Arch Microbiol (2015) 197:603–612
DOI 10.1007/s00203-015-1094-3

ORIGINAL PAPER

Microbial community structure in the gut of the New Zealand insect Auckland tree weta (*Hemideina thoracica*)

David W. Waite · Melissa Dsouza · Kristi Biswas · Darren F. Ward · Peter Deines · Michael W. Taylor

Received: 18 December 2013 / Revised: 8 February 2015 / Accepted: 11 February 2015 / Published online: 21 February 2015

Abstract The endemic New Zealand weta is an enigmatic insect. Although the insect is well known by its distinctive name, considerable size, and morphology, many basic aspects of weta biology remain unknown. Here, we employed cultivation-independent enumeration techniques and rRNA gene sequencing to investigate the gut microbiota of the Auckland tree weta (*Hemideina thoracica*). Fluorescence in situ hybridisation performed on different sections of the gut revealed a bacterial community of fluctuating density, while rRNA gene-targeted amplicon pyrosequencing revealed the presence of a microbial community containing high bacterial diversity, but an apparent absence of archaea. Bacteria were further studied using full-length 16S rRNA gene sequences, with statistical testing of bacterial community membership against publicly available termite- and cockroach-derived sequences, revealing that the weta gut microbiota is similar to that of cockroaches. These data represent the first analysis of the weta microbiota and provide initial insights into the potential function of these microorganisms.

Keywords Insect · Gut · Microbiota · Pyrosequencing · rRNA genes

Communicated by Andreas Brune.

Electronic supplementary material The online version of this article (doi:10.1007/s00203-015-1094-3) contains supplementary material, which is available to authorized users.

D. W. Waite · M. Dsouza · K. Biswas · P. Deines · M. W. Taylor (✉)
Centre for Microbial Innovation, School of Biological Sciences, The University of Auckland, Private Bag 92019, Auckland 1142, New Zealand
e-mail: mw.taylor@auckland.ac.nz

D. F. Ward
Landcare Research, Private Bag 92170, Auckland Mail Centre, Auckland 1142, New Zealand

Present Address:
P. Deines
Institute of Natural and Mathematical Sciences, Massey University, Private Bag 102904, Auckland 0745, New Zealand

Introduction

Insects represent one of the most successful branches of eukaryotic life, with an estimated 4–6 million extant species (Gaston 1991; Novotny et al. 2002). Insects have branched into a diverse range of niches and environments, occupying almost every trophic level as herbivores, carnivores, and decomposers. In addition to the innate interest in their biology, they have proved to be of relevance as a source of novel enzymes, capable of performing activities required in human industries (Matsui et al. 2009; Oppert et al. 2010; Willis et al. 2010). Due to New Zealand's ancient geographic isolation (Neall and Trewick 2008), it has developed native fauna that is unlike that of any other country, with approximately 80 % of native species being endemic (Gibbs 2006). Prior to human settlement, there was very little mammalian life, which allowed the resident avian and insect populations to expand into niches they do not traditionally occupy (Griffin et al. 2011b). Among these are the endemic weta, of the insect order Orthoptera. While the behavioural and physiological biology of the weta are well studied (Kelly 2011; Sinclair and Wharton 1997; Wehi et al. 2013; Wharton 2011), their diet is poorly understood (Cary 1983; Trewick and Morgan-Richards 1995; Wehi and Hicks 2010; Wilson and Jamieson 2005), although most species of weta are considered to be herbivores and opportunistic omnivores (Griffin et al. 2011a).

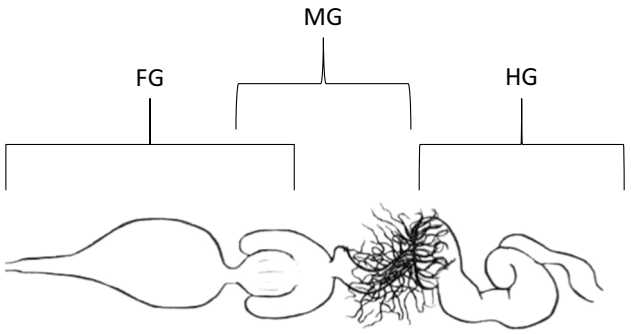


Fig. 1 Schematic drawing of the gut structure of the weta. Demarcations show dissected components: *FG* foregut, *MG* midgut, and *HG* hindgut

The role of microbes within their host organisms has been studied for a range of insects, including studies on orthopteran insects (crickets), the closest relative to weta, and dictyopteran insects (cockroaches and termites) (Idowu et al. 2009; Santo Domingo et al. 1998a, b), which share a broadly similar diet to tree weta (Broderick et al. 2004; Grünwald et al. 2010; Ohkuma 2003, 2008). Tree weta feed primarily on leaf litter, although they have been known to feed on seeds, fruit, and even prey on other insects in captivity (Griffin et al. 2011a; Trewick and Morgan-Richards 1995). The anatomy of the weta has been extensively documented, and the general compartmentalisation of the gut is well established; it is typically 'orthopteran' with the alimentary canal consisting of a foregut, midgut, and hindgut (Fontanetti and Zefa 2000; Maskell 1927) (Fig. 1). By contrast, nothing is known about the microbes that potentially inhabit the weta gut.

In this study, we sought to identify the microbial community density and membership within this iconic New Zealand insect, as well as compare the weta microbiota to that of other commonly studied invertebrates, such as termites and cockroaches. The data described here are the first of their kind for weta and provide a foundation for future studies into the activities of the gut microbiota and their potential roles in the ecology of the host.

Materials and methods

Sample collection and preparation

Adult Auckland tree weta were collected from a suburban garden in Meadowbank, Auckland, New Zealand, preserved in 100 % acetone (for DNA-based analyses). Insects were confirmed as Auckland tree weta through morphological identification then weighed and dissected under sterile conditions at the University of Auckland. For those weta individuals that were used for fluorescence in situ hybridisation (FISH), the gut was separated into four sections: foregut (including the crop and the proventriculus), midgut, and hindgut. Each gut section was weighed and then fixed for FISH by incubating in 4 % paraformaldehyde for 3 h, followed by washing twice with phosphate-buffered saline (PBS) and storing in 96 % ethanol/PBS [1:1 (v/v)] at −20 °C. Fixed samples were filtered onto 0.22-µm pore size polycarbonate membrane filters (diameter 25 mm, Millipore Ltd) and air-dried.

