Complementary symbiont contributions to plant decomposition in a fungus-farming termite

Michael Poulsen[a,1,2], Haofu Hu[b,1], Cai Li[b,c], Zhensheng Chen[b], Luohao Xu[b], Saria Otani[a], Sanne Nygaard[a], Tania Nobre[d,3], Sylvia Klaubauf[e], Philipp M. Schindler[f], Frank Hauser[g], Hailin Pan[b], Zhikai Yang[b], Anton S. M. Sonnenberg[h], Z. Wilhelm de Beer[i], Yong Zhang[b], Michael J. Wingfield[i], Cornelis J. P. Grimmelikhuijzen[g], Ronald P. de Vries[e], Judith Korb[f,4], Duur K. Aanen[d], Jun Wang[b,j], Jacobus J. Boomsma[a], and Guojie Zhang[a,b,2]

[a]Centre for Social Evolution, Department of Biology, University of Copenhagen, DK-2100 Copenhagen, Denmark; [b]China National Genebank, BGI-Shenzen, Shenzhen 518083, China; [c]Centre for GeoGenetics, Natural History Museum of Denmark, University of Copenhagen, DK-1350 Copenhagen, Denmark; [d]Laboratory of Genetics, Wageningen University, 6708 PB, Wageningen, The Netherlands; [e]Fungal Biodiversity Centre, Centraalbureau voor Schimmelcultures, Royal Netherlands Academy of Arts and Sciences, NL-3584 CT, Utrecht, The Netherlands; [f]Behavioral Biology, Fachbereich Biology/Chemistry, University of Osnabrück, D-49076 Osnabrück, Germany; [g]Center for Functional and Comparative Insect Genomics, Department of Biology, University of Copenhagen, DK-2100 Copenhagen, Denmark; [h]Department of Plant Breeding, Wageningen University and Research Centre, NL-6708 PB, Wageningen, The Netherlands; [i]Department of Microbiology, Forestry and Agricultural Biotechnology Institute, University of Pretoria, Pretoria SA-0083, South Africa; and [j]Department of Biology, University of Copenhagen, DK-2100 Copenhagen, Denmark

Edited by Ian T. Baldwin, Max Planck Institute for Chemical Ecology, Jena, Germany, and approved August 15, 2014 (received for review October 24, 2013)

Termites normally rely on gut symbionts to decompose organic matter but the Macrotermitinae domesticated *Termitomyces* fungi to produce their own food. This transition was accompanied by a shift in the composition of the gut microbiota, but the complementary roles of these bacteria in the symbiosis have remained enigmatic. We obtained high-quality annotated draft genomes of the termite *Macrotermes natalensis*, its *Termitomyces* symbiont, and gut metagenomes from workers, soldiers, and a queen. We show that members from 111 of the 128 known glycoside hydrolase families are represented in the symbiosis, that *Termitomyces* has the genomic capacity to handle complex carbohydrates, and that worker gut microbes primarily contribute enzymes for final digestion of oligosaccharides. This apparent division of labor is consistent with the *Macrotermes* gut microbes being most important during the second passage of comb material through the termite gut, after a first gut passage where the crude plant substrate is inoculated with *Termitomyces* asexual spores so that initial fungal growth and polysaccharide decomposition can proceed with high efficiency. Complex conversion of biomass in termite mounds thus appears to be mainly accomplished by complementary cooperation between a domesticated fungal monoculture and a specialized bacterial community. In sharp contrast, the gut microbiota of the queen had highly reduced plant decomposition potential, suggesting that mature reproductives digest fungal material provided by workers rather than plant substrate.

carbohydrate-active enzymes | eusocial | symbioses | cellulose | lignin

Interspecific mutualism usually allows partner species preferential access to complementary resources. Some hosts internalized microbial symbionts, leading to vertical transmission and varying degrees of genome loss (1), whereas others domesticated external partners that maintained independent reproduction (2). Understanding how such ectosymbioses remain evolutionarily stable is challenging (3) because prokaryote and eukaryote symbionts form interacting communities, which may be difficult for hosts to control when symbionts can achieve higher fitness by pursuing selfish reproductive strategies (4). Digestive symbiotic communities in animal guts provide excellent examples of such ambiguities; recent studies of human microbiotas show that gut communities vary by subject age, geography (5), and diet (6) and that deviating microbiotas can be associated with compromised health (7).

Given the continuous flow of food through animal guts, it is intriguing that adaptive microbiotas can normally be maintained (8–10) without invasion by less beneficial or harmful microbes (11). Insect lineages that have relied on nutritional symbioses have existed and adaptively radiated for tens of millions of years, suggesting that the benefits of these symbioses surpass the potential levels-of-selection conflicts that need to be regulated (12). However, beyond examples from humans and some domesticated ungulates, we lack fundamental insight into the genes involved, their expression, and their phenotypic functions. Termites provide a case in point, as they originated >150 Mya and have relied on protist and bacterial gut symbionts for the breakdown of lignocellulose throughout their evolutionary history (13), allowing them to become dominant decomposers in terrestrial ecosystems (13, 14).

Significance

Old World (sub)tropical fungus-growing termites owe their massive ecological footprints to an advanced symbiosis with *Termitomyces* fungi. They also have abundant gut bacteria, but the complementarity roles of these symbionts have remained unclear. We analyzed the genomic potential for biomass decomposition in a farming termite, its fungal symbiont, and its bacterial gut communities. We found that plant biomass conversion is mostly a multistage complementary cooperation between *Termitomyces* and gut bacteria, with termite farmers primarily providing the gut compartments, foraging, and nest building. A mature queen had highly reduced gut microbial diversity for decomposition enzymes, suggesting she had an exclusively fungal diet even though she may have been the source of the gut microbes of the colony's first workers and soldiers.

Author contributions: M.P., S.N., T.N., J.K., D.K.A., J.J.B., and G.Z. designed research; M.P., H.H., C.L., Z.C., L.X., S.O., S.N., S.K., P.M.S., H.P., Z.Y., Z.W.d.B., C.J.P.G., J.K., and D.K.A. performed research; M.P., F.H., A.S.M.S., Z.W.d.B., Y.Z., M.J.W., C.J.P.G., R.P.d.V., J.K., D.K.A., J.W., J.J.B., and G.Z. contributed new reagents/analytic tools; M.P., H.H., C.L., Z.C., L.X., S.O., S.N., S.K., P.M.S., F.H., H.P., Z.Y., A.S.M.S., Z.W.d.B., M.J.W., C.J.P.G., R.P.d.V., J.K., D.K.A., J.J.B., and G.Z. analyzed data; and M.P., H.H., C.L., S.N., T.N., M.J.W., C.J.P.G., R.P.d.V., J.K., D.K.A., J.J.B., and G.Z. wrote the paper.

