npg

ORIGINAL ARTICLE

The inconstant gut microbiota of *Drosophila* species revealed by 16S rRNA gene analysis

Adam C-N Wong[1,2], John M Chaston[1,2] and Angela E Douglas[1]
[1]Department of Entomology, Comstock Hall, Cornell University, Ithaca, NY, USA

The gut microorganisms in some animals are reported to include a core microbiota of consistently associated bacteria that is ecologically distinctive and may have coevolved with the host. The core microbiota is promoted by positive interactions among bacteria, favoring shared persistence; its retention over evolutionary timescales is evident as congruence between host phylogeny and bacterial community composition. This study applied multiple analyses to investigate variation in the composition of gut microbiota in drosophilid flies. First, the prevalence of five previously described gut bacteria (*Acetobacter* and *Lactobacillus* species) in individual flies of 21 strains (10 *Drosophila* species) were determined. Most bacteria were not present in all individuals of most strains, and bacterial species pairs co-occurred in individual flies less frequently than predicted by chance, contrary to expectations of a core microbiota. A complementary pyrosequencing analysis of 16S rRNA gene amplicons from the gut microbiota of 11 *Drosophila* species identified 209 bacterial operational taxonomic units (OTUs), with near-saturating sampling of sequences, but none of the OTUs was common to all host species. Furthermore, in both of two independent sets of *Drosophila* species, the gut bacterial community composition was not congruent with host phylogeny. The final analysis identified no common OTUs across three wild and four laboratory samples of *D. melanogaster*. Our results yielded no consistent evidence for a core microbiota in *Drosophila*. We conclude that the taxonomic composition of gut microbiota varies widely within and among *Drosophila* populations and species. This is reminiscent of the patterns of bacterial composition in guts of some other animals, including humans.

Introduction

The animal gut is a habitat for microorganisms, which are generally acquired orally with food. Nevertheless, the gut microbiota does not simply reflect the microorganisms in the food, but can be dominated by bacteria that are taxonomically distinct from bacteria in other environments (Ley *et al.*, 2008b; Tamames *et al.*, 2010; Chandler *et al.*, 2011). The distinctiveness of the gut microbiota can be attributed to the ecological conditions in the gut, including regions with extreme pH or redox potential, biologically active compounds (for example, digestive enzymes, immune effectors) and disturbance (for example, bulk flow of food, production of mucus or other extracellular secretions, epithelial cell turnover) (Karasov and Douglas, 2013). Furthermore, the gut is a living habitat, and coevolutionary interactions between the microbiota and the animal have been predicted, potentially resulting in the evolutionary divergence of gut-associated microorganisms from their free-living relatives, and codiversification of the microbiota and animal host (Dethlefsen *et al.*, 2007; Walter *et al.*, 2011). Sustained codiversification results incongruence between host phylogeny and composition of the gut microbiota.

A subset of the gut microbiota has been reported to be shared among host individuals within various animal species, including *Anopheles* mosquitoes, the honey bee *Apis mellifera*, zebrafish *Danio rerio* and the laboratory mouse (Mohr and Tebbe, 2006; Martinson *et al.*, 2011; Roeselers *et al.*, 2011; Wang *et al.*, 2011; Pedron *et al.*, 2012; Tang *et al.*, 2012). This subset has been described as the core microbiota (Hamady and Knight, 2009; Shade and Handelsman, 2012). Nevertheless, substantial temporal and among-individual variation in composition of the microbiota has been reported in some animals (Robinson *et al.*, 2010; Caporaso *et al.*, 2011; Lozupone *et al.*, 2012; The Human Microbiome Project Consortium, 2012), and it has been suggested that high variability in species composition may be characteristic of some microbial communities in animals and other habitats (Burke *et al.*, 2011).

The presence and abundance of microorganisms in a host can also be influenced by ecological

relationships among the gut microorganisms. The interactions may be antagonistic (competition ($-/-$), amensalism ($-/0$)) or positive (commensalism ($+/0$), mutualism ($+/+$)). Positive interactions would promote the persistence of a core microbiota, while negative interactions would reduce microbial co-occurrence, potentially leading to variation in microbiota composition among host individuals. Specific instances of competition, metabolite cross-feeding and other among-microbe interactions are known, (for example, Coyne *et al.,* 2005; Donohoe *et al.,* 2011; Rosenthal *et al.,* 2011), but the overall contribution of positive and negative interactions to the microbial community has rarely been considered. Exceptionally, Faust *et al.* (2012) found that most interactions in the human microbiota are negative, suggesting that processes such as competition and niche differentiation may be important determinants of community structure in this system.

The purpose of this study was to investigate whether drosophilid flies have a core set of gut-associated bacterial taxa. The gut microbiota in these insects has been reported to include Proteobacteria (especially Acetobacteraceae and Enterobacteriaceae) and Firmicutes of the order Lactobacillales (notably *Lactobacillus* and *Enterococcus* species). Despite regional variation in conditions (pH, redox potential and so on) in the gut (Shanbhag and Tripathi, 2009), bacteria occur in the crop, midgut and hindgut, with densities up to 10^6 cells per fly (Corby-Harris *et al.*, 2007; Cox and Gilmore, 2007; Ren *et al.*, 2007; Roh *et al.*, 2008; Sharon *et al.*, 2010; Chandler *et al.*, 2011; Storelli *et al.*, 2011; Wong *et al.*, 2011). Elimination of the gut bacteria can result in delayed larval development, altered lifespan and changes in nutrient allocation attributable to disruption in insect insulin signaling (Brummel *et al.*, 2004; Shin *et al.*, 2011; Storelli *et al.*, 2011; Ridley *et al.*, 2012). An important caveat to our understanding is whether the gut microbiota includes a common phylogenetic subset. Cox and Gilmore (2007) noted three taxa, *Acetobacter aceti*, *A. pasteurianus* and *Enterococcus faecalis*, in two laboratory strains and one wild population, but Corby-Harris *et al.* (2007) described 74 taxa that were 'unevenly spread' among wild populations of *D. melanogaster.* Chandler *et al.* (2011) reported that members of Enterobacteriaceae and Lactobacillales are very widely distributed, but apparently not universal, across 20 populations of multiple species. The shallow sampling available to the Sanger sequencing of the 16S rRNA gene used in these studies raises the possibility that some invariant taxa were undetected. This caveat can be addressed by high throughput sequencing as the bacterial communities in *D. melanogaster* are of low diversity, with saturation of rarefaction curves at $<20\,000$ pyrosequencing reads of 16S rRNA gene amplicons (Wong *et al.*, 2011).

In addition to the gut microbiota, some drosophilids possess bacteria, notably *Wolbachia* and *Spiroplasma*, which colonize internal organs, especially the gonads (Mateos *et al.*, 2006). These vertically transmitted bacteria can cause reproductive distortion, and confer protection against natural enemies (Hedges *et al.*, 2008; Werren *et al.*, 2008). They often have intermediate prevalence in populations and species, and do not contribute to the microbiota in the gut lumen (O'Neill *et al.*, 1997; Jaenike *et al.* 2010).

