Title: Continent-wide structure of bacterial microbiomes of European *Drosophila melanogaster* suggests host-control

Authors

Yun Wang¹,², Martin Kapun¹,³,⁴, Lena Waidele¹,², Sven Kuenzel⁵, Alan Bergland¹,⁶,⁷, and Fabian Staubach*¹,²

*to whom correspondence should be addressed: fabian.staubach@biologie.uni-freiburg.de

Author affiliations

1 The European *Drosophila* Population Genomics Consortium (DrosEU)

2 Department of Biology I, University of Freiburg, Germany

3 Department of Biology, University of Fribourg, CH-1700 Fribourg, Switzerland

4 University of Lausanne, Department of Ecology and Evolution, Lausanne, Switzerland

5 Max Planck Institute for Evolutionary Biology, Ploen, Germany

6 Department of Biology, University of Virginia, Charlottesville, VA

7 The *Drosophila* Real Time Evolution Consortium (DrosRTEC)

Competing interests

The authors declare no conflict of interest.

Abstract

The relative importance of host-control, environmental effects, and stochasticity in the assemblage of host-associated microbiomes has been much debated. With recent sampling efforts, the underpinnings of *D. melanogaster*’s microbiome structure have become tractable on larger spatial scales. We analyzed the microbiome among fly populations that were sampled across Europe by the *European Drosophila Population*
Genomics Consortium (DrosEU). We combined environmental data on climate and food-substrate, dense genomic data on host population structure, and microbiome profiling. Food-substrate, temperature, and host population-structure correlated with microbiome-structure. The microbes, whose abundance was co-structured with host populations, also differed in abundance between flies and their substrate in an independent survey, suggesting host-control. Patterns of enrichment and depletion of microbes between host and substrate were consistent with a model of host-control, where the host manipulates its microbiome for its benefit. Putative host-control was bacterial strain specific, supporting recent evidence for high specificity of D. melanogaster-microbe interaction.

Introduction

Species interactions, such as the interactions between microbes and their hosts are expected to be frequently driven by competition for resources that results in evolutionary conflict (Queller and Strassmann, 2018). In particular, horizontally transmitted microbes can leave an exploited host and move on to the next host, favoring conflict (Ebert, 2013a). Nonetheless, microbes that benefit higher organisms are prevalent (e.g. Jaenike et al., 2010; Ankrah and Douglas, 2018; Bang et al., 2018). The frequent benefits that hosts derive from their microbiome are thought to result from strong selection pressure on the host to evolve selectivity towards beneficial microbes (Foster and Wenseleers, 2006; Schluter and Foster, 2012). This selectivity is also termed host-control. However, evidence for strong host-control over microbiome is limited for many organisms and its importance compared to environmental factors and stochastic processes in shaping host-associated microbiomes is debated.

Recent data across a variety of hosts, including mammals and insects, suggest that host-control might not be a major driver of microbiome structure. First, environmental factors like host-diet often have a dominant effect on microbiome composition (Chandler et
al., 2011; Wu et al., 2011; Staubach et al., 2013; Wang et al., 2014; Waidele et al., 2017; Rothschild et al., 2018). Second, the discovery of substantial variation of the microbiome between individuals of the same species (Ley et al., 2008; The Human Microbiome Project Consortium et al., 2012; Linnenbrink et al., 2013; Wang and Staubach, 2018) is also difficult to explain under strong host-control. To explain this variation, stochastic, ecologically neutral processes have moved into focus. These processes comprise ecological drift, dispersal, and colonization history (Hubbell, 2001). A dominance of neutral microbiome structuring principles appears plausible for diverse species (Sieber et al., 2018), including Drosophila (Adair et al., 2018).

Without strong host control it is difficult to understand why higher organisms receive benefits from their microbiome so frequently. This is because host-control is supposed to be an important driver of the evolution of microbiome derived benefits in a setting that would otherwise frequently result in conflict between host and microbe (Foster and Wenseleers, 2006; Schluter and Foster, 2012; Ebert, 2013b; Queller and Strassmann, 2018). A model that resolves this contradiction has recently been proposed by Foster et al. (2017). In their 'ecosystem on a leash' model, the microbiome behaves similar to an ecosystem that is mainly shaped by microbe-microbe interactions. Host-control (the 'leash') acts only on a small subset of microbes that affect host fitness. Targeted fostering and exclusion of a relatively small set of microbes in the microbiome might suffice to favor beneficial microbiome function (Schluter and Foster, 2012; Agler et al., 2016; Foster et al., 2017).

Low but effective levels of host-control as proposed by the 'ecosystem on a leash'-model, also holds promise to help us to better understand the interaction of Drosophila with its microbiome. In D. melanogaster, the microbiome has well documented effects on somatic growth and reproductive output (Storelli et al., 2011; Shin et al., 2011; Téfit and Leulier, 2017; Sannino et al., 2018; Sommer and Newell, 2018). The effects of the
The microbiome on these traits is caused, at least in part, by an increase in the efficiency of nutrient acquisition by the host in the presence of specific microbial taxa. For instance, reproductive (Pais et al., 2018) and nutritional (Sannino et al., 2018) benefits as well as protection from pathogens (Shin et al., 2011) can be derived from members of the family Acetobacteraceae. This family dominates the natural microbiome (Corby-Harris et al., 2007; Cox and Gilmore, 2007; Chandler et al., 2011; Barata et al., 2012; Staubach et al., 2013; Adair et al., 2018; Walters et al., 2018). However, the role of host-control in the prevalence of these potentially beneficial bacteria is unclear because the ability of D. melanogaster to shape its associated microbiome might be limited (Wong et al., 2013; Blum et al., 2013; Broderick et al., 2014). Instead, probabilistic processes contribute to gut colonization (Obadia et al., 2017) and community structure in natural populations can be explained to a large extent by neutral ecological mechanisms (Adair et al., 2018). As in mammals and other organisms, environmental factors, such as the time of collection (Behrman et al., 2018; Adair et al., 2018) or diet (Chandler et al., 2011; Staubach et al., 2013; Erkosar et al., 2018; Wang and Staubach, 2018) have a strong effect on the Drosophila microbiome. The food-substrate that flies live on can be a more important driver of adaptation for the microbes than the host environment in D. melanogaster (Martino et al., 2018).

On the other hand, D. melanogaster microbiome structure is associated with host genotype (Unckless et al., 2015; Chaston et al., 2016; Behrman et al., 2018), indicating genotype dependent host-control. Microbial communities within flies differ from that in their stool (Fink et al., 2013), suggesting a selection process inside the fly. Further evidence that D. melanogaster exerts control over its microbiome comes from a recent study, which found that D. melanogaster larvae potentially foster Lactobacillus plantarum via excretions (Storelli et al., 2018). Host-control by D. melanogaster can be highly specific and finely tuned. Pais et al. (2018) showed that Acetobacter thailandicus can persist in the gut of D.
melanogaster, can be dispersed by the host, and provide a fitness benefit to the host, while closely related Acetobacter strains cannot. In lab-reared flies, dysregulation of antimicrobial effectors leads to highly specific changes in microbiome composition; Intriguingly, these changes in community composition can preferentially select for non-pathogenic taxa over pathogenic ones, despite a close phylogenetic relationship between these microbial species (Ryu et al., 2008).

Host-control of the microbiome that increases host fitness is a key parameter of the 'ecosystem on a leash'- model. The evidence for host-control in D. melanogaster suggests that this model might help us to understand the prevalence of benefits that D. melanogaster derives from its microbiome in the face of horizontal transmission, high intra-specific variation of microbiomes, and strong environmental effects. Given that the model was originally developed with the mammalian microbiome in mind (Foster et al., 2017), its applicability to D. melanogaster would create a common framework to understand mammalian and D. melanogaster microbiomes. As a consequence, results could become more transferable between the systems. Obviously, this is highly desirable because D. melanogaster is one of the best developed model systems in biology with many advantages over mammalian models. Together with the relatively simple microbiome (Erkosar et al., 2013) that is dominated by bacteria (Kapun et al., 2018) it holds promise for unraveling the driving forces of host-microbiome interactions (Erkosar et al., 2013; Douglas, 2018).

What is currently missing to assess whether the 'ecosystem on a leash' model is applicable to D. melanogaster is more information on the extent of host-control on larger ecological scales and under natural conditions. These are the conditions where potential host-control has originated by evolutionary processes. If host-control is important for D. melanogaster, imprints of host-control on the microbiome should also become apparent in a natural setting. Such imprints could be reflected by co-structure of host population
genetic variation and the microbiome because host-control in *D. melanogaster* varies between natural populations (Behrman et al., 2018; Walters et al., 2018) and depends on host genotype (Unckless et al., 2015; Chaston et al., 2016; Pais et al., 2018). Furthermore, if the natural *D. melanogaster* microbiome is subject to host-control, it should show general properties of host-associated microbiomes, for example elevated levels of 16S gene copies. Finally, microbes that are subject to host-control should differ in abundance between the host and its environment because the effects of host-control outside the host should be smaller or absent.