The authors declare no conflict of interest.

Data deposition: All raw reads reported in this paper have been deposited in the NCBI-SRA database [accession nos. SRA069856 (*Macrotermes natalensis* genome), SRA071609 (*Termitomyces* genome), and SRA071613 (gut metagenomes)], and assemblies are available from the GigaScience Database: http://dx.doi.org/10.5524/100055.

[1]M.P. and H.H. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: mpoulsen@bio.ku.dk or zhanggj@genomics.cn.

[3]Present address: Institute of Mediterranean Agrarian and Environmental Sciences (Pólo da Mitra), University of Évora, 7002-554 Évora, Portugal.

[4]Present address: Department of Evolutionary Biology and Animal Ecology, Institute of Biology I (Zoology), University Freiburg, D-79104 Freiburg, Germany.

A single monophyletic subfamily, the Macrotermitinae, realized a major evolutionary transition *ca*. 30 Mya, when they domesticated the ancestor of the fungal genus *Termitomyces* (15). They have radiated into 11 termite genera with more than 330 extant species (15, 16) to collectively obtain a massive ecological footprint in the Old World (sub)tropics, matched only by the fungus-growing (attine) ants of the New World (14, 17). Throughout their evolutionary history, the partnership with five major clades of *Termitomyces* has remained obligate, as no macrotermitine termite is known to have abandoned fungus farming or to rear other fungi than *Termitomyces* (15) (Fig. 1 *A* and *B*). Coinciding with the domestication of *Termitomyces*, the common ancestor of the Macrotermitinae underwent a major shift in the bacterial gut community (18, 19). The fungus-growing termites thus represent a major metazoan radiation based on a simultaneous tripartite life-history transition: insects becoming farmers, fungi becoming crops, and gut microbiotas adopting largely unknown complementary roles.

Fungus-growing termites rely on the external decomposition of plant substrate by their *Termitomyces* fungus garden symbiont. In *Macrotermes* species, the fungus comb is managed in a highly structured way, with older workers collecting crude forage material and bringing it back to the nest, where younger workers ingest it together with asexual *Termitomyces* spores (conidia) provided by fungal nodules from established "fungus-garden combs" to produce primary feces that is deposited as new layers of comb (17, 20) (Fig. 1*C*). This new substrate quickly develops dense hyphal networks and produces the next cohorts of nodules (2, 20), whereas older termites ultimately consume the old comb (Fig. 1*C*). This combination of substrate processing and inoculation at first gut passage followed by a second digestive phase makes the termite gut the central operational compartment of the symbiosis. It is here that the entire genetic potential of all members of the symbiosis comes together, presumably shaped by natural selection for optimal collective performance in two sequential digestive phases. To investigate functional complementarity of the three major components of the mutualism, we (*i*) obtained high-quality draft genome sequences of the fungus-growing termite *Macrotermes natalensis*, its *Termitomyces* sp. symbiont, and several caste-specific gut microbiotas; (*ii*) analyzed the genomic potential for lignocellulolytic enzyme potential to assess functional contributions across partners; and (*iii*) compared gut microbiotas across sterile and reproductive castes to evaluate functional gut specialization across termite family members.

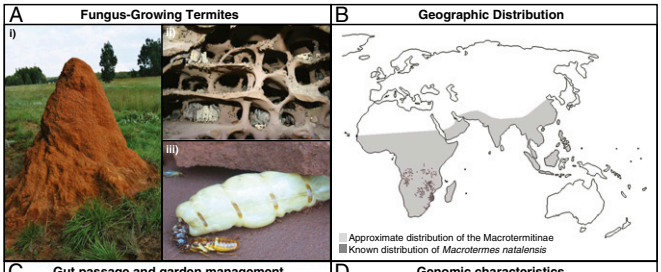

Fig. 1. The fungus-growing termite symbiosis and its genomic characteristics. (*A*) A *Macrotermes natalensis* colony in South Africa: (*i*) the underground fungus comb in which *Termitomyces* is maintained and (*ii* and *iii*) the royal chamber with the queen (*ii*) and the king (*iii*). (*B*) Geographic distribution of the Macrotermitinae (gray), with darker areas in southern Africa highlighting the known occurrences of *M. natalensis* (adapted from ref. 61). (*C*) The substrate and recurrent *Termitomyces* inoculation within a colony centered around the termite gut: Asexual *Termitomyces* spores from fungus comb nodules (*i*) and plant biomass substrate (*ii*) are mixed within the termite gut (*iii*, first gut passage) to become the new fungus comb substrate (*iv*) within which *Termitomyces* hyphae grow to maturity so that new nodules with asexual spores are produced (*v*) until the plant substrate is fully used and the old comb (*vi*) is consumed by the termites (*vii*, second gut passage). (*D*) To characterize the genetic potential of the fungus-growing termite symbiosis, we sequenced *M. natalensis* and *Termitomyces* and obtained worker, soldier, and queen gut metagenomes (*SI Appendix* and GigaScience Database, http://dx.doi.org/10.5524/100055).

Results and Discussion

Sequencing the Symbiotic Community. A *M. natalensis* colony was excavated in South Africa in 2011. DNA was extracted from the queen for genome sequencing of the termite on an Illumina platform, and RNA from the queen, the king, and workers and soldiers from four colonies of the same species was extracted for transcriptome sequencing. The *ca*. 1.3-Gb genome of *M. natalensis* is one of the largest insect genomes sequenced to date, but it was assembled to a high-quality (69× coverage) draft with >16,000 annotated genes, a scaffold N50 of 2 Mb, and with 246 of the 248 conserved eukaryotic genes (CEGs) (Fig. 1*D* and *SI Appendix*, Tables S1–S9 and S21).