The specific aims of this study on the gut microbiota of drosophilid flies were twofold. First, we tested for a common subset of the gut microbiota by two complementary methods: taxon-specific PCR assays of bacteria previously shown to account for $>90\%$ of the bacteria in *D. melanogaster* (Wong *et al.*, 2011) and pyrosequencing of the total bacterial community. Second, we investigated two ecological patterns likely associated with a core microbiota: positive co-occurrence of different bacteria in individual flies and congruence between host phylogeny and bacterial community composition. Most experiments were conducted on flies in laboratory culture. This enabled us to use aseptically dissected guts (not feasible with field-collected flies), giving assurance that the bacteria scored were members of the gut microbiota. Supplementary whole-body analyses of field-collected *D. melanogaster* compared the microbiota in wild and laboratory flies of one species.

This first comprehensive analysis of the gut microbiota in multiple *Drosophila* species revealed that the composition of the gut microbiota is remarkably inconstant, and does not vary in concordance with host phylogeny. In this respect, we found no evidence of microbial taxa that are shared in all *Drosophila* hosts.

Materials and methods

Drosophila samples
Samples of adult *Drosophila* were derived from: 11 *Drosophila* species reared at Cornell University on Y-G diet (Brewer's yeast (MP Biomedicals, Santa Ana, CA, USA) and glucose (Sigma, St Louis, MO, USA) (both at $83\,\mathrm{g\,l^{-1}}$), agar ($10\,\mathrm{g\,l^{-1}}$ (Genesee Scientific, San Diego, CA, USA) and preservatives (0.04% phosphoric acid, 0.42% propionic acid (Sigma)); seven *Drosophila* species maintained at University of Rochester on Formula 4–24 (Carolina Biological Supply Company); and samples of *D. melanogaster* adults (mixed age and sex) collected from three USA sites and fixed immediately in 70% ethanol (Table 1 and Supplementary Table S1).

DNA isolation
Total genomic DNA was extracted from isolated adult fly guts or whole-bodies (age and sex varying with experiment, as below) by the method of Cenis *et al.* (1993). Guts from surface-sterilized flies were

Table 1 *Drosophila* species and number of bacterial OTUs identified from 454 sequence analysis of 16S rRNA gene amplicons

Species	Strain	Source	Number of OTUs[a]
Set-1: reared at Cornell University			
D. ananassae	DSSC no. 14024-0371.13	Hawaii	42
D. erecta	DSSC no. 14021-0224.01	Not known	17
D. melanogaster	Canton-S	Not known	41, 48[b]
D. persimilis	DSSC no. 14011-0111.42	Port Coquitlam, British Columbia, Canada	35
D. pseudoobscura	SD02	Not known	130
D. santomea	DSSC no. 1402-0271.01	San Tome and Principe Island	46
D. sechellia	DSSC no. 14021-0248.03	Cousin Island, Seychelles	62
D. simulans	DSSC no. 14021-0251.001	Georgetown, Guyana	36
D. virilis	DSSC no. 15010-1051.87	Not known	64
D. willistoni	DSSC no. 14030-0811.24	Guadaloupe Island, France	48
D. yakuba	DSSC no. 14021-0261.01	Liberia	62
Set-2: reared at Rochester University			
D. deflecta	15130-2018.00	Skunk cabbage, Princeton, NJ, USA	318
D. duncanii	NJ-1 2009	Fungi, NJ, USA	108
D. falleni	Pittsford 2010-1	Fungi, Pittsford, NY, USA	223
D. munda	SWRS 2005	Fungi, SW Research Station, Portal, AZ, USA	195
D. neotestacea	W + S + Pittsford 2007	Fungi, Pittsford, NY, USA	71
D. quinaria	Pittsford 2010-1	Skunk cabbage, Pittsford, NY, USA	204, 224[b]
D. suboccidentalis	Diamond Lake 2005-19	Fungi, Diamond Lake, OR, USA	194
Set-3: field-collected			
D. melanogaster	NY1	Decaying vegetation at Pittsford, NY (09/7/10), USA	178
D. melanogaster	NY2	Apples at Apple Farm, Victor, NY (09/7/10), USA	61
D. melanogaster	AZ	Bananas at SW Research Station, AZ (09/10/11), USA	110

Abbreviation: OTUs, operational taxonomic units, defined with pair-wise 97% sequence identity.
[a]Samples comprised dissected guts for set-1 and set-2, and whole-insect bodies for set-3.
[b]Two technical replicates were analysed, with both values displayed.

dissected in sterile Ringer's solution as previously described (Wong *et al.*, 2011). Samples were homogenized in 180 µl lysis buffer (20 mM Tris-HCl, pH 8.0, 2 mM sodium EDTA, 1.2% Triton-X 100, 20 mg ml^{-1} lysozyme) and incubated at 37 °C for 1.5 h, with brief bead-beating at 45 min in a Disruptor Genie using 0.1 mm glass beads (Scientific Industries, Bohemia, NY, USA). Twenty microlitres 10 × extraction buffer (2 M Tris-HCl, pH 8.5, 2.5 M NaCl, 0.25 M EDTA, 5% w/v SDS) and 10 µl proteinase K (20 mg ml^{-1}) were added, samples were incubated at 55 °C for 1 h and precipitated with 100 µl 3 M sodium acetate (pH 5.2). The supernatant was mixed with equal volume 100% ice-cold isopropanol and incubated at room temperature for 30 min before centrifugation for 30 min at 18 000 *g*. After discarding the supernatant, each pellet was washed in 500 µl 70% ice-cold ethanol, dried and resuspended in 20 µl sterile water.

End-point PCR assays of bacterial prevalence
L. brevis, L. fructivorans, L. plantarum, Acetobacter pomorum and *A. tropicalis* in the guts of individual flies were scored by end-point PCRs using taxon-specific 16S rRNA gene primers (Supplementary Table S2a). The experimental samples were five; 5–7-days-old and 4–5-weeks-old adults of both

sexes, run in parallel with positive controls comprising DNA from pure culture of the corresponding bacteria and sterile water as negative control. The PCR reactions were as in Wong *et al.* (2011). PCR products were separated by gel electrophoresis using 1% agarose gel and visualized with SYBRSafe (Invitrogen, Carlsbad, CA, USA). Sanger's sequencing confirmed the identity of representative bands.

Multiplex 454 pyrosequencing of 16S rRNA gene sequences
Each sample comprised 50 guts (laboratory-reared flies) or 10 bodies of *D. melanogaster* (laboratory strain ZH26, wild samples), with a drop of Ringer's solution treated as for dissections but without insect material as the negative control. The laboratory fly samples comprised approximately equal numbers of males and females, and were of similar age range within set-1 and ZH26 (5–10-days-old) and set-2 (a broad age distribution for every species); the wild flies were of unknown age. 16S rRNA amplicons of the V2 16S rRNA region were prepared as previously described (Wong *et al.*, 2011), with primers mentioned in Supplementary Table S2b. Equal amounts (ng) of three reaction products per sample were mixed and purified using the QIAquick PCR purification kit (QIAGEN, Valencia, CA, USA), followed by Pico-Green quantification. Emulsion PCR was

conducted at 1.5 copies per bead using only 'A' beads for unidirectional 454 GS-FLX pyrosequencing with standard Titanium chemistry.