FISH-based counts of microbial cells

The aforementioned filters were cut into sections (~10 mm chord length) with a razor blade and put on a glass slide wrapped with Parafilm. Samples were hybridised with the Cy3-labelled *Bacteria* probe mix EUB338 I–III (Amann et al. 1990; Daims et al. 1999), and all probes were added at a concentration of 3 ng/µl, using a formamide concentration of 35 %. Hybridisation was performed in an isotonically equilibrated humidity chamber at 46 °C for 120 min. The filter pieces were then incubated for 10 min in a preheated washing buffer for 10 min at 48 °C. After rinsing filter pieces with distilled water and air-drying, samples were counterstained with a DAPI (4′,6-diamidino-2-phenylindole) solution (1 µmol/ml) for 10 min. After rinsing and drying, filter sections were mounted in a mixture of Citifluor (ProSciTech, Australia) and Vectashield (Vector Laboratories Inc., Canada). Hybridised filter sections were analysed using a Leica DMR epifluorescence microscope, with at least 300–500 cells manually counted for each sample.

DNA extraction and sequencing

Genomic DNA was extracted from whole gut homogenates by bead-beating in an ammonium acetate buffer (Taylor et al. 2004). In addition, DNA was extracted from the weta head using the same method and representative sequences of the weta 18S rRNA gene were amplified using the primer sets NS1 and EukA (Diez et al. 2001; White et al. 1990). The resulting amplicons were purified by gel extraction and sequenced directly. Sequences were identified using the NCBI online BLAST tool, classifying against the nucleotide collection (nr) database and uploaded to DDBJ/EMBL/GenBank databases under accession numbers KJ755445 and KJ755446.

For overall microbial identification, universal small-subunit rRNA gene amplification was performed using three primer pairs to separately target bacteria (27F/1391R), archaea (4aF/1391R), and eukaryotes (515F/1209R) (Woyke and Smith 2008). Roche 454 pyrosequencing was performed by the DOE Joint Genome Institute (California, USA). In addition, near-full-length bacterial 16S rRNA gene sequences were generated using the

previously described primers 616 V (targeting positions 8–25 of the *Escherichia coli* 16S rRNA gene) and 1492R, which amplify a ~1500-bp region of the gene (Polz and Cavanaugh 1998; Spring et al. 1998). Cycling conditions were as follows: initial denaturing of 94 °C for 5 min, then 30 cycles of denaturing at 94 °C for 45 s, annealing at 57 °C for 45 s, and elongation at 72 °C for 1.5 min. PCR was completed with a final elongation step at 72 °C for 10 min. PCR products were cloned with the pGEM T-easy vector (Promega) and *E. coli* DH5α chemically competent cells (Invitrogen) following the manufacturers' instructions. Inserts were sequenced from both ends by Macrogen Inc. (Seoul, South Korea).

Bioinformatic analysis

Following sequencing, clone inserts were assembled by aligning both ends of the gene, followed by manual quality curation in Geneious, version 7.1 (Kearse et al. 2012). Chimeras were removed using mothur, and the remaining 87 near-full-length 16S rRNA gene sequences were analysed using ARB with the SILVA 111 SSU database (Ludwig et al. 2004; Pruesse et al. 2007). High-quality sequences were uploaded to DDBJ/EMBL/GenBank under accession numbers KF318219–KF318305. Phylogenetic affiliations were analysed by constructing maximum likelihood trees, and robustness of branches was assessed by 5000 iterations of maximum parsimony bootstrapping. In order to compare the gut microbiota of weta to that of other insects, all bacterial sequences in the SILVA 111 SSU database that were obtained from cockroach or termite guts were exported and manually assigned to a host species based on the associated metadata. Sequences whose origin could not be assigned to (host) species level were discarded. Sequence data were then aligned, and unweighted UniFrac distances were calculated between the communities using 1000 iterations of subsampling to 30 sequences per sample (the smallest group containing 32 sequences). Principle coordinate analysis was performed on the resulting distance matrix and plotted in the R software environment (Team 2012).

For amplicon pyrosequencing, all bioinformatic analysis was performed using mothur, following the standard operating procedure (Schloss et al. 2011), with the exception of taxonomic classification. Flowgrams were denoised, and sequences were classified against the SILVA SSU database (version 119) using the inbuilt naïve Bayesian approach (Wang et al. 2007). Data were split according to domain-level match (prokaryote or eukaryote), and each group was analysed according to the mothur standard operating procedure using the appropriate alignment databases. Taxonomic classification of bacterial sequences was performed by augmenting the SILVA SSU database (version 119) with the sequences obtained in our clone libraries, and then trimming the taxonomic database to the gene region sequenced in our pyrosequencing data (Werner et al. 2012). Classification was then performed in QIIME using the default classification approach. Bacterial data were clustered into operational taxonomic units (OTUs) of 97 % sequence similarity for calculating diversity estimators. Eukaryotic sequences were clustered by taxonomic classification. Following the removal of Metazoa and Viridiplantae sequences, which were assumed to be host and food contaminants, samples were subsampled to the lowest coverage depth and the Shannon diversity estimator and evenness index were calculated. Raw flowgrams were uploaded to the NCBI Sequence Read Archive under accession numbers SAMN02382012–SAMN02382014.

Table 1 Summary of weta gut weight (wet w/w)

Sample	Weta #1	Weta #2	Weta #3
Whole insect	910	1437	828
Whole gut	228 (25.1)	436 (30.3)	283 (34.2)
Foregut	23 (2.5)	104 (7.2)	32 (3.9)
Midgut	85 (9.3)	121 (8.4)	92 (11.1)
Hindgut	120 (13.2)	211 (14.7)	159 (19.2)

Findings are reported both as the absolute gut weight (mg) and as a proportion (%) of the total insect weight (brackets). Slight discrepancies between the proportion of the whole gut and the summed total of the gut sections of Weta #1 occur due to rounding errors

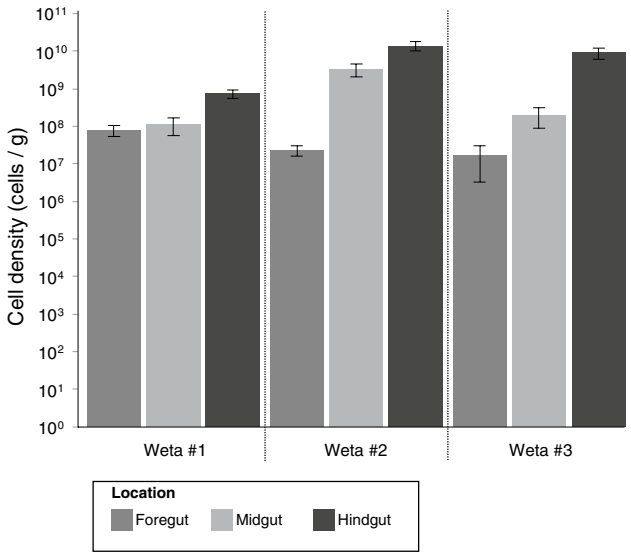


Fig. 2 Mean bacterial cell density along the length of the weta gastrointestinal tract. Samples reported based on FISH probing using the EUB338 mix, which targets essentially all known bacterial phyla. Cell counts are expressed per gram of gut (wet weight). *Error bars* represent 1 standard deviation