For *Termitomyces* genome sequencing, DNA was extracted from a laboratory-grown homokaryon obtained by protoplasting of a heterokaryon from a sympatric colony of the same termite species. The draft genome of *Termitomyces* (81× coverage; scaffold N50 = 262 kb; 244 CEGs) had an estimated size of 83.7 Mb and allowed the annotation of >11,000 genes (Fig. 1*D* and *SI Appendix*, Tables S2 and S10–S16) (21). Gut microbiotas were obtained using the same platform and similar analyses, resulting in scaffold assemblies of 446 Mb for young major worker guts (>1,200,000 complete or partial genes), 337 Mb for minor soldier guts (>900,000 genes), and 33 Mb for the queen gut (>50,000 genes) (Fig. 1*D* and *SI Appendix*, Tables S17–S21) (21).

Genetic Potential and Symbiotic Contributions to Plant Decomposition. We identified more than 2.2 million complete or partial genes from the symbiosis and focused our analyses on characterizing carbohydrate-active enzymes (CAZymes) involved in plant biomass decomposition, as this sequence-based family classification (www.cazy.org) (22) allows similarities and differences in plant substrate specificities to be mapped. Although precise molecular details linking CAZy genes to function are difficult to predict from CAZy family assignments alone, and specific CAZymes may have atypical activities relative to other members of their CAZy family, previous work has shown that families can be assigned to broad functional categories with reasonable accuracy (23). The *M. natalensis* symbiosis collectively produced 30,297 hits to partial or complete glycoside hydrolases (GHs), 11,622 glycoside transferases (GTs), 4,380 carbohydrate-binding modules (CBMs), 2,729 carbohydrate esterases (CEs), 793 polysaccharide lyases (PLs), and 79 auxiliary activities (AAs) (Dataset S1, Tables S27–S29). Within the most abundant CAZy category (GHs), we identified 85 partial or complete GHs in *M. natalensis*, including the likely functional GH9 cellulase, and 201 in *Termitomyces*, whereas worker and soldier gut microbiotas harbored 15,619 and 14,228 partial or complete GHs, respectively (Fig. 2 and Dataset S1, Tables S27–S29).

Whereas absolute numbers of GH hits may hide considerable functional redundancy, GH family-level diversity showed similar differences. Also here, the gut microbiota fraction was higher than the combined contributions of the eukaryote partners, with 98 GH families represented in worker guts, 48 in *Termitomyces*, and 27 in *M. natalensis* (Fig. 2). The total diversity of GH families present across the three symbiotic partners consisted of representatives of 111 (86.7%) of the 128 GH families (including GH61, which was recently renamed AA9, and unassigned GHs in the family GH0) that have been characterized in living organisms (Fig. 2). Although comparable analyses from other complex systems are lacking, these

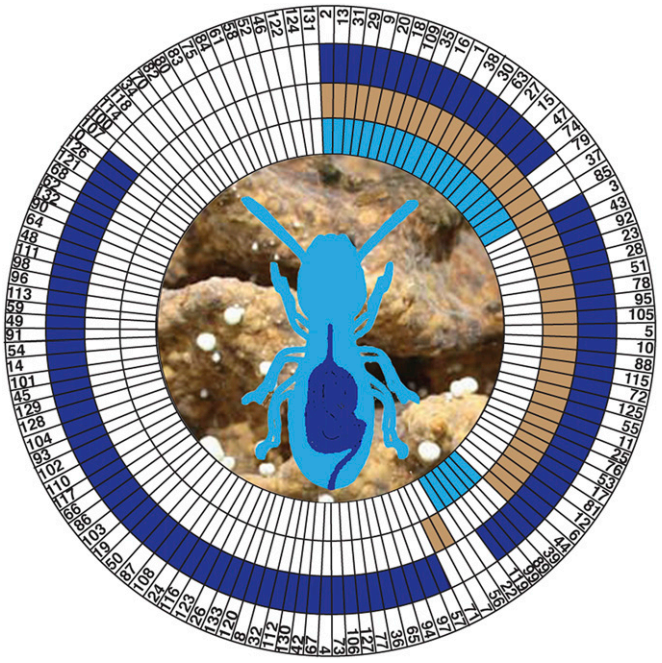

Fig. 2. Functional complementary contributions to biomass degradation. Using the carbohydrate-active enzyme database (www.cazy.org), we classified glycoside hydrolases (GHs), polysaccharide lyases (PLs), carbohydrate esterases (CEs), carbohydrate binding modules (CBMs), and glycoside transferases (GTs) in the *Macrotermes*, *Termitomyces*, and worker gut microbes (Dataset S1, Tables S27–S29). The GHs shown here were by far the most abundant enzyme class: 85 were identified in *M. natalensis* (27 GH families, light blue), 201 in *Termitomyces* (48 GH families, brown), and 15,619 in the worker gut microbiota (98 GH families, dark blue). The presence (color)/absence (white) pattern shows that the 111 GH families identified in the symbiosis represent 86.7% of all known GH families. Numbers for the CBMs (73.9%), GTs (68,4%), PLs (78,3%), and CEs (100%) were of similar magnitude (Dataset S1, Tables S27–S29). For enzyme names and key activities, including EC numbers, see Dataset S1, Table S31.

numbers indicate that the collective genetic potential of this tripartite symbiosis has a massive lignocellulolytic capacity with relatively minor termite contributions. Exploring only GH families directly involved in the targeted breakdown of plant polysaccharides, and adjusting for the total number of predicted genes in the three partners, the gut metagenome had a 20-fold higher (0.50% of total number of genes are GHs) and *Termitomyces* a 30-fold higher (0.78%) share of the total GH-family representation than the termite (0.02%) ($\chi^2 = 91.23$; df = 2; $P < 0.0001$).

Analyses of the *M. natalensis* Genome. The *M. natalensis* GH9 cellulases have previously been characterized from termites, cockroaches, and other insects (24). The recent publication of the genome of the dampwood termite *Zootermopsis nevadensis* (25) allowed for a comparison between CAZyme profiles of the two termite species. Although *M. natalensis* contained only 85 enzymes from 27 GH families compared with 97 CAZymes in 28 GH families in *Z. nevadensis*, the relative abundances were remarkably similar (Dataset S1, Table S28 and *SI Appendix*, Fig. S9). Although this reduction may imply that *M. natalensis* has reduced CAZy potential encoded, the sequencing of genomes of other higher termites will be needed to thoroughly assess expansions, contractions, and gains/losses of termite GH families associated with the emergence of fungiculture.