In order to test these basic predictions and further assess the role of host-control in natural *D. melanogaster* populations, we profiled the bacterial microbiome, in 50 samples across Europe, using 16S rRNA gene sequencing (Figure 1 and Table S1). The sampling range covered different climates and allowed us to address the effect of environmental factors on the microbiome. We combined the 16S profiles with population level allele frequency data for more than 20,000 neutral SNPs to test for co-structure of the microbiome with host genetic variation. For further exploration of potential host-control, we tested whether microbiomes showed typical properties of host-associated communities, specifically in terms of increased 16S gene copy number. Finally, we identified bacterial taxa that correlated with host population structure. These taxa were analyzed in an independent survey, comparing fly-associated microbiomes to that of their substrate, to test whether these taxa show different abundance that indicates of host-control.

**Results**

We analyzed a total of 5,217,762 16S rRNA reads after quality filtering (Table S2). 2,672,402 *Wolbachia* sequences were removed. In order to make the diversity assessment comparable between samples, we rarefied (subsampled) the number of sequences per sample to 5,135 per sample for the continental range analysis and to 898...
sequences per sample for the fly versus substrate survey. We grouped the sequences into
100% identity Operational Taxonomic Units (OTUs) for high resolution analysis, unless
otherwise noted. We chose 100% identity to resolve strain level differences because the
interaction with the fly host may differ for closely related bacteria (Ryu et al., 2008; Pais et
al., 2018). Please note that the sequences were rigorously quality filtered and sequencing
errors were removed (see Materials and Methods).

Community composition and diversity

The *Drosophila* microbiome across Europe was dominated by acetic acid bacteria
(Acetobacteraceae 63.6%, Figure S1). The three most common genera were *Acetobacter*
(26.1%), *Gluconobacter* (17.4%), and *Commensalibacter* (15.4%). Enterobacteriaceae
were also common (15.2%). Shannon diversity was 2.61 +/- 0.65 (SD) at the 100%
identity OTU level and 1.99 +/- 0.53 (SD) at the 97% identity OTU level. Composition and
diversity of the *D. melanogaster* microbiome were similar to those reported previously from
natural *D. melanogaster* isolates. For a comparison of alpha diversity between studies see
Staubach et al. (2013).

The natural *D. melanogaster* bacterial microbiome is structured on a continental scale

As a first step to better understand the structuring principles of the *D. melanogaster*
microbiome, we tested whether the natural *Drosophila* associated microbiome is structured
on a continental scale. The absence of continental structure would be consistent with a
stochastic distribution of microbes and speak against both, host-control, as well as
selection of microbial taxa by other environmental factors.

Bray-Curtis-Dissimilarities (BCD) of the bacterial communities increased rapidly with
geographic distance (r = 0.196, P = 0.0015, Mantel-Test, Figure 2), indicating geographic
structure of microbial communities associated with *D. melanogaster* on a continent-wide scale.

Host genetic differentiation and temperature correlate with microbiome structure

In order to identify the factors underpinning continental structure, we modeled microbiome composition in a Redundancy Analysis (RDA) framework. We selected temperature, precipitation, and substrate as candidate environmental variables that could affect microbiome structure. We chose temperature because it can affect microbial communities on geographic scales (Thompson et al., 2017). Substrate is a major determinant of natural *Drosophila* microbiomes (Chandler et al., 2011; Staubach et al., 2013; Wang and Staubach, 2018). A lack of precipitation might affect microbiome assembly by selecting for xerotolerant microbes. Because microbial communities might reflect long-term or short-term trends in temperature and precipitation, we included annual, as well as monthly means of temperature and precipitation in our model. The inclusion of monthly temperature and precipitation at the time of collection allows us to assess seasonal variation in these parameters that could affect microbiomes. Finally, we reasoned that *Drosophila* host population structure might explain interpopulation differences in microbial community composition. This would be expected if potential host-control of the microbiome varied between host populations or if microbes and fly host show patterns of co-dispersal and parallel demographic histories on a continent-wide scale.

The full model (including all factors described above) explained approximately half (46.3%) of the variance in the bacterial microbiome, indicating that the model contained factors that affect the bacterial microbiome. In order to select the most relevant explanatory variables from the full model, we employed a forward model selection approach. This resulted in host genetic differentiation (PC1 and PC2), mean annual temperature (T(y)), and substrate as relevant factors for *Drosophila* microbiome.
composition at the 100% OTU level (Table 1). We further reasoned that the microbiome might be structured at a higher taxonomic level. In particular, Acetobacteraceae comprise many bacteria that are dispersed by *D. melanogaster* and convey benefits to their hosts (Shin et al., 2011; Barata et al., 2012; Pais et al., 2018). Conversely, many Enterobacteriaceae are *Drosophila* pathogens. Susceptibility and virulence of these bacteria varies between natural host populations (Behrman et al., 2018). By applying the same model selection approach at the bacterial family level, we also identified host genetic differentiation (PC1) and annual mean temperature (T(y)) as relevant factors for microbiome composition (Table 2).

The abundance of OTU2 (*Commensalibacter*), and Enterobacteriaceae co-vary with host population structure

We were interested to identify bacteria underlying the correlation of microbiome composition with host population structure. These bacteria might respond to potential differences in host-control between natural host populations. Therefore, we tested whether the relative abundance of OTUs correlated with host genetic variation. At the 100% identity OTU level, only the abundance of OTU2 (*Commensalibacter, Acetobacteraceae*) correlated with PC1 of host genetic variation (Figure 3, \(P = 0.00017, q = 0.0065, r = -0.52\), Pearson's Product-Moment correlation). This suggested strain level specificity of host effects on the microbiome. At the family level, Enterobacteriaceae (\(P = 0.037, q = 0.17, r = -0.30\), Pearson's correlation), Leuconostocaceae (\(P = 0.047, q=0.17, r = 0.29\), Pearson's correlation), and Acetobacteraceae (\(P = 0.039 , q = 0.17, r = -0.30\), Pearson's correlation) were structured according to PC1. However, when removing all sequences from OTU2, Acetobacteraceae were not significantly correlated with host population structure anymore (\(P = 0.16, q = 0.38\)), suggesting that OTU2 contributed significantly to the correlation. A representative sequence of OTU2 perfectly matched *Commensalibacter intestini* strain...
A911 (Roh et al., 2008), a previously described commensal of *D. melanogaster*. No individual OTU correlated with PC2 of host genetic variation.

No evidence for pronounced dispersal limitation of bacteria that correlate with host population structure

We hypothesized above that the microbiome could be affected by host-control that varies between host populations. Alternatively, a correlation of microbiomes with host population structure could result from bacteria that are dispersal limited and depend on the fly host for dispersal. If the bacteria were severely dispersal limited on a global scale, we would expect the occupied geographic range of the bacteria in question to be rather limited. However, this is not the case; the bacterial groups that are structured according to host population structure in Europe (OTU2, Enterobacteraceae, Acetobacteraceae, Leuconostocaceae) can also be found along the East Coast and on the West Coast of the USA (Figure 4). Furthermore, these bacterial groups, were also previously found in association with wild-caught *D. takahashii* from Hawaii, *D. seychellia* collected from morinda fruit on the Seychelles, cactus feeding *D. mojavensis* and even in mushroom feeding *Microdrosophila* (Chandler et al., 2011). A representative sequence of OTU2 matched sequences from these diverse locations and species perfectly (Chandler et al., 2011). This suggested that there is no pronounced dispersal limitation on a global scale for these bacteria and that the bacteria in question are rather cosmopolitan. Hence, a scenario, in which the bacteria are severely dispersal limited and depend on *D. melanogaster* dispersal on the continental scale, appears implausible.

16S copy number of the natural *Drosophila* bacterial microbiome is typical for host-associated communities
Because the bacteria that are co-structured with their host populations on the continental scale are cosmopolitan, dispersal effects seemed insufficient to explain the co-structure of microbiomes and host population genetic variation. Therefore, we reasoned that host-control might contribute to the co-structure. If the Drosophila microbiomes that we analyzed were subject to host-control, they should differ from environmental microbiomes. Analyzing 16S rRNA gene copy numbers can help to distinguish between environmental and host-associated microbiomes: host-associated microbiomes have increased 16S rRNA gene copy numbers (Thompson et al., 2017) when compared to environmental microbiomes. The 16S gene copy number of our samples was in the typical range of host-associated communities, and significantly higher than that of non-host associated communities (P < 2.2e-16, Mann–Whitney U test, Figure 5). In an independent survey, where we compared the microbiomes of flies and their immediate substrate, we also found more copies of the 16S rRNA gene in the flies than in the substrate (P < 0.01, Mann–Whitney U test, one-sided, Figure 5). This distinguishes the D. melanogaster microbiome from purely environmental microbiomes and supports host-related structuring.

