne and Orr 1989; Kellermann et al. 2009). Drosophila have also played a central role in advances in genomics, including comparative work among fully sequenced genomes of numerous species (Ballard 2000; Bai et al. 2007; Clark et al. 2007; Stark et al. 2007).

Here, we measure a panel of metabolites across the genus Drosophila in both sexes at two ages. We find that patterns of variation in the metabolome are largely consistent with phylogenetic relatedness. Comparing two modes of evolution, we find that the simple Brownian motion (BM) model of evolution is a good fit to overall metabolome divergence, though sex and age are significant and conserved contributors to variation. The metabolome also shows evidence of modular coevolution, where groups of metabolites vary in concert across the phylogeny. Interestingly, the patterns of covariation are somewhat specific to each sex and age, highlighting the dynamic nature of metabolomic variation, its potential to explain variation in phenotype over sex and age, as well as the importance of common garden design in comparative studies. We then examine the variation within modules and find they show distinct patterns of evolution, including some that appear to be under stabilizing selection. Consistent with the idea that the evolution of pathway activity may explain metabolite coevolution, we find that some modules are enriched for specific biological pathways. Additionally, we find evidence for the coevolution of lifespan with metabolite modules, which suggests that lifespan evolution has a conserved molecular basis across a 50-My phylogeny. Our hope is that this work might inspire further explorations so that we can begin to understand how, over hundreds of millions of years, evolution has shaped systems of functional and structural building blocks that make up all of life.

Results

We raised one to three wild-type strains from each of 11 species of Drosophila in a common environment and collected age-matched samples of each sex at 5 days (young) and 31 days (old) after eclosion. Strain-specific mean lifespans in these experiments ranged from 20.1 to 85.7 days with a grand mean across all species of 50.5 days (supplementary fig. S8, Supplementary Material online). Targeted LC-MS/MS metabolome profiles of whole flies were measured at each age and in each sex, along with additional untargeted LC-TOF-MS profiling of young flies, with up to three replicates per strain, sex, age, and species. Our metabolomic data sets comprised a panel of 97 targeted metabolites and an untargeted data set with 4,419 features detected in at least one sample. Metabolomic analysis often includes missing values due to the absence of a metabolite in a sample or to limits of detection. All 97 targeted metabolites were detected in all samples, and the untargeted panel detected 590 features present in all or almost all samples, imputing 228 features that were absent in only one sample. Missingness, a measure of the number of missing values within a sample, lacked any evidence of phylogenetic signal ($K = 0.13$, $P = 0.90$), and was associated with signal intensity ($F_{1,4416} = 578.8$, $P = 2.78 \times 10^{-120}$), suggesting that many missing values are likely to represent features that fall below the limit of detection.

Phylogenetic and Selection Signatures in the Metabolome

For targeted metabolome profiles, the first and second principal components (PC1 and PC2) together explain 30.1% of the variance and, by visual inspection, capture latent variation by age and sex, respectively (fig. 1A). Along with the age and sex-related variation, we explored the influence of phylogeny on the multivariate metabolome by plotting the PCs in a phylogenetic context, and by measuring their phylogenetic signal. Visual inspection of PC1 and PC2 as well as statistical tests suggest phylogenetic signal in each PC for young flies, but not necessarily for older flies (fig. 1B and table 1). We also found that metabolome-derived phylogenies showed significant concordance with the genome-based phylogeny, where branch scores, the sum of the squares of the difference between each branch in the true and deduced trees, ranged
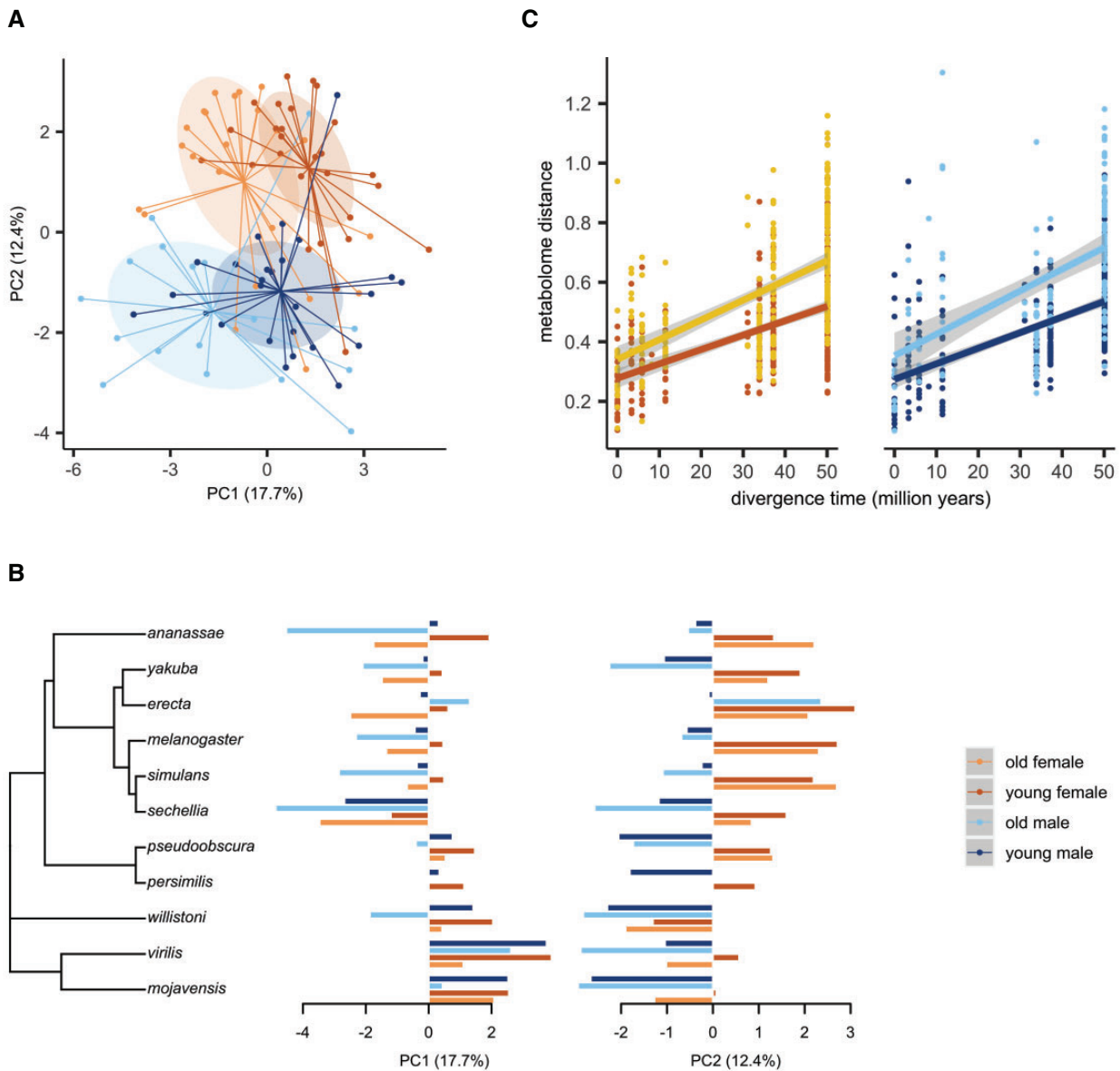
FIG. 1. Evolution of the Drosophila Metabolome. (A) Samples from each strain at each age (young 5 days, or old 31 days) and sex are plotted along principal components 1 (PC1) and 2 (PC2) of the mean-centered and scaled log-abundance of 97 targeted metabolites. Each sample is mapped by lines to the centroid of its respective sex and age group. The colored legend at bottom right refers to the sex and age groups in each figure. (B) The mean of PC1 (left) or PC2 (right) of each species is plotted along the phylogeny (Kumar et al., 2017). (C) The pairwise metabolome distance between all samples within each age and sex is plotted over the divergence time between species pairs. Intraspecies sample distances are plotted over time = 0. The metabolome distance was calculated as the average squared difference of the log-metabolite abundance for the 97 targeted metabolites between all pairs of samples. The lines in (C) represent the fit to an ordinary least squares linear model: metabolome divergence ~ divergence time within each group and are equivalent to the BM model of trait evolution.

from 0.131 to 0.178 for each sex, age, or detection methodology (Kuhner and Felsenstein 1994; Kumar et al. 2017). In all cases, the scores were significantly closer to the real phylogeny than were permutations of the genome phylogeny ($P \leq 0.017$, supplementary table S1, Supplementary Material online).

