



OPEN

Use of the proteomic tool MALDI-TOF MS in termite identification

Bouthaina Hasnaoui^{1,2}, Adama Zan Diarra^{1,2}, Jean-Michel Berenger^{1,2}, Hacène Medkour^{2,3}, Ahmed Benakhla⁴, Oleg Mediannikov^{2,3} & Philippe Parola^{1,2}✉

Matrix-assisted laser desorption/ionization mass spectrometry (MALDI-TOF MS) has proved effective for the identification of many arthropods. A total of 432 termite specimens were collected in Mali, Cote d'Ivoire, Togo, Senegal, Switzerland and France. Morphologically, 22 species were identified, including *Ancistrotermes cavithorax*, *Amitermes evuncifer*, *Cryptotermes brevis*, *Cubitermes orthognathus*, *Kaloterme flavicollis*, *Macrotermes bellicosus*, *Macrotermes herus*, *Macrotermes ivorensis*, *Macrotermes subhyalinus*, *Microcerotermes parvus*, *Microtermes* sp., *Odontotermes latericius*, *Procbitermes sjostedti*, *Promicrotermes holmgreni*, *Reticulitermes grassei*, *Reticulitermes lucifugus*, *Reticulitermes santonensis*, *Trinervitermes geminatus*, *Trinervitermes occidentalis*, *Trinervitermes togoensis*, *Trinervitermes* sp., *Trinervitermes trinervoides* and *Trinervitermes trinervius*. Analysis of MALDI-TOF MS spectra profiles from termites revealed that all were of high quality, with intra-species reproducibility and inter-species specificity. Blind testing of the spectra of 389 termites against our updated database with the spectra of 43 specimens of different termite species revealed that all were correctly identified with log score values (LSVs) ranging from 1.65 to 2.851, mean 2.290 ± 0.225 , median 2.299, and 98.4% (383) had LSVs > 1.8. This study is the first on the use of MALDI-TOF for termite identification and shows its importance as a tool for arthropod taxonomy and reinforces the idea that MALDI-TOF MS is a promising tool in the field of entomology.

Termites or white ants belong to the arthropods that have thrived on earth for over 300 million years^{1,2}. They belong to the class Insecta, infra-order Isoptera³. Phylogenetic studies have indicated that the nearest relatives are cockroaches, which explains their classification in the order of *Blattodea*⁴. Their distribution depends on climatic conditions, especially temperature and precipitation⁵. Termites live in colonies divided into two castes (reproductive or sexual and sterile or asexual)^{6,7}. Humans and termites live in close proximity⁸. Many termite species are recognized as harmful.

Termites attack the wooden parts of buildings and construction, causing damage costing more than three billion dollars each year, as well as in the agriculture field by eating fast-growing plants^{9,10}. At the same time, they are known as ecosystem engineers by influencing the distribution of natural resources, such as water and nutrients in the ecosystem¹¹, and that refers to their ability to generate valuable biogenesis that improves soil properties¹², which increases water infiltration rates^{13,14}. They are also of interest in traditional medicine, particularly for suturing wounds and treating angina, fever, burns and abscesses¹⁵. Termite species such as *Macrotermes bellicosus* have been shown to have anti-inflammatory and analgesic effects¹⁶. Termites can be used as bait to catch fish and birds¹⁵ and as a natural human food resource that has a significant value in protein and vitamins¹⁷. They participate in the beneficial chemical variation of the earth and its components¹⁸.

Morphological identification of termites, based mainly on the observation of morphological characteristics, is limited by the need for entomological expertise; i.e., the difficulty in identification up to the species level due to the ambiguity of their features, crypto-biotic social structure and their similarity, the availability of identification keys and the long time required for identification. Identification based on molecular amplification and sequencing of genes such as mitochondrial cytochrome oxidase subunits I and II (*cox1* and *cox2*), genes coding for the NADH-ubiquinone oxidoreductase chain 1 (*ND1*), *Internal Transcribed Spacer (ITS2)*, large and small ribosomal RNA subunits (*16S* and *12S rRNA*), and nuclear DNA such as *18S*, *Microsat* and genes for endo-beta-1,4-glucanase (*RsEG*), interactive domain-containing protein 1A (*AT-rich DNA*); these molecular markers

¹Aix Marseille Univ, IRD, SSA, AP-HM, VITROME, Marseille, France. ²IHU-Méditerranée Infection, Marseille, France. ³Aix Marseille Univ, IRD, AP-HM, MEPHI, Marseille, France. ⁴Département des Sciences Vétérinaire, Université Chadli Bendjdid, 36000 El Tarf, Algeria. ✉email: philippe.parola@univ-amu.fr

have proven to be an efficient alternative for species identification and overcoming morphological limitations¹⁹. However, the molecular identification approach is still limited by the high cost of reagents, the time consumed, and the absence of universal primers allowing for the amplification of a given gene in all species and sequences of all species on GenBank²⁰.

To overcome the difficulties of morphological and molecular identification of arthropods, MALDI-TOF MS has been proposed as an alternative identification technology to these two methods. The MALDI-TOF MS is a technique that allows identifying an organism from protein signals (borrowed protein) of molecular weight between 2000 and 20,000 Da. This method has been used in many studies to identify different arthropods, including ticks, mosquitoes, biting midges, fleas, lice, bedbugs, triatomines and phlebotomine sand flies, and also for the determination of their blood meal origin and to discriminate between the infectious status of some arthropod vectors²¹. In medical entomology, the use of MALDI-TOF MS requires a development of protocols such as the choice of the compartment to be used and the quantity of crushing mix allowing to have spectra with intra-species reproducibility and inter-species specificity. The part of the arthropod to generate reproducible and specific spectra by MALDI-TOF MS analysis varies between arthropod groups but also according to the developmental stages of the arthropod²¹. For example, the legs are used for mosquitoes and ticks, the cephalothorax for fleas, lice and bedbugs²¹. MALDI-TOF is a fast and easy technique that does not require expertise in entomology. However, the high cost of the machine, its maintenance, the choice of the compartment used for the analysis and the preservation methods are the limiting factors of this technique²¹. The machines used in entomology research are those used in the microbiology platform, at no extra cost, and the cost of the analysis is very low. At present, for termites the MALDI-TOF MS tool has been used to identify methanotrophic bacteria in the gut²², cuticular hydrocarbons²³, carboxy-methyl cellulose, crystalline celluloses or xylan from the gut of *Reticulitermes santonensis*²⁴ and the chemical profile and antimicrobial activity of *Macrotermes bellicosus* used in traditional medicine²⁵. However, no study has been done on the identification of termite species by MALDI-TOF MS. Hence, the aim of this study is to evaluate the ability of MALDI-TOF to identify different species of termites collected in four West African countries and in two countries in Europe.

Results

Termite collection and morphological identification. A total of 432 termite specimens were collected, including 135 in Marseille (32%), 123 in Senegal (28%), 87 in Cote d'Ivoire (20%), 79 in Mali (18%), four in Switzerland (1%) and four in Togo (1%). Morphologically, the termites were identified as belonging to 23 species, including five *Ancistrotermes cavithorax*, six *Amitermes evuncifer*, 12 *Cubitermes orthognathus*, four *Cryptotermes brevis*, 12 *Kaloterms flavicollis*, 28 *Macrotermes bellicosus*, four *Macrotermes herus*, three *Macrotermes ivorensis*, 13 *Macrotermes subhyalinus*, eight *Microcerotermes parvus*, one *Microtermes* sp., 75 *Odontotermes latericius*, eight *Promirotermes holmgreni*, two *Procubitermes sjostedti*, 14 *Reticulitermes grassei*, 84 *Reticulitermes lucifugus*, 26 *Reticulitermes santonensis*, 26 *Trinervitermes geminatus*, 13 *Trinervitermes occidentalis*, 26 *Trinervitermes* sp., 15 *Trinervitermes togoensis*, 34 *Trinervitermes trinervoides* and 13 *Trinervitermes trinervius* (Fig. 1 and Table 1). The largest number of species was identified in Senegal, with 12 different species (Fig. 1 and Table 1). The termites belonged to the sterile caste (workers and soldiers) and reproductive caste (winged). We could not identify several specimens of *Trinervitermes* and *Microtermes* termites up to the species level (Table 1).

