Title

Yeast Communities of Diverse Drosophila Species: Comparison of Two Symbiont Groups in the Same Hosts

Journal

Applied and Environmental Microbiology, 78(20)

ISSN

0099-2240

Authors

Chandler, James Angus
Eisen, Jonathan A
Kopp, Artyom

Publication Date

2012-10-15

DOI

10.1128/aem.01741-12

Peer reviewed

AEM
Journals.ASM.org

Yeast Communities of Diverse *Drosophila* Species: Comparison of Two Symbiont Groups in the Same Hosts

James Angus Chandler,[a] Jonathan A. Eisen,[a,b,c,d] and Artyom Kopp[a]

Department of Evolution and Ecology and Center for Population Biology, University of California Davis, Davis, California, USA[a]; University of California Davis Genome Center, University of California Davis, Davis, California, USA[b]; Department of Medical Microbiology and Immunology, School of Medicine, University of California Davis, Davis, California, USA[c]; and United States Department of Energy Joint Genome Institute, Walnut Creek, California, USA[d]

The combination of ecological diversity with genetic and experimental tractability makes *Drosophila* a powerful model for the study of animal-associated microbial communities. Despite the known importance of yeasts in *Drosophila* physiology, behavior, and fitness, most recent work has focused on *Drosophila*-bacterial interactions. In order to get a more complete understanding of the *Drosophila* microbiome, we characterized the yeast communities associated with different *Drosophila* species collected around the world. We focused on the phylum Ascomycota because it constitutes the vast majority of the *Drosophila*-associated yeasts. Our sampling strategy allowed us to compare the distribution and structure of the yeast and bacterial communities in the same host populations. We show that yeast communities are dominated by a small number of abundant taxa, that the same yeast lineages are associated with different host species and populations, and that host diet has a greater effect than host species on yeast community composition. These patterns closely parallel those observed in *Drosophila* bacterial communities. However, we do not detect a significant correlation between the yeast and bacterial communities of the same host populations. Comparative analysis of different symbiont groups provides a more comprehensive picture of host-microbe interactions. Future work on the role of symbiont communities in animal physiology, ecological adaptation, and evolution would benefit from a similarly holistic approach.

Animal microbiomes can include a wide variety of taxa, such as bacteria, archaea, fungi, and numerous clades of protists (18, 21, 31, 44, 54). Animal hosts have evolved lenient immune systems that tolerate, and possibly even encourage, the persistence of these microbiomes (63). The dynamics and composition of bacterial communities have been established in dozens of animal species from mammals (44) to insects (18). Other groups, such as archaea (54), fungi (21, 71), and protists (31, 66), have also been studied, although less extensively. In mammals, bacterial communities have been implicated in numerous aspects of host physiology (45, 79, 82) and fungi have recently been shown to influence host health and disease (32). Unfortunately, no study to date has explicitly compared the distribution of different symbiont groups in the same host samples. Thus, it remains unclear whether different symbiont communities show the same distribution patterns and community structure.

Drosophila (Diptera: Drosophilidae) flies are host to at least two different symbiont groups: bacteria and single-celled fungi (here referred to as yeasts). The bacterial communities of *Drosophila* consist primarily of three taxonomic groups: *Enterobacteriaceae*, *Acetobacteraceae*, and *Lactobacillales* (11, 14, 15, 83). Bacteria have been implicated in various aspects of fly physiology and fitness, including longevity (6), development time (70, 76), and mating success (69). Some of the immunological factors interacting with both pathogenic and commensal bacteria have also been identified (42, 46, 65).

Yeasts are also common *Drosophila* symbionts, in the broad sense of being consistently associated with live flies. The yeast-*Drosophila* interaction is generally thought of as a diffuse mutualism (74). Although originally proposed as simply a food source for *Drosophila* (5), yeasts can survive passage through the fly digestive tract so that the animal host acts as a vector for yeast dispersal and colonization of new habitats (20, 73, 75). In *Saccharomyces cerevisiae*, the dissolution of tetrad ascospores inside the *Drosophila* gut increases outbreeding (61), suggesting that association with flies can be an important factor in yeast evolution. Insect-yeast interactions have been examined for over a century, with early studies showing that yeasts living within the digestive tracts of flies multiply when the flies are fed a sterilized sugar solution (22).

In nature, *Drosophila* is associated primarily with yeasts in the phylum Ascomycota and the family Saccharomycetaceae (4, 28, 50, 51, 56). Different species of yeasts have different effects on *Drosophila melanogaster* survival and development time (2). *Drosophila* larvae show behavioral preferences for different yeast species, and in at least some cases, the preferred yeast species confers maximum fitness benefits on the host (13, 47). As adults, different fly species are differentially attracted to different yeast baits (19). Finally, much of the *Drosophila* immune system is devoted to the recognition of fungal infection and subsequent response (23, 43), suggesting that flies could potentially have substantial control over the yeast communities in their guts.

The structure of bacterial communities associated with different *Drosophila* species depends upon several factors. Host diet plays an especially important role, as evidenced by both field surveys and controlled laboratory experiments (11). Loss-of-function mutations in crucial host genes can shift bacterial micro-

Received 29 May 2012 **Accepted** 31 July 2012
Published ahead of print 10 August 2012
Address correspondence to James Angus Chandler, jschandler@ucdavis.edu.
Supplemental material for this article may be found at http://aem.asm.org.

doi:10.1128/AEM.01741-12
The authors have paid a fee to allow immediate free access to this article.

TABLE 1 Yeast community samples used in this study

Library name[a]	Bacterial data[b]	Host species	Host diet	Collection location
ELD	Yes	*D. elegans*	*Brugmansia* flowers	Hsinchu, Taiwan
FNS	Yes	*D. falleni*	*Russula* mushrooms	Stony Brook, NY
HCF	Yes	*D. hydei*	Citrus fruit	Wolfskill Orchard, Winters, CA
HPM	Yes	*D. hydei*	Pomegranates	Wolfskill Orchard, Winters, CA
HPP	Yes	*D. hydei*	*Opuntia* fruit	Arboretum, Davis, CA
ICF	Yes	*D. immigrans*	Citrus fruit	Wolfskill Orchard, Winters, CA
IMH	Yes	*D.* sp. aff. *immigrans*	Hibiscus flowers	Captain Cook, HI
MAG	Yes	*D. sulfurigaster*	Mango fruit	Waimanu, HI
MEC	No	*D. malerkotliana*	*Terminalia* fruit	Seychelles Islands
MOV	Yes	*D. mojavensis* and *D. arizonae*	*Agria* cactus	Sonora, Mexico
NNS	No	*D. neotestacea*	*Russula* mushrooms	Stony Brook, NY
PON	Yes	Unidentified *Drosophila* sp.	Pandanus fruit	Waimanu, HI
SPP	No	*D. melanogaster* and *D. simulans*	*Opuntia* fruit	Arboretum, Davis, CA
SCA	Yes	*Scaptodrosophila hibiscii*	Hibiscus flowers	Queensland, Australia
TBB	Yes	*D. melanogaster*	Citrus fruit	Turelli's Orchard, Winters, CA

[a] Samples HCF and ICF were collected simultaneously from the same pile of citrus fruit, and samples HPP and SPP were collected simultaneously from the same *Opuntia* fruit. Samples FNS and NNS, while collected from the same general area, were collected on different days in slightly different locations.
[b] Bacterial data are from reference 11. Further details are provided in Data Set S1 in the supplemental material.

biome composition (65), indicating that the host exercises some control. However, different fly species acquire essentially the same microbiome when raised together on the same food source, suggesting that there may be few if any interspecific differences in how *Drosophila* interacts with its gut bacteria (11). Geography also has little effect, as the same bacterial taxa are found with different *Drosophila* populations on several different continents (11).

With yeasts, the situation is less clear. Different species of *Drosophila*, despite living sympatrically and seemingly utilizing the same food sources, are often associated with different yeast species (16, 29, 56). Other studies, however, have found the effect of host species to be overshadowed by geography (39). When the physiological traits of the yeasts are taken into account, host ecology plays the most important role in shaping the *Drosophila*-associated yeast community (39). It should be noted that all these studies employed culture-based methods for isolating and identifying the yeasts. To our knowledge, no systematic culture-independent analysis of yeast symbiont communities has been conducted in *Drosophila* to date.

