MOLECULAR ECOLOGY

Molecular Ecology (2015) **24**, 5284–5295 doi: 10.1111/mec.13376

Diet is the primary determinant of bacterial community structure in the guts of higher termites

ARAM MIKAELYAN,*† CARSTEN DIETRICH,* TIM KÖHLER,* MICHAEL POULSEN,‡ DAVID SILLAM-DUSSÈS§¶ and ANDREAS BRUNE*†

Department of Biogeochemistry, Max Planck Institute for Terrestrial Microbiology, Karl-von-Frisch-Str. 10, 35043, Marburg, Germany, †LOEWE Center for Synthetic Microbiology, SYNMIKRO, Philipps-Universität Marburg, Marburg, Germany, ‡Section for Ecology and Evolution, Department of Biology, Centre for Social Evolution, University of Copenhagen, Copenhagen East, Denmark, §Laboratory of Experimental and Comparative Ethology (LEEC), University of Paris 13, Sorbonne Paris Cité, Villetaneuse, France, ¶Institute of Ecology and Environmental Sciences – Paris (iEES-Paris), Institute of Research for Development, Sorbonne Universités, Bondy, France

Abstract

The gut microbiota of termites plays critical roles in the symbiotic digestion of lignocellulose. While phylogenetically 'lower termites' are characterized by a unique association with cellulolytic flagellates, higher termites (family Termitidae) harbour exclusively prokaryotic communities in their dilated hindguts. Unlike the more primitive termite families, which primarily feed on wood, they have adapted to a variety of lignocellulosic food sources in different stages of humification, ranging from sound wood to soil organic matter. In this study, we comparatively analysed representatives of different taxonomic lineages and feeding groups of higher termites to identify the major drivers of bacterial community structure in the termite gut, using amplicon libraries of 16S rRNA genes from 18 species of higher termites. In all analyses, the wood-feeding species were clearly separated from humus and soil feeders, irrespective of their taxonomic affiliation, offering compelling evidence that diet is the primary determinant of bacterial community structure. Within each diet group, however, gut communities of termites from the same subfamily were more similar than those of distantly related species. A highly resolved classification using a curated reference database revealed only few genus-level taxa whose distribution patterns indicated specificity for certain host lineages, limiting any possible cospeciation between the gut microbiota and host to short evolutionary timescales. Rather, the observed patterns in the host-specific distribution of the bacterial lineages in termite guts are best explained by diet-related differences in the availability of microhabitats and functional niches.

Keywords: gut microbiota, insects, pyrosequencing, termites

Received 25 June 2015; revision received 20 August 2015; accepted 1 September 2015

Introduction

Termites are eusocial cockroaches that have specialized on a diet of lignocellulose in various stages of humification (Brune 2014). The evolutionary transition from an omnivorous to wood-feeding lifestyle, which occurred more than 150 million years ago, was accompanied by several digestive modifications, including a distinctive enlargement of the hindgut and the acquisition of cellulolytic flagellates (Brune & Dietrich 2015). The flagellates play a critical role in symbiotic digestion in the hindgut of all evolutionary 'lower' termite families, but were lost in the most derived family, the Termitidae or 'higher' termites, which appeared about 50 million years ago and whose gut microbiota is entirely prokaryotic (Brune & Ohkuma 2011; Bourguignon *et al.* 2015).

Representing over 85% of all termite genera, the higher termites are the most diverse of all termite families (Krishna *et al.* 2013). While higher termites have

Correspondence: Andreas Brune, Fax: +49 6421 178999; E-mail: brune@mpi-marburg.mpg.de

diversified in many aspects, the most important specialization concerns their diet, which extends far beyond the wood-feeding lifestyle of 'lower termites' (Eggleton & Tayasu 2001). Some higher termites feed on sound lignocellulose such as wood or dry grass, others on partially degraded material such as leaf litter, herbivore dung or humus, or in the case of the true soil feeders, on soil organic matter in an advanced stage of humification (Donovan *et al.* 2001; Eggleton & Tayasu 2001). A special case are the fungus-cultivating termites, which grow basidiomycete fungi (*Termitomyces* spp.) to predigest their lignocellulosic diet and, in addition, also consume the mycelia of the fungal garden (Poulsen 2015). While fungus feeders consist exclusively of Macrotermitinae, other feeding groups usually comprise representatives from various subfamilies (Donovan *et al.* 2001).

Higher termites offer the unique opportunity to study the co-evolution of gut microbiota and host. The tremendous diversity in their dietary specializations and corresponding digestive adaptations in gut anatomy and physiology are reflected in the composition of their bacterial communities (Brune & Ohkuma 2011; Brune 2014). Recent studies using deep-sequencing of 16S rRNA genes to examine the gut microbiota across a wide range of higher termites have detected obvious patterns in community structure, which were interpreted to reflect both phylogeny and diet of the respective host groups (Dietrich *et al.* 2014; Otani *et al.* 2014; Rahman *et al.* 2015). Nevertheless, limited taxon sampling within the higher termites has so far prevented an effective identification of the major drivers of community structure of their gut microbiota.

Therefore, to investigate potential drivers of community structure in higher termites, it is essential that the analysis include a critical number of representatives from the same feeding groups that belong to evolutionarily independent host lineages. Building on our previous analysis of bacterial community structure in a wide range of termites and cockroaches, which already included 10 species of higher termites (Dietrich *et al.* 2014), we investigated eight additional species from previously undersampled feeding groups. We sequenced the amplified 16S rRNA (V3–V4 region) genes using the Illumina MiSeq platform and analysed the communities across all samples including both similarity-based metrics and the identification of individual lineages using a highly resolved taxonomic framework (DictDb; Mikaelyan *et al.* 2015).

Materials and methods

Termites: collection, identification and dissection

Atlantitermes sp., *Cornitermes* sp., *Neocapritermes taracua*, *Termes fatalis* and *Velocitermes* sp. were collected near Petit-Saut dam in French Guiana (5°4′N, 52°59′W), *Termes hospes* and *Microcerotermes parvus* were collected near Pointe-Noire in the Democratic Republic of the Congo (4°41′S, 11°51′E), and *Promirotermes* sp. was collected in ARC-PPRI Rietondale Research Station, Pretoria in South Africa, in the year 2013. The origin of the remaining species has been previously described (Dietrich *et al.* 2014). Species identity was confirmed by sequence analysis of the gene encoding cytochrome oxidase subunit 2 (COII; for accession numbers, see Table 1). A maximum-likelihood tree was calculated for the COII gene sequences using FASTTREE (Price *et al.* 2009) under the general time-reversible (GTR) model of evolution to ascertain the phylogenetic position of the sequences with respect to each other and to close relatives in GenBank (Fig. S1, Supporting information). All termites (worker caste only) were dissected with fine-tipped forceps within a few days of arrival in the laboratory in Marburg (for details, see Table 1). Hindguts (10–20 per sample) were pooled in 2-mL tubes containing 750 μL sodium phosphate buffer (120 mM; pH 8.0) and homogenized. DNA was extracted and purified using a bead-beating protocol as previously described (Paul *et al.* 2012).