Results and discussion

Many aspects of weta biology are well studied, but until now, their microbiology has not been explored. Here, we investigated the gut microbiota of the Auckland tree weta using a variety of gene sequencing and microscopy-based approaches. BLAST analysis of the weta 18S rRNA genes identified the insect as being closely related to other sequences from the genus *Hemideina*, with high-confidence matches to previously published weta sequences obtained from *H. crassidens* and *H. maori* (100 and 98 % identity, respectively) (Pratt et al. 2008). The gut comprised approximately 30 % of the weta's body weight, with the majority of gut mass accounted for by the midgut and hindgut (Table 1). Bacterial cell density was highest in the hindgut, with the lowest bacterial densities in the foregut (Fig. 2). As bacterial morphology, as visualised by FISH, reveals little regarding the microbial diversity of a community, rRNA gene sequencing was utilised to more rigorously interrogate the microbial diversity of the weta gut.

Pyrosequencing of rRNA genes yielded a total of 102,591 reads, identified as 61,998 bacterial and 40,593 eukaryotic, with a median sequence length of 172 bp. Phylum-level taxonomic identification is reported in Table 2, and more detailed taxonomic classifications are provided in Table S1. The community of microorganisms present was consistent with those of most gut environments, with *Firmicutes* (59.7 % of all bacterial sequences) and *Bacteroidetes* (26.5 %) dominating among the bacteria. Members of the *Proteobacteria* were also prevalent (6.5 %), with a number of less abundant phyla including *Elusimicrobia* (originally described in termites, and often found in insect guts), *Verrucomicrobia*, and *Actinobacteria*. Figure 3 displays the phylum-level classification of bacteria within the weta gut, relative to the microbiota of cockroach and termite guts. A small proportion of sequences could not be classified at phylum level (0.7 %), although this number increased at finer taxonomic resolution with 4.7 % of sequences unable to be classified at the family level and 24.4 % at the genus level, using the classification method reported in methods (Table S2). Alternate classification approaches were employed,

Table 2 Relative abundance of the most abundant bacterial taxa in the 16S rRNA gene pyrosequencing data

Phylum	Class	Order	Family	Genus	Abundance (%)
Actinobacteria	*Coriobacteria*	*Coriobacteriales*	*Coriobacteriaceae*	Unclassified	1.8
Bacteroidetes[a]	*Bacteroidia*[a]	*Bacteroidales*[a]	*Bacteroidaceae*[a]	*Bacteroides*[a]	11.8
			Porphyromonadaceae	*Dysgonomonas*	3.2
				Tannerella	1.1
			Rikenellaceae[a]	*Alistipes*[a]	6.8
			3M1PL1-52 termite group	Unclassified	0.8
Deferribacteres[a]	*Deferribacteres*[a]	*Deferribacterales*[a]	*Deferribacteraceae*[a]	*Mucispirillum*[a]	0.8
Firmicutes[a]	*Bacilli*[a]	*Lactobacillales*[a]	*Lactobacillaceae*	*Lactobacillus*	2.8
			Leuconostocaceae[a]	*Weissella*[a]	15.4
	Clostridia[a]	*Clostridiales*[a]	*Christensenellaceae*[a]	Unclassified	1.8
			Clostridiaceae[a]	*Clostridium*[a]	8.0
			Defluviitaleaceae	Incertae Sedis	0.9
			Lachnospiraceae[a]	*Blautia*	0.8
				Incertae Sedis	4.4
				Unclassified	4.3
			Ruminococcaceae[a]	*Anaerofilum*[a]	4.5
				Anaerotruncus	0.8
				Intestinimonas[a]	2.0
				Incertae Sedis	1.2
				Unclassified	7.6
	Erysipelotrichia	*Erysipelotrichales*	*Erysipelotrichaceae*	Incertae Sedis	1.1
Proteobacteria[a]	*Deltaproteobacteria*[a]	*Desulfovibrionales*[a]	*Desulfovibrionaceae*[a]	*Desulfovibrio*[a]	3.4
		Rs-K70 termite group	Unclassified	Unclassified	0.9
Verrucomicrobia[a]	*Verrucomicrobiae*[a]	*Verrucomicrobiales*[a]	*Verrucomicrobiaceae*[a]	*Akkermansia*[a]	1.0

Values represent the average relative abundance (%) between weta individuals

[a] Lineages observed in the clone library data. Detailed description of inter-individual variation and complete phylotype listing is provided in Table S1