Host-specificity of microbes that are co-structured with host population genetic variation

16S gene copy numbers suggested that the natural D. melanogaster microbiome is a typical host-associated community. This encouraged us to further explore the possibility that interactions with the host underlie the co-structure of microbiomes with host genetic variation. In order to test this, we analyzed whether the bacteria that were co-structured with host genetic variation differed in abundance between flies and their substrate. Specifically, we hypothesized that potential host-control would lead to a depletion of Enterobacteraceae in the flies because flies might avoid or reduce contact with these bacteria for their frequent pathogenicity. For example, Enterobacteraceae of the genera Providencia, Serratia, Erwinia, Pseudomonas are Drosophila pathogens. Indeed,
Enterobacteraceae were more abundant in the substrate than in the flies ($P = 0.026$, paired Mann-Whitney test, one sided, Figure 6A). Furthermore, we expected to find OTU2 (*C. intestini*) at higher abundance in the fly than in the substrate because this OTU is a common member of the *D. melanogaster* associated community and contributes to healthy gut homeostasis (Ryu et al., 2008; Chandler et al., 2011). Indeed, OTU2 was enriched in flies (Figure 6B, $P = 0.022$, paired Mann-Whitney test, one-sided). Finally, we expected that Acetobacteraceae in general would be enriched in flies over substrate because this family contains several members that benefit *D. melanogaster* (Shin et al., 2011; Pais et al., 2018). This expectation was also confirmed (Figure 6C, $P = 0.034$, paired Mann-Whitney test, one-sided). However, when OTU2 was excluded from the analysis of Acetobacteraceae, Acetobacteraceae were not significantly enriched in flies anymore ($P = 0.21$, paired Mann-Whitney test, one sided), indicating that OTU2 contributed to family level differences. We found no difference between flies and substrate for Leuconostocaceae ($P = 0.27$, paired Mann-Whitney test, two-sided).

Discussion

We set out to test whether there is evidence for host-control over the microbiome by *D. melanogaster* in a natural setting. For this purpose, we combined a comprehensive analysis of the structuring principles of the *D. melanogaster* microbiome on a continent-wide scale (Figure 1 and Table S1) with an independent survey comparing the microbiome of flies to that of their substrate. This resulted in several lines of evidence that support the idea that *D. melanogaster* exerts limited, but detectable and highly specific control over its microbiome.
The correlation of host population genetic differentiation and the differentiation of microbiomes can be interpreted as evidence for host-control. This correlation is consistent with a model, in which stronger genetic differentiation leads, on average, to larger differences in host-control, and hence host-associated microbiomes. Given ample evidence for variation in host-control between natural populations that depends on genotype (Lazzaro et al., 2008; Corby-Harris and Promislow, 2008; Behrman et al., 2018; Walters et al., 2018) this seems a reasonable model.

It appears unlikely that co-structure resulted from environmental factors that affect both, the microbiome and host genetic variation for two reasons: First, we accounted for the most plausible environmental factors that could affect microbiomes and the host at the same time in our model (food-substrate, temperature, precipitation). Second, for assessing host genetic variation, we used SNPs from small introns that are considered least affected by natural selection (Parsch et al., 2010; Lawrie et al., 2013). Therefore, it is unlikely that selection exerted by environmental factors that also affect the microbiome strongly affects these SNPs and generates co-structure.

It is similarly difficult to explain the co-structure by co-dispersal of Drosophila and bacteria because we found no evidence for pronounced dispersal limitation of the bacteria that co-vary with host genetic differentiation on a global scale (Figure 4). Instead, our data and previous studies suggest that these bacteria are cosmopolitan (Cox and Gilmore, 2007; Chandler et al., 2011). Taken together that environmental variation was accounted for and that we found no evidence for dispersal limitation, a role for host-control in the observed co-structure appears plausible.

16S copy numbers and differences between flies and substrate microbiomes support host-control.
Host effects on the microbiome were further supported by 16S rRNA gene copy numbers that were in the typical range for host-associated communities and significantly different from that of non-host associated communities (Figure 5). As expected, the copy number in the substrate samples was smaller than that in fly samples. Interestingly, the copy number in substrate microbiomes was still larger than that of typical non-host associated microbiomes. This is consistent with Drosophila also affecting the microbiome of its immediate environment (Wong et al., 2015; Chaston et al., 2016; Storelli et al., 2018) and transforming it to appear more host-like.

Besides the increased number of 16S gene copies, host-control was evident from differences between the host microbiome and that of its substrate; three of the four bacterial groups (Acetobacteraceae, Enterobacteraceae, OTU2) that correlated with host genetic variation on a continental scale (Figure 3) also differed in abundance between flies and their substrate (Figure 6).

Fitness effects of microbes that show evidence of host structuring support host-control. The evidence above supports host-related structuring of the microbiome that is consistent with host-control. However, the term 'host-control' also implies that the effects of the host on the microbiome provide some fitness benefit to the host. The bacteria that are structured in the host environment and the direction of the structuring (enrichment or depletion) suggest such fitness benefits.

It seems reasonable to assume that the reduction of Enterobacteriaceae in the fly environment (Figure 6), is likely beneficial for the flies, because this family comprises a range of the most important D. melanogaster pathogens. Examples are Providencia (Galac and Lazzaro, 2011), Serratia (Flyg et al., 1980; Lazzaro et al., 2006), Erwinia (Basset et al., 2000), and Pseudomonas (Vodovar et al., 2005). A reduction of Enterobacteriaceae in the fly gut is in line with results from Ryu et al. (2008). These
Authors have shown that Enterobacteriaceae, including the highly pathogenic *Erwinia carotovora* carotovora-15 do not persist in the fly gut.

In contrast to Enterobacteriaceae, Acetobacteraceae were enriched in the host. This pattern was mainly driven by OTU2 (*C. intestini*). This OTU matches sequences from previous studies on fruit flies in the natural environment (Blast results Table S3) (Cox and Gilmore, 2007; Chandler et al., 2011; Wang and Staubach, 2018) and in the laboratory (Ryu et al., 2008). In particular, it perfectly matches *C. intestini strain A911* (Roh et al., 2008). This strain is sensitive to anti-microbial peptides (AMPs) (Ryu et al., 2008), and hence can be subject to host-control. In wild-type flies, it is a dominant member of the microbiome. When AMPs are misregulated it is replaced by *Gluconobacter morbifer* that has detrimental effects on flies. Thus, in the wild-type gut environment, *C. intestini strain A911* is favored by the host and has a protective function. Favoring of a protective microbe can be considered host-control.

The specificity of host-control

*G. morbifer* as well as *C. intestini* are Acetobacteraceae, and hence relatively closely related. That flies can favor one over the other, points towards highly specific host-control in *D. melanogaster*. Our results suggest that host-control can also be highly specific under natural conditions; Only OTU2 (perfect sequence match with *C. intestini*) was strongly co-structured with host genetic variation and at the same time enriched in flies over substrate (Figure 3 and 6). The evidence for high specificity in the interaction with bacteria that we found parallels recent results from Pais et al. (2018). These authors found that *Acetobacter thailandicus* colonizes *D. melanogaster* and persists in the gut, while a closely related *Acetobacter* strain does not persist. High specificity also fits in with results from Adair et al. (2018), who showed that the assembly of the natural bacterial microbiome in the *D. melanogaster* gut can be largely explained by neutral processes, except for a
specific set of bacteria. This specificity of the interaction of *D. melanogaster* with its microbiome is also fully compatible with recent advances in understanding the mechanistic principles of *D. melanogaster* immunity. A combination of highly specific regulation of the IMD pathway via different peptidoglycan recognition proteins (PGRPs) and specific regulation of the duox pathway (Ha et al., 2005; Lhocine et al., 2008; Bosco-Drayon et al., 2012; Lee et al., 2013; Guo et al., 2014; Iatsenko et al., 2016; Neyen et al., 2016) can lead to highly specific selection processes acting on bacterial communities in the fly gut.

