Genetica (2011) 139:541–550
DOI 10.1007/s10709-010-9498-0

SI - GOS

Comparative analysis of microbial diversity in *Longitarsus* flea beetles (Coleoptera: Chrysomelidae)

Scott T. Kelley · Susanne Dobler

Received: 30 April 2010 / Accepted: 4 September 2010 / Published online: 16 September 2010
© Springer Science+Business Media B.V. 2010

Abstract Herbivorous beetles comprise a significant fraction of eukaryotic biodiversity and their plant-feeding adaptations make them notorious agricultural pests. Despite more than a century of research on their ecology and evolution, we know little about the diversity and function of their symbiotic microbial communities. Recent culture-independent molecular studies have shown that insects possess diverse gut microbial communities that appear critical for their survival. In this study, we combined culture-independent methods and high-throughput sequencing strategies to perform a comparative analysis of *Longitarsus* flea-beetles microbial community diversity (MCD). This genus of beetle herbivores contains host plant specialists and generalists that feed on a diverse array of toxic plants. Using a deep-sequencing approach, we characterized the MCD of eleven *Longitarsus* species across the genus, several of which represented independent shifts to the same host plant families. Database comparisons found that *Longitarsus*-associated microbes came from two habitat types: insect guts and the soil rhizosphere. Statistical clustering of the *Longitarsus* microbial communities found little correlation with the beetle phylogeny, and uncovered discrepancies between bacterial communities extracted directly from beetles and those from frass. A Principal Coordinates Analysis also found some correspondence between beetle MCD and host plant family. Collectively, our data suggest that environmental factors play a dominant role in shaping *Longitarsus* MCD and that the root-feeding beetle larvae of these insects are inoculated by soil rhizosphere microbes. Future studies will investigate MCD of select *Longitarsus* species across their geographic ranges and explore the connection between the soil rhizosphere and the beetle MCD.

Keywords 16S rRNA · Bacteria · Biodiversity · Herbivory · Metagenomics · Microbial ecology · Phylogeny

S. T. Kelley (✉)
Department of Biology, San Diego State University,
5500 Campanile Drive, San Diego, CA 92182, USA
e-mail: skelley@sciences.sdsu.edu

S. Dobler
Institute of Zoology, University of Hamburg,
Martin-Luther-King-Platz 3, 20146 Hamburg, Germany
e-mail: susanne.dobler@uni-hamburg.de

Introduction

Herbivorous beetles comprise one of the most abundant groups of multi-cellular organisms on the planet and contribute significantly to overall biodiversity (Erwin 1982; Novotny et al. 2006). Their plant-feeding habits are associated with fascinating behavioral and physiological adaptations (Bernays and Chapman 1994), which can also make them devastating pests of agricultural and forestry crops (Haubruge and Arnaud 2001; Paine et al. 1997; Sexson and Wyman 2005; Wood 1982). The evolutionary radiation of herbivorous beetles coincided with the diversification of terrestrial plants and appears to have been influenced by the evolution of plant secondary compounds (Farrell 1998; Farrell and Mitter 1998; Futuyma and Scheffer 1993; Mitter et al. 1988). Over time, many beetle herbivores became highly specialized in their diets, feeding on a small number of related plants (Bernays and Graham 1988; Mitter et al. 1988; Mitter and Futuyma 1991). This degree of specialization appears to have had dramatic macroevolutionary consequences, accelerating rates of speciation and often restricting evolutionary host-shifts to

Tecan Ex. 2179, Page 1 of 11
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

related or chemically similar plants (Dobler and Farrell 1999; Ehrlich and Raven 1964; Farrell 1998; Futuyma and Moreno 1988; Futuyma and Scheffer 1993; Kelley and Farrell 1998).

Although numerous studies have investigated the molecular and physiological adaptations of insect herbivores to host plant secondary compounds, very little is known about the role gut microbial symbionts may play in this process. This is largely because of the difficulties inherent in culturing microbes from environmental samples, including animal host tissues. Studies of environmental microbial communities have estimated that we have cultured *less than one percent of the true microbial diversity* (Amann et al. 1995; Harris et al. 2004; Hugenholtz et al. 1998a; Hugenholtz and Pace 1996; Pace 1997). However, the development of culture-independent molecular techniques, based primarily on PCR and molecular cloning of small subunit (16S) ribosomal RNA gene sequences (Amann et al. 1995; Hugenholtz et al. 1998a; Pace et al. 1985), has revolutionized our ability to study previously uncultured microbes in an enormous range of environments, such as rainforest soils (Rondon et al. 1999), geothermal springs (Hugenholtz et al. 1998b), saturated salt pools (Ley et al. 2006a), and animal intestines (Andersson et al. 2008; Backhed et al. 2005; Suh et al. 2005). Not only are researchers discovering vast numbers of novel microorganisms, orders of magnitude more than previously thought (Lozupone and Knight 2007; Pace 1997; Rondon et al. 1999; Tringe and Hugenholtz 2008; Tringe et al. 2005; Warnecke et al. 2007), we are also discovering the critical roles microbes play in virtually every ecological setting. For instance, a steadily growing list of culture-independent and metagenomic studies have shown that animals harbor extremely complex communities of microorganisms (aka., "microbiota") (Ley et al. 2005; McKenna et al. 2008; Safaee et al. 2006; Suh et al. 2005; Warnecke et al. 2007). These gut microbes are critical for proper nutrition, immunity and development (Haverson et al. 2007; Ley et al. 2008; Rhee et al. 2004; Visotto et al. 2009; Warnecke et al. 2007).

Culture-independent studies of microbial communities associated with insects, including ants (Van Borm et al. 2002), beetles (Delalibera et al. 2005; Suh et al. 2005), moths (Broderick et al. 2004), termites (Warnecke et al. 2007), and flies (Behar et al. 2008b) have also discovered complex microbiota that appear to affect insect ecological adaptation. Recent studies of the Mediterranean fruit fly and the gypsy moth have shown that complex microbial communities can play important roles in the adaptation of insects to plant feeding. Female Med-flies vertically transfer bacteria during oviposition and these bacteria express enzymes that start the fruit rotting process and allow the larvae to feed and grow in the fruit (Behar 2005, 2008a, b). In gypsy moths, Broderick et al. (2004) used culture-independent methods to show that host plants had a dramatic effect on the insect gut microbial community diversity (MCD). Collectively, these studies of insects suggest that highly diverse gut microbial communities are involved in the adaptation of herbivores to host plants. However, the precise nature of insect-microbial interactions and how they might affect insect herbivore evolution, remains little understood.

