

A 454 Survey Reveals the Community Composition and Core Microbiome of the Common Bed Bug (*Cimex lectularius*) across an Urban Landscape

Matthew Meriweather¹, Sara Matthews¹, Rita Rio², Regina S. Baucom^{1*}

¹ Department of Biological Sciences, University of Cincinnati, Cincinnati, Ohio, United States of America, ² Department of Biology, West Virginia University, Morgantown, West Virginia, United States of America

Abstract

Elucidating the spatial dynamic and core constituents of the microbial communities found in association with arthropod hosts is of crucial importance for insects that may vector human or agricultural pathogens. The hematophagous *Cimex lectularius* (Hemiptera: Cimicidae), known as the human bed bug, has made a recent resurgence in North America, as well as worldwide, potentially owing to increased travel, climate change and resistance to insecticides. A comprehensive survey of the bed bug microbiome has not been performed to date, nor has an assessment of the spatial dynamics of its microbiome. Here we present a survey of internal and external bed bug microbial communities by amplifying the V4–V6 hypervariable region of the 16S rDNA gene region followed by 454 Titanium sequencing using 31 individuals from eight distinct collection locations obtained from residences in Cincinnati, OH. Across all samples, 97% of the microbial community is made up of two dominant OTUs, previously identified as the α -proteobacterium *Wolbachia* and an unnamed γ -proteobacterium from the Enterobacteriaceae. Microbial communities varied among host locations for measures of community diversity and exhibited structure according to collection location. This broad survey represents the most in-depth assessment, to date, of the microbes that associate with bed bugs.

Citation: Meriweather M, Matthews S, Rio R, Baucom RS (2013) A 454 Survey Reveals the Community Composition and Core Microbiome of the Common Bed Bug (*Cimex lectularius*) across an Urban Landscape. PLoS ONE 8(4): e61465. doi:10.1371/journal.pone.0061465

Editor: Colin Dale, University of Utah, United States of America

Received: October 12, 2012; **Accepted:** March 13, 2013; **Published:** April 9, 2013

Copyright: © 2013 Meriweather et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: Funding was provided by the University of Cincinnati. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: gina.baucom@gmail.com

Introduction

The interactions between microbes and their insect hosts can range from obligate mutualism to parasitic and ultimately detrimental to host fitness [1,2,3]. The microbes that insects carry can also negatively impact other organisms following vector transmission. Characterizing the microbial community of insect hosts is thus an important primary step for elucidating the nature of host/microbe interactions and the potential for the insect vectoring of important pathogens. Assessments of the core microbiome—defined as members of the microbial community that are common to two or more assemblages of a particular habitat type [4,5]—are often performed to identify key host/microbe associations. The microbes that appear in all hosts are thought to fulfill a functional niche within the community [6] and as such may provide valuable information on the ‘normal’ state of the community as well as on what members might be targeted for manipulation (e.g., in a scenario of insect biocontrol).

The microbiome of an insect host could vary among host populations, and such natural variation would represent one aspect of an ecological deviation from the core microbiome [6]. Although assessments of the microbiome of various arthropods is an area of very active research [7,8,9,10,11,12,13,14,15], the potential for natural, spatial variation in the microbiome has been elucidated in only a handful of insect species [8,10,16]—perhaps most thoroughly in the model species *Drosophila* [9,17]. Interro-

gating spatial variation in the core microbiome of a range of insects, and specifically arthropods that may vector disease is of interest for three main reasons: first, the maintenance of a cohort of microbial species across many insect populations may provide information about important components of host physiology; second, the potential for interaction between endosymbiont species may be elucidated by wide-ranging surveys of microbial taxon abundance and frequency of co-occurrence, and finally, surveys of the spatial variation in microbiomes may identify populations in regions of importance to human and/or agricultural health that may then be targeted for remediation.

Members of the Cimicidae family (bed bugs) are obligate ectoparasites obtaining blood meals from birds, bats and humans (Usinger 1966). The medically significant *Cimex lectularius*, the human bed bug, has made a now infamous resurgence in North America as well as worldwide [18,19,20,21]. Bed bug infestations, while not restricted to social or economic groups, can disproportionately impact the economically disadvantaged since the efficacy of less expensive methods of control, such as pesticide application, is low owing to widespread pyrethroid resistance [22,23]. While there is currently no evidence that bed bugs vector disease, human pathogens have been identified on the surface of the insects and in their excrement [24]. Further, the presence of this hematophagous insect in dwellings can lead to skin reactions, anaphylaxis, anxiety and psychological problems in humans [24,25,26], and thus bed

bugs are of broad and increasing concern from the standpoint of human health [27].

Endosymbiont microbial species in *C. lectularius* were first identified by microscopy in 1921 [28], and recent 16S rRNA surveys have identified two main endosymbionts of the bed bug to be *Wolbachia* [29,30,31] and an unnamed, γ -proteobacterium closely related to an endosymbiont species from the leafhopper *Euscelidius variegatus* (termed BEV-like symbiont by [29]). A high prevalence of *Wolbachia* infections has been found in North American and Japanese populations of *C. lectularius* (between 83–100%, [32]), and infection rates of the γ -proteobacterium are reported to vary from 0–100% among locations in Japan [33]. Further, 16S rRNA sequencing has identified the strain of *Wolbachia* found in *C. lectularius* to belong to the F supergroup [30,32,33]. Unlike the B supergroup of *Wolbachia* that parasitizes many arthropods, the F supergroup, described within both arthropods and nematodes, is not known to harbor any reproductive parasites [33]. Recent work has shown that *Wolbachia* is a nutritional mutualist of *C. lectularius* and provides the species with B vitamins essential for growth and reproduction [33]. For this reason, *Wolbachia* has been proposed as an obvious target for the biocontrol of this insect pest [33]. However, before such an endeavor can be expected to succeed, more information regarding the natural microbial community of *C. lectularius* is needed. A broad assessment and characterization of the core microbiome of bed bugs collected from an urban area is thus highly warranted.

Here, we surveyed the microbiome of bed bugs obtained from residences located in metropolitan Cincinnati, OH (Fig 1), which is currently listed as the second most bed bug infested US city [34,35]. To make our preliminary assessment as geographically comprehensive as possible, we surveyed between 2 and 5 *C. lectularius* individuals from each of eight geographically separated collection locations. We used culture-independent 16S rDNA PCR amplification followed by 454 Titanium sequencing to characterize the microbial community associated with each *C. lectularius* individual. We performed this survey to broadly characterize the core microbiome of naturally collected bed bug samples and to assess the following: Are there differences in the bed bug microbial community diversity and/or membership among collection locations? Are the two dominant and previously described endosymbionts (*Wolbachia* and an unnamed γ -proteobacterium) present across all collection locations? Is there any evidence for pathogenic microbial groups within collection locations of *C. lectularius* sampled from local residences?

Results

Data summary

Following DNA extraction from *C. lectularius* individuals, the hypervariable V4–V6 region of the bacteria-specific 16S rDNA gene was PCR amplified in triplicate for each individual. Libraries from the same individual were then combined, cleaned, normalized, and sequenced using 454 Titanium chemistry.

