

Fruit fly phylogeny imprints bacterial gut microbiota

Virginie Ravigné, Nathalie Becker, François Massol, Erwan Guichoux, Christophe Boury, Frédéric Mahé, Benoit Facon

▶ **To cite this version:**

Virginie Ravigné, Nathalie Becker, François Massol, Erwan Guichoux, Christophe Boury, et al.. Fruit fly phylogeny imprints bacterial gut microbiota. Evolutionary Applications, 2022, 15 (10), pp.1621-1638. 10.1111/eva.13352 . hal-03640460

HAL Id: hal-03640460

https://hal.univ-reunion.fr/hal-03640460

Submitted on 19 May 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Check for updates

Received: 13 August 2021 | Revised: 20 January 2022 | Accepted: 24 January 2022

DOI: 10.1111/eva.13352

Evolutionary Applications Open Access WILEY

SPECIAL ISSUE ORIGINAL ARTICLE

Fruit fly phylogeny imprints bacterial gut microbiota

Virginie Ravigné[1,2] | Nathalie Becker[3] | François Massol[4] | Erwan Guichoux[5] | Christophe Boury[5] | Frédéric Mahé[1,2] | Benoit Facon[6]

[1]CIRAD, UMR PHIM, Montpellier, France

[2]PHIM, Univ Montpellier, CIRAD, INRAE, Institut Agro, IRD, Montpellier, France

[3]UMR ISYEB, MNHN, Paris, France

[4]Inserm, CHU Lille, Institut Pasteur de Lille, U1019 – UMR 9017, Center for Infection and Immunity of Lille (CIIL), CNRS, Université de Lille, Lille, France

[5]INRAE - UMR 1202 BIOGECO - Plateforme Genome Transcriptome de Bordeaux, Cestas, France

[6]UMR CBGP, INRAE, Montferrier-sur-Lez, France

Correspondence
Virginie Ravigné, UMR PHIM, TA A-120 / K, Campus international de Baillarguet, 34398 Montpellier Cedex 5, France.
Email: virginie.ravigne@cirad.fr

Funding information
European Regional Development Fund, Grant/Award Number: GURDT I2016-1731-0006632; European Union FEDER, Grant/Award Number: 2003227; Conseil Régional d'Aquitaine, Grant/Award Number: 20030304002FA and 20040305003FA; Agence Nationale de la Recherche, Grant/Award Number: ANR-17-CE32-011; Investissements d'Avenir ANR, Grant/Award Number: ANR-10-EQPX-16-01; Centre de Coopération Internationale en Recherche Agronomique pour le Développement

Abstract

One promising avenue for reconciling the goals of crop production and ecosystem preservation consists in the manipulation of beneficial biotic interactions, such as between insects and microbes. Insect gut microbiota can affect host fitness by contributing to development, host immunity, nutrition, or behavior. However, the determinants of gut microbiota composition and structure, including host phylogeny and host ecology, remain poorly known. Here, we used a well-studied community of eight sympatric fruit fly species to test the contributions of fly phylogeny, fly specialization, and fly sampling environment on the composition and structure of bacterial gut microbiota. Comprising both specialists and generalists, these species belong to five genera from to two tribes of the Tephritidae family. For each fly species, one field and one laboratory samples were studied. Bacterial inventories to the genus level were produced using 16S metabarcoding with the Oxford Nanopore Technology. Sample bacterial compositions were analyzed with recent network-based clustering techniques. Whereas gut microbiota were dominated by the Enterobacteriaceae family in all samples, microbial profiles varied across samples, mainly in relation to fly identity and sampling environment. Alpha diversity varied across samples and was higher in the Dacinae tribe than in the Ceratitinae tribe. Network analyses allowed grouping samples according to their microbial profiles. The resulting groups were very congruent with fly phylogeny, with a significant modulation of sampling environment, and with a very low impact of fly specialization. Such a strong imprint of host phylogeny in sympatric fly species, some of which share much of their host plants, suggests important control of fruit flies on their gut microbiota through vertical transmission and/or intense filtering of environmental bacteria.

KEYWORDS
community ecology, long-read sequencing, metabarcoding

1 | INTRODUCTION

Agro-ecosystems comprise a significant proportion of land use and harbor a non-negligible fraction of biodiversity (Pimentel et al., 1992; Tilman et al., 2011). More than many others, these ecosystems suffer from intense structural anthropogenic alterations. Conflicting imperatives to intensify production while simultaneously reducing environmental impacts increasingly drive short-term and fine-scale

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2022 The Authors. *Evolutionary Applications* published by John Wiley & Sons Ltd.

ecological and evolutionary processes (Thrall et al., 2011), demanding greater capacity to predict and manage their consequences (Gilligan, 2008). One promising avenue for reconciling the goals of crop production and ecosystem preservation consists in manipulating quantitatively and/or qualitatively beneficial biotic interactions (Gaba et al., 2014; Massol & Petit, 2013). Over the last decade, this strategy has taken a new turn by considering risks and opportunities associated with plant and insect microbiota. In particular, microbes associated with phytophagous insect are thought to offer great potential for improved management of economically important pests (Crotti et al., 2012). For instance, gut bacteria can be used to reverse radiation-induced fitness decrease in sterile males used in the sterile insect technique, to produce new bacterial odoriferous attractants for insect traps, or to stimulate insect behaviors such as feeding or oviposition (Noman et al., 2020; Raza et al., 2020). Yet, identification of the associated microbial species, and of their respective role in plant–insect interactions and dynamics, is still far from complete.

There is now good agreement on the idea that microbes may play an important role in host adaptation (Macke et al., 2017). In particular, one of the major arenas for host–microbe interactions is the insect gut, which is typically colonized by a large number of diverse microbes, among which bacterial associations predominate (Engel & Moran, 2013). Empirical evidence accumulates, showing that insect gut microbiota can affect host fitness by contributing to development, host immunity, nutrition, or behavior (Kolodny et al., 2020). Gut microbes have even been suspected to be the hidden key player of plant exploitation by their insect pests, as, for example, for the olive fly *Bactrocera oleae* (Ben-Yosef et al., 2014) and the coffee berry borer *Hypothenemus hampei* (Ceja-Navarro et al., 2015).

Gut microbiota are complex, heterogeneous, and variable communities of microbes. First, they assemble within each host generation through different transmission routes. Specifically, gut microbes are mainly acquired via horizontal transfer from the surrounding environment (Broderick & Lemaitre, 2012). However, a number of mechanisms exist for inoculating progeny with microbial symbionts, increasing rates of vertical transmission, and enabling long-term associations (Engel & Moran, 2013). For example, in some flies eggshells are contaminated with parental bacteria (Capuzzo et al., 2005; Raza et al., 2020). Even when acquired horizontally at each generation, gut communities are not random assemblages of bacteria from the food or local environment, due to host filtering and promoting specific bacteria (Engel & Moran, 2013). Second, insect species vary immensely in their dependence on gut microbes: Some almost lack them entirely, while others have developed obligate dependence (Moran et al., 2019). Third, host–microbiota interactions extend along the parasite–mutualist continuum and the exact position may change according to the cost–benefit balance resulting from interactions between bacteria composing the microbiota (Mushegian & Ebert, 2016). Fourth, gut microbiota are often considered as having a multilayered structure (Shapira, 2016). One layer would be the so-called core microbiota, which would tend to be under host genetic and immune control, reliably transmitted across generations, and sharing evolutionary interests with the host (Macke et al., 2017). Some of these microbes may be beneficial to the host and contribute to essential functions or provide long-term adaptation to stable features of the host niche (Nougué et al., 2015). A second layer would be composed of a flexible, environment-modulated pool of microbes, varying within the course of individual life and exhibiting high interindividual variation. Because of possibly divergent evolutionary interests, microbes from this second layer could either be beneficial or detrimental to the host (Macke et al., 2017), potentially depending on the rest of the gut microbiota members (Mushegian & Ebert, 2016).

In relation to this important variability of insect–microbe associations, understanding the role of gut microbiota in plant–insect interactions may benefit from deciphering the determinants of gut microbiota composition and structure. Gut microbiota are affected by many factors, including host phylogeny and host ecology (Spor et al., 2011). First, the environment in which insects develop and live strongly determines the set of bacteria, with which they will have an opportunity to associate. In phytophagous insects, the environments encountered are not random. They depend on insect ecology, a major feature of which is host range, that is, the host plant species an insect uses. For instance, one could expect that insect species specialized on different host plants encounter different initial microbe pools and that generalist insect species encounter a more diverse set of microbes than specialist species (Deb et al., 2019). Second, host phylogeny could potentially structure insect gut microbiota through different mechanisms ranging from active filters (constrained by host development, immune function morphology, and physiology), to the sharing of similar microbe pools (through social interactions or similarity in diet; Brooks et al., 2016). While host phylogeny, host specialization, and sampling environment factors are all considered as potential determinants of gut microbial communities, their relative importance is still a matter of debate, not only because it probably varies across taxa but also because of the associated technical challenge. Studies generally compare gut microbiota among related host species with contrasting ecologies in natural environments (Ivens et al., 2018), and through broad phylogenetic sampling of animals with both divergent and convergent feeding ecologies (Nishida & Ochman, 2018). However, in addition to their differences in phylogenetic history and level of specialization, surveyed host species may differ in their geographic ranges, thus experiencing different microbial species pools in their local environment. Controlled or laboratory environments, used for studies of closely related host taxa (Erlandson et al., 2018; Kohl et al., 2018), may partially reduce this bias. However, sampled microbial pools are unlikely to be representative of those encountered in the wild. This limitation can be overcome by analyzing microbiota in sympatric species of known ecology and phylogenetic history (Martinson et al., 2017).

Reunion, a small island in South-West Indian Ocean, harbors a community of eight sympatric fruit flies, considered as the main actors in the local guild of fruit-eating phytophagous arthropods (Quilici & Jeuffrault, 2001), which could constitute a convenient system to tackle this question. These species belong to five genera from two tribes of the Tephritidae family (Moquet et al., 2021):

Three species are Ceratitinae (*Ceratitis capitata*, *Ceratitis quilicii*, and *Neoceratitis cyanescens*), and five species are Dacinae (closely related species *Bactrocera dorsalis*, *Bactrocera zonata*, and *Zeugodacus cucurbitae* on the one hand, and *Dacus ciliatus* and *Dacus demmerezi*, on the other). They differ in their level of specialization: Four are generalist species (the *Ceratitis* and *Bactrocera* species, commonly found on more than 30 plant species of several distant plant families), three are specialists of Cucurbitaceae (the *Dacus* and *Zeugodacus* species), and one is a specialist of Solanaceae (*N. cyanescens*). Most importantly, both tribes comprise specialist and generalist species.

Gut microbiota of Tephritidae have received great attention among those of phytophagous insects due to their promises for innovative pest management strategies (Deutscher et al., 2018; Noman et al., 2020), and because Tephritidae, which have a worldwide distribution, include some of the most economically damaging fruit and vegetable crop pests (Qin et al., 2015). The functional role of some particular bacterial taxa has been investigated within Tephritidae revealing links with nutritional provisioning (Behar et al., 2005), resistance to pathogenic bacteria (Behar, Yuval, et al., 2008), social interactions (Hadapad et al., 2016), pesticide resistance (Cheng et al., 2017), and foraging behavior (MacCollom et al., 2009). More recently, metabarcoding studies using next-generation sequencing have helped describe the diversity and structure of the gut bacterial communities associated with wild Tephritid flies (Noman et al., 2020). These studies have uncovered a substantial diversity of gut bacteria with a strong predominance of the Proteobacteria phylum, including many genera of the Enterobacteriaceae family. Some conclusions, such as the lower diversity of microbial communities harbored in laboratory-reared insects compared with field-collected ones (Liu et al., 2016; Ras et al., 2017), and a core microbiota found only at the family level (De Cock et al., 2020; Deutscher et al., 2019), were shared by most studies. However, these studies also came to contrasted conclusions about the relative importance of host plants (Behar, Jurkevitch, et al., 2008; Majunder et al., 2019; Malacrinò et al., 2018; Ventura et al., 2018) or fruit fly species (De Cock et al., 2020; Morrow et al., 2015) in determining the composition and variation of gut bacterial communities in natural populations.

Here, we aimed at using the fruit fly community of Reunion Island to test the contributions of fly phylogeny, fly specialization, and fly sampling environment on the composition and structure of their bacterial gut microbiota. To do so, for each species, whose precise host range and phylogenetic history are known, we studied bacterial gut communities in samples from two contrasted environments (field vs. laboratory). Assessing the amount of network variation driven by different environmental and biological factors is still an experimental and statistical challenge (Joffard et al., 2019). Here, bacterial inventories were conducted using 16S metabarcoding with the Oxford Nanopore Technology, reported to confer a greater taxonomic resolution than Illumina at the genus level (Matsuo et al., 2021; Nygaard et al., 2020), and hence a key feature to dig into the diversity of Enterobacteriaceae. Moreover, meaningful network analyses relied on the framework recently proposed by Massol et al. (2021), based on two methods: (i) group decomposition followed by canonical correspondence analysis (CCA) and (ii) singular value decomposition (SVD) followed by redundancy analysis (RDA).

