International Journal of Entomology Research
ISSN: 2455-4758; Impact Factor: RJIF 5.24
Received: 18-05-2020; Accepted: 20-05-2020; Published: 09-06-2020
www.entomologyjournals.com
Volume 5; Issue 3; 2020; Page No. 161-169

Revisiting DNA barcoding of ants (Formicidae: Hymenoptera) from India

Surya P[1], Sebastian CD[2]
[1, 2] Molecular Biology Laboratory, Department of Zoology, University of Calicut, Malappuram, Kerala, India

Abstract
Partial sequencing of mitochondrial cytochrome c oxidase subunit I (COI) gene was done for 15 ant specimens representing 11 species of Formicidae. Out of the 828 species of ant accounted from India so far, only less than twenty species were properly barcoded and published. The present study revisits the barcodes of ants from India. Of the 11 barcodes generated by the present study, 6 sequences are new barcode from India and out of these 3 sequences are the novel deposits to the GenBank. The DNA barcodes generated were compared and analysed with existing conspecific, congeneric and other ant barcodes reported from India. Divergence analysis was performed within and among various taxa. Mean interspecific distance (6.8%) was approximately twice greater than mean intraspecific divergence (2.5%). Two of the unambiguous earlier barcodes from India (that of *Paratrechina longicornis* and *Technomyrmex albipes*) were reanalyzed and compared with barcodes generated from the present study.

Keywords: Formicidae, Indian ants, mitochondrial COI gene, DNA barcode, K2P divergence

1. Introduction
The terrestrial biodiversity has ants as one of its most important component. Ants are one of the successful insect groups since they are omnipresent in almost all major terrestrial habitats (Holldobler and Wilson, 1990) [11]. According to the latest report, there are around 16,301 species described so far (http://www.antweb.org). Family Formicidae, is considered as the most species rich taxa of all social insect groups (Bolton *et al*. 2006) [4]. They are well discussed in the world of science on its social behaviors, ecology and associations with other organisms. The history of ant classification and taxonomy is too long and an abstract of this is well worked out in Brown (1955) [6] and is later updated in Bolton (2003) [3]. Though the ant systematics has come a long way, the diversity of these in its species level and the phylogenetic relationship between the groups is not well elucidated. Majority of the ant systematic works were focused on the worker caste which in turn is because of its easy availability in work fields but there are chances of misidentifications when the workers of closely related species are compared morphologically. Being a group showing high intraspecific morphological variations and worker polymorphisms, identification of ants to its species level is a tough task when compared to other group of insects. The recent interventions of molecular methods in taxonomy is now making the job of inferring phylogenetic relationships far easier and is also complementing the classical taxonomy (Ward, 2007) [19]. The state of ant systematics is being reviewed continuously in various parts of the world.
India, a country considered as mega-diverse, is being studied for its highly diverse flora and fauna. Though 828 valid species of ants belonging to 100 genera are listed from India with its geographical distributions (Bharti *et al*., 2016) [1], there are few studies from India elucidating the molecular details or DNA barcodes of these species. Barcoding of Formicidae is being done from 122 countries in the world and of these BOLD contains 73,583 identifiable (possessing species details) records representing 6142 species. Out of these BOLD has only 31 identifiable records from India.

2. Materials and Methods
Ants under study were collected during a period from November 2015 to December 2018. The sampling field spans various locations of six districts in Kerala state of India, which includes Thrissur, Malappuram, Kozhikode, Wayanad, Kannur and Kasaragod. Ants were collected by hand picking, pit fall traps, light traps and bait method (Brandon *et al*., 2000) [5]. The collected ants were then transferred to tubes containing 95% alcohol. The specimens were then cleaned and sorted. Each of the sorted morphospecies was then identified up to the species level with the help of ant taxonomic keys (Bingham, 1903) [2] and were reviewed and confirmed by expert consultation. Vouchers of identified specimen were documented and stored at ⁻20° C in Molecular Biology Laboratory, Department of Zoology, University of Calicut, Kerala.
DNA from the selected ant specimens was isolated using the commercially available DNA extraction kits following the respective instructions. The presence of DNA after the isolation procedure was confirmed by performing Agarose (1%) Gel Electrophoresis. On confirming the presence of DNA in the isolates, PCR amplification of the DNA samples were performed (Hebert *et al*., 2003) [10] using forward and reverse primers, LCO1490: 5'-ggtcaacaaatcataaagatattgg-3' and HC02198: 5'-taaacttcagggtgaccaaaaaatca-3', respectively (Folmer *et al*., 1994) [8]. The presence of amplified DNA in PCR product was confirmed using 2% Agarose gel electrophoresis. The nucleotide sequence of the PCR amplified region were sequenced by Sanger's method using automated sequencer ABI 3730XL (Sanger and Coulson, 1975) [18]. Sequencing resulted in 15 cytochrome oxidase subunit 1 (COI) gene sequences with length range of 444 – 703 bp size. Each of the sequences were then analysed for their translation

