Comparative Genomic and Metabolomic Analysis of Termitomyces Species Provides Insights into the Terpenome of the Fungal Cultivar and the Characteristic Odor of the Fungus Garden of Macrotermes natalensis Termites

Kreuzenbeck, Nina B.; Seibel, Elena; Schwitalla, Jan W.; Fricke, Janis; Conlon, Benjamin H.; Schmidt, Suzanne; Hammerbacher, Almuth; Köllner, Tobias G.; Poulsen, Michael; Hoffmeister, Dirk; Beemelmanns, Christine; Pupo, Monica T.

Published in:
mSystems

DOI:
10.1128/msystems.01214-21

Publication date:
2022

Document version
Publisher's PDF, also known as Version of record

Document license:
CC BY

Citation for published version (APA):
Comparative Genomic and Metabolomic Analysis of *Termitomyces* Species Provides Insights into the Terpenome of the Fungal Cultivar and the Characteristic Odor of the Fungus Garden of *Macrotermes natalensis* Termites

Nina B. Kreuzenbeck, Elena Seibel, Jan W. Schwitalla, Janis Fricke, Benjamin H. Conlon, Suzanne Schmidt, Almuth Hammerbacher, Tobias G. Köllner, Michael Poulsen, Dirk Hoffmeister, Christine Beemelmanns

*Group of Chemical Biology of Microbe-Host Interactions, Leibniz Institute for Natural Product Research and Infection Biology–Hans-Knöll Institute (HKI), Jena, Germany*

*Section for Ecology and Evolution, Department of Biology, University of Copenhagen, Copenhagen, Denmark*

*Department of Zoology and Entomology, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria, South Africa*

*Max Planck Institute for Chemical Ecology, Department of Biochemistry, Jena, Germany*

*Associated Group of Pharmaceutical Microbiology, Leibniz Institute for Natural Product Research and Infection Biology–Hans-Knöll Institute (HKI), Jena, Germany*

*Friedrich-Schiller-Universität Jena, Department of Pharmaceutical Biology at the Hans-Knöll Institute, Jena, Germany*

**ABSTRACT** Macrotremitinae termites have domesticated fungi of the genus *Termitomyces* as food for their colony, analogously to human farmers growing crops. Termites propagate the fungus by continuously blending foraged and predigested plant material with fungal mycelium and spores (fungus comb) within designated subterranean chambers. To test the hypothesis that the obligate fungal symbiont emits specific volatiles (odor) to orchestrate its life cycle and symbiotic relations, we determined the typical volatile emission of fungus comb biomass and *Termitomyces* nodules, revealing α-pinene, camphene, and D-limonene as the most abundant terpenes. Genome mining of *Termitomyces* followed by gene expression studies and phylogenetic analysis of putative enzymes related to secondary metabolite production encoded by the genomes uncovered a conserved and specific biosynthetic repertoire across strains. Finally, we proved by heterologous expression and *in vitro* enzymatic assays that a highly expressed gene sequence encodes a rare bifunctional mono-/sesquiterpene cyclase able to produce the abundant comb volatiles camphene and D-limonene.

**IMPORTANCE** The symbiosis between macrotremitinae termites and *Termitomyces* is obligate for both partners and is one of the most important contributors to biomass conversion in the Old World tropic’s ecosystems. To date, research efforts have dominantly focused on acquiring a better understanding of the degradative capabilities of *Termitomyces* to sustain the obligate nutritional symbiosis, but our knowledge of the small-molecule repertoire of the fungal cultivar mediating interspecies and interkingdom interactions has remained fragmented. Our omics-driven chemical, genomic, and phylogenetic study provides new insights into the volatilome and biosynthetic capabilities of the evolutionarily conserved fungal genus *Termitomyces*, which allows matching metabolites to genes and enzymes and, thus, opens a new source of unique and rare enzymatic transformations.