In this study, we use the flea-beetle genus *Longitarsus* (Coleoptera: Chrysomelidae) as a model system for studying the relationship of herbivorous beetle-associated MCD to host plant utilization and specialization. *Longitarsus* flea-beetles (Coleoptera: Chrysomelidae) comprise a morphologically well-defined group of species distributed primarily in Eurasia and North Africa (Palearctic). Evolutionary patterns of *Longitarsus* host plant utilization and host plant secondary chemistry have been extensively analyzed in the context of their phylogenetic relationships (Dobler 2001; Narberhaus et al. 2003; Willinger and Dobler 2001). Most of the *Longitarsus* species have a host range restricted to a few closely related plant species, with a few host plant family "generalists". However, the genus as a whole uses plants from many different families, including the Asteraceae, Boraginaceae, Lamiaceae, and Scrophulariaceae. The overall wealth of evolutionary and ecological information on *Longitarsus* enables the use of the comparative method to identify multiple independent shifts to plants with similar secondary chemistry, as well as many sister-species phylogenetic controls (Dobler 2001).

The primary goal of this preliminary study was to characterize *Longitarsus* MCD in the context of their phylogenetic relationships. While the majority of host-associated microbial community studies focus on a single animal species, we used a parallel-tagged deep-sequencing approach (Sogin et al. 2006) to survey multiple species across the genus. Specifically, we investigated the MCD of eleven different *Longitarsus* species feeding on members of different host plant families (Table 1) that were collected from locales across Germany. After obtaining the sequence data, we used rarefaction analysis and database matching to determine the number and types of bacteria that tend to comprise *Longitarsus* microbial communities. We then used phylogenetically-based statistical methods to determine whether the microbial communities of related *Longitarsus* species tended to be more similar in their microbial diversity than unrelated species. A high degree of similarity in the MCD of related *Longitarsus* species, regardless of host plant, would indicate that the beetle host plays an important role in determining the MCD. Alternatively, there may be little relationship between the beetle phylogeny and MCD, suggesting that other factors, such as diet or host plant chemistry, may play a stronger role in shaping the beetle gut microbiome. Finally, we used

Tecan Ex. 2179, Page 2 of 11
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

Table 1 Host plant, sample type and locale information for *Longitarsus* specimens collected in this study

Sample code[a]	*Longitarsus* species	Host plant species	Host plant family	Chemistry[b]	Sample type[c]	Collection locale[d]
AER#Eup#APF	*L. aeruginosus*	*Eupatorium cannabinum*	Asteraceae	PA	Frass	St. Valentin FR
AER#Eup#APG	*L. aeruginosus*	*Eupatorium cannabinum*	Asteraceae	PA	Gut	St. Valentin FR
ANC#Ech#BPF	*L. anchusae*	*Echium vulgare*	Boraginaceae	PA	Frass	Burkheim FR
ANC#Ech#BPG	*L. anchusae*	*Echium vulgare*	Boraginaceae	PA	Gut	Burkheim FR
ANC#Sym#BPF	*L. anchusae*	*Symphytum officinale*	Boraginaceae	PA	Frass	Freiburg
ECH#Ech#BPF	*L. echii*	*Echium vulgare*	Boraginaceae	PA	Frass	Burkheim FR
EXO#Ech#BPG	*L. exoletus*	*Echium vulgare*	Boraginaceae	PA	Gut	Burkheim FR
JAC#Sen#APG	*L. jacobaeae*	*Senecio jacobaea*	Asteraceae	PA	Gut	Duvenstedt HA
LAT#Pul#BPF	*L. lateripunctatus*	*Pulmonaria officinale*	Boraginaceae	PA	Frass	Niederweiler FR
MEM#Teu#LIF	*L. membranaceus*	*Teucrium scorodonia*	Lamiaceae	IG	Frass	St. Valentin FR
PRA#Pla#PIF	*L. pratensis*	*Plantago lanceolata*	Plantaginaceae	IG	Frass	Burkheim FR
PRA#Pla#PIG	*L. pratensis*	*Plantago lanceolata*	Plantaginaceae	IG	Gut	Burkheim FR
SUC#Ach#APG	*L. succineus*	*Achillea millefolium*	Asteraceae	PA	Gut	Schoenberg FR
SUC#Eup#APG	*L. succineus*	*Eupatorium cannabinum*	Asteraceae	PA	Gut	Zastler FR
SUT#Pet#APF	*L. suturellus*	*Petasites albus*	Asteraceae	PA	Frass	Zastler FR
SUT#Sen#APF	*L. suturellus*	*Senecio jacobaea*	Asteraceae	PA	Frass	Duvenstedt HA
TAB#Ver#SIG	*L. tabidus*	*Verbascum thapsus*	Scrophulariaceae	IG	Gut	Schoenberg FR

[a] The codes summarize the information in the table. The first three letters are same as the specific epithet; the middle three refer to the host plant genus; and the last three letters indicate the plant family, chemistry and sample type. (For example, APF = Asteraceae PA Frass.)

[b] Predominant host plant chemistry. *PA* Pyrrolizidine Alkaloids, *IG* Iridoid Glycosides

[c] The "Gut" samples indicate whole beetle extractions, which are presumably mostly gut in origin (see Methods)

[d] Nearest town of collection site in Germany. *FR* Site near Freiburg, *HA* Site near Hamburg

Principal Coordinates Analysis (PCoA) to investigate potential relationships between the beetle gut MCD and aspects of their resource use, such as host plant family affiliations and secondary chemistry. Altogether, our results shed considerable light on the origins, diversity and character of the *Longitarsus* gut-microbiome, and suggested new avenues of research concerning the relationship of gut microbes to resource use in herbivorous insects.

Materials and methods

Sample collection

Beetles were collected from host plants at numerous locales around Germany either via sweep-nets or through the used of a modified leaf blower (sucker) positioned over the plants. Table 1 lists the locales and host plants of the beetles collected over the course of this study. After collection, the beetles were transported live to the University of Hamburg, in 50 ml centrifuge tubes, where they were identified by S. Dobler. Some of the beetle species were frozen at −80°C prior to DNA extraction, while other specimens were kept alive and placed in sterile 15 ml centrifuge tubes for frass collection.