While we found support for specific interaction with OTU2 (*C. intestini*), more general mechanisms seemed to be at work for the interaction with Enterobacteriaceae. The reduction of Enterobacteriaceae in the fly, when compared to the substrate was not linked to any specific OTUs from this family. Likewise, the co-structure with host genetic variation, was only apparent for the family as a whole. This family level host-control could arise in response to signals that are common to Enterobacteriaceae or from their potential pathogenicity in the sense of a danger or damage signal (Matzinger, 2002). Alternatively, the *Drosophila* gut might be just less favorable in terms of its physical condition (e.g. pH) or presence of antimicrobial agents (e.g. AMPs).

Environmental factors and the *D. melanogaster* microbiome

In addition to host genetic structure, temperature as well as the substrate, the flies were collected from, correlated with microbiome structure. While the effects of substrate on the fly microbiome are well described (Chandler et al., 2011; Staubach et al., 2013; Wang and Staubach, 2018), a continental scale temperature effects on a host-associated microbiome has, to our knowledge, not been described before. Temperature affects environmental microbiomes on a global scale (Zhou et al., 2016; Thompson et al., 2017) and the effect in *Drosophila* might reflect the exposure of flies to different environmental microbiomes. Temperature dependence of the microbiome could also be the result of a temperature
dependent dietary switch (Brankatschk et al., 2018); small scale structure of the food sources might allow flies to acquire selectively more plant or yeast material, which might lead to changes in microbiome composition.

While there was a significant effect of annual temperature on the microbiome, the correlation with monthly temperature at the collection date only showed a trend ($P = 0.065$). Because our seasonal sampling was relatively limited (nine locations), more data is required to address the question whether seasonal temperature changes affect the microbiome. Seasonal variation in D. melanogaster associated microbiomes has been described by Behrman et al. (2018) and correlates with differences in pathogen susceptibility of the host. This points towards the possibility that seasonal changes in the microbiome could add to seasonal selective regimes and contribute to seasonal genome variation in D. melanogaster (Bergland et al., 2014; Machado et al., 2018).

The effects of temperature on the microbiome seemed more general as no specific OTU nor family was significantly correlated with temperature variation.

Conclusion

D. melanogaster lives in a microbe rich environment; rotting fruit. In this environment, it is essential for flies to foster beneficial microbes and avoid pathogens. Using continental scale data from natural populations, we presented evidence for specific host-control that favors a protective bacterium. This adds to the recent notion of high specificity of host-microbe interaction in D. melanogaster and shows that this specificity unfolds in an ecological and evolutionary context. At the same time, our study supports previous findings of strong environmental effects on the natural D. melanogaster microbiome. Strong environmental effects in combination with host-control of a relatively small subset of the partners is concordant with the 'ecosystem on a leash' model. Our results support the idea that this model might serve as a common framework to understand D. melanogaster and
mammalian microbiomes. A common framework for mammalian and *D. melanogaster*, increases the transferability and generalizability between systems. Hence, we see a bright future for the *D. melanogaster* microbiome as a model for other organisms, including mammals, in host-microbiome research.

**Materials and Methods**

**Fly and substrate samples**

European fly samples were collected as described in Kapun et al. (2018). In short, 50 samples of *D. melanogaster* were collected from 31 locations across the European continent with a joint effort of European research groups (Figure 1 and Table S1). Each sample contained a pool of 33-40 wild-caught males. We used males only because only males can be reliably distinguished from sympatric *D. simulans*. The effects of pooling on *D. melanogaster* microbiome profiling were assessed in detail by Wang and Staubach (2018). In short, pooling provides a more comprehensive picture of the population microbiome than an individual fly. While differences in microbiome structure between individuals tend to even out in a pool, differences between populations be well differentiated. Because we were interested in variation between populations here, a pooling approach is well suited. All 50 samples were included for analyzing *Drosophila*-associated bacterial community composition, diversity and dispersal patterns. Because data on host genetic differentiation for samples FR_Vil_14_06, UA_Yal_14_17, and DK_Kar_14_40 was not available, these samples were excluded from the analysis of continental scale community structure. The visualization of fly samples on the map in Figure 1 was generated with the R package ‘ggmap’ (Kahle and Wickham, 2013).

Twelve samples from the East Coast of the USA were collected in the same fashion as the European samples and represent population pools of males. Seven of these samples were already analyzed in Behrman et al. (2018) (see Table S1 for details). The
samples named NY and WI were described in Machado et al. (2018). However, the 16S data for these samples was generated here. Because we did not have detailed information on the substrate, these samples were collected from, we did not include them in our continental scale modeling. We used these samples for evaluating the global range of bacteria (Figure 4). For the same purpose, we included 13 fly samples from Wang and Staubach (2018) that were primarily collected at the West Coast of the USA (see Table S1 for details).

For the survey of the microbiome of flies and their substrate, pairs of pools of five flies and the corresponding substrate for a total of 24 samples were collected. The immediate substrate, on which the flies that we collected were sitting and feeding was collected with a sterile scalpel and transferred to a sterile microcentrifuge tube. The survey spanned 6 different substrates from 4 locations (Table S1).

DNA extraction, PCR and sequencing

DNA from the DrosEU samples was extracted by standard phenol-chloroform extraction after homogenization with 3 minutes of bead beating on QIAGEN TissueLyser II as described in Kapun et al. (2018). DNA from population pools from the USA were extracted as described in Bergland et al. (2014). DNA extraction for pools of five flies and the corresponding substrate was performed using the Qiagen QIAamp DNA extraction kit (Qiagen, Carlsbad, CA) combined with bead beating in the same way as for fly samples from Wang and Staubach (2018).

Barcoded bacterial broad range primers, 515F (5’GTGCCAGCMGCCGCGGTAA3’) and 806R (5’GGACTACHVGGGTWTCTAAT3’) from Caporaso et al. (2011) were used to amplify the V4 region of the bacterial 16S rRNA gene. DNA was amplified with Phusion® Hot Start DNA Polymerase (Finnzymes, Espoo, Finland) under the following conditions: 30 sec at 98°C; 30 cycles of 9 sec at 98°C, 60 sec at 50°C and 90 sec at 72°C; final
extension for 10 min at 72°C. In order to reduce PCR bias, amplification reactions were performed in duplicate and pooled. PCR products quantified on an agarose gel and pooled in equimolar amounts. Extraction control PCRs were negative and excluded. The resulting pool was gel extracted using the Qiaquick gel extraction kit (Qiagen, Carlsbad, CA) and sequenced on an illumina MiSeq sequencer reading 2 × 250bp.

Data analysis

We analyzed sequencing data using MOTHUR v1.40.0 (Schloss et al., 2009). Main processing steps in MOTHUR included alignment of paired reads, quality filtering, removal of PCR errors, removal of chimeric sequences, subsampling (rarefication), and alpha-diversity calculations (Kozich et al., 2013). Sequences were taxonomically classified using the SILVA reference database ‘Release 132’ (Pruesse et al., 2007) as implemented in MOTHUR. A detailed step by step analysis script with all commands executed can be found in the supplementary File Script1 for full reproducibility. For studying geographical microbiome structure, OTUs were clustered at 100% identity. Only for the comparison of Shannon diversity to previous studies, we also included clustering at 97% sequence identity.

To identify factors that shape microbial communities, we applied Redundancy Analysis (RDA). Following Borcard, Gillet and Legendre (2018), OTU count data was Hellinger transformed to allow analysis in the linear RDA framework. In order to reduce the effects of rare species on RDA and assuming that ecologically relevant species should be frequent, we focused the analysis on OTUs with more than 1000 reads across samples. Our candidate explanatory variables were temperature, precipitation, substrate and host genetic differentiation. Data of annual and monthly mean temperature (BIO 1 and tmean) and precipitation (BIO 12 and prec) were downloaded from WorldClim (Fick and Hijmans, 2017, see supplementary File Script2). Host genetic differentiation was represented by the
first two principle components of an allele frequency based Principle Components Analysis performed by Kapun et al. (2018). In short the data represents allele frequencies from more than 20,000 SNPs in short intronic sequences that evolve putatively neutral and best represent population structure. In order to select the variables that were most important for microbiome structure, we applied forward model selection of additive linear models. This was done with the ordistep function from the vegan R package (Oksanen et al., 2018). The Ordistep function provides a stepwise approach to select variables based on permutation P-values and Akaike's Information Criterion (AIC).

In order to test for potential spatial autocorrelation we followed the protocol by (Borcard et al., 2018) using the dbmem function. This protocol employs eigenvector analysis to detect autocorrelation at different scales. We found no evidence for significant autocorrelation in our data (see supplementary File Script3) after removal of the continent-wide trend in species distributions that we analyzed here. All algorithms were part of the vegan (Oksanen et al., 2018) and adespatial R packages (Dray et al., 2018). Geographic distances were computed with the gdist function from the Imap R package (Wallace, 2012, see supplementary File Script4).