products and confirmed for any gaps or non-sense codons. The final sequences were submitted to NCBI GenBank and accession numbers were obtained for the 15 sequences representing 11 species of ants. GenBank database were then searched for nearest neighbours of the deposited sequences that are either congeneric or conspecific with the sequences under study (Table 1). Further barcode sequences of ants from India (Ojha *et al*., 2014) [15], were retrieved from GenBank database for the analysis (Table 2). COI barcode sequences of *Vespa tropica*, *Phimenes flavopictum* and *Rhynchium brunneum* (Vespidae: Hymenoptera) were selected as outgroup. Sequences generated in the present study were then aligned with their nearest neighbor sequences, other congeneric or conspecific sequences (Table 1) and the already available barcodes of Indian Ants, with the help of BioEdit software (Hall, 1999) [9].

The aligned sequence data were worked out in MEGA X software (Kumar, 2018) [14] for the analysis of evolutionary status of selected ant species of India. The sequence divergence within the species and between the representative species, genus and subfamilies under study were also inferred using Kimura 2 parameter model (Kimura, 1980) [3]. Thus a clustering of all the selected sequences is produced by Neighbour Joining (NJ) method (Saitou and Nei, 1987) [17] adopting Kimura 2 Parameter with a nodal support test of 1000 bootstrap replication (Felsenstein, 1985) [7]. With this, the reliability of the dataset was checked and the final selected sequences were then used to construct the tree by Maximum Likelihood (ML) and Bayesian Inference (BI) (Huelsenbeck *et al*., 2001) [12]. Construction of Bayesian Inference tree was done using the UGENE graphical user interface tool (Okonechnikov, 2012) [16] in which the model used is GTR+G+I selected by the Bayesian Inference Criterion (BIC). The maximum likelihood tree was also constructed using the GTR + G+ I with lowest BIC and Akaike Information Criterion (AIC) as suggested by the best fitting model tool of MEGA X. A total of 50 sequences are compared including the 3 sequences selected to form the out group cluster in the Phylogenetic tree.

3. Results

BLAST search of every sequences produced in this study resulted in matching with species of same genus (or same subfamily, in the absence of sequences representing same genus). This proves the utility of these sequences in finding the genus if the database has enough representations of any particular genus.

The final alignment is of 50 sequences of which 15 sequences were newly generated in the present study. Partial COI sequences of 11 species of ants belonging to 11 genera were obtained. These sequences represent 5 subfamilies in Formicidae: Myrmicinae, Formicinae, Pseudomyrmecinae, Ponerinae and Dolichoderinae. Except *Myrmicaria brunnea*, *Carebara diversa*, *Anoplolepis gracilipes*, *Technomyrmex albipes* and *Paratrechina longicornis* all other six sequences were new deposits to barcode database from India. Three of these barcodes – sequences of *Brachyponera luteipes* (Figure 1.1), *Leptogenys peuqueti* (Figure 1.2) Polyrhachis *hippomanes ceylonensis* (Figure 1.3), were first submission to GenBank representing the above respective species.