Microbial DNA extraction

We extracted whole microbial community DNA from pooled insect samples and from frass samples. The small size of the flea-beetle species made gut-specific dissections untenable for most *Longitarsus* species. In order to study the *Longitarsus* MCD, we removed the legs, sterilized the outside of the insects, and extracted bacterial DNA from the whole specimen. These "whole beetle" extracted samples, therefore, represent bacteria from all the internal organs not just the gut. However, guts typically contain the vast majority of the animal microbiota. Furthermore, sequences of endosymbionts, such as *Wolbachia*, are easily identifiable and can be removed from select downstream analyses (e.g., Unifrac distance calculations, PCoA; see below). For the whole beetle extractions, we removed the wings and elytra from individual beetles and dipped them in 10% hydrogen peroxide for 10 s to remove external bacteria and DNA contamination. The beetles were then freeze-dried overnight and pulverized. We used 3–5 beetles per DNA extraction depending on the size of the species. Frass samples were scraped off the side of the 15 ml centrifuge frass-collections tubes with a flame-sterilized metal spatula and placed into a 1.5 ml tube for DNA extraction. Fecal (excrement) samples are commonly used to investigate gut communities of animals, particularly

Tecan Ex. 2179, Page 3 of 11
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

vertebrates (Ley et al. 2008), and they provide a ready source of bacteria that has passed through the gut. However, to our knowledge, direct comparisons of gut MCD to that found in frass have not been performed and our system allowed an easy comparison of these sample types within species. We also performed negative extraction (water) controls to ensure that the buffers, enzyme and solutions were not contaminated with microbes or DNA.

The DNA extraction mixture had a total volume of 200 µl and included the following final concentration: 20 M Tris, 2 mM EDTA (pH 8.0), 1.2% P40 detergent, 20 mg ml^{-1} lysozyme, and 0.2 µM filtered sterile water (Sigma Chemical Co., St. Louis, MO). Samples were incubated in a 37°C water bath for 30 min. Next, Proteinase K (DNeasy Tissue Kit, Qiagen Corporation, Valencia, CA) and AL Buffer (DNeasy Tissue Kit, Qiagen Corporation, Valencia, CA) were added to the tubes and gently mixed. Samples were incubated in a 70°C water bath for 10 min. All samples were subjected to purification using the DNeasy Tissue Kit. Following extraction, the DNA was quantified using a Pharmacia Ultrospec 2000 (Pharmacia Biotech, Cambridge, UK).

PCR reactions were performed in small lots (4–6 reactions plus a positive and negative extraction and PCR controls) to reduce the possibility of contamination. The PCR amplification was performed with "universal" bacterial primers. These primers were designed from regions of the 16S gene conserved in all bacteria (27F and 338R) and the same primer set were used in numerous other studies (Costello et al. 2009; Fierer et al. 2008; Ley et al. 2008). These primers also contained a unique 12-nucleotide "barcode" for each sample. The sequence barcode allowed us to pool the PCR products from all the samples into one 454 sequencing run. PCR reagent concentrations and reaction conditions followed Fierer et al. (2008), except that we performed fewer thermal cycles (30 instead of 35).

Pyrosequencing

Individual PCR products were purified using the AMPure purification kit (Agenourt, Beverly) following the manufacturer's protocol. After Ampure purification, each sample was quantified on an Agilent 2100 Bioanalyzer. All samples were diluted down to 2×10^{-5} mol/µl and were then pooled with a total combined concentration of 2×10^{-5} mol/µl (100 µl total volume). Pyrosequencing reactions were performed on a 454 GS FLX Titanium Sequencer (Roche). PCR purification, dilutions and pyrosequencing were all conducted by the core facility at the University of South Carolina (Environmental Genomics Core Facility).

Computational and statistical analyses

In-house programming scripts written in Python were used to identify barcodes for all sequences and sort the sequences by sample. Python scripts were also written to remove sequences that were too short (less than <130 bases in length), had low average quality scores (average score <25), or had unrecognizable barcodes. We also blasted against a database of well-curated 16S bacterial sequences obtained from Greengenes (DeSantis et al. 2006). Sequences that did not match a hit over their entire length were excluded. For example, we excluded a sequence of 200 bp that matched nearly identically to a sequence in the Greengenes database, but only aligned for 100 bp. These were considered possible chimeric sequences.

We also analyzed our data using the QIIME ("Quantitative Insights Into Microbial Ecology") software pipeline (Caporaso et al. 2010). QIIME is a software pipeline for performing microbial community analysis that, among many other things, performs quality checking and sorting of high-throughput sequence datasets. QIIME also integrates many of the third party tools which have become standard in the field and which we also used separately (e.g., RDP 10, UniFrac, PCoA analyses…etc.). The results of the QIIME analysis were indistinguishable from our in-house pipeline analysis, providing an independent confirmation of our sequence processing methods.

After quality checking, counting and sorting, we used the Ribosomal Database Project (RDP) Classifier to identify the most likely match of each sequence in our samples. Since all of our sequences were exact, or nearly exact (>98% identical) matches to a sequence in the Greengenes database, we used the Greengenes sequence instead of the original sequence with the RDP classifier. The RDP Classifier uses a naïve Bayesian algorithm to classify 16S rRNA gene sequences according to the Bergey's 2nd Edition classification system down to the genus level. At the time of our analysis, we used RDP version 10.13 (updated on July 28, 2009) which had a total of 1,049,433 sequences for comparison (Cole et al. 2009). The taxonomic sequence identifications were also cross-checked against the Greengenes database and NCBI nucleotide database using BLAST.

QIIME was used to produce the rarefaction plots (Caporaso et al. 2010). The Fast UniFrac program was used to: (1) Calculate weighted and un-weighted UniFrac distance measures between all pairs of microbial communities; (2) Cluster microbial communities based on their UniFrac distances; and (3) perform PCoA analyses. UniFrac distances are a measure of the phylogenetic distance between sets of taxa in a phylogenetic tree and is especially useful for comparing the similarity of microbial communities. Specifically, "UniFrac, measures the phylogenetic

Tecan Ex. 2179, Page 4 of 11
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

distance between sets of taxa in a phylogenetic tree as the fraction of the branch length of the tree that leads to descendants from either one environment or the other, but not both (Lozupone and Knight 2005)." In regards to the *Longitarsus* system, pair-wise UniFrac distances were calculated for all the samples analyzed in the study (Table 1). A small UniFrac distance indicated that a given pair of microbial communities were phylogenetically similar to each other, while a large UniFrac distance indicated the opposite. UniFrac distance calculations were performed both weighted by sequence abundance (beta-diversity) and un-weighted (alpha-diversity). The pairwise UniFrac measures were also used to cluster similar communities and for PCoA analyses.