For all the species that we identified morphologically, pictures of a body and mandibular were taken and represented in the supplementary Fig. 1.

Validation of morphological identification by molecular tools. A total of 47 randomly selected termite specimens morphologically identified for the creation of our database were submitted to standard PCR and sequencing for molecular identification using both 12S rRNA and COI genes. BLAST analysis of sequences obtained from specimens identified as *M. subhyalinus* showed that they were 100% and 99.4% identical to the corresponding sequence of *M. subhyalinus* (GenBank: DQ441726 and AY127708) respectively for 12S rRNA and COI genes. Those identified as *M. bellicosus* were 99.7% and 93.7% identical to the corresponding sequence of *M. bellicosus* (AY127702) and *Pseudacanthotermes militaris* (KY224517). Sequences from termites identified as *M. herus*, *M. ivorensis* and *M. subhyalinus* were 99.2 to 99.7% identical to *Macrotermes* sp. (KY224531) and 99 to 99.4% to *M. subhyalinus* (DQ441726) for 12S rRNA gene and they were also 98.3 to 98.7% identical to *M. subhyalinus* (FJ207427) and 98 to 98.7% to *M. herus* (FJ207441) for COI gene, but we don't obtained sequences using COI gene for *M. ivorensis*. Specimens morphologically identified as *T. geminatus* and *T. trinervoides* were 99.5% to 100% and 97.2% identical to the corresponding sequence of *T. geminatus* (DQ441829 and JF923343), respectively for both genes. The sequences from termites identified as *T. occidentalis*, *T. togoensis* and *Trinervitermes* sp. were 96.4% and 95.9% to 96.3% identical to the corresponding sequence of *Cortaritermes intermedius* (MH574830) and *T. occidentalis* (JF923333). Sequences from termites identified as *T. trinervius* were 96.1% to 96.7% and 95.4% identical to the corresponding sequence of *T. geminatus* (DQ441829) and *T. geminatus* (JF923343), respectively for both genes. The sequences obtained from termites morphologically identified as *R. santonensis*, *R. grassei* and *R. lucifugus* were 99.7% to 99.5% identical to the corresponding sequence of *R. lucifugus* (NC_045240), respectively for both genes. The specimens identified as *K. flavicollis* were 99.2% and 91.8% identical to *K. flavicollis* (DQ441716; EU253842), respectively for both genes. The sequences from termites identified as *O. latericius* were 97% and 99.3% identical to sequence of *Odontotermes obesus* (KY224493) and *O. latericius* (AY127714), and those identified as *C. orthognathus* were identical to 99.5% to sequences of 98.4% to *Cubitermes* sp. (EU253865) and *Nitiditermes* sp. (MN646731). Termites identified as *A. evuncifer* and *Mi. parvus* were 97.7% and 99.6% identical to the corresponding sequence of *A. evuncifer* (DQ441626; AY127718) and 96.6% and 96.3% identical to *Microcerotermes* sp. (KY224578; F923261), respectively for both species using two genes. The sequences from termites identified as *A. cavithorax* and *P. sjostedti* were 98.7% identical to the

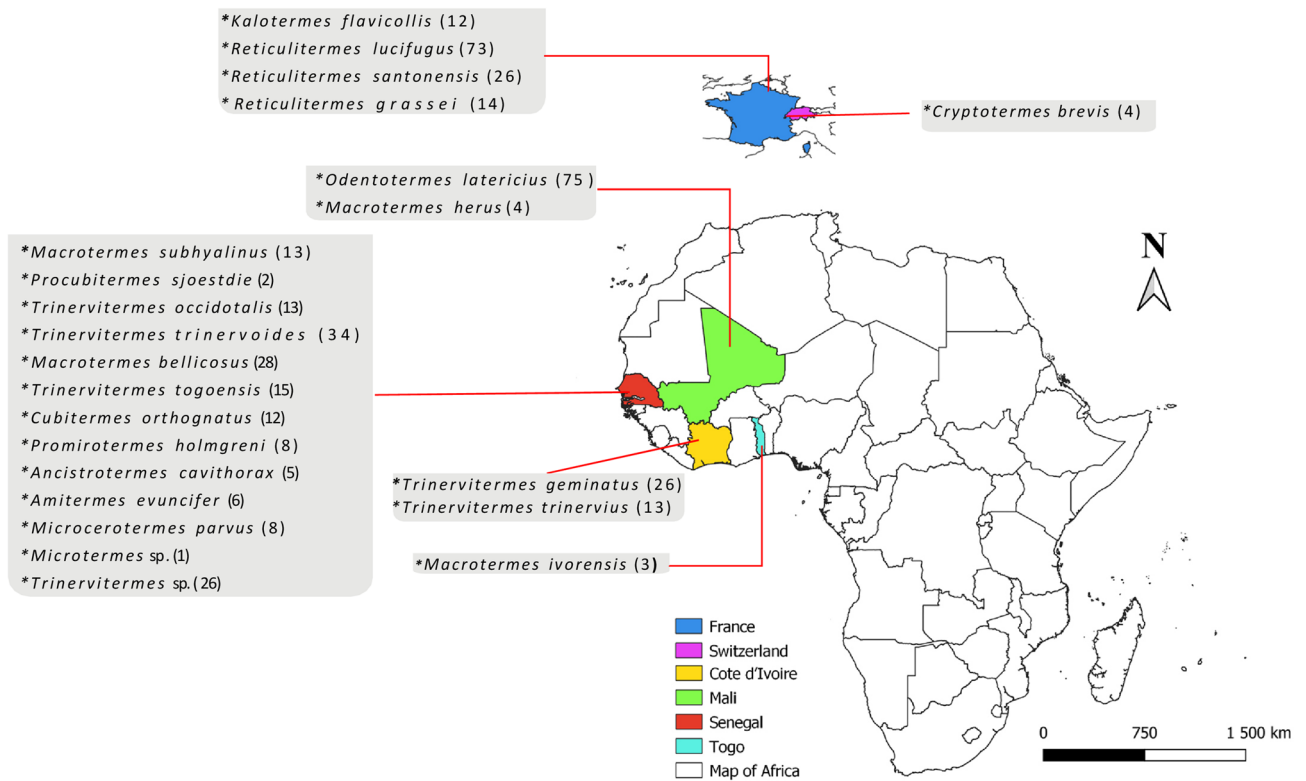


Figure 1. Map of Africa and part of Europe showing termite collection sites, morphologically identified termite species names and numbers of each species.

sequence of *A. crucifer* (AY127715) and 97% to the sequence of *Cubitermes ugandensis* (KP091689), respectively for the 12S rRNA gene, 99.50% identical to the sequence of *An. cavithorax* (DQ441630), and 99.52% to the sequence of *Nitiditermes proximatus* (MN646731) for the COI gene. Sequences from termites identified as *Microtermes* sp. and *Trinervitermes* sp. were 96.8% and 99.50% identical to the sequence of *Microtermes* sp. (KY224390 and AY127713) and 96.9% *Nasutitermes corniger* (DQ441746) and 95.9% to *T. occidentalis* (JF923309), respectively for both termite species. Those from termites identified as *P. holmgreni* were 96.9% and 91.5% identical to *Promirotermes pygmaeus* (KY224529) and *Promirotermes* sp. (KY224554) respectively for both genes. Those from *C. brevis* were 100% and 99.6% identical to *C. brevis* (KM618724 and MT040337) respectively for both genes. The results of the molecular identification of the termites in our study are summarized in Table 1. The phylogenetic position of termites in this study is shown in Fig. 2A and B. The sequences of the 12S rRNA and COI genes obtained in this study were deposited in the GenBank (National Centre for Biotechnology Information, NCBI) under the following accession numbers: MW078935 to MW078965 and MZ029056 to MZ029087, respectively for both genes (supplementary Table 1).