**KEYWORDS** *Termitomyces*, secondary metabolism, symbiosis, terpenes

*Macrotermes natalensis*, a fungus-growing termite species, has undergone an obligate mutualistic interaction with species of the basidiomycete genus *Termitomyces*. This fungus is maintained as a monoculture and is used as the primary food source by...
the termite colony (1, 2). Fungus-farming *M. natalensis* colonies have developed high levels of social complexity with polymorphism and division of labor across castes, similar to other nonfarming insect species. While old workers leave the nest to forage for plant material in the surrounding environment, young workers propagate the fungal monoculture of *Termitomyces* on predigested plant material crafted into comb-like structures (comb material) belowground (Fig. 1) (3–7). *Termitomyces* species mostly grow as fungal mycelium on the predigested plant biomass (fungus comb), thereby decomposing the lignocellulose-rich biomass, and occasionally differentiate into fungal nodules (1, 7). While the ecology and biology of social termites is well described, our current understandings of the molecular mechanisms that drive and orchestrate the sophisticated organization of the agricultural insect society remain fragmented (8). Communication by volatiles represents one possible mechanism that might be used by *Termitomyces* to coordinate its obligate symbiotic relation with the termite colony and its fungal life cycle, as volatiles can quickly diffuse within the subterranean fungus chambers (5). Similar to bacteria (9), fungi are known to emit a complex mixture of volatile organic compounds (VOCs), including alcohols, hydrocarbons, aldehydes, ketones, acids, esters, and terpenes that contribute to their characteristic odor (10, 11) and can be linked to their phylogenetic placement, sexual life cycle (12, 13), and parasitic, saprophytic, or symbiotic lifestyles (14–16). Only recently, first reports on the volatilome of *Termitomyces heimii* (17) and *Termitomyces* sp. strain J132 (18, 19) hinted toward the enormous capacity of *Termitomyces* species to produce a diverse set of volatiles, including terpenes. To test the hypothesis that VOCs in general, and fungal VOCs in particular, are produced within the comb environment and could serve as communication signals within the agricultural society, we analyzed the volatilome of fungal comb material and fungal nodules of *Termitomyces* species from *M. natalensis* colonies using gas chromatography-mass spectrometry (GC-MS). To identify the biosynthetic origin of the detected VOCs, we then analyzed the genomes of *Termitomyces* strains and found that members of this genus harbor a distinct set of secondary metabolite gene

![FIG 1](A) Termite mound of a *Macrotermes natalensis* colony; (B) belowground garden chamber filled with fungus comb biomass; (C) closeup picture of fungus comb material pervaded with fungal biomass and pointy nodules (scale, 10 mm); and (D) schematic sketch of fungus comb emitting volatiles including monoterpenes.
clusters, including an extraordinary number of genes annotated as terpene cyclases, many of which were differentially transcribed in different biological samples. Finally, we verified the enzymatic origin for the most abundant terpenes in vitro, thereby linking the obtained metabolomic and transcriptomic data sets.

RESULTS AND DISCUSSION

Volatilome of the fungus comb. In a first experimental setup, we collected different types of biosamples (fungus comb, fungal nodules, and termite workers, with soil and air samples serving as controls) from six different *M. natalensis* colonies (see Table S1 at https://doi.org/10.6084/m9.figshare.16702471). Volatiles were collected from pre-weighed biosamples (*n* = 3) using solid-phase microextraction (SPME) and then analyzed by GC-MS. Obtained data sets were dereplicated using the National Institute of Standards Mass Spectral Library (NIST 2017), and peak intensities (with a match quality of at least 90%) were quantified relative to the measured sample weight and averaged signal intensities (see Fig. S1 to S3 and Tables S2 and S3 at https://doi.org/10.6084/m9.figshare.16702471).

As depicted in Fig. 2, fungus comb samples emitted a diverse set of volatiles, including alkanes, alcohols, aliphatic aldehydes (e.g., undecane, 3-octanol, hexanal), ketones (e.g., undecanone, 3-octanone), furanes, terpenes, and aromatic metabolites. In comparison, fungal nodules emitted a less diverse volatile blend containing, e.g., isoamyl alcohol, oxidized octane derivatives (20), butanoic acid derivatives, terpenes, and, in particular, a specific set of aromatic volatiles (benzyl acetic acid methyl ester and phenylethyl alcohol), while control samples (air and soil) emitted only a few detectable volatiles, most dominantly β-pinene and naphthalene (see Tables S4 to S71 at https://doi.org/10.6084/m9.figshare.16702471).