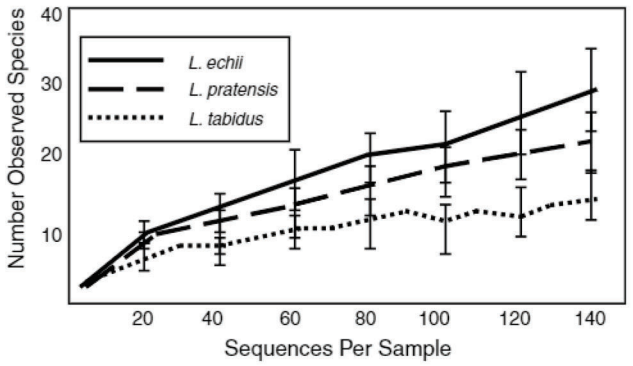

Fig. 1 Rarefaction curve of three representative species indicating the operational taxonomic unit (OTU) richness of *Longitarsus* gut microbial samples. All sequences less than 1% divergence from one another were grouped into an OTU. The curves include the most (*L. echii*) and least (*L. tabidus*) diverse of the *Longitarsus* communities sampled

Results and discussion

We were able to successfully amplify bacterial 16S rRNA gene sequence from all *Longitarsus* species and sample types (whole beetle or frass) collected in the study. Purified and normalized PCR amplicons were pooled from seventeen different samples collected from eleven *Longitarsus* species. A half-run of the final pooled PCR amplicons on a 454 GS FLX Titanium Sequencer yielded a total of approximately 250,000 sequences (71 MB of data). Post-quality checking (length, chimera, barcode and primer sequence removal and base-quality checks) resulted in 145,000 sequences of more than 150 bp in length.

The sequences were also screened for eukaryotic rRNA contamination. We found no evidence of insect nuclear or mitochondrial rRNA contamination. However, the primer set we used amplified an inordinate amount of plant chloroplast rRNA sequences. Over half the sequences generated from our pyrosequencing run turned out to be plant chloroplast in origin. These were missed at first because of errors in the Greengenes databases that identified plant chloroplast sequences as "uncultured environmental bacteria" and "cyanobacteria". The fact that this primer set so readily picks up plant chloroplast sequences was not reported in earlier studies that used this primer set, likely due to their subject matter (Costello et al. 2009; Fierer et al. 2008). Interestingly, we found hundreds of sequences from environmental 16S rRNA studies labeled as "environmental bacteria" in Genbank that were perfect matches to plant chloroplast rRNA. A recent study by Redford et al. (2010) on the bacteria present on leaf surfaces reported similar problems. These authors designed a new primer set that does not appear to amplify chloroplast sequences and we will be using this alternative set in future studies of *Longitarsus*.

Although we had to discard half the sequences due to this contamination, we still retained more than enough bacterial community sequences to analyze the diversity of *Longitarsus* microbial communities. Figure 1 shows the results of a 16S rRNA OTU (operational taxonomic unit)-based rarefaction analysis. In this study, we conservatively defined OTUs as sequences less than 1% divergence from one another. The numbers of bacterial OTUs tended to be quite limited compared to other animal microbial communities and the rarefaction curves were not steep. However, none of the curves reached a clear asymptote, suggesting the presence of many rare bacterial species.

Table 2 lists the taxonomic identification of the most commonly occurring bacteria in *Longitarsus* microbial communities based on RDP and BLAST matches to Genbank. Despite the low overall MCD per sample, the types of bacteria found associated with these insects were highly interesting and informative concerning the *Longitarsus* microbiome. The three most common organisms found in most or all of the samples were identical to bacteria belonging to the Enterobacteriaceae (γ-Proteobacteria), and have also been discovered in the guts of other animals, particularly insects. The most commonly occurring sequence (found in all samples) was an identical match to an uncultured Enterobacteriaceae found in mouse intestines (Ley et al. 2008). The next two most commonly occurring organisms were also Enterobacteriaceae, *Rahnella* sp. and *Serratia* sp., and members of both these genera have been found in insect gut studies (Grimont et al. 1979; Yu et al. 2008). The role of *Rahnella* sp. in insect guts is not known, but these bacteria are common in the guts of fish (salmon and trout) where they are involved in nitrate reduction and the fermentation of complex carbohydrates (Yu et al. 2008). *Serratia* sp. are known to metabolize both arthropod and fungal chitins (Jones et al. 1986) as well as a complex assortment of carbohydrates (Grimont et al. 1979).

Tecan Ex. 2179, Page 5 of 11
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

Table 2 Most commonly occurring bacterial genera in gut and frass samples among *Longitarsus* species surveyed and their typical habitat associations

Bacterial lineage	Habitat
Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Uncultured	Mouse Gut
Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Rahnella	Insect Gut
Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Serratia	Insect Gut
Gammaproteobacteria; environmental	Soils/Sediments
Actinobacteria; environmental samples	Soils
Gammaproteobacteria; Pseudomonadaceae; Pseudomonas	Guts
Gammaproteobacteria; Pseudomonadaceae; Pseudomonas	Guts
Gammaproteobacteria; Pseudomonadales; Acinetobacter	Soils
Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; environmental samples	Soils
Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Serratia	Insect Gut
Alphaproteobacteria; Sphingomonadales; Sphingomonadaceae; Citromicrobium	Unknown Environmental
Gammaproteobacteria; Xanthomonadales; Xanthomonadaceae; Frateuria	Plants
Actinobacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Micrococcineae; Microbacteriaceae; Microbacterium	Soils
Alphaproteobacteria; Rhizobiales; Bradyrhizobiaceae; Blastobacter	Rhizosphere
Proteobacteria; Alphaproteobacteria; Rickettsiales; Anaplasmataceae; Wolbachia	Endosymbiont
Proteobacteria; Alphaproteobacteria; Rickettsiales; Rickettsiaceae; Rickettsia	Endosymbiont

After the gut-associated bacteria, the second most common group of organisms found in *Longitarsus* samples matched sequences of bacteria found in soils and, specifically, the roots of plants (i.e., the "rhizosphere"). For example, the fourth and fifth most common bacterial matches were to an uncultured γ-Proteobacteria found in soil sediments (Genbank Accession number AB188783) and to an uncultured *Actinobacteria* associated with the Aspen rhizosphere (Lesaulnier et al. 2008; Genbank Accession number EF020313). Other common soil bacteria associated with *Longitarsus* species included several other γ-Proteobacteria (*Acinetobacter* sp., Xanthomonads), α-Proteobacteria (Sphingomonads, members of the Rhizobiales) and other *Actinobacteria* spp. Both *Actinobacteria* and α-Proteobacteria are extremely common in soils and in the rhizosphere.