Pyrosequencing flowgrams were converted to sequence reads using 454 Life Science software (www.454.com). Reads with ambiguous nucleotides (N) and <270 nucleotides after the forward primer, and mismatches with the 16S rRNA gene primers were excluded in the initial filtering. Multiplexed samples from two half-plate runs were combined before downstream analyses by modifying the barcodes in the fna files and concatenating the two fna and qual files into a single fna and qual file, respectively. The QIIME 1.4.0 virtualbox package was used to split the multiplexed sequences, discard chimeras, denoise the data, bin sequences at 97% sequence identity and make taxonomy calls to genus level (Caporaso *et al.*, 2010). Default parameters were used except that the denoising cutoff was set to retain doubletons, and the RDP classifier was applied using a custom Greengenes database to assign class through genus designations. Species identity of each operational taxonomic units (OTU) was assigned by local BLAST (Stand-alone MEGA-BLAST program) with the 16S Microbial database (June 2012). OTUs with either single reads or fewer reads than in the negative controls were excluded. For comparison, OTU tables were generated in Pyrotagger (http://pyrotagger.jgi-psf.org/release). Reads assigned to *Wolbachia* were excluded because, first, this bacterium is not a member of the gut microbiota (it has weak tropism for the gut, and does not inhabit the gut lumen); and, second, the *D. ananassae* genome includes laterally acquired *Wolbachia* sequences (Dunning Hotopp *et al.*, 2007), such that *Wolbachia* reads are a measure of host DNA in the gut samples (*D. ananassae* accounted for 90% of *Wolbachia* reads across all gut samples assayed). For consistency, *Wolbachia* reads were removed from data sets for whole-body samples of wild flies. Reads assigned to *Wolbachia* are shown as 'excluded sequences' in Supplementary Table S3, and the minimal effect of their exclusion on our analysis is indicated by PCA plots in Supplementary Figure S1. The samples included technical replicates for two *Drosophila* species (*D. melanogaster* in set-1, *D. quinaria* in set-2).

PCA plots of the bacterial communities were created using pcaMethods (Stacklies *et al.*, 2007) in R (R Development Core Team, 2012), following log-transformation of number of reads per OTU. Correlation matrices derived from the OTU tables were used to create dendrograms of the bacterial communities using pvclust (CRAN.R-project.org/package = pvclust) and ape (Paradis *et al.*, 2004) in R, and compared with *Drosophila* phylogenetic trees built in BioEdit from a clustalX alignment of concatenated DNA sequences obtained from NCBI. Trees were manipulated in FigTree v1.3.1. Graphical taxonomy networks were created using the make_otu_network.py QIIME script and visualized as an unweighted forced-directed layout with Cytoscape v2.8.2 (Smoot *et al.*, 2011) using default QIIME instructions. The analyses shown were conducted with the full microbiota; the patterns were equivalent when OTUs representing <1% or <0.1% of reads were excluded (data not shown).

Bacterial co-occurrence analyses
The likelihood of co-occurrence of bacterial species in individual flies was analysed by *C*-score test (Stone and Roberts, 1990) using data obtained by PCR with taxon-specific primers for each fly. *C*-score calculates the mean number of instances where two bacterial species co-occur, across all fly species pairs. The computed *C*-score is significantly greater than the null distribution if the bacteria co-occur less frequently than predicted by chance, and less than the null distribution for positive co-occurrence. The prevalence data sets were arranged in presence-absence matrices with the five bacterial species as rows and individual flies as columns. The most appropriate null model for these data, in which the presence/absence of each bacterial species in each fly is known, is the 'fixed-fixed' null model (SIM9 of Gotelli, 2000). The observed data matrices were compared with 5000 randomly generated matrices using EcoSim 7.72 (Gotelli and Entsminger, 2012).

Results

Prevalence of bacteria in laboratory Drosophila *populations*
Our first approach to investigate the taxonomic composition of the gut microbiota in *Drosophila* was to score for five bacterial taxa in individual flies of 21 strains in 10 *Drosophila* species (Figure 1a). The five bacteria have previously been shown to account for >90% of the bacteria in multi-individual samples *D. melanogaster* strain Canton-S in our laboratory (Wong *et al.*, 2011). No bacterial taxon was detected in every individual of every fly strain. One bacterium, *L. fructivorans*, was detected in at least one fly of every *Drosophila* strain; *A. pomorum, A. tropicalis* and *L. plantarum* were detected in every strain except *D. melanogaster* ZH26 (strain-4 in Figure 1a) and *L. brevis* was detected in 13 (62%) of the strains. Overall, the frequency of each bacterium did not vary significantly with age (5–7-days-old versus 4–5-weeks-old) or sex ($P > 0.05$), but the frequency of *A. pomorum*, *L. brevis* and *L. plantarum* varied significantly among strains ($P < 0.001$).

In a few of the 21 *Drosophila* strains, every individual scored positive for a bacterial taxon: nine (43%) strains for *A. tropicalis*, five (24%) for *L. fructivorans*, four (19%) for *A. pomorum*, one (5%) for *L. plantarum* and none for *L. brevis*. Furthermore, each of the bacteria was at intermediate prevalence (that is, at least one fly scored positive and one fly scored negative) in more than half of the 21 strains
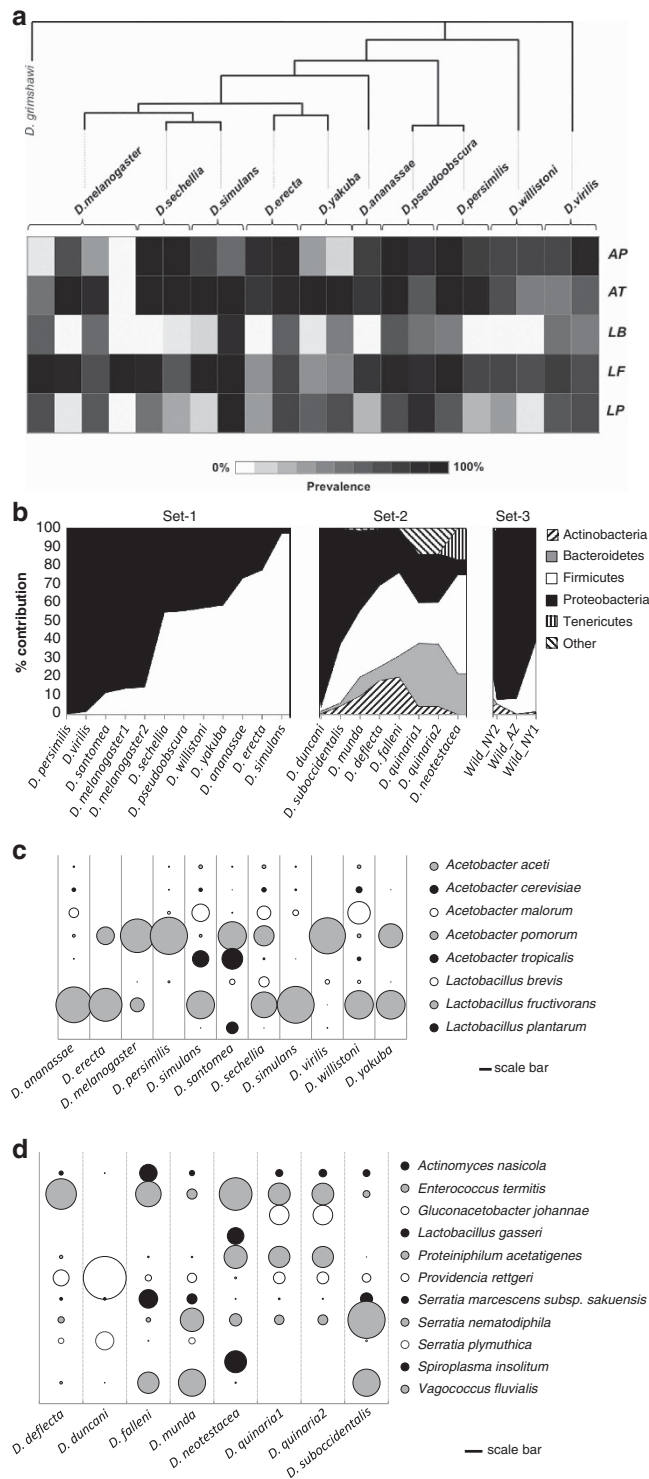
Figure 1 Analysis of the composition of bacterial communities in *Drosophila* species. (**a**) Prevalence of 5 bacterial taxa in 21 *Drosophila* strains (*Drosophila* strain details provided in Supplementary Table S1). (**b**) Abundance of bacterial phyla in pyrosequence analyses. (**c**) Abundance of dominant species in *Drosophila* species set-1 based on 97% similarity OTU assignments. (**d**) Abundance of dominant species in *Drosophila* species set-2 based on 97% similarity OTU assignments.