It was particularly noteworthy that the emission of several monoterpenes was detectable from the volatile blend released by the fungus comb as well as nodules. While α-limonon and camphene were emitted from both samples, α-pinene was predominately detected from comb samples. Interestingly, β-pinene and 3-carene, common plant metabolites of pine trees, were detected in soil and air samples and might have originated from the predigested plant material of the comb or surrounding trees (21). However, most of the detected chemical features from comb and nodule samples are known fungal infochemicals emitted during fungal growth and differentiation (22), some of which were also detected in previous volatile studies of *Termitomyces* and, thus, may contribute to the characteristic scent of the fungal cultivar (18, 23). While isoamyl alcohol is known to induce morphological changes in yeast (24, 25), phenylethyl alcohol serves as an antifungal agent (26) and might have additional morphogenic functions during the fungal life cycle, as shown in the yeast *Candida albicans* (27). Both α-pinene and β-limonene are known microbial volatiles with antimicrobial properties (9, 16, 28) and serve as communication signals in other termite species (29, 30), while camphene is reported to have antioxidant activities (31). Only a few of the detected chemical features were deduced to be of anthropogenic or unknown origin (e.g., 2-ethylhexyl salicylate, cyclohexane, and trichloromethane) (29). To corroborate these findings and to pinpoint major termite volatiles, we also measured VOCs of 30 major workers (Tables S71 to S73). Here, several known sesquiterpenes were detectable, with β-gurjunene, gymnomintrine, and aristolene as the most abundant features, as well as smaller amounts of four terpenes, *trans*-α-bergamotene, α-pinene, α-barbatene, and β-chamigrene that were also produced by, e.g., *Termitomyces* sp. strain J132 and could be due to insects feeding on fungus biomass (see Tables S4 to S71 and Fig. S4 at https://doi.org/10.6084/m9.figshare.16702471).

In silico analysis of biosynthetic gene clusters in *Termitomyces* sp. To correlate detected volatiles to the biosynthetic capacity of *Termitomyces*, we analyzed the draft genomes of eight *Termitomyces* strains using the platform fungiSMASH (v6.0) and web service NRPSpredictor2 (see Table S73 at https://doi.org/10.6084/m9.figshare.16702471) (32, 33). Overall, each fungal genome encoded more than 20 detectable biosynthetic features related to secondary metabolism, including predominately fungal terpene cyclases...
TCs), one nonreducing iterative type I polyketide synthase (NR-PKS) with conserved domain architecture (SAT-KS-AT-PT-ACP-ACP-TE), and a nonribosomal peptide synthetase (NRPS) with a conserved domain architecture (A1-T1-C1-T2-C2-T3-C3) and an adjacent putative cytochrome P450 monoxygenase (34). Phylogenetic analysis of the NR-PKS sequences revealed their close relationship to known orsellinic acid-producing synthases described from other basidiomycetes (35, 36), while an analysis of the detected NRPS sequences indicated their close relationship to one of the most abundant type VI basidiomycete siderophore synthetases that is assumed to produce the trimeric siderophore basidioferrin, derived from N$_5$-acylated N$_5$-hydroxy-L-ornithine (L-AHO) (see Table S73 and Fig. S5 and S6 at https://doi.org/10.6084/m9.figshare.16702471) (37). Thus, we deduced that both PKS and NRPS are unlikely to be accountable for any of the observed chemical features in the volatilome.

Consequently, we focused on the analysis and categorization of identified

<table>
<thead>
<tr>
<th></th>
<th>comb</th>
<th>nodules</th>
<th>Termimitomyces sp.</th>
<th>soil</th>
</tr>
</thead>
<tbody>
<tr>
<td>Saturated alkenes, alkenes, alcohols, aldehydes and ketones</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonane</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Undecane</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-Octene</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-Nonene</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-Decene</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-Methyl-2-pentene</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-Ethyl-2-methyl-1,3-hexadiene</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hexanoyl alcohol</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-Octanol</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-Octen-3-ol</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-Octanaline</td>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>2-Undecane</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hexanal</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Octanal</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-Hexenal</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-Heptenal</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-Nonenal</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-Decenal</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Butanoic acid ethyl ester</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-Methyl butanoic acid ethyl ester</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hexanoic acid methyl ester</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**FIG 2** Heatmap of assigned volatiles detected from comb, nodules, and soil from termitite mounds (signal intensities based on averaged peak area (n = 3) per g of sample) and axenic Termimitomyces sp. strain T153 and J132 cultures (footnote a, others, combined results from literature reported volatile studies; see references 17, 18, and 23).
putative fungal terpene cyclase (TC) sequences based on the following general classification: (i) monoterpene cyclases (MTCs), which cyclize geranylpyrophosphate (GPP), yielding monoterpenes; (ii) sesquiterpene cyclases (STCs), which produce sesquiterpenes from farnesyl pyrophosphate (FPP); (iii) diterpene cyclases (DTCs), which use geranylgeranylpyrophosphate for product formation; and (iv) triterpene cyclases (Tri-TCs), also known as oxidosqualenecyclases (OSC)/lanosterol synthases, which catalyze cyclization of oxidosqualene into triterpenes.