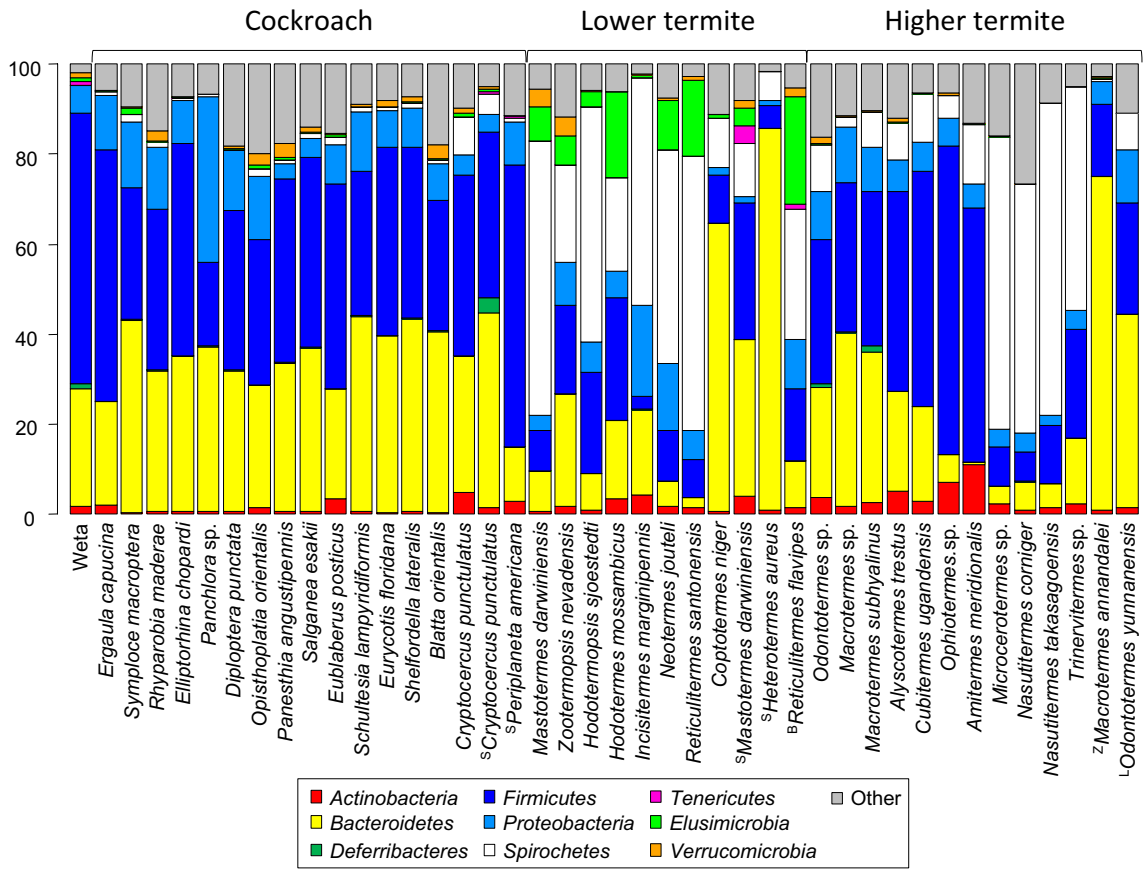
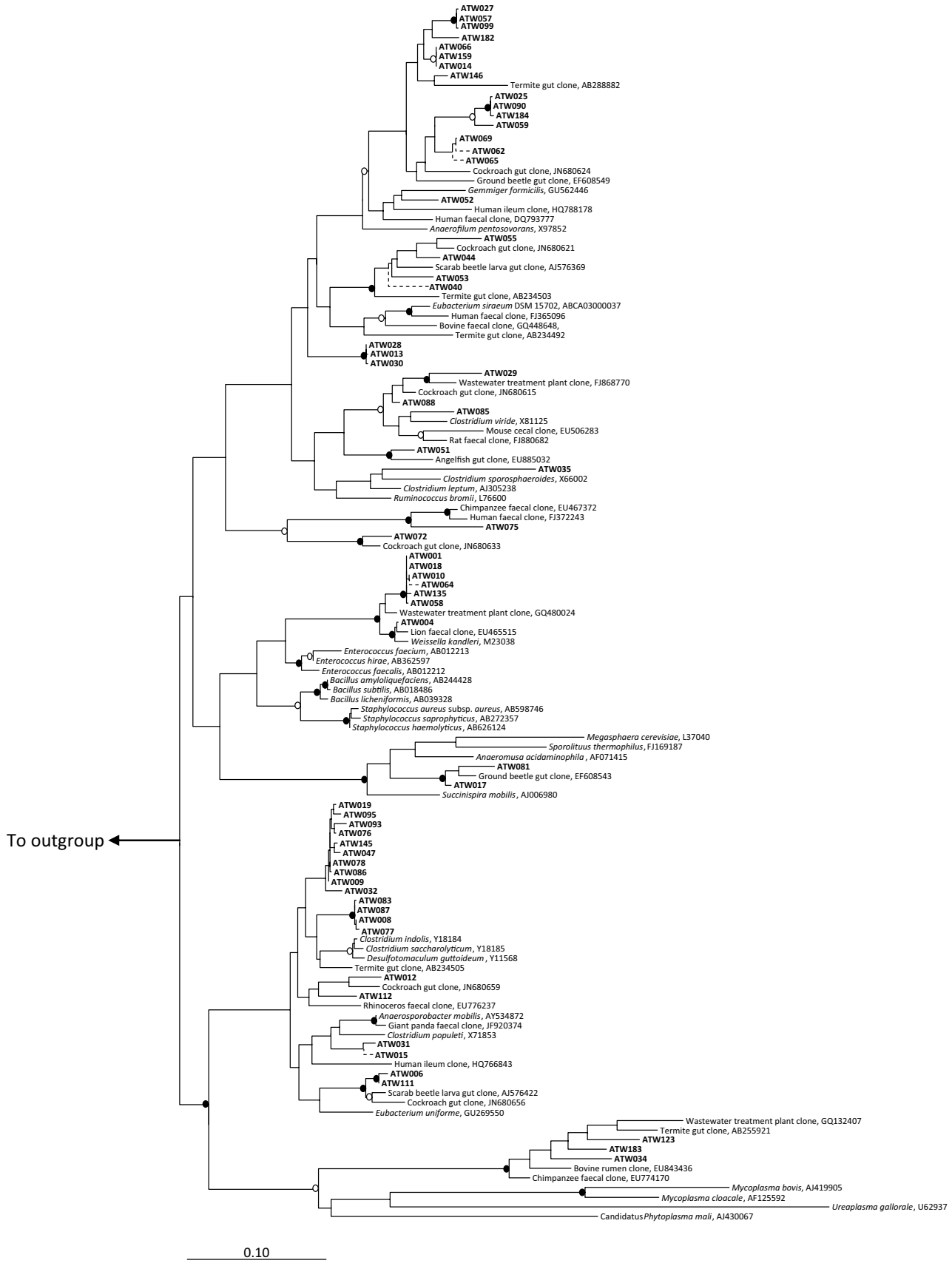


Fig. 3 Comparison of weta gut microbiota to that of cockroaches and termites. Phylum-level abundances of the dominant bacterial phyla in the microbial communities of weta, cockroach, and termite gut samples. Superscript notation identifies original data from published studies by Boucias et al. (2013) (B), Sabree and Moran (2014) (S), Liu et al. (2013) (L) and Zhang et al. (2014) (Z). Unmarked cockroach and termites samples were obtained from Dietrich et al. (2014)

performing naive Bayesian classification against the base SILVA SSU 119 database and the weta-augmented version of SILVA SSU 119, but these yielded a high proportion of unclassified sequences at phylum level (10.4 and 5.6 %, respectively). Those 16S rRNA sequences which could not be assigned to phylum level using the QIIME method were extracted from the main data set and analysed using the NCBI online BLAST tool, comparing these sequences to the nucleotide collection (nr/nt) database. All sequences were matched to bacterial clone sequences, primarily of *Firmicutes* and *Bacteroidetes* origins, although a low sequence similarity was observed between these matches and the reference database (~90 % sequence identity, data not shown).