Having established a strong phylogenetic signature in the metabolome, we then sought to determine the mode of metabolome evolution. We compared two relatively simple models of evolution, the BM and the Ornstein–Uhlenbeck (OU) models. The BM model posits that traits diverge linearly

with respect to time in a direction that is independent of the current trait values, whereas the OU model extends the BM model with the addition of a parameter representing stabilizing selection and thus can model limits on the extent of divergence (Cressler et al. 2015). We find that the BM model is a better fit than the OU model to the divergence of the metabolome, where the r^2 of the BM fit ranged from 0.16 to 0.27 across both ages and sexes, and in each case ΔAIC analysis favored BM (supplementary fig. S2, Supplementary Material online). Additionally, we fit a linear model,

Table 1. Phylogenetic Signal in the *Drosophila* Metabolome. Phylogenetic Signal in the First Two Metabolome PCs at Each Sex and Age (Group) Was Measured, Either as Pagel's λ or Blomberg's K, and Tested by the Likelihood Ratio Test, or by 1×10^5 Simulations, Respectively.

PC	Group	Pagel's λ	Blomberg's K
1	Young female	0.890*	1.378**
1	Young male	0.754*	0.861*
1	Old female	0.586	0.586
1	Old male	0.313	0.529
2	Young female	1.062**	2.031**
2	Young male	0.977	1.123*
2	Old female	0.773	0.983
2	Old male	0.000	0.387

*$P < 0.05$,
**$P < 0.005$.

equivalent to BM with all data and tested for effects of age, sex, or their interaction on the metabolome distance and the rate of divergence. The metabolome distance was greater ($\beta_{age} = 0.063$, $P = 0.022$), and the divergence rate higher ($\beta_{time \times age} = 1.80 \times 10^{-3}$, $P = 0.012$) in the metabolome of older flies (fig. 1C).

Evidence of Modular Coevolution in the *Drosophila* Metabolome

We explored the possibility that the levels of metabolites coevolve by measuring the pairwise covariance among phylogenetically independent contrast scores (PICs, Materials and Methods). PICs remove the confounding effect of phylogeny while preserving the correlations that may exist between traits as they diverge at each node in a phylogeny (Felsenstein 2008), and have been applied to multivariate comparative morphometric data in similar ways (Klingenberg 2014). We find substantial covariance between PICs of metabolite levels in samples of each age and sex. In comparison with randomized data, where species labels are permuted such that the relationships within species are preserved, the original data show a much higher frequency of highly positive and negative pairwise correlations among the PICs of metabolites than expected by chance (supplementary fig. S3, Supplementary Material online). Clustering of pairwise PIC correlations among the metabolites indicated a high degree modularity (fig. 2). We evaluate the significance of the modularity in these networks by measuring edge betweenness community detection (Girvan and Newman 2002), and comparing them to rewired networks of the same degree distribution. In each sex and age group, the real network is significantly modular ($P < 0.001$).

To define the metabolites within each module, we used weighted gene correlation analysis (WGCNA), which led to over 88% of the 97 targeted metabolites being placed in a module, with five to six modules in each sex and age group (fig. 2 and supplementary fig. S4, Supplementary Material online). The metabolite members of modules in each sex and age are somewhat distinct, where for example, two metabolites might be members of module A in young males, and be members of two different modules in older males,

though pairwise comparison between sex and age groups revealed from two to four modules with significant intersections in each comparison (Fisher's exact test, $P < 0.05$, supplementary fig. S5 and table S2, Supplementary Material online).

Rather than metabolite levels evolving under direct selection, we hypothesized that the modularity in metabolite coevolution reflects selection acting more directly on biological pathways (Hartwell et al. 1999; Wagner et al. 2007; Noda-Garcia et al. 2018). In support of this hypothesis, we made two observations. First, we tested the hypothesis that selection may operate at the level of metabolomic module. We compared the fit of the OU and BM models on the divergence of metabolites within each module (Cressler et al. 2015). Across the 23 modules, we find that eight are better fit by the OU model, suggesting that some modules are evolving under stabilizing selection (supplementary table S3 and fig. S6, Supplementary Material online). Second, we found by enrichment analysis that the metabolites within five of the modules are more connected to subgraphs of the KEGG database than we would expect by chance (Picart-Armada et al. 2018), with at least one KEGG pathway in each of these five modules showing such enrichment (FDR ≤ 0.2, supplementary table S5, Supplementary Material online). Thus, there is evidence of extensive modularity in the metabolome at the phylogenetic level, and the patterns of covariation are consistent with adaptive coevolution of metabolites that may share biological function.

Sex and Age Affect Interspecific Metabolome Variation

To examine the effects of sex and age on the metabolome in a phylogenetic context, we use a Bayesian mixed model (Materials and Methods). Among the targeted metabolites, we found 44 of 97 metabolites with significant sex effects, 38 with age effects, and 2 with sex-by-age interactions (FDR < 0.05, fig. 3). Analysis of 590 untargeted features at young age found effects of sex for 228 metabolite features at FDR < 0.05. Thus, we find evidence that a substantial portion of the metabolome varies with sex and age in ways that have persisted over at least 50 My.

Along with the conserved effects of sex, we also find evidence for evolution in sexual dimorphism within the metabolome. At a multivariate level, measuring the distance along PC1 and PC2 for each sex within each species, we see that the nine species of the *Sophophora* subgenus separate with the female samples having consistently higher PC values than the males, whereas in the two species of the *Drosophila* subgenus, the male and female samples remain clustered along this axis (supplementary fig. S7, Supplementary Material online).

Lifespan and the Metabolome Coevolve

Lastly, we analyzed the lifespan of the 26 strains with an average of 95.3 flies for each sex and strain (± 30 SD, $n = 16$–133, supplementary fig. S8, Supplementary Material online). To identify metabolites and modules that covary with lifespan, we took two approaches. First, regression of PICs of
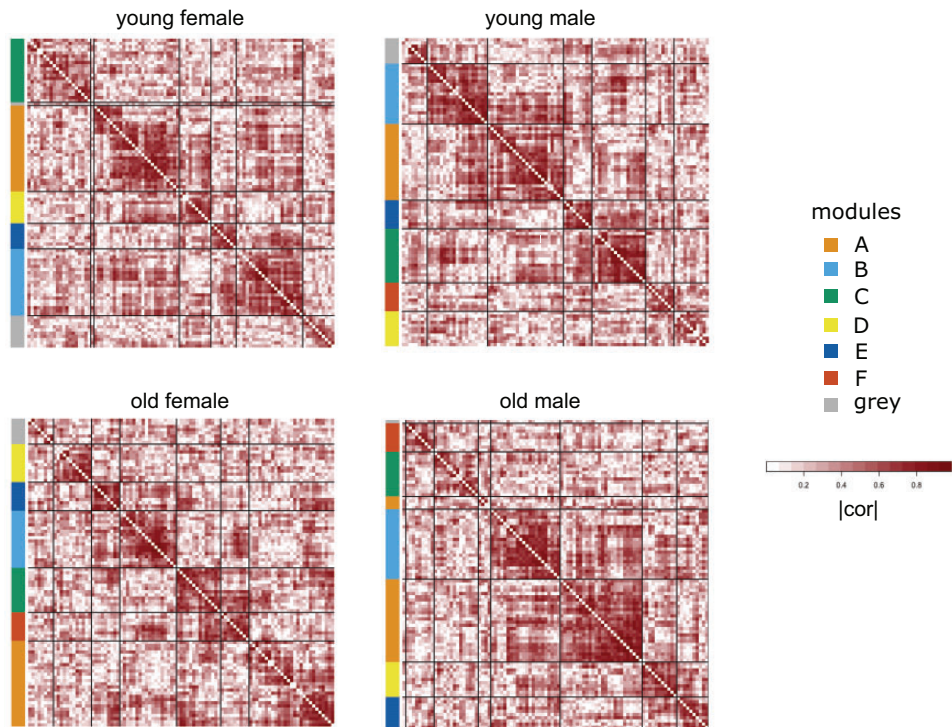
FIG. 2. Modular Evolution of the Metabolome. Heatmaps showing the pairwise absolute correlation between PICs for the 97 metabolites in each sex and age group (young 5 days, or old 31 days). The modules identified by WGCNA are shown on the left axis and partitions are included in the heatmap corresponding to these modules. Metabolites not fitting the criteria for modularity are in the "gray" category. Hierarchical clustering was done in WGCNA (Materials and Methods) and is shown in supplementary figure S4, Supplementary Material online.

lifespan on each targeted metabolite identified an association between lifespan and oxaloacetate (FDR = 0.007, $P < 1 \times 10^{-4}$, fig. 4C). The PICs of several other metabolites were significant at less conservative FDR (fig. 4 and supplementary table S6, Supplementary Material online). We also regressed lifespan PICs on the eigenmodules within each sex and age and identified module D in the older male metabolome associated with lifespan PICs ($r^2 = 0.80$, FDR = 0.01, table 2), and module B in the older female metabolome at a more modest FDR of 0.2 ($r^2 = 0.52$, table 2). Although neither of these modules was enriched for KEGG pathways after FDR correction, KEGG pathways whose enrichment had a nominal $P < 0.05$ are shown in supplementary table S5, Supplementary Material online.