Library preparation and sequencing

The V3–V4 region of the 16S rRNA gene was amplified from each sample using the primers M13-343Fmod and M13-784Rmod, which were based on the universal bacterial primers 343Fmod and 784Rmod (Köhler *et al.* 2012) respectively, but additionally included universal M13-specific priming sites on their 5′ ends (Daigle *et al.* 2011). The cycling conditions for this PCR step were as described previously (Köhler *et al.* 2012). The resulting amplicons, tagged with the M13 tails, were used as template for a subsequent PCR step using the Herculase II Fusion DNA Polymerase Kit (Agilent Technologies, USA). The following PCR conditions were used: 94 °C for 3 min, followed by 28 cycles (94 °C for 20 s, 58 °C for 20 s and 72 °C for 50 s), and finally 72 °C for 2 min. Purified PCR products were mixed in equimolar amounts and sequenced commercially (paired-end; 2 × 350 nt; Illumina MiSeq; GATC Biotech, Konstanz, Germany). The quality of the final products was checked by gel electrophoresis. The quality-checked MiSeq data sets were submitted to the MG-RAST server (Meyer *et al.* 2008) and can be accessed under the numbers 4639074.3–4639081.3.

Processing of sequence data

Both iTag (Degnan & Ochman 2012) libraries (this study) and pyrotag libraries (Dietrich *et al.* 2014; GenBank accession nos SAMN02228091–101) were

Table 1 Higher termites used in the current study, their phylogenetic affiliations and dietary preferences

ID	Host species	COII gene	Feeding group*	Diet†
Macrotermitinae				
1	*Macrotermes* sp.	KT184483	F	Lignocellulose, fungus†
2	*Macrotermes subhyalinus*	KT184482	F	Lignocellulose, fungus†
3	*Odontotermes* sp.	KT184484	F	Lignocellulose, fungus†
Apicotermitinae				
4	*Alyscotermes trestus*	KT184481	S	Soil organic matter†
Syntermitinae				
5	*Cornitermes* sp.	AIZ68247‡	L	Litter§
Termitinae				
6	*Microcerotermes parvus*	AIZ68273‡	W	Wood†
7	*Microcerotermes* sp.	KT184480	W	Wood†
8	*Neocapritermes taracua*	AIZ68299‡	H	Humus†
9	*Cubitermes ugandensis*	AIZ68260‡	S	Soil organic matter†
10	*Ophiotermes* sp.	KT184477	S	Soil organic matter†
11	*Termes hospes*	AIZ68312‡	H	Humus†
12	*Termes fatalis*	KT184478	H	Humus†
13	*Promirotermes* sp.	KT184479	H	Humus†
Nasutitermitinae				
14	*Atlantitermes* sp.	KT184476	H	Humus†
15	*Velocitermes* sp.	KT184475	L	Litter†
16	*Trinervitermes* sp.	KT184474	W	Dry grass†
17	*Nasutitermes corniger*	AIZ68286‡	W	Wood†
18	*Nasutitermes takasagoensis*	KT184473	W	Wood†

*Feeding groups: F, fungus feeders; W, wood or grass feeders; L, litter feeders; H, humus feeders; S, 'true' soil feeders.
†Based on dietary information in Bignell *et al.* (2011).
‡COII genes annotated in the mitochondrial genomes reconstructed by Dietrich & Brune (2014).
§Based on dietary information in Gontijo & Domingos (1991).

processed using the *UPARSE* pipeline (Edgar 2013). Only reads with a minimum length of 250 bp and a maximum expected error of 0.5 were selected and separated by sample into different fastq files using the sample-specific barcodes included in the sequences. For the iTag libraries, the *UPARSE* pipeline was used to merge paired reads; pairs with mismatches in the overlapping region were discarded. After removal of barcodes and primers, the reads from each sample were clustered at a threshold of 99% sequence similarity to form operational taxonomic units (OTUs) using *UPARSE*. A representative sequence from each OTU was selected and aligned with the MOTHUR aligner, using the *Silva* reference alignment (SSURef release 119) as a template.

Analysis of community structure

For the analysis of taxonomic composition of each sample, the OTUs in the de-replicated data sets were classified using the *RDP* classifier (Wang *et al.* 2007) implemented in the MOTHUR software suite (Schloss *et al.* 2009), using a confidence cut-off of 80% and the Dictyoptera taxonomic reference database (DICTDB) v. 3.0 (Mikaelyan *et al.* 2015). Genus-level lineages were ranked by determining their cumulative contribution to the principal component analysis (PCA) of the gut communities as previously described (Dietrich *et al.* 2014; Otani *et al.* 2014). Bray–Curtis distances between all communities at the genus level were calculated using the VEGAN package (Oksanen *et al.* 2015) in the R statistical software suite (R Core Team 2015). Data sets were subsampled to a thousand sequences, and distances were visualized using principal coordinate analysis (PCoA) implemented in *vegan*. The Bray–Curtis distances were additionally subjected to hierarchical cluster analysis and visualized as a dendrogram with the PVCLUST package (Suzuki & Shimodaira 2006) of R.

Similarities in the phylogenetic structure of the communities were determined using the taxonomy-independent weighted UniFrac metric (Lozupone *et al.* 2007) implemented in MOTHUR. A maximum-likelihood tree was constructed using FASTTREE (Price *et al.* 2009) with subsamples of 1100 sequences per sample (corresponding to the number of sequences in the smallest library). This tree served as input for the calculation of pairwise distances between all 18 samples with UniFrac, which were ordinated using PCoA.

The significance of clustering at the community level was tested for both metrics using the *adonis* function implemented in the VEGAN package (Oksanen *et al.* 2015), including taxonomy, diet and geography as possible variables. Additionally, the Kruskal–Wallis nonparametric test implemented in R was used to assess whether differences in the relative abundance of individual bacterial genera can be explained by differences in diet or host taxonomy.

Phylogenetic analysis of short reads

Representatives of each OTU from each sample were subjected to phylogenetic analysis using FASTTREE (Price *et al.* 2009) under maximum-likelihood criteria and the GTR model. The relative abundance of each OTU in the tree was annotated as circles using the APE package (Paradis *et al.* 2004) written for the R software suite (R Core Team 2015).

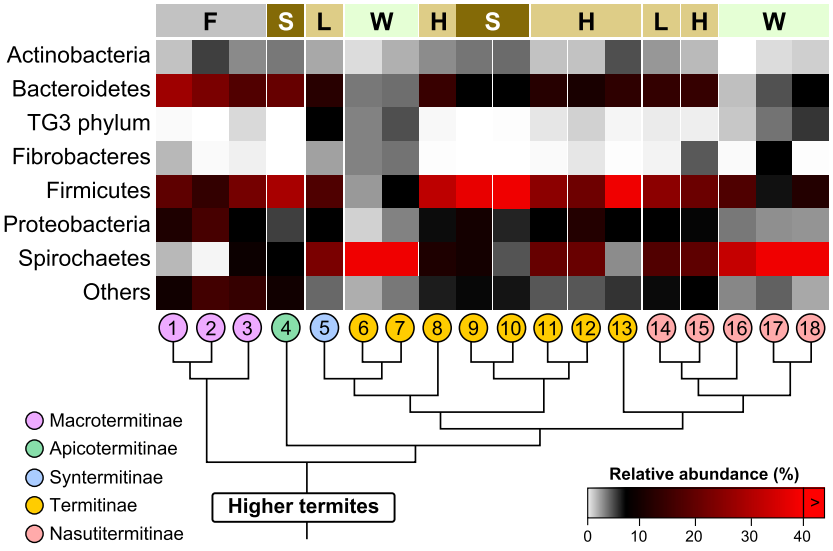

Fig. 1 Abundance of dominant bacterial phyla in the gut microbiota of higher termites. Numbers below the heatmap correspond to hosts in Table 1. The cladogram represents the phylogenetic relationship among the hosts and is based on the latest multilocus analysis of concatenated mitochondrial genes (Bourguignon *et al.* 2015). The subfamily assignment is based on the taxonomy proposed by Inward *et al.* (2007). The letters above the heatmap indicate the feeding groups: F, fungus-feeding; S, soil-feeding; W, wood or grass; H, humus; L, litter (see Table 1 for more information on the hosts' dietary preferences).