The weta gut community contained a large proportion of 16S rRNA gene sequences belonging to the *Ruminococcaceae* (16.7 %) and *Lachnospiraceae* (10.4 %), bacterial families which are commonly associated with the guts of animals such as ruminants, cockroaches, and termites (Dietrich et al. 2014; Gosables et al. 2011; Kittelmann et al. 2013; Meehan and Beiko 2014; Sabree and Moran 2014; Thompson et al. 2012). Approximately half of the sequences associated with each of these families could not be classified to genus level using our approach (Table S1). These sequences were represented by 663 unique sequences, which were extracted from the data set and compared directly to the clone library data. The sequences generally exhibited high identity to the clone data in the 155 *Lachnospiraceae* (minimum 82.4 %, median 96.2 %, and maximum 100 %) and 508 *Ruminococcaceae* (minimum 80.5 %, median 94.0 %, and maximum 100 %) sequences were analysed. While it is likely that some of these low-identity sequences reflect sequencing error, we concluded that the majority of these sequence clusters were comprised of high-quality sequences that reflect the genuine weta microbiota. The bacterial diversity was consistent among individuals, with Shannon diversity indices approximately equal (WT#1 = 4.5, WT#2 = 4.0, WT#3 = 4.0), but representative of an uneven community profile in all individuals (Shannon evenness index WT#1 = 0.47, WT#2 = 0.41, WT#3 = 0.41).

Although methanogenic archaea perform well-described roles in the termite (Ohkuma 2003; Tokura et al. 2000) and cockroach hindguts (Gijzen and Barughare 1992; Gijzen


ATW027
ATW057
ATW099
ATW182
ATW066
ATW159
ATW014
ATW146
Termite gut clone, AB288882
ATW025
ATW090
ATW184
ATW059
ATW069
ATW062
ATW065
Cockroach gut clone, JN680624
Ground beetle gut clone, EF608549
Gemmiger formicilis, GU562446
ATW052
Human ileum clone, HQ788178
Human faecal clone, DQ793777
Anaerofilum pentosovorans, X97852
ATW055
Cockroach gut clone, JN680621
ATW044
Scarab beetle larva gut clone, AJ576369
ATW053
ATW040
Termite gut clone, AB234503
Eubacterium siraeum DSM 15702, ABCA03000037
Human faecal clone, FJ365096
Bovine faecal clone, GQ448648,
Termite gut clone, AB234492
ATW028
ATW013
ATW030
ATW029
Wastewater treatment plant clone, FJ868770
Cockroach gut clone, JN680615
ATW088
ATW085
Clostridium viride, X81125
Mouse cecal clone, EU506283
Rat faecal clone, FJ880682
ATW051
Angelfish gut clone, EU885032
ATW035
Clostridium sporosphaeroides, X66002
Clostridium leptum, AJ305238
Ruminococcus bromii, L76600
Chimpanzee faecal clone, EU467372
Human faecal clone, FJ372243
ATW075
ATW072
Cockroach gut clone, JN680633
ATW001
ATW018
ATW010
ATW064
ATW135
ATW058
Wastewater treatment plant clone, GQ480024
ATW004
Lion faecal clone, EU465515
Weissella kandleri, M23038
Enterococcus faecium, AB012213
Enterococcus hirae, AB362597
Enterococcus faecalis, AB012212
Bacillus amyloliquefaciens, AB244428
Bacillus subtilis, AB018486
Bacillus licheniformis, AB039328
Staphylococcus aureus subsp. aureus, AB598746
Staphylococcus saprophyticus, AB272357
Staphylococcus haemolyticus, AB626124
Megasphaera cerevisiae, L37040
Sporolituus thermophilus, FJ169187
Anaeromusa acidaminophila , AF071415
ATW081
Ground beetle gut clone, EF608543
ATW017
Succinispira mobilis, AJ006980
To outgroup
ATW019
ATW095
ATW093
ATW076
ATW145
ATW047
ATW078
ATW086
ATW009
ATW032
ATW083
ATW087
ATW008
ATW077
Clostridium indolis, Y18184
Clostridium saccharolyticum, Y18185
Desulfotomaculum guttoideum , Y11568
Termite gut clone, AB234505
ATW012
Cockroach gut clone, JN680659
ATW112
Rhinoceros faecal clone, EU776237
Anaerosporobacter mobilis, AY534872
Giant panda faecal clone, JF920374
Clostridium populeti, X71853
ATW031
ATW015
Human ileum clone, HQ766843
ATW006
ATW111
Scarab beetle larva gut clone, AJ576422
Cockroach gut clone, JN680656
Eubacterium uniforme, GU269550
Wastewater treatment plant clone, GQ132407
Termite gut clone, AB255921
ATW123
ATW183
ATW034
Bovine rumen clone, EU843436
Chimpanzee faecal clone, EU774170
Mycoplasma bovis, AJ419905
Mycoplasma cloacale, AF125592
Ureaplasma gallorale, U62937
Candidatus Phytoplasma mali, AJ430067
0.10

◀**Fig. 4** 16S rRNA-based phylogeny of *Firmicutes* sequences obtained from weta gut. Branch lengths were generated using maximum likelihood calculations on sequences with length >1200 bp (*solid lines*) using a 50 % conservation filter. Shorter reads were subsequently added using the Parsimony Interactive tool in ARB (*dashed lines*). Sequences from this study appear in bold type with the prefix 'ATW'. Bootstrap values were calculated using maximum parsimony with 5000 samplings. Solid junctions represent a branch with >90 % support, and hollow junctions >75 %. *Scale bar* 10 % divergence

et al. 1991), we did not observe the presence of methanogenic lineages, or indeed any archaea, in the weta. Archaeal 16S rRNA sequences were apparently absent from the pyrosequencing data set, and exhaustive efforts to amplify full-length archaeal 16S rRNA gene sequences from the gut homogenate using a previously reported method were also unsuccessful (Webster et al. 2004).