From the initial 69,854 sequences, 11,147 were removed for being <200 bp or low quality, 3,927 were identified as chimeric, and 666 were identified as either mitochondrial or chloroplast. After filtering the sequence reads for quality scores, sequencing errors and chimeras, our dataset consisted of 54,114 sequences distributed across 31 *C. lectularius* individuals (also referred to as libraries) (Table S1) with an average read length of 284 bases. The 31 libraries varied in size from 74 to 4092 sequences, and averaged 1745.6 ± 171.2 sequences. Most libraries (25) contained at least 1000 sequences. Clustering the data with mothur [36] at the 3% level of sequence divergence returned 359 OTUs across all

libraries. Two OTUs overwhelmingly dominated the sequence data, with one OTU containing 34,354 sequences and the other prominent OTU containing 18,085 sequences, which cumulatively constitute approximately 97% of all sequences. On average, OTUs contained 150 sequences, but many OTUs (239) were represented by 1 or 2 sequences. Twenty-nine OTUs were represented by 10 or more sequences. The taxonomic membership of collection locations (described below) was assessed using reads from all libraries and reported in Table 1, whereas analyses of community structure and diversity were performed following a random sampling of sequences from the libraries that contained at least 1000 sequences (see Methods; Table S1).

Bacterial taxa associated with wild *Cimex* populations

Approximately 97% of all sequences within our dataset were assigned to two families within the Proteobacteria by the RDP classifier: Anaplasmataceae (63.48%) and the Enterobacteriaceae (33.57%). Two genera—*Wolbachia* and an unclassified genus from the Enterobacteriaceae (hereafter ‘unnamed γ -proteobacterium’)—comprised the majority of these families, at 63.48% and 33.42%, respectively (Table 1). A blastn analysis of a representative sequence from the unnamed γ -proteobacterium shows high similarity (99%, e-value = $3e^{-139}$) to a previously identified maternally transmitted γ -proteobacteria endosymbiont from *C. lectularius* that was termed ‘BEV-like’ endosymbiont [29,33]. The remaining OTUs from our dataset comprise roughly 3.10% of the sequences (Table S2). All 31 individual *C. lectularius* libraries contained the dominant *Wolbachia* and the unnamed γ -proteobacterium; all individuals from eight collection locations had a shared OTU from each of the following genera at very low frequency: *Shinella*, *Sphingomonas*, *Methylobacterium*, *Pseudomonas*, a second unclassified genus from the Enterobacteriaceae (OTU 11), *Acinetobacter*, and *Cloacibacterium*, as well as a shared OTU from the suborder Propionibacterineae (Table 1 and Table S2). Thus, a core microbiome in wild-collected populations of *Cimex* could be expected to contain *Wolbachia* and the γ -proteobacterium and possibly OTUs from the low-frequency groups listed above. We screened our data for putative pathogens that were previously found to associate with *Cimex* individuals [24], and while the length of the 16S fragment that we analyzed precludes the identification of bacteria at the species level, we detected five bacterial OTUs from the same genera (*Bacillus*, *Coxiella*, *Staphylococcus*, and *Streptococcus*) as a potentially human-pathogenic bacterial species (Table 2).

Community structure of the *Cimex* microbial associates

While the core microbiome of *Cimex lectularius* individuals collected from urban Cincinnati is dominated by two OTUs from Proteobacteria (*Wolbachia* and the unnamed γ -proteobacterium), their relative abundances differed among collection locations (Fig 2A). *Wolbachia* was found to be highly abundant (84–85%) in locations 5 and 8, respectively, whereas the unnamed γ -proteobacterium was found at only 11–13% in these locations. In contrast, locations 6 and 3 exhibited slightly more of the unnamed γ -proteobacterium (56 and 49%) than *Wolbachia* (40 and 43%) (Fig 2A). Locations 2, 4, 7 and 9 showed similar relative abundances of these two genera on average (~60% *Wolbachia* and ~34% unnamed γ -proteobacterium). It is important to note that the relative abundance of *Wolbachia* and the γ -proteobacterium among collection locations reported here could be influenced by copy number variation of the 16S rDNA gene, either within or among populations. Although the 16S rDNA gene of *Wolbachia* is reported to be single copy from annotated genomes within the A, B and D supergroups [37,38,39] there is no record of copy



Figure 1. Map of *Cimex lectularius* collection locations from the greater Cincinnati metropolitan area in Cincinnati, OH, USA.
doi:10.1371/journal.pone.0061465.g001

number in the F supergroup of *Wolbachia*, nor is there a record for 16S rRNA copy number in the γ -proteobacterium.

The next 8 most abundant OTUs were detected in almost all locations but at very low frequency (Table 1). An OTU classified

as *Shinella*, from the Rhizobiales order, was found in location 3 at 1.63% and yet at only 0.2% in location 6 and at lower frequency in all other locations (Fig 2B). An OTU from the Frankineae suborder was present at relatively similar frequency among

Table 1. The relative abundance as the proportion of 16S rRNA sequences that mapped to the 10 most abundant OTUs shared across all or most locations.

| OTU ID Number | Phylum/Class/Suborder or Family | Genus | Abundance | Number of libraries OTU present | Number of locations OTU present |
|-----------------------|---------------------------------|--|-----------|---------------------------------|---------------------------------|
| Actinobacteria | | | | | |
| Actinobacteria | | | | | |
| 29 | Frankineae | Unclassified | 0.23 | 18 | 7 |
| 4 | Propionibacterineae | Unclassified | 0.11 | 21 | 8 |
| 31 | Streptomycineae | Unclassified | 0.07 | 17 | 7 |
| Proteobacteria | | | | | |
| Alphaproteobacteria | | | | | |
| 1 | Anaplasmataceae | <i>Wolbachia</i> | 63.48 | 31 | 8 |
| 26 | Rhizobiaceae | <i>Shinella</i> genera (<i>incertae sedis</i>) | 0.22 | 16 | 8 |
| 23 | Sphingomonadaceae | <i>Sphingomonas</i> | 0.13 | 22 | 8 |
| 15 | Methylobacteriaceae | <i>Methylobacterium</i> | 0.10 | 25 | 8 |
| Gammaproteobacteria | | | | | |
| 2 | Enterobacteriaceae | Unclassified | 33.42 | 31 | 8 |
| 85 | Pseudomonadaceae | <i>Pseudomonas</i> | 0.16 | 16 | 8 |
| 11 | Enterobacteriaceae | Unclassified | 0.15 | 20 | 8 |

The number of individual *C. lectularius* (i.e. libraries) that exhibit each OTU is presented along with the number of locations with the OTU.
doi:10.1371/journal.pone.0061465.t001

Table 2. Genera detected in the wild populations of *C. lectularius* previously detected in either wild or laboratory strains of bed bugs (¹from Delaunay et al 2011).