For the correlation of host genetic differentiation with the relative abundance of individual OTUs and bacterial families, we calculated q-values with the p.adjust function in R to account for multiple testing. Following the recommendation by Efron et al. (2007), only significant correlations ($P < 0.05$) with bacterial groups with q-values smaller than 0.2 were considered significant.

Average community 16S rRNA gene copy number (ACN) was predicted from 16S rRNA gene amplicon data using PICRUSt (Langille et al., 2013). The method for calculating ACN was adapted from Thompson et al. (2017). We first classified sequences using the Greengenes reference database and generated an OTU table in the Biom-format using the make.biom function in MOTHUR. The resulting biom-formated table served as
input for the normalize_by_copy_number.py command implemented in PICRUSt. The
output file is a normalized observation table. ACN for each sample were calculated as the
raw sample sum divided by the normalized sample sum.

Data availability

Raw sequence data is in the submission process to the ncbi short read archive (SRA).

Acknowledgements

We thank the DrosEU consortium for the coordinated sampling and fruitful collaboration.

We thank the Paul Schmidt (University of Pennsylvania), John Pool (University of
Wisconsin), Lazzaro (Cornell University) labs, as well as the Drosophila Real Time
Evolution Consortium (DrosRTEC) for samples. We thank Carsten Dormann (University of
Freiburg) for statistical support. The authors acknowledge support by the state of Baden-
Württemberg through bwHPC. This work was funded by a DFG grant to FS (STA1154/4-1,
408908608) and the Landesgraduiertenfoerterung of the state of Baden-Württemberg.
Table 1 RDA model selection for factors that explain community composition at the 100% identity OTU level. The forward selection approach starts with the null model, adding the best explanatory factors one by one until adding the next factor fails to significantly improve the model. Significant p-values and the best model are in bold. P-values are based on permutation tests. PC1 = Axis 1 of host genetic variation, PC2 = Axis 1 of host genetic variation, substrate = substrate the flies were collected from, T(y) = mean annual temperature, T(m) = mean monthly temperature.

| model | factors             | add       | AIC   | P     |
|-------|---------------------|-----------|-------|-------|
| M0    | NULL                |           | -20.757 |       |
| M1    | PC1                 | +PC1      | -21.802 | 0.001 |
| M2    | PC1+substrate       | +bait     | -18.470 | 0.007 |
| M3    | PC1+substrate+PC2   | +PC2      | -19.076 | 0.013 |
| M4    | PC1+substrate+PC2+T(y) | +T(y) | -19.905 | 0.011 |
| M5    | PC1+substrate+PC2+T(y)+T(m) | +T(m) | -20.130 | 0.065 |

Table 2 RDA model selection for factors that explain community composition at the family level. The forward selection approach starts with the null model, adding the best explanatory factors one by one until adding the next factor fails to significantly improve the explanatory power of the model. Significant p-values and the best model are in bold. P-values are based on permutation tests. PC1 = Axis 1 of host genetic variation, PC2 = Axis of host genetic variation, substrate = substrate the flies were collected from, T(y) = mean annual temperature.

| model | factors           | add       | AIC   | P     |
|-------|-------------------|-----------|-------|-------|
| M0    | NULL              |           | -73.782 |       |
| M1    | T(y)              | +T(y)     | -76.095 | 0.002 |
| M2    | T(y)+PC1          | +PC1      | -77.237 | 0.012 |
| M3    | T(y)+PC1+PC2      | +PC2      | -77.488 | 0.067 |
Figure 1 Overview of sampling locations. The map shows the geographic locations of 50 samples for bacterial community analysis in the 2014 DrosEU dataset. The color of the circles indicates the sampling season for each location.
Figure 2 Correlation between pairwise geographic distance and BCD of the bacterial communities. P-value according to Mantel test; r is Pearson's correlation coefficient.
Figure 3 OTU2 (A) and three bacterial families (B-D) correlate with host genetic differentiation (PC1). P-value and correlation coefficient according to Pearson's Product-Moment Correlation.
Figure 4: The bacteria that were structured according to host genetic variation, are common in Europe, the East Coast (EC), and the West Coast (WC) of and North America. Gray scale indicates relative abundance. Samples marked with * were described in Wang & Staubach (2018).
Figure 5 Comparison of 16S rRNA gene average copy number (ACN) between flies and substrates. ACN was higher in the fly than in the substrate communities (**P < 0.01, Mann-Whitney-U test, one-sided). Barplot (blue and green) in the background shows ACN from (Thompson et al. 2017). Fly as well as substrate samples were in the typical range of animal-associated microbiomes.
Figure 6 Comparison of relative abundance of (A) OTU2, (B) Acetobacteraceae and (C) Enterobacteraceae between flies and their substrate. P-values according to paired Mann-Whitney-U test, one-sided.
References

Adair KL, Wilson M, Bost A, Douglas AE. 2018. Microbial community assembly in wild populations of the fruit fly Drosophila melanogaster. *The ISME Journal* **12**:959–972. doi:10.1038/s41396-017-0020-x

Agler MT, Ruhe J, Kroll S, Morhenn C, Kim S-T, Weigel D, Kemen EM. 2016. Microbial Hub Taxa Link Host and Abiotic Factors to Plant Microbiome Variation. *PLoS Biology* **14**:e1002352. doi:10.1371/journal.pbio.1002352

Ankrah NYD, Douglas AE. 2018. Nutrient factories: metabolic function of beneficial microorganisms associated with insects. *Environmental Microbiology* **20**:2002–2011. doi:10.1111/1462-2920.14097

Bang C, Dagan T, Deines P, Dubilier N, Duschl WJ, Fraune S, Hentschel U, Hirt H, Hülter N, Lachnit T, Picazo D, Pita L, Pogoreutz C, Rädecker N, Saad MM, Schmitz RA, Schuelenburg H, Voolstra CR, Weiland-Bräuer N, Ziegler M, Bosch TCG. 2018. Metaorganisms in extreme environments: do microbes play a role in organismal adaptation? *Zoology* **127**:1–19. doi:10.1016/j.zool.2018.02.004

Barata A, Santos SC, Malfeito-Ferreira M, Loureiro V. 2012. New Insights into the Ecological Interaction Between Grape Berry Microorganisms and Drosophila Flies During the Development of Sour Rot. *Microb Ecol* **64**:416–430. doi:10.1007/s00248-012-0041-y

Basset A, Khush RS, Braun A, Gardan L, Boccard F, Hoffmann JA, Lemaitre B. 2000. The phytopathogenic bacteria Erwinia carotovora infects Drosophila and activates an immune response. *PNAS* **97**:3376–3381. doi:10.1073/pnas.97.7.3376

Behrman EL, Howick VM, Kapun M, Staubach F, Bergland AO, Petrov DA, Lazzaro BP, Schmidt PS. 2018. Rapid seasonal evolution in innate immunity of wild Drosophila melanogaster. *Proceedings of the Royal Society B: Biological Sciences* **285**:20172599. doi:10.1098/rspb.2017.2599

Bergland AO, Behrman EL, O’Brien KR, Schmidt PS, Petrov DA. 2014. Genomic Evidence of
Rapid and Stable Adaptive Oscillations over Seasonal Time Scales in *Drosophila*. *PLOS Genetics* **10**:e1004775. doi:10.1371/journal.pgen.1004775

Blum JE, Fischer CN, Miles J, Handelsman J. 2013. Frequent Replenishment Sustains the Beneficial Microbiome of *Drosophila melanogaster*. *mBio* **4**:e00860-13. doi:10.1128/mBio.00860-13

Borcard D, Gillet F, Legendre P. 2018. Numerical Ecology with R, 2nd ed, Use R! Springer International Publishing.