Using computational analysis of gene family evolution (CAFE) and subsequent manual checking, we found that three gene families were reduced in gene number (contracted) in the termite relative to other genomes: esterase FE4, trypsin, and the short-chain dehydrogenase/reductase (SDR) superfamily (*SI Appendix*, Table S22 and Fig. S2). The former two gene families are associated with digestion (26), so their contractions may be associated with a rather uniform diet compared with many other insects, consistent with the esterase FE4 gene family also being contracted in the genome of the fungus-growing ant *Acromyrmex echinatior* (27). The SDR superfamily includes genes from pathways for lipid, amino acid, carbohydrate, cofactor, hormone, and xenobiotics metabolism, as well as redox sensor mechanisms (28). Also this contraction may thus relate to nutrition, but further work will be needed to clarify such connections and explore possible links to concomitant evolutionary change in functional genes of *Termitomyces* and the termite gut microbiota (cf. ref. 29). Finally, a unique presence/absence spectrum of 39 neuropeptides, protein hormones, and biogenic amines and their receptors involved in central physiological processes was characterized (details in Dataset S1, Table S36).

The *Termitomyces* Genome and Its Genetic Potential for Plant Decomposition. CAFE analysis of the *Termitomyces* genome showed 10 gene family expansions and 4 gene family contractions (*SI Appendix*, Tables S23 and S24 and Figs. S3–S5). The Chitinase 1 (GH19) family expansion may be related to the high growth rates of *Termitomyces* within the fungus comb, and the increased presence of feruloyl esterases (CE1) and unsaturated rhamnogalacturonyl hydrolases (GH88) may be associated with selection for rapid breakdown of predigested plant material (30). The contraction of the MAL32 α-glucosidase family (part of GH13; *SI Appendix*, Table S25) suggests that *Termitomyces* has a reduced capacity for the breakdown of oligosaccharides, which is likely taken care of by the gut microbiota, where these genes are abundantly present (Figs. 3*B* and 4*C* and Dataset S1, Tables S29 and S31).

Termitomyces has a very broad range of plant polysaccharide degrading enzymes, indicating that it does not depend on a specific substrate provided by the termite host. Consistent with these inferences, subsequent CAZyme analyses showed that *Termitomyces* is not particularly enriched or reduced for many CAZy families compared with other fungi (Fig. 3*B* and Dataset S1, Table S30). The most pronounced exceptions to this rule were the enrichments of GH49s (dextranases), GH79 (glucuronidase, heparanase), and GH10 (xylanases) and the largest contractions observed were in GH47 (mannosidase) and GH13, containing enzymes involved in the utilization of inulin and sucrose (Fig. 3*B* and Dataset S1, Table S30).

In Vitro Growth Profiles of *Termitomyces* Support Inferences from Genome Analyses. Growth rates of plated cultures of *Termitomyces* and other fungi on a range of relevant substrates (*SI Appendix*, Fig. S8*B*) confirmed that *Termitomyces* can degrade complex polysaccharides, as it grew very well on cellobiose and particularly cellulose, relative to free-living fungi. However, CAZyme analysis showed that the total number of cellulases in the *Termitomyces* genome was not greatly enhanced, suggesting higher specific activity of these enzymes, up-regulation of cellulase gene expression similar to *Trichoderma reesei*, higher synergy between the cellulolytic components, or the presence of a better and larger spectrum of polysaccharide monooxygenases than in *Trichoderma* (Dataset S1, Tables S30 and S31) (31).

Although *Termitomyces* has a moderate suite of starch/maltose degrading enzymes (Dataset S1, Table S30), it grew on starch but not on maltose, indicating that it is unable to split short oligosaccharides (*SI Appendix*, Fig. S8*A*). Fungal growth on cottonseed hulls is normally associated with proficient growth on lignin, but *Termitomyces* did not grow when lignin was the sole carbon source (*SI Appendix*, Fig. S8). This suggests that lignin cleavage is accomplished (ref. 32 and this study), consistent with lignin content decreasing from young to old comb (33), but that this merely facilitates access to other plant components rather than utilization of lignin as a specific carbon source.

Complementary Carbohydrate-Active Enzymes in Worker and Soldier Guts. Consistent with GH13 genes being markedly reduced in abundance, *Termitomyces* did not grow well on simple oligosaccharides

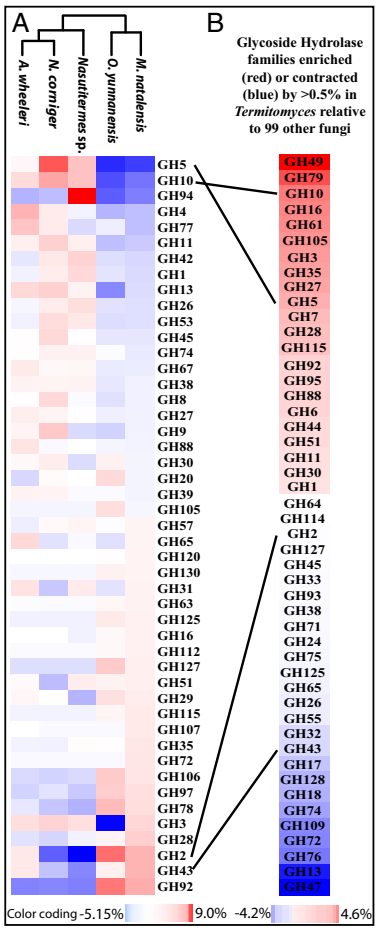

Fig. 3. Complementary contributions to the spectrum of carbohydrate-active enzymes in *Termitomyces* and termite worker gut microbiota. (*A*) A heat map of GH families enriched (red) or contracted (blue) in relative abundance across five termite species: the dung-feeding higher termite *Amitermes wheeleri* (34), two species of wood-feeding higher termites [*Nasutitermes corniger* (34) and *Nasutitermes* sp (35)], and two fungus-growing termite species [*Odontotermes yunnanensis* (36) and *M. natalensis* (this study)]. Only GH families with at least one termite species exhibiting >0.25% enrichment or contraction are shown (Dataset S1, Table S33). Cluster analyses showed that the two fungus-growing termite species were more similar to each other in GH composition than to nonfarming termites (non-parametric *P* value = 0.03 after 10,000 Monte Carlo permutations; details in *SI Appendix*). (*B*) GH families enriched (red) or contracted (blue) by >0.5% in the *Termitomyces* fungal symbiont relative to 99 fungi (62) (Dataset S1, Table S30). GH families connected with lines were enriched in *Termitomyces* and contracted in the *M. natalensis* worker gut microbiota or vice versa.