In this paper, we pursue two related goals. First, we present the first culture-independent characterization of the yeast communities associated with ecologically and phylogenetically diverse *Drosophila* species. Given the importance of yeasts in *Drosophila* physiology and fitness, description of these communities is necessary for a full understanding of *Drosophila* ecology and evolution. Second, we compare the distribution of yeasts and bacteria in the same host populations. Since neither symbiont group shows any evidence of vertical transmission, both are likely environmentally acquired, and therefore, some similarities in their community structure are inevitable. However, host-symbiont interactions could also create significant differences in community structure, as bacteria and yeasts interact with different components of the *Drosophila* innate immune system.

MATERIALS AND METHODS

Fly collection, dissection, and DNA extraction. *Drosophila* samples were collected with the help of many colleagues around the world (see Acknowledgments) (Table 1; see Data Set S1 in the supplemental material). All samples were obtained from naturally occurring substrates, and no artificial baits were used to attract flies. For collections done in northern California, adults were immediately transferred to sterile no-nutrient medium (2% agar in water) and transported to the University of California Davis (UC-Davis) for dissection, which occurred within 2 h of collection. For more remote field collections, flies were stored in 100% ethanol for transport.

For freshly collected flies, the entire gut was dissected in sterile insect saline and placed in sterile TES buffer (10 mM Tris-HCl [pH 7.5], 1 mM EDTA, 100 mM NaCl). For flies stored in ethanol, dissection was not feasible because weakening of the fly tissues caused the gut to fragment. For these samples, the entire fly bodies were washed twice in 1 ml 2.5% sodium hypochlorite and twice in 1 ml sterile water. Each wash consisted of 30 s of vortexing at maximum speed with 0.5 ml of 0.1-mm glass beads. To ensure adequate removal of external fungi, each final wash was confirmed to be free of fungal cells by PCR with primers NL1 and NL4 (38) (see below). In no case did the final wash show evidence of fungal contamination. Seven to 20 fly bodies or guts were combined for each sample. A previous study did not find any difference in bacterial community composition in dissected intestines and externally sterilized whole bodies (11), and we therefore treated both types of samples equivalently here. Further details regarding sample collection dates, locations, and contents can be found in Data Set S1 in the supplemental material.

DNA was extracted from samples using a modification of the Bead Beater protocol (48). The tissue was homogenized by grinding and three freeze/thaw cycles on dry ice. Samples were then incubated with 50 units/ml of lysozyme for 15 min. Next, physical disruption was performed in a Bead Beater (BioSpec Products, Inc., Bartlesville, OK) on the homogenize setting for 3 min. An overnight incubation with 1% SDS and 2 mg/ml proteinase K was followed by extraction with an equal volume of 25:24:1 phenol-chloroform-isoamyl alcohol. The aqueous phase was incubated at room temperature for at least 15 min with 0.7 volume of 100% isopropanol and 0.1 volume of 3 M sodium acetate before centrifugation at 16,000 × *g* for 30 min at 4°C. The DNA pellet was washed with cold 70% ethanol and allowed to air dry before resuspension in TE (10 mM Tris-HCl, pH 7.5, 1 mM EDTA). A caveat to this method is that it captures all yeast DNA present within the host, whether it belongs to live yeasts or to dietary yeasts that have already been partially digested. While the ratio of live to digested yeasts cannot be ascertained using our methods, studies using culture-based methods found that at least some proportion of yeasts survive passage through the digestive tract of adult flies (28).

Library creation and sequencing. The D1/D2 loop of the rDNA large subunit (LSU) was amplified using the standard fungal primers NL1 and NL4 (38). This region and these primers were chosen for several reasons. First, other commonly used diagnostic genes for fungi (such as the internal transcribed spacer [ITS] region) are much more variable in length, and the sequencing method used would bias results toward shorter PCR fragments. Second, all yeast type strains are required to have their LSU sequenced and publicly available (36). Third, these primers are known to successfully amplify the LSU region from a wide variety of yeasts cultured from *Drosophila* (28).

The 454 adaptor A, a 4-bp key, and a 6-bp barcode were added to the NL1 primer (see Data Set S2 in the supplemental material). The barcodes were a shortened version of 8-bp error-correcting barcodes developed for multiplex pyrosequencing (27). Each barcode differs from all others by at least two nucleotide substitutions (see Data Set S2).

D1/D2 loop fragments were amplified from each sample using the following PCR protocol: 94°C for 5 min; 30 or 35 cycles of 94°C for 1 min, 50°C for 1 min, and 72°C for 1.5 min; and 72°C for 7 min. Three independent 50-μl PCRs were done for each library. Each PCR product was run individually on a 1% agarose gel, and bands of the correct size were excised. To ensure enough template DNA for emPCR amplification (emulsion-based clonal amplification), all three gel slices were pooled, and DNA was isolated using the MN Nucleospin Extract II kit. Because of the longer PCR product, we used a modified emPCR amplification protocol developed specifically for longer fragments (61a). The samples were sequenced on a Roche GS Junior Titanium machine in the laboratory of Jonathan Eisen with the assistance of the UC-Davis Microarray Core Facility.

Sequence alignment. A total of 12,819 reads were generated. The sequencing run did not produce as many reads as expected, but budgetary constraints precluded any additional sequencing. The lengths of these reads followed a bimodal distribution, with most reads (7,040) falling below 90 bp (see Data Set S3 in the supplemental material). All 12,819 reads were checked using the QIIME platform (9). Default parameters were used in all cases, except for the total number of mismatches allowed for the primer sequence, which was changed from 0 to 2. The 3′ end of the NL1 primer ends with GGAAAAG, and it was found that 646 reads would be excluded because of the addition or loss of an A in this homopolymer run and 52 reads would be excluded because of the addition or loss of a G in the paired Gs. Relaxing the parameters to allow 2 mismatches kept 780 additional reads in the data set. A total of 7,942 reads were removed for the following reasons: a minimum length of less than 200 bp (7,489 reads), at least one ambiguous base (87 reads), a mean quality score below 25 (31 reads), a maximum homopolymer run size exceeding 6 (75 reads), the number of mismatches in the primer sequence exceeding the limit of 2 (256 reads), or a barcode that could not be assigned to a library (4 reads). The remaining 4,877 reads had an average length of 529 bp (minimum, 200 bp; maximum, 611 bp).

These 4,877 reads were classified using the Ribosomal Database Project's Fungal Large Subunit Classifier (12). Only the 2,466 reads that were assigned with 100% confidence to the phylum Ascomycota were included in the final analysis. We chose to focus on the Ascomycota because previous studies had identified members of the phylum as important to *Drosophila* survival and fitness (2, 4). The non-Ascomycota reads included 424 from the phylum Basidiomycota (see Data Set S4 in the supplemental material). A majority of the Basidiomycota reads were from a genus of fruiting mushrooms (*Lentinus*) and are found with the mushroom-feeding *Drosophila* (Data Set S4). These reads likely represent the diet that is passing through the fly intestinal tract. A total of 1,210 reads were identified as possible nonfungal unicellular eukaryotes and will be described, along with any possible interactions with bacterial and yeast communities, in a separate paper. The remaining reads include mainly *Drosophila* and host plant sequences and will not be discussed further. However, it should be noted that the primers used in this study (NL1 and NL4) were not specific to fungi and amplified the LSU region from Euglenozoa, Arthropoda, and plants.

Of the 2,466 Ascomycota sequences, all but 1 were between 440 and 581 bp in length (average length, 514 bp; standard deviation [SD], 10.1). The short sequence (259 bp) was removed to facilitate good alignment. Nineteen sequences were assumed to be chimeric and were removed because an NCBI BLAST search found that the 5′ and 3′ ends were most closely related to a different taxonomic class. After this filtering, 6 libraries contained fewer than 50 sequences each and were removed from subsequent analysis (total, 126 sequences).

The remaining 2,316 sequences were aligned using MAFFT and the genafpair option, producing an alignment of 867 columns (34, 35). Since many positions toward the end of the alignment (roughly corresponding to the D2 region) contained mostly gaps, the final 379 columns were removed. Despite removing these 379 columns from the alignment, an average of only 128 nucleotides was actually removed from each read (minimum, 91; maximum, 166; SD, 7.7).

Seventy-five sequences were identified as chimeric using the UCHIME chimera checker in mothur (67). The remaining 2,241 sequences were realigned using MAFFT, producing an alignment of 484 positions. Nineteen additional sequences were identified as chimeric in this alignment. After their removal, MAFFT was used to produce the final alignment of 2,222 sequences and 475 columns. The final data set consisted of 15 libraries with an average of 148 sequences each (minimum, 69; maximum, 264; SD, 67). All sequences and alignments are available through BioTorrents (http://www.biotorrents.net/details.php?id=263) (41).