MS spectra analysis. In total, the legs of 432 termite specimens were subjected to MALDI-TOF/MS analysis. The visualization of the MS spectra obtained from all samples showed that they were of high quality (no smoothing, baseline subtraction corrects and peak intensity > 3000 arbitrary units.) (Fig. 3). A cluster analysis (dendrogram) was performed with two to five MS spectra from the legs of each species to evaluate the reproducibility and specificity of the spectra according to species. A perfect clustering of specimens of the same species on the same branch reveals intra-species reproducibility and inter-species specificity for the different termite species (Fig. 4). Interestingly, the MS spectra of sterile and reproductive caste specimens of the same species were specific according to caste as shown in the PCA and dendrogram made with two castes (reproductive and sterile) of *K. flavicollis* and *R. lucifugus* (Supplementary Fig. 2).

All the reference MS spectra included in our lab database have been deposited in a public repository in order to be shared with the entire research community, they are available and can be downloaded with the following DOI number: <https://doi.org/10.35088/q281-st29>.

Blind test for validation of termite identification. The accuracy of MALDI-TOF MS identification of 17 termite species was evaluated by querying 389 specimens morphologically, identified against our updated MALDI-TOF MS database with 43 spectra of one to ten spectra per species confirmed by molecular biology (Table 2). However, the termite identified as *Microtermes* sp. was not included in the blind tested because there was only one specimen.

The result of the interrogation showed that 100% (389) of the species were correctly identified; i.e., they agreed with our morphological identification, with an LSV ranging from 1.65 to 2.851, mean of 2.290 ± 0.225 , median

Morphological identification (number identified)	Number tested by molecular biology	Molecular identification			
		Closest GenBank match (% identity) 12S rRNA	Accession number	Closest GenBank match (% identity) COI	Accession number
<i>Ancistrotermes cavithorax</i> (5)	1	<i>Ancistrotermes crucifer</i> (98.7%)	DQ441630	<i>A. cavithorax</i> (99.3%)	AY127715
<i>Amitermes evuncifer</i> (6)	1	<i>A. evuncifer</i> (97.7%)	DQ441626	<i>A. evuncifer</i> (99.3%)	AY127718
<i>Cryptotermes brevis</i> (4)	2	<i>C. brevis</i> (100%)	KM618724	<i>C. brevis</i> (99.6%)	MT040337
<i>Cubitermes orthognathus</i> (12)	2	<i>Cubitermes ugandensis</i> (98.5%)	KP091689	<i>Nitiditermes</i> sp. (99.5%)	MN646703
<i>Kalotermes flavicollis</i> (12)	2	<i>K. flavicollis</i> (99.2%)	DQ441716	<i>K. flavicollis</i> (92.6%)	EU253842
<i>Macrotermes bellicosus</i> (28)	3	<i>Pseudacanthotermes militaris</i> (93.6%)	KY224517	<i>M. bellicosus</i> (98.9–99.5%)	AY127702
<i>Macrotermes herus</i> (4)	1	<i>Macrotermes</i> sp. (99.7%) <i>M. subhyalinus</i> (99.4%)	KY224531 DQ441726	<i>M. subhyalinus</i> (98.7%) <i>M. herus</i> (98.7%)	FJ207427 FJ207441
<i>Macrotermes ivorensis</i> (3)	1	<i>Macrotermes</i> sp. (99.4%) <i>M. subhyalinus</i> (99.2%)	KY224531 JX144937	–	
<i>Macrotermes subhyalinus</i> (13)	2	<i>Macrotermes</i> sp. (99.2%) <i>M. subhyalinus</i> (99%)	KY224531 JX144937	<i>M. subhyalinus</i> (98.3%) <i>M. herus</i> (98%)	FJ207427 FJ207441
<i>Microcerotermes parvus</i> (8)	1	<i>Microcerotermes</i> sp. (96.5%)	KY224578	<i>Microcerotermes</i> sp. (92.5%)	KY224573
<i>Microtermes</i> sp. (1)	1	<i>Microtermes</i> sp. (98.7%)	KY224390	<i>Microtermes</i> sp. (99.5%)	AY127713
<i>Odontotermes latericius</i> (75)	4	<i>Odontotermes obesus</i> (97.0%)	KY224493	<i>O. latericius</i> (99.5%)	AY127714
<i>Procubitermes sjostedti</i> (2)	1	<i>Cubitermes ugandensis</i> (96.9%)	KP091689	<i>Nitiditermes proximatus</i> (99.5%)	MN646731
<i>Promirotermes holmgreni</i> (8)	1	<i>Promirotermes pygmaeus</i> (96.1%)	KY224529	<i>Promirotermes</i> sp. (91.5%)	KY224554
<i>Reticulitermes grassei</i> (14)	2				
<i>Reticulitermes lucifugus</i> (84)	6	<i>R. lucifugus</i> (99.7%)	NC045240	<i>R. lucifugus</i> (99.5%)	NC045240
<i>Reticulitermes santonensis</i> (26)	2				
<i>Trinervitermes geminatus</i> (26)	2				JF923343
<i>Trinervitermes trinervoides</i> (34)	2	<i>T. geminatus</i> (99.5–100%)	DQ441829	<i>T. geminatus</i> (96.8–97.2%)	JF923317
<i>Trinervitermes occidentalis</i> (13)	3				
<i>Trinervitermes togoensis</i> (15)	2	<i>Cortaritermes intermedius</i> (96.4%)	MH574830	<i>T. occidentalis</i> (95.9–96.3%)	JF923333 JF923309
<i>Trinervitermes</i> sp. (26)	3				
<i>Trinervitermes trinervius</i> (13)	2	<i>T. geminatus</i> (96.1–96.7%)	DQ441829	<i>T. geminatus</i> (97.2%)	JF923343

Table 1. Morphological and molecular identification of termites, number of each termite species, number tested by molecular and 12S rRNA and COI accession number.

of 2.299 (Table 2) and 98.4% (382) had LSVs > 1.8 (Supplementary Fig. 3). Looking at the level of each species for *A. cavithorax*, LSV identification ranged from 1.98 to 2.752, with a mean of 2.325 ± 0.236 ; for *A. evuncifer* the LSVs ranged from 1.919 to 2.744 (2.2 ± 0.306); for *C. orthognathus*, from 1.938 to 2.813 (2.508 ± 0.271); for *C. brevis*, from 2.041 to 2.350 (2.181 ± 0.112); for *K. flavicollis*, from 1.729 to 2.547 (2.265 ± 0.231); for *M. bellicosus*, from 2.062 to 2.737 (2.459 ± 0.115); for *M. herus*, from 2.585 to 2.78 (2.678 ± 0.068); for *M. ivorensis*, from 2.173 to 2.471 (2.322 ± 0.149) and for *M. subhyalinus* it ranged from 1.933 to 2.572 (2.309 ± 0.195). Similarly, *Mi. parvus* was identified, with LSVs ranging from 1.818 to 2.849 and a mean of 2.393 ± 0.225 ; *O. latericius*, with 1.963 to 2.581 (2.204 ± 0.111); *P. holmgreni*, with 1.719 to 2.846 (2.277 ± 0.257); *R. lucifugus*, with 1.797 to 2.851 (2.329 ± 0.156); *T. geminatus*, with 1.65 to 2.693 (2.259 ± 0.165); *T. occidentalis*, with 1.727 to 2.642 (2.226 ± 0.171); *Trinervitermes* sp., with 1.817 to 2.515 (2.066 ± 0.200) and one *P. sjostedti* identified with 2.806.