| Genus detected in this study | Number of OTUs identified | Potential pathogen previously detected in <i>C. lectularius</i> or <i>C. hemipterus</i> ¹ |
|------------------------------|---------------------------|--|
| <i>Bacillus</i> | 2 | <i>Bacillus anthracis</i> |
| <i>Coxiella</i> | 2 | <i>Coxiella burnetii</i> (Q fever) |
| <i>Staphylococcus</i> | 1 | <i>Staphylococcus aureus</i> |
| <i>Streptococcus</i> | 2 | <i>Streptococcus pneumoniae</i> |
| <i>Wolbachia</i> | 3 | <i>Wolbachia</i> spp |

doi:10.1371/journal.pone.0061465.t002

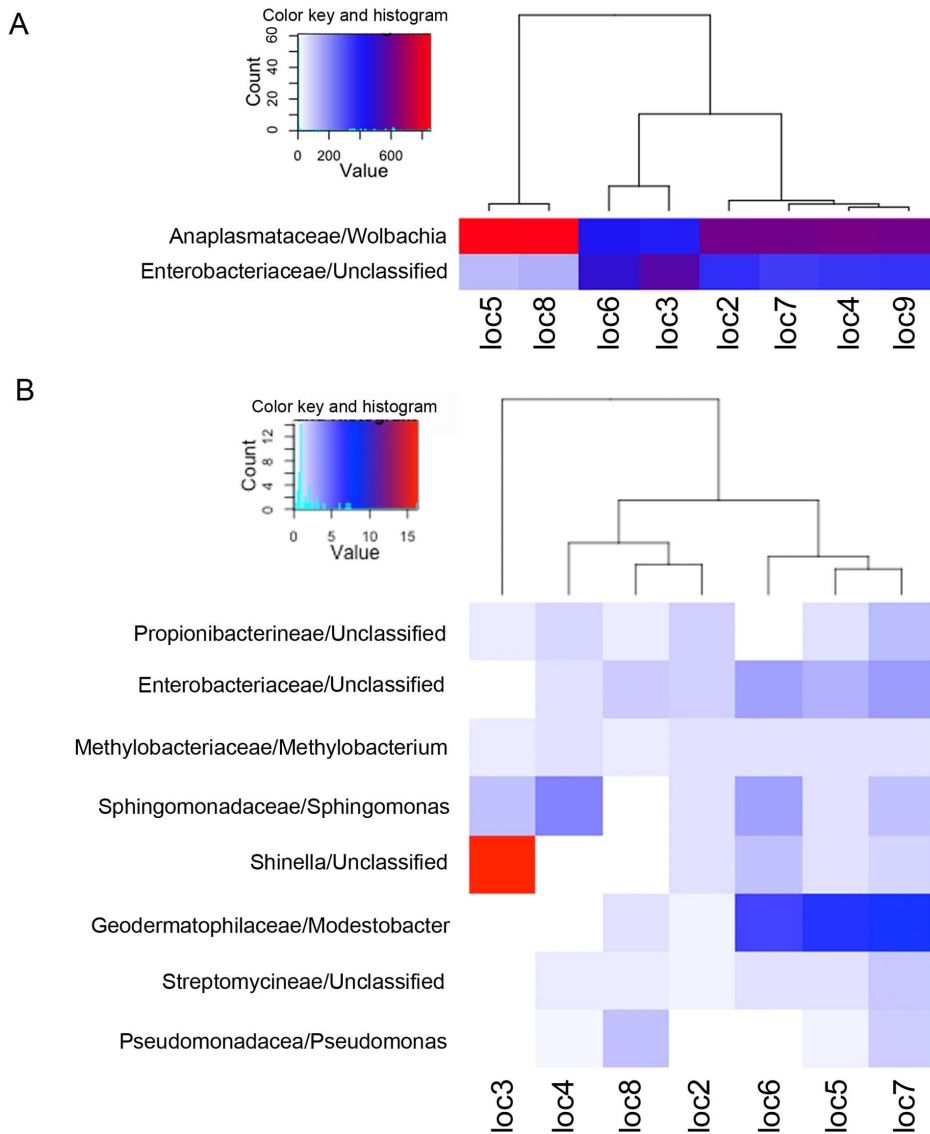


Figure 2. Composition and abundance of the *C. lectularius* microbiome within eight wild populations collected from Cincinnati, OH, USA. Each individual *C. lectularius* library was randomly subsampled to 1000 sequences, and the abundance of each OTU was averaged according to collection location. Note that location 6 is represented by a single individual. A. Relative abundance of the two most abundant bacterial families and genera within the locations. B. Relative abundance of the next 8 most abundant bacterial families and genera. Note the difference in scale in the respective keys for panels A and B.
 doi:10.1371/journal.pone.0061465.g002

locations 5, 6, and 7 (~0.68%). These same locations also exhibited a relatively similar frequency as an OTU that was another member of the Enterobacteriaceae. The majority of our bootstrap values from the RDP classifier provided 100% support, especially for the high-frequency OTUs, with the taxonomic assignment of a few low-frequency OTUs between 80–90% (see Table S2).

To determine if the patterns of abundance represent significant microbial community structuring across collection locations, we estimated the Yue and Clayton similarity coefficient (θ) [40] among the individual *Cimex* microbiomes. The Yue and Clayton θ index is a similarity index that includes species proportions of both the shared and non-shared species in each population. The index ranges from 0 to 1, with 1 = complete similarity and 0 = complete dissimilarity. We then visualized the resulting distance matrices using a principle components analysis (PCoA) and performed an analysis of molecular variance (AMOVA) to determine if the spatial separation in the PCoA plot was statistically significant, *i.e.*, individuals of the same location shared similar communities and at similar abundances. The PCoA plot (Fig 3; Movie S1) of this metric of community similarity shows evidence of structure in that individual microbiomes from the same location tended to cluster with one another; further, the AMOVA uncovered evidence that the variation among location spatial structure was greater than variation between individuals collected from the same location ($F_7 = 11.25$, $P = 0.008$).

Patterns of microbial diversity among the *Cimex* microbiomes

On average, 22.56 ± 2.48 (SE) OTUs were uncovered across the locations, with four libraries either approaching or exhibiting approximately 60 OTUs (Fig 4). On the lower end, three libraries appeared to asymptote between 10–15 OTUs. Rarefaction analysis shows that individual *Cimex* microbiomes vary in richness (Fig 4), and that the libraries were sampled at different depths. Those that were not sampled to completion could harbor rare and potentially important bacterial taxa. Coverage of the *Cimex* microbiomes averaged at 98.5%, suggesting that, on average,

1.5 unique OTUs would be expected for every additional 100 sequences.

Indices of diversity and evenness varied significantly among locations (Fig 5; Tables S3A&B, Table S4), and this variation appeared to be driven largely by the relative low diversity and evenness of locations 5 and 8. These two locations were dominated by *Wolbachia* relative to the unnamed γ -proteobacterium, and exhibited an average of 18.5 (location 5) and 14.3 (location 8) OTUs. Rarefaction analysis of the Inverse Simpson index shows that unequal sampling among *Cimex* libraries did not excessively influence this estimate of diversity (Figure S1).

Discussion

The core *Cimex* microbiome

We found two dominant OTUs to make up 97% of the *C. lectularius* microbiome. The RDP classifier assigned these OTUs to the α -proteobacteria *Wolbachia*, an obligate endosymbiont, and an unnamed BEV-like γ -proteobacterium, closely related to a symbiont of the leafhopper *Euscelidius variegatus*. Both taxa have been described from previous surveys of *C. lectularius*—they are transovarially transmitted, persist in bacteriocytes (both taxa) [29,33,41] and/or are found in many cell types of the bed bug (unnamed γ -proteobacterium [33]). *Wolbachia* comprised the largest fraction of our microbiome survey (~63%) and was present in all locations and individuals. Broader geographical assessments have previously indicated that *Wolbachia* infections are present in ~95% of surveyed *C. lectularius* populations [32], and earlier experimental work suggested a role for this symbiont in *C. lectularius* fertility [42,43,44]. More recently, it has been shown that the presence of *Wolbachia* in bed bugs provides a fertility benefit through the production of B vitamins [33] supporting the idea that this taxon is an obligate nutritional mutualist of the bed bug. Thus, the high frequency of *Wolbachia* uncovered in this study is relatively unsurprising, but provides evidence that our short-read 16S rDNA assessment is largely congruent with expectations.