Both MTCs and STCs are known to initiate terpene formation by a metal-dependent ionization and diphosphate release mechanism and are identified by the characteristic active-site motifs (DDXXD and NSE) that stabilize and guide the reaction pathway (9, 10). However, in silico predictions of fungal MTCs are still ambiguous due to their high similarities to fungal STC (similar active motifs, catalytic pocket, and sequence length) and the low numbers of biochemical characterized MTCs in basidiomycetes (38). In contrast, numerous fungal STCs have been characterized in basidiomycetes, and comparative studies uncovered a correlation between their phylogenetic placement and cyclization mechanism (clades I to V) (Fig. 3) (16, 39–41). In short, STCs belonging to clade I have been found to cyclize FPP via the formation of a germacadienyl cation, while STCs belonging to clade II induce a 1,10-cyclization mechanism of nerolidyl diphosphate (NPP), resulting either in bicyclogermacrene (clade IIa) or the germacadienyl cation (clade IIb), which undergoes further transformations. In contrast, enzymes belonging to clade III generally produce humulyl cation intermediates via 1,11-cyclization of FPP. Finally, enzymes of clade IV catalyze a 1,6- or 1,7-cyclization of NPP, which leads to an intermediate bisabolyl or cycloheptenyl cation, while clade V putatively includes TCs that favor only a 1,6-cyclization mechanism.

To the best of our knowledge, only 10 DTCs have been biochemically characterized so far from basidiomycetes and were found to be either mono- or bifunctional enzymes (16). While monofunctional DTCs cyclize GGPP via a type I ionization-dependent mechanism centered at the N terminus, triggering the second cyclization step by diphosphate removal (ionization-triggered reaction) (16, 39), bifunctional enzymes use both class I and class II mechanisms to form diterpenes. Here, class II activity is located at the C-terminal site and catalyzes the cyclization of GGPP by protonation of the substrate. Only recently, a third class of fungal DTC was discovered that contains proteins from the UbiA family, which can have class I diterpene cyclase activity besides prenylation activity (42, 43). Tri-TCs, which are key enzymes in ergosterol biosynthesis, carry a QW motif that determines its cyclization specificity in the transformation of squalene epoxide via a type II cyclization mechanism (9, 10). So far only four Tri-TC candidates involved in the biosynthesis of triterpenoid natural products have been characterized from different fungal species (44).
Based on these considerations, we extended our fungiSMASH survey by dedicated HMM profile searches of predicted protein sequences of TCs (45). Overall, we identified, on average, 22 distinct putative MTC/STC protein sequences per strain, which were predicted from the open reading frames in Termitomyces genomes, while slightly lower numbers of protein sequences for putative DTCs and Tri-TCs were identified (Table 1 and Tables S74 to S82 at https://doi.org/10.6084/m9.figshare.16702471). The number of predicted TCs varied across Termitomyces species, which is likely a result of various sequencing qualities and annotation approaches as well as lacking transcriptomic data, which impairs the prediction of putative proteins. Despite these ambiguities, it was important to note that genomes of free-living basidiomycetes generally encoded lower numbers of terpene cyclases, a finding suggestive of the importance of terpene-based communication mechanisms in a termite-associated lifestyle.

**Phylogenetic analysis and prediction of chemical scaffolds.** To more precisely predict the putative cyclization mechanism and product scope of the identified TCs, we performed a phylogenetic analysis of all candidate sequences identified from the nine Termitomyces spp., as depicted in Fig. 4.