Clearly, *Longitarsus* MCD shares important similarities with other insect gut communities, but the significant proportion of soil-associated organisms is quite remarkable and deserving of further investigation. Like many other herbivorous beetle species, *Longitarsus* larvae feed on the roots of their host plants before emerging and feeding on the leaf tissues. Thus, it seems logical to suggest that the larvae are inoculated and colonized by soil microbes surrounding the rhizosphere that may persist in the guts of the adults. Needless to say, soil communities are far more biologically complex than the microbial communities of the insects (Borneman et al. 1996), indicating only a few of the soil organisms persist long-term to the adult stage. Based on the findings of this study, one main goal of future research in this system will be to determine how the life-cycle of the insect, and root feeding in particular, contributes to the MCD of the beetles.

The soil-inoculation hypothesis, if true, may explain the general lack of association between *Longitarsus* MCD and the phylogeny or life-history traits of the beetles. Figure 2 shows a clustering analysis of the *Longitarsus* microbial communities analyzed in this study. The clusters indicate which of the microbial communities were most similar based on their UniFrac distances, a standard metric in microbial ecology. The clustering analysis appeared to reject the hypothesis that the phylogeny of the insects provides the best explanation of MCD patterns in these insects. Although the MCD of some related beetles clustered together (e.g., *L. aeruginosus* (AER) and *L. suturellus* (SUC)), we found other examples of unrelated species clustering together. For example, *L. aeruginosus*, *L. jacobaeae* (JAC), *L. tabidus* (TAB) and *L. suturellus* (SUT) samples formed a cluster that was well-supported by jackknife re-sampling (Fig. 2) despite the fact that only *L. aeruginosus* and *L. suturellus* are close relatives (Dobler 2001). Similarly, we find clustering of distant relatives *L. anchusae* (ANC) and *L. echii* (ECH), while the MCD of close relatives *L. echii* and *L. aeruginosus* was not similar (Fig. 2). Perhaps supportive of the soil-inoculation hypothesis, we note that the *L. anchusae* and *L. echii* samples (ANC#Ech#BPF and ECH#Ech#BPF; Table 1, Fig. 2) were collected from beetles found feeding on the

Tecan Ex. 2179, Page 6 of 11
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

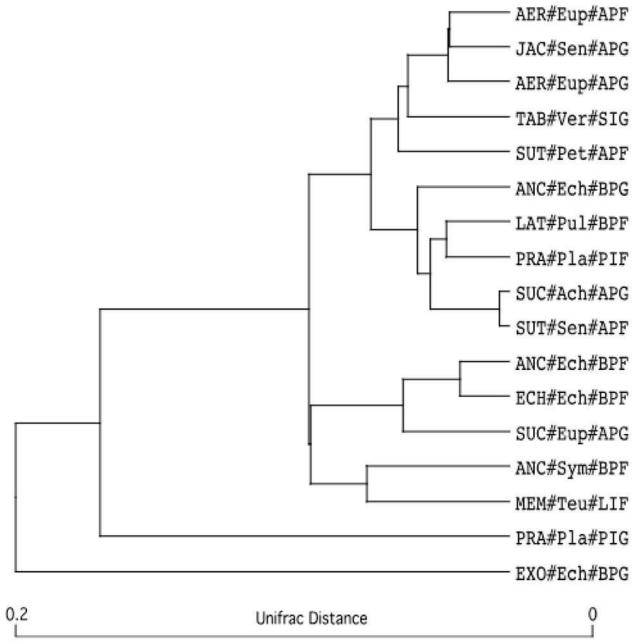

Fig. 2 UPGMA clustering of weighted pair-wise Unifrac distances between *Longitarsus* microbial communities (Table 1) based on their distribution of bacterial 16S rRNA gene sequences. Weighted UniFrac distances incorporate sequence abundance (beta diversity). The codes refer to samples in Table 1. The *first three letters* before the # refer to the beetle species (e.g., *AER L. aeruginosus*), the *middle three letters* refer to the host plant genus (e.g., *Eup Eupatorium*) and the *last three letters* indicate host plant family, chemistry and sample type. For example, *APF* Asteraceae PA Frass. In some cases, related species has similar communities, but in many cases related beetle species were highly dissimilar. We even found strong differences within the same species on different hosts. In two cases, *L. anchusae* (ANC) and *L. praetensis* (PRA), we found stark differences in the microbial communities between gut and frass samples collected from the same species on the same hosts in the same locale

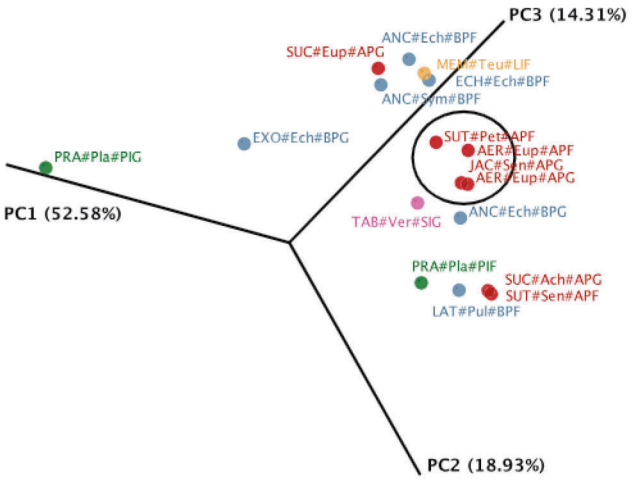

Fig. 3 *Longitarsus* gut microbial communities clustered using PCoA of the pair-wise weighted Unifrac distances. The *ball colors* indicated the host plant family associations. (*Red* Asteraceae, *Blue* Boraginaceae, *Yellow* Lamiaceae, *Purple* Scrophulariaceae, *Green* Plantaginaceae). There appeared to be some clustering of the Asteraceae feeders indicated by the circle. (Color figure online)

same plants and probably developed in the same soil. Collectively, these results indicate that some aspect of the beetle's environment (e.g., host-use or soil environment) may be playing a stronger role in determining *Longitarsus* microbial community diversity than the beetle itself, a hypothesis that deserves further testing as suggested above.