2 | MATERIALS AND METHODS

2.1 | Sample collection and DNA extraction

Details on each sample are provided in Table S1, Appendix S2. Field samples were collected in several localities between April and June 2018. When possible, flies were caught with pheromone traps in places where several host plants coexist. For species with no efficient trap, flies were collected from sets of infested fruits from a given locality (details in Table S1, Appendix S2). As pheromone traps only attract males, only male individuals were included in the study. Differences in gut composition between the sexes have been found nonsignificant in a preliminary study on *C. capitata* and *B. dorsalis* (not shown), and in previous studies on *B. dorsalis* (Andongma et al., 2015; Liu et al., 2018) and on another Tephritid species (*Bactrocera carambolae*, Yong et al., 2017). Laboratory flies were collected using mouth aspirator in populations maintained in the laboratory of Plant Populations and Bio-agressors in Tropical Ecosystems Joint Research Unit (Saint-Pierre, Reunion Island). All flies were stored for at least 48 h in fresh 90% ethanol at −30°C in a 10× liquid/fly volume ratio to optimize washing and dilution of any external bacteria. One hour prior to dissection, flies were rinsed at ambient temperature by successive buffers providing three more washes (75% ethanol, 50% ethanol, and 25% ethanol, 5 min each), while ensuring a progressive rehydration of the abdominal tissues for dissection. Dissection of the abdominal gut portion was performed on a sterilized glass slide with a pair of sterile tweezers under a stereomicroscope. The abdominal gut portion includes the midgut and the ileum of the hindgut, excluding anterior thoracic crop, foregut, and posterior rectum. For each sample, guts from around 30 males were dissected under sterile conditions and pooled.

DNA extraction from dissected guts was performed using the DNeasy Blood & Tissue Kit (Qiagen) following the manufacturer's instructions, adding 0.5% N-lauroyl sarkosyl (Merck KGaA) for 30', 65°C at the end of the lysis step. DNA was subsequently checked for quantity and quality with a NanoDrop 2000 (Thermo Fisher Scientific Inc.).

2.2 | Gene amplification and MinION sequencing

For each sample, ~10 ng of DNA was amplified using specific primers that target the whole 16S rRNA gene (27F 5′-AGAGTTTG GATCMTGGCTCAG-3′; 1492R 5′-GGTTACCTTGTTACGACTT-3′), as well as subsequent specific barcodes using a 16S Barcoding Kit (SQK-RAB204; Oxford Nanopore Technologies). After bead purification for removal of excess primers, amplification products were attached to rapid sequencing adapters before being loaded on a MinION flow cell for real-time sequencing. Samples were analyzed in three separate

experiments (RUN1, RUN2C, and RUN3 in barcodes cited in Table S1), each containing a mock community sample (more details in Appendix S1).

2.3 | Bioinformatics

Basecalling, demultiplexing, and chimera removal were performed using Guppy v4.0.11 (https://community.nanoporetech.com). Reads were trimmed (only nucleotides between positions 60 and 1460 bp of the 16S rRNA gene were kept) and filtered (only sequences longer than 900 bp and above quality score Q10 were kept) using Nanofilt (De Coster et al., 2018), leading to a total of 268,960 sequences (ranging from 4693 to 36,902 across the 16 samples). Taxonomy was assigned by confronting reads to the Silva 138 database (Quast et al., 2013; Yilmaz et al., 2014) using VSEARCH 2020.8.0 (Rognes et al., 2016) embedded in QIIME 2 2020.8 (Bolyen et al., 2019), with a percentage of identity of 90%. A PHYLOSEQ object was produced and imported in R (McMurdie & Holmes, 2013; R Core Team, 2020). Examining mock samples revealed correct identification of mock taxa at all taxonomic levels, with relative abundances both very constant across runs and very close to the expectation (Figure S1 in Appendix S1). Among all reads, the percentage of successful assignment (proportion of total reads assigned to a taxon identified in the reference database) was 78.1% at phylum, class, order, and family levels. It dropped to 74.9% at the genus level and 34.0% at the species level. For further analyses, features were merged at the genus level, constituting an incidence table of 105 genera in 16 fruit fly samples. As the maximal relative abundance of a false-positive taxon was 0.001 in mock community samples, the incidence table was filtered of taxa below this threshold relative abundance. This led to a final incidence table of 46 genera (list provided in Table S2) in 16 fruit fly samples (Table S3), used for all following statistical analyses.

2.4 | Community diversity analyses

Community diversity was described as "effective numbers" (Hill, 1973; Jost, 2006, 2007) of bacterial genera within and among sample groups. The total (gamma) diversity of each group was multiplicatively partitioned into two components: (i) alpha diversity, the within-group component; and (ii) beta diversity, the among-group component, that is, the effective number of completely distinct communities present (Jost, 2006). Diversity decomposition was performed using iNEXT (Hsieh et al., 2016) and the multipart() function of package VEGAN in R (Oksanen et al., 2020). To approximate uncertainty around diversity estimates, hierarchical bootstrapping was used. Further exploration of the variability of gut microbiota was conducted by nonmetric multidimensional scaling (NMDS) applied to the Bray–Curtis dissimilarity scores (Bray & Curtis, 1957).

2.5 | Network analyses

To determine to what extent gut community structure is driven by fruit fly phylogeny, specialization, or sampling environment, we applied two network analysis methods exposed in Massol et al. (2021). To account for fly phylogeny, samples were divided into four groups based on fly genus: *Neoceratitis*, *Ceratitis*, *Dacus*, and the group formed by *Bactrocera* and *Zeugodacus*. The two latter genera are considered very close, to the point that until recently *Z. cucurbitae* was called *Bactrocera cucurbitae* (Virgilio et al., 2015; Zhang et al., 2010). Specialization groups are based on known host ranges in Reunion as inferred from long-term observational data (Moquet et al., 2021) and divide samples into three groups: generalists (*Bactrocera* and *Ceratitis* species), specialists of Cucurbitaceae (*Z. cucurbitae* and *Dacus* species), and specialist of Solanaceae (*N. cyanescens*). Sampling environment opposes laboratory versus field samples.

The first method is based on inferring groups within the observed network. It compares this grouping of nodes (here samples) with groups based on factors at stake (here fly phylogeny, fly specialization, and fly sampling environment). The second method assesses the link between multivariate explanatory variables and network structure using redundancy analyses after SVD of the incidence matrix. In both methods, the significance of effects can be gauged through randomization.

Read counts can be poor proxies of abundances due to distortions inherited from the PCR process itself, and to representation biases of bacteria in reference databases (Brooks et al., 2016; Pollock et al., 2018). Therefore, it is generally considered safer to use presence–absence data. Here mock samples suggested both repeatable and moderate biases in abundance estimates from read counts. Thus we systematically conducted all community analyses on two versions of the sample × bacterial taxa incidence matrix: the weighted matrix, containing raw read counts; and presence–absence matrices, obtained by applying a threshold after rarefaction of the weighted matrix. While presence–absence matrices are generally considered to enable coping with uncertainty on relative abundance inference, they give rare taxa more weight into the analysis, as compared to weighted matrices. Importantly, because rarefaction is a random process, all analyses were applied on a distribution of presence–absence matrices, a safety step rarely done in microbiome studies.

2.5.1 | Binary incidence matrices

Presence–absence matrices can be obtained from read count data by setting a read count threshold below which a taxon is considered absent. Such threshold will only be meaningful if samples are first rarefied to a common total read count. However, rarefaction is a random process generating different matrices each time it is applied (examples are provided in Figure S2, Appendix S2). To account for this variability, we conducted community analyses on 1000 binary

matrices. Each binary matrix was obtained by rarefying the read count matrix to 3000 reads (the smallest read count was 3250 for a *C. capitata* sample) and applying a threshold of three reads. This threshold value was determined by rarefying the mock samples to 3000 reads as well, and observing that false positives were never above three reads. For each observed binary matrix, a search of groups was conducted by maximizing network modularity with the leading eigenvector algorithm (Newman, 2006) using the R package IGRAPH (Csardi & Nepusz, 2006). The membership of each fruit fly sample to inferred groups was summed up into a binary adjacency matrix (16 samples × 16 samples) with zero if samples belonged to two different groups and one if they belonged to the same group. The probability that two samples belong to the same group was then obtained as the proportion of the 1000 binary matrices leading to group these samples together.

2.5.2 | Null models for binary bipartite networks

Subsequent analyses required producing null distributions of network statistics. Following Massol et al. (2021), we produced a null model, called the configuration model, using the "curveball" algorithm (Strona et al., 2014), with functions "simulate" and "nullmodel" of R package VEGAN. In theory, each observed binary matrix is associated with a specific null distribution, which can only be approached by simulating multiple networks. For the sake of computation time, in the following, each "curveball"-based test was performed using 1000 simulated networks for each of 100 observed binary matrices.

2.5.3 | Sample classification-based tests

We assessed the effect of sampling environment, fly phylogeny, and fly specialization on gut bacterial community structure. We first proceeded one factor at a time and tested the congruence of sample classifications obtained through community-search algorithms with those associated with external categorical variables, using the Normalized Mutual Information Index (NMI) (Astegiano et al., 2017) available through the function "compare" in the R package IGRAPH (Csardi & Nepusz, 2006). The NMI takes values between zero, indicating no congruence, and one, corresponding to perfect congruence. One NMI value can be obtained for each rarefied matrix, and its associated significance can be inferred from 1000 corresponding matrices simulated under the null model as explained above. Here, mean NMI values were obtained on 1000 rarefied matrices, and the mean associated *p*-value was obtained by comparing 100 rarefied matrices with 1000 corresponding simulations each.

To extend the same logic to multiple factors, we used CCA (ter Braak, 1986) using the function "cca" in the R package VEGAN. CCA allowed decomposing the variation of the community-based classification of samples relatively to fly phylogeny, fly specialization and sampling environment. CCA can classically test the significance of a given "fraction" (e.g., chi-square explained by factors *X* or *Y* once the effect of *Z* has been removed) by comparing the obtained *F*-statistic with those yielded by randomizations of data rows (Peres-Neto et al., 2006). Using the null model matrices, we could further test whether an effect that is deemed significant based on classical row permutations is purely due to heterogeneity in node degrees between communities (i.e., not significantly different from edge-permuted expectation; richness effect) or not (affinity effect). Again, the whole CCA was conducted on 100 rarefied matrices, using 1000 corresponding simulations each.

2.5.4 | Singular value decomposition-based tests

As a complementary approach, we also modeled the effects of fly phylogeny, fly specialization, and sampling environment on network structure using SVD coupled with RDA as explained in Massol et al. (2021). Any given $n \times p$ bipartite network can be approximated as two matrices (**L** and **R**) with a low number of columns and as many rows as nodes (n in **L**, p in **R**). Matrices **L** and **R** can be analyzed through a RDA to gauge how much variation among rows is explained by external variables. The number of vectors to keep after SVD was fixed after examining the congruence between communities inferred from SVD-approximated networks with those inferred from the original network. SVD-approximated networks were obtained by multiplying matrices **L** and **R** and setting a threshold for interaction prediction. Congruence between communities was obtained using the NMI between module partitions on a number of rarefied matrices.

2.5.5 | Weighted incidence matrix

A similar approach was applied to the weighted (read counts) incidence matrix, with the following differences. First, with weighted matrices, it is recommended to proceed through latent block models (LBMs) rather than modularity maximization to look for groups of nodes in networks (Leger et al., 2015). We therefore inferred groups using LBM with the R package SBM (Chiquet et al., 2021). We used a Gaussian distribution to model log-transformed read counts. The best grouping was selected based on ICL criterion (Integrated Complete-data Likelihood, a penalized likelihood criterion suited for clustering; Biernacki et al., 2000). Second, as no rarefaction step was used, analyses were conducted only once. Third, the null model comprised 10,000 matrices produced by Gaussian sampling on the outer product of margins of the log-transformed weighted incidence matrix.

3 | RESULTS

3.1 | Descriptive analyses

The full bacterial composition of samples is provided in Figure 1 and Table S3 (Appendix S2). Rarefaction curves for each sample are provided in Figure S3 (Appendix S2). The 46 bacterial genera identified