Eight terpene cyclases (TTCs; TTC21 to TTC28) showed high homologies to sequences belonging to STCs of clade I, with TTC21 to TTC26 appearing to build a Termitomyces-specific subclade. TTC28 from Termitomyces sp. strain J132 (KNZ77998; previously named STC15) (19) was biochemically characterized in previous studies and found to cyclize FPP in a 1,10-cyclization fashion via the (E,E)-germacradienyl cation to (1)-germacrene D-4-ol. We also identified two groups of TCs belonging to clade II (TTC1 and TTC13), with TTC13 being more closely related to members of clade IV, while nine TTCs (TTC2 to TTC10) were categorized as representatives of clade III and again appeared to belong to a Termitomyces-specific subclade. Earlier studies revealed that one representative of TTC13 (previously named STC9, from Termitomyces strain J132, KNZ74377) acts as a (−)-γ-cadinene synthase by cyclizing NPP at positions 1 and 10, a typical cyclization mode for clade II terpene cyclases (38). TTC13 is phylogenetically related to CpSTS18 (Clitopilus pseudopinus) and ShSTS5 (Stereum hirsutum), both known producers of γ-cadinene. Seven TTCs (TTC14 to TTC20) were located in clade IV, and sequence alignments of all representatives showed that the NSE motif is well conserved between Cop6, Omp9, Omp10, and all Termitomyces sequences, while the aspartate-rich motif (DDXXD/

### TABLE 1 Numbers of identified putative TC protein sequences in Termitomyces species and other model basidiomycetes

<table>
<thead>
<tr>
<th>Termitomyces strain</th>
<th>Host</th>
<th>No. of sequences</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>M/STC$^a$</td>
<td>DTC</td>
<td>Tri-TC</td>
</tr>
<tr>
<td>T153</td>
<td><em>Macrotermes natalensis</em></td>
<td>36</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>T112</td>
<td><em>Macrotermes natalensis</em></td>
<td>36</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>J132</td>
<td><em>Macrotermes natalensis</em></td>
<td>37</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>GCA_003316525</td>
<td>Not defined$^d$</td>
<td>12</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>GCA_003313785</td>
<td>Not defined$^d$</td>
<td>20</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>GCA_003313675</td>
<td>Not defined$^d$</td>
<td>8</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>GCA_003313075</td>
<td>Not defined$^d$</td>
<td>16</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>GCA_003313055</td>
<td>Not defined$^d$</td>
<td>20</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>GCA_001972325</td>
<td><em>Macrotermes gilvus</em></td>
<td>13</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Avg</td>
<td></td>
<td>22</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Other$^b$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>abp</td>
<td></td>
<td>10</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>lbc</td>
<td></td>
<td>10</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>mpi</td>
<td></td>
<td>18</td>
<td>13</td>
<td>4</td>
</tr>
<tr>
<td>scm</td>
<td></td>
<td>4</td>
<td>19</td>
<td>2</td>
</tr>
</tbody>
</table>

$^a$Due to ambiguous in silico analyses, putative MTC and STC counts are combined.


Figure 4: Phylogenetic tree of characterized fungal STC and putative *Termitomyces* STC protein sequences showing similarities across species (TTCs), which were grouped in clades according to their predicted cyclization mechanism (unrooted tree based on maximum likelihood method) with branch values indicating bootstrap support of 1,000 pseudoreplicates.
metabolites (e.g., ganoderic acids) (10, 48). to produce lanosterol (46), the precursor for ergosterol (47) and other triterpene-based seq) data obtained from cultured ural product biosynthetic-related gene sequences in transcriptome sequencing (RNA- and S8 athttps://doi.org/10.6084/m9. and old fungus comb and nodules on which young workers feed (Fig. 6 and Fig. S7 (14, 15, 16, 18, and 19) as well as two minor products [(E)-β-farnesene (11) and β-sesquiphellandrene (17)] (Fig. 7B and Fig. S11 at https://doi.org/10.6084/m9 .gshare.16702471). The six identified putative DTC genes were grouped as TTC33 to TTC34 (Fig. 5A) and classified according to the classification in Li et al. (42, 43). We tentatively categorized the first group, TTC33, as UbiA-type DTCs (clade III), while TTC34 was assigned to an unclassified phylogenetic branch. Lastly, we also identified seven homologous Tri- TC candidate sequences based on our similarity searches (Fig. 5B), which were grouped as TTC35 and found to be closely related to OSC from Ganoderma lucidum and known to produce lanosterol (46), the precursor for ergosterol (47) and other triterpene-based metabolites (e.g., ganoderic acids) (10, 48).