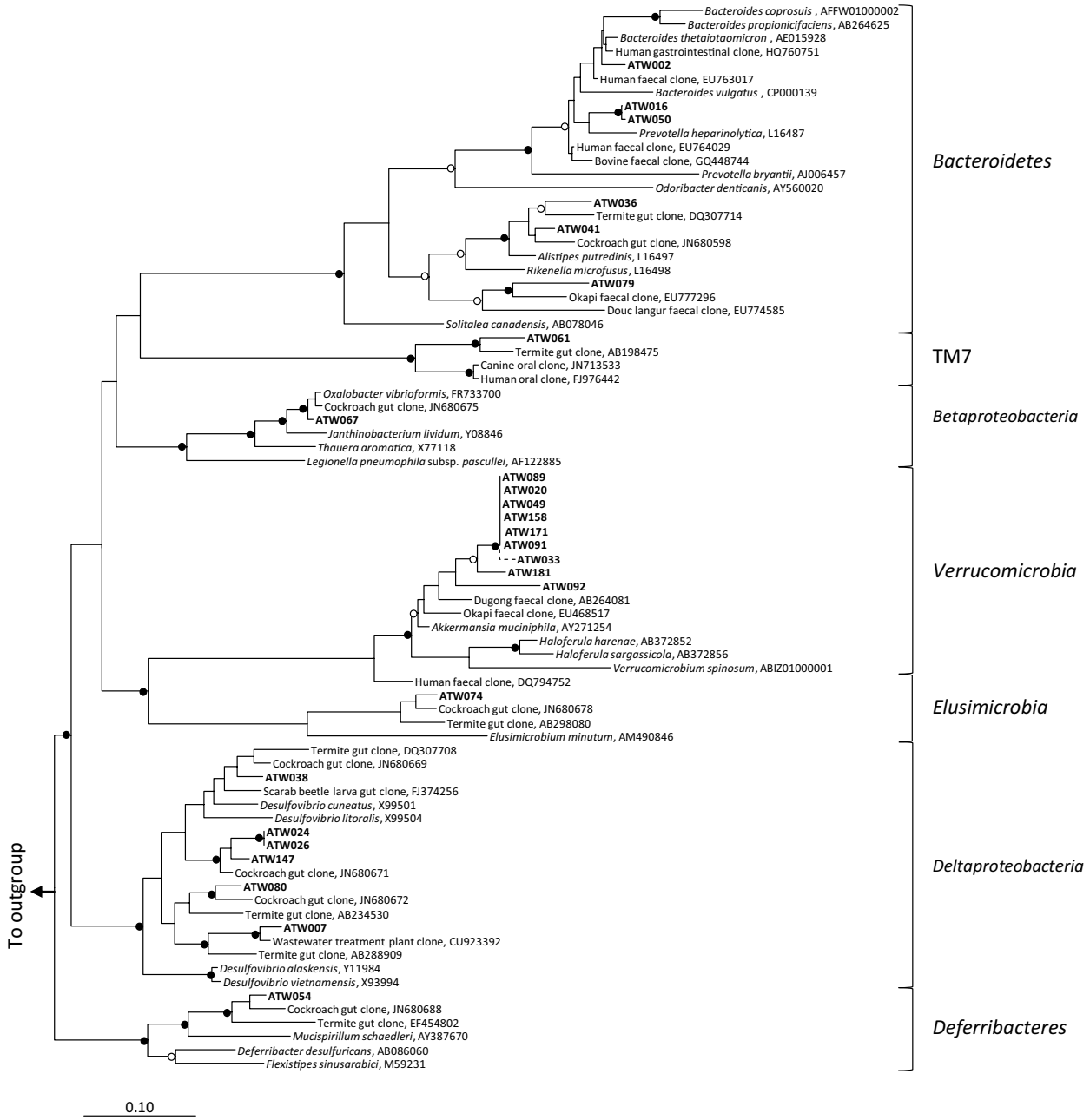


Fig. 5 16S rRNA-based phylogeny of non-*Firmicutes* bacteria sequences obtained from weta gut. Branch lengths were generated using maximum likelihood calculations on sequences with length >1200 bp (*solid lines*) using a 50 % conservation filter. Shorter reads were subsequently added using the Parsimony Interactive tool in ARB (*dashed lines*). Sequences from this study appear in bold type with the prefix 'ATW'. Bootstrap values were calculated using maximum parsimony with 5000 samplings. Solid junctions represent a branch with >90 % support, and hollow junctions >75 %. *Scale bar* 10 % divergence

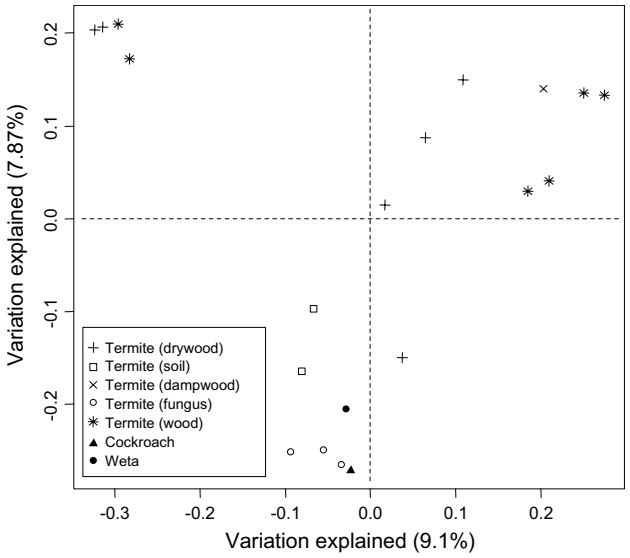


Fig. 6 Principle coordinate analysis of the bacterial communities obtained from termites, a cockroach, and the Auckland tree weta. Unweighted UniFrac distances were calculated from a neighbour-joining tree generated from pairwise distances between aligned sequences. Host species identities were as reported in the sample metadata

Classification of eukaryotic gene fragments revealed that most of the 18S rRNA gene sequences were affiliated with the Orthoptera (Table S2). Classifications of representative sequences from this group were identified as belonging to the genus *Hemideina* (99 % identity to *H. crassidens* and *H. maori*) through use of the NCBI online BLAST tool. These likely belong to the host itself and therefore were not useful in identifying the eukaryotic microbiota of the weta gut. In addition, approximately 2.6 % of eukaryotic sequence reads from weta were classified as plant material (Table S3), which was presumed to be food. Following the removal of these 18S rRNA gene sequences, the eukaryotic community was of low diversity (mean Shannon index = 0.49), but this may be the result of a large proportion of recovered 18S rRNA gene sequences belonging to the host insect, thus drastically lowering the sampling depth for the remaining eukaryotic microbes.