Discussion

Evolution of the *Drosophila* Metabolome

Here, we examine the *Drosophila* metabolome and its associations with sex, age, and lifespan across 11 species. We find strong phylogenetic signal, with both Blomberg's K and Pagel's λ (Revell et al. 2008) significant in the metabolome of younger flies (table 1), and concordant with the genome-based phylogeny (figs. 1B and supplementary fig. S1, Supplementary Material online). Surprisingly, we see little evidence of selection acting to constrain divergence in overall metabolome variation (fig. 1C). However, there is considerable evidence for selection acting on metabolite levels when we analyze variation within coevolving modules (supplementary

table S3 and fig. S6, Supplementary Material online). Some of the coevolutionary modules enrich biological pathways, and modules in the older male and female metabolome associate with evolution in lifespan.

Before we discuss each of these points below, there are several possible caveats to keep in mind. First, to control for effects of environment, we have used a common garden design. However, it is perhaps inevitable that what is a viable environment for all species in the study might be far from ideal for some species. Second, although this study explores variation within as well as between species, we are capturing a very small snapshot of within-species variation, with only three strains per species. Moreover, the individual strains used are likely to be inbred, which could have a strong impact on phenotypic variation in general, and sexual dimorphism more specifically (Connallon and Clark 2014; Yassin et al. 2016; Ruzicka et al. 2019). Third, with only 11 species, we lack the power to fully explore the full evolutionary range of the metabolome, lifespan, and sexual dimorphism (Cressler et al. 2015) within *Drosophila*, let alone more broadly. Finally, using whole body samples likely obscures natural variation that manifests in organs, tissues, or cell types—variation that might ultimately drive phenotypic variation (Chintapalli et al. 2013).

The linear metabolome-wide divergence pattern that we see suggests that the *Drosophila* metabolome is not broadly constrained by stabilizing selection. The linearity we see contrasts with the plateau that some (Bedford and Hartl 2009; Ma et al. 2018), but not all (Khaitovich et al. 2004), have
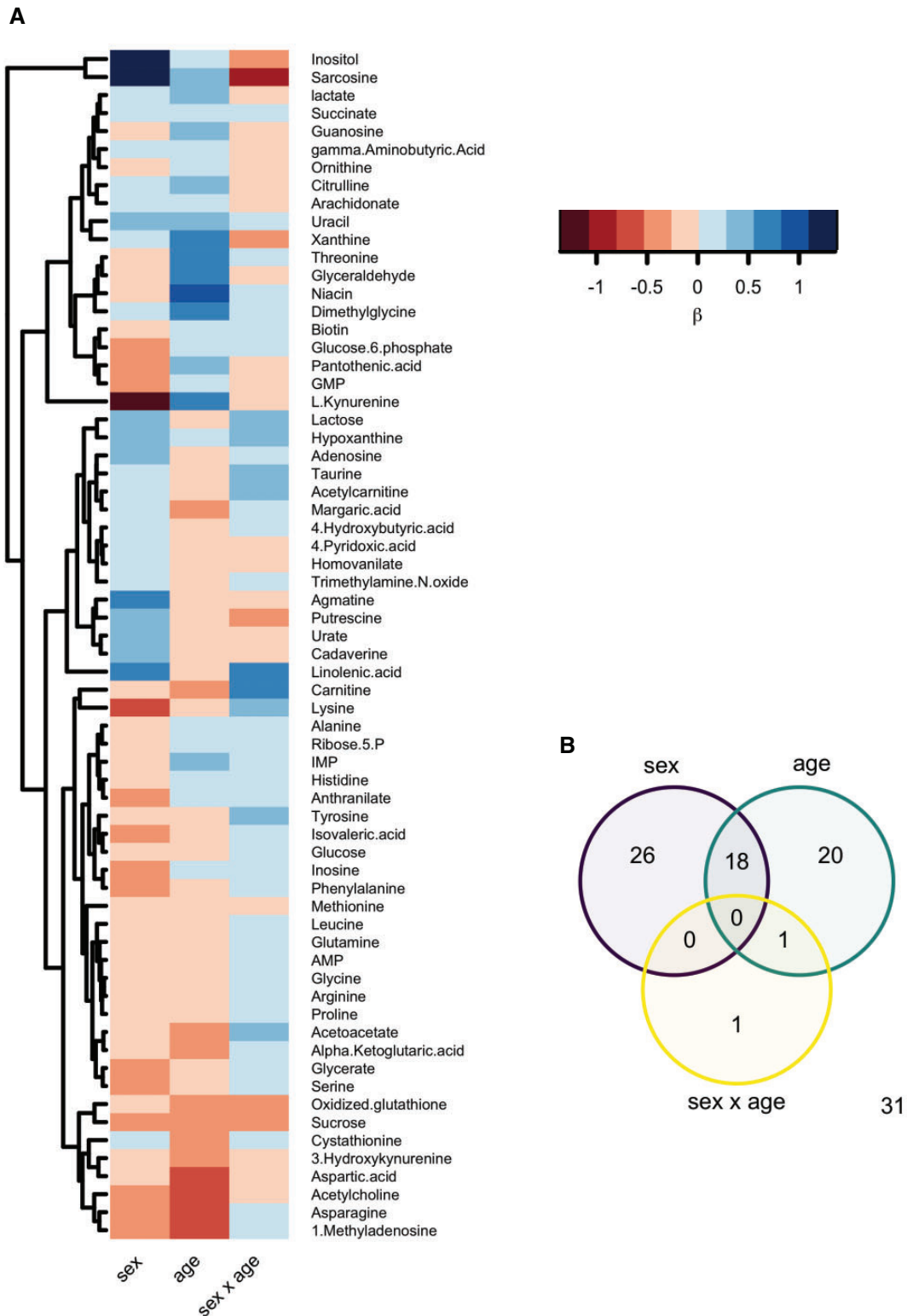
Fig. 3. Metabolite Variation by Sex and Age. (A) The effects (β) of sex, age, or their interaction on metabolite levels in *Drosophila* estimated in a phylogenetic mixed model. For sex, metabolites that are more abundant in males have a positive β. The effects are clustered by metabolite (row). (B) The number of metabolites effected by sex, age, or their interaction ($n = 97$, FDR ≤ 0.05).

observed in transcriptome divergence among the *Drosophila*. There may be some difference in the nature of metabolome evolution compared with the transcriptome, and our results are similar to an analysis of metabolome divergence in