Bosco-Drayon V, Poidevin M, Boneca IG, Narbonne-Reveau K, Royet J, Charroux B. 2012. Peptidoglycan Sensing by the Receptor PGRP-LE in the *Drosophila* Gut Induces Immune Responses to Infectious Bacteria and Tolerance to Microbiota. *Cell Host & Microbe* **12**:153–165. doi:10.1016/j.chom.2012.06.002

Brankatschk M, Gutmann T, Knittelfelder O, Palladini A, Prince E, Grzybek M, Brankatschk B, Shevchenko A, Coskun Ü, Eaton S. 2018. A Temperature-Dependent Switch in Feeding Preference Improves *Drosophila* Development and Survival in the Cold. *Developmental Cell* **46**:781-793.e4. doi:10.1016/j.devcel.2018.05.028

Broderick NA, Buchon N, Lemaitre B. 2014. Microbiota-Induced Changes in *Drosophila melanogaster* Host Gene Expression and Gut Morphology. *mBio* **5**:e01117-14. doi:10.1128/mBio.01117-14

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, Fierer N, Knight R. 2011. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *PNAS* **108**:4516–4522. doi:10.1073/pnas.1000080107

Chandler JA, Lang JM, Bhatnagar S, Eisen JA, Kopp A. 2011. Bacterial Communities of Diverse *Drosophila* Species: Ecological Context of a Host–Microbe Model System. *PLOS Genetics* **7**:e1002272. doi:10.1371/journal.pgen.1002272

Chaston JM, Dobson AJ, Newell PD, Douglas AE. 2016. Host Genetic Control of the Microbiota Mediates the *Drosophila* Nutritional Phenotype. *Appl Environ Microbiol* **82**:671–679.
doi:10.1128/AEM.03301-15
Corby-Harris V, Pontaroli AC, Shimkets LJ, Bennetzen JL, Habel KE, Promislow DEL. 2007. Geographical Distribution and Diversity of Bacteria Associated with Natural Populations of Drosophila melanogaster. Appl Environ Microbiol 73:3470–3479. doi:10.1128/AEM.02120-06
Corby-Harris V, Promislow DEL. 2008. Host ecology shapes geographical variation for resistance to bacterial infection in Drosophila melanogaster. Journal of Animal Ecology 77:768–776. doi:10.1111/j.1365-2656.2008.01399.x
Cox CR, Gilmore MS. 2007. Native Microbial Colonization of Drosophila melanogaster and Its Use as a Model of Enterococcus faecalis Pathogenesis. Infection and Immunity 75:1565–1576. doi:10.1128/IAI.01496-06
Douglas AE. 2018. Which experimental systems should we use for human microbiome science? PLOS Biology 16:e2005245. doi:10.1371/journal.pbio.2005245
Dray S, Bauman D, Blanchet G, Borcard D, Clappe S, Guenard G, Jombart T, Larocque G, Legendre P, Madi N, Wagner HH. 2018. adespatial: Multivariate Multiscale Spatial Analysis.
Ebert D. 2013a. The Epidemiology and Evolution of Symbionts with Mixed-Mode Transmission. Annu Rev Ecol Evol Syst 44:623–643. doi:10.1146/annurev-ecolsys-032513-100555
Ebert D. 2013b. The Epidemiology and Evolution of Symbionts with Mixed-Mode Transmission. Annu Rev Ecol Evol Syst 44:623–643. doi:10.1146/annurev-ecolsys-032513-100555
Efron B. 2007. Size, power and false discovery rates. Ann Statist 35:1351–1377. doi:10.1214/009053606000001460
Erkosar B, Storelli G, Defaye A, Leulier F. 2013. Host-Intestinal Microbiota Mutualism: “Learning on the Fly.” Cell Host & Microbe 13:8–14. doi:10.1016/j.chom.2012.12.004
Erkosar B, Yashiro E, Zajitschek F, Friberg U, Maklakov AA, Meer JR van der, Kawecki TJ. 2018. Host diet mediates a negative relationship between abundance and diversity of Drosophila
gut microbiota. *Ecology and Evolution* 8:9491–9502. doi:10.1002/ece3.4444

Fick SE, Hijmans RJ. 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology* 37:4302–4315. doi:10.1002/joc.5086

Fink C, Staubach F, Kuenzel S, Baines JF, Roeder T. 2013. Noninvasive Analysis of Microbiome Dynamics in the Fruit Fly *Drosophila melanogaster*. *Appl Environ Microbiol* 79:6984–6988. doi:10.1128/AEM.01903-13

Flyg C, Kenne K, Boman HG. 1980. Insect pathogenic properties of *Serratia marcescens*: phage-resistant mutants with a decreased resistance to Cecropia immunity and a decreased virulence to *Drosophila*. *J Gen Microbiol* 120:173–181. doi:10.1099/00221287-120-1-173

Foster KR, Schluter J, Coyte KZ, Rakoff-Nahoum S. 2017. The evolution of the host microbiome as an ecosystem on a leash. *Nature* 548:43–51. doi:10.1038/nature23292

Foster KR, Wenseleers T. 2006. A general model for the evolution of mutualisms. *Journal of Evolutionary Biology* 19:1283–1293. doi:10.1111/j.1420-9101.2005.01073.x

Galac MR, Lazzaro BP. 2011. Comparative pathology of bacteria in the genus Providencia to a natural host, *Drosophila melanogaster*. *Microbes and Infection* 13:673–683. doi:10.1016/j.micinf.2011.02.005

Guo L, Karpac J, Tran SL, Jasper H. 2014. PGRP-SC2 Promotes Gut Immune Homeostasis to Limit Commensal Dysbiosis and Extend Lifespan. *Cell* 156:109–122. doi:10.1016/j.cell.2013.12.018

Ha E-M, Oh C-T, Bae YS, Lee W-J. 2005. A Direct Role for Dual Oxidase in *Drosophila* Gut Immunity. *Science* 310:847–850. doi:10.1126/science.1117311

Hubbell SP. 2001. The Unified Neutral Theory of Biodiversity and Biogeography. Princeton University Press.

Iatsenko I, Kondo S, Mengin-Lecreulx D, Lemaitre B. 2016. PGRP-SD, an Extracellular Pattern-Recognition Receptor, Enhances Peptidoglycan-Mediated Activation of the *Drosophila* Imd Pathway. *Immunity* 45:1013–1023. doi:10.1016/j.immuni.2016.10.029
Jaenike J, Unckless R, Cockburn SN, Boelio LM, Perlman SJ. 2010. Adaptation via Symbiosis: Recent Spread of a *Drosophila* Defensive Symbiont. *Science* **329**:212–215. doi:10.1126/science.1188235

Kahle D, Wickham H. 2013. ggmap: Spatial Visualization with ggplot2. *The R Journal* **5**:144–161.

Kapun M, Aduriz MGB, Staubach F, Vieira J, Obbard D, Goubert C, Stabelli OR, Kankare M, Haudry A, Wiberg RAW, Waidele L, Kozeretska I, Pasyukova E, Loeschcke V, Pascual M, Vieira CP, Serga S, Montchamp-Moreau C, Abbott J, Gibert P, Porcelli D, Posnien N, Grath S, Sucena E, Bergland A, Guerreiro MPG, Onder BS, Argyridou E, Guio L, Schou MF, Deplancke B, Vieira C, Ritchie MG, Zwaan B, Tauber E, Orenge D, Puerma E, Aguade M, Schmidt P, Parsch J, Betancourt A, Flatt T, Gonzalez J. 2018. Genomic analysis of European *Drosophila* populations reveals longitudinal structure and continent-wide selection. *bioRxiv* 313759. doi:10.1101/313759

Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD. 2013. Development of a Dual-Index Sequencing Strategy and Curation Pipeline for Analyzing Amplicon Sequence Data on the MiSeq Illumina Sequencing Platform. *Appl Environ Microbiol* **79**:5112–5120. doi:10.1128/AEM.01043-13

Langille MGI, Zaneveld J, Caporaso JG, McDonald D, Knights D, Reyes JA, Clemente JC, Burkepile DE, Vega Thurber RL, Knight R, Beiko RG, Huttenhower C. 2013. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nature Biotechnology* **31**:814–821. doi:10.1038/nbt.2676

Lawrie DS, Messer PW, Hershberg R, Petrov DA. 2013. Strong Purifying Selection at Synonymous Sites in *D. melanogaster*. *PLOS Genetics* **9**:e1003527. doi:10.1371/journal.pgen.1003527

Lazzaro BP, Flores HA, Lorigan JG, Yourth CP. 2008. Genotype-by-Environment Interactions and Adaptation to Local Temperature Affect Immunity and Fecundity in *Drosophila melanogaster*. *PLOS Pathogens* **4**:e1000025. doi:10.1371/journal.ppat.1000025

Lazzaro BP, Sackton TB, Clark AG. 2006. Genetic Variation in *Drosophila melanogaster*
Resistance to Infection: A Comparison Across Bacteria. *Genetics* **174**:1539–1554. 
doi:10.1534/genetics.105.054593

Lee K-A, Kim S-H, Kim E-K, Ha E-M, You H, Kim B, Kim M-J, Kwon Y, Ryu J-H, Lee W-J. 2013. Bacterial-Derived Uracil as a Modulator of Mucosal Immunity and Gut-Microbe Homeostasis in *Drosophila*. *Cell* **153**:797–811. doi:10.1016/j.cell.2013.04.009

Ley RE, Hamady M, Lozupone C, Turnbaugh PJ, Ramey RR, Bircher JS, Schlegel ML, Tucker TA, Schrenzel MD, Knight R, Gordon JJ. 2008. Evolution of Mammals and Their Gut Microbes. *Science* **320**:1647–1651. doi:10.1126/science.1155725