The phylogenetic relationship of nearly full-length 16S rRNA gene sequences is reported in Figs. 4 and 5, displaying members of the *Firmicutes* and the remainder of the bacterial microbiota, respectively. The sequences retrieved from the clone library were broadly congruent with those obtained by pyrosequencing, though the relative proportions of the various bacterial phyla did differ somewhat. Of the 87 long 16S rRNA gene sequences, 62 were affiliated with the *Firmicutes* (71.2 %), with *Bacteroidetes* comprising 6.9 %, *Verrucomicrobia* 10.3 %, and *Deltaproteobacteria* 6.9 %. Sequences belonging to the latter were related to the genus *Desulfovibrio*, members of which are capable of sulphate reduction. In the absence of measurements of sulphur concentrations and/or sulphate reduction rates, one can merely speculate as to a potential involvement of these bacteria in sulphur cycling within the weta, as has been implicated in the guts of other insects (Dröge et al. 2005; Sato et al. 2009). Weta-derived sequences appeared to cluster with clone sequences previously obtained from internationally collected cockroach and termite gut samples (order *Blattodea*), of which some species share a broadly similar diet to that of the tree weta. In order to test this relationship, unweighted UniFrac distances were calculated to test the phylogenetic membership of the bacterial community found in weta, termites, and cockroaches. Weta samples appeared to cluster with those from the cockroach *Shelfordella lateralis* and the soil- and fungus-feeding termites (Fig. 6). The single exception to this observation was that of the bacterial community associated with the drywood termite *Coptotermes formosanus*, although it is noted that this termite builds nests in the soil, which may influence the gut microbiota (Cabrera et al. 2005). When considering the distant phylogenetic relationship and broad geographic distribution of the insects sampled, we speculate that this clustering could reflect the influence of diet on the gut microbiota of these insects. Samples obtained from wood-feeding termites form clusters separate from samples of insects with different diet.

In summary, we have performed the first analysis of the gut microbiota of tree weta. We have shown that the gut of the Auckland tree weta harbours a diverse bacterial community of varying density along the gastrointestinal tract. In addition, we have shown that the weta gut microbiota is broadly similar to that of the cockroach and some termites, potentially suggesting a convergence of the gut microbiota.

Acknowledgments The work conducted by the U.S. Department of Energy Joint Genome Institute is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. We thank S. Tringe, S. Malfatti, and T. Glavina del Rio (JGI) for their helpful advice, M. Dhami for providing the 18S rRNA gene primers, and R. Kleinpaste and G. Holwell for supplying weta. DWW and MD were supported by University of Auckland Doctoral Scholarships.

References

Amann RI, Binder BJ, Olson RJ, Chisholm SW, Devereux R, Stahl DA (1990) Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. Appl Environ Microbiol 56:1919–1925

Boucias DG, Cai Y, Sun Y, Lietze V-U, Sen R, Raychoudhury R, Scharf ME (2013) The hindgut lumen prokaryotic microbiota of the termite *Reticulitermes flavipes* and its responses to dietary lignocellulose composition. Mol Ecol 22:1836–1853

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 70:292–300

Cabrera BJ, Su N-Y, Scheffrahn RH, Oi FM, Koehler PG (2005) Formosan Subterranean Termite. ENY-216. 2001. Florida Cooperative Extension Service, Institute of Food and Agricultural Sciences, University of Florida, Gainesville

Cary PRL (1983) Diet of the ground weta *Zealandosandrus gracilis* (Orthoptera: Stenopelmatidae). NZ J Zool 10:295–298

Daims H, Brühl A, Amann R, Schleifer KH, Wagner M (1999) The domain-specific probe EUB338 is insufficient for the detection of all Bacteria: development and evaluation of a more comprehensive probe set. Syst Appl Microbiol 22:434–444

Dietrich C, Köhler T, Brune A (2014) The cockroach origin of the termite gut microbiota: patterns in bacterial community structure reflect major evolutionary events. Appl Environ Microbiol 80:2261–2269

Diez B, Pedrós-Alió C, Marsh TL, Massana R (2001) Application of Denaturing Gradient Gel Electrophoresis (DGGE) to study the diversity of marine picoeukaryotic assemblages and comparison of DGGE with other molecular techniques. Appl Environ Microbiol 67:2942–2951

Dröge S et al (2005) In vitro and in vivo sulfate reduction in the gut contents of the termite *Mastotermes darwiniensis* and the rose-chafer *Pachnoda marginata*. J Gen Appl Microbiol 51:57–64

Fontanetti CS, Zefa E (2000) Morphological characterization of the proventriculus of *Gryllus assimilis* Fabricius (Orthoptera, Gryllidae). Revista Brasileira de Zoologia 17:193–198

Gaston KJ (1991) The magnitude of global insect species richness. Conserv Biol 6:293–296

Gibbs GW (2006) Ghosts of Gondwana : the history of life in New Zealand. Craig Potton Publishing, Nelson

Gijzen HJ, Barughare M (1992) Contribution of anaerobic protozoa and methanogens to hindgut metabolic activities of the American cockroach, *Periplaneta americana*. Appl Environ Microbiol 58:2565–2570

Gijzen HJ, Broers CAM, Barughare M, Stumm CK (1991) Methanogenic bacteria as endosymbionts of the ciliate *Nyctotherus ovalis* in the cockroach. Appl Environ Microbiol 57:1630–1634

Gosables MJ et al (2011) Metatranscriptomic approach to analyze the functional human gut microbiota. PLoS One 6:e17447

Griffin MJ, Morgan-Richards M, Trewick SA (2011a) Is the tree weta *Hemideina crassidens* an obligate herbivore? NZ Nat Sci 36:11–19

Griffin MJ, Trewick SA, Wehi PM, Morgan-Richards M (2011b) Exploring the concept of niche convergence in a land without rodents: the case of weta as small mammals. NZ J Ecol 35:302–307

Grünwald S, Pilhofer M, Höll W (2010) Microbial associations in gut systems of wood- and bark-inhabiting longhorned beetles (Coleoptera: Cerambycidae). Syst Appl Microbiol 33:25–34

Idowu AB, Edema MO, Oyedepo MT (2009) Extracellular enzyme production by microflora from the gut region of the variegated grasshopper *Zonocerus variegatus* (Orthoptera: Pyrgomorphidae). Int J Trop Insect Sci 29:229–235

Kearse M et al (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647–1649

Kelly CD (2011) Reproductive and physiological costs of repeated immune challenges in female Wellington tree weta (Orthoptera: Anostostomatidae). Biol J Linn Soc 104:38–46

Kittelmann S, Seedorf H, Walters WA, Clemente JC, Knight R, Gordon JI, Janssen PH (2013) Simultaneous amplicon sequencing to explore co-occurrence patterns of bacterial, archaeal and eukaryotic microorganisms in rumen microbial communities. PLoS One 8:e47879