While different sub-types of *Wolbachia* are known to inhabit many arthropods and can range in association with their hosts along a parasitic to mutualistic continuum [45,46,47], there is

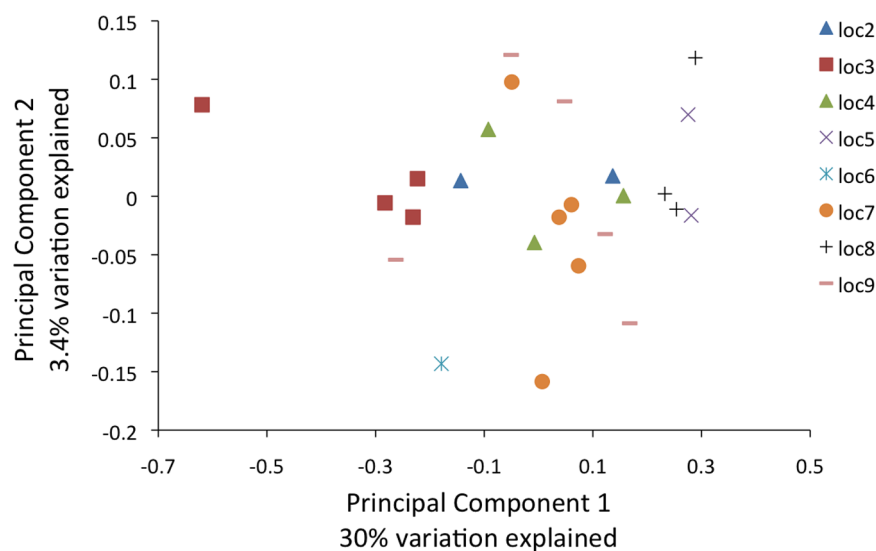


Figure 3. Principle component analysis of the *Cimex* microbiome. The Yue and Clayton measure of dissimilarity between the structures of the microbiome communities (θ -YC) was estimated and visualized using the `dist.shared` and `pcoa` commands of `mothur` [36]. doi:10.1371/journal.pone.0061465.g003

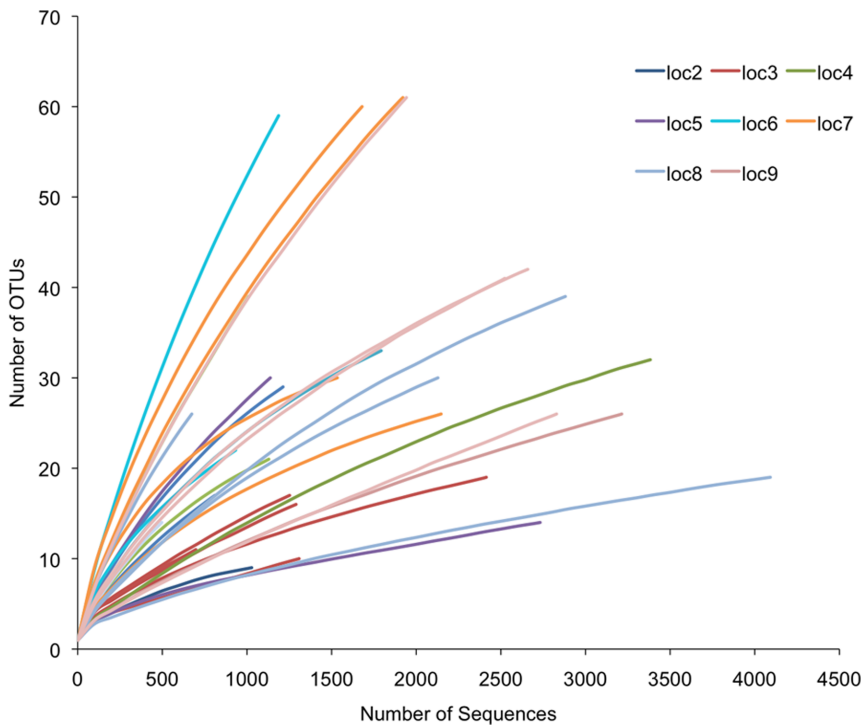


Figure 4. Rarefaction analysis of observed richness of the microbiome of *C. lectularius* individuals. All calculations were performed with mothur [36]. OTUs were defined at the 3% level of sequence divergence. Individuals collected from the same location are represented by the same color.

doi:10.1371/journal.pone.0061465.g004

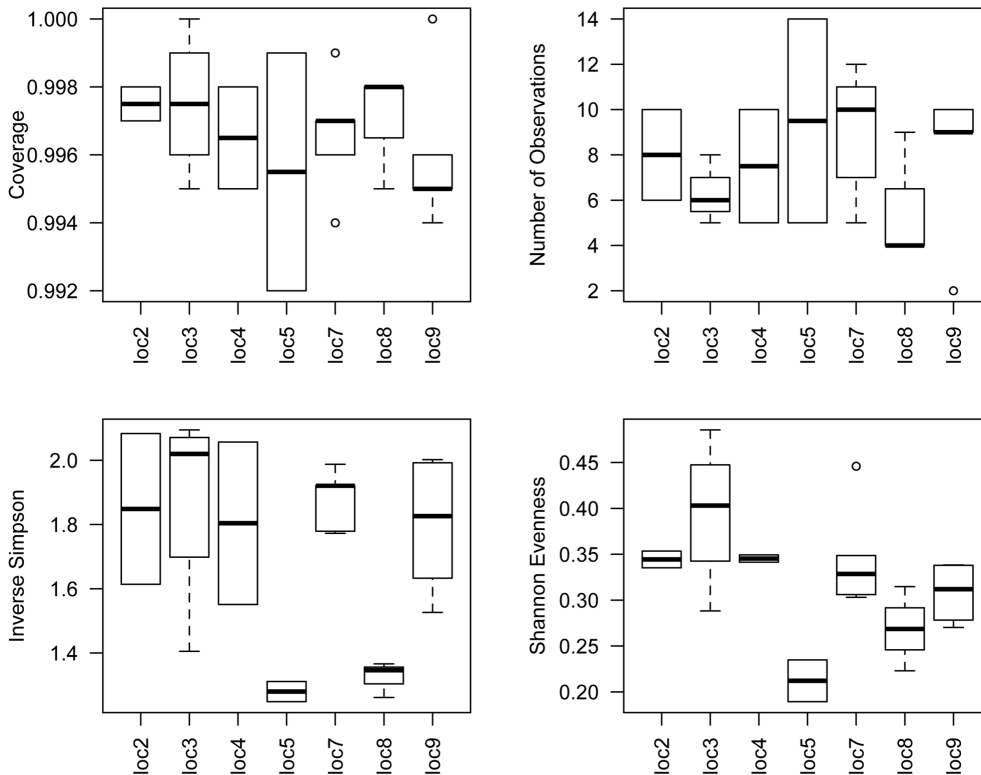


Figure 5. Boxplots of the diversity and evenness of the bacterial communities found in association with *C. lectularius* individuals, averaged according to collection location. Individual libraries were randomly subsampled to 1000 sequences each prior to estimating diversity parameters. Calculations were performed with mothur [36] at the 3% level of sequence divergence. The number of individuals within each population is presented in Table S1.

doi:10.1371/journal.pone.0061465.g005

comparatively less known about the lifestyle of the second most common OTU in our dataset, the unnamed γ -proteobacterium. Phylogenetic assessment of this taxon found it to exhibit greater than 99% similarity to BEV [29], a facultative endosymbiont reported from the leafhopper *Euscelidius variegatus* [48] and closely related to the plant pathogenic genera *Dickeya*, *Pectobacterium*, and *Erwinia* [49]. Close phylogenetic relationships between endosymbionts found in plant-feeding and blood-feeding insect hosts have been uncovered in other bacterial clades-*Rickettsia*, *Sodalis* and *Arsenophonus* [50,51,52]- and may represent putative examples of horizontal transfer, potentially due to their ability to persist outside of hosts [49]. While the phenotypic effects of this γ -proteobacteria on *C. lectularius* are largely unknown, there are accounts of pen-strep treatment (which does not affect *Wolbachia*) leading to reductions in egg production in *C. lectularius*, suggesting that the removal of the γ -proteobacteria can lead to reductions in fertility [41] and thus it may prove to be a nutritional mutualist as well.