OTU generation, α-diversity measurements, and β-diversity PCA. The software package mothur was used to generate operational taxonomic units (OTUs) from the chimera-checked alignment (67). OTUs were formed at the 3% divergence level (97% similarity) using the average neighbor-clustering algorithm and the "countends = F" option during the calculation of the distance matrix. This produced 51 OTUs with an average size of 43.6 sequences (SD, 135.4). The largest OTU contained 828 sequences, 21 OTUs contained only 1 sequence, and 41 OTUs contained 10 or fewer sequences (Table 2; see Data Set S5 in the supplemental material). A representative sequence for each OTU was chosen using the "get.oturep" function, which selects the sequence that has the minimum total distance to all the other sequences within that OTU. Alpha diversity measurements and rarefaction curves were then calculated using mothur (Fig. 1 and Table 3; see Fig. S1 and Data Set S6 in the supplemental material). QIIME was used for Jaccard β-diversity principal component analysis (PCA) using the OTUs generated by mothur (Fig. 2C and D). The number of OTUs as a function of genetic distance was also calculated (see Fig. S2 in the supplemental material).

OTUs were also calculated at the 1% divergence level (99% similarity). Even at this more stringent cutoff, the largest OTUs remained relatively unchanged, though their sizes were reduced slightly. For example, the 10 OTUs that contain 10 or more sequences at the 3% divergence level retain an average of 90% of their size at the 1% divergence level. The largest OTU contains 828 sequences at the 3% divergence level and 807 sequences at the 1% level. None of the 3% OTUs are split into roughly equal-size OTUs at the 1% level. Instead, substantially more singleton OTUs are generated. Since the 1% level does not change the major OTUs and 3% divergence was used in our analysis of bacterial communities in the same samples (11), we used 3% divergence for OTU clustering in this study.

Taxonomy assignment. A taxonomic assignment was given to each OTU using two separate approaches. First, the identities of all sequences within each OTU were determined. Second, to resolve inter-OTU taxonomic conflicts, a phylogenetic tree was constructed containing the representative sequences of all OTUs and their closest relatives.

The full-length, untrimmed, unaligned reads representing the 2,222 sequences in the trimmed alignment were used for taxonomy assignment. We chose to use the full-length sequences because, although the final parts of the reads (roughly corresponding to the D2 region) could not be satisfactorily aligned, the trimmed bases still represent useful taxonomic information. All sequences were queried against the SILVA Large Subunit Parc Database release 108 (59) using the blastn algorithm (1). We chose to

TABLE 2 Distribution of the 10 most common yeast OTUs within 15 *Drosophila* populations

	No. of sequences in indicated population															
	Fruit									Flower			Mushroom		Cactus	
Yeast	HCF	HPM	HPP	ICF	MAG	MEC	PON	SPP	TBB	ELD	IMH	SCA	FNS	NNS	MOV	Total
Candida amapae		1	21					54								76
Candida diversa		1			96	1										98
Candida panamericana group					1									135		136
Candida restingae						7				77		2				86
Candida tritomae group						1							56	45		102
Hanseniaspora occidentalis	227	154		66	23	1			1			1				473
Hanseniaspora uvarum group	15	34	63		68	121	260	113	60		91	2	1			828
Pichia kluyveri		8	12					17	3							40
Pichia kudriavzevii									72						1	73
Saccharomyces cerevisiae/S. paradoxus	4	9	2			29	1		13	32	1	67	2	1	50	211
All other taxa (41 OTUs)	1	2	2	7	4	8	3	6	5	4	3	3	10	20	21	99
Total	247	209	100	73	192	168	264	190	154	113	95	75	69	201	72	2,222

use this database over the Ribosomal Database Project's Classifier and the SILVA LSU reference database because preliminary taxonomy searches revealed that the last two databases do not contain some of the taxa present in our data set. The taxonomic identity of each sequence within each OTU was then used to determine a consensus taxonomy for the OTU. If all the taxonomy assignments agreed, the OTU was given that assignment. In cases of disagreement or with small OTUs (i.e., with one sequence, internal OTU disagreement is impossible), phylogenetic information was used to determine taxonomy, as described below.

To test whether using a more stringent OTU cutoff would affect our results, the above-mentioned taxonomy assignment was also performed at the 1% divergence level. In particular, we noted whether the large OTUs that contain several conflicting taxonomy assignments would be split into two or more smaller groups that would each agree internally. We found that at 1% divergence, more singletons were created, but no large OTUs with consistent taxonomic classification emerged from the larger 3% OTUs.

Although the SILVA LSU database is regularly updated, it is possible that newly defined taxa may not be present in the most recent release. Therefore, taxonomy assignment was also repeated using the entire NCBI database (as of 21 October 2011). The NCBI and SILVA assignments agreed, except for the following major disagreement: 217 sequences in the largest OTU termed *Hanseniaspora* sp. by SILVA are *Hanseniaspora thailandica* (33) according to the NCBI assignment (see Data Set S7 in the supplemental material). The remaining few disagreements involved OTUs of a single sequence and were corrected during the phylogenetic analysis described below.

Representative sequences from each of the 51 OTUs were queried against the SILVA LSU Parc database release 108 (59) using the blastn

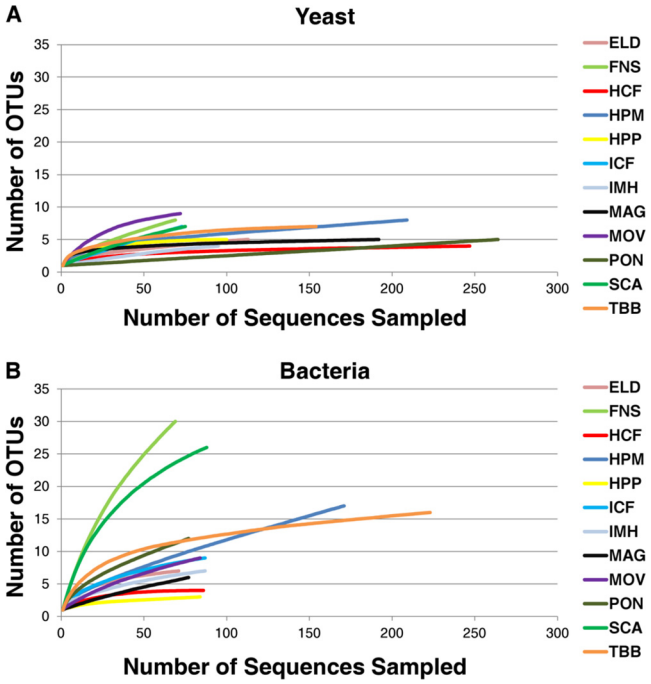

FIG 1 Comparison of yeast and bacterial rarefaction curves. All calculations were performed using mothur (7). OTUs were defined at the 3% divergence threshold using the average neighbor-clustering algorithm. Only the 12 populations for which both yeast and bacterial data (3) are available are shown. Library identifiers are given in Table 1. Rarefaction curves for populations MEC, NNS, and SPP are shown in Fig. S1 in the supplemental material. (A) Yeast rarefaction curves (this study). (B) Bacterial rarefaction curves (11).

TABLE 3 Comparison of α-diversity measurements of yeast and bacterial communities[a]

	Observed richness		Chao1 richness		Shannon diversity	
Library	Yeast	Bacteria	Yeast	Bacteria	Yeast	Bacteria
ELD	5	7	5.5	7.5	0.77	1.32
FNS	8	30	13	57.2	0.79	3.04
HCF	4	4	4	4	0.34	0.54
HPM	8	17	14	50	0.88	1.34
HPP	5	3	5	3	1.03	0.32
ICF	7	9	12	9.75	0.48	1.22
IMH	4	7	4.5	8.5	0.22	1.07
MAG	5	6	5	12	1.08	0.44
MOV	9	9	12	12.33	1.20	0.76
PON	5	12	11	22.5	0.10	1.48
SCA	7	26	9	30	0.52	2.9
TBB	7	16	7	26	1.17	1.86
Avg	6.17	12.17	8.50	20.23	0.72	1.36
SD	1.70	8.58	3.74	17.80	0.38	0.88

[a] All calculations were performed using mothur (67). Only the 12 populations for which both yeast and bacterial data are available are shown here. Library identifiers are given in Table 1. Further yeast diversity information and the data for samples MEC, NNS, and SPP are available in Data Set S6 in the supplemental material.