primates (Bozek et al. 2014). Although the whole metabolome may lack evidence of selective constraint, we also consider if selection acts on subsets of the metabolome differently and/or with varying strength. Our observation of
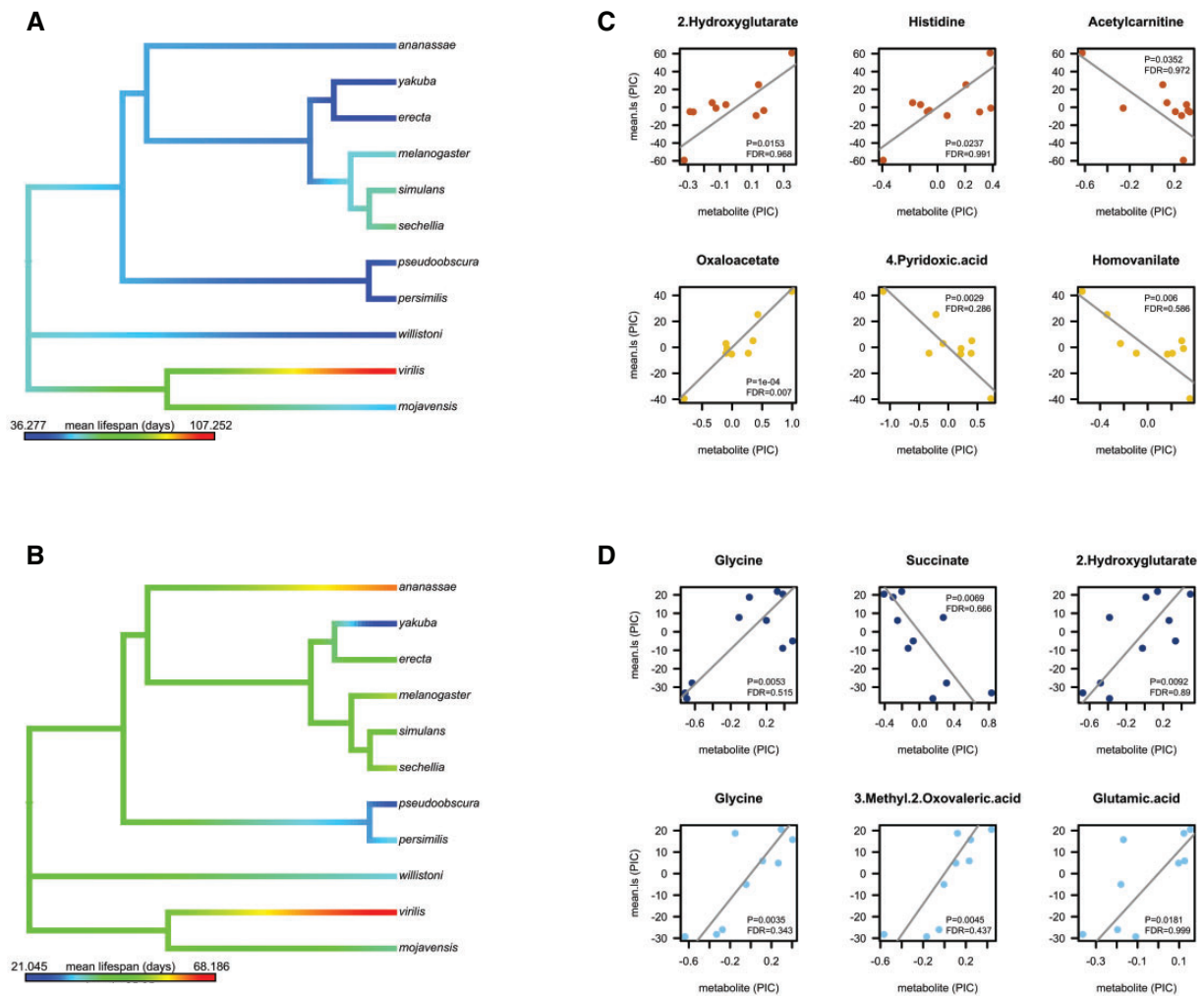
FIG. 4. Lifespan and Metabolite Coevolution. Mean lifespans of females (*A*) and males (*B*) from 11 *Drosophila* species are mapped on the phylogeny and a color scale is added based on estimated rates of continuous trait evolution. Means are from 16 to 378 individuals from each of 1–3 strains per species per sex. At each sex and age, PICs of mean lifespan for each species were regressed on PICs for 97 targeted metabolites using major axis regression forced through the origin. Plots of the regression for the three lowest *P* values from females (*C*) and males (*D*) of each age group, and the FDR of each association is show inside the plots. The regression data for all metabolites are shown in supplementary table S6, Supplementary Material online.

better fits for the OU model to the divergence of some modules is consistent with this idea. Selection that affects subsets of the metabolome is an expectation of metabolic adaptation and the evolution of biochemical pathways (Hartwell et al. 1999; Flowers et al. 2007; Wagner et al. 2007; Wagner 2009).

Several theories predict that covarying or coevolving traits will tend to include components of functional modules, either in development, cellular interactions, or other biological processes (Cheverud 1984; Hartwell et al. 1999; Wagner et al. 2007; Wagner and Zhang 2011; Collet et al. 2018). To our knowledge, this is the first study to examine the coevolution of the metabolome. Most studies of modularity in the evolution of endophenotypes have compared gene content, where the evolutionary persistence of homologs is an indication of evolutionary conservation (von Mering et al. 2003; Snel and Huynen 2004; Li et al. 2014). Comparative analyses of gene coexpression have also detected modular structure, either in

the genes encoding interacting proteins, or among members of biological pathways (Martin and Fraser 2018; Cope et al. 2020), and others find conservation of within-species gene coexpression across taxa (Stuart et al. 2003; Oldham et al. 2006). There is evidence that gene coexpression within *Drosophila* species can predict the axes of variation between *Drosophila* species (Innocenti and Chenoweth 2013). However, Martin and Fraser (2018) find no evidence that genes that covary across environmental or genetic backgrounds within species also covary over evolutionary time. These results indicate that such analyses are complementary ways to gain insight into patterns of endophenotypic covariation (Dunn et al. 2018), and ultimately, both approaches may shed light on the functional interrelations among genes, their products, and organismal phenotypes. Similar to comparative analysis of gene coexpression, our study sought to identify covariation in metabolite levels across the *Drosophila* to shed light on the nature of evolutionary change in this phylogeny.

Table 2. Associations between Metabolite Coevolutionary Modules and PICs of Mean Lifespan at Each Age and Sex (Group), Major Axis Regression through the Origin Was Used to Test the Association between PICs of Mean Lifespan and the Eigenmodules of the Metabolite PICs.

Group	Module	r sq	P value	FDR
Young female	E	0.204	0.1637	1
Young female	D	0.123	0.2902	1
Young female	C	0.091	0.3678	1
Young female	A	0	0.9962	1
Young female	B	0	0.9502	1
Old female	B	0.52	0.0186	0.2
Old female	F	0.227	0.1641	0.9
Old female	E	0.142	0.2824	1
Old female	A	0.06	0.4944	1
Old female	C	0.023	0.6751	1
Old female	D	0.004	0.8706	1
Young male	C	0.391	0.0395	0.47
Young male	B	0.184	0.1877	0.96
Young male	D	0.149	0.2405	0.96
Young male	A	0.075	0.4167	1
Young male	E	0.022	0.6606	1
Young male	F	0.015	0.7189	1
Old male	D	0.798	0.0005	0.01
Old male	E	0.412	0.0454	0.27
Old male	C	0.177	0.2262	0.68
Old male	B	0.046	0.5528	1
Old male	A	0.042	0.5707	1
Old male	F	0.005	0.8444	1

Note.—The model r-squared (r sq) and P value (P value) are shown, and multiple comparisons were FDR corrected (FDR).

Effects of Sex on Metabolome Profiles

In addition to the modular nature of covariation in metabolite levels, we also find two interesting patterns of sex-specific variation. These include sex differences in coevolutionary patterns among metabolites (fig. 1B), and the difference in metabolome sexual dimorphism between the *Sophophora* and *Drosophila* subgenera (supplementary fig. S7, Supplementary Material online). Given that metabolites are the building blocks of downstream traits, the relative lack of dimorphism in these metabolomic principal components for the *Sophophora* subgenus might reflect a downstream trait or set of traits that are sexually monomorphic in this lineage but not the other. Further work will be needed to determine what that might be. There are several examples of evolved sexual dimorphism in organismal traits in the *Drosophila* (Kopp et al. 2000; Luo et al. 2019), and our results suggest that LC-MS techniques could provide mechanistic insights into such evolutionary change. In using principal components to identify dimorphic phenotypes, we are biased toward detecting the largest sources of latent variation and so we do not suggest that sexual dimorphism is not present in the metabolome of the *Sophophora*. We also note that in sampling whole flies for metabolomics analysis, at least for some metabolites we are likely detecting variation that reflects sex differences in reproductive structures of the abdomen. We took care to sample only virgins in our analysis to avoid the more profound effects of egg development in inseminated females; however, a future analysis of tissues without obvious sexual dimorphism would allow us to investigate sexual dimorphism in the metabolome while minimizing the influence of structural morphology.

Age, Lifespan, and the Metabolome

The evolutionary forces shaping phenotypic variation across the lifespan are central to theories of aging (Medawar 1946; Williams 1957; Hamilton 1966). We find that the metabolite composition of coevolutionary modules differs by age (supplementary fig. S5, Supplementary Material online), implying that metabolites that are within a coevolving module in the context of a young *Drosophila* might be a part of a separate coevolutionary module in older *Drosophila*. We emphasize that selection is most likely acting on the biological pathways, even on the activity of single enzymes, and not the level of individual metabolites per se, so it is not surprising to see metabolites whose levels reflect different evolutionary patterns in flies at different ages. Interestingly, we see evidence of coevolution of lifespan and metabolic modules in the older metabolome of both sexes, whereas we fail to detect such association in the younger metabolome (table 2).

We also find larger between-species divergence in the whole metabolome of old flies when compared with that of young flies (fig 1C). This pattern mirrors the increase in metabolome divergence with age observed within each species seen in mammals (Ivanisevic et al. 2016; Dansereau et al. 2019), and is consistent with predictions from evolutionary theory that age-related genetic variation should increase with age (Charlesworth and Hughes 1996; Moorad and Promislow 2009).

Most of what we know about the molecular mechanisms of aging is derived from lab studies of inbred lines in single species. In recent years, comparative studies have begun to probe evidence for genes and metabolites associated with inter-specific variation in lifespan (Ma et al. 2015, 2018; Cui et al. 2019; Kowalczyk et al. 2020). Previous work across a broad mammalian phylogeny identified metabolites

associated with lifespan, albeit among samples from dissimilar environments and with lifespans estimated in other studies (Ma et al. 2015). Our work is the first to use metabolomic approaches to study correlates of lifespan in a common garden design, where animal rearing is done together in a controlled environment and samples are taken from the same population in which lifespan is measured.