Results

Quality processing of the amplicon libraries yielded 1097–21 828 sequence reads for each pyrotag and 57 028–222 116 sequence reads for each iTag library. Classification with the taxonomic framework of DictDb successfully assigned 99% of the reads at the phylum level. Classification success decreased with taxonomic depth (48–90% at the genus level; for details, see Table S1, Supporting information). The lowest classification success was observed with some humus-feeding taxa, which is indicative of the bacterial diversity that still remains to be explored in this diet group, and the need to add more full-length 16S rRNA reference sequences to DictDb. For a detailed overview of the classification results, see Tables S2 and S3 (Supporting information).

At the phylum level, termites within a subfamily differed considerably in community structure, with the distribution of dominant phyla being strongly reflective of the termite feeding group. Irrespective of their phylogenetic affiliation, termites belonging to the same feeding group showed striking similarities in the distribution of dominant bacterial phyla (Fig. 1). Members of Spirochaetes, Fibrobacteres and/or the TG3 phylum were more prevalent among termites feeding on sound wood or grass than among termites of other feeding groups, although the proportion of each phylum differed among host species. Large proportions of Spirochaetes, Fibrobacteres and TG3 phylum were also found among litter-feeding termites. However, with the prevalence of Firmicutes and Bacteroidetes encountered in these taxa, the overall pattern among the litter feeders resembled that of the humus feeders. Humus feeders, soil feeders and fungus feeders also shared similarities in community structure, particularly in the large proportions of Firmicutes, Bacteroidetes and Proteobacteria, but differed in the abundance of Spirochaetes, which was lower in soil feeders and almost absent from fungus feeders (Fig. 1). Soil feeders also harboured a larger proportion of Actinobacteria than most other species. Fibrobacteres and TG3 phylum, which were characteristic of wood-feeding and litter-feeding termites, were virtually absent from most representatives of the other feeding groups.

The diet-specific pattern in gut community structure observed at the phylum level was even more evident at higher taxonomic or phylogenetic resolution. PCoA of community structure showed a clustering of samples by host diet with both Bray–Curtis (at the genus level) and UniFrac metrics, with the first two axes explaining approximately half of the total variability in the data sets (Fig. 2). In addition, hierarchical cluster analysis revealed a signal of host taxonomy within each diet group; that is, phylogenetically related termites are more similar in community structure than termites from different subfamilies (Fig. 3). Both Bray–Curtis and UniFrac dissimilarities in community structure were explained by host diet and taxonomy (considering a significance threshold at $P \leq 0.05$), and not by geographical location (see Table S5, Supporting information for details). Typically, differences in the relative abundance of individual bacterial genera were better explained by differences in diet in comparison with host taxonomy (Tables S6 and S7, Supporting information).

To identify the genus-level taxa that contribute most to the observed clustering pattern shown in Fig. 2, we conducted a PCA and ranked the genus-level taxa in descending order of their cumulative contribution to the loading factors. The relative abundances of the 20 taxa contributing the most to the clustering observed in PCA illustrate a preferential enrichment of different

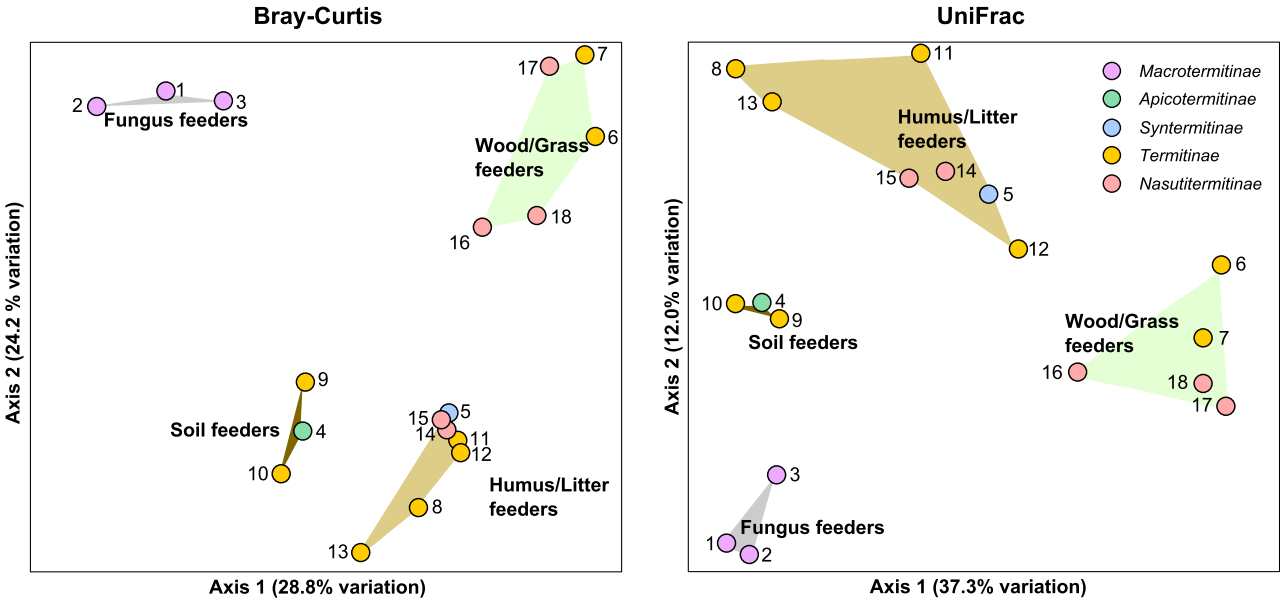

Fig. 2 Principal coordinate analysis (PCoA) of pairwise distances among bacterial communities based on the Bray–Curtis and UniFrac metrics. Each data point in the plots represents the community harboured by a given termite species; colours indicate the different host subfamilies. For more details on the samples, see Table 1.

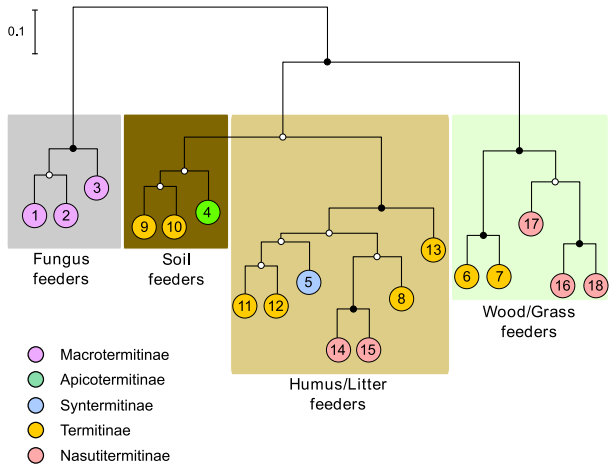

Fig. 3 Hierarchical cluster analysis of pairwise distances between bacterial communities based on the Bray–Curtis metric. Nodes in the dendrogram that are strongly supported by high approximately uniform probability values are marked by open (>70%) and closed (>90%) circles.

bacterial lineages among certain termite feeding groups (Fig. 4). With only a few exceptions, the reads classified as Spirochaetes (assigned to the genus-level lineages *Treponema* Ia, Ic or If) decreased in abundance with increasingly humified host diets. *Treponema* Ia was most abundant in wood/grass feeders, and more abundant in soil feeders than in humus/litter feeders (Fig. 4; Table S2, Supporting information). The relative abundance of *Treponema* Ic rose from 0.1% to 1.3% in the fungus-feeding termites, through 0.25–2.1% in the soil feeders, to as high as 16.0% in humus/litter feeders. The highest abundance of *Treponema* Ic, however, was observed among termites feeding on sound lignocellulose, where the reads accounted for 25.1–48.7% of the communities; an exception is the grass-feeding *Trinervitermes* sp. (4.2% of the community). Also *Treponema* If was significantly more abundant among the wood/grass feeders than the other diet groups, with the exception of the wood-feeding *Nasutitermes takasagoensis*.