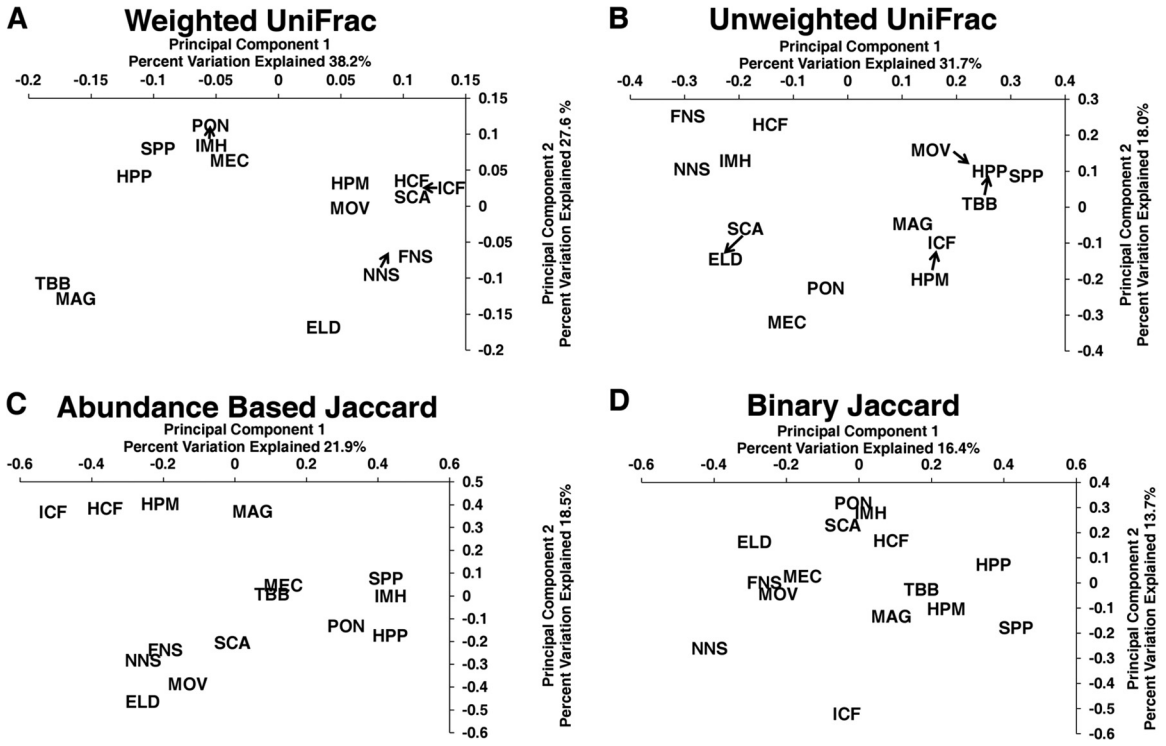

FIG 2 Principal component analysis of yeast communities. (A and B) A representative sequence from each OTU was generated using mothur (67), and a tree of the sequences was generated with FastTree (57) using *Amanita muscaria* as an outgroup. Principal component analysis was done with the FastUniFrac Web application (26) using both the weighted (A) and the unweighted (B) algorithms. (C and D) β-Diversity principal component analysis was performed in QIIME (9) for both the abundance-based (C) and binary (D) Jaccard metrics. For clarity, overlapping data points were moved; the arrows indicate their true positions. Library identifiers are given in Table 1.

algorithm (1). The 10 closest hits to each sequence were extracted from the SILVA LSU fasta file, and duplicate sequences were removed. The 401 unique sequences were aligned with the representative OTU sequences using the genafpair algorithm in MAFFT. To root this tree of Ascomycota, the basidiomycete *Amanita muscaria* was included as an outgroup. Since this database contains some sequences that include the small-subunit (SSU) and ITS regions in addition to the LSU, the resulting alignment was trimmed to retain only the LSU region contained in all representative OTU sequences. This trimmed alignment was realigned using MAFFT. A phylogenetic tree was created using FastTree and the GTR option (58). The results were visualized using Dendroscope (30) (see Fig. S3 in the supplemental material). The NEWICK tree file is available on BioTorrents (http://www.biotorrents.net/details.php?id=263; 41).

This phylogenetic tree was used to refine the taxonomic assignment of each OTU. If there was disagreement within an OTU or between the SILVA and NCBI assignments, the tree usually showed a polytomy or near polytomy between several different yeast species, suggesting that the amount of sequence data was not sufficient to determine the OTU to the species level. In such cases, the OTU was given a more general name representative of this ambiguity. In a few cases, the questionable OTU branched outside a monophyletic group composed of several different species. Such OTUs were also named to represent their phylogenetic positions. Finally, in one case, both the SILVA and the NCBI assignments found an OTU composed of two separate genera. The phylogenetic analysis found that the OTU was equally related to both genera. This OTU was named based on the shared family name of the genera. Final taxonomy assignments, along with the SILVA, NCBI, and phylogenetic information, can be found in Data Set S7 in the supplemental material.

Tree building and UniFrac principal component analysis. Each of the 51 representative OTU sequences generated using mothur were aligned to the basidiomycete outgroup *Amanita muscaria* using MAFFT, the genafpair algorithm, and the add function. This produced an alignment of 1,140 positions, which was trimmed to include only the LSU region, as described above. This trimmed alignment was realigned using MAFFT and the genafpair algorithm. A tree was built from the final alignment using FastTree and the GTR option (58). Phylogenetic analysis of the community composition was then performed using the FastUniFrac Web application (26) (Fig. 2A and B).

Comparison to bacterial communities. The bacterial communities associated with 12 of the *Drosophila* populations used in this study have been characterized previously (11). These 12 libraries, containing 1,203 sequences in the bacterial data set and 1,663 sequences in the yeast data set, were used to provide a direct comparison of the bacterial and yeast communities in the same hosts. Both phylogenetically based and OTU-based approaches were used to compare the two microbial communities.

The phylogenetically based comparisons were done using the yeast tree as described above, although the UniFrac environment file was modified to contain only the sequences from the 12 comparison populations. Similarly, the bacterial tree was created using the representative bacterial OTU sequences from these 12 *Drosophila* populations (11). Briefly, the sequences were aligned using Infernal and a high-quality bacterium-specific covariance model (55). OTUs were created at the 3% divergence level using mothur and the average neighbor-clustering algorithm (67). In total, 101 representative sequences are present in these 12 populations. A tree was built with the final alignment using FastTree and the GTR option (58). PCA was then performed for both the yeast and the bacterial microbiomes using FastUniFrac (26).

OTU-based comparisons were done using various β-diversity measurements calculated using QIIME (9) (Table 4). These particular measurements were chosen because our data were intended to measure the variation in community structure among a set of samples (3). Additional measurements were included to test for the role of joint absences in com-

TABLE 4 Procrustes *P* values comparing yeast and bacterial community structures

Type of comparison[a]	*P* value[b]
Unweighted UniFrac	0.426
Weighted, nonnormalized UniFrac	0.086
Weighted, normalized UniFrac	0.076
Jaccard	0.672
Binary Jaccard	0.478
Bray-Curtis	0.282
Sorensen-Dice	0.428
Euclidean	0.270
Binary Euclidean	0.117
Gower	0.139
Manhattan	0.088

[a] UniFrac PCA was performed using the FastUniFrac Web application (26). β-Diversity calculations and PCA were done with QIIME (9).
[b] Procrustes significance values were calculated with QIIME as described in the text. Nonsignificant values indicate no correlation between the distribution patterns of the yeast and bacterial communities associated with *Drosophila*.

munity structure (3). Each β-diversity measurement was calculated for both the yeast and the bacterial communities. PCA was then performed on each measurement using QIIME.

Procrustes analysis was then used to compare the yeast and bacterial principal component (PC) results from both FastUniFrac and QIIME. Procrustes analysis rotates and stretches the points within a matrix so that they most closely align with the points of a second matrix (53). The distance between corresponding points in the superimposed matrices is used to test whether variation is similarly distributed in each of the two matrices. Measuring the distance between the corresponding points in the transformed, superimposed matrices gives an M^2 value. By randomly changing the identity of the points in one matrix, one can calculate how often an M^2 value smaller than the observed M^2 value would arise by chance, thus giving a *P* value measuring the similarity of the two matrices (10).