One additional, and potentially important, finding of the clustering analysis was the fact that in two of the three comparisons made, whole beetle and frass communities from the same insect species collected from the same host and locales were not similar in their MCD. While *L. aeruginosus* whole beetle and frass sample MCD clustered together, whole beetle and frass samples of two other species, *L. pratensis* (PRA) and *L. anchusae*, were quite dissimilar (Fig. 2). Because the whole beetle extraction included both gut symbionts and symbionts from other organ tissues, these comparisons were not exclusively gut versus frass. Nevertheless, we were surprised to see such a dramatic difference between whole beetle and frass samples from the same species and even the same individuals. This finding suggests that frass MCD may not always be reflective of gut MCD, perhaps because bacteria that tend to be found in the gut remain there or that, after defecation, certain bacteria grow better outside the anaerobic gut environment and dominate in the frass. This finding has potential ramifications for not only this system but also other systems, as most researchers prefer to use fecal samples in metagenomic studies as a proxy for gut MCD because of the ease of sampling (Breitbart et al. 2003, 2008; Hackstein et al. 1995; Ley et al. 2005, 2006b, 2008; Mittal et al. 2005). However, at least for herbivorous beetles, this strategy may be inappropriate or should be cautiously undertaken.

While the lack of association between the phylogeny of the insects and their gut MCD indicated that the environment plays a dominant role in shaping these communities, a PCoA analysis did not shed much light on the factors that associate with MCD (Fig. 3). Despite the fact that the first three principal components explained 84% of the variation in pair-wise UniFrac distances, we did not see strong associations with life-history characters, such as host plant family, or secondary chemistry. However, we did see clustering of *Longitarsus* samples associated with Asteraceaous plants and potential clustering of the Boraginaceae-feeders (Fig. 3). These results suggest that host plants may play a role in determining *Longitarsus* MCD, also that some host plant types may have a much stronger impact on insect MCD than others.

In summary, our results show that *Longitarsus*-associated microbial communities are comprised of a complex mix of insect gut-associated and rhizosphere-associated

Tecan Ex. 2179, Page 7 of 11
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

bacteria. The general lack of association between environmental factors or life-history traits (e.g., host plant family and chemistry) may reflect random soil inoculation at the larval stage. However, our results did suggest that host plants may influence *Longitarsus* MCD, and that this effect may vary in strength depending on the host plant family. Clearly, a great deal more sampling needs to be done in order to understand the origin and complexity of the *Longitarsus* gut ecosystem. The next phase of the study will include a much deeper intra-specific analysis of MCD within a few easily accessible *Longitarsus* species across their range, including individuals on the same host plants in different locales and comparisons of unrelated sympatric species from the same host plants. We will also explore the soil microbial diversity around the roots of the plants and compare it to that of the adults just after they emerge from the soils. Future studies will also use other primer sets that target bacterial housekeeping genes (Santos and Ochman 2004) and focus on whole beetle extractions to avoid issues associated with frass samples. Finally, since fungi are known to be important symbionts of many insects, and can be very diverse in beetles (e.g., Suh et al. 2005), we will be using "universal" 18S primers (Fierer et al. 2007) to target *Longitarsus*-associated fungal diversity alongside the bacterial diversity.

Acknowledgments We wish to thank all the members of the Dobler lab who helped make this study possible, especially K. Meyer, S. Marzez and C. Baden. We thank V. Thackray, the editor B. Normark, and two anonymous reviewers for their many helpful suggestions. This study was funded by a grant from the Alexander von Humboldt Foundation. Finally, as first author, I would like to extend a very special thank you to Professor Richard G. Harrison. Rick's lectures on evolutionary biology were wonderful and inspired me during a very difficult period of my undergraduate education at Cornell University. After taking Rick's class, I knew for certain that I would become an evolutionary biologist. I also thank Rick for taking a chance and hiring me as a research technician in his lab after my difficult stint as an elementary school teacher in Houston, Texas. Once again, this experience with Rick proved to be tremendously inspirational and guided me towards my eventual career as a biology professor. During my time in Rick's lab I learned molecular biology skills from the world's best (Steven Bogdanowicz), uncovered my passion for phylogenetics, discovered the joys of working with insects, came to know my future Ph.D. advisor, and met some of the best graduate students and scientists in evolutionary biology. Thus, I credit much of my career success, and indeed happiness, to Rick's kindness and generosity.

References

Amann RI, Ludwig W, Schleifer KH (1995) Phylogenetic identification and in situ detection of individual microbial cells without cultivation. Micro Rev 59:143–169

Andersson AF, Lindberg M, Jakobsson H, Backhed F, Nyren P, Engstrand L (2008) Comparative analysis of human gut microbiota by barcoded pyrosequencing. PLoS ONE 3:e2836

Backhed F, Ley RE, Sonnenburg JL, Peterson DA, Gordon JI (2005) Host-bacterial mutualism in the human intestine. Science 307:1915–1920

Behar A, Yuval B, Jurkevitch E (2005) Enterobacteria-mediated nitrogen fixation in natural populations of the fruit fly *Ceratitis capitata*. Mol Ecol 14:2637–2643

Behar A, Jurkevitch E, Yuval B (2008a) Bringing back the fruit into fruit fly-bacteria interactions. Mol Ecol 17:1375–1386

Behar A, Yuval B, Jurkevitch E (2008b) Gut bacterial communities in the Mediterranean fruit fly (*Ceratitis capitata*) and their impact on host longevity. J Insect Phys 54:1377–1383

Bernays EA, Chapman RF (1994) Host-plant selection by phytophagous insects. Chapman & Hall, London

Bernays EA, Graham M (1988) On the evolution of host specificity in phytophagous arthropods. Ecology 69:886–892

Borneman J, Skroch PW, O'Sullivan KM, Palus JA, Rumjanek NG, Jansen JL, Nienhuis J, Triplett EW (1996) Molecular microbial diversity of an agricultural soil in Wisconsin. Appl Environ Microbiol 62:1935–1943