Our assessment of the gut communities at greater taxonomic depth revealed patterns of preferential association of genus-level groups in Fibrobacteres and TG3 phylum with particular host lineages (Fig. 4). Subcluster Ia of Fibrobacteres and Subcluster IIIb of TG3 phylum were preferentially associated with *Microcerotermes* spp. and *Termes* spp., both members of Termitinae. Contrastingly, Subcluster Ib (of Fibrobacteres) and Subcluster IIIa (of TG3 phylum) were found associated with *Nasutitermes corniger* and *Trinervitermes* sp., both members of Nasutitermitinae. Reads from *Cornitermes* sp., a member of Syntermitinae, were classified to Subcluster IIIb, while those from *Macrotermes* sp. that were assigned to Fibrobacteres could be binned to genus-level clusters within 'Cockroach cluster II'.

One of the most predominant genus-level taxa among fungus-feeding termites was '*Alistipes* II' (family: Rikenellaceae), which represented 6–19% of the reads obtained from the three Macrotermitinae but was absent from most other host species (Fig. 4; Table S2,

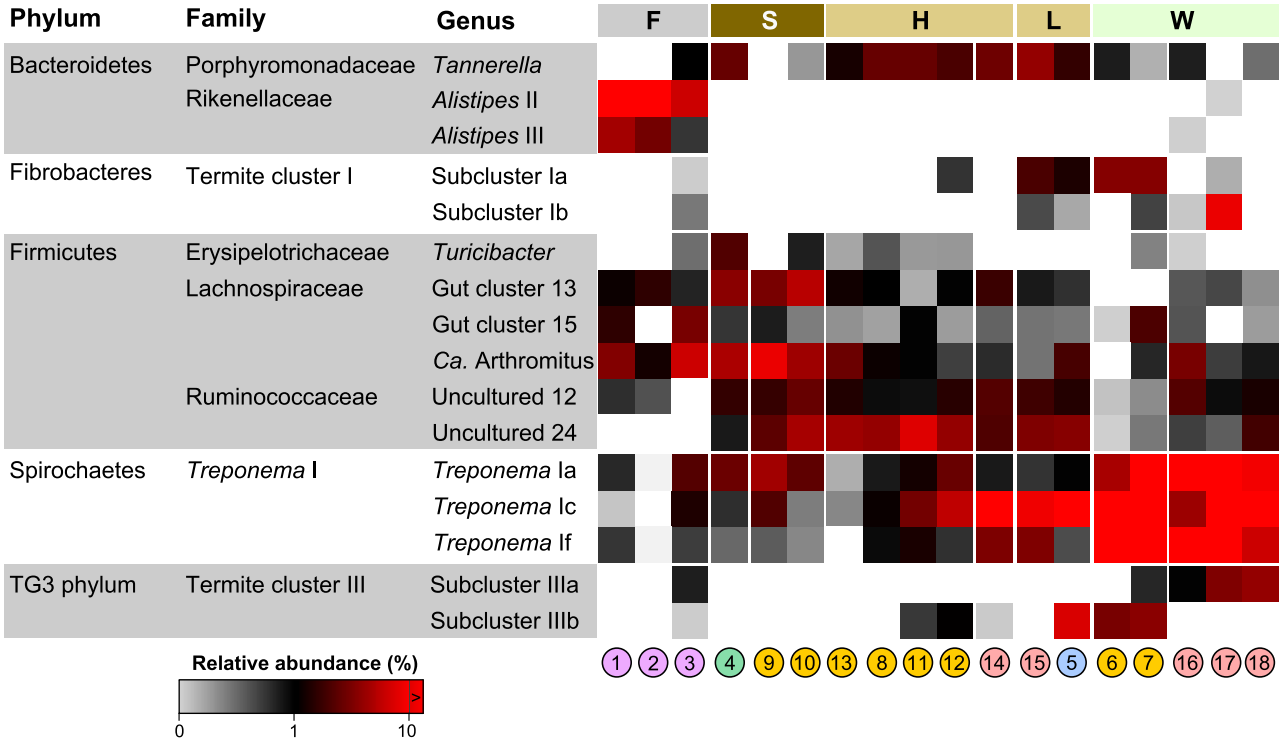

Fig. 4 Abundance of selected bacterial genus-level groups in higher termites. Numbers below the heatmap correspond to hosts in Table 1; colours of circles represent the different subfamilies of higher termites (see Fig. 1 legend for colour key). The letters above the heatmap indicate the different feeding groups: F, fungus-feeding; S, soil-feeding; W, wood or grass; H, humus; L, litter. See Tables S2 and S3 (Supporting information) for a more detailed taxonomic overview of the communities.

Supporting information). In addition, the fungus feeders were characterized by the abundant presence of several genus-level groups within Lachnospiraceae and Ruminococcaceae. The former include the genus-level 'Gut cluster 13' and '*Ca.* Arthromitus', which were represented also in other feeding groups and most abundant among the soil feeders (Fig. 4).

An analysis of the representative phylotypes from selected genus-level lineages in the *Treponema* I clade revealed that the highly divergent phylotypes clustered according to the subfamilies of their respective host species (e.g. *Treponema* Ia, Ic; Fig. 5A). A tendency of the phylotypes to cluster by host subfamily was observed also in the Fibrobacteres (Cluster I) and the related TG3 phylum (Subcluster III) (Fig. 5B). However, such host-specific clustering was observed mostly with the dominant phylotypes, whereas many of the rarer phylotypes clustered with sequences from other subfamilies (Fig. S2, Supporting information).

Discussion

Previous studies have attempted to assess in how far host phylogeny or dietary specialization of termites determines the composition of their gut microbiota (Colman *et al.* 2012; Dietrich *et al.* 2014; Rahman *et al.* 2015). However, because dietary diversification of higher termites is part of their evolutionary history, the influence of diet on gut community structure cannot be easily disentangled from that of host phylogeny (Brune & Dietrich 2015). Dietary diversification apparently involved concerted adaptations in mandible morphology (Donovan & Jones 2000), intestinal anatomy (Noirot 2001) and physicochemical gut conditions (Bignell & Eggleton 1995; Brune *et al.* 1995; Brune & Kühl 1996). At the same time, there is considerable phylogenetic conservatism in the dietary habits of closely related species, for example wood-feeding within the genus *Microcerotermes*, and soil-feeding within the '*Cubitermes* group' (Bourguignon *et al.* 2015).