To compare the yeast and bacterial communities, each set of coordinate matrices (PCs) was transformed to place individual data points (representing the two separate microbial communities of a single population) as near each other as possible in 12-dimensional space. These 12 dimensions represent nearly 100% of the variation present in either microbial community. Next, 1,000 Monte Carlo simulations were performed to determine the significance of each yeast-bacterial comparison (Table 4) (10). Finally, a plot of each comparison was created (Fig. 3; see Fig. S4 to S13 in the supplemental material) and visualized using KiNG (http://kinemage.biochem.duke.edu/software/king.php). The KiNG data file for Fig. 3 and the other comparisons (Table 4) is available through BioTorrents (http://www.biotorrents.net/details.php?id=263).

In addition to PCA-based Procrustes analysis, alpha diversity measurements of the *Drosophila* bacterial communities (see Table S3 in reference 11) were compared to those of the yeast communities (Table 3).

Nucleotide sequence accession numbers. The 2,222 sequences used in the final analysis were submitted to NCBI under accession numbers JX407386 to JX409607.

RESULTS

Dominant yeast taxa and the role of host diet in shaping *Drosophila*-associated yeast communities. *Drosophila* populations were chosen to maximize the breadth of their phylogenetic, geographic, and ecological diversity. Although greater host sampling may reveal additional features of *Drosophila*-associated yeast communities, the characterized samples nonetheless provide valuable information regarding this host-microbe system. In this study, we found that *Hanseniaspora* is the dominant yeast genus associated with *Drosophila*, representing 59% of the yeast sequences in our samples (Table 2). Members of this yeast genus are considered physiological specialists known to colonize decaying fruits (74). *Hanseniaspora* was present in all but two *Drosophila* populations. It was not detected in one of the three flower-feeding *Drosophila* species (*Drosophila elegans*, collected from *Brugmansia* flowers; population ELD) and in the cactus-feeding *Drosophila mojavensis* (population MOV). (Note that a single *Hanseniaspora vineae* sequence was found in *Drosophila neotestaceae*, collected from *Russula* mushrooms [population NNS] [see Data Set S5 in the supplemental material]) *Hanseniaspora* species are also relatively rare in the two mushroom feeders, representing only 10% and 1% of the yeast communities in these populations.

The most abundant OTU (referred to in Table 2 as the *Hanseniaspora uvarum* group) is closely related to the species complex composed of *Hanseniaspora uvarum*, *Hanseniaspora lachancei*, *Hanseniaspora guilliermondii*, *Hanseniaspora pseudoguilliermondii*, *Hanseniaspora opuntiae*, *Hanseniaspora meyeri*, and *Hanseniaspora clermontiae* (8). The recently described *H. thailandica* is also closely related to this OTU (33). The OTU is present in all but four of the *Drosophila* populations and is the dominant OTU (>50%) in five of them (populations HPP, IMH, MEC, PON, and SPP) (Table 2). The abundance of the OTU leads to the close association of these samples in PCAs that take abundance into account (Fig. 2A and C). The five samples were collected at geographically distant locations, including Hawaii, North America, and the East African island nation of Seychelles. Members of this *Hanseniaspora* group are known to inhabit food sources that *Drosophila* commonly interacts with and have previously been isolated from insects, including *Drosophila*. For example, *H. uvarum*

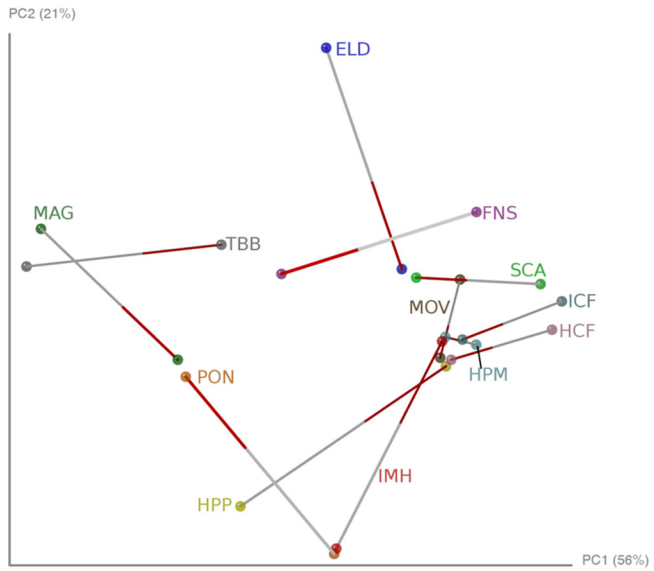

FIG 3 Plot showing Procrustes analyses of transformed PCAs of yeast and bacterial communities. Each pair of dots connected by an edge represents a paired set of yeast and bacterial data points from the same host population. The gray half of the edge is connected to the yeast data point, while the red half is connected to the bacterial data point. Note that the yeast and bacterial communities from the same hosts are no closer than a random pair of yeast and bacterial communities. The plot represents weighted, normalized UniFrac data ($P = 0.076$). The KiNG data file for this and the other comparisons (Table 3) is available through BioTorrents (http://www.biotorrents.net/details.php?id=263).

was isolated from *Parahancornia amapa* fruits and several Amazonian *Drosophila* species (50), as well as from green June beetles feeding on peaches (80). *H. uvarum* has also been found with several *Drosophila* species collected in and around California's Yosemite National Park (56) and with *Drosophila suzukii* collected in Northern California (28). *H. opuntiae* was first described from isolates from prickly pear cactus grown in Hawaii (7), and indeed, this *Hanseniaspora* OTU is prevalent (>35%) in all three Hawaiian *Drosophila* populations (populations IMH, MAG, and PON).

Another common *Hanseniaspora* OTU is closely related to *Hanseniaspora occidentalis*. Ninety-five percent of this OTU comes from three fruit-feeding *Drosophila* populations collected at the Wolfskill experimental orchard outside Davis, CA (populations HPM, HCF, and ICF) (Table 2). *H. occidentalis* is rarely found with flower-feeding *Drosophila* and never with mushroom or cactus feeders. It has been suggested that *H. occidentalis* var. *citrica* is a habitat specialist on oranges and their products (8), which is consistent with our finding that *H. occidentalis* represented >90% of the yeast communities found with both populations of *Drosophila* collected from citrus fruit (populations HCF and ICF). Additionally, *H. occidentalis* var. *occidentalis* was isolated from *D. melanogaster* feeding on mangos in Brazil (49), and this study found it associated with Hawaiian *Drosophila sulfurigaster* collected from mangos (population MAG).

An OTU closely related to *Saccharomyces paradoxus* and *Saccharomyces cerevisiae* is the third most abundant yeast OTU associated with *Drosophila*. It is found with more *Drosophila* populations than any other OTU (12 of 15). It is the only OTU found in flies of all four feeding types, although it represents only a small proportion (1%) of the yeast communities associated with mushroom feeders (Table 2). It is also found on every continent, although it is relatively rare (<1%) in the Hawaiian *Drosophila*. Although *S. cerevisiae* is usually thought of as a domesticated strain closely tied to viticulture, it and *S. paradoxus* can be found on tree bark and decaying leaves and in the soil (40, 60, 72). These taxa have previously been found with *Drosophila* in Brazil (49) and California (56).

Two abundant *Candida* OTUs were found almost exclusively with two mushroom-feeding *Drosophila* populations collected in New York (populations FNS and NNS) (Table 2). Sixty-seven percent of the yeast community associated with *D. neotestacea* is an OTU very closely related to *Candida panamericana*, *Candida atakaporum*, *Candida tanzawaensis*, and *Candida ambrosiae*. Another OTU related to *Candida tritomae*, *Candida pallodes*, and *Candida lycoperdinae* makes up 81% and 22% of the *Drosophila falleni* and *D. neotestacea* yeast communities, respectively. All six of these *Candida* species are known associates of basidiocarp-feeding beetles (77, 78).

A substantial portion (68%) of the yeast community associated with flower-feeding *D. elegans* is represented by an OTU that is closely related to *Candida restingae* (population ELD) (Table 2). This yeast was originally isolated from nitidulid beetles feeding on *Pilosocereus arrabidae* flowers in Brazil (62). An OTU closely related to *Candida diversa* makes up 50% of the yeast community associated with mango-feeding *D. sulfurigaster* (population MAG) (Table 2); this species has been previously isolated from *Drosophila* (49–51). Both *Drosophila* species collected from *Opuntia* fruits (populations HPP and SPP) are associated with an OTU closely related to *Candida amapae* (a possible anamorph of *Saccharomycopsis crataegensis*) (Table 2). This species of *Candida* was originally isolated from the Amazonian fruit *Parahancornia amapa* (52) but was not found with *Drosophila* feeding on the same fruit (50).