(*SI Appendix*, Fig. S8), but microbial CAZymes amply cover these functions (Figs. 2 and 3). The functional CAZyme spectra of *Termitomyces* and the gut microbiota overlapped, but there was also substantial complementarity, with numerous enzymes in the worker gut microbiota targeting oligosaccharides, thus complementing the genetic potential of *Termitomyces*. To further test whether the gut microbiota have indeed shifted toward digesting simpler carbohydrates after *Termitomyces* was domesticated as a crop fungus, we compared the GH composition in *M. natalensis* workers with that in workers from the dung-feeding higher termite *Amitermes wheeleri* (34), two species of *Nasutitermes* wood-feeding higher termites [*Nasutitermes corniger* (34) and *Nasutitermes* sp. (35)], and the fungus-growing termite *Odontotermes yunnanensis* (36) (Dataset S1, Table S33 and Fig. 3*A*).

Clustering analysis revealed that GH profiles in fungus-growing termite guts were significantly more similar to each other than to non-fungus–growing termites and that GH families with reduced relative abundance in fungus-growing termites included enzymes targeting complex polysaccharides (e.g., GH5, GH10, and GH94), whereas enzymes from enriched families tended to be involved in the breakdown of relatively simple oligosaccharides (e.g., GH92, GH43, and GH2; Fig. 3*A* and Dataset S1, Table S33). This cumulative evidence suggests that the *Macrotermes* gut microbiota do most of the final digestion during the second gut passage of the comb material, whereas the first gut passage (Fig. 1) mainly functions to mix the crude substrate with *Termitomyces* conidia, so that initial fungal growth and polysaccharide decomposition can proceed at high rates.

Functional Diversity of Worker and Soldier Gut Metagenomes. Phylogenetic classifications of caste-specific termite microbiotas revealed the presence of representatives from 420 bacterial genera, although only 239 of these were present in more than 0.02% relative abundance in at least one caste gut metagenome (Dataset S1, Table S26 and Fig. 4*A*). Some bacterial genera were unique in a single caste or present in only two of three castes, but none of these had appreciable abundances: 0.001 ± 0.0009% (mean ± SE) in workers, 0.0004 ± 0.0001% in soldiers, and 0.0009 ± 0.003% in the queen (Dataset S1, Table S26). Workers and soldiers shared all bacterial genera with abundances >0.02% in at least one caste, but the queen gut contained less than half of these genera (Fig. 4*A*). Rarefaction curves confirmed that we had performed sufficient sampling to capture the vast majority of genera in the symbiosis, including the queen gut, despite an order

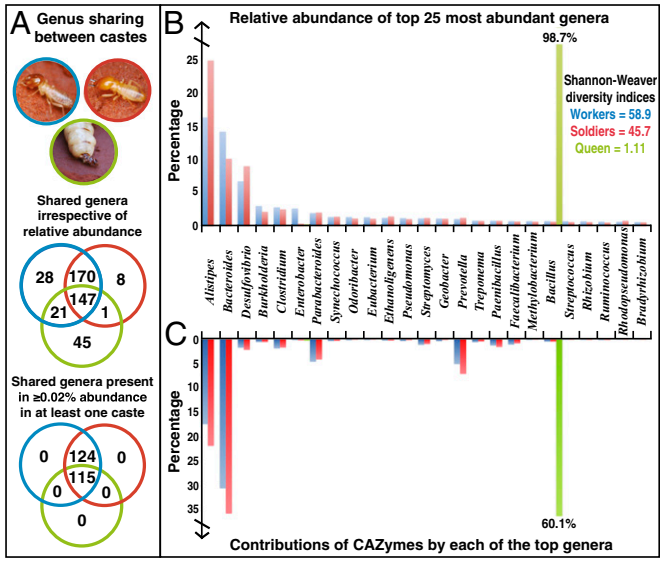

Fig. 4. Diversity, distribution, and CAZy potential of gut microbiotas from workers (blue), soldiers (red), and the queen (green). (*A*) Venn diagrams of the number of genera shared between the three gut metagenomes, identified using a combination of PhymmBL and BLASTn (details in *Methods* and *SI Appendix*). *Upper* diagram used all genera irrespective of their relative abundance within gut communities, whereas *Lower* diagram represents a similar analysis using only genera for which at least one of the castes had ≥0.02% relative abundance, showing that none of the hits unique to only one or two castes were abundant. (*B*) The percentage of paired reads for each of the 25 most abundant bacterial genera, comprising a major portion of the total number of paired reads in workers (65.4%), soldiers (68.1%), and the queen (99.1%). Workers and soldiers shared the dominant genera *Alistipes*, *Bacteroides*, *Desulfovibrio*, *Burkholderia*, and *Clostridium* and had relatively even distributions of reads across genera, as illustrated by similar Shannon–Weaver diversity indexes. In contrast, the queen microbiota was skewed toward a dominant genus (*Bacillus*), resulting in a diversity index of only 1.11. (*C*) The percentage of CAZymes identified to originate from the 25 most abundant genera, corresponding to 68.6% of identified CAZymes in workers, 79.2% in soldiers, and 60.4% in the queen (Dataset S1, Table S33).

of magnitude fewer assembled sequences than obtained for worker and soldier guts (Fig. 1*D* and *SI Appendix*, Fig. S7).

The worker and soldier gut microbiotas exhibited large resemblance and both were dominated by the genera *Alistipes*, *Bacteroides*, *Desulfovibrio*, *Clostridium*, and *Burkholderia* (Fig. 4*B*). These genera collectively comprised 43.3% and 48.9% of sequence reads in workers and soldiers, respectively, and their dominance is consistent with findings from other fungus-growing termite gut studies (e.g., refs. 18 and 19). *Alistipes* is generally not abundant in other termites [mean 0.62% relative abundance across eight lower termite species and mean 0.28% across eight higher non-fungus–growing termite species (19), but is a large component of cockroach gut microbiotas, mean 11.2% across 15 species (19, 37)]. This suggests that the functional roles of macrotermitine gut microbes are more similar to those of the distantly related noneusocial sister clade of the termites than to those of the more closely related non-fungus–growing termites (19). This is consistent with the underrepresentation of, e.g., *Treponema* (<1% relative abundance in workers and soldiers) compared with non-fungus–growing termites (19, 35), where this genus has been suggested to be the source of the relatively abundant GH5 and GH94 enzymes in *Nasutitermes* (34, 35).