Discussion

According to our morphological identification, 22 species belonging to 12 genera were identified in this study. The presence of most of the species identified in West Africa had already been reported in this area²⁶. Termites belonged to the reproductive or biological caste (with wings) and sterile caste composed of workers and soldiers (without wings). Each caste ensures a specific role within the colony; thus, the soldiers are involved in defense, workers in elementary tasks (gathering food, taking care of the queen and king and constructing or repairing the nest) and the reproductive caste ensure the function of reproduction²⁷.

The molecular biology of social insects has always been contentious²⁸; we used both gene systems including 12SrRNA and COI in our study. Based on previous studies the 12SrRNA gene was suggested as more discriminating compared to the COI gene^{29,30}. However, molecular identification of Isoptera has most often used the COI gene^{29,31,32}. Additionally, 12S is characterized by the narrow spectrum of sequences for each species where each species is presented by one or two sequence.

Based on the results of molecular biology, our morphological identification was confirmed for some species, such as *K. flavicollis*, *M. subhyalinus*, *R. lucifugus*, *T. geminatus*, *T. occidentalis*, *C. brevis*, *A. evuncifer*, *A. cavithorax*, *M. bellicosus*, *M. subhyalinus*, *M. herus* and *O. latericius* by both genes or only by the COI gene, whose homologous sequences were available on Genbank. However, for ten species (*M. ivorensis*, *Mi. parvus*, *P. sjostedti*, *R. santonensis*, *R. grassei*, *P. holmgreni*, *T. trinervoides*, *T. togoensis*, *T. trinervius*) morphological identification

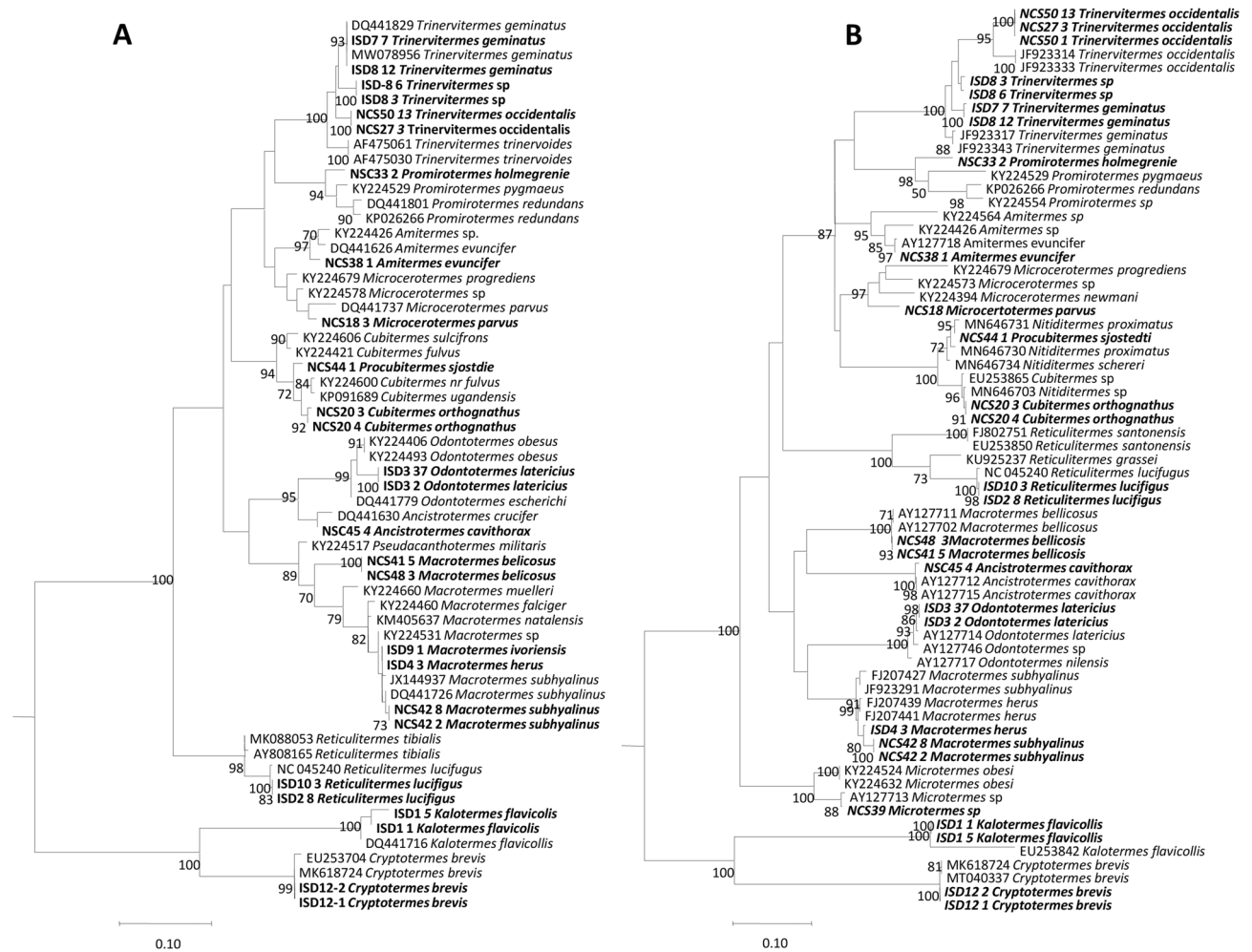


Figure 2. Phylogenetic tree based on *12S rRNA* gene sequences (A) and *COI* gene sequences (B). Phylogenetic tree highlighting the position of studied termites (in bold) relative to other termite sequences available via Genbank. The partial sequences of the mitochondrial *cox1* gene and the *12SrRNA* gene were aligned using CLUSTALW, and phylogenetic inferences were obtained from a maximum likelihood phylogenetic analysis with the Tamura and Nei 1993 (TrN) submodel for *12SrRNA* gene (A) and with the general time-reversible (GTR) submodel (B) proposed by TOPALi 2.5 software. The GenBank accession numbers are indicated at the beginning. The numbers at the nodes are the bootstrap values obtained by repeating the analysis 100 times to generate a majority consensus tree.

did not match with molecular identification. Most probable explication is the morphological misidentifications that we made due to close anatomical similarities among species from *Reticulitermes* and *Trinervitermes* genera^{33,34}. On the other hand, this discrepancy between our morphological and molecular identification would be due to the updating of the systematics of the termites with a change in the name of the genus as in the case of the genus *Cubitermes* changed to *Nitiditermes*^{35,36} or by the lack of reference sequence of some species in the GenBank database, as in the case of the species *M. ivorensis* and *P. holmgreni*, which is one of the limitations of the molecular method²¹.