Breitbart M, Hewson I, Felts B, Mahaffy JM, Nulton J, Salamon P, Rohwer F (2003) Metagenomic analyses of an uncultured viral community from human feces. J Bacteriol 185:6220–6223

Breitbart M, Haynes M, Kelley S, Angly F, Edwards RA, Felts B, Mahaffy JM, Mueller J, Nulton J, Rayhawk S, Rodriguez-Brito B, Salamon P, Rohwer F (2008) Viral diversity and dynamics in an infant gut. Res Microbiol 159:367–373

Broderick NA, Raffa KF, Goodman RM, Handelsman J (2004) Census of the bacterial community of the gypsy moth larval midgut by using culturing and culture-independent methods. Appl Environ Microbiol 70:293–300

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Pena AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R (2010) QIIME allows analysis of high-throughput community sequencing data. Nat Methods 7:335–336

Cole JR, Wang Q, Cardenas E, Fish J, Chai B, Farris RJ, Kulam-Syed-Mohideen AS, McGarrell DM, Marsh T, Garrity GM, Tiedje JM (2009) The ribosomal database project: improved alignments and new tools for rRNA analysis. Nucleic Acids Res 37:D141–145

Costello EK, Lauber CL, Hamady M, Fierer N, Gordon JI, Knight R (2009) Bacterial community variation in human body habitats across space and time. Science 326:1694–1697

Delalibera I, Handelsman J, Raffa KF (2005) Contrasts in cellulolytic activities of gut microorganisms between the Wood Borer, *Saperda vestita* (Coleoptera: Cerambycidae), and the Bark Beetles, *Ips pini* and *Dendroctonus frontalis* (Coleoptera: Curculionidae). Environ Entomol 34:541–547

DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, Andersen GL (2006) Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Appl Environ Microbiol 72:5069–5072

Dobler S (2001) Evolutionary aspects of defense by recycled plant compounds in herbivorous insects. Basic Appl Ecol 2:15–26

Dobler S, Farrell BD (1999) Host use evolution in *Chrysochus* milkweed beetles: evidence from behaviour, population genetics and phylogeny. Mol Ecol 8:1297–1307

Ehrlich PR, Raven PH (1964) Butterflies and plants: a study in coevolution. Evolution 18:586–608

Erwin TL (1982) Tropical forests: their richness in Coleoptera and other arthropod species. Coleop Bull 36:74–75

Farrell BD (1998) "Inordinate Fondness" explained: why are there so many beetles? Science 281:555–559

Tecan Ex. 2179, Page 8 of 11
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

Farrell BD, Mitter C (1998) The timing of insect/plant diversification: might *Tetraopes* (Coleoptera: Cerambycidae) and *Asclepias* (Asclepiadaceae) have co-evolved? Biol J Linn Soc 63:553–577

Fierer N, Breitbart M, Nulton J, Salamon P, Lozupone C, Jones R, Robeson M, Edwards RA, Felts B, Rayhawk S, Knight R, Rohwer F, Jackson RB (2007) Metagenomic and small-subunit rRNA analyses reveal the genetic diversity of bacteria, archaea, fungi, and viruses in soil. Appl Environ Microbiol 73:7059–7066

Fierer N, Hamady M, Lauber CL, Knight R (2008) The influence of sex, handedness, and washing on the diversity of hand surface bacteria. Proc Natl Acad Sci USA 105:17994–17999

Futuyma DJ, Moreno G (1988) The evolution of ecological specialization. Ann Rev Ecol Syst 19:207–233

Futuyma DJMCK, Scheffer SJ (1993) Genetic constraints and the phylogeny of insect-plant associations: responses of *Ophraella* communa (Coleoptera: Chrysomelidae) to host-plants of its congeners. Evolution 47:888–905

Grimont PA, Grimont F, Le Minor S, Davis B, Pigache F (1979) Compatible results obtained from biotyping and serotyping in *Serratia marcescens*. J Clin Microbiol 10:425–432

Hackstein JH, Van Alen TA, Op Den Camp H, Smits A, Mariman E (1995) Intestinal methanogenesis in primates—a genetic and evolutionary approach. Dtsch Tierarztl Wochenschr 102: 152–154

Harris JK, Kelley ST, Pace NR (2004) New perspective on uncultured bacterial phylogenetic division OP11. Appl Environ Microbiol 70:845–849

Haubruge E, Arnaud L (2001) Fitness consequences of malathion-specific resistance in red flour beetle (Coleoptera: Tenebrionidae) and selection for resistance in the absence of malathion. J Econ Entomol 94:552–557

Haverson K, Rehakova Z, Sinkora J, Sver L, Bailey M (2007) Immune development in jejunal mucosa after colonization with selected commensal gut bacteria: a study in germ-free pigs. Vet Immunol Immunopathol 119:243–253

Hugenholtz P, Pace NR (1996) Identifying microbial diversity in the natural environment: a molecular phylogenetic approach. Trends Biotech 14:190–197

Hugenholtz P, Goebel BM, Pace NR (1998a) Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. J Bacteriol 180:4765–4774

Hugenholtz P, Pitulle C, Hershberger KL, Pace NR (1998b) Novel division level bacterial diversity in a Yellowstone hot spring. J Bacteriol 180:366–376

Jones JD, Grady KL, Suslow TV, Bedbrook JR (1986) Isolation and characterization of genes encoding two chitinase enzymes from *Serratia marcescens*. EMBO 5:467–473

Kelley ST, Farrell BD (1998) Is specialization a dead end? The phylogeny of host use in *Dendroctonus* bark beetles (Scolytidae). Evolution 52:1731–1743

Lesaulnier C, Papamichail D, McCorkle S, Ollivier B, Skiena S, Taghavi S, Zak D, van der Lelie D (2008) Elevated atmospheric CO_2 affects soil microbial diversity associated with trembling aspen. Environ Microbiol 10:926–941

Ley RE, Backhed F, Turnbaugh P, Lozupone CA, Knight RD, Gordon JI (2005) Obesity alters gut microbial ecology. Proc Natl Acad Sci USA 102:11070–11075

Ley RE, Harris JK, Wilcox J, Spear JR, Miller SR, Bebout BM, Maresca JA, Bryant DA, Sogin ML, Pace NR (2006a) Unexpected diversity and complexity of the Guerrero Negro hypersaline microbial mat. Appl Environ Microbiol 72: 3685–3695