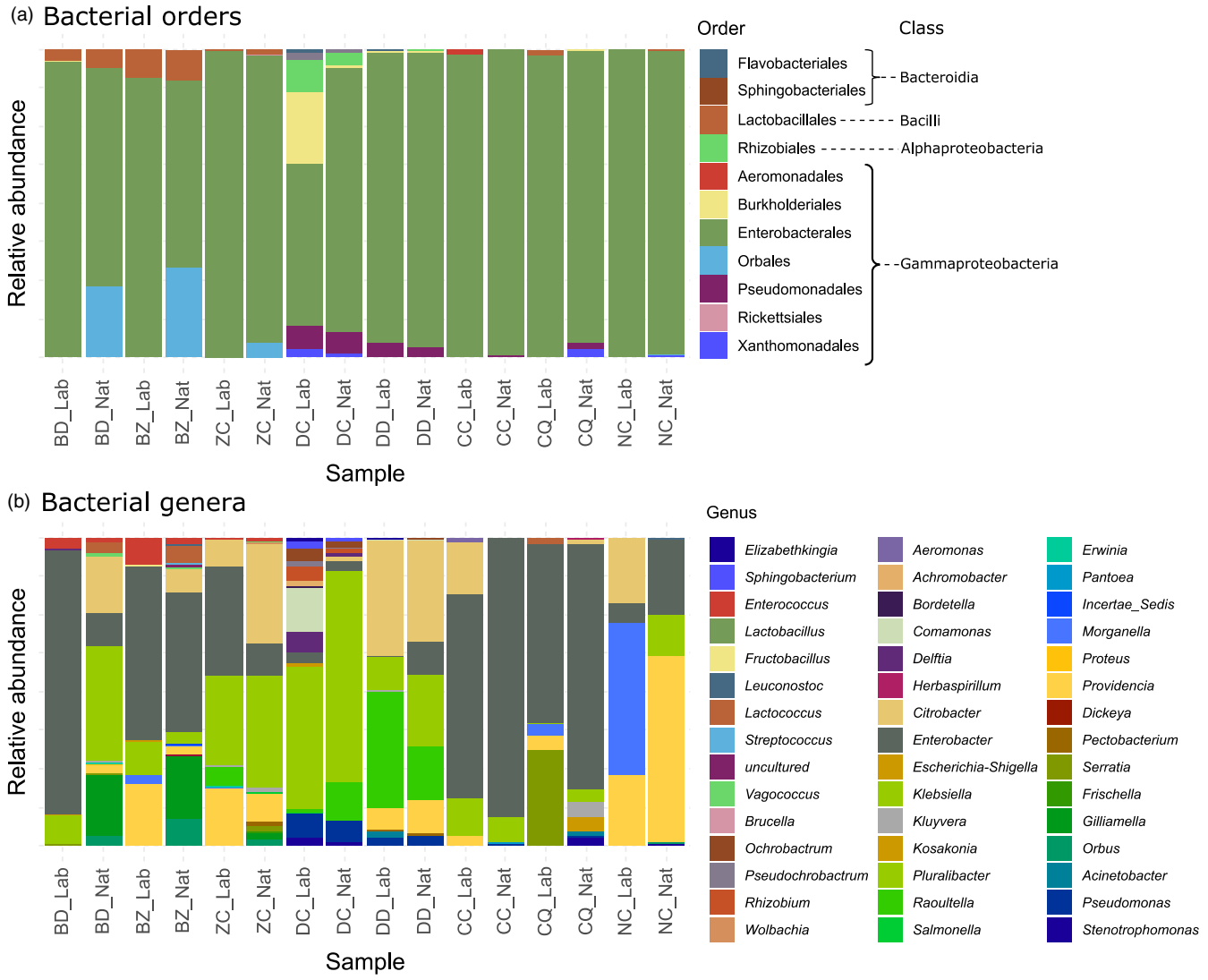


FIGURE 1 Relative abundances of bacterial taxa in the 16 samples. Within bars, colored areas indicate the proportion of each bacterial taxon in each sample from the weighted incidence matrix (Table S3), and white lines between two colored areas separate different genera. (a) Colors correspond to bacterial orders. (b) Colors indicate bacterial genera. Sample names are composed by the initials of fly species names (BD, *Bactrocera dorsalis*; BZ, *Bactrocera zonata*; CC, *Ceratitis capitata*; CQ, *Ceratitis quilicii*; DC, *Dacus ciliatus*; DD, *Dacus demmerezi*; NC, *Neoceratitis cyanescens*; and ZC, *Zeugodacus cucurbitae*) and the mention of sampling environment (Lab, laboratory; Nat, field, full details in Table S1)

in the global dataset belonged to three phyla: Proteobacteria (97.7%), including two classes, eight orders, 15 families, and 36 genera; Firmicutes (2.0%), all of class Bacilli and order Lactobacillales, with eight genera in six families; and Bacteroidota (0.3%), only represented by two genera (Table S2). The genera above 1% in total abundance belonged to two classes (Bacilli and Gamma-Proteobacteria), with an overrepresentation of the latter (97.1%). Among the 11 detected orders, four were above 1% in total abundance (Enterobacterales, Lactobacillales, Orbales, and Pseudomonadales), with an overrepresentation of Enterobacterales in all samples (52.5%–99.9%). Only five families (Enterobacteriaceae, Enterococcaceae, Morganellaceae, Orbaceae, and Pseudomonadaceae) were above 1% in total abundance, with an overrepresentation of Enterobacteriaceae (69.1%) in all fly species but *N. cyanescens*, dominated by the phylogenetically close Morganellaceae. Only nine genera were above the 1% threshold (*Enterobacter*, *Klebsiella*, *Citrobacter*, *Providencia*, *Morganella*, *Raoultella*, *Gilliamella*, *Pseudomonas*, and *Enterococcus*; Table S3). Some bacterial taxa have preferential associations with fly phylogenetic groups. The Bacteroidota phylum and the Alpha-Proteobacteria class tended to associate with samples of *Dacus*. The Firmicutes phylum associated with *Bactrocera* and *Zeugodacus* samples. Some bacterial taxa had variable prevalence across sampling environments as well. Examples of genera with variable prevalence between laboratory and field samples include *Enterobacter*, *Morganella*, and *Citrobacter*. Finally, for some bacteria, the prevalence seemed determined by both fly phylogeny and sampling environment, such as the Orbales class, mainly found in field samples of *Bactrocera* and *Zeugodacus*.

3.2 | Diversity partitioning

The total (gamma) diversity of the 16 samples was 8.40 (95% CI 5.72–10.18) genus equivalents.

The alpha diversity of samples ranged from 1.41 (95% CI 1.41–1.45, for field *C. capitata*) to 6.86 (95% CI 6.84–7.11, laboratory *D. ciliatus*) genus equivalents, with a mean of 4.01 (SE 0.41) (Figure 2). Average alpha diversity of laboratory (3.87, SE 0.54) and field samples (4.14, SE 0.65) was close. For all Dacinae samples but *D. ciliatus*, the field sample was more diverse than the laboratory sample, whereas in Ceratitinae, the laboratory sample was more diverse than the field sample (Figure 2). Among laboratory samples, there was no clear link between alpha diversity and the number of generations spent by populations in the laboratory prior to sampling (Appendix S2, Figure S4). Alpha diversity did not seem to particularly correlate with specialization (Figure 2): Diversity was not greater in generalists (3.45, SE 0.63) than in specialists of Cucurbitaceae and Solanaceae (5.06, SE 0.48 and 3.08, SE 0.31, respectively). In contrast, sample diversity tended to differ between phylogenetic groups (Figure 2). Dacinae samples had an average of 4.83 (SE 0.47) genus equivalents, while Ceratitinae samples had only 2.64 (SE 0.27).

Pairwise beta diversity between samples ranged between 1.03 (between laboratory *B. dorsalis* and field *C. capitata*) and 1.91 (between laboratory *N. cyanescens* and *D. ciliatus*). Differentiation among bacterial communities was not particularly structured by sampling environment, as beta diversity between laboratory and field samples was 1.12. In contrast, beta diversity, even though estimated on the whole dataset (i.e., with both laboratory and field samples), tended to be higher between specialization groups (1.70) and between host phylogenetic groups (1.83).

Nonmetric multidimensional scaling attained a stress value of 0.1932. It tended to group samples by phylogenetic group, rather than by sampling environment or fly specialization (Figure 3), a result also observed in NMDS ordination of presence–absence matrices (Figure S5 in Appendix S2). *Dacus* samples seemed to distinguish from other samples by higher relative abundance of Bacteroidota (genera *Elizabethkingia* and *Sphingobacterium*), lower relative abundance of Firmicutes (eight genera, all of class Bacilli, order Lactobacillales), and higher relative abundance of several genera from two orders of Alpha-Proteobacteria (Rhizobiales and Burkholderiales). Field *Bactrocera* and *Zeugodacus* samples tended to preferentially associate with the Lactobacillales *Streptococcus*, *Lactobacillus*, and *Vagococcus*, and among Gamma-Proteobacteria, with three Orbales

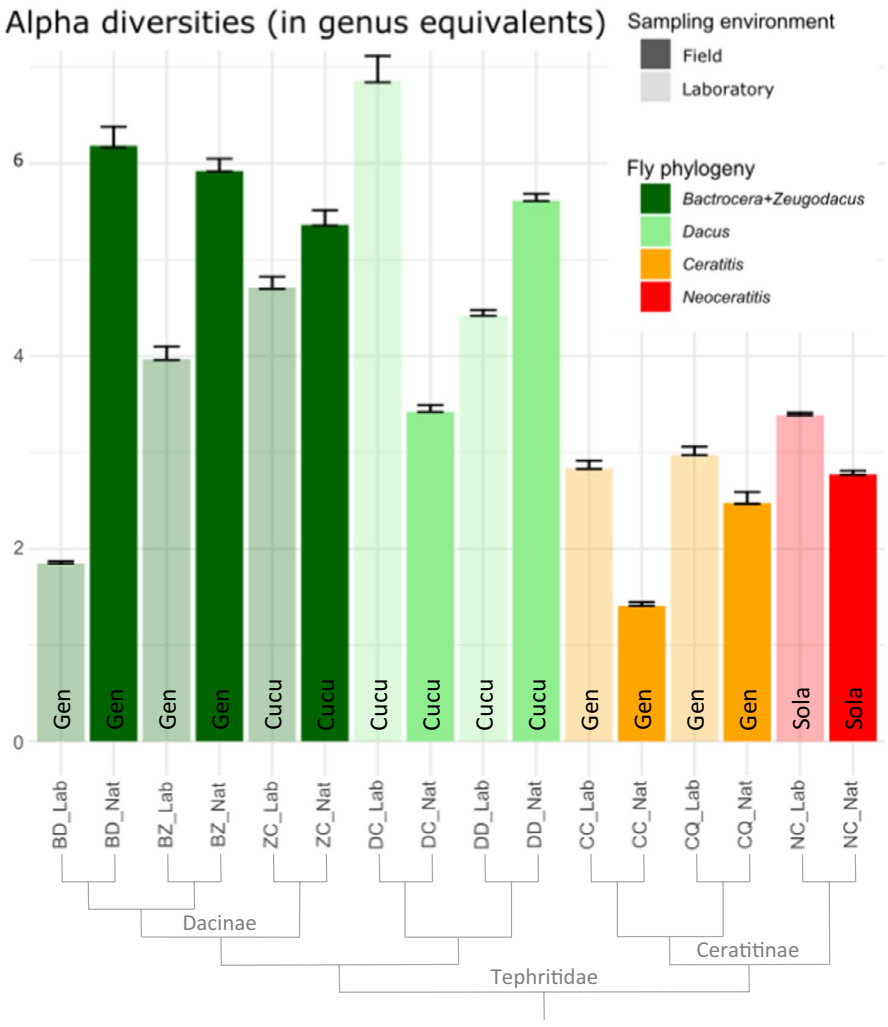


FIGURE 2 Alpha diversity for each sample. Bars indicate Hill numbers of order one (exponential of Shannon diversity, in genus equivalents). Bar colors correspond to fly phylogenetic groups, and shading stands for sampling environment. Codes at the bottom of each bar designate fly specialization groups: "Gen" for generalist species, "Cucu" for specialists of Cucurbitaceae, and "Sola" for the specialist of Solanaceae. Error bars represent 95% confidence intervals estimated by bootstrapping