Identification of putative human pathogens in *Cimex*

One of the goals of this survey was to determine if human pathogenic bacteria could be detected within *C. lectularius* collected in a dense urban setting. Because *C. lectularius* colonies infest walls and other cracks and crevices of structures, live in high density in their excrement, and shuttle around dust, the environmental microbes (*i.e.* those that can be found on the exoskeleton) were of interest as were potential pathogenic endosymbionts. Although bed bugs have not been definitively shown to transmit disease, 45 human pathogens with the potential for transmission through bed bug vectors, most likely mechanical in contrast to true vector, have been previously identified within or on the insects and/or their excrement [24]. Furthermore, bed bug bites and exposure to bed bug excrement can trigger systemic reactions such as asthma and anaphylaxis [25,26]. Our survey of the whole insect microbiome uncovered at least 5 genera to which known or putative human pathogen species belong. For example, two OTUs of *Bacillus* were detected in three of our eight collection locations surveyed-other blood-feeding arthropods (mosquitos, horse and deer flies) have been documented to vector *B. anthracis*, the bacteria that causes anthrax [53]. We also found two OTUs in one collection location from the family that *Coxiella burnetii* (Q-fever) belongs to, along with two *Streptococcus* OTUs found in five populations and one *Staphylococcus* OTU found in four collection locations. While the use of short-read 16S rDNA sequences does not enable identification of the bacterial species and strains of interest to human health, the use of such technology across a broad area, followed by either 16S rDNA Sanger sequencing and/or high-depth metagenomic sequencing could identify human pathogens at the level of the species.

Similar bacterial lineages across populations suggest host-microbe association

Our distance-based screen of 54,114 sequences from 31 individual *Cimex* libraries uncovered 359 OTUs-a number that is high relative to other surveys of arthropod bacterial communities (*e.g.*, 139 OTUs found across species and wild populations of *Drosophila*, [9]; 74 OTUs uncovered across 11 *Drosophila* populations, [17]; both Sanger sequenced, long-read surveys). Our use of multiple individuals collected across geographically separate locations, the method of sample preparation, and 454 sequencing error are all factors that likely influenced this high number. To reduce sequencing error on our estimates, we followed the standard operating procedure from the Schloss lab (http://www.mothur.org/wiki/Main_Page), which included sequence quality checks and trimming of low-quality sequence and homopolymers,

as well as a pre-clustering step that bins rare sequences that are within 2 bp of a more common OTU into the more common group. We also searched for and removed chimeric sequences within the database-such screens are reported to reduce the overall chimera rate to 1% [54]. Further, we utilized techniques that reduce PCR amplification errors such as using a high-fidelity Taq polymerase and the pooling of triplicate amplifications of each individual's DNA. With these checks in place, we still uncovered a large number of OTUs that were very low-copy (239 OTUs represented by 1 or 2 sequences). We removed these low-copy sequences from the overall dataset and performed each step of the presented analysis with OTUs represented by three or more sequences. This highly conservative treatment of the data uncovered 49 OTUs at the 3% level of divergence-many fewer than the 359 uncovered with the full dataset. However, all of our major conclusions were the same: collection locations exhibited variable diversity and we found a similar relative abundance of the 10 most abundant OTUs (data not shown). Although we cannot distinguish rare, environmental OTUs from 454 sequencing error with our full dataset, if we were to assume an error rate of 1% from non-identified chimeras, and another 18% error from 454 sequencing chemistry (as identified using the V4 region in [55]), we would reduce our number of OTUs by 68 to 291. Thus, it is reasonable to predict that many of the low-copy OTUs from across the eight collection locations are from environmental sources, such as bacteria picked up and transported on the surface of the individuals.

In comparison, the OTUs that we uncovered across all or most individuals and/or collection locations, and are represented by more than a few sequences, are likely to be either endosymbionts or consistently associated with the *C. lectularius* exoskeleton rather than strictly environmental. Ten such OTUs were uncovered in all locations, and 11 OTUs were found in 15 or more individuals, albeit most at very low frequency. Of note, *Methylobacterium* was found in 25 of the individual libraries and in all locations-this group of bacteria is often associated with plants. One species in particular, *M. mesophilicum*, is harbored by glassy-winged sharpshooters [56] and has been suggested as a sharpshooter biocontrol agent to control associated vector-borne plant diseases. An OTU classified as belonging to the *Sphingomonas* genus, having a widespread ecological distribution within water, soil and in association with plants [57], was uncovered in 22 individual bed bug libraries. Finally, in 21 libraries, we detected the presence of *Propionibacterium*, a genus of bacteria to which *P. acnes* belongs-these bacteria, which cause acne in humans, are often found on skin and infecting human blood [58,59].

Variation in the structure, diversity and evenness of the *Cimex* microbiome

We uncovered evidence of variation in diversity and evenness among the *Cimex* microbial collection locations. Whereas most of the Inverse Simpson values range from 1.8–2.1 among locations, locations 5 and 8 show comparatively lower values, from \sim 1.3 to 1.4. These results are likely influenced by the relative abundances of the most abundant OTUs-*Wolbachia* dominated these two locations, whereas the other locations exhibited less extreme relative abundances of *Wolbachia* versus the unnamed γ -proteobacterium. Further, locations 5 and 8 did not exhibit wide differences in the abundance of the low-copy OTUs that were similarly uncovered in most locations, suggesting that the significant variation among locations in diversity and evenness are driven largely by the differences in abundance of *Wolbachia* and the unnamed γ -proteobacterium. A survey of the geographical distribution of the microbial community from natural *Drosophila*

melanogaster populations has also uncovered population variation in community composition and taxon richness [17], and further, there is evidence that the major taxa of the lone star tick microbiome show heterogeneity among populations [10]. This survey of the *Cimex* microbiome is relatively unique compared to previous assessments [33] in that between 2 and 5 individuals were assessed separately per collection location. The use of multiple sequenced microbiomes per population thus allows us to definitely test for the presence of location differentiation among wild-caught *C. lectularius*, which has implications for understanding endosymbiont patterns of co-occurrence and for generating expectations regarding the frequency of potential human pathogenic bacteria.