Termite guts are small but complex ecosystems with a wide range of physicochemically distinct microhabitats, the distribution of which strongly reflects adaptations to different diets (Brune & Dietrich 2015). In this section, we discuss how differences in the availability of these microhabitats and the dynamics of ecological niches for the gut microbiota in the course of dietary diversification can explain the differences in the distribution and abundance of many bacterial lineages and provide tenable ecological and evolutionary explanations

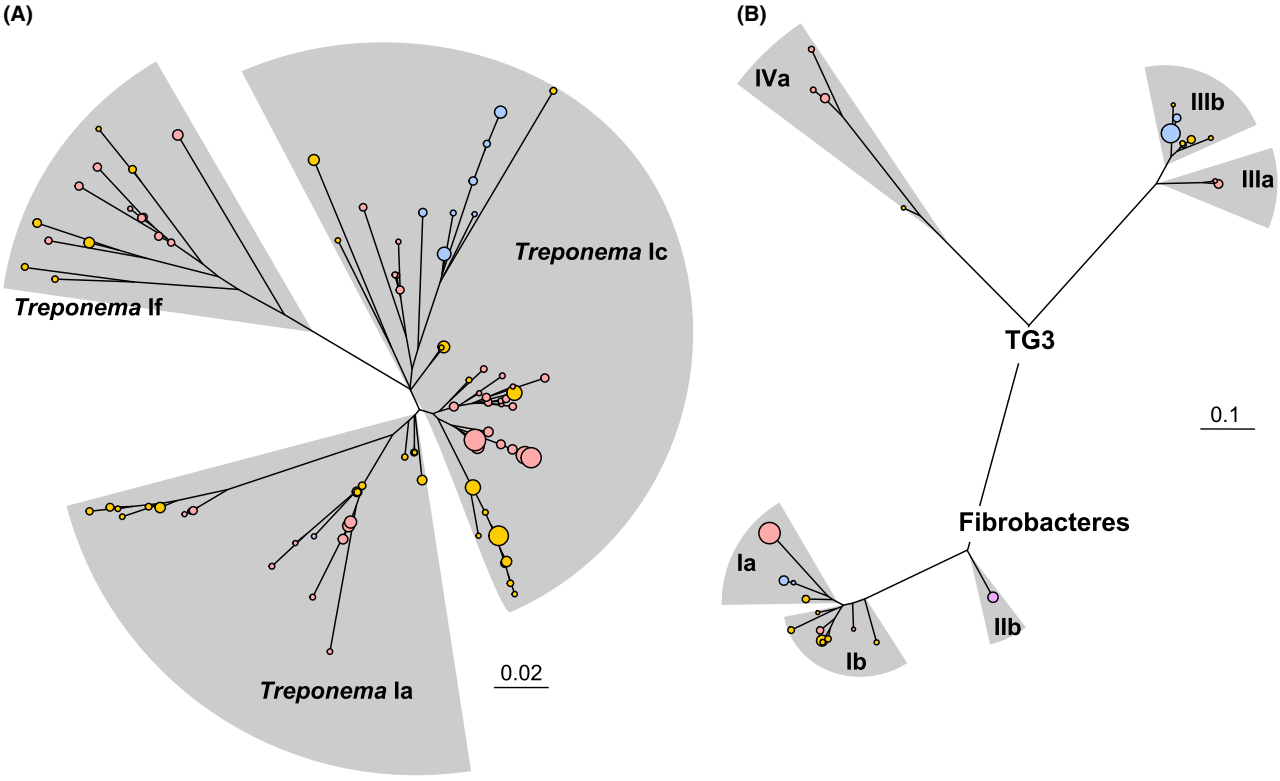

Fig. 5 Maximum-likelihood tree of 16S rRNA phylotypes from the different subfamilies of higher termites (see Fig. 1 legend for colour key and Fig. S2, Supporting information for tip labels). (A) *Treponema* Ia–c genus-level groups. (B) Phylum Fibrobacteres and TG3. Size of the circles at the tips of the tree is proportional to the relative abundance of the phylotypes in the respective hosts. Colours of the circles correspond to different subfamilies of higher termites.

that reconcile the artificial dichotomy in the effects of diet and host phylogeny on gut community structure in higher termites.

Bacterial communities in wood feeders

The remarkable convergence of bacterial community structure among wood-feeding Nasutitermitinae and Termitinae serves to illustrate how diet can engineer major changes in the gut microbiota through a redistribution of niches and microhabitats in the gut. In the gut of wood-feeding 'lower termites', microbial niches connected to lignocellulose digestion are mostly occupied by gut flagellates, which sequester wood particles in their digestive vacuoles (Brugerolle & Radek 2006; Ni & Tokuda 2013). In higher termites, the loss of flagellates opened up these niches to fibre-digesting bacteria (Brune & Dietrich 2015). This hypothesis is sparked by the upshift in abundance in wood-feeding higher termites of bacterial lineages from the Fibrobacteres and the TG3 phylum, which are extremely rare in the gut microbiota of 'lower termites' and in higher termites of other feeding groups (Hongoh *et al.* 2006; Dietrich *et al.* 2014; Rahman *et al.* 2015).

The deep taxon sampling of higher termites and the highly resolved classification of the present study revealed that bacterial lineages from Termite cluster I (phylum: Fibrobacteres), Termite cluster III (TG3 phylum) and the *Treponema* clusters Ic and If (phylum: Spirochaetes) are more abundant in wood-feeding representatives than in other feeding groups. Members of these bacterial lineages dominate the cellulolytic community that is specifically associated with the wood particles in the hindgut of *Nasutitermes* spp. (Mikaelyan *et al.* 2014), and are also abundantly represented in the guts of other wood-feeding termites (Hongoh *et al.* 2006; Warnecke *et al.* 2007; He *et al.* 2013). The high abundance of these cellulose-digesting, fibre-associated bacterial lineages among wood and litter feeders identifies the important structuring role played by both niche (cellulose degradation) and microhabitat (wood fibres) selection in the guts of higher termites.

Bacterial communities in litter, humus and soil feeders

One of the most notable adaptations to humus or soil based diet is the increased alkalinity of the gut fluid, particularly in the first-proctodeal compartment of the

hindgut (Bignell & Eggleton 1995; Brune & Kühl 1996). We observed the association of several genus-level lineages among Ruminococcaceae and Lachnospiraceae (phylum Firmicutes) that are specifically enriched in the alkaline first-proctodeal (P1) compartments of wood feeders (Thongaram *et al.* 2005; Köhler *et al.* 2012), humus feeders (Thongaram *et al.* 2005) and soil feeders (Schmitt-Wagner *et al.* 2003), clearly suggesting a preference of these clostridial lineages for alkaline habitats in the gut. Moreover, in comparison with the generally tubular compartments of wood feeders, humus- and soil-feeding termites are characterized by dilated P1 compartments that probably further magnify the relative abundance of these alkali-adapted clostridia in their hindguts. Interestingly, guts of humivorous *Pachnoda* scarab beetle larvae possess alkaline midguts (Lemke *et al.* 2003), and many phylotypes in clone libraries from *Pachnoda* spp. (Egert *et al.* 2003; Andert *et al.* 2010) have been found to cluster in close phylogenetic neighbourhood of clones from P1 hindgut compartments of humivorous higher termites (Schmitt-Wagner *et al.* 2003; Thongaram *et al.* 2005). The convergent enrichment of related bacterial lineages in distantly related orders of insects further underscores the importance of habitat selection in the structuring of gut communities in analogous gut environments, both characterized by elevated intestinal pH.