In this study we developed the MALDI-TOF MS tool to identify different termite species. Although some studies have been performed on the identification of the gut microbiota and chemical composition of termites using MALDI-TOF MS^{26,37,38}, to our knowledge our study is the first to use this tool for the identification of these insects. MALDI-TOF MS is a technique that allows the identification of biomolecular protein contained in a sample by soft ionization according to the mass/charge ratio (m/z)²¹. The measurement of the m/z ratio is determined by the time it takes for an ion to travel the flight path and thus generate a spectral profile specific to the composition of the sample being analyzed. Over several years, this technique has been developed for the routine diagnosis of a large number of microorganisms (bacteria, archaea, yeasts, filamentous fungi, helminths and intestinal protozoa) of medical and/or veterinary importance²¹. Over the last 15 years, MALDI-TOF MS has been widely used in entomology for the identification of a large number of arthropod vectors and non-vectors, as well as for the determination of blood meal origin and the discrimination of infected and non-infected arthropods. Although this technique is less expensive, reliable, fast and does not require knowledge of entomology, its use in entomology requires the development of specific protocols to generate reproducible and species-specific spectra²¹. The price of the machine, the choice of the part of the arthropod used for MALDI-TOF MS analysis

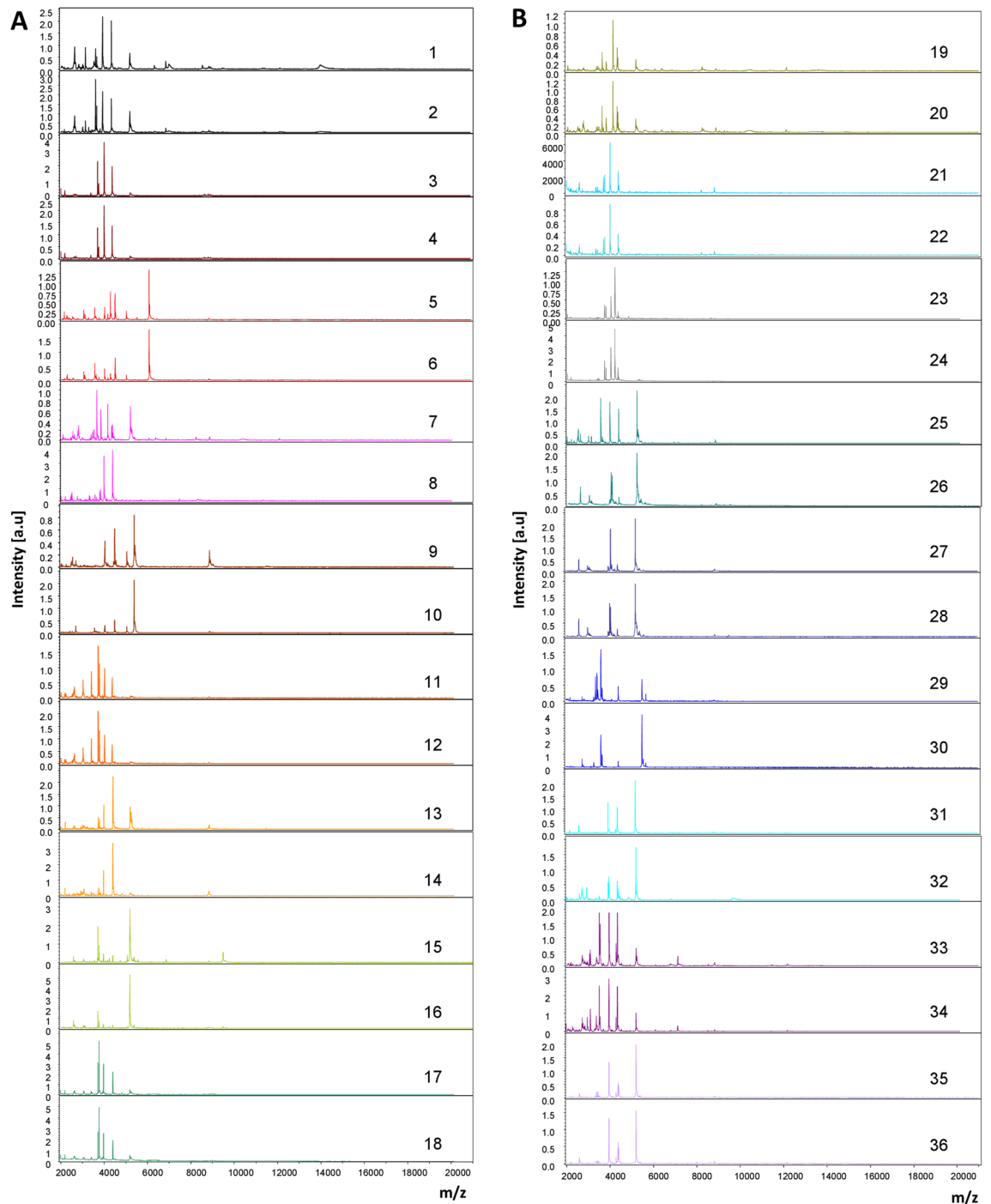


Figure 3. Representative MS profiles obtained for each termite species using the Flex analysis 3.3 software (A) and (B). (A) represents the spectra obtained from the leg protein of *Ancistrotermes cavithorax* (1–2), *Amitermes evuncifer* (3–4), *Cubitermes orthognatus* (5–6), *Cryptotermes brevis* (7–8), *Kaloterms flavicollis* (9–10), *Macrotermes bellicosus* (11–12), *Macrotermes herus* (13–14), *Macrotermes ivorensis* (15–16), *Macrotermes subhyalinus* (17–18) and (B) the spectra from *Microcerotermes parvus* (19–20), *Microtermes sp* (21–22), *Odontotermes latericius* (23–24), *Procutitermes sjostedti* (25–26), *Promirotermes holmgreni* (27–28), *Reticulitermes lucifugus* (29–30), *Trinervitermes geminatus* (31–32), *Trinervitermes occidentalis* (33–34), *Trinervitermes trinervius* (35–36). a.u.: arbitrary units; m/z: mass-to-charge ratio.

and the method of preservation of the arthropods, are limiting elements of this technique²¹. In our study, 100% of the species were correctly identified; i.e., they agreed with our morphological identification, with an LSV ranging from 1.705 to 2.862 (mean: 2.302 ± 0.254) and 97% of specimens had LSVs > 1.8. The high rate of correct