Overall, our results are consistent with previous records of yeasts associated with insects and their food sources and suggest that the yeast communities of *Drosophila* populations are shaped to a large extent by their diet.

The relatively small role of host species in shaping the *Drosophila*-associated yeast communities. In two instances, two different *Drosophila* species were collected simultaneously from the same substrate. *Drosophila simulans/D. melanogaster* (these closely related taxa were pooled, since females are hard to distinguish morphologically) and *Drosophila hydei* were collected from prickly pear fruit (populations SPP and HPP, respectively), while *D. hydei* and *Drosophila immigrans* were collected from citrus fruit (HCF and ICF, respectively). In both cases, the dominant yeast taxon is the same for both species (Table 2), and abundance-based PCA confirms the similarity of their yeast communities (Fig. 2A and C). A comparison of the less abundant taxa does reveal differences between the two populations collected from citrus fruit. *H. occidentalis* was not detected in *D. immigrans* but comprises 6% of the yeast community of *D. hydei* (Table 2), and many of the very low-abundance taxa are not shared between the two fly species (see Data Set S5 in the supplemental material). PCA using unweighted UniFrac and binary Jaccard β-diversity confirms the dissimilarity of the two populations based on their rare taxa (Fig. 2B and D).

Yeast diversity in comparison to bacteria. Observed and expected (Chao1) yeast OTU richness is low in all *Drosophila* samples (Table 3; see Data Set S6 in the supplemental material). For nearly half of the populations, Chao1 estimates suggest that fewer than 10 OTUs would be expected had they been sampled exhaustively. No more than 20 OTUs are estimated for any of the samples. In the 12 populations for which both yeast and bacterial data are available, observed and expected richness is lower for the yeasts than for bacteria (Table 3). A comparison of rarefaction curves for these populations confirms that the bacterial microbiome has greater diversity (Fig. 1).

Analysis of rarefaction curves showed that different populations were sampled to different depths. The rarefaction curves for several fruit-feeding *Drosophila* leveled off, suggesting that nearly all the yeast taxa present in these hosts have been detected. In contrast, some rare but potentially important taxa associated with other fly populations remain to be identified.

Different rates of sequence evolution could affect OTU level comparisons of yeast and bacterial communities. For example, the lower observed richness in yeast communities could, in principle, be the result of slower divergence of LSU than of SSU sequences, although a survey of 500 ascomycetous yeasts found that the divergence rate of LSU sequences is comparable to that of SSU sequences and is potentially greater in certain closely related taxa (37). Sequence diversity within microbial communities can be measured by calculating the number of OTUs as a function of genetic distance (17). Using this method to compare the yeast and bacterial communities, we found that strain level differences (the point where the curve representing the relationship between genetic distance and the number of OTUs increases in slope [17]) begin at roughly the same divergence (1%) for both the yeast LSU and the bacterial SSU data (see Fig. S2 in the supplemental material). Although OTUs based upon sequence divergence are not

always directly related to taxonomic boundaries, Fig. S2 suggests that an OTU comparison using the yeast LSU and the bacterial SSU is valid.

Comparison of yeast and bacterial communities. Procrustes analysis revealed no significant correlation between the compositions of the yeast and bacterial communities across 12 *Drosophila* populations. Neither phylogenetically informed (UniFrac) nor OTU-based comparisons between the distributions of bacterial and yeast symbionts showed significant similarities (Table 4). Visual inspection of the first two dimensions of the superimposed, transformed PCA plots confirms the incongruent distribution of the two symbiont groups (Fig. 3; see Fig. S4 to S13 in the supplemental material).

DISCUSSION

***Drosophila*-associated yeast and bacterial communities show similar distribution patterns.** Previous analysis of bacterial communities associated with different *Drosophila* species and populations has revealed several general patterns (11). Namely, *Drosophila* has taxonomically restricted bacterial communities; the same bacterial lineages are associated with different host species, diets, and locations; and host diet has a greater effect on bacterial community composition than host species. Here, we found similar patterns in the *Drosophila*-associated yeast communities.

Natural *Drosophila* populations carry a very limited number of yeast taxa. Despite the phylogenetic, ecological, and geographic diversity of the host populations surveyed in this study, only a few OTUs comprise a large proportion of the total yeast community (Table 2). The overall OTU richness is even smaller for the *Drosophila*-associated yeast communities than for the bacterial communities (Table 3 and Fig. 1). This conclusion is conservative, since the 12 yeast samples that came from the same hosts as the bacterial data (11) were sequenced to greater depth (1,663 yeast reads compared to 1,203 reads for bacteria). Although the yeast LSU and the bacterial SSU may not offer the most direct comparison, our analysis is robust to this caveat (see Fig. S2 in the supplemental material).

Similar to the bacterial communities, the same yeast lineages are present in many populations regardless of host species, diet, or location (Table 2). Although 21 OTUs are present in only one population each, these OTUs represent only ~1% of all yeasts associated with *Drosophila*. On the other hand, the dominant OTUs from each host population are usually found in other, often geographically distant populations as well. Several particularly wide-ranging OTUs are seen in nearly every population. Furthermore, many of the yeast taxa identified in this study have previously been isolated in association with different *Drosophila* species in different regions (49–51, 56).

Finally, our results suggest that diet plays a more important role than host species in structuring the *Drosophila*-associated yeast communities. Several *Candida* OTUs are found almost exclusively with mushroom-feeding flies, while the yeast taxa that are dominant in other populations are rarely present in the mushroom feeders (Table 2). Conversely, when different *Drosophila* species are simultaneously collected from a single food source, their yeast communities are very similar (Table 2 and Fig. 2). Although manipulative experiments will be needed to confirm these observations, this work nonetheless provides suggestive evidence of the importance of host diet compared to host species.

Little correlation between the distribution of yeast and bacterial taxa. In the 12 *Drosophila* populations for which both yeast and bacterial data are available, we found no significant correlation between these two components of the host symbiont communities (Table 4 and Fig. 3; see Fig. S4 to S13 in the supplemental material). The reasons for this are not clear, but we can envision several possible explanations. The simplest of these is that there is no correlation between the yeast and bacterial communities that inhabit the fruits, flowers, mushrooms, cacti, and other substrates that serve as *Drosophila* food sources. This could occur if different physicochemical attributes of the substrate (pH, salinity, moisture content, phytochemicals, etc.) have different effects on yeast and bacteria. For example, soil bacterial communities are more strongly affected by soil pH than fungal communities (64). It is also possible that yeasts and bacteria have different turnover times or growth rates in the *Drosophila* digestive tract. Even though the same host individuals were used to characterize both communities, controlling for stochastic interactions with the habitat, one of the symbiont groups may take longer to pass through the gut or longer to grow to a detectable density. This would lead to a situation where one symbiont group represents the substrate that the flies visited most recently while the other group reflects a substrate visited at an earlier time. To test these hypotheses, future studies will need to characterize the microbial communities of the *Drosophila* food sources and compare them to the host-associated communities.

Alternatively, the host could be independently controlling the persistence of the two symbiont groups in its gut. In the simplest case, one symbiont group could be a random sample of the ingested microbes while the second could be subject to host control. For example, the structure of symbiont communities may depend on pH. The *Drosophila* digestive tract contains distinct regions of varying pH, most notably a very acidic midgut (68). If yeasts can tolerate a wider range of pH conditions than bacteria (81), the bacteria may be subject to stronger host control than the yeasts. Another important physicochemical aspect of the *Drosophila* intestine is the level of reactive oxygen species (ROS), which has been shown to affect both bacterial and yeast growth (24, 25). If environmentally acquired yeasts have lower resistance to ROS than similarly acquired bacteria, then the intestinal yeast community will be a constrained subset of the possible community while the intestinal bacterial community will more closely resemble that of the environment. Finally, since bacteria and fungi are recognized and cleared by semi-independent innate immunity pathways (42), a more nuanced control of the two symbiont groups is also possible. As both bacteria and yeasts can have significant effects on fly physiology and fitness (2, 70, 76), the host could be under selection to maintain only the most beneficial symbionts for that particular diet or location.

In summary, our results provide the first study of multiple symbiont groups associated with natural populations of *Drosophila*. Future *Drosophila*-microbe work, whether it is focused on genetics, metabolism, or evolution, and animal-microbe research in general, should take a similarly holistic view and consider all possible symbionts that are associated with a given host.