Although the exact role of these alkali-adapted clostridia remains unknown, their presence in humus- and soil-feeding termite species may suggest an important role in the digestion of peptides. This hypothesis is supported by the presence of alkaline proteolytic activity (Ji & Brune 2005) and the accumulation of large amounts of ammonia (Ngugi & Brune 2012) in the anterior hindgut of soil feeders. Several genes encoding proteases in the metagenome of the humivorous (dung-feeding) *Amitermes wheeleri* have been binned to the *Clostridiales* (He *et al.* 2013). Interestingly, the distribution patterns of genus-level groups within the Lachnospiraceae and Ruminococcaceae observed in *A. wheeleri* are quite similar to those in termites feeding on litter or humus (see Table S4, Supporting information), suggesting important roles of these lineages in the digestion of peptides in dung and humus.

In addition to the proteolytic activities (Ji & Brune 2005), the hindgut content of soil-feeding termites also shows small amounts of xylanolytic and cellulolytic activities (Rouland & Chararas 1986), which have been suggested to be of bacterial origin (Brauman 2000). Again, the abundance of clostridial genes encoding xylanases in the hindgut metagenome and metatranscriptome of *A. wheeleri* (He *et al.* 2013) suggests that the clostridial lineages in the soil-feeding and humus-feeding termites investigated in the current study play a role also in the digestion of residual polysaccharides of lignocellulose.

Gut content analyses have also shown that members of the humus-feeding termites consume a greater proportion of plant material and/or wood fibres than the true soil feeders (Sleaford *et al.* 1996; Donovan *et al.* 2001), and the composition of artificial lignocellulosic diets strongly affects community structure in wood-feeding *N. takasagoensis*, particularly the abundance of the phyla Fibrobacteres, Spirochaetes and TG3 (Miyata *et al.* 2007). Therefore, differences in the proportion of plant material in the diets of higher termites could explain the higher abundance of fibre-associated bacterial lineages in humus feeders compared to soil feeders, that is *Treponema* clusters Ic and If (Spirochaetes), Termite cluster I (Fibrobacteres) and Termite cluster III (TG3 phylum). Considerable differences in the distribution of these lineages even among the humus feeders agree with differences in the degree of humification in their lignocellulosic diets (Bourguignon *et al.* 2011).

Bacterial communities in fungus feeders

The similarity in gut community structure we observed among fungus-feeding termites from different genera confirms the presence of a core set of bacterial lineages that are either absent or rare in the communities of other higher termites (Dietrich *et al.* 2014; Otani *et al.* 2014). This similarity in community structure among the fungus feeders is driven by their highly specialized diet, which has been shown to include fungal mycelia, in addition to wood or litter (Donovan *et al.* 2001). The shift to a relatively proteinaceous diet among the fungus feeders could be also responsible for the convergence of community structure with that of omnivorous cockroaches (Dietrich *et al.* 2014), particularly in the preferential enrichment of genus-level lineages such as *Alistipes* clusters II, III and IV in both distantly related insect groups (Mikaelyan *et al.* 2015). However, the role of the *Alistipes* clusters in the guts of the termites is still unknown (Poulsen *et al.* 2014). As another result of possible habitat selection, *Macrotermes* sp. harbours Cockroach cluster II of Fibrobacteres, a lineage so far detected only among litter-feeding cockroaches (Mikaelyan *et al.* 2015), which could again be a consequence of dietary similarities between the two insect groups.

Evidence for co-evolution or cospeciation?

Our analyses clearly identify diet as the primary determinant of community structure in higher termites. Nevertheless, each diet group conceals a significant signal of host taxonomy, which is reflected in the clustering of bacterial phylotypes by termite subfamily (Mikaelyan

et al. 2015; Fig. 5). This explains the previously reported patterns of host phylogeny in the bacterial gut microbiota in a broader selection of termite and cockroach hosts (Dietrich *et al.* 2014; Rahman *et al.* 2015), which led to the suggestion of vertical transmission being the primary force shaping the composition of the termite gut microbiota (Rahman *et al.* 2015). However, even a thorough assessment of phylogenetic similarity of gut communities in different hosts (e.g. by the unweighted UniFrac metric) may be misleading, because not all members of the gut community are necessarily co-evolving with the host (Brune & Dietrich 2015).

Using a wider selection of host taxa that comprises members of the same feeding groups from different subfamilies, we were able to demonstrate that the co-evolutionary signal in community structure is masked by an overwhelming signal of host diet. Although a preferential association of certain bacterial lineages with particular host groups is evident already in the classification (Fig. 4) and phylogenetic analyses of short reads (Fig. 5), the co-evolutionary signal is better resolved in phylogenetic analyses of full-length 16S rRNA sequences. For instance, detailed analyses of Fibrobacteres and the TG3 phylum revealed several monophyletic, genus-level lineages that are specifically associated with termites of the same subfamilies (Hongoh *et al.* 2006; Mikaelyan *et al.* 2015).

Although co-evolution may lead to cospeciation, the degree of cocladogenesis will depend on the intimacy of interaction among the partners. In contrast to the hereditary symbioses between insects and their intracellular bacteria (e.g. *Buchnera*, *Blattabacterium*), which have resulted in perfectly congruent phylogenies of symbiont and host (Clark *et al.* 2000; Lo *et al.* 2003), the vertical inheritance of gut microbiota is less reliable (Brune & Dietrich 2015). The exchange of stomodeal or proctodeal fluids among nestmates (trophallaxis) and the consumption of faeces (coprophagy) or nest material (which is composed of faeces) increase the possibility of an uptake of bacteria from the environment, including gut microbiota of other termite species (Nalepa *et al.* 2001).

However, the contribution of horizontal transfer of gut microbiota among different termite lineages is apparently not very large. Already an early study of several *Reticulitermes* and *Microcerotermes* species indicated that the bacterial communities in the guts of congeneric termites are very similar, irrespective of geographical location (Hongoh *et al.* 2005). Also the clustering patterns in our larger and more diverse data set (Fig. 2) cannot be explained by the geographical origin of the samples (Table S5, Supporting information). Nonetheless, it would be important to investigate whether sympatric termite species with more frequent opportunities to exchange gut microbiota, for example inquilines such as *Velocitermes* spp. that cohabit the termitaria of other termites (Florencio *et al.* 2013), share gut microbiota with their respective host. A horizontal transmission of microbial symbionts, for example during territory defence, has been used to explain high similarities in the flagellate assemblages of Termopsidae and certain Rhinotermitidae (*Reticulitermes* spp.), which sometimes occur even within the same log of wood (Kitade 2004).

Although the phylogenetic continuity in the identity of other microbial lineages that occupy important niches in the gut is in agreement with the vertical transmission of symbionts, some of the co-evolutionary patterns observed at higher taxonomic levels may merely be the result of a diffuse selective pressure that the gut habitat exerts on microbial lineages that are transferred via the environment (Won *et al.* 2003; Fontanez & Cavanaugh 2014). Convincing evidence for cospeciation cannot be obtained from the analysis of short reads but will require the highly resolved phylogenetic analysis of candidate symbionts from a sufficient number of closely related termites.

Acknowledgements

We thank Katja Meuser for excellent laboratory assistance. This study was funded by a grant of the Deutsche Forschungsgemeinschaft (DFG) in the Collaborative Research Center SFB 987 and the LOEWE Programme of the State of Hessen (Synmikro). AM was supported by the Synmikro Post-Doc Programme. CD received a scholarship from the International Max Planck Research School for Environmental, Cellular and Molecular Microbiology (IMPRS-MIC). MP was supported by a STENO grant from the Danish Agency for Science, Technology and Innovation.