As in case of typical arthropods, a high A-T (69%) content was observed with A = 29.8%, G = 11.5%, C = 19.4%, T = 39.2%. The K2P mean divergence among different taxonomic levels showed a hierarchical increase in values (Table 3). The mean intraspecific divergence observed is 2.5% with a range of 0-4% whereas interspecific mean divergence was observed as 6.8% with a range of 1 – 17%. The divergence between genera within a subfamily has an average of 16% with a range of 2 – 24%. The sequence divergence distance between the subfamilies was in the range of 8.75 – 36.54% with an average of 19.5%. The barcode gap found with divergence in sequences was checked as they clustered in the tree constructed using NJ method (Figure 2). The overlaps in sequence divergence at every taxonomic level were resolved by constructing phylogenetic trees using Maximum Likelihood (Figure 3) and Bayesian Inference methods (Figure 4). Clustering of the congeneric and conspecific taxa was found cohesive in all trees whereas at the subfamily level there was slight difference in congruency. Dolichoderinae and Ponerinae showed monophyly whereas Myrmicinae, Pseudomyrmecinae and Formicinae showed sort of paraphyletic relation within the subfamilies in all the three phylogenetic trees.

Table 1: List of specimens analysed in the present study with their GenBank accession and percentage of sequence similarity (BLAST) to the nearest neighbour in the database

S. No.	Name of species	Voucher No.	GenBank Accession No.	GenBank accession of nearest neighbour	Percentage of similarity
1	*Carebara diversa*	CUAN7	MN599029	MN607156	99.12%
2	*Carebara diversa*	CUANZ3	MN607156	MN599029	99.12%
3	*Trichomyrmex mayri*	CUANZ2	MN607154	KJ847497	96.49%
4	*Myrmicaria brunnea*	CUAN8	MN590430	MN607152	99.85%
5	*Myrmicaria brunnea*	CUANV3	MN607155	MN590430	100%
6	*Myrmicaria brunnea*	CUAN99	MN607152	MN590430	99.85%
7	*Tetraponera rufonigra*	CUAN4	MN234143	BK010387	99.55%
8	*Paratrechina longicornis*	CUANV6	MN613240	KP232173	97.83%
9	*Anoplolepis gracilipes*	CUANc	MN590429	NC039576	98.54%
10	*Polyrhachis hippomanes ceylonensis*	CUANZ6	MN613239	KM348215	86.71%
11	*Leptogenys peuqueti*	CUANV2	MN590431	KY000663	88.80%
12	*Brachyponera luteipes*	CUANV4	MN613241	MF673717	88.74%
13	*Odontomachus simillimus*	CUAN2	KX587511	MN613242	100%
14	*Odontomachus simillimus*	CUANV5	MN613242	KU504909	100%
15	*Technomyrmex albipes*	CUANZ5	MN234144	HQ925235	100%

Table 2: Accession numbers of COI sequences (other than nearest neighbours) of Indian ants (Ojha *et al*., 2014) retrieved from NCBI GenBank

S. No.	Name of Species	GenBank Accession No.
1	*Myrmicaria brunnea*	JN886029
2	*Carebara diversa*	JN987859
3	*Anoplolepis gracilipes*	JN987860
4	*Technomyrmex albipes*	JN886038
5	*Paratrechina longicornis*	JN886034
6	*Monomorium destructor*	GU709851
7	*Trichomyrmex destructor*	GU710447
8	*Anoplolepis gracilipes*	NC039576
9	*Tetraponera rufonigra*	MF716964
10	*Paratrechina longicornis*	KP232061
11	*Paratrechina longicornis*	FJ982476
12	*Polyrhachis aurea*	KM348211
13	*Polyrhachis ornata*	KM348255
14	*Odontomachus laticeps*	HM919582
15	*Odontomachus bauri*	KU985513
16	*Pachycondyla sp*	MF673719
17	*Pachycondyla sp*	MF673716
18	*Leptogenys angusta*	HM418986
19	*Leptogenys alatapia*	HM419328
20	*Leptogenys lavavava*	HQ925205
21	*Leptogenys truncatirostris*	HM418996
22	*Technomyrmex vitiensis*	KY000670
23	*Technomyrmex albipes*	EF610584
24	*Technomyrmex albipes*	EF610581
25	*Technomyrmex albipes*	EF610580

Table 3: Sequence divergence (K2P) at various taxonomic levels

Sequence divergence at various taxonomic levels	Range (%)	Mean Distance (%)
Within species	0 – 4	2.5
Amongst species	1 – 17	6.8
Amongst genera	2 – 24	16
Amongst subfamilies	8.75 – 36.54	19.5