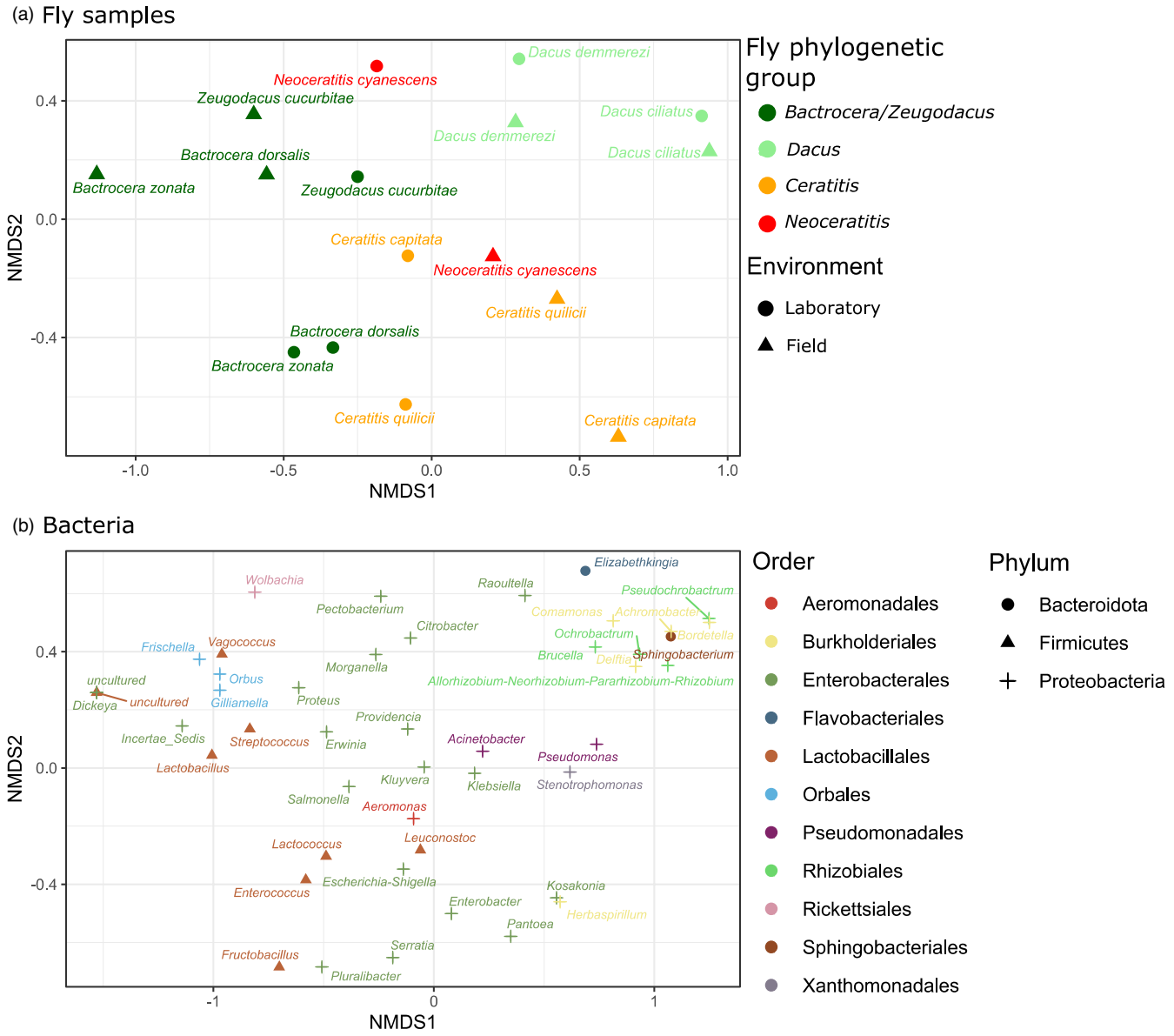


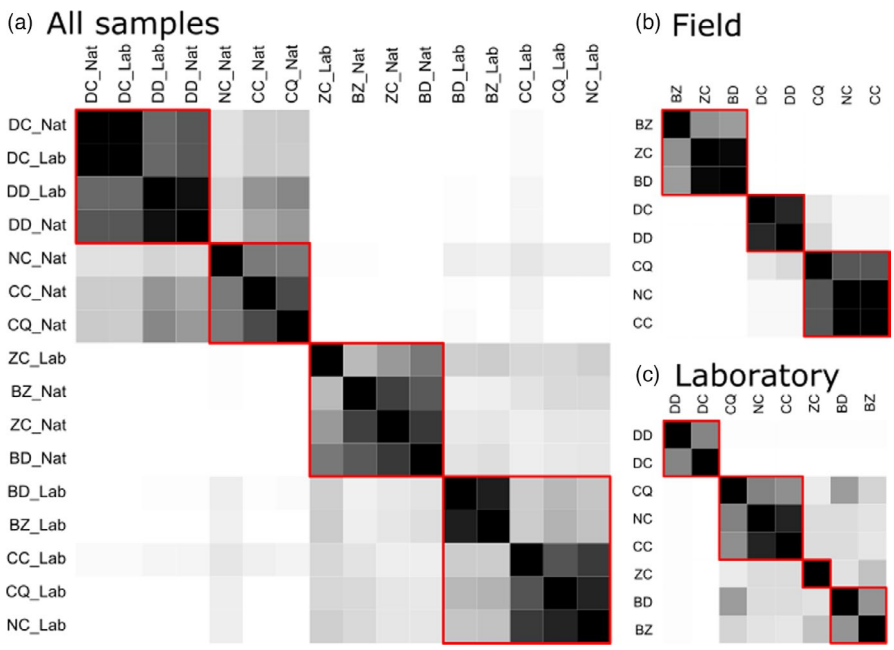
FIGURE 3 Nonmetric multidimensional scaling (NMDS) ordination of fruit fly gut microbiome communities compared using the Bray–Curtis similarity coefficient of bacterial genera relative abundances. (a) Sample ordination. Colors represent fly phylogeny. Symbols represent sampling environment. (b) Ordination of bacterial genera. Colors represent bacterial orders. Symbols correspond to phyla

genera (*Frischella*, *Gilliamella*, and *Orbus*) and some Enterobacterales genera. Field Ceratitinae mainly differed from others by their association with Enterobacterales genera such as *Kosakonia* and *Pantoea* and with the Burkholderiales genus *Herbaspirillum*.

3.3 | Network analysis

3.3.1 | Sample groups

Applying the leading eigenvector community-search algorithm to the 16 samples over 1000 observed presence–absence matrices led to identify 4.203 groups of nodes in the network on average (SE = 0.057, Figure 4a and Figure S6 in Appendix S2), with a relatively high and significant modularity score (Q = 0.301, SE = 0.0005, 95% PI = 0.262–0.328, left panel of Figure S7, Appendix S2). Over a random subset of 100 observed binary matrices, the *p*-value of the observed modularity had a mean of 0.005 (SE = 0.002, 95% PI = 0.000–0.0461, right panel of Figure S7, Appendix S2), suggesting that observed matrices were more structured than expected under the null model. All binary matrices separated at least two relatively stable groups (Figure 4a). The first group tended to split into two subgroups: (i) all *Dacus* samples, whatever their environment, most frequently grouped together (72%–95% of observed binary matrices), and (ii) field samples of Ceratininae species (genera *Ceratitis* and *Neoceratitis*). Field *Ceratitis* species were associated with 80% of observed binary matrices. *Neoceratitis* was less frequently associated with them (68% of observed binary matrices). Samples of



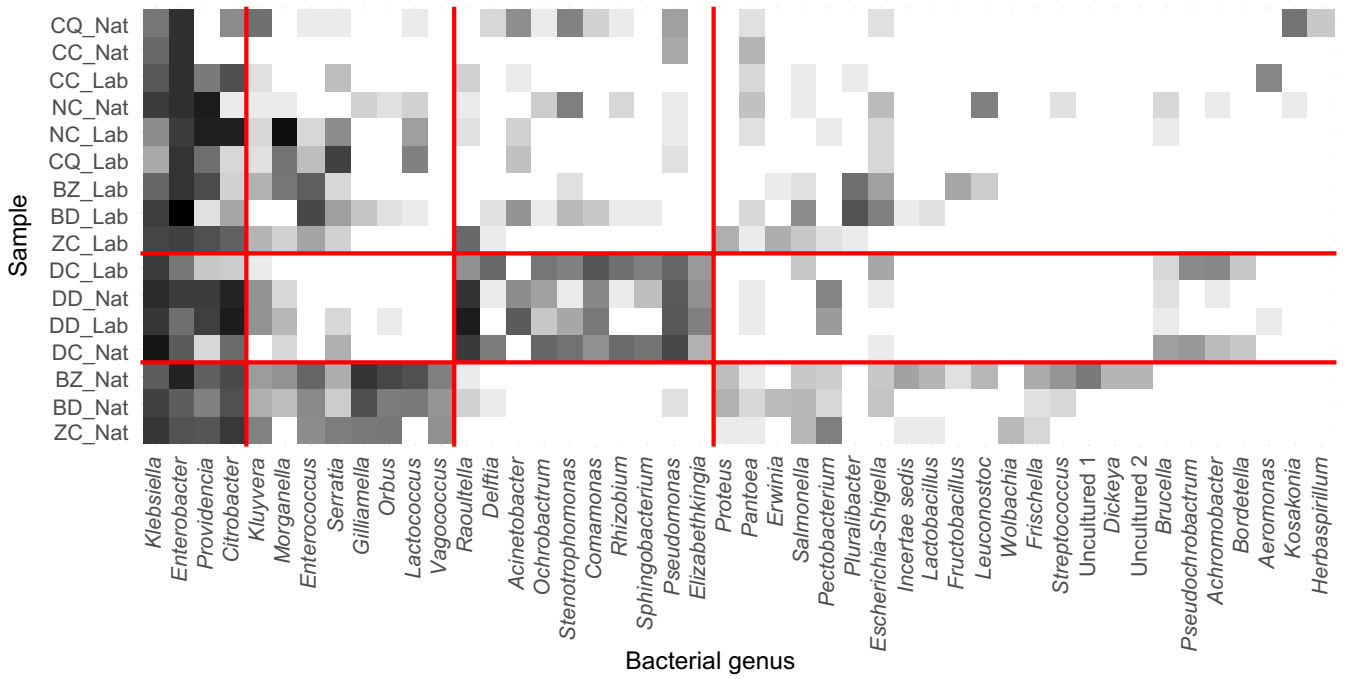


FIGURE 4 Mean clustering of samples based on gut microbial groups identified from 1000 rarefied presence–absence matrices. Color gradient corresponds to the percentage of rarefied matrices in which two samples are found in the same cluster (white = 0%, black = 100%). Red contours indicate the most common clustering. (a) All 16 samples. (b) Only field samples. (c) Only laboratory samples

FIGURE 5 Groups identified from the full read count matrix. Log-transformed read counts are represented on a continuous gradient from white for log10(reads + 1) = 0, to black for log10(reads + 1) = 4.4. Red lines delimit clusters identified under the best latent block model

both subgroups (*Dacus* and field Ceratitinae) were associated with 22.4%–63.6% of observed binary matrices. The second group was also composed of two main subgroups, with more variable composition: (i) *Zeugodacus* samples and field *Bactrocera* samples (percentages varying from 51% to 84%), and (ii) all remaining samples, that is, laboratory *Bactrocera* and Ceratitinae samples.

The same community-search algorithm also revealed an average of 3.45 (SE = 0.02) groups among field samples only and 3.81 (SE = 0.03) groups among laboratory samples, with congruent compositions with the 16-sample grouping (Figure 4b,c).

On the whole weighted incidence matrix, LBM identified three groups of samples (Table S4, Figure S8 in Appendix S2, and Figure 5): one with field samples of *Zeugodacus* and *Bactrocera* species, one with all *Dacus* samples, and the remaining samples (all Ceratitinae samples and laboratory samples of *Zeugodacus* and *Bactrocera*). On field samples only, two groups of samples were found: one with Dacinae species (*Bactrocera*, *Dacus*, and *Zeugodacus*) and one with all Ceratitinae species (*Ceratitis*, and *Neoceratitis*) (Figure S9, Appendix S2). On laboratory samples, no group was identified (Figure S9, Appendix S2).

3.3.2 | Congruence of classifications

Distributions of the congruence indices (NMI) are provided in Table 1. The communities found in the whole network were most congruent with genus-level fly phylogeny (mean p-value < 0.05). Other classifications of samples, that is, based on higher-level fly phylogeny (Ceratitinae vs. Dacinae), sampling environment, or fly specialization, were not statistically more congruent with gut microbiota-based clustering than expected by chance (see Figure 6 for an illustration).

3.3.3 | Canonical correspondence analyses

The results of CCA applied to the communities found on both presence–absence and read count data confirmed the results found by congruence comparisons. Fly phylogeny significantly explained gut bacterial communities, irrespective of the removal of the effects of environment, host specialization, or both (Table 2A,B). On presence–absence data, none of the models omitting phylogeny or removing the effect of phylogeny was significant (Table 2A). Significant models were all doubly significant (with permutation tests on rows and edges), indicating both node richness and affinity differences between groups. On read count data, the models associated with the lowest p-values included both phylogeny and sampling environment (Table 2B). Most significant effects were not significant under edge permutations, indicating an effect mainly driven by differences in gut microbiota richness between groups of nodes.

3.3.4 | Singular value decomposition and redundancy analyses

As a first step, the number of vectors required to faithfully approximate incidence matrices was determined by estimating the congruence between groups obtained from the approximated matrices with groups obtained on the full matrix. On presence–absence data, congruence tended to increase with the number of vectors retained, but the first local maximum occurred between two and four vectors retained depending on the rarefied matrix. On read counts, a single local maximum was observed at four vectors. Adjusted R^2 values of individual fractions are given for these various options (Figure 7) and are all very congruent. Residual error (i.e., variance not explained by fly phylogeny, fly specialization of sampling environment) increased steadily with the number of vectors, but remained high (>36%). For any given number of vectors, fly phylogeny had the highest adjusted R^2, followed by the interaction between fly phylogeny and fly specialization. Sampling environment explained a marginal part of variance on read counts only.

4 | DISCUSSION

The gut bacterial microbiota of eight Tephritidae species were described using Oxford Nanopore MinION full-length 16S metabarcoding. At taxonomic levels ranging from phylum to family, the abundance of bacterial taxa was found congruent with former descriptions obtained with Illumina MiSeq data from other Tephritidae species (for a review, see Noman et al., 2020 and Raza et al., 2020), and from some of these species in other geographic area (De Cock et al., 2020; Hadapad et al., 2016; Malacrinò et al., 2018). Enterobacteriaceae, identified as the most prevalent family in nearly all samples, are reportedly transferred vertically in some species (Aharon et al., 2013; Lauzon et al., 2009; Majunder et al., 2019) and thus are considered important for Tephritid development and physiology. At genus level, existing published studies exhibit substantial variability in descriptions of abundant bacteria. Here, thanks to the higher resolution of long-read metabarcoding, 46 genera were found, the most abundant of which have also been described in other Tephritidae studies, including *Enterobacter*, *Klebsiella*, *Citrobacter*, *Providencia*, *Morganella*, and *Raoultella* (for a review, see Noman et al., 2020 and Raza et al., 2020). In contrast, some genera mentioned as frequent

TABLE 1 Congruence between classifications of samples based on gut bacteria presence–absence data, and potential determinants of community structure

Classification of samples		NMI with clustering based on gut bacteria			p-Value	
Factor	**Modalities**	**Mean**	**SE**	**Percentile interval**	**Mean**	**SE**
Phylogeny	Dacinae–Ceratitinae	0.169	0.002	0.055–0.302	0.462	0.030
	Ceratitis–Neoceratitis–(*Bactrocera* + *Zeugodacus*)*–Dacus*	0.572	0.003	0.427–0.746	0.044	0.007
Sampling environment	Field–laboratory	0.267	0.004	0.056–0.497	0.210	0.026
Specialization	Generalists—specialists of Cucurbitaceae—specialist of Solanaceae	0.304	0.003	0.111–0.466	0.372	0.029

Note: Congruence was estimated by NMI obtained from 1000 observed rarefied matrices. NMI ranges from zero for no congruence, to one for perfect congruence. Significance of these NMI values was obtained from 100 comparisons between one observed rarefied matrix and 1000 random null model matrices.

Abbreviation: NMI, Normalized Mutual Information Index.