The significant community structure among *C. lectularius* microbiomes from different collection locations that we uncovered in this survey may be due, in part, to the maternally inherited portion of the microbiome. Recent work shows that the population genetic structure of *C. lectularius* is high across the eastern US, with little variation within populations [60]. This pattern suggests multiple introductions of *C. lectularius* into the United States—and hence high diversity among populations—followed by subsequent inbreeding within populations due to the establishment of the population by a single mated female [60]. Our results of significant structure within the microbiome of *C. lectularius* among collection locations preliminarily supports this pattern since the dominant members of the microbiome are known to be maternally transmitted, and thus, the establishment of a new population via a single female would establish a maternally-inherited microbiome. That we uncovered community structure in the microbiomes between locations is even more striking in light of our experimental decision to assess the whole-insect microbiome, *i.e.*, environmental and endosymbiont bacteria together. However, this finding may also reflect differences among households in the bacteria present in the environment, and our data thus far cannot disentangle the relative influence of the environmentally-derived bacteria from the internal community on the overall microbiome structure.

Conclusions

The data presented here provide a novel view of the microbial constituents of bed bugs across collection locations. That the relative abundances of *Wolbachia* and the γ -proteobacterium appeared to vary across locations suggest a range of interesting questions, such as: Might the two dominant members of the bed bug microbiome competitively exclude one another, or, does host fitness depend on the presence of both types in an additive fashion? Are the patterns of community structure and relative abundance of the two dominant groups from naturally collected locations strictly due to inheritance from a founding maternal individual, or might there be unexamined environmental influences? How might these patterns of community structure impact potential attempts to eradicate or control bed bug infestations? Further assessments of the core microbiome across multiple established bed bug populations will have broad implications for our understanding of bed bug physiology, human health, and bed bug biocontrol.

Methods

Sample collection and DNA isolation

C. lectularius samples from distinct collection locations were obtained from residences in the greater Cincinnati (OH, USA) area by Scherzinger Pest Control and generously donated to the Baucom lab for this project (see Fig 1 for locations). Technicians

recorded the perceived severity of each infestation, which ranged from 5 (most severe) to 1 (least severe). Most donated populations were scored '5,' with the exception of a single population that was scored as a '3' (see Table S1 for specific information). The sizes of populations from collection locations were not otherwise characterized. In accordance with residential contracts, the pest company did not divulge the exact location of each residence; however, a general address was supplied for our records. Other identifying information about each residence, such as whether it was from an apartment building or single-family residence was not provided. Each *C. lectularius* population was collected from within a single residence. Between 10–50 individuals per collection location were randomly collected from the residence and placed in sterile 50 mL falcon tubes. Once collected, each sample of insects was starved for ~7 days before being snap frozen in liquid nitrogen and stored at -20°C until DNA isolation. Five individuals from each of the eight collection locations were randomly selected and weighed in sterile conditions. DNA was isolated from each individual using the Powersoil DNA Isolation Kit (MO Bio, Carlsbad, CA, USA) and tested for purity on a NanoDrop 2000 Spectrophotometer (ThermoScientific, Waltham, MA, USA). Across collection locations, average DNA yield was $3.97 (\pm 0.62 \text{ SE}) \text{ ng/uL}$ with an average 260/280 score of $2.30 (\pm 0.11 \text{ SE})$. Individuals were not surface sterilized prior to DNA isolation since the whole insect microbial community, and the community the insects potentially transport on their exteriors and/or in their feces were of interest.

16S Library and Sequencing

16S rDNA genes were PCR amplified in triplicate using the bacterial universal primers 530F (5'-GTGCCAGCMG-CNGCGG-3') and 1100R (5'-GGGTTNCGNTCGTTG-3') which amplified ~600 bp corresponding to the V4–V6 hypervariable region of *Escherichia coli* [61]. DNA from each individual was amplified by PCR in order to create amplicon libraries for pyrosequencing using Titanium GS FLX chemistry on a 454 sequencing machine. Primers thus included the Titanium Fusion A or B primers and a 15 bp error-correcting mid-tag optimized for individual amplicon library identification from a single pyrosequencing run (Roche, Penzberg, DE); the general sequence of each primer was LinkerA-midtag-530F and LinkerB-1100R (see Table S2 for specific primers utilized in this study [62]). The HotStar HiFidelity Polymerase Kit (Qiagen, Valencia, CA, USA) was used for PCR under the following conditions: 95°C for 5 min followed by 40 cycles of 94°C for 15 s; 60°C for 1 min; 72°C for 1 min; and a final elongation of 72°C for 10 min. Each sample was amplified in triplicate, pooled and cleaned using the QIAquick PCR Purification Kit (Qiagen, Valencia, CA, USA) followed by normalization of each library to 25 ng using the SequalPrep Normalization Plate Kit (Invitrogen, Carlsbad, CA, USA). After amplification, the PCR products were run on a 1% agarose gel, and samples that did not amplify well in triplicate were not included in subsequent steps. This reduced the total number of libraries to 31 (see STable1 for the number of libraries per collection location). Once normalized, libraries were pooled and sent to the Purdue Genomics Lab for pyrosequencing according to standard protocols (Roche, Penzberg, Germany).

Sequence analysis

Sequence quality control was performed using mothur [36] in the following manner: first, the trim.seqs command was used to trim the sequence when the average quality score over a 50 bp window dropped below 35, and sequences with ambiguous bases (N), homopolymers (>8 bases), or sequences that were <200 bp were removed. The pdiffs=2 and bdiffs=1 options of this

command were utilized in order to allow a maximum of 2 and 1 nucleotide differences in the primer and barcode sequences, respectively. Sequences were aligned to the SILVA-compatible alignment database using `align.seqs`, and then trimmed to a common region. To further reduce the influence of pyrosequencing error on downstream analysis, sequences were clustered using the `pre.cluster` command, and those that were within 2 bp of similarity to a more abundant sequence were merged with the more abundant type. Next, potential chimeric sequences were removed from the full dataset by utilizing the SILVA gold database as the reference in the `chimera.uchime` command. Finally, the `classify.seqs` and `remove.lineage` commands were utilized to identify and remove potential mitochondrial and chloroplast contaminants.

Bacterial taxonomy was assigned to each sequence in the improved dataset using the `classify.seqs` command, which implements the Naïve Bayesian classifier of RDP, as implemented in `mothur` [36]. Following taxonomic assignment, sequences were assigned to OTUs using the `cluster` command, following which the consensus taxonomy for each OTU at the 3% level of divergence was determined using the `classify.otu` command [36]. The abundance of each OTU_{0.03} within each individual and collection location was determined using the `shared.otu` command. Broad patterns of the presence or absence of specific OTUs across all *Cimex* populations were considered utilizing the improved, full dataset. To compare estimates of diversity and community structure among collection locations, we first normalized the data such that 1000 sequences were randomly subsampled from each individual. Individuals with <1000 sequences following quality control were thus excluded from further analyses (Table S1).

Commands within the program `mothur` were used to produce rarefaction curves and assess both alpha- and beta-diversity of the communities. The coverage, number of observed OTU_{0.03}, the Shannon evenness and the Inverse Simpson statistics were calculated for each individual, and an ANOVA using Type I SS was performed in the R statistical programming language (<http://www.R-project.org>) to determine if collection locations varied in these estimates of microbial community diversity, after \log_{10} transforming dependent variables to improve the normality of residuals. To describe the community dissimilarity, (and conversely, similarity in community membership and structure across locations) the Yue and Clayton theta statistic was estimated via the `dist.shared` command. The resulting distance matrix was visualized in a principle components ordination plot using the first two axes, and then a rotation of the 3D matrices was performed using the `ggobi` package in R [63].