Uniformity of the Queen Gut Microbiome. The majority of bacterial genera were absent or grossly underrepresented in the queen gut (Fig. 4*B*). None of the three dominant genera in workers or soldiers (*Alistipes*, *Bacteroides*, or *Desulfovibrio*) were detected, whereas a single genus accounted for the vast majority (98.7%) of sequence reads. A total of 84.9% of these reads mapped to two *Bacillus* sp. genomes from isolates obtained from *M. natalensis* (38) (Fig. 4*B*). This striking contrast to worker and soldier gut microbiomes may imply that the queen gut microbiota undergoes substantial compositional change between colony founding, when she (or her cofounding king) is expected to inoculate the first worker guts, and later colony life when workers inoculate each other and queens have become massive egg-laying machines, possibly requiring a special royal diet.

The queen gut microbes appeared to encode mainly enzymes from GH13 (e.g., amylase, glucosidases, and pullanase), GH1 (e.g., glucosidases, mannosidases, and galatosidases), GH4 (e.g., glucosidases), and GH18 (e.g., chitinases) families (Dataset S1, Table S29), suggesting that she is fed fungus material and simple sugars by oral trophallaxis. The dominant queen gut microbe (*Bacillus* sp.) may contribute many of these enzymes (66.3%) (Fig. 4*C*) and the queen gut thus appears to be completely decoupled from the plant decomposition functions of the guts of other colony members. Minor soldiers had gut microbiomes that were functionally similar to those of workers, with *Alistipes* and *Bacteroides* as dominant GH-contributing bacterial genera (Fig. 4*C*), in contrast to earlier suggestions that they are fed by workers and do not contribute to plant biomass degradation.

Conclusion

Our findings shed significant new light on the fungus-growing termite symbiosis. Genomic data show that these eusocial insects mostly provide the mound and foraging infrastructure to manage two vastly different, spatially segregated mutualisms. The domestication of *Termitomyces* apparently allowed for an increase in carbohydrate decomposition capacity relative to that of other higher termites, facilitated by partial functional complementarity between the prime decomposition targets of *Termitomyces* and those of the worker and soldier gut microbiotas. This integrated the services of sterile helper castes and symbionts into a very high level of somatic organismality (39), while exempting the colony germ line from litter or comb digestion.

Methods

Sequencing and Analyses of the *M. natalensis* Genome. Nine libraries of different insert sizes were constructed and more than 130 Gb of paired sequence reads [typically 90–100 bp for short insert libraries (insert size <2 kb), and 49 bp for long insert libraries (insert size ≥2 kb)] were generated using the HiSeq 2000 Illumina platform from DNA from a single *M. natalensis* queen. To aid genome assembly and annotation, we extracted RNA from workers, soldiers, king, and queen from four colonies of the same species. After assembly with SOAPdenovo (40) all original reads were aligned to the genome sequence with SOAPaligner (41). Coverage was then estimated based on short-read alignments, and GC content was determined. A GC vs. depth scatter plot indicated no apparent GC bias. Finally, absence of contaminated sequences in the assembly was confirmed using BLASTp (42) and transposable elements were identified.

A final set of protein-coding genes was determined using homology-based annotation with GeneWise (43), de novo annotation with AUGUSTUS (44) and SNAP (45), and 4.17 Gb of transcriptome data (*SI Appendix*). SwissProt (46) annotations were assigned according to the best match of the alignments generated by BLASTp (42). InterproScan (47) was then used to annotate motifs and domains of translated proteins. Gene sequences were searched against SUPERFAMILY, Pfam, PRINTS, PROSITE, ProDom, Gene3D, PANTHER, and SMART in Interpro. Gene Ontology terms for each gene were obtained from the Interpro database and KEGG annotations (48) were done using the KAAS online server (49). We clustered genes from 13 insect genomes and *Caenorhabditis elegans* and used Treefam (50) to construct gene families, after which CAFÉ (51) was used to detect gene family expansions or contractions. Using BLASTp (42), we identified genes involved in immune defense, antimicrobial peptides, neuropeptides, protein hormones, and biogenic amines and their receptors (*SI Appendix*).

Sequencing and Analyses of the *Termitomyces* Genome. A homokaryotic strain of *Termitomyces* was obtained from a heterokaryon in a colony of *M. natalensis*, using a standard protoplasting procedure (52, 53). DNA was extracted from pure culture material, after which 12.5 Gb raw paired reads were generated for five insert libraries with HiSeq 2000, which allowed for de novo assembly of the 83.7-Mbp genome. Reads were aligned to the genome to obtain a coverage estimate, and the GC vs. depth scatter plot indicated no obvious GC bias. Transposable elements were identified as described above. A final gene set was obtained using GLEAN after homology-based annotation against seven fungal species, combining AUGUSTUS and SNAP de novo predictions from a training set of 500 randomly selected genes with complete ORFs from homology annotation to *Saccharomyces cerevisiae*. The quality of the assembly and gene annotation was assessed by aligning 1,382 ESTs of a *Termitomyces* transcriptome from *Macrotermes gilvus* (54). Functional annotation, gene family construction, and gene family expansions and contractions were performed as described above.

Sequencing and Analyses of Caste-Specific Metagenomes. Whole guts were dissected from 50 major workers, 50 minor soldiers, and the queen from a single *M. natalensis* colony under sterile conditions. After DNA extraction, paired-end libraries with insert size of 350 bp were constructed, and reads were mapped to the termite and *Termitomyces* assemblies to filter out eukaryotic reads before assembly with SOAPdenovo. Read use and depth of assembly were evaluated by mapping the clean reads from each gut to their respective assemblies. Gene predictions for bacteria and archaea in the three gut microbiotas were done using the combined GeneMark-P* and GeneMark.hmm-P with precomputed models based on 265 sequenced genomes from the National Center for Biotechnology Information (NCBI) (55). KEGG annotations were done using the KAAS online server and BLASTp (42) was used to determine COG annotations.

PhymmBL (56) and BLASTn (42) were used to classify assembled reads to genus level, using the NCBI database of complete and draft genomes of bacteria, archaea, fungi, and protozoa, in addition to 12 bacteria draft genomes (Dataset S1, Table S26). Rarefaction curves indicated that all metagenomes had been sufficiently sampled to recover the expected number of genera (*SI Appendix*, Fig. S7). The relative abundance of different operational taxonomic units present in the three metagenomes was estimated by counting the number of paired reads that were assigned to each bacterial genus (Dataset S1, Table S26 and Fig. 4). Shannon–Weaver indexes (57) were calculated to assess genus-level differences in abundance profile between castes.