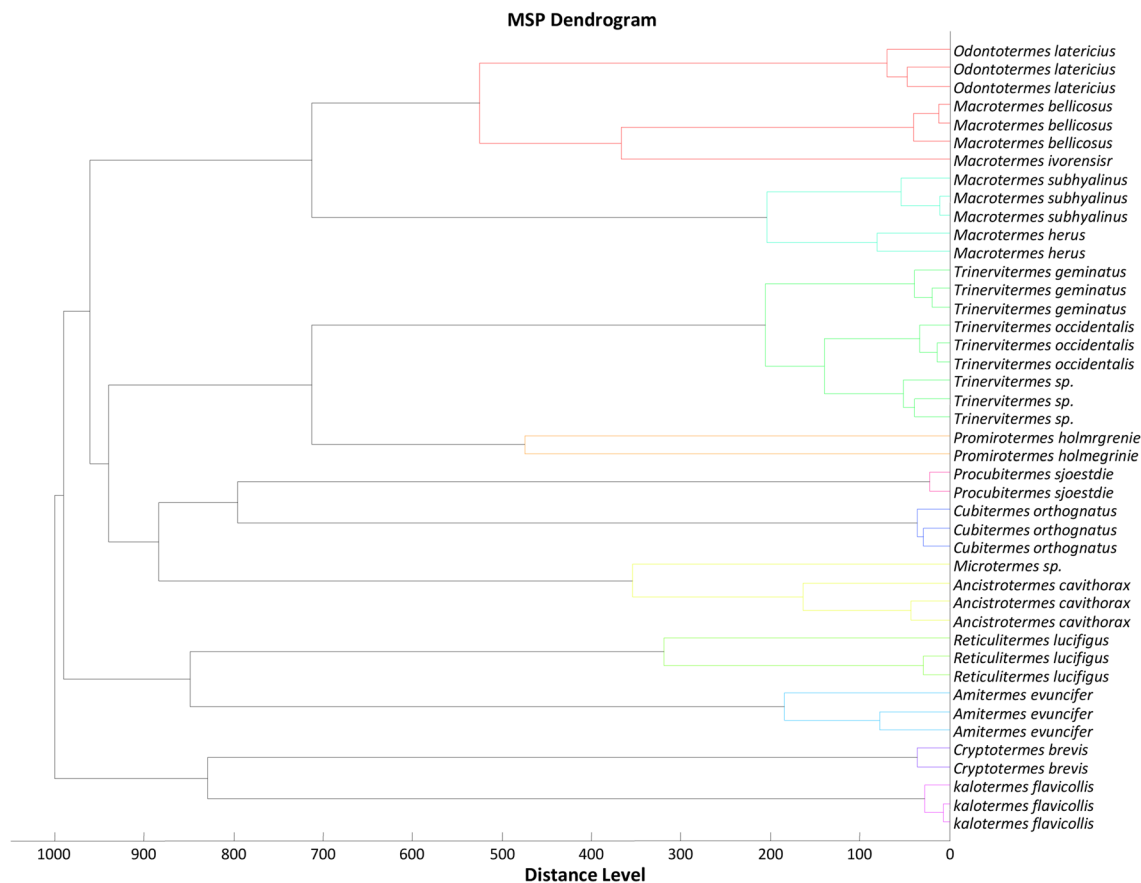


Figure 4. Dendrogram using 2 to 3 representative MS spectra from 16 distinct species treated with MalDI Biotyper v 3.0 software and the distance units correspond to the relative similarity of MS spectra.

Morphological identification	Collection country	Collection years	Number tested	Number added DB	MALDI-TOF MS identification (%)	Score rang LSV
<i>Ancistrotermes cavithorax</i>	Senegal	2019	4	1	<i>A. cavithorax</i> (100%)	[2.093–2.257]
<i>Amitermes evuncifer</i>	Senegal	2019	5	1	<i>A. evuncifer</i> (100%)	[1.993–2.421]
<i>Cubitermes orthognathus</i> (<i>Nitiditermes</i>)	Senegal	2019	10	2	<i>C. orthognathus</i> (<i>Nitiditermes</i>) (100%)	[1.705–2.253]
<i>Cryptotermes brevis</i>	Switzerland	2021	3	1	<i>C. brevis</i> (100%)	[2.041–2.350]
<i>Kaloterme flavicollis</i>	France	2013	10	2	<i>K. flavicollis</i> (100%)	[1.8–2.46]
<i>Macrotermes bellicosus</i>	Senegal	2019	25	3	<i>M. bellicosus</i> (100%)	[1.818–2.577]
<i>Macrotermes herus</i>	Mali	2016	3	1	<i>M. herus</i> (100%)	[2.780–2.322]
<i>Macrotermes ivorensis</i>	Togo	2010	2	1	<i>M. ivorensis</i> (100%)	[2.226–2.024]
<i>Macrotermes subhyalinus</i>	Senegal	2019	11	2	<i>M. subhyalinus</i> (100%)	[1.795–2.572]
<i>Microcerotermes parvus</i>	Senegal	2019	7	1	<i>Mi. parvus</i> (100%)	[2.531–2.853]
<i>Microtermes sp.</i>	Senegal	2019	0	1	–	–
<i>Odontotermes latericius</i>	Mali	2019	71	4	<i>O. latericius</i> (100%)	[2.001–2.581]
<i>Procubitermes sjoestedti</i>	Senegal	2019	2	1	<i>Procubitermes sjoestedti</i> (100%)	[2.448–2.455]
<i>Promirotermes holmgreni</i>	Senegal	2019	6	1	<i>P. holmgreni</i> (100%)	[2.492–2.599]
<i>Reticulitermes lucifugus</i>	France	2014–2021	113	11	<i>R. lucifugus</i> (100%)	[2.002–2.851]
<i>Trinervitermes geminatus</i>	Côte d'Ivoire/Senegal	2016/2019	56	4	<i>T. geminatus</i> (100%)	[1.924–2.403]
<i>Trinervitermes occidentalis</i>	Senegal/Cote d'Ivoire	2019	50	4	<i>T. occidentalis</i> (100%)	[1.93–2577]
<i>Trinervitermes trinervius</i>	Cote d'Ivoire	2016	11	2	<i>Trinervitermes sp.</i> (100%)	[1.921–2.342]

Table 2. Termite collection information (location, year of collection, morphological identification) and MALDI-TOF identification with log score value obtained by blind test.

identification, proving the performance of the MALDI-TOF MS tool in distinguishing the different arthropod species, has already been reported in several studies²¹. It is interesting to note that very closely related species, often difficult to distinguish morphologically, (species of the genus *Macrotermes* and *Trinervitermes*) and specimens of different castes (reproductive and sterile), could be distinguished by MALDI-TOF MS. The ability of MALDI-TOF MS to distinguish closely related tick species has already been reported³⁹.

To summarize, our study reliably shows that MALDI-TOF/MS is a promising tool that may make termite studies much easier and specific. It would be interesting to apply this innovative tool on termites from other sites in the world in order to have a maximum of species to enrich our MALDI-TOF MS database.

Materials and methods

Study area and collection periods. Termites were collected in two European countries (France and Switzerland) and in four West African countries: (Mali, Ivory Coast, Togo and Senegal). The termites from France were collected in a heavily infested house in an old building located in Châteauneuf-Les-Martigues near Marseille. The pictures taken show the damage caused by the termites. The termites from Marseille were sampled in 2013, 2014 and 2021, those from Togo in 2010, Mali and Ivory Coast in 2016, Senegal in 2019 and Switzerland in 2021.

The collection in Senegal was made at four sites; Niokolo-Koba National Park (sites of Simenti, Dar Salam and Niokolo Poste) and in Southeastern Senegal (region of Kédougou) in the Dindefello forest, were the termitaries were identified visually, and samples collected using a shovel and pincer. From each termitary, the soil substrate, the fungus combs (if available) and adult termites (soldiers and workers) were collected in a plastic ventilated box where they were stored during the transport at ambient temperature. On arrival, termites were separated from the substrate and fungus. The maps showing the collection sites were made with QGIS software version 3.20 (Fig. 1). All termites were stored in 70% ethanol except for termites collected from Marseille in 2020. They were stored at $-20\text{ }^{\circ}\text{C}$ because they were collected in the proximity of the laboratory.

Morphological identification. Morphological identification of termites was done down to the genus and/or species level using different keys based on the specific and discriminating characteristics of soldiers, workers and alates, or biological, such as the keys of Sands⁴⁰, Clay⁴¹, Adlard et al.⁴², Ifan⁴³, Bouillon⁴⁴.

The termite collection site was also an important element that was considered and helped us in the taxonomy.

For workers of some species, we dissected the mandible to identify termites down to the species. The criteria were observed with the optical microscope, the binocular loupe, and the ZEISS Axio Zoom V16, and then photographed with the digital Canon E05 7D supplied with a Canon MP-E 65 mm Lens (French).

Molecular identification of termites. To confirm the morphological identification of termites whose spectra were to be introduced into our MS spectra database (DB), molecular analyses were performed. DNA was extracted from a small part of the termite abdomen, which was subjected to enzymatic lysis by incubation at $56\text{ }^{\circ}\text{C}$ overnight in $180\text{ }\mu\text{L}$ of lysis buffer G2 (QIAGEN, Hilden, Germany) and $20\text{ }\mu\text{L}$ of proteinase K (QIAGEN, Hilden, Germany). Total DNA was extracted into $100\text{ }\mu\text{L}$ of eluate using EZ1 Tissue Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions and stored at $-20\text{ }^{\circ}\text{C}$ before use.