We next performed an analysis of molecular variance using the `amova` command of `mothur` to determine if there was significant structure among collection locations, followed by the `metastats` command to identify OTU_{0.03} that were differentially represented

in the separate *Cimex* populations. The presence of a core microbiome was assessed by first using the `make.shared` command to determine which OTUs were shared across all individuals and by then visualizing the abundance of OTUs using the `heatmap.2` command of R. Sequences from this study are deposited in Dryad (doi:10.5061/dryad.rd240) and are also available from RS Baucom upon request.

Supporting Information

Figure S1 Rarefaction of the Inverse Simpson index for each *Cimex* library.

(TIF)

Table S1 General location of populations and the number of individuals used in the taxonomic, diversity and abundance screens.

(DOCX)

Table S2 The abundance of all OTUs within the *Cimex* populations.

(XLSX)

Table S3 Analysis of variance for assessing the effect of location on variation in estimates of A. Inverse Simpson and B. Shannon Evenness indices.

(DOCX)

Table S4 The coverage, number of observed OTUs, and the Inverse Simpson, Shannon Evenness, and the non-parametric Shannon diversity index.

(XLSX)

Movie S1 A rotating PCoA plot of community similarity (Yue and Clayton theta) among individual bed bug microbiomes.

The rotation of the first 2 axes from the principle components ordination plot was performed by `ggobi` package in R. (MOV)

Acknowledgments

We thank E. Buschbeck for providing general advice on culturing an arthropod, P. Boland from Scherzinger Pest Control for performing the collections and providing the samples, and V. Corby-Harris, members of the Baucom lab, and two anonymous reviewers for critiques that helped improve a previous version of this manuscript. Funding was provided by the University of Cincinnati.

Author Contributions

Conceived and designed the experiments: MM SM RSB RR. Performed the experiments: MM SM RSB RR. Analyzed the data: MM RSB RR. Contributed reagents/materials/analysis tools: MM SM RR. Wrote the paper: RR RSB.

References

- Buchner P (1965) Endosymbiosis of animals with plant microorganisms. New York: Interscience Publishers.
- Dillon RJ, Dillon VM (2004) The gut bacteria of insects: Nonpathogenic interactions. *Annual Review of Entomology* 49: 71–92.
- Weiss B, Aksoy S (2011) Microbiome influences on insect host vector competence. *Trends in Parasitology* 27: 514–522.
- Hamady M, Knight R (2009) Microbial community profiling for human microbiome projects: Tools, techniques, and challenges. *Genome Research* 19: 1141–1152.
- Turnbaugh PJ, Ley RE, Hamady M, Fraser-Ligggett CM, Knight R, et al. (2007) The Human Microbiome Project. *Nature* 449: 804–810.
- Shade A, Handelsman J (2012) Beyond the Venn diagram: the hunt for a core microbiome. *Environmental Microbiology* 14: 4–12.
- Andreotti R, de Leon AAP, Dowd SE, Guerrero FD, Bendele KG, et al. (2011) Assessment of bacterial diversity in the cattle tick *Rhipicephalus (Boophilus) microplus* through tag-encoded pyrosequencing. *BMC Microbiology* 11(1):6.
- Carpi G, Cagnacci F, Wittekindt NE, Zhao F, Qj J, et al. (2011) Metagenomic profile of the bacterial communities associated with *Ixodes ricinus* ticks. *PLoS one* 6: e25604–e25604.
- Chandler JA, Lang JM, Bhatnagar S, Eisen JA, Kopp A (2011) Bacterial communities of diverse *Drosophila* species: Ecological context of a host-microbe model system. *Plos Genetics* 7(9): e1002272.
- Clay K, Klyachko O, Grindle N, Civitello D, Oleske D, et al. (2008) Microbial communities and interactions in the lone star tick, *Amblyomma americanum*. *Molecular Ecology* 17: 4371–4381.
- Hail D, Lauziere I, Dowd SE, Bextine B (2011) Culture independent survey of the microbiota of the glassy-winged sharpshooter (*Homalodisca vitripennis*) using 454 pyrosequencing. *Environmental Entomology* 40: 23–29.