RNA-seq analysis. Subsequently, we analyzed the relative expression levels of natural product biosynthetic-related gene sequences in transcriptome sequencing (RNA-seq) data obtained from cultured Termitomyces sp. strains 153 and J132 as well as fresh and old fungus comb and nodules on which young workers feed (Fig. 6 and Fig. S7 and S8 at https://doi.org/10.6084/m9.figshare.16702471) (49). Intriguingly, a differential expression pattern of TTCs was observed, with some candidate sequences mostly expressed in comb and lab cultures but not in nodules, while others were expressed in all samples. Intriguingly, one candidate sequence (TTC15) was predominately expressed in fungal nodules, while only moderate to low expression levels were detectable within the comb and agar cultures. In contrast, NRPS- and NR-PKS-related gene sequences were expressed with similar abundance under all conditions and confirmed by RT-PCR of a fungal monoculture (Fig. S9 at https://doi.org/10.6084/m9 .figshare.16702471).

Biochemical analysis of TTC15-T153. We then hypothesized that some of the terpenes detected in the headspace of comb and nodules (e.g., limonene/camphene, Fig. 2) were produced from proteins encoded by the highly expressed TTC15 sequence (Fig. 6). To test our hypothesis, we heterologously expressed a homologous gene sequence of TTC15 identified in Termitomyces sp. strain T153 (TTC15-T153). The coding sequence was synthesized from mRNA predictions, cloned in a pET28a vector, and heterologously expressed in Escherichia coli BL21 cells. The hexahistidine fusion protein of enzyme TTC15-T153 (36.5 kDa; native protein, 34.3 kDa) was purified on a Ni2+-NTA-affinity column with a yield of 14 mg protein per 100 ml expression culture (Fig. S10 at https://doi.org/10.6084/m9.figshare.16702471). The purified protein was incubated at 30°C for 2 h with GPP, FPP, or GGPP in the presence of Mg2+ as a cofactor. Following the reaction, the produced compounds were extracted with hexane and analyzed by GC-MS, and detected products were tentatively assigned using the NIST 2017 database. While incubation with GGPP resulted in no detectable product formation, transformation of GPP by TTC15-T153 resulted in the formation of several major monoterpenes, such as camphene (2), β-myrcene (4), α-limonene (6), and linalool (8) (Fig. 7A and Table S6 and Fig. S10). Additionally, α-pinene (1), β-pinene (3), α-phellandrene (5), terpinolene/4-carene (7), 3-methyl camphenilanol (9), and α-terpineol (10) were formed as minor products (see Fig. S11 at https://doi.org/10.6084/m9.figshare .16702471).

In contrast, incubation of TTC15-T153 with FPP yielded eight major terpenes, including β-epi-β-santalene (12), α-bergamotenes (13a/b), and a mixture of bisabolene isomers (14, 15, 16, 18, and 19) as well as two minor products [(E)-β-farnesene (11) and β-sesquiphellandrene (17)] (Fig. 7B and Fig. S11 and Tables S87 and S88 at https://doi.org/10.6084/ m9.figshare.16702471). The observed bifunctional activity of TTC15-T153 was intriguing,
FIG 5 (A and B) Phylogenetic analysis of DTCs (A) and Tri-TCs (B) grouped in clades according to similarity (unrooted tree [based on maximum likelihood method] with branch values indicating bootstrap support of 1,000 pseudoreplicates).
as only one other enzyme from basidiomycetes (Agr4 from Agrocybe aegerita) has been characterized to convert GPP as well as FPP into cyclic terpenes (40).

The observed product spectrum of GPP conversion by TTC15-T153 can be explained by formation of the linalyl cation (Fig. 8A), which either serves as a precursor for the linear monoterpenes linalool (8) and myrcene (4) or is further converted into the α-terpinyl cation, the intermediate precursor of terpinolene, α-phellandrene, α-terpineol (10), 4-carene (7), and limonene (6) (50, 51). Once the α-terpinyl cation is formed, it could also be further enzymatically converted via a 3,7-ring closure to the bornyl cation, which, after Wagner-Meerwein rearrangement, deprotonation, or addition of water, yields camphene (2) or 3-camphenilanol (9) (52). Furthermore, enzymatic

**FIG 6** Heatmap of expression levels of putative natural product biosynthetic genes from *Termitomyces* species from *Macrotermes* colony Mn156 (49) under natural conditions (nODULES, fresh, and old comb) and plate culture of *Termitomyces* sp. strains J132 and T153 (green). Transcript abundances are depicted as log10 gene expression values, and color schemes were generated by viridis (version 0.5.1).
transformation of \(\alpha\)-terpinyl cation to the pinyl cation results in the formation of the observed \(\alpha\)– and \(\beta\)-pinene (1 and 3) (53).