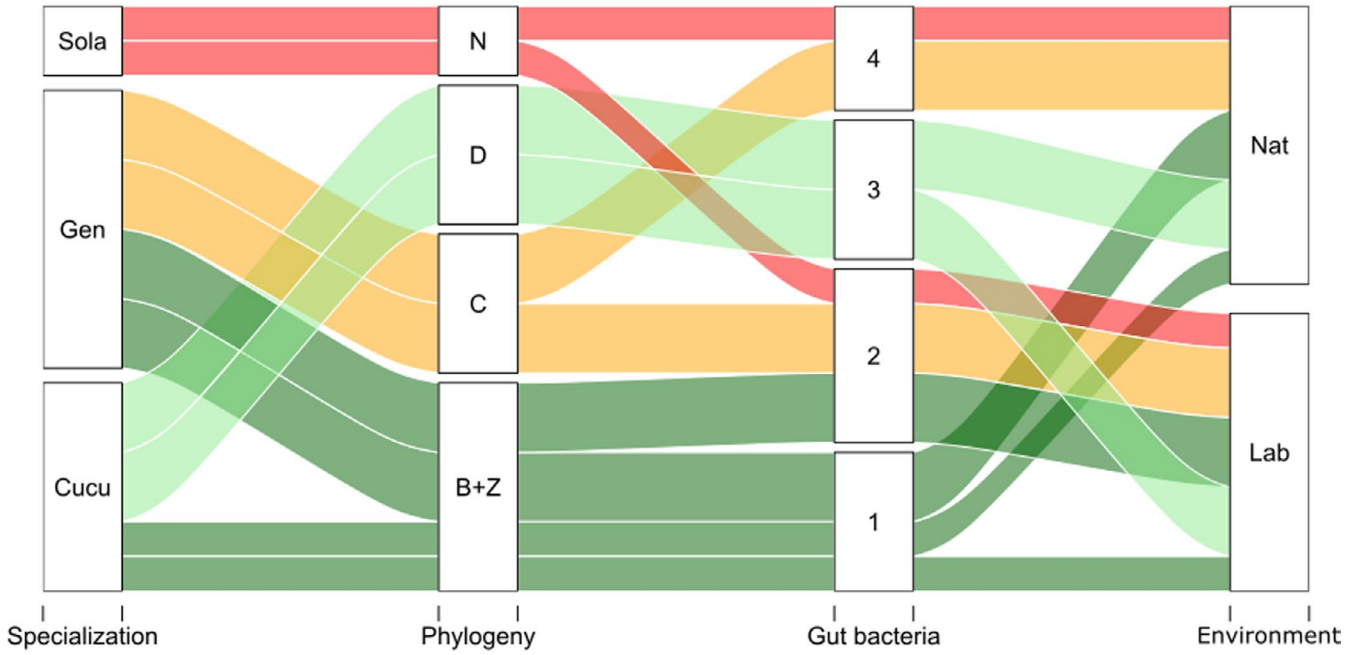


FIGURE 6 Alluvial plot showing how the 16 samples group according to fly specialization, fly phylogeny, fly sampling environment, and gut bacterial composition. Fly specialization counts three groups of samples: the specialists of Solanaceae (the two *Neoceratitis cyanescens* samples), the generalists (the eight *Bactrocera* and *Ceratitis* samples), and specialists of Cucurbitaceae (*Dacus* and *Zeugodacus* samples). The factor "Phylogeny" groups samples by fly genus: N, D, C, and B+Z stand for *Neoceratitis*, *Dacus*, *Ceratitis*, and *Bactrocera* or *Zeugodacus*, respectively. Environment separates laboratory samples (Lab) from field samples (Nat). The last grouping factor (Gut bacteria) is the mean clustering of samples based on gut microbial subcommunities identified from 1000 rarefied presence–absence matrices (the four red squares in Figure 4a). Colors correspond to fly phylogeny

in other Tephritid studies were only found at low abundances here, as for example, *Acetobacter*, *Escherichia*, *Pectobacterium*, and *Serratia*. Whether these discrepancies are due to methodological issues or natural variability cannot be fully deciphered here. In the present study, only one pooled sample by fly species and fly sampling environment was studied, hampering considerations on natural intraspecific variability in gut microbiome composition. For some abundant taxa, results of functional studies monitoring fruit fly fitness are worth mentioning. For instance, *Enterobacter* and *Klebsiella* enhance larval nutrition (Noman et al., 2020 and references herein). An addition of *Klebsiella* in controlled conditions increases pathogen resistance of *C. capitata* (Ben-Ami et al., 2010). In the same way, Cheng et al. (2017) have described the resistance of *Citrobacter* to resist trichlorfon insecticide in *B. dorsalis*. Finally, *Enterobacter*, *Raoultella*, *Klebsiella*, *Citrobacter*, and *Providencia* may also play a role in sexual and host plant attractiveness (Raza et al., 2020 and references herein). In contrast, *Providencia* and *Morganella* have been described as potential pathogens of fruit flies (M'Saad Guerfali et al., 2018; Salas et al., 2017), thus able to decrease fruit fly fitness.

The recent accumulation of sequence data from microbial communities has made some authors plead for an extension of community analyses beyond the exploration of alpha- and beta-diversity patterns in order to detect robust associations between microorganisms and hosts (Barberán et al., 2012; Burns et al., 2016). Here, classic diversity analyses were supplemented with network-based clustering techniques (Massol et al., 2021) using either the leading eigenvector of presence–absence matrices (Csardi & Nepusz, 2006) or LBMs for the read count matrix (Chiquet et al., 2021). Such techniques may help cluster bacterial taxa according to their pattern of association with host flies and gut samples based on their microbial community profiles. Clustering methods may thus provide a natural way of revisiting the notion of core microbiome. Here, the use of various clustering analyses (on all, only laboratory or only field samples, and on read count vs. presence–absence data) supported at least three congruent groups of samples: all *Dacus* samples, field *Bactrocera* and *Zeugodacus* samples, and other samples (Ceratitinae and laboratory *Bactrocera* and *Zeugodacus*). Within this latter group, presence–absence matrices suggested possible subgrouping of Ceratitinae vs. the Dacinae *Bactrocera* and *Zeugodacus*. Clustering of bacteria highlighted a group of bacterial genera accounting for more than half of the bacterial prevalence in all samples: the Enterobacteriaceae *Citrobacter*, *Enterobacter*, and *Klebsiella*, and the Morganellaceae *Providencia*. This group, also supported by numerous studies of Tephritidae microbiota (Behar, Jurkevitch, et al., 2008; Hadapad et al., 2016; Liu et al., 2016; Morrow et al., 2015; Ventura et al., 2018; for a review, see Noman et al., 2020 and Raza et al., 2020), could be considered as a core microbiota at the scale of the Tephritidae family. A second group of bacterial genera, common in field *Bactrocera* and *Zeugodacus* samples, was rare in *Dacus* samples and of variable abundance in other samples. This group included Enterobacterales (*Kluyveria*, *Morganella*, *Serratia*), two Orbales (*Gilliamella* and *Orbus*), and all the Lactobacillales (represented

TABLE 2 Canonical correspondence analyses (CCA) between groups based on gut microbiota and sampling environment (samp env), fly specialization (fly spe), and fly phylogeny (fly phy)

	(A) Presence–absence data					(B) Read counts				
	p-Value (row perm.)		*p*-Value (edge perm.)							
Formulas	Mean	SD	Mean	SD		Chi^2	*F*	*p*-Value (row perm.)	*p*-Value (edge perm.)	
samp env	0.163	0.204	—	—	NS	0.250	2.000	0.269	—	NS
samp env + Cond (fly spe)	0.118	0.163	—	—	NS	0.250	2.678	0.119	—	NS
samp env + Cond (fly phylo)	0.071	0.127	—	—	NS	0.250	7.200	0.044	0.027	**
samp env + Cond (fly spe) + Cond (fly phy)	0.070	0.127	—	—	NS	0.250	6.600	0.041	0.031	**
fly spe	0.255	0.218	—	—	NS	0.630	3.216	0.049	0.389	*
fly spe + Cond (samp env)	0.202	0.186	—	—	NS	0.630	3.653	0.023	0.395	*
fly spe + Cond (fly phy)	0.410	0.241	—	—	NS	0.000	0.000	NA	—	NS
Cond (samp env) + fly spe + Cond (fly phy)	0.334	0.252	—	—	NS	0.000	0.000	NA	—	NS
fly phy	0.018	0.031	0.020	0.042	**	1.333	10.000	0.000	0.283	*
fly phy + Cond (samp env)	0.010	0.026	0.014	0.035	**	1.333	14.933	0.000	0.113	*
fly phy + Cond (fly spe)	0.024	0.033	0.031	0.057	**	0.704	6.861	0.007	0.486	*
Cond (samp env) + Cond (fly spe) + fly phy	0.015	0.027	0.021	0.048	**	0.704	10.133	0.001	0.248	*
samp env + fly spe	0.096	0.127	—	—	NS	0.880	3.664	0.017	0.211	*
samp env + fly phy	0.008	0.022	0.010	0.032	**	1.583	14.250	0.000	0.067	*
fly spe + fly phy	0.035	0.055	0.036	0.068	**	1.333	7.500	0.002	0.471	*
samp env + fly spe + fly phy	0.014	0.037	0.016	0.049	**	1.583	11.400	0.000	0.244	*

Note: Significance of individual fractions was tested by row permutations or edge permutations. Significance based on row permutations is evaluated based on the corresponding *p*-value, and estimated as the probability that a randomized version of the explained contingency table, once removed the effect of conditioning variables, obtains a *F*-statistic equal or larger to the one obtained with real data. Significance based on row permutations is indicated with a star. NS stands for not significant. Significance based on edge permutations is only computed for effects significant with the first test. It is obtained as the probability that a randomized version of the contingency table, keeping node degrees constant, obtains a *F*-statistic equal or larger to the one obtained with real data. Double significance is indicated with two stars. (A) Mean-*p*-values associated with any given combination of factors (with SD) obtained from 100 rarefied presence–absence matrices, with 1000 simulated null matrices each. (B) Chi-square, and *F* and *p*-values associated with any given combination of factors obtained for the read count matrix.

by *Enterococcus*, *Lactococcus*, and *Vagococcus*). Associations between *Lactococcus* and *B. zonata* have already been described (De Cock et al., 2020). The third group of bacteria was preferentially associated with *Dacus* samples: the Alpha-Proteobacteria genera *Rhizobium* and *Ochrobactrum*, the Bacteroidia genera *Elizabethkingia* and *Sphingobacterium*, and among Gamma-Proteobacteria, genera belonging to diverse orders (the Burkholderiales *Comamonas* and *Delftia*, the Pseudomonadales *Acinetobacter* and *Pseudomonas*, the Xanthomonadales *Stenotrophomonas*, and the Enterobacterales *Raoultella*). Other bacteria fell in a fourth cluster, with no obvious association profile, likely due to their low abundances. These non-random associations of bacterial taxa with fly samples were further confirmed by NMDS. Interestingly, some preferential associations occurred at higher taxonomic scales. For instance, Bacteroidota and Alpha-Proteobacteria were mainly associated with *Dacus* samples. In contrast, Firmicutes were completely absent from *Dacus* samples, as well as from field *Ceratitis* samples. Many preferential associations involved different families of Gamma-Proteobacteria and different genera within the Enterobacteriaceae family, raising the need for a finer taxonomic resolution within this key bacterial family. Because of the genuine sympatry of the eight species, the highlighted clusters could not be considered as determined by geographic differentiation in microbial pools, and therefore offer candidate taxa for subsequent functional analyses.

The different methods used to evaluate the relative importance of fly phylogeny, fly specialization, and fly sampling environment converged to the conclusion that fly phylogeny was the main factor explaining microbial profile. In contrast, host ecology (i.e., fly specialization and sampling environment) did not imprint significantly gut microbial communities. For instance, samples of the species *Z. cucurbitae*, a specialist of Cucurbitaceae host plants, systematically grouped with the *Bactrocera* samples, which correspond to generalist species, and not with *Dacus* samples, which share the same host range but are more distant phylogenetically. Samples of both *Ceratitis* generalist species tended to group with the other Ceratitinae species, *N. cyanescens*, a specialist of Solanaceae, rather than with *Bactrocera* samples, which share this generalist niche. The methodological robustness of our results was achieved by the observation of both read count (which tend to give more weight to very abundant species) and of presence–absence (which are more affected by rare taxa) data. Our results, suggesting that microbial profiles are affected by host phylogeny rather than host ecology,

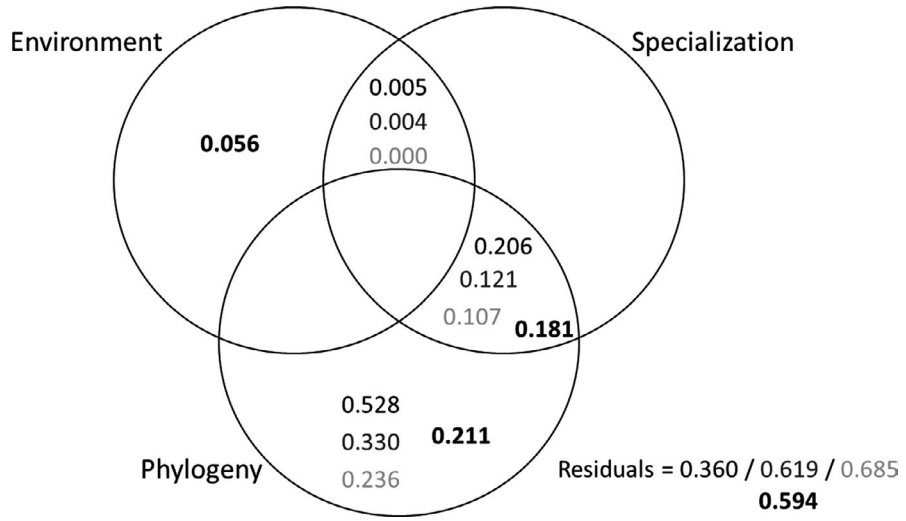


FIGURE 7 Venn diagrams representing the partition of variation (redundancy analysis) within the reduced matrices obtained by singular value decomposition of incidence matrices. Given values are non-negative adjusted R^2 of individual fractions. Negative adjusted R^2 are omitted for clarity. Bold numbers were obtained on the read count matrix approximated with four retained vectors. Nonbold numbers were obtained from presence–absence matrices approximated with two (upper black numbers), three (middle dark gray numbers), and four (bottom light gray numbers) vectors. These latter values are means from 100 presence–absence matrices. All standard errors ranged between 0.002 and 0.004

are thus unsupportive of the hypothesis formulated by Zhao et al. (2016), according to which the Tephritidae gut community membership would be controlled by host genetics, while bacterial abundance would be driven by nongenetic factors.