(ranging from 52% for *A. tropicalis* to 90% for *L. plantarum*). By the criterion of diagnostic PCR assay, most of the five bacteria are not members of the core microbiota in most of the *Drosophila* strains, and none was core to every strain.

To investigate the pattern of occurrence of the five bacteria across the individual flies, the data set was analysed by *C*-scores. The *C*-score for the full data set, 4114.5, was significantly higher than expected by chance ($P < 0.001$), indicating that the bacterial

Table 2 Temporal variation in frequency of bacterial taxa in individual adult *D. melanogaster*

Taxon-specific PCR assay	Strain ZH26 (10 flies per sample)			X^2 (2 df)	Strain Canton-S (10 flies per sample)			X^2 (2 df)
	Number of flies positive for bacterium				Number of flies positive for bacterium			
	Feb 2011 (time-0)	Nov 2011 (9 months)	Apr 2012 (12 months)		May 2010 (time-0)	Feb 2011 (9 months)	Feb 2012 (21 months)	
A. pomorum	0	10	10	10, $P = 0.007$[a]	10	10	6	1.23, $P = 0.54$
A. tropicalis	0	10	10	10, $P = 0.007$[a]	10	4	10	3, $P = 0.223$
L. brevis	0	10	0	20, $P < 0.001$[a]	10	5	0	10, $P = 0.007$[a]
L. fructivorans	10	10	10	0, $P = 1$	9	10	9	0.07, $P = 0.966$
L. plantarum	0	10	0	20, $P < 0.001$[a]	10	9	0	10, $P = 0.007$[a]

[a]Significant after Bonferroni correction for five samples (critical probability = 0.01).

species co-occurred less often than in random distribution. Significantly elevated *C*-scores were also obtained for young males ($P < 0.001$) and females ($P < 0.002$), and old males ($P < 0.01$), indicative of segregation among the bacteria in these samples. The *C*-score for old females was not significant ($P > 0.05$). In general, significant scores were associated with negative relationships between *L. fructivorans* and *Acetobacter* species. The observed segregation among these bacteria would tend to hinder the assembly of a core microbiota.

One *D. melanogaster* strain, ZH26, was unique; in that every fly was colonized with only one of the five tested taxa: *L. fructivorans* (Figure 1a). In a complementary 454 analysis (Supplementary Table S3a), *L. fructivorans* accounted for >99% of the 55 683 reads, confirming the PCR data and indicating that strain ZH26 does not bear a highly divergent bacterial community. This colonization status was not consistent across fly generations: when the five taxon-specific PCR assays were repeated on the same stock of ZH26 9 months later, all five bacteria were universally present; but *L. brevis* and *L. plantarum* were absent after a further 3 months (Table 2, Supplementary Figure S2a).

To assess whether variability in the composition of the gut bacteria was unique to ZH26, we determined the prevalence of the five dominant gut taxa in *D. melanogaster* strain Canton-S (in which the five taxa were originally identified (Wong *et al.*, 2011)). All five bacteria were detected, but none was universally present, in the three samples of 10 flies analysed over 21 months. The prevalence of every bacterium shifted between the three sampling periods, and *L. brevis* and *L. plantarum* varied between being present in all and none of the 10 flies tested (Table 2, Supplementary Figure S2b). We conclude that variation in bacterial prevalence is not unique to strain ZH26.

These results indicate that the five bacteria previously identified as major constituents of the gut microbiota under our laboratory rearing conditions are not universally present in all individual flies, and they vary in prevalence across generations.

Pyrosequencing of bacterial communities in Drosophila
As an alternative approach to investigate the bacterial communities in *Drosophila* guts, we quantified the total gut microbiota by pyrosequencing 16S rRNA gene amplicons from three independent sets of drosophilid flies (Table 1). In total, 26 811–62 138 reads of 16S rRNA gene amplicons per sample were identified in QIIME, after quality filtering and removal of chimeras and single reads (Supplementary Table S3b–d). All the rarefaction curves tended to saturation (Supplementary Figure S3), indicating that the OTUs were representative of the total bacterial community in each sample. Close correspondence in the number and identity of the OTUs between two technical replicate samples (samples of the same genomic DNA) were obtained for both *D. melanogaster* in set-1 (Pearson's correlation coefficient $r = 0.998$, $P < 0.0001$) and *D. quinaria* in set-2 ($r = 0.959$, $P < 0.0001$) (Supplementary Figure S4), showing that random sampling effects, which have constrained the reproducibility of pyrosequencing data in certain complex bacterial communities (Zhou *et al.*, 2011), were not significant in this study. The combined data for set-1 and set-2 (Supplementary Table S3e) were also processed by Pyrotagger, an alternative program used in our previous research on the gut microbiota of *D. melanogaster* (Wong *et al.*, 2011). The correlation between the outputs of QIIME and Pyrotagger was highly significant for numbers of reads (Pearson's correlation coefficient, $r = 0.988$, $P < 0.001$) and OTUs ($r = 0.972$, $P < 0.001$), although, on average, 18% fewer reads were obtained by Pyrotagger than QIIME (Supplementary Table S4).