Lhocine N, Ribeiro PS, Buchon N, Wepf A, Wilson R, Tenev T, Lemaitre B, Gstaiger M, Meier P, Leulier F. 2008. PIMS Modulates Immune Tolerance by Negatively Regulating *Drosophila* Innate Immune Signaling. *Cell Host & Microbe* **4**:147–158. doi:10.1016/j.chom.2008.07.004

Linnenbrink M, Wang J, Hardouin EA, Künzel S, Metzler D, Baines JF. 2013. The role of biogeography in shaping diversity of the intestinal microbiota in house mice. *Molecular Ecology* **22**:1904–1916. doi:10.1111/mec.12206

Machado H, Bergland AO, Taylor R, Tilk S, Behrman E, Dyer K, Fabian D, Flatt T, Gonzalez J, Karasov T, Kozeretska I, Lazzaro B, Merritt T, Pool J, O’Brien K, Rajpurohit S, Roy P, Schaeffer S, Serga S, Schmidt P, Petrov D. 2018. Broad geographic sampling reveals predictable and pervasive seasonal adaptation in *Drosophila*. *bioRxiv* 337543. doi:10.1101/337543

Martino ME, Joncour P, Leenay R, Gervais H, Shah M, Hughes S, Gillet B, Beisel C, Leulier F. 2018. Bacterial Adaptation to the Host’s Diet Is a Key Evolutionary Force Shaping *Drosophila*-Lactobacillus Symbiosis. *Cell Host & Microbe* **24**:109-119.e6. doi:10.1016/j.chom.2018.06.001

Matzinger P. 2002. The Danger Model: A Renewed Sense of Self. *Science* **296**:301–305. doi:10.1126/science.1071059
Neyen C, Runchel C, Schüpfer F, Meier P, Lemaitre B. 2016. The regulatory isoform rPGRP-LC induces immune resolution via endosomal degradation of receptors. *Nature Immunology* 17:1150–1158. doi:10.1038/ni.3536

Obadia B, Güvener ZT, Zhang V, Ceja-Navarro JA, Brodie EL, Ja WW, Ludington WB. 2017. Probabilistic Invasion Underlies Natural Gut Microbiome Stability. *Current Biology* 27:1999-2006.e8. doi:10.1016/j.cub.2017.05.034

Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O’Hara RB, Simpson GL, Solymos P, Stevens MHH, Szoecs E, Wagner H. 2018. vegan: Community Ecology Package.

Pais IS, Valente RS, Sporniak M, Teixeira L. 2018. *Drosophila melanogaster* establishes a species-specific mutualistic interaction with stable gut-colonizing bacteria. *PLOS Biology* 16:e2005710. doi:10.1371/journal.pbio.2005710

Parsch J, Novozhilov S, Saminadin-Peter SS, Wong KM, Andolfatto P. 2010. On the Utility of Short Intron Sequences as a Reference for the Detection of Positive and Negative Selection in *Drosophila*. *Mol Biol Evol* 27:1226–1234. doi:10.1093/molbev/msq046

Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J, Glöckner FO. 2007. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res* 35:7188–7196. doi:10.1093/nar/gkm864

Queller DC, Strassmann JE. 2018. Evolutionary Conflict. *Annual Review of Ecology, Evolution, and Systematics*. doi:10.1146/annurev-ecolsys-110617-062527

Roh SW, Nam Y-D, Chang H-W, Kim K-H, Kim M-S, Ryu J-H, Kim S-H, Lee W-J, Bae J-W. 2008. Phylogenetic Characterization of Two Novel Commensal Bacteria Involved with Innate Immune Homeostasis in *Drosophila melanogaster*. *Appl Environ Microbiol* 74:6171–6177. doi:10.1128/AEM.00301-08

Rothschild D, Weissbrod O, Barkan E, Kurilshikov A, Korem T, Zeevi D, Costea PI, Godneva A, Kalka IN, Bar N, Shilo S, Lador D, Vila AV, Zmora N, Pevsner-Fischer M, Israeli D,
Kosower N, Malka G, Wolf BC, Avnit-Sagi T, Lotan-Pompan M, Weinberger A, Halpern Z, Carmi S, Fu J, Wijmenga C, Zhernakova A, Elinav E, Segal E. 2018. Environment dominates over host genetics in shaping human gut microbiota. Nature 555:210–215. doi:10.1038/nature25973

Ryu J-H, Kim S-H, Lee H-Y, Bai JY, Nam Y-D, Bae J-W, Lee DG, Shin SC, Ha E-M, Lee W-J. 2008. Innate Immune Homeostasis by the Homeobox Gene Caudal and Commensal-Gut Mutualism in Drosophila. Science 319:777–782. doi:10.1126/science.1149357

Sannino DR, Dobson AJ, Edwards K, Angert ER, Buchon N. 2018. The Drosophila melanogaster Gut Microbiota Provisions Thiamine to Its Host. mBio 9:e00155-18. doi:10.1128/mBio.00155-18

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Horn DJV, Weber CF. 2009. Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. Appl Environ Microbiol 75:7537–7541. doi:10.1128/AEM.01541-09

Schluter J, Foster KR. 2012. The Evolution of Mutualism in Gut Microbiota Via Host Epithelial Selection. PLOS Biology 10:e1001424. doi:10.1371/journal.pbio.1001424

Shin SC, Kim S-H, You H, Kim B, Kim AC, Lee K-A, Yoon J-H, Ryu J-H, Lee W-J. 2011. Drosophila Microbiome Modulates Host Developmental and Metabolic Homeostasis via Insulin Signaling. Science 334:670–674. doi:10.1126/science.1212782

Sieber M, Pita L, Weiland-Bräuer N, Dirksen P, Wang J, Mortzfeld B, Franzenburg S, Schmitz RA, Baines JF, Fraune S, Hentschel U, Schulenburg H, Bosch TCG, Traulsen A. 2018. The Neutral Metaorganism. bioRxiv 367243. doi:10.1101/367243

Sommer AJ, Newell PD. 2018. Metabolic Basis for Mutualism Between Gut Bacteria and Its Impact on Their Host Drosophila melanogaster. Appl Environ Microbiol AEM.01882-18. doi:10.1128/AEM.01882-18
Staubach F, Baines JF, Künzel S, Bik EM, Petrov DA. 2013. Host Species and Environmental Effects on Bacterial Communities Associated with *Drosophila* in the Laboratory and in the Natural Environment. *PLOS ONE* 8:e70749. doi:10.1371/journal.pone.0070749

Storelli G, Defaye A, Erkosar B, Hols P, Royet J, Leulier F. 2011. Lactobacillus plantarum Promotes *Drosophila* Systemic Growth by Modulating Hormonal Signals through TOR-Dependent Nutrient Sensing. *Cell Metabolism* 14:403–414. doi:10.1016/j.cmet.2011.07.012

Storelli G, Strigini M, Grenier T, Bozonnet L, Schwarzer M, Daniel C, Matos R, Leulier F. 2018. *Drosophila* Perpetuates Nutritional Mutualism by Promoting the Fitness of Its Intestinal Symbiont Lactobacillus plantarum. *Cell Metabolism* 27:362-377.e8. doi:10.1016/j.cmet.2017.11.011

Téfit MA, Leulier F. 2017. Lactobacillus plantarum favors the early emergence of fit and fertile adult *Drosophila* upon chronic undernutrition. *Journal of Experimental Biology* 220:900–907. doi:10.1242/jeb.151522