Here we detect metabolites that coevolve with lifespan, suggesting the potential of metabolomics to identify longevity-regulating pathways that are conserved across species. A similar comparative analysis of lifespan and the transcriptome in 14 *Drosophila* species identified few individual genes of strong effect, but provided evidence that sets of genes with marginally significant coevolutionary association with lifespan might be enriched for a small number of biological pathways (Ma et al. 2018). Ad hoc comparison of the lifespan-associated pathways identified by Ma et al. (2018) and those identified here finds little evidence of overlap, indicating either that the study designs were different enough to obscure commonality that might exist between the metabolome and transcriptome, or possibly that lifespan coevolves with sets of coexpressed genes and metabolites that are somehow involved in nonoverlapping processes. Regardless, the coevolution of metabolite modules and lifespan supports the theory that lifespan variation in diverse species is influenced at least in part by common biological pathways. Our results in no way exclude the possibility that species-specific mechanisms also shape the evolution of life history traits (Martin et al. 1996; Partridge and Gems 2002).

We see clues to conserved mechanisms of lifespan regulation in the relative enrichment of KEGG pathways (supplementary table S5, Supplementary Material online). With the caveat that FDR correction indicates that none of the pathways are enriched, in the metabolome of older females, the pathway dme04213, annotated as a "longevity regulating pathway—multiple species" reaches an empirical P value of 0.0093 (FDR = 0.911) and includes insulin-like signaling, mTOR signaling, and superoxide dismutase expression (supplementary table S5, Supplementary Material online). In addition to the modular analysis, lifespan associates with oxaloacetate in the older female metabolome. There is no direct connection between oxaloacetate and lifespan in *Drosophila* that we are aware of. However, oxaloacetate is part of the TCA cycle which has previously been implicated in lifespan variation in *D. melanogaster* (Talbert et al. 2015; Jin et al. 2020). Interestingly, oxaloacetate supplementation extends the lifespan of *Caenorhabditis elegans* (Williams et al. 2009). The association of lifespan variation and oxaloacetate levels, and a module that enriches known lifespan regulating pathways in various species, suggest that lifespan evolution may have common underlying mechanisms that are reflected in the metabolome.

Conclusion

Overall, we have shown that evolution in *Drosophila* co-occurs with interconnected effects on the metabolome. We find both broadly conserved effects of sex and age, as well as dynamic and phylogenetically independent correlated evolutionary change. The modular nature of coevolution among the metabolites measured here points not only to the individual metabolites, but also the broader biological processes, that underlie evolution of organismal phenotypes.

The metabolome describes the basic structural and functional building blocks of all organisms on the planet. In this light, it is surprising how little is known about the patterns and process that underlie its evolution over billions of years. The work we present here, althyough focused on questions related to the evolution of aging, is also an effort to stimulate a much deeper and broader exploration of metabolome evolution.

Materials and Methods

Species and Strains

This study included 11 species of *Drosophila*: *D. melanogaster, D. simulans, D. sechellia, D. ananassae, D. erecta, D. yakuba, D. willistoni, D. pseudoobscura, D. persimilis, D. mojavensis, and D. virilis*. We obtained all species from the Drosophila Species Stock Center then at the University of California San Diego, except for *D. mojavensis* lines, which were provided as a gift from Luciano Matzkin (University of Arizona). We obtained three wild-type lines from each species with the exception of *D. yakuba* and *D. erecta*, for which we obtained one wild-type strain (supplementary table S7, Supplementary Material online). This design enabled us to study intra- and interspecies variation in metabolite levels and lifespan.

Media and Fly Culture

To limit the effect of diet on metabolome profiles, all lines were reared on the same diet, banana medium, consisting of 14% peeled banana, 4% molasses, 4.75% corn syrup, 2.75% Brewer's yeast, 1% methylparaben, and 3% ethanol, solidified in 1.4% agar. For survival assays and metabolomics, flies were placed on banana medium in bottles and allowed to mate and lay eggs for 48 h at 24 °C and 50–60% RH, after which adults were removed, and offspring were allowed to develop. We collected virgin males and females into vials with banana medium under light CO_2 anesthesia within 8 h of eclosion, except *D. virilis*, which was collected within 12 h of eclosion.

Metabolomics Assay and Data Normalization
Targeted Metabolomics

For targeted metabolomic analysis, *Drosophila* were reared under the conditions described above, and three flies from each sex/strain combination at 5 and 31 days of age were flash frozen and stored at −80 °C. One to three samples were collected per sex for each species at each age, for a total of 93 targeted LC-MS/MS samples. Metabolites were extracted by homogenizing samples in 200 μl HPLC water (Sigma) using a TissueLyser II (Qiagen) for 6 min at 25 Hz at 4 °C. We then added 800 μl methanol and incubated for 30 min on dry ice. The suspension was homogenized again for 10 min at 25 Hz, then spun at 14,000 rcf for 10 min at 4 °C. The supernatant was transferred to a new tube and dried in a Speed-Vac at 30 °C.

LC-MS/MS experiments were performed on an Agilent 1260 LC (Agilent Technologies, Santa Clara, CA)-AB Sciex QTrap 5500 mass spectrometer (AB Sciex, Toronto, ON, Canada) system at the University of Washington Northwest Metabolomics Research Center (UWNMRC). Each sample was injected twice, 10 µl for analysis using negative ionization mode and 2 µl for analysis using positive ionization mode. Both chromatographic separations were performed in hydrophilic interaction chromatography (HILIC) mode on two Waters XBridge BEH Amide columns (150 × 2.1 mm, 2.5 µm particle size, Waters Corporation, Milford, MA) connected in parallel. The flow rate was 0.300 ml/min, autosampler temperature was kept at 4 °C, and the column compartment was set at 40 °C. The mobile phase was composed of Solvents A (5 mM ammonium acetate in 90% H_2O/ 10% acetonitrile + 0.2% acetic acid) and B (5 mM ammonium acetate in 90%acetonitrile/10% H_2O + 0.2% acetic acid). After the initial 2 min isocratic elution of 90% B, the percentage of Solvent B decreased to 50% at $t = 5$ min. The composition of Solvent B maintained at 50% for 4 min ($t = 9$ min), and then the percentage of B gradually went back to 90%, to prepare for the next injection. Targeted data acquisition was performed in multiple reaction monitoring (MRM) mode. The LC-MS system was controlled by Analyst 1.5 software (AB Sciex). The extracted MRM peaks were integrated using MultiQuant 2.1 software (AB Sciex). Samples were spiked with ^{13}C internal standards, and two types of LC-MS/MS quality control (QC) samples were run at 11 evenly spaced intervals throughout the run to track potential drift in the assay. One QC sample was a pool of 10 fly samples, and the other was a sample of human serum. The CV for these QCs was 8.2% and 7.7%, respectively. Detailed LC-MS methods and data are available on www.metabolomicsworkbench.org doi: 10.21228/M8J11K.

Untargeted Metabolomics

Flies were raised as above and at 5 days old, 1–3 biological replicates of 10 flies from each strain and sex were flash frozen in 1.5-ml tubes in liquid N_2 and stored at −80 °C. This gave a total of 168 samples, with 3 replicates for 53 of the 58 strain and sex combinations. Metabolite extraction was identical to the procedure for targeted LC-MS-MS described above. Untargeted analysis was completed using an Agilent 1200 SL LC-6520 Quadrupole-Time of Flight (TOF)-MS system (Agilent Technologies) at the UWNMRC. The separation conditions for the LC-TOF-MS experiments were the same as those for the LC-MS/MS described above. The ESI voltage was 3.8 kV, and the m/z scan range was 60–1,000. The LC-TOF-MS data were extracted using Agilent MassHunter Qualitative Analysis (version B.07.00), Quantitative Analysis (version B.07.01), and Mass Profiler Professional (MPP, version B.13.00) software. The absolute intensity threshold for the LC-TOF-MS data extraction was 1,000, and the mass accuracy limit was set to 10 ppm. Missingness was assigned to peaks below 4,500 counts per second. Detailed methods and data are available at www.metabolomicsworkbench.org doi: 10.21228/M8J11K.

All statistical analyses were conducted in R version 4.0.3 (R Core Team 2018) unless otherwise stated. The targeted and untargeted metabolome data were analyzed separately. All targeted metabolites and untargeted features were log_e-normalized and the data from each sample were centered by subtracting the sample mean. For the untargeted LC-TOF-MS, the positive and negative mode data were combined, giving 4,419 features. Only 362 features were detected in every strain ($n = 1$–3 for each sex). In the 228 features with a single missing observation among all strains, missing values were imputed by ten nearest-neighbor imputation, resulting in 590 complete features. Principal components analysis (PCA) was performed using prcomp on the sample centered and scaled log abundance of 590 untargeted metabolite features, and on the 97 targeted metabolites.