References

Andert J, Marten A, Brandl R, Brune A (2010) Inter- and intraspecific comparison of the bacterial assemblages in the hindgut of humivorous scarab beetle larvae (*Pachnoda* spp.). *FEMS Microbiology Ecology*, **74**, 439–449.

Bignell DE, Eggleton P (1995) On the elevated intestinal pH of higher termites (Isoptera: Termitidae). *Insectes Sociaux*, **42**, 57–69.

Bignell DE, Roisin Y, Lo N (2011) Global biogeography of termites. In: *Biology of Termites: A Modern Synthesis*(eds Bignell DE, Roisin Y, Lo N), pp. 477–498. Springer Netherlands, Dodrecht.

Bourguignon T, Šobotník J, Lepoint G *et al.* (2011) Feeding ecology and phylogenetic structure of a complex neotropical termite assemblage, revealed by nitrogen stable isotope ratios. *Ecological Entomology*, **36**, 261–269.

Bourguignon T, Lo N, Cameron SL *et al.* (2015) The evolutionary history of termites as inferred from 66 mitochondrial genomes. *Molecular Biology and Evolution*, **32**, 406–421.

Brauman A (2000) Effect of gut transit and mound deposit on soil organic matter transformations in the soil feeding termite: a review. *European Journal of Soil Biology*, **36**, 117–125.

Brugerolle G, Radek R (2006) Symbiotic protozoa of termites. In: *Intestinal Microorganisms of Termites and Other Invertebrates* (eds König H, Varma A), pp. 243–269. Springer, Berlin, Heidelberg.

Brune A (2014) Symbiotic digestion of lignocellulose in termite guts. *Nature Reviews Microbiology*, **12**, 168–180.

Brune A, Dietrich C (2015) The gut microbiota of termites: digesting the diversity in the light of ecology and evolution. *Annual Review of Microbiology*, **69**, in press. doi: 10.1146/annurev-micro-092412-155715.

Brune A, Kühl M (1996) pH profiles of the extremely alkaline hindguts of soil-feeding termites (Isoptera: Termitidae) determined with microelectrodes. *Journal of Insect Physiology*, **42**, 1121–1127.

Brune A, Ohkuma M (2011) Role of the termite gut microbiota in symbiotic digestion. In: *Biology of Termites*(eds Bignell DE, Roisin Y, Lo N), pp. 439–475. Springer, Dodrecht.

Brune A, Emerson D, Breznak JA (1995) The termite gut microflora as an oxygen sink: microelectrode determination of oxygen and pH gradients in guts of lower and higher termites. *Applied and Environmental Microbiology*, **61**, 2681–2687.

Clark MA, Moran NA, Baumann P *et al.* (2000) Cospeciation between bacterial endosymbionts (*Buchnera*) and a recent radiation of aphids (*Uroleucon*) and pitfalls of testing for phylogenetic congruence. *Evolution*, **54**, 517–525.

Colman DR, Toolson EC, Takacs-Vesbach CD (2012) Do diet and taxonomy influence insect gut bacterial communities? *Molecular Ecology*, **21**, 5124–5137.

Daigle D, Simen BB, Pochart P (2011) High-throughput sequencing of PCR products tagged with universal primers using 454 life sciences systems. *Current Protocols in Molecular Biology*, **96**, 1–14.

Degnan P, Ochman H (2012) Illumina-based analysis of microbial community diversity. *The ISME Journal*, **6**, 183–194.

Dietrich C, Brune A (2014) The complete mitogenomes of six higher termite species reconstructed from metagenomic datasets (*Cornitermes* sp., *Cubitermes ugandensis*, *Microcerotermes parvus*, *Nasutitermes corniger*, *Neocapritermes taracua*, and *Termes hospes*). *Mitochondrial DNA*. doi: 10.3109/19401736.2014.987257.

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

Donovan SE, Jones DT (2000) Morphological phylogenetics of termites (Isoptera). *Biological Journal of the Linnean Society*, **70**, 467–513.

Donovan SE, Eggleton P, Bignell DE (2001) Gut content analysis and a new feeding group classification of termites. *Ecological Entomology*, **26**, 356–366.

Edgar RC (2013) UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nature Methods*, **10**, 996–998.

Egert M, Wagner B, Lemke T *et al.* (2003) Microbial community structure in midgut and hindgut of the humus-feeding larva of *Pachnoda ephippiata* (Coleoptera: Scarabaeidae). *Applied and Environmental microbiology*, **69**, 6659–6668.

Eggleton P, Tayasu I (2001) Feeding groups, lifetypes and the global ecology of termites. *Ecological Research*, **16**, 941–960.

Florencio DF, Marins A, Rosa CS *et al.* (2013) Diet segregation between cohabiting builder and inquiline termite species. *PLoS ONE*, **8**, e66535.

Fontanez KM, Cavanaugh CM (2014) Evidence for horizontal transmission from multilocus phylogeny of deep-sea mussel (Mytilidae) symbionts. *Environmental Microbiology*, **16**, 3608–3621.

Gontijo TA, Domingos DJ (1991) Guild distribution of some termites from cerrado vegetation in south-east Brazil. *Journal of Tropical Ecology*, **7**, 523–529.

He S, Ivanova N, Kirton E *et al.* (2013) Comparative metagenomic and metatranscriptomic analysis of hindgut paunch microbiota in wood- and dung-feeding higher termites. *PLoS ONE*, **8**, e61126.

Hongoh Y, Deevong P, Inoue T *et al.* (2005) Intra- and interspecific comparisons of bacterial diversity and community structure support coevolution of gut microbiota and termite host. *Applied and Environmental Microbiology*, **71**, 6590–6599.

Hongoh Y, Deevong P, Hattori S *et al.* (2006) Phylogenetic diversity, localization, and cell morphologies of members of the candidate phylum TG3 and a subphylum in the phylum Fibrobacteres, recently discovered bacterial groups dominant in termite guts. *Applied and Environmental Microbiology*, **72**, 6780–6788.

Inward DJG, Vogler AP, Eggleton P (2007) A comprehensive phylogenetic analysis of termites (Isoptera) illuminates key aspects of their evolutionary biology. *Molecular Phylogenetics and Evolution*, **44**, 953–967.

Ji R, Brune A (2005) Digestion of peptidic residues in humic substances by an alkali-stable and humic-acid-tolerant proteolytic activity in the gut of soil-feeding termites. *Soil Biology and Biochemistry*, **37**, 1648–1655.

Kitade O (2004) Comparison of symbiotic flagellate faunae between termites and a wood-feeding cockroach of the genus *Cryptocercus*. *Microbes and Environments*, **19**, 215–220.

Köhler T, Dietrich C, Scheffrahn RH, Brune A (2012) High-resolution analysis of gut environment and bacterial microbiota reveals functional compartmentation of the gut in wood-feeding higher termites (*Nasutitermes* spp.). *Applied and Environmental Microbiology*, **78**, 4691–4701.

Krishna K, Grimaldi DA, Krishna V, Engel MS (2013) Treatise on the Isoptera of the world. *Bulletin of the American Museum of Natural History*. American Museum of Natural History, New York.

Lemke T, Stingl U, Egert M, Friedrich MW, Brune A (2003) Physicochemical conditions and microbial activities in the highly alkaline gut of the humus-feeding larva of *Pachnoda ephippiata* (Coleoptera: Scarabaeidae). *Applied and Environmental Microbiology*, **69**, 6650–6658.