All 16S rRNA gene amplicon reads in set-1 could be assigned to two phyla: Proteobacteria and Firmicutes (Figure 1b). Two genera, *Lactobacillus* (Firmicutes) and *Acetobacter* (α-Proteobacteria) accounted for 94–100% of the reads in every sample (Supplementary Table S3b). The most abundant bacterium in every *Drosophila* species was either *L. fructivorans* OTU179 or *A. pomorum* OTU630, which accounted for up to 63% and 82%, respectively, of all 16S reads per sample (Figure 1c). Nonetheless, none

of the 209 OTUs or 124 bacterial species were present in every *Drosophila* species (Supplementary Table S3b). We conclude that no bacterial taxon at the level of OTU or species is present at detectable levels in all 11 *Drosophila* species.

The data for set-1 were investigated by PCA. Phylogenetically related *Drosophila* species were not clustered by the first two axes, which together accounted for 73% of the variance (Figure 2a), or any other axis combination tested. The implication that the bacterial communities were not patterned according to host phylogeny was confirmed by the poor correspondence between the phylogenetic relationship among the 11 *Drosophila* species and the relatedness of host-associated gut bacterial community taxonomic composition (Figure 2b). Furthermore, the bacterial communities could not be differentiated between fly samples possessing and lacking *Wolbachia* (Supplementary Figure S1).

As an independent test for the relationship between bacterial community composition and host phylogeny, we investigated the bacterial community in guts dissected from *Drosophila* species of set-2 (Table 1). These bacteria included representatives of seven phyla (Figure 1b, Supplementary Table S3c) and were dominated by *Enterococcus termitis* OTU659 and *Vagococcus fluvialis* OTU4 in the Firmicutes (Lactobacillales), and *Providencia rettgeri* OTU937 and *Serratia nematodiphila* OTU3 in the γ-Proteobacteria (Figure 1d). Ten (1%) of the 997 OTUs were detected in all seven *Drosophila* species (Supplementary Table S3c), accounting in total for 1–70% of the reads (median 9%), but the prevalence of these OTUs among the individual flies contributing to each samples (that included both sexes and a broad age range) is unknown. As with set-1, the relationship among bacterial communities did not map onto the phylogeny of their *Drosophila* hosts (Figures 2c and d).

Our final analysis tested for bacterial OTUs or species shared across field-collected and laboratory samples of a single *Drosophila* species, *D. melanogaster*. The three field-collected samples included representatives of three bacterial phyla: Firmicutes, Proteobacteria and at <5%, Actinobacteria (Figure 1b). The dominant Firmicutes included *Leuconostoc mesenteroides* OTU5 and *Lactococcus lactis* OTU121, and the abundant Proteobacteria were Acetobacteraceae, specifically *Gluconobacter japonicus* OTU4 and *Gluconobacter albidus* OTU6 and the γ-proteobacterium *Tatumella ptyseos* OTU1 (Supplementary Table S3d).

The bacterial communities in the three wild samples were compared with five data sets for laboratory cultures of *D. melanogaster*. No OTU or species was detected in each sample (Supplementary Table S3f), offering no support for bacterial taxa universally present in the guts of *D. melanogaster*. The wild samples grouped together closely on the first two axes of the PCA, and were separated from the laboratory samples (Figure 2e). The difference between wild and laboratory samples

and the greater variability among laboratory samples are confirmed by the bipartite graph, in which the edges connect each host sample node to every bacterial OTU in that sample (Figure 2f).

Discussion

Immigration with food and emigration with feces are important processes shaping the microbial community in animal guts, including variation in community composition among host individuals and over time within one host. Despite this continual flux of microorganisms through the gut habitat, a subset of the microorganisms is consistently recovered from certain animal taxa. This subset, sometimes described as the 'core microbiota', is of special interest because it is predicted to be ecologically-distinctive and may have coevolved with the host.

The concept of the core microbiota has been applied in multiple ways. In some studies, specific bacterial taxa has been detected in all samples, each of which comprised multiple hosts, but the prevalence of the bacteria in each individual was not tested (Mohr and Tebbe, 2006; Martinson *et al.*, 2011; Roeselers *et al.*, 2011; Wang *et al.*, 2011). Other investigations have tested individual hosts, often with study-specific criteria for a core, for example, relaxation of the detected prevalence of the bacteria to 80 or 50% of hosts, or use of variable or low (<97%) OTU-call cutoffs (Qin *et al.*, 2010; Boissiere *et al.*, 2012; Moran *et al.*, 2012; Nelson *et al.*, 2012; Salonen *et al.*, 2012). Such relaxation can be justified for technical reasons, including the artifactual inflation of community diversity from contamination, error in sequencing and sequence alignment and incomplete sampling, especially for highly diverse bacterial communities (Huse *et al.*, 2010; Kunin *et al.*, 2010; Sun *et al.*, 2012; Wylie *et al.*, 2012). Nevertheless, the variation in criteria adopted across different studies and animal systems hinders systematic analysis of the degree of partner fidelity between animals and their gut microbiota.