Drosophila Phylogeny

We use the topology, branch lengths, and estimated divergence time of the consensus tree available at http://www.timetree.org/ (accessed January 2021; Kumar et al. 2017). The Newick format is available in Supplementary Material online .

Phylogenetic Signal and Multivariate Clustering

To measure the phylogenetic signal in the metabolomic profile, we estimated Pagel's λ and Blomberg's K, using the phylosig function within the phytools R package (Revell 2012). We used the strain-level data ($n = 1$–3) to estimate the standard error of the species-level means and used these errors as a measure of intraspecies variance (Ives et al. 2007; Revell 2012). For the two species with only a single strain, D. erecta and D. yakuba, we used the maximum standard error among the remaining species as the standard error. The significance of K was determined by 10^5 randomizations, and of λ using the likelihood ratio test.

For hierarchical clustering of the metabolome, we separately analyze the sexes and, for the targeted data, both ages as well. In all cases, the species means ($n = 1$–3) of each metabolite feature were used. We constructed trees using the complete linkage method of hierarchical clustering in R and evaluated node support by bootstrapping 1,000 times. We also compare each tree to the genome-based phylogeny by the branch score method (Kuhner and Felsenstein 1994) with the dist.topo function of the ape R package, using relative branch lengths of the phylogenies which were calculated by dividing branch lengths of each tree by the respective total branch lengths (Kuhner and Felsenstein 1994). We used a permutation approach to test the significance of each comparison by calculating the branch scores of 10^5 randomized phylogenies, made using the rtree function of the ape package, in comparison to the real phylogeny.

Modeling Metabolome Evolution

We modeled the divergence of the metabolome using a similar method used for divergence in the transcriptome (Bedford and Hartl 2009; Ma et al. 2018), by measuring the metabolome distance (y) as the squared difference between

each $\log_e$-metabolite among pairs of strains from the same age and sex. We first fit a linear model, analogous to the BM model, and to evaluate the effect of sex and age—young or old as a categorical variable—on the rate of metabolome-wide divergence:

$$y \sim a + \delta^2 x + \beta_{age} + \beta_{sex} + \beta_{x*\ age} + \beta_{x*\ sex} \\ + \beta_{x*\ age*sex} + \varepsilon, \tag{1}$$

where y is the metabolome distance, a is the intercept, δ^2 is analogous to the force of mutation/drift in the BM model, x is the evolutionary divergence time, ε is the residual error, and the separate β values represent the main or interaction effects of each model term, which were tested for significance by ANOVA.

For model comparison, the following BM model was fit by maximum likelihood using the `nlreg` R package:

$$y = a + \delta^2 x, \tag{2}$$

where δ^2, the slope, is analogous to the force of mutation/drift; and a is the intercept. For the OU model we used the formula of Ma et al. (2018):

$$y = \frac{\delta^2}{2\alpha}(1 - e^{-2\alpha x}). \tag{3}$$

The selection (α) and drift (δ^2) parameters were estimated using maximum likelihood. Because the OU model is undefined where divergence time (t) is zero, the intraspecies variance, we excluded data from $t = 0$ while fitting both BM and OU. The BM and OU models were compared using the Akaike information criterion (AIC).

Metabolome Coevolution and Modularity
The evolutionary relationship between metabolites was measured as the Pearson correlation among PICs for each metabolite pair. Within each sex and age group, PICs were estimated for each metabolite using the `pic.ortho` function in the `ape` package using multiple measurements per species (Felsenstein 2008). We compared the distribution of correlations between PICs to the distributions among 100 permutations of the species labels, which maintained the relationships between metabolites within each species, and effectively randomized the trait values across the phylogeny. We measured modularity by first defining networks of coevolving metabolites, where edges correspond to metabolite pairs whose PICs correlated at $r > 0.7$, and then identified modules using the `edge.betweeness.community` function in the `iGraph` package (Girvan and Newman 2002; Csardi and Nepusz 2006). Modularity was measured using the modularity function and was then compared with 1,000 rewired networks with the same degree distribution, and empirical P values were calculated.

To identify coevolutionary modules, we used the `WGCNA` package to find covarying PICs within each sex and age group (Langfelder and Horvath 2008). A topological overlap matrix for network construction was made using the

`TOMsimilarity` function on a matrix of pairwise Pearson correlations between all metabolite PICs raised to the power of 7. Modules were identified by the `cutreeDynamic` function using a deepSplit of 3 and a minimum module size of 8. The first principal component of the PICs of each module (eigenmodule) was used as module-level vector in regression models.

Enrichment Analysis
The 97 targeted metabolites measured here are distributed broadly over many metabolic pathways and thus we lack power to detect enrichment of the majority of individual pathways using hypergeometric testing. Instead, we used the network diffusion-based analysis in the `FELLA` package to evaluate enrichment within the *D. melanogaster* KEGG Release 97.0: 128 pathways, 165 modules, 749 enzymes, 5,417 reactions, and 3,961 compounds (Kanehisa and Goto 2000; Picart-Armada et al. 2018). To evaluate the significance of the enrichment, we ran 10^5 iterations where the significant metabolites were permuted over the full set of 97 metabolites. We used this analysis to evaluate both metabolites with effects of sex, age, or their interaction, as well as the enrichment among metabolites of each coevolving module in each sex and age. Multiple comparisons within each node type (pathway or KEGG module) were corrected using the Benjamini–Hochberg FDR method (Benjamini and Hochberg 1995).

Mixed Model Analysis
To determine the effects of sex and age on individual metabolites, we used the MCMCglmm package to analyze a phylogenetic mixed model, with sex and age and their interaction as fixed effects, and random effects of species and strain (Hadfield 2010):

$$\text{Metabolite} \sim \beta_{sex} + \beta_{age} + \beta_{sex\ *\ age} + \text{Species} \\ + \text{Strain} + \Sigma^{-1} + \varepsilon, \tag{4}$$

where β are the fixed effects of sex and age and their interaction, and Σ^{-1}, the inverse of the phylogenetic correlation matrix, was calculated using the *Drosophila* phylogeny, after resolving the *D. virilus*, *D. mojavensis*, and *D. willistoni* polytomy at the root by adding 1×10^{-4} My to all edges, which adds a trivial distance to the *D. willistoni* edge. Errors were assumed to be Gaussian, priors were set at $V = 1$ and $\nu = 0.02$, and $\geq 5 \times 10^5$ iterations were run to estimate posterior effects.

Lifespan Analysis
At the time of virgin collection (see Media and Fly Culture), approximately 20 flies of each sex/genotype were placed in individual vials (2–7 replicates, mean = 4.8). Flies were transferred onto new food three times over a 7-day period before the start of the longevity assay. Deaths during this period were not recorded as they may simply be due to extrinsic mortality from handling. On the seventh day, flies were transferred into randomized vials and the vials were censused three times per week thereafter while transferring flies to fresh food vials. Data

were recorded using the dLife software (Linford et al. 2013), and recording continued until all flies in all vials were dead. There were no censorship events during the experiments, so mean lifespan for each strain was simply the arithmetic mean of age at death for all individuals in that strain.

Coevolution between lifespan and metabolites was evaluated by major axis regression forced through the origin in the `smatr` package, where PICs of lifespan were regressed either on PICs of individual metabolites, or on the eigenmodules of the coevolutionary modules identified previously (see Metabolome Coevolution and Modularity). The metabolome at the two ages was compared with the same mean lifespan for that strain and sex. Therefore, multiple comparisons were handled by controlling FDR for tests of all predictors at both ages; $n = 194$ metabolites for the univariate analysis, and $n = 10$–11 eigenmodules for multivariate analysis.

Supplementary Material

Supplementary data are available at *Molecular Biology and Evolution* online.

Acknowledgments

Data Availability

The data underlying this article and the code used in the analysis are available on GitHub, at https://github.com/ben6uw/HarrisonBR_HoffmanJM_2021. The LC-MS data and detailed methods are available on Metabolomics Workbench: doi: 10.21228/M8J11K

References

Bai Y, Casola C, Feschotte C, Betrán E. 2007. Comparative genomics reveals a constant rate of origination and convergent acquisition of functional retrogenes in *Drosophila*. *Genome Biol*. 8(1):R11.

Ballard JWO. 2000. Comparative genomics of mitochondrial DNA in members of the *Drosophila melanogaster* subgroup. *J Mol Evol*. 51(1):48–63.

Bedford T, Hartl DL. 2009. Optimization of gene expression by natural selection. *Proc Natl Acad Sci U S A*. 106(4):1133–1138.