ACKNOWLEDGMENTS

This work was supported by a UC-Davis Center for Population Biology Graduate Student Research Award to J.A.C., by an Alfred P. Sloan Foundation grant to J.A.E., and by NSF grant IOS-0815141 to A.K. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

This work would not have been possible without the help of many colleagues around the world. Wild *Drosophila* samples were collected for us by Olga Barmina, Marie-Louise Cariou, Bill Etges, Jenna Morgan Lang, Delphine Legrand, Shane McEvey, Chen-Siang Ng, and John True. We thank Jenna Morgan Lang and David Plachetzki for advice on phylogenetic and bioinformatic methods and Olga Barmina and Matt Rolston for technical help. Kyria Boundy-Mills and Marc-André Lachance provided helpful suggestions for improving the manuscript.

REFERENCES

1. **Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ.** 1990. Basic Local Alignment Search Tool. J. Mol. Biol. **215**:403–410.
2. **Anagnostou C, Dorsch M, Rohlfs M.** 2010. Influence of dietary yeasts on Drosophila melanogaster life-history traits. Entomol. Exp. Appl. **136**:1–11.
3. **Anderson MJ, et al.** 2011. Navigating the multiple meanings of beta diversity: a roadmap for the practicing ecologist. Ecol. Lett. **14**:19–28.
4. **Barker JSF, Starmer WT (ed).** 1982. Ecological genetics and evolution: the cactus-yeast-drosophila model system. Academic Press Australia, New York, NY.
5. **Begon M.** 1982. Yeast and *Drosophila*, p 345–384. *In* Ashburner M, Carson H, Thompson JN (ed), The genetics and biology of *Drosophila*. Academic Press, London, United Kingdom.
6. **Brummel T, Ching A, Seroude L, Simon AF, Benzer S.** 2004. Drosophila lifespan enhancement by exogenous bacteria. Proc. Natl. Acad. Sci. U. S. A. **101**:12974–12979.
7. **Cadez N, Poot GA, Raspor P, Smith MT.** 2003. Hanseniaspora meyeri sp nov., Hanseniaspora clermontiae sp nov., Hanseniaspora lachancei sp nov and Hanseniaspora opuntiae sp nov., novel apiculate yeast species. Int. J. Syst. Evol. Microbiol. **53**:1671–1680.
8. **Cadez N, Raspor P, Smith MT.** 2006. Phylogenetic placement of Hanseniaspora-Kloeckera species using multigene sequence analysis with taxonomic implications: descriptions of Hanseniaspora pseudoguilliermondii sp. nov and Hanseniaspora occidentalis var. citrica var. nov. Int. J. Syst. Evol. Microbiol. **56**:1157–1165.
9. **Caporaso JG, et al.** 2010. QIIME allows analysis of high-throughput community sequencing data. Nat. Methods **7**:335–336.
10. **Caporaso JG, et al.** 2011. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proc. Natl. Acad. Sci. U. S. A. **108**: 4516–4522.
11. **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 Genet. **7**:e1002272. doi:10.1371/journal.pgen.1002272.
12. **Cole JR, et al.** 2009. The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. Nucleic Acids Resh. **37**:D141–D145.
13. **Cooper DM.** 1960. Food preferences of larval and adult Drosophila. Evolution **14**:41–55.
14. **Corby-Harris V, et al.** 2007. Geographical distribution and diversity of bacteria associated with natural populations of Drosophila melanogaster. Appl. Environ. Microbiol. **73**:3470–3479.
15. **Cox CR, Gilmore MS.** 2007. Native microbial colonization of Drosophila melanogaster and its use as a model of Enterococcus faecalis pathogenesis. Infect. Immun. **75**:1565–1576.
16. **Dacunha AB, Shehata AMET, Deoliveira W.** 1957. A study of the diets and nutritional preferences of tropical species of Drosophila. Ecology **38**: 98–106.
17. **Dethlefsen L, McFall-Ngai M, Relman DA.** 2007. An ecological and evolutionary perspective on human-microbe mutualism and disease. Nature **449**:811–818.
18. **Dillon RJ, Dillon VM.** 2004. The gut bacteria of insects: nonpathogenic interactions. Annu. Rev. Entomol. **49**:71–92.
19. **Dobzhansky T, Cooper DM, Phaff HJ, Knapp EP, Carson HL.** 1956. Studies on the ecology of Drosophila in the Yosemite region of California. 4. Differential attraction of species of Drosophila to different species of yeasts. Ecology **37**:544–550.
20. **Ganter PF.** 1988. The vectoring of cactophilic yeasts by Drosophila. Oecologia **75**:400–404.
21. **Ganter PF.** 2006. Yeast and invertebrate associations, p 303–370. *In* Péter G, Rosa C (ed), Biodiversity and ecophysiology of yeasts. Springer, Berlin, Germany.
22. **Gigioli I.** 1897. Insects and yeasts. Nature **56**:575–577.
23. **Gottar M, et al.** 2006. Dual detection of fungal infections in Drosophila via recognition of glucans and sensing of virulence factors. Cell **127**:1425–1437.
24. **Ha EM, et al.** 2009. Regulation of DUOX by the G alpha q-phospholipase C beta-Ca2+ PAthway in Drosophila gut immunity. Dev. Cell **16**:386–397.
25. **Ha EM, Oh CT, Bae YS, Lee WJ.** 2005. A direct role for dual oxidase in Drosophila gut immunity. Science **310**:847–850.
26. **Hamady M, Lozupone C, Knight R.** 2010. Fast UniFrac: facilitating high-throughput phylogenetic analyses of microbial communities including analysis of pyrosequencing and PhyloChip data. ISME J. **4**:17–27.
27. **Hamady M, Walker JJ, Harris JK, Gold NJ, Knight R.** 2008. Error-correcting barcoded primers for pyrosequencing hundreds of samples in multiplex. Nat. Methods **5**:235–237.
28. **Hamby KA, Hernandez A, Boundy-Mills K, Zalom FG.** 2012. Associations of yeasts with spotted-wing Drosophila (Drosophila suzukii; Diptera: Drosophilidae) in cherries and raspberries. Appl. Environ. Microbiol. **78**:4869–4873.
29. **Heed WB, Starmer WT, Miranda M, Miller MW, Phaff HJ.** 1976. Analysis of yeast flora associated with cactiphilic Drosophila and their host plants in Sonoran-desert and its relation to temperate and tropical associations. Ecology **57**:151–160.
30. **Huson DH, et al.** 2007. Dendroscope: an interactive viewer for large phylogenetic trees. BMC Bioinformatics **8**:460.
31. **Ikeda-Ohtsubo W, Brune A.** 2009. Cospeciation of termite gut flagellates and their bacterial endosymbionts: Trichonympha species and 'Candidatus Endomicrobium trichonymphae'. Mol. Ecol. **18**:332–342.
32. **Iliev ID, et al.** 2012. Interactions between commensal fungi and the C-type lectin receptor dectin-1 influence colitis. Science **336**:1314–1317.
33. **Jindamorakot S, et al.** 2009. Three new species of bipolar budding yeasts of the genus Hanseniaspora and its anamorph Kloeckera isolated in Thailand. FEMS Yeast Res. **9**:1327–1337.
34. **Katoh K, Misawa K, Kuma K, Miyata T.** 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. **30**:3059–3066.
35. **Katoh K, Toh H.** 2010. Parallelization of the MAFFT multiple sequence alignment program. Bioinformatics **26**:1899–1900.
36. **Kurtzman CP, Fell JW, Boekhout T.** 2011. The yeasts: a taxonomic study, 5th ed. Elsevier, New York, NY.
37. **Kurtzman CP, Robnett CJ.** 1998. Identification and phylogeny of ascomycetous yeasts from analysis of nuclear large subunit (26S) ribosomal DNA partial sequences. Antonie Van Leeuwenhoek **73**:331–371.
38. **Kurtzman CP, Robnett CJ.** 1997. Identification of clinically important ascomycetous yeasts based on nucleotide divergence in the 5′ end of the large-subunit (26S) ribosomal DNA gene. J. Clin. Microbiol. **35**:1216–1223.
39. **Lachance MA, Gilbert DG, Starmer WT.** 1995. Yeast communities associated with Drosphila species and related flies in an eastern oak-pine forest— a comparison with Western communities. J. Ind. Microbiol. **14**: 484–494.
40. **Landry CR, Townsend JP, Hartl DL, Cavalieri D.** 2006. Ecological and evolutionary genomics of Saccharomyces cerevisiae. Mol. Ecol. **15**:575–591.
41. **Langille MGI, Eisen JA.** 2010. BioTorrents: a file sharing service for scientific data. PloS One **5**:e10071. doi:10.1371/journal.pone.0010071.
42. **Lemaitre B, Hoffmann J.** 2007. The host defense of Drosophila melanogaster. Annu. Rev. Immunol. **25**:697–743.
43. **Lemaitre B, Reichhart JM, Hoffmann JA.** 1997. Drosophila host defense: differential induction of antimicrobial peptide genes after infection by various classes of microorganisms. Proc. Natl. Acad. Sci. U. S. A. **94**: 14614–14619.
44. **Ley RE, et al.** 2008. Evolution of mammals and their gut microbes. Science **320**:1647–1651.
45. **Ley RE, Turnbaugh PJ, Klein S, Gordon JI.** 2006. Human gut microbes associated with obesity. Nature **444**:1022–1023.
46. **Lhocine N, et al.** 2008. PIMS modulates immune tolerance by negatively regulating Drosophila innate immune signaling. Cell Host Microbe **4**:147–158.
47. **Lindsay SL.** 1958. Food preferences of Drosophila larvae. Am. Nat. **92**: 279–285.
48. **McOrist AL, Jackson M, Bird AR.** 2002. A comparison of five methods