Despite these general difficulties, consistent patterns in the taxonomic diversity of the gut microbiota in drosophilid flies are emerging. The bacterial communities are very predictable at high-phylogenetic levels, dominated by one, two or all of the order Lactobacillales (phylum Firmicutes) and the families Enterobacteriaceae and Acetobacteraceae (phylum Proteobacteria) (References in Introduction); but they vary irregularly at the level of genus, species and OTU. This inconstancy is evident at multiple phylogenetic scales of the host, among-species, within-species and even within single laboratory lines, and no OTU was detected in every sample analysed in this study. Although technical artifacts can inflate among-sample differences (see above), the severity of these limitations is much ameliorated in this study of the *Drosophila* system by the use of whole-gut samples and the near-saturation of sequence reads.
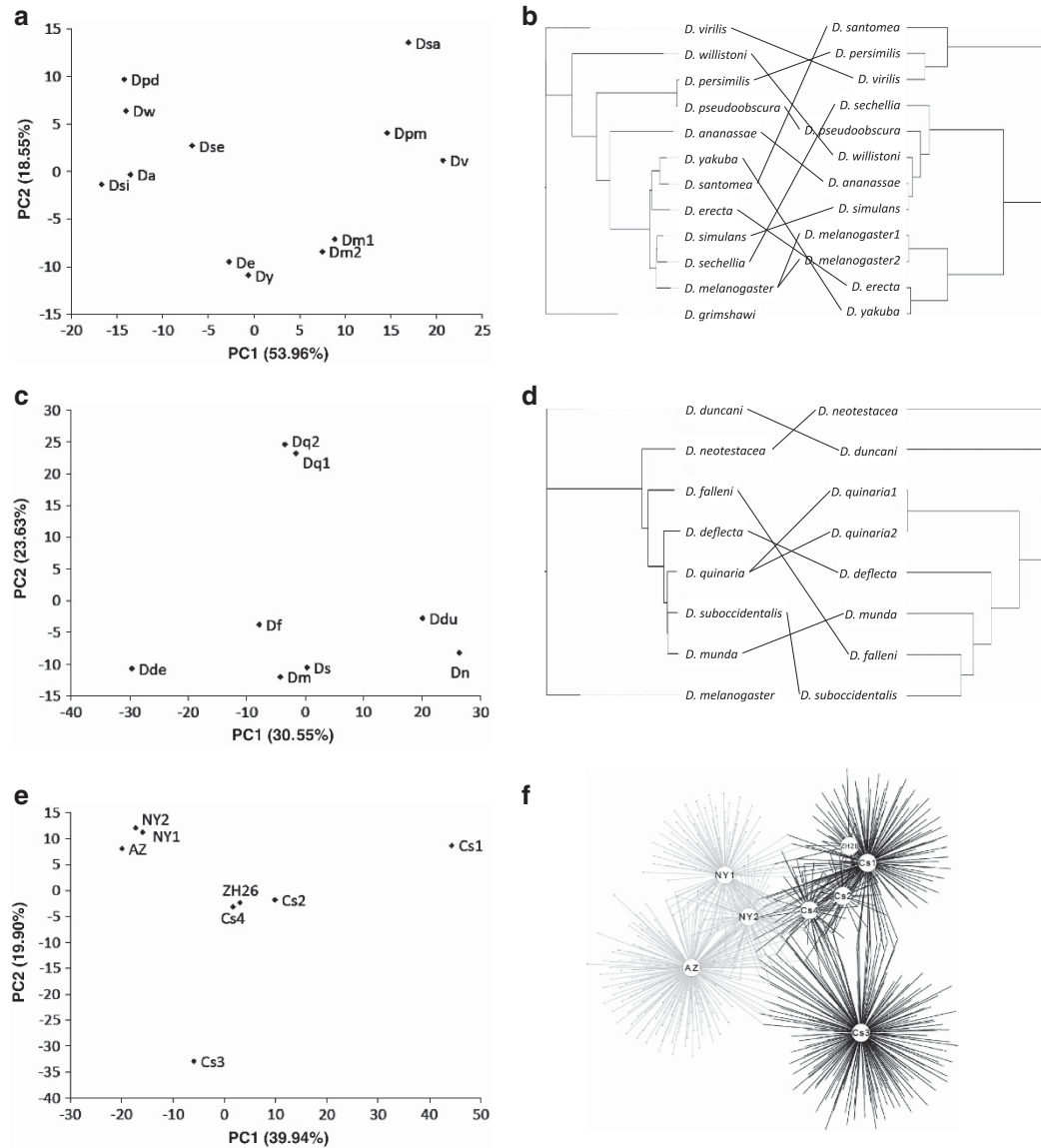
Figure 2 Relationship between bacterial community composition and *Drosophila* species based on 97% similarity OTU assignments. (**a**) Principal components analysis (PCA) of the bacterial community and (**b**) correspondence between dendrograms of bacterial communities and the phylogeny of *Drosophila* in set-1. (**c**) PCA and (**d**) dendrogram correspondence in *Drosophila* set-2. (**e**) PCA and (**f**) bipartite graph of *D. melanogaster* from wild (AZ, NY1, NY2: see Table 1c) and laboratory (CS1-4 (Canton-S isolates) and ZH26: see legend to Supplementary Table S3f). (Abbreviations in (**a**) and (**c**) indicate species name, as provided in (**b**) and (**d**), respectively, for example, Dq is *D. quinaria*, Dde is *D. deflecta*).

The composition of the gut microbiota can also be affected strongly by rearing conditions. In particular, repeated environmental perturbations (including variation in food consumed) in the field may prevent the realization of the full core microbiota in some individuals, while laboratory-reared animals may not have access to key members of the core microbiota occurring in the natural habitat. In this study, individual microbial taxa were not generally found to be shared universally, either within or among drosophilid species in laboratory conditions. In particular, the data do not substantiate the common bacterial taxa found by Cox and Gilmore (2007) across laboratory and field conditions for one species, *D. melanogaster*. Our results complement and extend the research of Chandler *et al.* (2011), in which shallow sampling with Sanger sequencing failed to yield a common subset of bacterial OTUs among field samples of multiple *Drosophila* species.

The incongruence between drosophilid phylogeny and bacterial community composition suggests weak partner fidelity, and that a consistent microbiota does not operate across evolutionary timescales in this system. The implication is that, in terms of taxonomic composition, the gut microbiota in *Drosophila* has neither coevolved with the host over evolutionary time, nor tracked evolutionary changes in gut physiology that may vary according to

phylogenetic relatedness between different host taxa. In this respect, *Drosophila* appears to parallel mammals, for which no phylogenetic pattern in the composition of the gut microbiota has been found (Ley *et al.*, 2008a; Muegge *et al.*, 2011). Our results differ from the evidence for congruence between host phylogeny and gut microbiota composition obtained, for example, for bacterial community composition in laboratory cultures of jewel wasps *Nasonia* (Brucker and Bordenstein, 2012), and wild populations of both great apes/humans (Ochman *et al.*, 2010) and termites (Hongoh *et al.*, 2005); and the genotypes of one bacterial species, *Lactobacillus reuteri*, in studies that included inbred lab mice and rats (Oh *et al.*, 2010; Frese *et al.*, 2011). An important issue for future work is the ecological factors that dictate the variation in the congruence of host-microbiota phylogenies across different animal groups.

The inconstancy in *Drosophila* gut microbiota composition raises two broad issues: the population processes that dictate whether a microbial community includes a consistent subset, and how taxonomic composition influences the functional traits of the bacterial community. A consistent subset of the microbiota would be promoted by high rates of transmission between conspecific hosts (including parent-to-offspring and among kin), and by extended residence time within individual hosts. As gut microorganisms are routinely shed in feces, residence time of an ingested microorganism and its descendants is shaped by the relationship between the rates of proliferation and emigration (Savage, 1977; Costello *et al.*, 2012). Variation in these key ecological parameters among different animal groups has yet to be compared systematically. The second issue, the relationship between taxonomic and functional diversity of animal-associated bacterial communities, has been studied in mammalian gut associations, with evidence that taxonomically diverse bacterial communities can be functionally equivalent, for example, that a taxonomically-variable microbiota can potentially support a core microbiome (Turnbaugh *et al.*, 2008; Muegge *et al.*, 2011; The Human Microbiome Project Consortium, 2012; Morgan *et al.*, 2013). Research to date on *Drosophila* has focused on the relationship between taxonomic composition of the bacteria and host phenotype. There is persuasive evidence that individual members of the gut microbiota vary in their impact on the phenotype of *D. melanogaster* (Shin *et al.*, 2011; Storelli *et al.*, 2011), but the effects of natural variation among bacteria on the phenotype and fitness of *Drosophila* in laboratory culture and field remain to be studied.