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R. Stat Soc. Ser B Methodol*. 57(1):289–300.

Blekhman R, Perry GH, Shahbaz S, Fiehn O, Clark AG, Gilad Y. 2014. Comparative metabolomics in primates reveals the effects of diet and gene regulatory variation on metabolic divergence. *Sci Rep*. 4:5809.

Bozek K, Khrameeva EE, Reznick J, Omerbasic D, Bennett NC, Lewin GR, Azpurua J, Gorbunova V, Seluanov A, Regnard P, et al. 2017. Lipidome determinants of maximal lifespan in mammals. *Sci Rep*. 7(1):5.

Bozek K, Wei Y, Yan Z, Liu X, Xiong J, Sugimoto M, Tomita M, Paabo S, Pieszek R, Sherwood CC, et al. 2014. Exceptional evolutionary divergence of human muscle and brain metabolomes parallels human cognitive and physical uniqueness. *PLoS Biol*. 12(5):e1001871.

Bozek K, Wei Y, Yan Z, Liu X, Xiong J, Sugimoto M, Tomita M, Paabo S, Sherwood CC, Hof PR, et al. 2015. Organization and evolution of brain lipidome revealed by large-scale analysis of human, chimpanzee, macaque, and mouse tissues. *Neuron* 85(4):695–702.

Brawand D, Soumillon M, Necsulea A, Julien P, Csardi G, Harrigan P, Weier M, Liechti A, Aximu-Petri A, Kircher M, et al. 2011. The evolution of gene expression levels in mammalian organs. *Nature* 478(7369):343–348.

Carson HL, Kaneshiro KY. 1976. *Drosophila* of Hawaii: systematics and ecological genetics. *Annu Rev Ecol Syst*. 7(1):311–345.

Charlesworth B, Hughes KA. 1996. Age-specific inbreeding depression and components of genetic variance in relation to the evolution of senescence. *Proc Natl Acad Sci U S A*. 93(12):6140–6145.

Cheverud JM. 1984. Quantitative genetics and developmental constraints on evolution by selection. *J Theor Biol*. 110(2):155–171.

Chintapalli VR, Al Bratty M, Korzekwa D, Watson DG, Dow JA. 2013. Mapping an atlas of tissue-specific *Drosophila melanogaster* metabolomes by high resolution mass spectrometry. *PLoS One*. 8(10):e78066.

Clark AG, Eisen MB, Smith DR, Bergman CM, Oliver B, Markow TA, Kaufman TC, Kellis M, Gelbart W, Iyer VN, et al.; Drosophila 12 Genomes Consortium. 2007. Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature* 450(7167):203–218.

Collet JM, McGuigan K, Allen SL, Chenoweth SF, Blows MW. 2018. Mutational pleiotropy and the strength of stabilizing selection within and between functional modules of gene expression. *Genetics* 208(4):1601–1616.

Connallon T, Clark AG. 2014. Balancing selection in species with separate sexes: insights from Fisher's geometric model. *Genetics* 197(3):991–1006.

Cope AL, O'Meara BC, Gilchrist MA. 2020. Gene expression of functionally-related genes coevolves across fungal species: detecting coevolution of gene expression using phylogenetic comparative methods. *BMC Genomics*. 21(1):370.

Coyne JA, Orr HA. 1989. Patterns of speciation in *Drosophila*. *Evolution* 43(2):362–381.

Cressler CE, Butler MA, King AA. 2015. Detecting adaptive evolution in phylogenetic comparative analysis using the Ornstein-Uhlenbeck model. *Syst Biol*. 64(6):953–968.

Csardi G, Nepusz T. 2006. The igraph software package for complex network research. *InterJ Compl Syst*. 1695:1–9.

Cui R, Medeiros T, Willemsen D, Iasi LNM, Collier GE, Graef M, Reichard M, Valenzano DR. 2019. Relaxed selection limits lifespan by increasing mutation load. *Cell* 178(2):385–399.e320.

Dansereau G, Wey TW, Legault V, Brunet MA, Kemnitz JW, Ferrucci L, Cohen AA. 2019. Conservation of physiological dysregulation signatures of aging across primates. *Aging Cell*. 18(2):e12925.

Dobzhansky T. 1946. Genetics of natural populations; recombination and variability in populations of *Drosophila pseudoobscura*. *Genetics* 31:269–290.

Dunn CW, Zapata F, Munro C, Siebert S, Hejnol A. 2018. Pairwise comparisons across species are problematic when analyzing functional genomic data. *Proc Natl Acad Sci U S A*. 115(3):E409–E417.

Felsenstein J. 2008. Comparative methods with sampling error and within-species variation: contrasts revisited and revised. *Am Nat.* 171(6):713–725.

Flowers JM, Sezgin E, Kumagai S, Duvernell DD, Matzkin LM, Schmidt PS, Eanes WF. 2007. Adaptive evolution of metabolic pathways in *Drosophila. Mol Biol Evol.* 24(6):1347–1354.

Fraser HB, Hirsh AE, Wall DP, Eisen MB. 2004. Coevolution of gene expression among interacting proteins. *Proc Natl Acad Sci U S A.* 101(24):9033–9038.

Fu X, Giavalisco P, Liu X, Catchpole G, Fu N, Ning Z-B, Guo S, Yan Z, Somel M, Pääbo S, et al. 2011. Rapid metabolic evolution in human prefrontal cortex. *Proc Natl Acad Sci U S A.* 108(15):6181–6186.

Ge H, Liu Z, Church GM, Vidal M. 2001. Correlation between transcriptome and interactome mapping data from Saccharomyces cerevisiae. *Nat Genet.* 29(4):482–486.

Girvan M, Newman MEJ. 2002. Community structure in social and biological networks. *Proc Natl Acad Sci U S A.* 99(12):7821–7826.

Gordon KL, Ruvinsky I. 2012. Tempo and mode in evolution of transcriptional regulation. *PLoS Genet.* 8(1):e1002432.

Hadfield JD. 2010. MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *J Stat Softw.* 33(2):1–22.

Hamilton WD. 1966. The moulding of senescence by natural selection. *J Theor Biol.* 12(1):12–45.

Hartwell LH, Hopfield JJ, Leibler S, Murray AW. 1999. From molecular to modular cell biology. *Nature* 402(6761 Suppl):C47–C52.

Hoffman JM, Soltow QA, Li S, Sidik A, Jones DP, Promislow DE. 2014. Effects of age, sex, and genotype on high-sensitivity metabolomic profiles in the fruit fly, *Drosophila melanogaster. Aging Cell.* 13(4):596–604.

Innocenti P, Chenoweth SF. 2013. Interspecific divergence of transcription networks along lines of genetic variance in *Drosophila*: dimensionality, evolvability, and constraint. *Mol Biol Evol.* 30(6):1358–1367.

Ivanisevic J, Stauch KL, Petrascheck M, Benton HP, Epstein AA, Fang M, Gorantla S, Tran M, Hoang L, Kurczy ME, et al. 2016. Metabolic drift in the aging brain. *Aging* 8(5):1000–1020.

Ives AR, Midford PE, Garland T. Jr. 2007. Within-species variation and measurement error in phylogenetic comparative methods. *Syst Biol.* 56(2):252–270.

Jin K, Wilson KA, Beck JN, Nelson CS, Brownridge GW 3rd, Harrison BR, Djukovic D, Raftery D, Brem RB, Yu S, et al. 2020. Genetic and metabolomic architecture of variation in diet restriction-mediated lifespan extension in *Drosophila. PLoS Genet.* 16(7):e1008835.

Jones DP, Park Y, Ziegler TR. 2012. Nutritional metabolomics: progress in addressing complexity in diet and health. *Annu Rev Nutr.* 32:183–202.

Kambysel MP, Heed WB. 1971. Studies of oogenesis in natural populations of drosophilidae. 1. Relation of ovarian development and ecological habitats of Hawaiian species. *Am Nat.* 105(941):31–49.

Kanehisa M, Goto S. 2000. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28(1):27–30.

Kellermann V, van Heerwaarden B, Sgro CM, Hoffmann AA. 2009. Fundamental evolutionary limits in ecological traits drive *Drosophila* species distributions. *Science* 325(5945):1244–1246.

Khaitovich P, Lockstone HE, Wayland MT, Tsang TM, Jayatilaka SD, Guo AJ, Zhou J, Somel M, Harris LW, Holmes E, et al. 2008. Metabolic changes in schizophrenia and human brain evolution. *Genome Biol.* 9(8):R124.

Khaitovich P, Weiss G, Lachmann M, Hellmann I, Enard W, Muetzel B, Wirkner U, Ansorge W, Paabo S. 2004. A neutral model of transcriptome evolution. *PLoS Biol.* 2(5):E132.