for extraction of bacterial DNA from human faecal samples. J. Microbiol. Methods **50**:131–139.
49. **Morais PB, Hagler AN, Rosa CA, Mendonca-Hagler LC, Klaczko LB.** 1992. Yeasts associated with Drosophila in tropical forests of Rio de Janeiro, Brazil. Can. J. Microbiol. **38**:1150–1155.
50. **Morais PB, Martins MB, Klaczko LB, Mendonca-Hagler LC, Hagler AN.** 1995. Yeast succession in the Amazon fruit Parahancornia amapa as resource partitioning among Drosophila spp. Appl. Environ. Microbiol. **61**:4251–4257.
51. **Morais PB, Rosa CA, Hagler AN, Mendonca-Hagler LC.** 1995. Yeast communities as descriptors of habitat use by the Drosophila-Fasciola subgroup (Repleta group) in Atlantic rainforests. Oecologia **104**:45–51.
52. **Morais PB, Rosa CA, Meyer SA, Mendonca-Hagler LC, Hagler AN.** 1995. Candida amapae, a new amino acid-requiring yeast from the Amazonian fruit Parahancornia amapa. J. Ind. Microbiol. **14**:531–535.
53. **Muegge BD, et al.** 2011. Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. Science **332**: 970–974.
54. **Nam YD, et al.** 2008. Bacterial, archaeal, and eukaryal diversity in the intestines of Korean people. J. Microbiol. **46**:491–501.
55. **Nawrocki EP, Kolbe DL, Eddy SR.** 2009. Infernal 1.0: inference of RNA alignments. Bioinformatics **25**:1335–1337.
56. **Phaff HJ, Miller MW, Recca JA, Shifrine M, Mrak EM.** 1956. Studies on the ecology of Drosophila in the Yosemite region of California. 2. Yeasts found in the alimentary canal of Drosophila. Ecology **37**:533–538.
57. **Price MN, Dehal PS, Arkin AP.** 2010. FastTree 2—approximately maximum-likelihood trees for large alignments. PloS One **5**:e9490. doi: 10.1371/journal.pone.0009490.
58. **Price MN, Dehal PS, Arkin AP.** 2009. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. Mol. Biol. Evol. **26**:1641–1650.
59. **Pruesse E, et al.** 2007. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Res. **35**:7188–7196.
60. **Replansky T, Koufopanou V, Greig D, Bell G.** 2008. Saccharomyces sensu stricto as a model system for evolution and ecology. Trends Ecol. Evol. **23**:494–501.
61. **Reuter M, Bell G, Greig D.** 2007. Increased outbreeding in yeast in response to dispersal by an insect vector. Curr. Biol. **17**:R81–R83.
61a. **Roche.** 2011. Roche technical bulletin, TCB no. 2011-001. Roche, Basel, Switzerland.
62. **Rosa CA, et al.** 1999. Kodamaea nitidulidarum, Candida restingae and Kodamaea anthophila, three new related yeast species from ephemeral flowers. Int. J. Syst. Bacteriol. **49**:309–318.
63. **Round JL, O'Connell RM, Mazmanian SK.** 2010. Coordination of tolerogenic immune responses by the commensal microbiota. J. Autoimmun. **34**:J220–J225.
64. **Rousk J, et al.** 2010. Soil bacterial and fungal communities across a pH gradient in an arable soil. ISME J. **4**:1340–1351.
65. **Ryu JH, et al.** 2008. Innate immune homeostasis by the homeobox gene Caudal and commensal-gut mutualism in Drosophila. Science **319**:777–782.
66. **Scanlan PD, Marchesi JR.** 2008. Micro-eukaryotic diversity of the human distal gut microbiota: qualitative assessment using culture-dependent and -independent analysis of faeces. ISME J. **2**:1183–1193.
67. **Schloss PD, et al.** 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl. Environ. Microbiol. **75**:7537–7541.
68. **Shanbhag S, Tripathi S.** 2009. Epithelial ultrastructure and cellular mechanisms of acid and base transport in the Drosophila midgut. J. Exp. Biol. **212**:1731–1744.
69. **Sharon G, et al.** 2010. Commensal bacteria play a role in mating preference of Drosophila melanogaster. Proc. Natl. Acad. Sci. U. S. A. **107**: 20051–20056.
70. **Shin SC, et al.** 2011. Drosophila microbiome modulates host developmental and metabolic homeostasis via insulin signaling. Science **334**:670–674.
71. **Six DL, Klepzig KD.** 2004. Dendroctonus bark beetles as model systems for studies on symbiosis. Symbiosis **37**:207–232.
72. **Sniegowski PD, Dombrowski PG, Fingerman E.** 2002. Saccharomyces cerevisiae and Saccharomyces paradoxus coexist in a natural woodland site in North America and display different levels of reproductive isolation from European conspecifics. FEMS Yeast Res. **1**:299–306.
73. **Starmer WT, Fogleman JC.** 1986. Coadaptation of Drosophila and yeasts in their natural habitat. J. Chem. Ecol. **12**:1037–1055.
74. **Starmer WT, Lachance MA.** 2011. Yeast ecology, p 65–83. *In* Kurtzman CP, Fell JW, Boekhout T (ed), The yeasts: a taxonomic study, 5th ed, vol 1. Elsevier, New York, NY.
75. **Starmer WT, Peris F, Fontdevila A.** 1988. The transmission of yeasts by Drosophila-Buzzatii during courtship and mating. Anim. Behav. **36**: 1691–1695.
76. **Storelli G, et al.** 2011. Lactobacillus plantarum promotes Drosophila systemic growth by modulating hormonal signals through TOR-dependent nutrient sensing. Cell Metab. **14**:403–414.
77. **Suh SO, McHugh JV, Blackwell M.** 2004. Expansion of the Candida tanzawaensis yeast clade: 16 novel Candida species from basidiocarp-feeding beetles. Int. J. Syst. Evol. Microbiol. **54**:2409–2429.
78. **Suh SO, Nguyen NH, Blackwell M.** 2006. A yeast clade near Candida kruisii uncovered: nine novel Candida species associated with basidioma-feeding beetles. Mycol. Res. **110**:1379–1394.
79. **Turnbaugh PJ, et al.** 2006. An obesity-associated gut microbiome with increased capacity for energy harvest. Nature **444**:1027–1031.
80. **Vishniac HS, Johnson DT.** 1990. Development of a yeast flora in the adult green June beetle (Cotinis-Nitida, Scarabaeidae). Mycologia **82**:471–479.
81. **Wheeler KA, Hurdman BF, Pitt JI.** 1991. Influence of pH on the growth of some toxigenic species of Aspergillus, Penicillium and Fusarium. Int. J. Food Microbiol. **12**:141–150.
82. **Willing BP, et al.** 2010. A pyrosequencing study in twins shows that gastrointestinal microbial profiles vary with inflammatory bowel disease phenotypes. Gastroenterology **139**:1844–1854.
83. **Wong CNA, Ng P, Douglas AE.** 2011. Low-diversity bacterial community in the gut of the fruitfly Drosophila melanogaster. Environ. Microbiol. **13**:1889–1900.