Lo N, Bandi C, Watanabe H, Nalepa C, Beninati T (2003) Evidence for cocladogenesis between diverse dictyopteran lineages and their intracellular endosymbionts. *Molecular Biology and Evolution*, **20**, 907–913.

Lozupone CA, Hamady M, Kelley ST, Knight R (2007) Quantitative and qualitative β diversity measures lead to different insights into factors that structure microbial communities. *Applied and Environmental Microbiology*, **73**, 1576–1585.

Meyer F, Paarmann D, D'Souza M *et al.* (2008) The metagenomics RAST server—a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics*, **9**, 386.

Mikaelyan A, Strassert JFH, Tokuda G, Brune A (2014) The fibre-associated cellulolytic bacterial community in the hindgut of wood-feeding higher termites (*Nasutitermes* spp.). *Environmental Microbiology*, **16**, 2711–2722.

Mikaelyan A, Köhler T, Lampert N *et al.* (2015) Classifying the bacterial gut microbiota of termites and cockroaches: a curated phylogenetic reference database (DictDb). *Systematic and Applied Microbiology*, in press. doi: 10.1016/j.syapm.2015.07.004.

Miyata R, Noda N, Tamaki H *et al.* (2007) Influence of feed components on symbiotic bacterial community structure in the gut of the wood-feeding higher termite *Nasutitermes takasagoensis*. *Bioscience, Biotechnology, and Biochemistry*, **71**, 1244–1251.

Nalepa CA, Bignell DE, Bandi C (2001) Detritivory, coprophagy, and the evolution of digestive mutualisms in Dictyoptera. *Insectes Sociaux*, **48**, 194–201.

Ngugi DK, Brune A (2012) Nitrate reduction, nitrous oxide formation, and anaerobic ammonia oxidation to nitrite in the gut of soil-feeding termites (*Cubitermes* and *Ophiotermes* spp.). *Environmental Microbiology*, **14**, 860–871.

Ni J, Tokuda G (2013) Lignocellulose-degrading enzymes from termites and their symbiotic microbiota. *Biotechnology Advances*, **31**, 838–850.

Noirot C (2001) The gut of termites (Isoptera). *Annales de la Société Entomologique de France*, **37**, 431–471.

Oksanen JF, Blanchet G, Kindt R *et al.* (2015) *vegan: Community ecology package*. R package version 2.3-0. Available from http://CRAN.R-project.org/package=vegan.

Otani S, Mikaelyan A, Nobre T *et al.* (2014) Identifying the core microbial community in the gut of fungus-growing termites. *Molecular Ecology*, **23**, 4631–4644.

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

Paul K, Nonoh JO, Mikulski L, Brune A (2012) "Methanoplasmatales", Thermoplasmatales-related archaea in termite guts and other environments, are the seventh order of methanogens. *Applied and Environmental Microbiology*, **78**, 8245–8253.

Poulsen M (2015) Towards an integrated understanding of the consequences of fungus domestication on the fungus-growing termite gut microbiota. *Environmental Microbiology*. doi: 10.1111/1462-2920.12765.

Poulsen M, Hu H, Li C *et al.* (2014) Complementary symbiont contributions to plant decomposition in a fungus-farming termite. *Proceedings of the National Academy of Sciences of the USA*, **111**, 14500–14505.

Price MN, Dehal PS, Arkin AP (2009) FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Molecular Biology and Evolution*, **26**, 1641–1650.

R Core Team (2015) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. Available from http://www.R-project.org/.

Rahman NA, Parks DH, Willner DL *et al.* (2015) A molecular survey of Australian and North American termite genera indicates that vertical inheritance is the primary force shaping termite gut microbiomes. *Microbiome*, **3**, 1–16.

Rouland C, Chararas C (1986) Étude comparée des osidases de trois espèces de termites africains à régime alimentaire différent. *Comptes rendus de l'Académie des sciences*, **302**, 341–345.

Schloss PD, Westcott SL, Ryabin T *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.

Schmitt-Wagner D, Friedrich MW, Wagner B, Brune A (2003) Phylogenetic diversity, abundance, and axial distribution of Bacteria in the intestinal tract of two soil-feeding termites (*Cubitermes* spp.). *Applied and Environmental Microbiology*, **69**, 6007–6017.

Sleaford F, Bignell DE, Eggleton P (1996) A pilot analysis of gut contents in termites from the Mbalmayo Forest Reserve, Cameroon. *Ecological Entomology*, **21**, 279–288.

Suzuki R, Shimodaira H (2006) Pvclust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics*, **22**, 1540–1542.

Thongaram T, Hongoh Y, Kosono S *et al.* (2005) Comparison of bacterial communities in the alkaline gut segment among various species of higher termites. *Extremophiles*, **9**, 229–238.

Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology*, **73**, 5261–5267.

Warnecke F, Luginbühl P, Ivanova N *et al.* (2007) Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. *Nature*, **450**, 560–565.

Won YJ, Hallam SJ, O'Mullan GD *et al.* (2003) Environmental acquisition of thiotrophic endosymbionts by deep-sea mussels of the genus *Bathymodiolus*. *Applied and Environmental Microbiology*, **69**, 6785–6792.

A.M., T.K. and A.B. conceived and designed the experiments. M.P. and D.S.-D. collected and identified termites. A.M. and C.D. performed the experiments and analysed the data. C.D. and D.S.-D. contributed to the writing the manuscript. A.M. and A.B. wrote the manuscript.

Data accessibility

The de-replicated, quality-checked MiSeq data sets have been archived at MG-RAST under the accession numbers 4639074.3–4639081.3. The COII nucleotide alignment and tree, the phylogenetic tree of OTU representatives (encoding abundance and sample information in the labels), the OTU table for each sample, and the Bray–Curtis and UniFrac dissimilarity matrices are archived at Dryad (http://dx.doi.org/10.5061/dryad.v46f0).

Supporting information

Additional supporting information may be found in the online version of this article.

Fig. S1 Maximum-likelihood tree constructed using COII DNA sequences from 18 termite hosts included in this study in the context of close relatives in public databases.

Fig. S2 Maximum-likelihood tree of 16S rRNA phylotypes from the different subfamilies of higher termites.

Table S1 Characteristics of the 16S rRNA amplicon libraries and the classification success obtained for each library with DictDb v 3.0 at different taxonomic levels.

Table S2 Relative abundance of bacterial taxa in iTag and pyrotag libraries of 16S rRNA amplicons from different higher termites (arranged by host phylogeny).

Table S3 Relative abundance of bacterial taxa in the iTag and pyrotag amplicon libraries of 16S rRNA genes from the guts of different higher termites (arranged by host diet).

Table S4 Comparison of the distribution of genus-level clusters in Lachnospiraceae and Ruminococcaceae (order Clostridiales) in a library from *Amitermes wheeleri* (He *et al*. 2013) and the termites investigated in the current study.

Table S5 The correlation of the diet, host taxonomy and geographical origin with Bray–Curtis and UniFrac distances tested for significance tested with *adonis*.

Table S6 Pairwise Kruskal–Wallis tests for significant differences in the distribution of bacterial lineages for different diet groups of termites (F, fungus feeders; S, soil feeders; H, humus/litter feeders; W, wood/grass feeders).

Table S7 Pairwise Kruskal–Wallis tests for significant differences in the distribution of bacterial lineages for different combinations of taxonomic groups of termites (At, Apicotermitinae; Mt, Macrotermitinae; Nt, Nasutitermitinae; Tt, Termitinae; St, Syntermitinae).