In the presence of FPP, TTC15-T153 likely converts FPP via NPP into a bisabolyl cat-ion (Fig. 8B), which then results in the formation of a \(\beta\)-sesquiphellandrene and bisabolene blend (54). Starting from the bisabolyl cation, a 7,2-ring closure and deprotonation event can also lead to the formation of \((E)-\alpha\)-bergamotene, while \(\beta\)-santalene is formed after a sequence of rearrangements and a final deprotonation step (54–57).

Comparative analysis of TTC15-T153 and related sequences from clade TTC15 enzymes revealed that the first metal binding motif differs from the common DDXXD/E amino acid composition, which might cause the acceptance of both GPP and FPP as substrates by TTC15-T153, causing the formation of the different terpene blends.

**Conclusions.** In this study, we elaborated on the hypothesis that the fungal cultivar of *Macrotermitinae* termites emits specific volatiles (odor) that allows *Termitomyces* to orchestrate its complex life cycle within the fungal garden and its symbiotic relation with the termite host. To test the hypothesis, we combined volatile studies of fungal
FIG 8 Putative enzymatic conversion of GPP (A) and FPP (B) by terpene cyclase TTC15-T153.
nODULES AND COMB, IN WHICH THE FUNGAL CULTIVAR Termitomyces RESIDES, WITH PHYLOGENETIC ANALYSES, GENE EXPRESSION, AND MOLECULAR BIOLOGICAL STUDIES.

OUR FINDING THAT Termitomyces THRIVING IN FUNGUS COMB MATERIAL EMITS A SPECIFIC VOLATILOMEX a-pinene, camphene, and D-limonene among the most abundant terpenes (5–10) correlated with the identification of an above average repertoire of different types of terpene cyclases with conserved candidate sequences across Termitomyces species. While TCs from different Termitomyces strains showed high homologies to each other, they grouped into distinct phylogenetic branches compared to other TCs reported from other fungal genera. The detailed comparative analysis also uncovered that phylogenetic relatedness was not always correlated with the predicted cyclization mode, and more detailed biochemical studies on fungal TCs are required to solidify the correlations or to uncover the characteristics of their deviations. Based on RNA-seq data analysis, we verified that a congener of the highly expressed enzyme TTC15-T153 accepts GPP as well as FPP as substrates, causing the formation of more than 20 different terpenes, most of which were also detected in the volatile blends of collected biosamples and may contribute significantly to the characteristic odor of Termitomyces. From a chemical-ecological perspective, these findings were highly intriguing, as the two major GPP-derived products of TTC15-T153, D-limonene and camphene, were identified within the fungus comb volatilome and are known to be metabolites exchanged within symbiotic systems (14, 58). This study represents, to the best of our knowledge, the second characterized example of a bifunctional mono-/sesquiterpene cyclase in basidiomycetes. The high expression level of the coding sequence TTC15 in comb and nodules as well as natural abundance of these characteristic monoterpenes points toward a central role in termite symbiosis. Overall, our combined studies allowed us, for the first time, to link and verify produced metabolites within the complex community with their genetic basis, transcriptional pattern, and conserved biosynthetic origin and improve our current understanding of the chemical language orchestrating this complex farming symbiosis.

MATERIALS AND METHODS

Volatile analysis. For volatile sampling, different parts of M. natalensis nests from different areas of Pretoria (South Africa) were collected in preweighed amber glass vials directly in the field. Soil, air, and fungus comb samples were collected in 40-ml amber glass vials, and nodules were separated from the fungus comb using clean forceps and collected in 1-ml amber glass vials, both equipped with a cap containing a silicone septum. The vials were closed tightly and kept at 4°C until volatile extraction. Each sample was collected in triplicates, if not stated differently. Headspace sampling was performed using solid-phase microextraction (SPME). After penetrating the silicone septum of the vial cap using an SPME fiber holder, a conditioned SPME fiber coated with 100 μm polydimethylsiloxane (Supelco) was exposed to the headspace in the vial for 20 min at room temperature. Room temperature and sample weight were recorded each time before measuring. The SPME fiber was then directly injected into the inlet of an Agilent 7890B gas chromatograph coupled to a 7977 MSD quadrupole mass spectrometer. Thermal desorption was achieved at 250°C. Compound separation was achieved using a DB wax GC column (length, 20 m; inner diameter, 0.18 mm; and film thickness, 0.18 μm [J&W]). Separation was achieved with a 5-min hold at 50°C, followed by a linear temperature increase of 10°C/min to 250°C. The column was reconditioned with a 2-min hold at 250°C. Data analysis and peak integration were performed using the program MSD ChemStation (version F.01.03.2357). Metabolites were tentatively identified based on comparison of mass spectra and retention times to the National Institute of Standards mass spectral library (NIST 2017). After peak detection and integration, metabolites with a database match quality of at least 90% were taken for further analysis. Thereafter, the peak area was normalized to the sample weight for each metabolite (area per gram) and integration values used to generate a heat map of averaged values (see the supplemental material at https://doi.org/10.6084/m9.figshare.16702471).