Relevant to these considerations, multiple aspects of insect function can be altered by experimental elimination of the gut microbiota, including intestinal cell proliferation, nutrient content, metabolic rate, insulin signaling, larval developmental rates and lifespan (Brummel *et al.*, 2004; Buchon *et al.*, 2009; Shin *et al.*, 2011; Storelli *et al.*, 2011; Ridley

et al., 2012). These data suggest *Drosophila* is adapted to the presence of microorganisms in the gut, even though the taxonomic composition of the microbiota is variable and partner fidelity is weak at timescales ranging from a few generations in a single laboratory culture to millions of years of *Drosophila* evolution. As noted above, this may be reminiscent of evolution with a core microbiome rather than a core microbiota. Furthermore, many drosophilid flies are additionally associated with vertically-transmitted bacteria (especially *Wolbachia* and *Spiroplasma*) localized to the reproductive organs and other internal tissues. It is an open question whether these bacteria may influence the composition and function of the spatially distinct gut microbiota. A full understanding of the ecology of the inconstant gut microbiota of drosophilids will require further research on the interaction of host traits with the composition and activities of the bacterial taxa.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgements

References

Boissiere A, Tchioffo MT, Bachar D, Abate L, Marie A, Nsango SE *et al.* (2012). Midgut microbiota of the malaria mosquito vector *Anopheles gambiae* and interactions with *Plasmodium falciparum* infection. *PLoS Pathog* **8**: e1002742.

Brucker RM, Bordenstein SR. (2012). The roles of host evolutionary relationships (genus: *Nasonia*) and development in structuring microbial communities. *Evolution* **66**: 349–362.

Brummel T, Ching A, Seroude L, Simon AF, Benzer S. (2004). *Drosophila* lifespan enhancement by exogenous bacteria. *Proc Natl Acad Sci USA* **101**: 12974–12979.

Buchon N, Broderick NA, Chakrabarti S, Lemaitre B. (2009). Invasive and indigenous microbiota impact intestinal stem cell activity through multiple pathways in *Drosophila*. *Genes Dev* **23**: 2333–2344.

Burke C, Thomas T, Lewis M, Steinberg P, Kjelleberg S. (2011). Composition, uniqueness and variability of the epiphytic bacterial community of the green alga *Ulva australis*. *ISME J* **5**: 590–600.

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK *et al.* (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **7**: 335–336.

Caporaso JG, Lauber CL, Costello EK, Berg-Lyons D, Gonzalez A, Stombaugh J *et al.* (2011). Moving pictures of the human microbiome. *Genome Biol* **12**: R50.

Cenis JL, Perez P, Fereres A. (1993). Identification of aphid (Homoptera, Aphididae) species and clones by random amplified polymorphic DNA. *Ann Entomol Soc Amer* **86**: 545–550.

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.

Corby-Harris V, Pontaroli AC, Shimkets LJ, Bennetzen JL, Habel KE, Promislow DE. (2007). Geographical distribution and diversity of bacteria associated with natural populations of *Drosophila melanogaster. Appl Environ Microbiol* **73**: 3470–3479.

Costello EK, Stagaman K, Dethlefsen L, Bohannan BJ, Relman DA. (2012). The application of ecological theory toward an understanding of the human microbiome. *Science* **836**: 1255–1262.

Cox CR, Gilmore MS. (2007). Native microbial colonization of *Drosophila melanogaster* and its use as a model of *Enterococcusfaecalis* pathogenesis. *Infect Immun* **75**: 1565–1576.

Coyne MJ, Reinap B, Lee MM, Comstock LE. (2005). Human symbionts use a host-like pathway for surface fucosylation. *Science* **307**: 1778–1781.

Dethlefsen L, McFall-Ngai M, Relman DA. (2007). An ecological and evolutionary perspective on human-microbe mutualism and disease. *Nature* **449**: 811–818.

Donohoe DR, Garge N, Zhang X, Sun W, O'Connell TM, Bunger MK *et al.* (2011). The microbiome and butyrate regulate energy metabolism and autophagy in the mammalian colon. *Cell Metab* **13**: 517–526.

Dunning Hotopp JC, Clark ME, Oliveira DC, Foster JM, Fischer P, Munoz Torres MC *et al.* (2007). Widespread lateral gene transfer from intracellular bacteria to multicellular eukaryotes. *Science* **317**: 1753–1756.

Faust K, Sathirapongsasuti JF, Izard J, Segata N, Gevers D, Raes J *et al.* (2012). Microbial co-occurrence relationships in the human microbiome. *PLoS Comput Biol* **8**: e1002606.

Frese SA, Benson AK, Tannock GW, Loach DM, Kim J, Zhang M *et al.* (2011). The evolution of host specialization in the vertebrate gut symbiont. *PLoS Genet* **7**: e1001314.

Gotelli NJ, Entsminger GL. (2012). EcoSim 7.72.Acquired Intelligence, Inc. http://www.uvm.edu/~ngotelli/EcoSim/EcoSim.html.

Gotelli NJ. (2000). Null model analysis of species co-occurrence patterns. *Ecology* **81**: 2606–2621.

Hamady M, Knight R. (2009). Microbial community profiling for human microbiome projects: Tools, techniques, and challenges. *Genome Res* **19**: 1141–1152.

Hedges LM, Brownlie JC, O'Neill SL, Johnson KN. (2008). *Wolbachia* and virus protection in insects. *Science* **322**: 702.

Hongoh Y, Deevong P, Inoue T, Moriya S, Trakulnaleamsai S, Ohkuma M *et al.* (2005). Intra- and interspecific comparisons of bacterial diversity and community structure support coevolution of gut microbiota and termite host. *Appl Environ Microbiol* **71**: 6590–6599.

Huse AM, Welch DM, Morrison HG, Sogin ML. (2010). Ironing out the wrinkles in the rare biosphere through improved OTU clustering. *Environ Microbiol* **12**: 1889–1898.

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.

Karasov WH, Douglas AE. (2013). Gastrointestinal physiology. In: JW H, Garland T, Wang T (eds.) *APS Handbook of Physiology: Comparative and Evolutionary Physiology*. Wiley-Blackwell: Hoboken, NJ.

Kunin V, Engelbrektson A, Ochman H, Hugenholz P. (2010). Wrinkles in the rare biosphere: pyrosequencing errors can lead to artificial inflation of diversity estimates. *Environ Microbiol* **12**: 118–123.

Ley RE, Hamady M, Lozupone C, Turnbaugh PJ, Ramey RR, Bircher JS *et al.* (2008a). Evolution of mammals and their gut microbes. *Science* **320**: 1647–1651.

Ley RE, Lozupone CA, Hamady M, Knight R, Gordon JI. (2008b). Worlds within worlds: evolution of the vertebrate gut microbiota. *Nat Rev Microbiol* **6**: 776–788.

Lozupone CA, Stombaugh JI, Gordon JI, Jansson JK, Knight R. (2012). Diversity, stability and resilience of the human gut microbiota. *Nature* **489**: 220–230.

Martinson VG, Danforth BN, Minckley RL, Rueppell O, Tingek S, Moran NA. (2011). A simple and distinctive microbiota associated with honey bees and bumble bees. *Mol Ecol* **20**: 619–628.

Mateos M, Castrezana SJ, Nankivell BJ, Estes AM, Markow TA, Moran NA. (2006). Heritable endosymbionts of *Drosophila. Genetics* **174**: 363–376.

Mohr KI, Tebbe CC. (2006). Diversity and phylotype consistency of bacteria in the guts of three bee species (Apoidea) at an oilseed rape field. *Environ Microbiol* **8**: 258–272.