Phylogenetic determinism of gut microbial communities has been observed in a diversity of taxonomic groups, including nematodes, numerous insect clades, fish, mammals, and hominids (Moran et al., 2019). Such a pattern may indicate a shared, faithful history between hosts and their microbes (Brooks et al., 2016). This process, sometimes referred to as "phylosymbiosis," has been observed in *Nasonia* wasps, and is prone to co-adaptations between hosts and microbes (Brucker & Bordenstein, 2012). Alternatively, this same pattern may very well be driven by physiological, morphological, ecological, or behavioral similarities in closely related hosts that lead to similar environmental filtering of microbial pools (Moran & Sloan, 2015). In Tephritidae species of Reunion Island, host ecology likely determines social interactions. *Bactrocera* and *Ceratitis* species on the one hand, and *Dacus* and *Zeugodacus* species on the other hand, are often found developing in the same fruits (Facon et al., 2021). We do not find such clustering when analyzing their gut microbiota, which suggests that such social interactions unlikely contribute to the structure of gut microbiota in the studied species.

The present results do not conform to the hypothesis that generalist species should have more diverse gut microbial communities (Deb et al., 2019), as confirmed in scavengers and omnivores (Shukla et al., 2016; Yadav et al., 2015; Yun et al., 2014). The generalist *Ceratitis* species had lowest gut microbial diversity (around two genus equivalents). Their relative specialist of Solanaceae, *N. cyanescens*, had slightly higher microbial diversity, noticeably due to a relatively high abundance of Morganellaceae. The specialists of Cucurbitaceae (*Dacus* and *Zeugodacus* species) had the highest microbial diversity (around five genus equivalents), whereas their relative generalists of genera *Bactrocera* had less diverse gut content (around four genus equivalents). The observation that fruit fly specialization does not significantly imprint gut microbial communities is rather a surprise (but see De Cock et al., 2020, for a first mention). Plants present numerous nutritional and defensive challenges to phytophagous insects. A growing body of research emphasizes the potential contribution of symbiotic microbes to phytophagous diets (Feldhaar, 2011; Felton & Tumlinson, 2008; Oliver et al., 2010). Nevertheless, the accumulated evidence is mixed and requires further sampling and functional analyses of the fruit fly gut microbiota. Gut microbiota respond more to host phylogeny rather than to host ecology in aphids (McLean et al., 2019) and in lycaenid butterflies (Whitaker et al., 2016). The reverse has been observed in both fruit-feeding and mycophagous drosophilid species (Adair et al., 2020), ants (Anderson et al., 2012), and beetles (Blankenchip et al., 2018).

In interaction with phylogeny, the environment of sampling (here field vs. laboratory) had a detectable moderate effect on gut communities, in terms of both diversity and composition. In most Dacinae species (all but *D. ciliatus*, the most recent laboratory population), laboratory populations had less diverse gut microbiota as compared to natural populations. This observation has been made repeatedly in Tephritidae, such as *B. tryoni* (Morrow et al., 2015), *B. oleae* (Ras et al., 2017), and a range of arthropod species (Belda et al., 2011; Ng et al., 2018; Pérez-Cobas et al., 2015; Staubach et al., 2013; Xiang et al., 2006). In clear contrast with these observations, we found that Ceratitinae laboratory populations were more diverse than field ones. Besides, except for *Dacus* samples, laboratory and field populations tended to differ strongly in terms of composition. In the present study, laboratory populations almost missed the class of Orbales (as already

observed by Martinson et al., 2017), several Enterobacterales and Lactobacillales. Laboratory samples were also less dominated by the genus *Klebsiella*. Three genera were only present in the laboratory (*Aeromonas*, *Fructobacillus*, and *Pluralibacter*), and some genera very rare in nature had important relative abundance in the laboratory, such as the Yersiniaceae *Serratia* and the Morganellaceae *Morganella*. Some differences between laboratory and field populations contrasted with former observations, for example, describing a dominance of *Providencia* or *Acinetobacter* in laboratory populations (Ben-Yosef et al., 2015; Kounatidis et al., 2009).

The laboratory populations are occasionally supplemented with field individuals so that this differentiation cannot be explained by pure drift. This suggests a genuine recomposition of gut microbiota in laboratory populations in response to local conditions (missing nutriments, antifungal treatment…). Interestingly, while laboratory populations did share very similar conditions, the constraint of phylogeny on microbial communities was still much apparent. Such a result has implications for fruit fly management strategies based on sterile insect techniques, as well as for ecological and evolutionary studies using laboratory populations. Many studies have mentioned a loss of competitiveness of laboratory flies vs. field individuals. It is possible that part of this lesser fitness is due to gut microbiota modifications, that it could be attenuated by working of microbiota restoration, and that the intensity of this effect is species-specific.

Overall, gut microbiota were strongly imprinted by fly phylogeny, but could be subject to important restructuring in the face of new environmental conditions. As a consequence, the observed lack of correlation between gut microbiota and both fly specialization and fly sampling environment is a surprise and needs to be addressed. It is possible that most gut microbes have functions other than fruit digestion (Ben-Ami et al., 2010; Cheng et al., 2017; Hadapad et al., 2016) or that there is functional redundancy; that is, microbial functions can be ensured by different taxa (Moya & Ferrer, 2016). Importantly only adults were studied here. In fruit flies, adults do not eat much, and only larvae feed on fruit. Yet, adult gut bacteria are the ones with a chance to be vertically transmitted. It could then be advantageous for flies that adults keep and transmit bacteria beneficial to larvae, including bacteria associated with plant use. Alternatively, it is possible that some useful gut bacteria are transitorily acquired by larvae in the fruit they grow in, before being eliminated at metamorphosis. Such ability to select and breed useful bacteria in the environment would confer an adaptive plastic response to host plants. These bacteria, which would likely differ across sampling environments, would not be detectable in studies focused on adults as here. And studying the contribution of gut microbiota to fly host range would require studying larvae as well. In Tephritidae, comparisons between larval and adult gut content are too rare and divergent for any conclusion to be drawn as to whether or not larvae acquire essential bacteria in the fruit, which would be released upon metamorphosis. Evidence from comparisons between larvae and fruits does not point toward this hypothesis. In *B. tryoni*, larval gut microbiota were more diverse than those of fruits and not influenced by fruit (Majunder et al., 2019). But in other flies, such as drosophilid flies, host ecology seems to have detectable impact on larval gut microbiota (Chandler et al., 2011). Another possible factor affecting gut microbiota composition and transmission might be the effect of larval diet on adult immunity (Fellous & Lazzaro, 2010). Adult immunity is likely the final gate filtering microbial taxa inherited by their progeny, and thus factors affecting immunity, including diet and other environmental conditions, could explain phylosymbiosis (or the lack thereof). Besides, the interactions between the host and a given microbe could be highly dependent on the other microbes constituting the microbiota. In such cases, a high rate of vertical transmission for a given microbe could greatly influence the rest of the microbiota. Dissecting the contribution of niche-based processes in the assembly of the gut microbiota is therefore still an important challenge for future research using both field samples and gnotobiotic animals in controlled conditions.

ACKNOWLEDGMENTS

We thank Jim Payet and Serge Glénac for clues as to where and when to sample wild fruit flies and for kindly providing samples from the laboratory populations they have been maintaining for years. Monique Mussard and Mathieu Parmanum helped with fly dissections during their internships. The authors are also grateful to Julien Chiquet and Sophie Donnet for their advice on block model clustering methods, and to Charlotte Mouden for her help on early versions of bioinformatic analyses. BF, CB, NB, EG, FM, and VR received the financial support of the French "Agence Nationale de la Recherche" project NGB (ANR-17-CE32-011). VR was funded by the European Union (European Regional Development Fund, ERDF contract GURDT I2016-1731-0006632), the Conseil Régional de La Réunion, and the Centre de Coopération Internationale en Recherche Agronomique pour le Développement (CIRAD). This study used the facilities provided by the Plant Protection Platform (3P, IBISA), Saint Pierre, Réunion, France. MinION sequencing was performed at the PGTB (grants from the Conseil Régional d'Aquitaine n° 20030304002FA and 20040305003FA, from the European Union FEDER n° 2003227, and from Investissements d'Avenir ANR-10-EQPX-16-01).

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

DATA AVAILABILITY STATEMENT

Sequences were deposited in the NCBI SRA database with under the Bio-Project PRJNA781104. Data and scripts are available online in the Zenodo repository using https://doi.org/10.5281/zenodo.5710318.

ORCID

Virginie Ravigné https://orcid.org/0000-0002-4252-2574
François Massol https://orcid.org/0000-0002-4098-955X
Frédéric Mahé https://orcid.org/0000-0002-2808-0984
Benoit Facon https://orcid.org/0000-0001-8970-6840

REFERENCES

Adair, K. L., Bost, A., Bueno, E., Kaunisto, S., Kortet, R., Peters-Schulze, G., Martinson, V. G., & Douglas, A. E. (2020). Host determinants of among-species variation in microbiome composition in drosophilid flies. *ISME Journal*, *14*(1), 217–229. https://doi.org/10.1038/s41396-019-0532-7

Aharon, Y., Pasternak, Z., Yosef, M. B., Behar, A., Lauzon, C., Yuval, B., & Jurkevitch, E. (2013). Phylogenetic, metabolic, and taxonomic diversities shape Mediterranean fruit fly microbiotas during ontogeny. *Applied and Environmental Microbiology*, *79*(1), 303–313. https://doi.org/10.1128/AEM.02761-12

Anderson, K. E., Russell, J. A., Moreau, C. S., Kautz, S., Sullam, K. E., Hu, Y., Basinger, U., Mott, B. M., Buck, N., & Wheeler, D. E. (2012). Highly similar microbial communities are shared among related and trophically similar ant species. *Molecular Ecology*, *21*(9), 2282–2296. https://doi.org/10.1111/j.1365-294X.2011.05464.x

Andongma, A. A., Wan, L., Dong, Y. C., Li, P., Desneux, N., White, J. A., & Niu, C.-Y. (2015). Pyrosequencing reveals a shift in symbiotic bacteria populations across life stages of *Bactrocera dorsalis*. *Scientific Reports*, *5*, 9470. https://doi.org/10.1038/srep09470

Astegiano, J., Altermatt, F., & Massol, F. (2017). Disentangling the co-structure of multilayer interaction networks: Degree distribution and module composition in two-layer bipartite networks. *Scientific Reports*, *7*, 15465. https://doi.org/10.1038/s41598-017-15811-w

Barberán, A., Bates, S. T., Casamayor, E. O., & Fierer, N. (2012). Using network analysis to explore co-occurrence patterns in soil microbial communities. *ISME Journal*, *6*(2), 343–351. https://doi.org/10.1038/ismej.2011.119

Behar, A., Jurkevitch, E., & Yuval, B. (2008). Bringing back the fruit into fruit fly-bacteria interactions. *Molecular Ecology*, *17*(5), 1375–1386. https://doi.org/10.1111/j.1365-294X.2008.03674.x

Behar, A., Yuval, B., & Jurkevitch, E. (2005). Enterobacteria-mediated nitrogen fixation in natural populations of the fruit fly *Ceratitis capitata*. *Molecular Ecology*, *14*(9), 2637–2643. https://doi.org/10.1111/j.1365-294X.2005.02615.x

Behar, A., Yuval, B., & Jurkevitch, E. (2008). Gut bacterial communities in the Mediterranean fruit fly (*Ceratitis capitata*) and their impact on host longevity. *Journal of Insect Physiology*, *54*, 1377–1383. https://doi.org/10.1016/j.jinsphys.2008.07.011

Belda, E., Pedrola, L., Peretó, J., Martínez-Blanch, J. F., Montagud, A., Navarro, E., Urchueguia, J., Ramon, D., Moya, A., & Porcar, M. (2011). Microbial diversity in the midguts of field and lab-reared populations of the European Corn Borer *Ostrinia nubilalis*. *PLoS One*, *6*(6), e21751. https://doi.org/10.1371/journal.pone.0021751

Ben-Ami, E., Yuval, B., & Jurkevitch, E. (2010). Manipulation of the microbiota of mass-reared Mediterranean fruit flies *Ceratitis capitata* (Diptera: Tephritidae) improves sterile male sexual performance. *ISME Journal*, *4*, 28–37. https://doi.org/10.1038/ismej.2009.82

Ben-Yosef, M., Pasternak, Z., Jurkevitch, E., & Yuval, B. (2014). Symbiotic bacteria enable olive flies (*Bactrocera oleae*) to exploit intractable sources of nitrogen. *Journal of Evolutionary Biology*, *27*(12), 2695–2705. https://doi.org/10.1111/jeb.12527

Ben-Yosef, M., Pasternak, Z., Jurkevitch, E., & Yuval, B. (2015). Symbiotic bacteria enable olive fly larvae to overcome host defences. *Royal Society Open Science*, *2*, 150170. https://doi.org/10.1098/rsos.150170

Biernacki, C., Celeux, G., & Govaert, G. (2000). Assessing a mixture model for clustering with the integrated completed likelihood. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *22*(7), 719–725. https://doi.org/10.1109/34.865189

Blankenchip, C. L., Michels, D. E., Elizabeth Braker, H., & Goffredi, S. K. (2018). Diet breadth and exploitation of exotic plants shift the core microbiome of Cephaloleia, a group of tropical herbivorous beetles. *PeerJ*, *6*, e4793. https://doi.org/10.7717/peerj.4793

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., Alexander, H., Alm, E. J., Arumugam, M., Asnicar, F., Bai, Y., Bisanz, J. E., Bittinger, K., Brejnrod, A., Brislawn, C. J., Brown, C. T., Callahan, B. J., Caraballo-Rodríguez, A. M., Chase, J., … Caporaso, J. G. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature Biotechnology*, *37*(8), 852–857. https://doi.org/10.1038/s41587-019-0209-9

Bray, J. R., & Curtis, J. T. (1957). An ordination of upland forest communities of southern Wisconsin. *Ecological Monographs*, *27*, 325–349.