Identification of putative NRPS and PKS and phylogenetic analysis. Identified enzymes were aligned to characterized fungal sequences from the same class in MEGAX by MUSCLE algorithm. Phylogenetic trees were generated with IQ TREE, calculating 1,000 bootstrap replicates (see the supplemental material at https://doi.org/10.6084/m9.figshare.16702471).

Identification of putative terpene cyclases and phylogenetic analysis. Putative terpene cyclases in Termitomyces species were identified by HMMsearch of HMMprofi les from characterized fungal terpene cyclases against predicted proteins from different Termitomyces species. Identified terpene cyclases that matched specific criteria for each class of terpene cyclases were aligned to characterized fungal sequences from the same class in MEGAX by MUSCLE algorithm. Phylogenetic trees were generated with IQ TREE, calculating 1,000 bootstrap replicates. Bioinformatic analyses were performed on the Galaxy EU webserver (59) (see the supplemental material at https://doi.org/10.6084/m9.figshare.16702471).
RNA-seq data acquisition and processing. Termitomyces RNA-seq data for fresh comb (SRR5944783), old comb (SRR5944781), and nodules (SRR5944782) from Macrotermes colony Mn156 were downloaded from the European Nucleotide Archive, and RNA-seq data were from cultured Termitomyces sp, strain T153 (PRJNAS29327) and J132 (PRJNA193471) (18). Identified terpene cyclase sequences from Termitomyces genomes T153 and J132 were prepared as separate references for Bowtie version 1.2.2 14 using rsemPrepareReference in RSEM version 1.3.1.15. RNA-seq reads were then aligned to each reference genome individually, and expression levels were calculated using rsem-calculate-expression in RSEM version 1.2.2.15 (see the supplemental material at https://doi.org/10.6084/m9.figshare.16702471).

Expression analysis. Expression levels calculated by RSEM were imported separately for each reference genome into R version 3.6.3 [60] using tximport (version 1.12.3) [61] and DESeq2 (version 1.24.0) [14] to identify differentially expressed genes. Expression analysis was performed by immobilized metal affinity chromatography for investigation of the enzyme function. This yielded the successful isolation of a 36.5-kDa protein. For assaying, the enzyme was incubated with GPP, (E,E)-FPP, and GGPP as a substrate and 10 mM Mg2+ as the cofactor. Hexane extracts from the reaction were directly investigated by GC-MS (see the supplemental material at https://doi.org/10.6084/m9.figshare.16702471).

Heterologous expression. The gene sequence of TTC15-153 was codon optimized for E. coli and synthesized in a pET28a (+) vector (Biocat). After transformation of E. coli BL21 cells with the plasmid pET28-TTC15, the protein was heterologously produced and purified by immobilized metal affinity chromatography for investigation of the enzyme function. This yielded the successful isolation of a 36.5-kDa protein. For assaying, the enzyme was incubated with GPP, (E,E)-FPP, and GGPP as a substrate and 10 mM Mg2+ as the cofactor. Hexane extracts from the reaction were directly investigated by GC-MS (see the supplemental material at https://doi.org/10.6084/m9.figshare.16702471).

ACKNOWLEDGMENTS
This study was funded by the Deutsche Forschungsgemeinschaft (DFG; German Research Foundation) under project ID 239748522 (CRC 1127) (project A6), project BE 4799/3-1, and Germany’s Excellence Strategy under project ID 390713860-EXC 2051 to C.B. This study was also funded by the European Research Council (ERC-CoG; 771349) and The Danish Council for Independent Research (DFF; 7014-00178) to M.P. Forestry and The Danish Council for Independent Research (DFF; 7014-00178) to M.P. Forestry South Africa funded the volatile analysis of biological samples. We acknowledge funding by the International Leibniz Research School for Microbial and Biomolecular Interactions to J.S.

REFERENCES


}