Khrameeva E, Kurochkin I, Bozek K, Giavalisco P, Khaitovich P. 2018. Lipidome evolution in mammalian tissues. *Mol Biol Evol.* 35(8):1947–1957.

Klingenberg CP. 2014. Studying morphological integration and modularity at multiple levels: concepts and analysis. *Philos Trans R Soc Lond B Biol Sci.* 369(1649):20130249–20130249.

Kopp A, Duncan I, Godt D, Carroll SB. 2000. Genetic control and evolution of sexually dimorphic characters in *Drosophila. Nature* 408(6812):553–559.

Kowalczyk A, Partha R, Clark NL, Chikina M. 2020. Pan-mammalian analysis of molecular constraints underlying extended lifespan. *eLife* 9:e51089.

Kuhner MK, Felsenstein J. 1994. A simulation comparison of phylogeny algorithms under equal and unequal evolutionary rates. *Mol Biol Evol.* 11(3):459–468.

Kumar S, Stecher G, Suleski M, Hedges SB. 2017. TimeTree: a resource for timelines, timetrees, and divergence times. *Mol Biol Evol.* 34(7):1812–1819.

Langfelder P, Horvath S. 2008. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9:559.

Li Q, Bozek K, Xu C, Guo Y, Sun J, Paabo S, Sherwood CC, Hof PR, Ely JJ, Li Y, et al. 2017. Changes in lipidome composition during brain development in humans, chimpanzees, and macaque monkeys. *Mol Biol Evol.* 34(5):1155–1166.

Li Y, Calvo SE, Gutman R, Liu JS, Mootha VK. 2014. Expansion of biological pathways based on evolutionary inference. *Cell* 158(1):213–225.

Linford NJ, Bilgir C, Ro J, Pletcher SD. 2013. Measurement of lifespan in *Drosophila melanogaster. J Vis Exp.* (71):50068.

Luo Y, Zhang Y, Farine JP, Ferveur JF, Ramírez S, Kopp A. 2019. Evolution of sexually dimorphic pheromone profiles coincides with increased number of male-specific chemosensory organs in *Drosophila prolongata. Ecol Evol.* 9(23):13608–13618.

Ma S, Avanesov AS, Porter E, Lee BC, Mariotti M, Zemskaya N, Guigo R, Moskalev AA, Gladyshev VN. 2018. Comparative transcriptomics across 14 *Drosophila* species reveals signatures of longevity. *Aging Cell.* 17(4):e12740.

Ma S, Yim SH, Lee SG, Kim EB, Lee SR, Chang KT, Buffenstein R, Lewis KN, Park TJ, Miller RA, et al. 2015. Organization of the mammalian metabolome according to organ function, lineage specialization, and longevity. *Cell Metab.* 22(2):332–343.

Markow TA, Fogleman JC, Heed WB. 1983. Reproductive isolation in Sonoran desert *Drosophila. Evolution* 37(3):649–652.

Martin GM, Austad SN, Johnson TE. 1996. Genetic analysis of ageing: role of oxidative damage and environmental stresses. *Nat Genet.* 13(1):25–34.

Martin T, Fraser HB. 2018. Comparative expression profiling reveals widespread coordinated evolution of gene expression across eukaryotes. *Nat Commun.* 9(1):4963.

Medawar PB. 1946. Old age and natural death. *Modern Q.* 1:30–56.

Moorad JA, Promislow DE. 2009. What can genetic variation tell us about the evolution of senescence? *Proc Biol Sci.* 276(1665):2271–2278.

Noda-Garcia L, Liebermeister W, Tawfik DS. 2018. Metabolite-enzyme coevolution: from single enzymes to metabolic pathways and networks. *Annu Rev Biochem.* 87:187–216.

Oldham MC, Horvath S, Geschwind DH. 2006. Conservation and evolution of gene coexpression networks in human and chimpanzee brains. *Proc Natl Acad Sci U S A.* 103(47):17973–17978.

Park YH, Lee K, Soltow QA, Strobel FH, Brigham KL, Parker RE, Wilson ME, Sutliff RL, Mansfield KG, Wachtman LM, et al. 2012. High-performance metabolic profiling of plasma from seven mammalian species for simultaneous environmental chemical surveillance and bioeffect monitoring. *Toxicology* 295(1–3):47–55.

Partridge L, Gems D. 2002. Mechanisms of aging: public or private? *Nat Rev Genet.* 3(3):165–175.

Partridge L, Hoffmann A, Jones JS. 1987. Male size and mating success in *Drosophila melanogaster* and *Drosophila pseudoobscura* under field conditions. *Anim Behav.* 35(2):468–476.

Picart-Armada S, Fernández-Albert F, Vinaixa M, Yanes O, Perera-Lluna A. 2018. FELLA: an R package to enrich metabolomics data. *BMC Bioinformatics* 19(1):538.

R Core Team. 2018. R: a language and environment for statistical computing. Vienna (Austria): R Foundation for Statistical Computing.

Revell LJ. 2012. phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol Evolut.* 3(2):217–223.

Revell LJ, Harmon LJ, Collar DC. 2008. Phylogenetic signal, evolutionary process, and rate. *Syst Biol.* 57(4):591–601.

Ruzicka F, Hill MS, Pennell TM, Flis I, Ingleby FC, Mott R, Fowler K, Morrow EH, Reuter M. 2019. Genome-wide sexually antagonistic variants reveal long-standing constraints on sexual dimorphism in fruit flies. *PLoS Biol.* 17(4):e3000244.

Schnebel EM, Grossfield J. 1983. A comparison of life-span characteristics in *Drosophila*. *Exp Gerontol.* 18(5):325–337.

Snel B, Huynen MA. 2004. Quantifying modularity in the evolution of biomolecular systems. *Genome Res.* 14(3):391–397.

Spirin V, Gelfand MS, Mironov AA, Mirny LA. 2006. A metabolic network in the evolutionary context: multiscale structure and modularity. *Proc Natl Acad Sci U S A.* 103(23):8774–8779.

Stark A, Lin MF, Kheradpour P, Pedersen JS, Parts L, Carlson JW, Crosby MA, Rasmussen MD, Roy S, Deoras AN, et al.; Berkeley Drosophila Genome Project. 2007. Discovery of functional elements in 12 *Drosophila* genomes using evolutionary signatures. *Nature* 450(7167):219–232.

Steuer R. 2006. Review: on the analysis and interpretation of correlations in metabolomic data. *Brief Bioinform.* 7(2):151–158.

Stuart JM, Segal E, Koller D, Kim SK. 2003. A gene-coexpression network for global discovery of conserved genetic modules. *Science* 302(5643):249–255.

Talbert ME, Barnett B, Hoff R, Amella M, Kuczynski K, Lavington E, Koury S, Brud E, Eanes WF. 2015. Genetic perturbation of key central metabolic genes extends lifespan in *Drosophila* and affects response to dietary restriction. *Proc R Soc B.* 282(1815):20151646.

von Mering C, Zdobnov EM, Tsoka S, Ciccarelli FD, Pereira-Leal JB, Ouzounis CA, Bork P. 2003. Genome evolution reveals biochemical networks and functional modules. *Proc Natl Acad Sci U S A.* 100(26):15428–15433.

Wagner A. 2009. Evolutionary constraints permeate large metabolic networks. *BMC Evol Biol.* 9:231.

Wagner GP, Pavlicev M, Cheverud JM. 2007. The road to modularity. *Nat Rev Genet.* 8(12):921–931.

Wagner GP, Zhang J. 2011. The pleiotropic structure of the genotype-phenotype map: the evolvability of complex organisms. *Nat Rev Genet.* 12(3):204–213.

Whitkus R, Doebley J, Lee M. 1992. Comparative genome mapping of Sorghum and maize. *Genetics* 132(4):1119–1130.

Wilinski D, Winzeler J, Duren W, Persons JL, Holme KJ, Mosquera J, Khabiri M, Kinchen JM, Freddolino PL, Karnovsky A, et al. 2019. Rapid metabolic shifts occur during the transition between hunger and satiety in *Drosophila melanogaster*. *Nat Commun.* 10(1):4052.

Williams DS, Cash A, Hamadani L, Diemer T. 2009. Oxaloacetate supplementation increases lifespan in *Caenorhabditis elegans* through an AMPK/FOXO-dependent pathway. *Aging Cell.* 8(6):765–768.

Williams GC. 1957. Pleiotropy, natural selection, and the evolution of senescence. *Evolution* 11(4):398–411.

Yassin A, Bastide H, Chung H, Veuille M, David JR, Pool JE. 2016. Ancient balancing selection at tan underlies female colour dimorphism in *Drosophila* erecta. *Nat Commun.* 7:10400.