Molecular identification of termites at the species level was performed by sequencing the standard PCR product of a fragment of the small subunit ribosomal RNA (*12S rRNA*) gene (12S-F SR-J-141995: 5'TACTAT GTTACGACTTAT-3'/12S-R SR-N-14594: 5'AAACTAGGATTAGATACCC-3') and the cytochrome c oxidase I gene (*COI*) (LCO1490: 5'-GGTCAACAAATCATAAAGAYATYGG-3' and dgHCO2198: 5'-TAAACTTCAGGG TGACCAARAAYCA-3') as previously described^{29,45}. PCR products were separated and visualized by electrophoresis in 1.5% agarose. All positive samples were purified and directly sequenced using the commercial Big Dye Terminator Cycle Sequencing Kit (Perkin Elmer Applied Biosystems, Foster City, CA, USA) with an ABI automated sequencer (Applied Biosystems). Obtained sequences were edited using ChromasPro software (ChromasPro 1.7, Technelysium Pty Ltd., Tewantin, Australia) and compared for similarity to sequences available in the GenBank (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>). Molecular phylogenetic and evolutionary analyses were conducted in TOPALi2.5 (<http://www.topali.org/>).

Termite preparation for MALDI-TOF analysis. After rinsing and drying on sterile filter paper, three legs from each of the termites were individually placed in an Eppendorf tube and dried at $37\text{ }^{\circ}\text{C}$ overnight for those stored in alcohol. Specimens that were at $-20\text{ }^{\circ}\text{C}$ were not dried at $37\text{ }^{\circ}\text{C}$ overnight. The legs were then homogenized with the TissueLyser (Qiagen) with a pinch of glass beads and $20\text{ }\mu\text{L}$ of a mixture of 70% formic acid and 50% acetonitrile (Fluka, Buchs, Switzerland) in a three-minute cycle at a frequency of 30 Hertz as already described^{20,46,47}. The legs of *Aedes albopictus* reared in our laboratory were used as a positive control in all manipulations.

MALDI-TOF/MS parameters. Protein mass profiles were obtained using a Microflex LT MALDI-TOF Mass Spectrometer (Bruker Daltonics, Germany), with detection in the linear positive-ion mode at a laser frequency of 50 Hz within a mass range of 2–20 kDa. The setting parameters of the MALDI-TOF/MS apparatus were identical to those previously used^{20,48,49}. Briefly, the acceleration voltage was 20 kV, and the extraction delay time was 200 ns. Each spectrum corresponds to ions obtained from 240 laser shots performed in six regions of the same spot and automatically acquired using the AutoXecute of the Flex Control v.2.4 software (Bruker Daltonics).

Spectra analysis. The MS spectra were then exported to flex Analysis v3.3, ClinProTools v2.2 and MALDI-Biotyper v3.0. (Bruker Daltonics) software for data processing (smoothing, baseline subtraction, peak picking). The quality of MS spectra was evaluated by visualization of spectra obtained from the four spots for each sample with the flex Analysis v3.3 software (Bruker Daltonics). Cluster analyses (MSP dendrogram) and principal component analysis (PCA) were performed to verify intra-species reproducibility and inter-species specificity as well as variability within different castes (soldiers and workers) of the same species. Cluster analyses were performed based on the comparison of the MSPs given by the MALDI-Biotyper v3.0. software and grouped according to the mass profile of the proteins (i.e., their mass signals and intensities) and it reflects how tick specimens are related to each other. The setting parameters were as follows: distance measure by correlation, linkage by average; the score threshold value for a single organism was 300 (arbitrary unit) and for related organisms was 0 (arbitrary unit).

Database creation and blind test for validation of termite identification. Reference MS spectra were created from the spectra of each termite species, where available, using MALDI-Biotyper v3.0. (Bruker Daltonics). MS spectra of legs from 42 specimens of termites identified morphologically and molecularly were added into our MS spectra database (DB)⁵⁰, *Microtermes* sp. spectra were not added to the DB because we had only one specimen.

All remaining spectra were blind tested against our MS spectra database for termite identification. The level of the identification significance was established using the log score values (LSVs) given by the MALDI-Biotyper v3.3 software that correlated with a corresponding degree of signal intensity of the request and reference mass spectra. The LSVs, ranging from 0 to 3, were obtained for each spectrum of the tested samples. The results of identification were considered reliable and relevant when the LSVs were greater than or equal to 1.8, as previously established in many studies^{20,49}.

Data availability

All relevant data are within the manuscript and its Supporting Information files.

Received: 20 September 2021; Accepted: 8 December 2021

Published online: 14 January 2022

References

- Šobotník, J. & Dahlsjö, C. A. L. Isoptera. (2017).
- Sekhar, C. & Vidhyavathi, A. *Termite and Food Security*. (AkiNik Publication, 2018).
- Krishna, K., Grimaldi, D. A., Krishna, V. & Engel, M. S. Treatise on the Isoptera of the world. (Bulletin of the American Museum of Natural History, no. 377). *Isoptera*. (2013).
- Daley, J. Termites Are Moving in With Cockroaches, Taxonomically. *Smithsonian Magazine* <https://www.smithsonianmag.com/smart-news/termites-are-moving-cockroaches-taxonomically-180968332/>.
- Andrieu, D. *et al.* Géographie des termites souterrains en région Centre-Val de Loire: le risque d'une espèce invasive. *Cybergeo Eur. J. Geogr.* <https://doi.org/10.4000/cybergeo.28412> (2017).
- Harun, Y. *The Miracle of Termites*. (2007).
- Hussain, A., Tian, M.-Y. & Wen, S.-Y. Exploring the caste-specific multi-layer defense mechanism of Formosan subterranean termites, *Coptotermes formosanus* Shiraki. *Int. J. Mol. Sci.* **18**, 2694 (2017).
- Jacques Renoux. Les termites et l'homme. *Laboratoire ÉBÉNA - Université Paris XII, Val de Marne*.
- Mitchell, J. D. Termites as pests of crops, forestry, rangeland and structures in southern Africa and their control. *Sociobiology* **40**, 47–69 (2002).
- Lax, A. R. & Osbrink, W. L. A. United States Department of Agriculture-Agriculture Research Service research on targeted management of the Formosan subterranean termite *Coptotermes formosanus* Shiraki (Isoptera: Rhinotermitidae). *Pest Manag. Sci.* **59**, 788–800 (2003).
- Jouquet, P., Traoré, S., Choosai, C., Hartmann, C. & Bignell, D. Influence of termites on ecosystem functioning. Ecosystem services provided by termites. *Eur. J. Soil Biol.* **47**, 215–222 (2011).
- Cheik, S. *et al.* Effects of termite foraging activity on topsoil physical properties and water infiltration in Vertisol. *Appl. Soil. Ecol.* **133**, 132–137 (2019).
- Mettrop, I. S., Cammeraat, L. H. & Verbeeten, E. The impact of subterranean termite activity on water infiltration and topsoil properties in Burkina Faso. *Ecohydrology* **6**, 324–331 (2013).
- Shanbhag, R. R. *et al.* Litter quality affects termite sheathing production and water infiltration in the soil. *Sociobiology* **66**, 491–499 (2019).
- Lawal, O. A. & Banjo, A. D. Survey for the usage of arthropods in traditional medicine in southwestern Nigeria. *J. Entomol.* **4**, 104–112 (2007).
- Thierry, K. *et al.* Antioxidant activity of the major soldier salivary glands extracts of the termite *Macrotermes bellicosus*. *IJSR* **5**, 4 (2013).
- Klass, K.-D., Nalepa, C. & Lo, N. Wood-feeding cockroaches as models for termite evolution (Insecta: Dictyoptera): *Cryptocercus* vs. *Parasphaeria boleiriana*. *Mol. Phylogenet. Evol.* **46**, 809–817 (2008).
- Scharf, M. E. Termites as targets and models for biotechnology. *Annu. Rev. Entomol.* **60**, 77–102 (2015).
- Uva, P. *et al.* Origin of a new Reticulitermes termite (Isoptera, Rhinotermitidae) inferred from mitochondrial and nuclear DNA data. *Mol. Phylogenet. Evol.* **30**, 344–353 (2004).
- Ouarti, B. *et al.* Development of MALDI-TOF mass spectrometry for the identification of lice isolated from farm animals. *Parasite* **27**, 28 (2020).
- Sevestre, J., Diarra, A. Z., Laroche, M., Almeras, L. & Parola, P. Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry: an emerging tool for studying the vectors of human infectious diseases. *Future Microbiol.* **16**, 323–340 (2021).
- Reuss, J. *et al.* Isolation of methanotrophic bacteria from termite gut. *Microbiol. Res.* **179**, 29–37 (2015).
- Cvacka, J., Jiros, P., Šobotník, J., Hanus, R. & Svatos, A. Analysis of insect cuticular hydrocarbons using matrix-assisted laser desorption/ionization mass spectrometry. *J. Chem. Ecol.* **32**, 409–434 (2006).
- Brasseur, C. *et al.* MALDI-TOF MS analysis of cellooligosaccharides and xylo-oligosaccharides produced by hindgut homogenates of *Reticulitermes santonensis*. *Molecules* **19**, 4578–4594 (2014).