Ley RE, Turnbaugh PJ, Klein S, Gordon JI (2006b) Microbial ecology: human gut microbes associated with obesity. Nature 444:1022–1023

Ley RE, Hamady M, Lozupone C, Turnbaugh PJ, Ramey RR, Bircher JS, Schlegel ML, Tucker TA, Schrenzel MD, Knight R, Gordon JI (2008) Evolution of mammals and their gut microbes. Science 320:1647–1651

Lozupone C, Knight R (2005) UniFrac: a new phylogenetic method for comparing microbial communities. Appl Environ Microbiol 71:8228–8235

Lozupone CA, Knight R (2007) Global patterns in bacterial diversity. Proc Natl Acad Sci USA 104:11436–11440

McKenna P, Hoffmann C, Minkah N, Aye PP, Lackner A, Liu Z, Lozupone CA, Hamady M, Knight R, Bushman FD (2008) The macaque gut microbiome in health, lentiviral infection, and chronic enterocolitis. PLoS Path 4:e20

Mittal RD, Kumar R, Bid HK, Mittal B (2005) Effect of antibiotics on *Oxalobacter formigenes* colonization of human gastrointestinal tract. J Endourol 19:102–106

Mitter C, Futuyma DJ (1991) Phylogenetic studies of insect-plant interactions: insights into the genesis of diversity. Trends Ecol Evol 6:290–293

Mitter C, Farrell B, Wiegmann B (1988) The phylogenetic study of adaptive zones: has phytophagy promoted insect diversification? Am Nat 132:107–128

Narberhaus I, Theuring C, Hartmann T, Dobler S (2003) Uptake and metabolism of pyrrolizidine alkaloids in *Longitarsus* flea beetles (Coleoptera: Chrysomelidae) adapted and non-adapted to alkaloid-containing host plants. J Comp Phys 173:483–491

Novotny V, Drozd P, Miller SE, Kulfan M, Janda M, Basset Y, Weiblen GD (2006) Why are there so many species of herbivorous insects in tropical rainforests? Science 313: 1115–1118

Pace NR (1997) A molecular view of microbial diversity and the biosphere. Science 276:734–740

Pace NR, Stahl DA, Lane DJ, Olsen GJ (1985) Analyzing natural microbial populations by rRNA sequences. ASM News 51:4–12

Paine TD, Raffa KF, Harrington TC (1997) Interactions among scolytid bark beetles, their associated fungi, and live host conifers. Ann Rev Entomol 42:179–206

Redford AJ, Bowers RM, Knight R, Linhart Y, Fierer N (2010) The ecology of the phyllosphere: geographic and phylogenetic variability in the distribution of bacteria on tree leaves. Environ Micro (Epub) doi: 10.1111/j.1462-2920.2010.02258.x

Rhee KJ, Sethupathi P, Driks A, Lanning DK, Knight KL (2004) Role of commensal bacteria in development of gut-associated lymphoid tissues and preimmune antibody repertoire. J Immunol 172:1118–1124

Rondon MR, Goodman RM, Handelsman J (1999) The Earth's bounty: assessing and accessing soil microbial diversity. Trends Biotech 17:403–409

Safaee S, Weiser GC, Cassirer EF, Ramey RR, Kelley ST (2006) Microbial diversity in bighorn sheep revealed by culture-independent methods. J Wild Dis 42:545–555

Santos SR, Ochman H (2004) Identification and phylogenetic sorting of bacterial lineages with universally conserved genes and proteins. Environ Microbiol 6:754–759

Sexson DL, Wyman JA (2005) Effect of crop rotation distance on populations of Colorado potato beetle (Coleoptera: Chrysomelidae): development of areawide Colorado potato beetle pest management strategies. J Econ Entomol 98:716–724

Sogin ML, Morrison HG, Huber JA, Mark Welch D, Huse SM, Neal PR, Arrieta JM, Herndl GJ (2006) Microbial diversity in the deep sea and the underexplored "rare biosphere". Proc Natl Acad Sci USA 103:12115–12120

Suh SO, McHugh JV, Pollock DD, Blackwell M (2005) The beetle gut: a hyperdiverse source of novel yeasts. Mycol Res 109:261–265

Tecan Ex. 2179, Page 9 of 11
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

Tringe SG, Hugenholtz P (2008) A renaissance for the pioneering 16S rRNA gene. Curr Opin Microbiol 11:442–446

Tringe SG, von Mering C, Kobayashi A, Salamov AA, Chen K, Chang HW, Podar M, Short JM, Mathur EJ, Detter JC, Bork P, Hugenholtz P, Rubin EM (2005) Comparative metagenomics of microbial communities. Science 308:554–557

Van Borm S, Billen J, Boomsma JJ (2002) The diversity of microorganisms associated with *Acromyrmex* leafcutter ants. BMC Evol Biol 2:9

Visotto LE, Oliveira MG, Guedes RN, Ribon AO, Good-God PI (2009) Contribution of gut bacteria to digestion and development of the velvetbean caterpillar, *Anticarsia gemmatalis*. J Insect Phys 55:185–191

Warnecke F, Luginbuhl P, Ivanova N, Ghassemian M, Richardson TH, Stege JT, Cayouette M, McHardy AC, Djordjevic G, Aboushadi N, Sorek R, Tringe SG, Podar M, Martin HG, Kunin V, Dalevi D, Madejska J, Kirton E, Platt D, Szeto E, Salamov A, Barry K, Mikhailova N, Kyrpides NC, Matson EG, Ottesen EA, Zhang X, Hernandez M, Murillo C, Acosta LG, Rigoutsos I, Tamayo G, Green BD, Chang C, Rubin EM, Mathur EJ, Robertson DE, Hugenholtz P, Leadbetter JR (2007) Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. Nature 450:560–565

Willinger G, Dobler S (2001) Selective sequestration of iridoid glycosides from their host plants in *Longitarsus* flea beetles. Biochem Syst Ecol 29:335–346

Wood SL (1982) The bark and ambrosia beetles (Coleoptera: Scolytidae) of North and Central America, a taxonomic monograph. Great Basin Naturalist Memoirs 6

Yu H, Wang Z, Liu L, Xia Y, Cao Y, Yin Y (2008) Analysis of the intestinal microflora in *Hepialus gonggaensis* larvae using 16S rRNA sequences. Curr Microbiol 56:391–396

Tecan Ex. 2179, Page 10 of 11
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029