Liu N et al (2013) Metagenomic insights into metabolic capacities of the gut microbiota in a fungus-cultivating termite (*Odontotermes yunnanensis*). PLoS One 8:e69184

Ludwig W et al (2004) ARB: a software environment for sequence data. Nucleic Acids Res 32:1363–1371

Maskell FG (1927) The anatomy of *Hemideina thoracica*. Trans NZ Inst 57:637–670

Matsui T, Tokuda G, Shinzato N (2009) Termites as functional gene resources. Recent Pat Biotechnol 3:10–18

Meehan CJ, Beiko RG (2014) A phylogenomic view of ecological specialization in the Lachnospiraceae, a family of digestive tract-associated bacteria. Genome Biol Evol 6:703–713

Neall VE, Trewick SA (2008) The age and origin of the Pacific islands: a geological overview. Philos Trans Royal Soc B 33:3293–3308

Novotny V, Basset Y, Miller SE, Weiblen GD, Bremer B, Cizek L, Drozd P (2002) Low host specificity of herbivorous insects in a tropical forest. Nature 416:841–844

Ohkuma M (2003) Termite symbiotic systems: efficient bio-recycling of lignocellulose. Appl Microbiol Biotechnol 61:1–9

Ohkuma M (2008) Symbioses of flagellates and prokaryotes in the gut of lower termites. Trends Microbiol 16:345–352

Oppert C, Klingerman WE, Willis JD, Oppert B, Jurat-Fuentes JL (2010) Prospecting for cellulolytic activity in insect digestive fluids. Comp Biochem Physiol B 155:145–154

Polz MF, Cavanaugh CM (1998) Bias in template-to-product ratios in multitemplate PCR. Appl Environ Microbiol 64:3724–3730

Pratt RC, Morgan-Richards M, Trewick SA (2008) Diversification of New Zealand weta (Orthoptera: Ensifera: Anostostomatidae) and their relationships in Australasia. Philos Trans Royal Soc Lond B 363:3427–3437

Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J, Glöckner FO (2007) SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nuclic Acids Res 35:7188–7196

R Core Team (2012) R: a language and environment for statistical computing. http://www.R-project.org/

Sabree ZL, Moran NA (2014) Host-specific assemblages typify gut microbial communities of related insect species. SpringerPlus 3:138

Santo Domingo JW, Kaufman MG, Klug MJ, Holben WE, Harris R, Tiedje JM (1998a) Influence of diet on the structure and function of the bacterial hindgut community of crickets. Mol Ecol 7:761–767

Santo Domingo JW, Kaufman MG, Klug MJ, Tiedje JM (1998b) Characterization of the cricket hindgut microbiota with fluorescently labeled rRNA-targeted oligonucleotide probes. Appl Environ Microbiol 64:752–755

Sato T, Hongoh Y, Noda S, Hattori S, Ui S, Ohkuma M (2009) *Candidatus* Desulfovibrio trichonymphae, a novel intracellular symbiont of the flagellate *Trichonympha agilis* in termite gut. Environ Microbiol 11:1007–1015

Schloss PD, Gevers D, Westcott SL (2011) Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies. PLoS One 6:e27310

Sinclair BJ, Wharton D (1997) Avoidance of Intracellular Freezing by the Freezing-Tolerant New Zealand Alpine Weta *Hemideina maori* (Orthoptera: Stenopelmatidae). J Insect Physiol 43:621–625

Spring S, Lins U, Amann R, Schleifer KH, Ferreira LCS, Esquivel DMS, Farina M (1998) Phylogenetic affiliation and ultrastructure of uncultured magnetic bacteria with unusually large magnetosomes. Arch Microbiol 169:136–147

Taylor MW, Schupp PJ, Dahllöf I, Kjelleberg S, Steinberg PD (2004) Host specificity in marine sponge-associated bacteria, and potential implications for marine microbial diversity. Environ Microbiol 6:121–130

Thompson CL, Vier R, Mikaelyan A, Wienemann T, Brune A (2012) 'Candidatus Arthromitus' revised: segmented filamentous bacteria in arthropod guts are members of *Lachnospiraceae*. Environ Microbiol 14:1454–1465

Tokura M, Ohkuma M, Kudo T (2000) Molecular phylogeny of methanogens associated with flagellated protists in the gut and with the gut epithelium of termites. FEMS Microbiol Ecol 33:233–240

Trewick SA, Morgan-Richards M (1995) On the distribution of tree weta in the North Island, New Zealand. J Royal Soc NZ 25:485–493

Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl Environ Microbiol 73:5261–5267

Webster NS, Negri AP, Munro MMHG, Battershill CN (2004) Diverse microbial communities inhabit Antarctic sponges. Environ Microbiol 6:288–300

Wehi PM, Hicks BJ (2010) Isotopic fractionation in a large herbivorous insect, the Auckland tree weta. J Insect Physiol 56:1877–1882

Wehi PM, Jorgensen M, Morgan-Richards M (2013) Sex- and season-dependent behaviour in a flightless insect, the Auckland tree weta (*Hemideina thoracica*). NZ J Ecol 37:75–83

Werner JJ et al (2012) Impact of training sets on classification of high-throughput bacterial 16 s rRNA gene surveys. ISME J 6:94–103

Wharton D (2011) Cold tolerance of New Zealand alpine insects. J Insect Physiol 57:1090–1095

White TJ, Bruns T, Lee S, Taylor J (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis M, Gelfand DH, Shinsky JJ, White TJ (eds) In PCR protocols: a guide to methods and applications. Academic Press, New York, pp 315–322

Willis JD, Klingerman WE, Oppert C, Oppert B, Jurat-Fuentes JL (2010) Characterization of cellulolytic activity from digestive fluids of *Dissosteira carolina* (Orthoptera: Acrididae). Comp Biochem Physiol B 157:267–272

Wilson GC, Jamieson IG (2005) Does melanism influence the diet of the mountain stone weta *Hemideina maori* (Orthoptera: Anostostomatidae)? NZ J Ecol 29:149–152

Woyke T, Smith D (2008) 16S/18S rRNA PCR Library Creation http://my.jgi.doe.gov/general/protocols/SOP_16S18S_rRNA_PCR_Library_Creation.pdf. Accessed 26 Nov 2013

Zhang M et al (2014) Phylogenetic and functional analysis of gut microbiota of a fungus-growing higher termite: *Bacteroidetes* from higher termites are a rich source of β-glucosidase genes. Microb Ecol 68:416–425