12. Ishak HD, Miller JL, Sen R, Dowd SE, Meyer E, et al. (2011) Microbiomes of ant castes implicate new microbial roles in the fungus-growing ant *Trachymyrmex septentrionalis*. *Scientific Reports* 1:204.
13. Schloss PD, Delalibera I, Handelsman J, Raffa KF (2006) Bacteria associated with the guts of two wood-boring beetles: *Anoplophora glabripennis* and *Saperda vestita* (Cerambycidae). *Environmental Entomology* 35: 625–629.
14. Suen G, Scott JJ, Aylward FO, Adams SM, Tringe SG, et al. (2010) An insect herbivore microbiome with high plant biomass-degrading capacity. *PLoS Genetics* 6(9): e1001129.
15. Terenius O, Lindh JM, Eriksson-Gonzales K, Bussiere L, Laugen AT, et al. (2012) Midgut bacterial dynamics in *Aedes aegypti*. *FEMS Microbiology Ecology* 80: 556–565.
16. Estes AM, Hearn DJ, Burrack HJ, Rempoulakis P, Pierson EA (2012) Prevalence of candidatus *Erwinia dacicola* in wild and laboratory olive fruit fly populations and across developmental stages. *Environmental Entomology* 41: 265–274.
17. Corby-Harris V, Pontaroli AC, Shimkets LJ, Bennetzen JL, Habel KE, et al. (2007) Geographical distribution and diversity of bacteria associated with natural populations of *Drosophila melanogaster*. *Applied and Environmental Microbiology* 73: 3470–3479.
18. Gangloff-Kaufmann J, Hollingsworth C, Hahn J, Hansen L, Kard B, et al. (2006) Bed bugs in America: A pest management industry survey. *American Entomologist* 52: 105–106.
19. Doggett SL, Geary M, Russell R (2004) The resurgence of bed bugs in Australia: with notes on their ecology and control. *Environmental Health* 4: 30–38.
20. Boase CJ (2004) Bed bugs-reclaiming our cities. *Biologist* 51: 9–12.
21. Lee IH, Ree H, An S, Linton JA, Yong T (2008) Reemergence of the bed bug *Cimex lectularius* in Seoul, Korea. *Korean Journal of Parasitology* 46: 269–271.
22. Potter MF (2005) A bed bugs state of mind: emerging issues in bed bug management. *Pest Control Technology* 33: 82–97.
23. Reinhardt K, Siva-Jothy MT (2007) Biology of the bed bugs (Cimicidae). *Annual Review of Entomology*. pp. 351–374.
24. Delaunay P, Blanc V, Del Giudice P, Levy-Bencheton A, Chosidow O, et al. (2011) Bedbugs and infectious diseases. *Clinical Infectious Diseases* 52: 200–210.
25. Goddard J, deShazo R (2009) Bed bugs (*Cimex lectularius*) and clinical consequences of their bites. *JAMA-Journal of the American Medical Association* 301: 1358–1366.
26. Pinto LJ, Cooper R, Kraft SK (2007) Bed bug hand-book: the complete guide to bed bugs and their control. Mechanicsville, MD: Pinto & Associates, Inc.
27. Rossi L, Jennings S (2010) Bed bugs: a public health problem in need of a collaborative solution. *Journal of Environmental Health* 72: 34–35.
28. Arkwright JA, Atkin EE, Bacot A (1921) An hereditary *Rickettsia*-like parasite of the bed bug (*Cimex lectularius*). *Parasitology* 13: 27–36.
29. Hypsa V, Aksoy S (1997) Phylogenetic characterization of two transovarially transmitted endosymbionts of the bedbug *Cimex lectularius* (Heteroptera: Cimicidae). *Insect Molecular Biology* 6: 301–304.
30. Rasgon JL, Scott TW (2004) Phylogenetic characterization of *Wolbachia* symbionts infecting *Cimex lectularius* L. and *Oeciacus vicarius* Horvath (Hemiptera: Cimicidae). *Journal of Medical Entomology* 41: 1175–1178.
31. Sakamoto JM, Feinstein J, Rasgon JL (2006) *Wolbachia* infections in the Cimicidae: Museum specimens as an untapped resource for endosymbiont surveys. *Applied and Environmental Microbiology* 72: 3161–3167.
32. Sakamoto JM, Rasgon JL (2006) Geographic distribution of *Wolbachia* infections in *Cimex lectularius* (Heteroptera: Cimicidae). *Journal of Medical Entomology* 43: 696–700.
33. Hosokawa T, Koga R, Kikuchi Y, Meng XY, Fukatsu T (2010) *Wolbachia* as a bacteriocyte-associated nutritional mutualist. *Proceedings of the National Academy of Sciences of the United States of America* 107: 769–774.
34. Anonymous ODoH (2011) Ohio Bedbug Workgroup. In: Program ZD, editor. Cincinnati, OH.
35. Eddy C, Jones SC (2011) Bed bugs, public health, and social justice: Part 1, A call to action. *Journal of Environmental Health* 73: 8–14.
36. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, et al. (2009) Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology* 75: 7537–7541.
37. Foster J, Ganatra M, Kamal I, Ware J, Makarova K, et al. (2005) The *Wolbachia* genome of *Brugia malayi*: Endosymbiont evolution within a human pathogenic nematode. *PLoS Biology* 3: 599–614.
38. Klason L, Westberg J, Sapountzis P, Naesiund K, Lutnaes Y, et al. (2009) The mosaic genome structure of the *Wolbachia* wRi strain infecting *Drosophila simulans*. *Proceedings of the National Academy of Sciences of the United States of America* 106: 5725–5730.
39. Wu M, Sun LV, Vamathevan J, Riegler M, Deboy R, et al. (2004) Phylogenomics of the reproductive parasite *Wolbachia pipiensis* wMel: A streamlined genome overrun by mobile genetic elements. *PLoS biology* 2: 327–341.
40. Yue JC, Clayton MK (2005) A similarity measure based on species proportions. *Communications in Statistics-Theory and Methods* 34: 2123–2131.
41. Sakamoto JM, Rasgon JL (2006) Endosymbiotic bacteria of bed bugs: Evolution, ecology and genetics. *American Entomologist* 52: 119–122.
42. Chang KP (1974) Effects of elevated temperature on the mycetome and symbiotes of bed bug *Cimex lectularius* (Heteroptera). *Journal of Invertebrate Pathology* 23: 333–340.
43. De Meillon B, Golberg L (1947) Preliminary studies on the nutritional requirements of the bedbug (*Cimex lectularius* L.) and the tick *Ornithodoros moubata* Murray. *The Journal of experimental biology* 24: 41–63.
44. Takano-Lee M, Velten RK, Edman JD, Mullens BA, Clark JM (2003) An automated feeding apparatus for in vitro maintenance of the human head louse, *Pediculus capitis* (Anoplura: Pediculidae). *Journal of Medical Entomology* 40: 795–799.
45. Weeks AR, Turelli M, Harcombe WR, Reynolds KT, Hoffmann AA (2007) From parasite to mutualist: Rapid evolution of *Wolbachia* in natural populations of *Drosophila*. *PLoS biology* 5: 997–1005.
46. Werren JH (1997) Biology of *Wolbachia*. *Annual Review of Entomology* 42: 587–609.
47. Werren JH, Zhang W, Guo LR (1995) Evolution and phylogeny of *Wolbachia* - reproductive parasites of arthropods. *Proceedings of the Royal Society B-Biological Sciences* 261: 55–63.
48. Campbell BC, Purcell AH (1993) Phylogenetic affiliation of BEV, a bacterial parasite of the leafhopper *Euscelidius variegatus*, on the basis of 16S rRNA sequences. *Current Microbiology* 26: 37–41.
49. Degnan PH, Bittleston LS, Hansen AK, Sabree ZL, Moran NA, et al. (2011) Origin and examination of a leafhopper facultative endosymbiont. *Current Microbiology* 62: 1565–1572.
50. Novakova E, Hypsa V (2007) A new *Sodalis* lineage from bloodsucking fly *Craterina melbae* (Diptera, Hippoboscoidea) originated independently of the tsetse flies symbiont *Sodalis glossinidius*. *FEMS Microbiology Letters* 269: 131–135.
51. Novakova E, Hypsa V, Moran NA (2009) *Arsenophonus*, an emerging clade of intracellular symbionts with a broad host distribution. *BMC Microbiology* 9:143.
52. Perlman SJ, Hunter MS, Zchori-Fein E (2006) The emerging diversity of *Rickettsia*. *Proceedings of the Royal Society B-Biological Sciences* 273: 2097–2106.
53. Jensen GB, Hansen BM, Eilenberg J, Mahillon J (2003) The hidden lifestyles of *Bacillus cereus* and relatives. *Environmental Microbiology* 5: 631–640.
54. Schloss PD, Gevers D, Westcott SL (2011) Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies. *PLoS ONE* 6(12):27310.
55. Huse SM, Huber JA, Morrison HG, Sogin ML, Mark Welch D (2007) Accuracy and quality of massively parallel DNA pyrosequencing. *Genome Biology* 8:R143.
56. Gai CS, Lacava PT, Quecine MC, Auric M-C, Spotti Lopes JR, et al. (2009) Transmission of *Methylobacterium mesophilicum* by *Bucephalogramma xanthophis* for paratransgenic control strategy of citrus variegated chlorosis. *Journal of Microbiology* 47: 448–454.
57. White D, Sutton S, DB R (1996) The genus *Sphingomonas*: physiology and ecology. *Current Opinion in Biotechnology* 7: 301–306.
58. Park HJ, Na S, Park SY, Moon SM, Cho O-H, et al. (2011) Clinical significance of *Propionibacterium acnes* recovered from blood cultures: Analysis of 524 episodes. *Journal of Clinical Microbiology* 49: 1598–1601.
59. Williams HC, Dellavalle RP, Garner S (2012) Acne vulgaris. *Lancet* 379: 361–372.
60. Saenz VL, Booth W, Schal C, Vargo EL (2012) Genetic analysis of bed bug populations reveals small propagule size within individual infestations but high genetic diversity across infestations from the eastern United States. *Journal of Medical Entomology* 49: 865–875.
61. Baker G, Smith J, Cowan D (2003) Review and re-analysis of domain-specific 16S primers. *Journal of Microbiological Methods* 55: 541–555.
62. Acosta-Martinez V, Dowd S, Sun Y, Allen V (2008) Tag-encoded pyrosequencing analysis of bacterial diversity in a single soil type as affected by management and land use. *Soil Biology & Biochemistry* 40: 2762–2770.
63. Swayne DF, Lang DT, Buja A, Cook D (2003) GGobi: evolving from XGobi into an extensible framework for interactive data visualization. *Computational Statistics & Data Analysis* 43: 423–444.