Moran NA, Hansen AK, Powell JE, Sabree ZL. (2012). Distinctive gut microbiota of honey bees assessed using deep sampling from individual worker bees. *PLoS One* **7**: e36393.

Morgan XC, Segata N, Hutttenhower C. (2013). Biodiversity and functional genomics in the human microbiome. *Trends Genet* **29**: 51–58.

Muegge BD, Kuczynski J, Knights D, Clemente JC, Gonzalez A, Fontana L *et al.* (2011). Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. *Science* **332**: 970–974.

Nelson TM, Rogers TL, Carlini AR, Brown MV. (2012). Diet and phylogeny shape the gut microbiota of Antarctic seals: a comparison of wild and captive animals. *Environ Microbiol* **15**: 1132–1145.

Ochman H, Worobey M, Kuo CH, Ndjango JB, Peeters M, Hahn BH *et al.* (2010). Evolutionary relationships of wild hominids recapitulated by gut microbial communities. *PLoS Biol* **8**: e1000546.

Oh PL, Benson AK, Peterson DA, Patil PB, Moriyama EN, Roos S *et al.* (2010). Diversification of the gut symbiont *Lactobacillusreuteri* as a result of host-driven evolution. *ISME J* **4**: 377–387.

O'Neill SL, Hoffmann AA, Werren JH (eds). (1997). *Influential Passengers—Inherited Microorganisms and Arthropod Reproduction*. Oxford University Press: Oxford, UK.

Paradis E, Claude J, Strimmer K. (2004). APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* **20**: 289–290.

Pedron T, Mulet C, Dauga C, Frangeul L, Chervaux C, Grompone G *et al.* (2012). A crypt-specific core microbiota resides in the mouse colon. *MBio* **3**: e00116–12.

1932

Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C *et al.* (2010). A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **464**: 59–65.

R Development Core Team (2012). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing: Vienna, Austria, ISBN 3-900051-07-0.

Ren C, Webster P, Finkel SE, Tower J. (2007). Increased internal and external bacterial load during *Drosophila* aging without life-span trade-off. *Cell Metab* **6**: 144–152.

Ridley EV, Wong AC, Westmiller S, Douglas AE. (2012). Impact of the resident microbiota on the nutritional phenotype of *Drosophila melanogaster. PLoS One* **7**: e36765.

Robinson CJ, Schloss P, Ramos Y, Raffa K, Handelsman J. (2010). Robustness of the bacterial community in the cabbage white butterfly larval midgut. *Microb Ecol* **59**: 199–211.

Roeselers G, Mittge EK, Stephens WZ, Parichy DM, Cavanaugh CM, Guillemin K *et al.* (2011). Evidence for a core gut microbiota in the zebrafish. *ISME J* **5**: 1595–1608.

Roh S, Nam Y, Chang H, Kim K, Kim M, Ryu JH *et al.* (2008). Phylogenetic characterization of two novel commensal bacteria related to innate immune homeostasis in *Drosophila. Appl Environ Microbiol* **74**: 6171–6177.

Rosenthal AZ, Matson EG, Eldar A, Leadbetter JR. (2011). RNA-seq reveals cooperative metabolic interactions between two termite-gut spirochete species in co-culture. *ISME J* **5**: 1133–1142.

Salonen A, Salojarvi J, Lahti L, de Vos WM. (2012). The adult intestinal core microbiota is determined by analysis depth and health status. *Clin Microbiol Infect* **18**(Suppl 4): 16–20.

Savage DC. (1977). Microbial ecology of the gastrointestinal tract. *Annu Rev Microbiol* **31**: 107–133.

Shade A, Handelsman J. (2012). Beyond the Venn diagram: the hunt for a core microbiome. *Environ Microbiol* **14**: 4–12.

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.

Sharon G, Segal D, Ringo JM, Hefetz A, Zilber-Rosenberg I, Rosenberg E. (2010). Commensal bacteria play a role in mating preference of *Drosophila melanogaster. Proc Natl Acad Sci USA* **107**: 20051–20056.

Shin SC, Kim SH, You H, Kim B, Kim AC, Lee KA *et al.* (2011). *Drosophila* microbiome modulates host developmental and metabolic homeostasis via insulin signaling. *Science* **334**: 670–674.

Smoot ME, Ono K, Ruscheinski J, Wang PL, Ideker T. (2011). Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* **27**: 431–432.

Stacklies W, Redestig H, Schol M, Walthe D, Selbig J. (2007). pcaMethods—a Bioconductor package providing PCA methods for incomplete data. *Bioinformatics* **23**: 1164–1167.

Stone L, Roberts A. (1990). The checkerboard score and species distributions. *Oecologia* **85**: 74–79.

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 Metab* **14**: 403–414.

Sun Y, Cal Y, Huse SM, Knight R, Farmerie WG, Wang X *et al.* (2012). A large-scale benchmark study of existing algorithms for taxonomy-independent microbial community analysis. *Brief Bioinform* **13**: 107–121.

Tamames J, Abellan JJ, Pignatelli M, Camacho A, Moya A. (2010). Environmental distribution of prokaryotic taxa. *BMC Microbiol* **10**: 85.

Tang X, Freitak D, Vogel H, Ping L, Shao Y, Cordero EA *et al.* (2012). Compexity and variability of gut commensal microbiota in polyphagous lepidopteran larvae. *PLoS One* **7**: 336978.

The Human Microbiome Project Consortium (2012). Structure, function and diversity of the healthy human microbiome. *Nature* **486**: 207–214.

Turnbaugh PJ, Hamady M, Yatsunenko T, Cantarel BL, Duncan A, Ley RE *et al.* (2008). A core gut microbiome in obese and lean twins. *Nature* **457**: 480–484.

Walter J, Britton RA, Roos S. (2011). Host-microbial symbiosis in the vertebrate gastrointestinal tract and the *Lactobacillus reuteri* paradigm. *Proc Natl Acad Sci USA* **108**(Suppl 1): 4645–4652.

Wang Y, Gilbreath TM, Kukutla P, Yan G, Xu J. (2011). Dynamic gut microbiome across life history of the malaria mosquito *Anopheles gambiae* in Kenya. *PLoS One* **6**: e24767.

Werren JH, Baldo L, Clark ME. (2008). *Wolbachia*: master manipulators of invertebrate biology. *Nat Rev Microbiol* **6**: 741–751.

Wong CN, Ng P, Douglas AE.. (2011). Low-diversity bacterial community in the gut of the fruitfly *Drosophila melanogaster. Environ Microbiol* **13**: 1889–1900.

Wylie KM, Truty RM, Sharpton TJ, Mihindukulasuriya KA, Zhou Y, Gao H *et al.* (2012). Novel bacterial taxa in the human microbiome. *PLoS One* **7**: e35294.

Zhou J, Wu L, Deng Y, Zhi X, Jiang YH, Tu Q *et al.* (2011). Reproducibility and quantitation of amplicon sequencing-based detection. *ISME J* **5**: 1303–1313.

Supplementary Information accompanies this paper on The ISME Journal website (http://www.nature.com/ismej)