Broderick, N. A., & Lemaitre, B. (2012). Gut-associated microbes of *Drosophila melanogaster*. *Gut Microbes*, *3*(4), 307–321. https://doi.org/10.4161/gmic.19896

Brooks, A. W., Kohl, K. D., Brucker, R. M., van Opstal, E. J., & Bordenstein, S. R. (2016). Phylosymbiosis: Relationships and functional effects of microbial communities across host evolutionary history. *PLoS Biology*, *14*(11), 1–29. https://doi.org/10.1371/journal.pbio.2000225

Brucker, R. M., & Bordenstein, S. R. (2012). The roles of host evolutionary relationships (genus: *Nasonia*) and development in structuring microbial communities. *Evolution*, *66*(2), 349–362. https://doi.org/10.1111/j.1558-5646.2011.01454.x

Burns, A. R., Stephens, W. Z., Stagaman, K., Wong, S., Rawls, J. F., Guillemin, K., & Bohannan, B. J. M. (2016). Contribution of neutral processes to the assembly of gut microbial communities in the zebrafish over host development. *ISME Journal*, *10*(3), 655–664. https://doi.org/10.1038/ismej.2015.142

Capuzzo, C., Firrao, G., Mazzon, L., Squartini, A., & Girolami, V. (2005). '*Candidatus Erwinia dacicola*', a coevolved symbiotic bacterium of the olive fly *Bactrocera oleae* (Gmelin). *International Journal of Systematic and Evolutionary Microbiology*, *55*(4), 1641–1647. https://doi.org/10.1099/ijs.0.63653-0

Ceja-Navarro, J. A., Vega, F. E., Karaoz, U., Hao, Z., Jenkins, S., Lim, H. C., Kosina, P., Infante, F., Northen, T. R., & Brodie, E. L. (2015). Gut microbiota mediate caffeine detoxification in the primary insect pest of coffee. *Nature Communications*, *6*, 7618. https://doi.org/10.1038/ncomms8618

Chandler, J. A., Lang, J., Bhatnagar, S., Eisen, J. A., & Kopp, A. (2011). Bacterial communities of diverse Drosophila species: Ecological context of a host-microbe model system. *PLoS Genetics*, *7*(9), e1002272. https://doi.org/10.1371/journal.pgen.1002272

Cheng, D., Guo, Z., Riegler, M., Xi, Z., Liang, G., & Xu, Y. (2017). Gut symbiont enhances insecticide resistance in a significant pest, the oriental fruit fly *Bactrocera dorsalis* (Hendel). *Microbiome*, *5*(1), 13. https://doi.org/10.1186/s40168-017-0236-z

Chiquet, J., Donnet, S., & Barbillon, P. (2021). *sbm: Stochastic Block models*. R package version 0.4.0-9250. https://grosssbm.github.io/sbm/

Crotti, E., Balloi, A., Hamdi, C., Sansonno, L., Marzorati, M., Gonella, E., Favia, G., Cherif, A., Bandi, C., Alma, A., & Daffonchio, D. (2012). Microbial symbionts: A resource for the management of insect-related problems. *Microbial Biotechnology*, *5*(3), 307–317. https://doi.org/10.1111/j.1751-7915.2011.00312.x

Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal*, *Complex Systems*, 1695. https://igraph.org

De Cock, M., Virgilio, M., Vandamme, P., Bourtzis, K., De Meyer, M., & Willems, A. (2020). Comparative microbiomics of Tephritid frugivorous pests (Diptera: Tephritidae) from the field: A tale of high variability across and within species. *Frontiers in Microbiology*, *11*, 1890. https://doi.org/10.3389/fmicb.2020.01890

De Coster, W., D'Hert, S., Schultz, D. T., Cruts, M., & Van Broeckhoven, C. (2018). NanoPack: Visualizing and processing long-read sequencing data. *Bioinformatics*, *34*(15), 2666–2669. https://doi.org/10.1093/bioinformatics/bty149

Deb, R., Nair, A., & Agashe, D. (2019). Host dietary specialization and neutral assembly shape gut bacterial communities of wild dragonflies. *PeerJ*, *7*, e8058. https://doi.org/10.7717/peerj.8058

Deutscher, A. T., Burke, C. M., Darling, A. E., Riegler, M., Reynolds, O. L., & Chapman, T. A. (2018). Near full-length 16S rRNA gene

next-generation sequencing revealed Asaia as a common midgut bacterium of wild and domesticated Queensland fruit fly larvae. *Microbiome*, *6*(1), 85. https://doi.org/10.1186/s40168-018-0463-y

Deutscher, A. T., Chapman, T. A., Shuttleworth, L. A., Riegler, M., & Reynolds, O. L. (2019). Tephritid-microbial interactions to enhance fruit fly performance in sterile insect technique programs. *BMC Microbiology*, *19*(Suppl 1), 287. https://doi.org/10.1186/s12866-019-1650-0

Engel, P., & Moran, N. A. (2013). The gut microbiota of insects – Diversity in structure and function. *FEMS Microbiology Reviews*, *37*, 699–735. https://doi.org/10.1111/1574-6976.12025

Erlandson, S., Wei, X., Savage, J., Cavender-Bares, J., & Peay, K. (2018). Soil abiotic variables are more important than Salicaceae phylogeny or habitat specialization in determining soil microbial community structure. *Molecular Ecology*, *27*(8), 2007–2024. https://doi.org/10.1111/mec.14576

Facon, B., Hafsi, A., Charley de la Masselière, M., Robin, S., Massol, F., Dubart, M., Chiquet, J., Frago, E., Chiroleu, F., Duyck, P. F., & Ravigné, V. (2021). Joint species distributions reveal the combined effects of host plants, abiotic factors and species competition as drivers of community structure in fruit flies. *Ecology Letters*, *24*, 1905–1916. https://doi.org/10.1101/2020.12.07.414326

Feldhaar, H. (2011). Bacterial symbionts as mediators of ecologically important traits of insect hosts. *Ecological Entomology*, *36*(5), 533–543. https://doi.org/10.1111/j.1365-2311.2011.01318.x

Fellous, S., & Lazzaro, B. P. (2010). Larval food quality affects adult (but not larval) immune gene expression independent of effects on general condition. *Molecular Ecology*, *19*, 1462–1468. https://doi.org/10.1111/j.1365-294X.2010.04567.x

Felton, G. W., & Tumlinson, J. H. (2008). Plant–insect dialogs: Complex interactions at the plant–insect interface. *Current Opinion in Plant Biology*, *11*(4), 457–463. https://doi.org/10.1016/j.pbi.2008.07.001

Gaba, S., Bretagnolle, F., Rigaud, T., & Philippot, L. (2014). Managing biotic interactions for ecological intensification of agroecosystems. *Frontiers in Ecology and Evolution*, *2*, 29. https://doi.org/10.3389/fevo.2014.00029

Gilligan, C. A. (2008). Sustainable agriculture and plant diseases: An epidemiological perspective. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1492), 741–759. https://doi.org/10.1098/rstb.2007.2181

Hadapad, A. B., Prabhakar, C. S., Chandekar, S. C., Tripathi, J., & Hire, R. S. (2016). Diversity of bacterial communities in the midgut of *Bactrocera cucurbitae* (Diptera: Tephritidae) populations and their potential use as attractants. *Pest Management Science*, *72*(6), 1222–1230. https://doi.org/10.1002/ps.4102

Hill, M. O. (1973). Diversity and evenness: A unifying notation and its consequences. *Ecology*, *54*(2), 427–432. https://doi.org/10.2307/1934352

Hsieh, T. C., Ma, K. H., & Chao, A. (2016). iNEXT: An R package for rarefaction and extrapolation of species diversity (Hill numbers). *Methods in Ecology and Evolution*, *7*(12), 1451–1456. https://doi.org/10.1111/2041-210X.12613

Ivens, A. B. F., Gadau, A., Kiers, E. T., & Kronauer, D. J. C. (2018). Can social partnerships influence the microbiome? Insights from ant farmers and their trophobiont mutualists. *Molecular Ecology*, *27*(8), 1898–1914. https://doi.org/10.1111/mec.14506

Joffard, N., Massol, F., Grenié, M., Montgelard, C., & Schatz, B. (2019). Effect of pollination strategy, phylogeny and distribution on pollination niches of Euro-Mediterranean orchids. *Journal of Ecology*, *107*(1), 478–490. https://doi.org/10.1111/1365-2745.13013

Jost, L. (2006). Entropy and diversity. *Oikos*, *113*(2), 363–375. https://doi.org/10.1111/j.2006.0030-1299.14714.x

Jost, L. (2007). Partitioning diversity into independent alpha and beta components. *Ecology*, *88*(10), 2427–2439. https://doi.org/10.1890/06-1736.1

Kohl, K. D., Dearing, M. D., & Bordenstein, S. R. (2018). Microbial communities exhibit host species distinguishability and phylosymbiosis along the length of the gastrointestinal tract. *Molecular Ecology*, *27*(8), 1874–1883. https://doi.org/10.1111/mec.14460

Kolodny, O., Callahan, B. J., & Douglas, A. E. (2020). The role of the microbiome in host evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*, 20190588. https://doi.org/10.1098/rstb.2019.0588

Kounatidis, I., Crotti, E., Sapountzis, P., Sacchi, L., Rizzi, A., Chouaia, B., Bandi, C., Alma, A., Daffonchio, D., Mavragani-Tsipidou, P., & Bourtzis, K. (2009). *Acetobacter tropicalis* is a major symbiont of the olive fruit fly (*Bactrocera oleae*). *Applied and Environmental Microbiology*, *75*(10), 3281–3288. https://doi.org/10.1128/AEM.02933-08

Lauzon, C. R., McCombs, S. D., Potter, S. E., & Peabody, N. C. (2009). Establishment and vertical passage of *Enterobacter* (*Pantoea*) *agglomerans* and *Klebsiella pneumoniae* through all life stages of the Mediterranean fruit fly (Diptera: Tephritidae). *Annals of the Entomological Society of America*, *102*(1), 85–95. https://doi.org/10.1603/008.102.0109

Leger, J.-B., Daudin, J.-J., & Vacher, C. (2015). Clustering methods differ in their ability to detect patterns in ecological networks. *Methods in Ecology and Evolution*, *6*, 474–481. https://doi.org/10.1111/2041-210X.12334

Liu, L. J., Martinez-Sañudo, I., Mazzon, L., Prabhakar, C. S., Girolami, V., Deng, Y. L., Dai, Y., & Li, Z. H. (2016). Bacterial communities associated with invasive populations of *Bactrocera dorsalis* (Diptera: Tephritidae) in China. *Bulletin of Entomological Research*, *106*(6), 718–728. https://doi.org/10.1017/S0007485316000390

Liu, S.-H., Chen, Y., Li, W., Tang, G.-H., Yang, Y., Jiang, H.-B., Dou, W., & Wang, J.-J. (2018). Diversity of bacterial communities in the intestinal tracts of two geographically distant populations of *Bactrocera dorsalis* (Diptera: Tephritidae). *Journal of Economic Entomology*, *111*, 2861–2868. https://doi.org/10.1093/jee/toy231

MacCollom, G. B., Lauzon, C. R., Sjogren, R. E., Meyer, W. L., & Olday, F. (2009). Association and attraction of blueberry maggot fly Curran (Diptera: Tephritidae) to *Pantoea* (*Enterobacter*) *agglomerans*. *Environmental Entomology*, *38*(1), 116–120. https://doi.org/10.1603/022.038.0114

Macke, E., Tasiemski, A., Massol, F., Callens, M., & Decaestecker, E. (2017). Life history and eco-evolutionary dynamics in light of the gut microbiota. *Oikos*, *126*(4), 508–531. https://doi.org/10.1111/oik.03900

Majunder, R., Sutcliffe, B., Taylor, P. W., & Chapman, T. A. (2019). Next-generation sequencing reveals relationship between the larval microbiome and food substrate in the polyphagous Queensland fruit fly. *Scientific Reports*, *9*, 14292. https://doi.org/10.1038/s41598-019-50602-5

Malacrinò, A., Campolo, O., Medina, R. F., & Palmeri, V. (2018). Instar- and host-associated differentiation of bacterial communities in the Mediterranean fruit fly *Ceratitis capitata*. *PLoS One*, *13*(3), e0194131. https://doi.org/10.1371/journal.pone.0194131

Martinson, V. G., Douglas, A. E., & Jaenike, J. (2017). Community structure of the gut microbiota in sympatric species of wild Drosophila. *Ecology Letters*, *20*(5), 629–639. https://doi.org/10.1111/ele.12761

Massol, F., Macke, E., Callens, M., & Decaestecker, E. (2021). A methodological framework to analyse determinants of host–microbiota networks, with an application to the relationships between Daphnia magna's gut microbiota and bacterioplankton. *Journal of Animal Ecology*, *90*(1), 102–119. https://doi.org/10.1111/1365-2656.13297

Massol, F., & Petit, S. (2013). Interaction networks in agricultural landscape mosaics. *Advances in Ecological Research*, *49*, 291–338. https://doi.org/10.1016/B978-0-12-420002-9.00005-6