The Human Microbiome Project Consortium, Huttenhower C, Gevers D, Knight R, Abubucker S, Badger JH, Chinwalla AT, Creasy HH, Earl AM, FitzGerald MG, Fulton RS, Giglio MG, Hallsworth-Pepin K, Lobos EA, Madupu R, Magrini V, Martin JC, Mitreva M, Muzny DM, Sodergren EJ, Versalovic J, Wollam AM, Worley KC, Wortman JR, Young SK, Zeng Q, Aagaard KM, Abolude OO, Allen-Vercoe E, Alm EJ, Alvarado L, Andersen GL, Anderson S, Appelbaum E, Arachchi HM, Armitage G, Arze CA, Ayvaz T, Baker CC, Begg L, Belachew T, Bhonagiri V, Bihan M, Blaser MJ, Bloom T, Bonazzi V, Paul Brooks J, Buck GA, Buhay CJ, Busam DA, Campbell JL, Canon SR, Cantarel BL, Chain PSG, Chen I-MA, Chen L, Chhibba S, Chu K, Ciulla DM, Clemente JC, Clifton SW, Conlan S, Crabtree J, Cutting MA, Davidovics NJ, Davis CC, DeSantis TZ, Deal C, Delehaunty KD, Dewhirst FE, Deych E, Ding Y, Dooling DJ, Dugan SP, Michael Dunne W, Scott Durkin A, Edgar RC, Erlich RL, Farmer CN, Farrell RM, Faust K, Feldgarden M, Felix VM, Fisher S, Fodor AA, Forney LJ, Foster L, Di Francesco V, Friedman J, Friedrich DC, Fronick CC, Fulton LL,
Gao H, Garcia N, Giannoukos G, Giblin C, Giovanni MY, Goldberg JM, Goll J, Gonzalez A, Griggs A, Gujja S, Kinder Haake S, Haas BJ, Hamilton HA, Harris EL, Hepburn TA, Herter B, Hoffmann DE, Holder ME, Howarth C, Huang KH, Huse SM, Izard J, Jansson JK, Jiang H, Jordan C, Joshi V, Katancik JA, Keitel WA, Kelley ST, Kells C, King NB, Knights D, Kong HH, Koren O, Koren S, Kota KC, Kovar CL, Kyrpides NC, La Rosa PS, Lee SL, Lemon KP, Lennon N, Lewis CM, Lewis L, Ley RE, Li K, Liolios K, Liu B, Liu Y, Lo C-C, Lozupone CA, Dwayne Lunsford R, Madden T, Mahurkar AA, Mannion PJ, Mardis ER, Markowitz VM, Mavromatis K, McCorrison JM, McDonald D, McEwen J, McGuire AL, McInnes P, Mehta T, Mihindukulasuriya KA, Miller JR, Minx PJ, Newsham I, Nusbaum C, O’Laughlin M, Orvis J, Pagani I, Palaniappan K, Patel SM, Pearson M, Peterson J, Podar M, Pohl C, Pollard KS, Pop M, Priest ME, Proctor LM, Qin X, Raes J, Ravel J, Reid JG, Rho M, Rhodes R, Riehle KP, Rivera MC, Rodriguez-Mueller B, Rogers Y-H, Ross MC, Russ C, Sanka RK, Sankar P, Fah Sathirapongsasuti J, Schloss JA, Schloss PD, Schmidt TM, Scholz M, Schriml L, Schubert AM, Segata N, Segre JA, Shannon WD, Sharp RR, Sharpton TJ, Shenoy N, Sheth NU, Simone GA, Singh I, Smillie CS, Sobel JD, Sommer DD, Spicer P, Sutton GG, Sykes SM, Tabbaa DG, Thiagarajan M, Tomlinson CM, Torralba M, Treangen TJ, Truty RM, Vishnivetskaya TA, Walker J, Wang L, Wang Z, Ward DV, Warren W, Watson MA, Wellington C, Wetterstrand KA, White JR, Wilczek-Boney K, Wu Y, Wylie KM, Wylie T, Yandava C, Ye L, Ye Y, Yooshef S, Youmans BP, Zhang L, Zhou Y, Zhu Y, Zoloth L, Zucker JD, Birren BW, Gibbs RA, Highlander SK, Methé BA, Nelson KE, Petrosino JF, Weinstock GM, Wilson RK, White O. 2012. Structure, function and diversity of the healthy human microbiome. *Nature* **486**:207–214. doi:10.1038/nature11234

Thompson LR, Sanders JG, McDonald D, Amir A, Laday J, Locey KJ, Prill RJ, Tripathi A, Gibbons SM, Ackermann G, Navas-Molina JA, Janssen S, Kopylova E, Vázquez-Baeza Y, González A, Morton JT, Mirarab S, Zech Xu Z, Jiang L, Haroon MF, Kanbar J, Zhu Q, Jin Song S, Kosciolk T, Bokulich NA, Lefler J, Brislawn CJ, Humphrey G, Owens SM, Hampton-
Castellano YO, Osuolale O, Oswald W, Parnell J, Peralta-Sánchez JM, Petraitis P, Pfister C, Pilon-Smits E, Piombino P, Pointing SB, Pollock FJ, Potter C, Prithiviraj B, Quince C, Rani A, Ranjan R, Rao S, Rees AP, Richardson M, Riebesell U, Robinson C, Rockne KJ, Rodriguezl SM, Rohwer F, Roundstone W, Safran RJ, Sangwan N, Sanz V, Schренк M, Schrenzel MD, Scott NM, Seger RL, Seguin-Orlando A, Seldin L, Seyler LM, Shaksheer B, Sheets GM, Shen C, Shi Y, Shin H, Shogan BD, Shutler D, Siegel J, Simmon S, Sjöling S, Smith DP, Soler JJ, Sperling M, Steinberg PD, Stephens B, Stevens MA, Taghavi S, Tai V, Tait K, Tan CL, Tas, N, Taylor DL, Thomas T, Timling I, Turner BL, Urich T, Ursell LK, van der Lelie D, Van Treuren W, van Zwieten L, Vargas-Robles D, Thurber RV, Vitaglione P, Walker DA, Walters WA, Wang S, Wang T, Weaver T, Webster NS, Wehrle B, Weisenhorn P, Weiss S, Werner JJ, West K, Whitehead A, Whitehead SR, Whittingham LA, Willerslev E, Williams AE, Wood SA, Woodhams DC, Yang Y, Zaneveld J, Zarraonaindia I, Zhang Q, Zhao H. 2017. A communal catalogue reveals Earth’s multiscale microbial diversity. *Nature* **551**:457–463. doi:10.1038/nature24621

Unckless RL, Rottschaefer SM, Lazzaro BP. 2015. The Complex Contributions of Genetics and Nutrition to Immunity in *Drosophila melanogaster*. *PLOS Genetics* **11**:e1005030. doi:10.1371/journal.pgen.1005030

Vodovar N, Vinals M, Liehl P, Basset A, Degrouard J, Spellman P, Boccard F, Lemaitre B. 2005. *Drosophila* host defense after oral infection by an entomopathogenic *Pseudomonas* species. *PNAS* **102**:11414–11419. doi:10.1073/pnas.0502240102

Waidele L, Korb J, Voolstra CR, Künzel S, Dedeine F, Staubach F. 2017. Differential Ecological Specificity of Protist and Bacterial Microbiomes across a Set of Termite Species. *Front Microbiol* **8**. doi:10.3389/fmicb.2017.02518

Wallace JR. 2012. Imap: Interactive Mapping.

Walters AW, Matthews MK, Hughes RC, Malcolm J, Rudman S, Newell PD, Douglas AE, Schmidt PS, Chaston J. 2018. The microbiota influences the *Drosophila melanogaster* life history
strategy. *bioRxiv* 471540. doi:10.1101/471540

Wang J, Linnenbrink M, Künzel S, Fernandes R, Nadeau M-J, Rosenstiel P, Baines JF. 2014. Dietary history contributes to enterotype-like clustering and functional metagenomic content in the intestinal microbiome of wild mice. *PNAS* **111**:E2703–E2710. doi:10.1073/pnas.1402342111

Wang Y, Staubach F. 2018. Individual variation of natural *D. melanogaster*-associated bacterial communities. *FEMS Microbiol Lett* **365**. doi:10.1093/femsle/fny017

Wong AC-N, Chaston JM, Douglas AE. 2013. The inconstant gut microbiota of *Drosophila* species revealed by 16S rRNA gene analysis. *The ISME Journal* **7**:1922–1932. doi:10.1038/ismej.2013.86

Wong AC-N, Luo Y, Jing X, Franzenburg S, Bost A, Douglas AE. 2015. The Host as the Driver of the Microbiota in the Gut and External Environment of *Drosophila melanogaster*. *Appl Environ Microbiol* **81**:6232–6240. doi:10.1128/AEM.01442-15

Wu GD, Chen J, Hoffmann C, Bittinger K, Chen Y-Y, Keilbaugh SA, Bewtra M, Knights D, Walters WA, Knight R, Sinha R, Gilroy E, Gupta K, Baldassano R, Nessel L, Li H, Bushman FD, Lewis JD. 2011. Linking Long-Term Dietary Patterns with Gut Microbial Enterotypes. *Science* **334**:105–108. doi:10.1126/science.1208344

Zhou J, Deng Y, Shen L, Wen C, Yan Q, Ning D, Qin Y, Xue K, Wu L, He Z, Voordekers JW, Nostrand JDV, Buzzard V, Michaletz ST, Enquist BJ, Weiser MD, Kaspari M, Waide R, Yang Y, Brown JH. 2016. Temperature mediates continental-scale diversity of microbes in forest soils. *Nature Communications* **7**:12083. doi:10.1038/ncomms12083