Carbohydrate-Active Enzyme Analyses for Genomes and Metagenomes. Encoded proteins from *M. natalensis* and *Termitomyces* genomes were first compared with the full-length sequences of the CAZy database, using BLASTp (42). Subsequently, each protein with a hit was subjected to two methods: (*i*) a BLASTp search against a library built by cutting the full sequences in CAZy into their respective GH, PL, CE, GT, AA, and CBM domains and (*ii*) a HMMer (58) search using hidden Markov models built by aligning partial sequences corresponding to each CAZy family. A sequence was considered reliably assigned when it was placed in the same family with both methods. The metagenomes were analyzed using the FASTY routine of the FASTA package (59) against sequence libraries derived from the CAZy database. To assign genus-level bacterial origins

of CAZymes, we combined the results of metagenome classification and CAZy and BLAST analyses of identified genes (Dataset S1, Table S32).

Fungus Growth Profiling. We compared growth performance of *Termitomyces* and a series of free-living fungi on different carbohydrate substrates, using Serpula minimal medium (60) adjusted to pH 6.0 and containing 1.5% agar (Invitrogen; 30391-049). Carbon sources were added at concentrations as indicated at www.fung-growth.org and were evaluated after growth at 25 °C for 10 d (complete growth profiles at www.fung-growth.org).

ACKNOWLEDGMENTS. We thank T. Gruntjes, Z. Human, M. Kolarik, J. Ringelberg, and S. Taerum for help with field work; J. Mitchell and the Agricultural Research Council-Plant Protection Research Institute (ARC-PPRI) Rietondale Research Centre for permission to collect termites; the Oerlemans family for access to their property; A. Brune for comments on gut microbiota compositions; A. Visser for worker and soldier images in Figs. 3 and 4; and S. Taerum for the queen image in Fig. 4. This work was supported by a STENO grant from The Danish Council for Independent Research | Natural Sciences (to M.P.), Danish National Research Foundation Centre of Excellence Grant DNRF57 (to J.J.B.), and a Marie Curie International Incoming Fellowship (300837; to G.Z.).