Matsuo, Y., Komiya, S., Yasumizu, Y., Yasuoka, Y., Mizushima, K., Takagi, T., & Hirota, K. (2021). Full-length 16S rRNA gene amplicon analysis of human gut microbiota using MinION™ nanopore sequencing confers species-level resolution. *BMC Microbiology*, *21*, 35. https://doi.org/10.1186/s12866-021-02094-5

McLean, A. H. C., Godfray, H. C. J., Ellers, J., & Henry, L. M. (2019). Host relatedness influences the composition of aphid microbiomes. *Environmental Microbiology Reports*, *11*(6), 808–816. https://doi.org/10.1111/1758-2229.12795

McMurdie, P. J., & Holmes, S. (2013). Phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One*, *8*(4), e61217. https://doi.org/10.1371/journal.pone.0061217

Moquet, L., Payet, J., Glénac, S., & Delatte, H. (2021). Niche shift of tephritid species after the Oriental fruit fly (*Bactrocera dorsalis*) invasion in La Réunion. *Diversity and Distributions*, *27*(1), 109–129. https://doi.org/10.1111/ddi.13172

Moran, N. A., Ochman, H., & Hammer, T. J. (2019). Evolutionary and ecological consequences of gut microbial communities. *Annual Review of Ecology Evolution and Systematics*, *50*(1), 451–475. https://doi.org/10.1146/annurev

Moran, N. A., & Sloan, D. B. (2015). The hologenome concept: Helpful or hollow? *PLoS Biology*, *13*(12), e1002311. https://doi.org/10.1371/journal.pbio.1002311

Morrow, J. L., Frommer, M., Shearman, D. C. A., & Riegler, M. (2015). The microbiome of field-caught and laboratory-adapted Australian Tephritid fruit fly species with different host plant use and specialisation. *Microbial Ecology*, *70*(2), 498–508. https://doi.org/10.1007/s00248-015-0571-1

Moya, A., & Ferrer, M. (2016). Functional redundancy-induced stability of gut microbiota subjected to disturbance. *Trends in Microbiology*, *24*(5), 402–413. https://doi.org/10.1016/j.tim.2016.02.002

M'Saad Guerfali, M., Djobbi, W., Charaabi, K., Hamden, H., Fadhl, S., Marzouki, W., Dhaouedi, F., & Chevrier, C. (2018). Evaluation of *Providencia rettgeri* pathogenicity against laboratory Mediterranean fruit fly strain (*Ceratitis capitata*). *PLoS One*, *13*, e0196343. https://doi.org/10.1371/journal.pone.0196343

Mushegian, A. A., & Ebert, D. (2016). Rethinking "mutualism" in diverse host-symbiont communities. *BioEssays*, *38*, 100–108. https://doi.org/10.1002/bies.201500074

Newman, M. E. J. (2006). Finding community structure in networks using the eigenvectors of matrices. *Physical Review E*, *74*, 036104. https://doi.org/10.1103/PhysRevE.74.036104

Ng, S. H., Stat, M., Bunce, M., & Simmons, L. W. (2018). The influence of diet and environment on the gut microbial community of field crickets. *Ecology and Evolution*, *8*(9), 4704–4720. https://doi.org/10.1002/ece3.3977

Nishida, A. H., & Ochman, H. (2018). Rates of gut microbiome divergence in mammals. *Molecular Ecology*, *27*(8), 1884–1897. https://doi.org/10.1111/mec.14473

Noman, M. S., Liu, L., Bai, Z., & Li, Z. (2020). Tephritidae bacterial symbionts: Potentials for pest management. *Bulletin of Entomological Research*, *110*(1), 1–14. https://doi.org/10.1017/S0007485319000403

Nougué, O., Gallet, R., Chevin, L.-M., & Lenormand, T. (2015). Niche limits of symbiotic gut microbiota constrain the salinity tolerance of brine shrimp. *The American Naturalist*, *186*, 390–403. https://doi.org/10.1086/682370

Nygaard, A. B., Tunsjø, H. S., Meisal, R., & Charnock, C. (2020). A preliminary study on the potential of Nanopore MinION and Illumina MiSeq 16S rRNA gene sequencing to characterize building-dust microbiomes. *Scientific Reports*, *10*(1), 3209. https://doi.org/10.1038/s41598-020-59771-0

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., Szoecs, E., & Wagner, H. (2020). *vegan: Community ecology package*. R package version 2.5-7. https://CRAN.R-project.org/package=vegan

Oliver, K. M., Degnan, P. H., Burke, G. R., & Moran, N. A. (2010). Facultative symbionts in aphids and the horizontal transfer of ecologically important traits. *Annual Review of Entomology*, *55*, 247–266. https://doi.org/10.1146/annurev-ento-112408-085305

Peres-Neto, P. R., Legendre, P., Dray, S., & Borcard, D. (2006). Variation partitioning of species data matrices: Estimation and comparison of fractions. *Ecology*, *87*(10), 2614–2625. https://doi.org/10.1890/0012-9658(2006)87[2614:VPOSDM]2.0.CO;2

Pérez-Cobas, A. E., Maiques, E., Angelova, A., Carrasco, P., Moya, A., & Latorre, A. (2015). Diet shapes the gut microbiota of the omnivorous cockroach *Blattella germanica*. *FEMS Microbiology Ecology*, *91*(4), 1–14. https://doi.org/10.1093/femsec/fiv022

Pimentel, D., Stachow, U., Takacs, D. A., Brubaker, H. W., Dumas, A. R., Meaney, J. J., O'Neil, J. A. S., Onsi, D. E., & Cozilius, D. B. (1992). Conserving biological diversity in agricultural forestry systems – Most biological diversity exists in human-managed ecosystems. *BioScience*, *42*(5), 354–362. https://doi.org/10.2307/1311782

Pollock, J., Glendinning, L., Wisedchanwet, T., & Watson, M. (2018). The madness of microbiome: Attempting to Find consensus "best practice" for 16S microbiome studies. *Applied and Environmental Microbiology*, *84*(7), e02627-17. https://doi.org/10.1128/AEM.02627-17

Qin, Y., Paini, D. R., Wang, C., Fang, Y., & Li, Z. (2015). Global establishment risk of economically important fruit fly species (Tephritidae). *PLoS One*, *10*(1), e0116424. https://doi.org/10.1371/journal.pone.0116424

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., & Glöckner, F. O. (2013). The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Research*, *41*(D1), 590–596. https://doi.org/10.1093/nar/gks1219

Quilici, S., & Jeuffrault, E. (2001). *Plantes-hôtes des mouches des fruits: Maurice, Réunion, Seychelles*. PRMF/COI, Imp. Graphica.

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Ras, E., Beukeboom, L. W., Cáceres, C., & Bourtzis, K. (2017). Review of the role of gut microbiota in mass rearing of the olive fruit fly, *Bactrocera oleae*, and its parasitoids. *Entomologia Experimentalis et Applicata*, *164*(3), 237–256. https://doi.org/10.1111/eea.12609

Raza, M. F., Yao, Z., Bai, S., Cai, Z., & Zhang, H. (2020). Tephritidae fruit fly gut microbiome diversity, function and potential for applications. *Bulletin of Entomological Research*, *110*(4), 423–437. https://doi.org/10.1017/S0007485319000853

Rognes, T., Flouri, T., Nichols, B., Quince, C., & Mahé, F. (2016). VSEARCH: A versatile open source tool for metagenomics. *PeerJ*, *4*, e2584. https://doi.org/10.7717/peerj.2584

Salas, B., Conway, H. E., Schuenzel, E. L., Hopperstad, K., Vitek, C., & Vacek, D. C. (2017). *Morganella morganii* (Enterobacteriales: Enterobacteriaceae) is a lethal pathogen of Mexican fruit fly (Diptera: Tephritidae) larvae. *Florida Entomologist*, *100*, 743–751. https://doi.org/10.1653/024.100.0422

Shapira, M. (2016). Gut microbiotas and host evolution: Scaling up symbiosis. *Trends in Ecology and Evolution*, *31*(7), 539–549. https://doi.org/10.1016/j.tree.2016.03.006

Shukla, S. P., Sanders, J. G., Byrne, M. J., & Pierce, N. E. (2016). Gut microbiota of dung beetles correspond to dietary specializations of adults and larvae. *Molecular Ecology*, *25*(24), 6092–6106. https://doi.org/10.1111/mec.13901

Spor, A., Koren, O., & Ley, R. (2011). Unravelling the effects of the environment and host genotype on the gut microbiome. *Nature Reviews Microbiology*, *9*(4), 279–290. https://doi.org/10.1038/nrmicro2540

Staubach, F., Baines, J. F., Künzel, S., Bik, E. M., & Petrov, D. A. (2013). Host species and environmental effects on bacterial communities associated with Drosophila in the laboratory and in the natural environment. *PLoS One*, *8*(8), e70749. https://doi.org/10.1371/journal.pone.0070749

Strona, G., Nappo, D., Boccacci, F., Fattorini, S., & San-Miguel-Ayanz, J. (2014). A fast and unbiased procedure to randomize ecological binary matrices with fixed row and column totals. *Nature Communications*, *5*, 4114. https://doi.org/10.1038/ncomms5114

Ter Braak, C. J. F. (1986). Canonical correspondence analysis: A new eigenvector technique for multivariate direct gradient analysis. *Ecology*, *67*(5), 1167–1179. https://doi.org/10.2307/1938672

Thrall, P. H., Oakeshott, J. G., Fitt, G., Southerton, S., Burdon, J. J., Sheppard, A., Russell, R. J., Zalucki, M., Heino, M., & Ford Denison, R. (2011). Evolution in agriculture: The application of evolutionary approaches to the management of biotic interactions in agro-ecosystems. *Evolutionary Applications*, *4*(2), 200–215. https://doi.org/10.1111/j.1752-4571.2010.00179.x

Tilman, D., Balzer, C., Hill, J., & Befort, B. L. (2011). Global food demand and the sustainable intensification of agriculture. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(50), 20260–20264. https://doi.org/10.1073/pnas.1116437108

Ventura, C., Briones-Roblero, C. I., Hernández, E., Rivera-Orduña, F. N., & Zúñiga, G. (2018). Comparative analysis of the gut bacterial community of four Anastrepha fruit flies (Diptera: Tephritidae) based on pyrosequencing. *Current Microbiology*, *75*(8), 966–976. https://doi.org/10.1007/s00284-018-1473-5

Virgilio, M., Jordaens, K., Verwimp, C., White, I. M., & De Meyer, M. (2015). Higher phylogeny of frugivorous flies (Diptera, Tephritidae, Dacini): Localised partition conflicts and a novel generic classification. *Molecular Phylogenetics and Evolution*, *85*, 171–179. https://doi.org/10.1016/j.ympev.2015.01.007

Whitaker, M. R. L., Salzman, S., Sanders, J., Kaltenpoth, M., & Pierce, N. E. (2016). Microbial communities of lycaenid butterflies do not correlate with larval diet. *Frontiers in Microbiology*, *7*, 1920. https://doi.org/10.3389/fmicb.2016.01920

Xiang, H., Wei, G. F., Jia, S., Huang, J., Miao, X. X., Zhou, Z., Zhao, L. P., & Huang, Y. P. (2006). Microbial communities in the larval midgut of laboratory and field populations of cotton bollworm (*Helicoverpa armigera*). *Canadian Journal of Microbiology*, *52*(11), 1085–1092. https://doi.org/10.1139/w06-064

Yadav, K. K., Bora, A., Datta, S., Chandel, K., Gogoi, H. K., Prasad, G. B. K. S., & Veer, V. (2015). Molecular characterization of midgut microbiota of *Aedes albopictus* and *Aedes aegypti* from Arunachal Pradesh, India. *Parasites and Vectors*, *8*, 641. https://doi.org/10.1186/s13071-015-1252-0

Yilmaz, P., Parfrey, L. W., Yarza, P., Gerken, J., Pruesse, E., Quast, C., Schweer, T., Peplies, J., Ludwig, W., & Glöckner, F. O. (2014). The SILVA and "all-species Living Tree Project (LTP)" taxonomic frameworks. *Nucleic Acids Research*, *42*(D1), 643–648. https://doi.org/10.1093/nar/gkt1209

Yong, H. S., Song, S. L., Chua, K. O., & Lim, P. E. (2017). High diversity of bacterial communities in developmental stages of *Bactrocera carambolae* (Insecta: Tephritidae) revealed by Illumina MiSeq sequencing of 16S rRNA gene. *Current Microbiology*, *74*, 1076–1082. https://doi.org/10.1007/s00284-017-1287-x

Yun, J.-H., Roh, S. W., Whon, T. W., Jung, M.-J., Kim, M.-S., Park, D.-S., Yoon, C., Nam, Y.-D., Kim, Y.-J., Choi, J.-H., Kim, J.-Y., Shin, N.-R., Kim, S.-H., Lee, W.-J., & Bae, J.-W. (2014). Insect gut bacterial diversity determined by environmental habitat, diet, developmental stage, and phylogeny of host. *Applied and Environmental Microbiology*, *80*, 5254–5264. https://doi.org/10.1128/AEM.01226-14

Zhang, B., Liu, Y. H., Wu, W. X., & Wang, Z. L. (2010). Molecular phylogeny of *Bactrocera* species (Diptera: Tephritidae: Dacini) inferred from mitochondrial sequences of 16s rDNA and COI sequences. *The Florida Entomologist*, *93*(3), 369–377.

Zhao, P., Irwin, D. M., & Dong, D. (2016). Host genetics is associated with the gut microbial community membership rather than the structure. *Molecular BioSystems*, *12*, 1676–1686. https://doi.org/10.1039/C5MB00850F

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Ravigné, V., Becker, N., Massol, F., Guichoux, E., Boury, C., Mahé, F., & Facon, B. (2022). Fruit fly phylogeny imprints bacterial gut microbiota. *Evolutionary Applications*, 00, 1–18. https://doi.org/10.1111/eva.13352