25. Hammoud Mahdi, D. *et al.* Chemical profile and antimicrobial activity of the fungus-growing termite strain *Macrotermes bellicosus* used in traditional medicine in the Republic of Benin. *Molecules* **25**, 5015 (2020).
26. Korb, J. *et al.* Termite taxonomy, challenges and prospects: West Africa, a case example. *Insects* **10**, 32 (2019).
27. Nalepa, C. & Bandi, C. *Termites: Evolution, Sociality, Symbioses, Ecology*, in 53–75. https://doi.org/10.1007/978-94-017-3223-9_3 (2000).
28. Donovan, S. E., Jones, D. T., Sands, W. A. & Eggleton, P. Morphological phylogenetics of termites (Isoptera). *Biol. J. Lin. Soc.* **70**, 467–513 (2000).
29. Alajmi, R., Mashaly, A., Al-Otaibi, N., Mahmoud, A. & Ayaad, T. Efficiency of three mitochondrial genes in molecular identification and phylogenetic analysis of termites. *Res. J. Biotechnol.* **14**, 10 (2019).
30. Legendre, F. *et al.* The phylogeny of termites (Dictyoptera: Isoptera) based on mitochondrial and nuclear markers: Implications for the evolution of the worker and pseudergate castes, and foraging behaviors. *Mol. Phylogenet. Evol.* **48**, 615–627 (2008).
31. Khaustov, A. A., Hugo-Coetzee, E. A. & Ermilov, S. G. A new species of Scutacarus (Acari: Heterostigmata: Scutacaridae) associated with *Trinervitermes trinervoides* (Isoptera: Termitidae) from South Africa. *Int. J. Acarol.* **44**, 59–67 (2018).
32. Alotaibi, N., Mashaly, A., Alajmi, R., Ahmed, A. & Ayaad, T. Genetic diversity of termites from Ta'if city Saudi Arabia | KSU Faculty. (2018).
33. Bordereau, C., Clément, J.-L., Jequel, M. & Vieau, F. *Les termites: Biologie lutte réglementation*. (Centre Technique du Bois et de l'Ameublement, 2003).
34. Ke, Y., Wenjing, Wu., Zhang, S. & Li, Z. Morphological and genetic evidence for the synonymy of reticulitermes species: Reticulitermes dichrous and Reticulitermes guangzhouensis (Isoptera: Rhinotermitidae). *Fla. Entomol.* **100**(1), 101–108. <https://doi.org/10.1653/024.100.0115> (2017).
35. Hellemans, S., Deligne, J., Roisin, Y. & Josens, G. Phylogeny and revision of the 'Cubitermes complex' termites (Termitidae: Cubitermitinae). *Syst. Entomol.* **46**, 224–238 (2021).
36. Josens, G. & Makatia Wango, S. Niche differentiation between two sympatric Cubitermes species (Isoptera, Termitidae, Cubitermitinae) revealed by stable C and N isotopes. *Insects* **10**, 38 (2019).
37. Tokuda, G. *et al.* Major alteration of the expression site of endogenous cellulases in members of an apical termite lineage: Evolution of cellulose digestion in termites. *Mol. Ecol.* **13**, 3219–3228 (2004).
38. Inward, D. J. G., Vogler, A. P. & Eggleton, P. A comprehensive phylogenetic analysis of termites (Isoptera) illuminates key aspects of their evolutionary biology. *Mol. Phylogenet. Evol.* **44**, 953–967 (2007).
39. Boyer, P. H. *et al.* Identification of closely related Ixodes species by protein profiling with MALDI-TOF mass spectrometry. *PLoS ONE* **14**, e0223735 (2019).
40. Sands, W. A. The Identification of Worker Castes of Termite Genera from Soils of Africa and the Middle East (1998).
41. Clay, T. *Bulletin of the British Museum (Natural History) Botany*. vol. Vol. 24 ((Natural History) Entomology, 1969).
42. Clay, T. *Bulletin of the British Museum (Natural History) Botany*. Vol. XXIV 1–452 (The Museum, 1971).
43. Binder, E. *Le Parc National du Niokolo-Koba (Sénégal)*. Vol. IV (1969).
44. Bouillon, G. & Mathot, A. *Quel est ce termite africain? Edition de l'Université Léopoldville: Léopoldville, Congo* (1965).
45. Meyer, C. P. Molecular systematics of cowries (Gastropoda: Cypraeidae) and diversification patterns in the tropics. *Biol. J. Linn. Soc.* **79**, 401–459 (2003).
46. Laroche, M. *et al.* MALDI-TOF MS as an innovative tool for detection of Plasmodium parasites in Anopheles mosquitoes. *Malar. J.* **16**, 1–10 (2017).
47. Tandina, F. *et al.* Use of MALDI-TOF MS and culturomics to identify mosquitoes and their midgut microbiota. *Parasit. Vectors* **9**, 495 (2016).
48. El Hamzaoui, B. *et al.* Detection of Bartonella spp. in fleas by MALDI-TOF MS. *PLoS Negl. Trop. Dis.* **12**, e0006189 (2018).
49. Diarra, A. Z. *et al.* Molecular and MALDI-TOF identification of ticks and tick-associated bacteria in Mali. *PLoS Negl. Trop. Dis.* **11**, e0005762 (2017).
50. Boyer, P. H. *et al.* Assessment of MALDI-TOF MS biotyping for *Borrelia burgdorferi* sl detection in *Ixodes ricinus*. *PLoS ONE* **12**, e0185430 (2017).

Acknowledgements

We thank the field crew for providing termite samples. We are grateful to Abdoulaye Baila Ndiaye for his help. We also thank our laboratory colleagues for their generous help and support in this study.

Author contributions

Conceived and designed the experiments: A.Z.D., P.P. O.M. Performed the experiments: B.H., A.Z.D. Analysed the data: A.Z.D., B.H. Contributed reagents/materials/analysis tools: B.H., A.Z.D., O.M., J.M.B., H.M., A.B. Took pictures: J.M.B., B.H. Acquired funding: P.P. Project administration: P.P., A.Z.D., O.M. Supervision: P.P. Drafted the paper: B.H., A.Z.D. Critically reviewed the paper: P.P., A.Z.D., O.M. All authors reviewed and approved the final version.

Funding

This study was supported by the Institut Hospitalo-Universitaire (IHU) Méditerranée Infection, the National Research Agency under the “Investissements d’avenir” programme, Reference ANR-10-IAHU-03, the Région Provence Alpes Côte d’Azur and European ERDF PRIM1 funding. BH received a grant of PhD scholarship from IHU Méditerranée Infection. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-04574-0>.

Correspondence and requests for materials should be addressed to P.P.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022