1. McCutcheon JP, Moran NA (2012) Extreme genome reduction in symbiotic bacteria. *Nat Rev Microbiol* 10(1):13–26.
2. Aanen DK (2006) As you reap, so shall you sow: Coupling of harvesting and inoculating stabilizes the mutualism between termites and fungi. *Biol Lett* 2(2): 209–212.
3. Frank SA (1996) Host control of symbiont transmission: The separation of symbionts into germ and soma. *Am Nat* 148:1113–1124.
4. Schluter J, Foster KR (2012) The evolution of mutualism in gut microbiota via host epithelial selection. *PLoS Biol* 10(11):e1001424.
5. Yatsunenko T, et al. (2012) Human gut microbiome viewed across age and geography. *Nature* 486(7402):222–227.
6. Turnbaugh PJ, et al. (2009) The effect of diet on the human gut microbiome: A metagenomic analysis in humanized gnotobiotic mice. *Sci Transl Med* 1:6ra14.
7. Kinross JM, Darzi AW, Nicholson JK (2011) Gut microbiome-host interactions in health and disease. *Genome Med* 3(3):14.
8. Hess M, et al. (2011) Metagenomic discovery of biomass-degrading genes and genomes from cow rumen. *Science* 331(6016):463–467.
9. Hildebrand F, et al. (2012) A comparative analysis of the intestinal metagenomes present in guinea pigs (*Cavia porcellus*) and humans (*Homo sapiens*). *BMC Genomics* 13:514.
10. Boucias DG, et al. (2013) The hindgut lumen prokaryotic microbiota of the termite *Reticulitermes flavipes* and its responses to dietary lignocellulose composition. *Mol Ecol* 22(7):1836–1853.
11. Liévin-Le Moal V, Servin AL (2006) The front line of enteric host defense against unwelcome intrusion of harmful microorganisms: Mucins, antimicrobial peptides, and microbiota. *Clin Microbiol Rev* 19(2):315–337.
12. Suh SO, McHugh JV, Pollock DD, Blackwell M (2005) The beetle gut: A hyperdiverse source of novel yeasts. *Mycol Res* 109(Pt 3):261–265.
13. Bignell DE, Roisin Y, Lo N, eds (2011) *Biology of Termites: A Modern Synthesis* (Springer, Dordrecht, The Netherlands), 2nd Ed, pp 211–253.
14. Wood TG, Sands WA (1978) The role of termites in ecosystems. *Production Ecology of Ants and Termites*, ed Brian MV (Cambridge Univ Press, Cambridge, UK), pp 245–292.
15. Aanen DK, et al. (2002) The evolution of fungus-growing termites and their mutualistic fungal symbionts. *Proc Natl Acad Sci USA* 99(23):14887–14892.
16. Aanen DK, et al. (2009) High symbiont relatedness stabilizes mutualistic cooperation in fungus-growing termites. *Science* 326(5956):1103–1106.
17. Nobre T, Aanen DK (2012) Fungiculture or termite husbandry? The ruminant hypothesis. *Insects* 3:307–323.
18. Hongoh Y, et al. (2006) Intracolony variation of bacterial gut microbiota among castes and ages in the fungus-growing termite *Macrotermes gilvus*. *Mol Ecol* 15(2):505–516.
19. 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(7):2261–2269.
20. Leuthold RH, Badertscher S, Imboden H (1989) The inoculation of newly formed fungus comb with *Termitomyces* in *Macrotermes* colonies (Isoptera, Macrotermitinae). *Insectes Soc* 36:328–338.
21. Poulsen M, et al. (2014) Complementary symbiont contributions to plant decomposition in a fungus-farming termite. *GigaScience*. Available at http://dx.doi.org/10.5524/100055, 100056, 1000557, and 100058.
22. Cantarel BL, et al. (2009) The Carbohydrate-Active EnZymes database (CAZy): An expert resource for glycogenomics. *Nucleic Acids Res* 37(Database issue):D233–D238.
23. Cantarel BL, Lombard V, Henrissat B (2012) Complex carbohydrate utilization by the healthy human microbiome. *PLoS ONE* 7(6):e28742.
24. Watanabe H, Tokuda G (2010) Cellulolytic systems in insects. *Annu Rev Entomol* 55: 609–632.
25. Terrapon N, et al. (2014) Molecular traces of alternative social organization in a termite genome. *Nat Commun* 5:3636.
26. Barrett AJ (1994) Classification of peptidases. *Methods Enzymol* 244:1–15.
27. Nygaard S, et al. (2011) The genome of the leaf-cutting ant *Acromyrmex echinatior* suggests key adaptations to advanced social life and fungus farming. *Genome Res* 21(8):1339–1348.
28. Kavanagh KL, Jörnvall H, Persson B, Oppermann U (2008) Medium- and short-chain dehydrogenase/reductase gene and protein families: The SDR superfamily: Functional and structural diversity within a family of metabolic and regulatory enzymes. *Cell Mol Life Sci* 65(24):3895–3906.
29. Zhang X, Leadbetter JR (2012) Evidence for cascades of perturbation and adaptation in the metabolic genes of higher termite gut symbionts. *MBio* 3(4):e00223–e12.
30. Li J, Cai S, Luo Y, Dong X (2011) Three feruloyl esterases in *Cellulosilyticum ruminicola H1* act synergistically to hydrolyze esterified polysaccharides. *Appl Environ Microbiol* 77(17):6141–6147.
31. Martinez D, et al. (2008) Genome sequencing and analysis of the biomass-degrading fungus *Trichoderma reesei* (syn. *Hypocrea jecorina*). *Nat Biotechnol* 26(5):553–560.
32. Yang F, Xu B, Li J, Huang Z (2012) Transcriptome analysis of *Termitomyces albuminosus* reveals the biodegradation of lignocellulose. *Wei Sheng Wu Xue Bao* 52(4):466–477.
33. Hyodo F, et al. (2003) Differential role of symbiotic fungi in lignin degradation and food provision for fungus-growing termites (Macrotermitinae: Isoptera). *Funct Ecol* 17(2): 186–193.
34. He S, et al. (2013) Comparative metagenomic and metatranscriptomic analysis of hindgut paunch microbiota in wood- and dung-feeding higher termites. *PLoS ONE* 8(4):e61126.
35. Warnecke F, et al. (2007) Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. *Nature* 450(7169):560–565.
36. Liu N, et al. (2013) Metagenomic insights into metabolic capacities of the gut microbiota in a fungus-cultivating termite (*Odontotermes yunnanensis*). *PLoS ONE* 8(7):e69184.
37. Schauer C, Thompson CL, Brune A (2012) The bacterial community in the gut of the cockroach *Shelfordella lateralis* reflects the close evolutionary relatedness of cockroaches and termites. *Appl Environ Microbiol* 78(8):2758–2767.
38. Um S, Fraimout A, Sapountzis P, Oh D-C, Poulsen M (2013) The fungus-growing termite *Macrotermes natalensis* harbors bacillaene-producing *Bacillus* sp. that inhibit potentially antagonistic fungi. *Sci Rep* 3:3250.
39. Queller DC, Strassmann JE (2009) Beyond society: The evolution of organismality. *Philos Trans R Soc Lond B Biol Sci* 364(1533):3143–3155.
40. Li R, et al. (2010) De novo assembly of human genomes with massively parallel short read sequencing. *Genome Res* 20(2):265–272.
41. Li R, et al. (2009) SOAP2: An improved ultrafast tool for short read alignment. *Bioinformatics* 25(15):1966–1967.
42. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215(3):403–410.
43. Birney E, Clamp M, Durbin R (2004) GeneWise and Genomewise. *Genome Res* 14(5): 988–995.
44. Stanke M, Waack S (2003) Gene prediction with a hidden Markov model and a new intron submodel. *Bioinformatics* 19(Suppl 2):ii215–ii225.
45. Korf I (2004) Gene finding in novel genomes. *BMC Bioinformatics* 5:59.
46. Boeckmann B, et al. (2003) The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. *Nucleic Acids Res* 31(1):365–370.
47. Zdobnov EM, Apweiler R (2001) InterProScan—an integration platform for the signature-recognition methods in InterPro. *Bioinformatics* 17(9):847–848.
48. Kanehisa M, Goto S (2000) KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 28(1):27–30.
49. Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M (2007) KAAS: An automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res* 35(Web Server issue):W182–W185.
50. Li H, et al. (2006) TreeFam: A curated database of phylogenetic trees of animal gene families. *Nucleic Acids Res* 34(Database issue):D572–D580.
51. De Bie T, Cristianini N, Demuth JP, Hahn MW (2006) CAFE: A computational tool for the study of gene family evolution. *Bioinformatics* 22(10):1269–1271.
52. Sonnenberg AS, Wessels JG, van Griensven LJ (1988) An efficient protoplasting / regeneration system for *Agaricus bisporus* and *Agaricus bitorquis*. *Curr Microbiol* 17:285–291.
53. Turgeon BG, Condon B, Liu J, Zhang N (2010) Protoplast transformation of filamentous fungi. *Methods Mol Biol* 638:3–19.
54. Johjima T, Taprab Y, Noparatnaraporn N, Kudo T, Ohkuma M (2006) Large-scale identification of transcripts expressed in a symbiotic fungus (*Termitomyces*) during plant biomass degradation. *Appl Microbiol Biotechnol* 73(1):195–203.
55. Lukashin AV, Borodovsky M (1998) GeneMark.hmm: New solutions for gene finding. *Nucleic Acids Res* 26(4):1107–1115.
56. Brady A, Salzberg SL (2009) Phymm and PhymmBL: Metagenomic phylogenetic classification with interpolated Markov models. *Nat Methods* 6(9):673–676.
57. Shannon CE, Weaver W (1949) *Mathematical Theory of Communication* (Univ of Illinois Press, Urbana, IL).
58. Eddy SR (1998) Profile hidden Markov models. *Bioinformatics* 14(9):755–763.
59. Pearson WR, Wood T, Zhang Z, Miller W (1997) Comparison of DNA sequences with protein sequences. *Genomics* 46(1):24–36.
60. Eastwood DC, et al. (2011) The plant cell wall-decomposing machinery underlies the functional diversity of forest fungi. *Science* 333(6043):762–765.
61. Ruelle JE, Coaton WGH, Sheasby JL (1975) National survey of the Isoptera of southern Africa. 8. The genus *Macrotermes* Holmgren (Termitidae: Macrotermitinae). *Cimbebasia* 3:73–94.
62. Zhao Z, Liu H, Wang C, Xu JR (2013) Comparative analysis of fungal genomes reveals different plant cell wall degrading capacity in fungi. *BMC Genomics* 14:274.