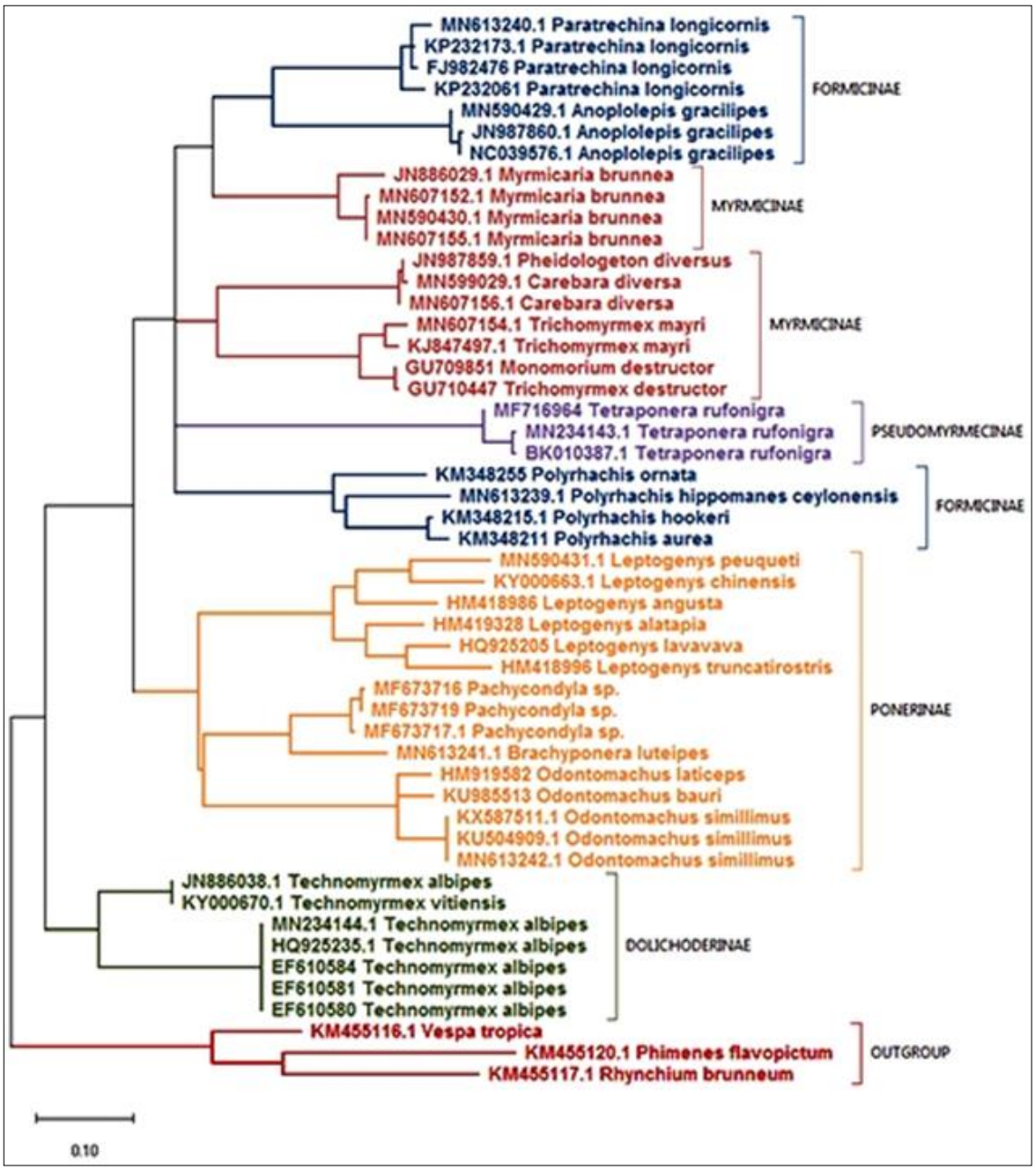

Fig 1: Illustrations of three newly barcoded ant species. (1.1) *Brachyponera luteipes*, (1.2) *Leptogenys peuqueti and (1.3) Polyrhachis hippomanes ceylonensis.*

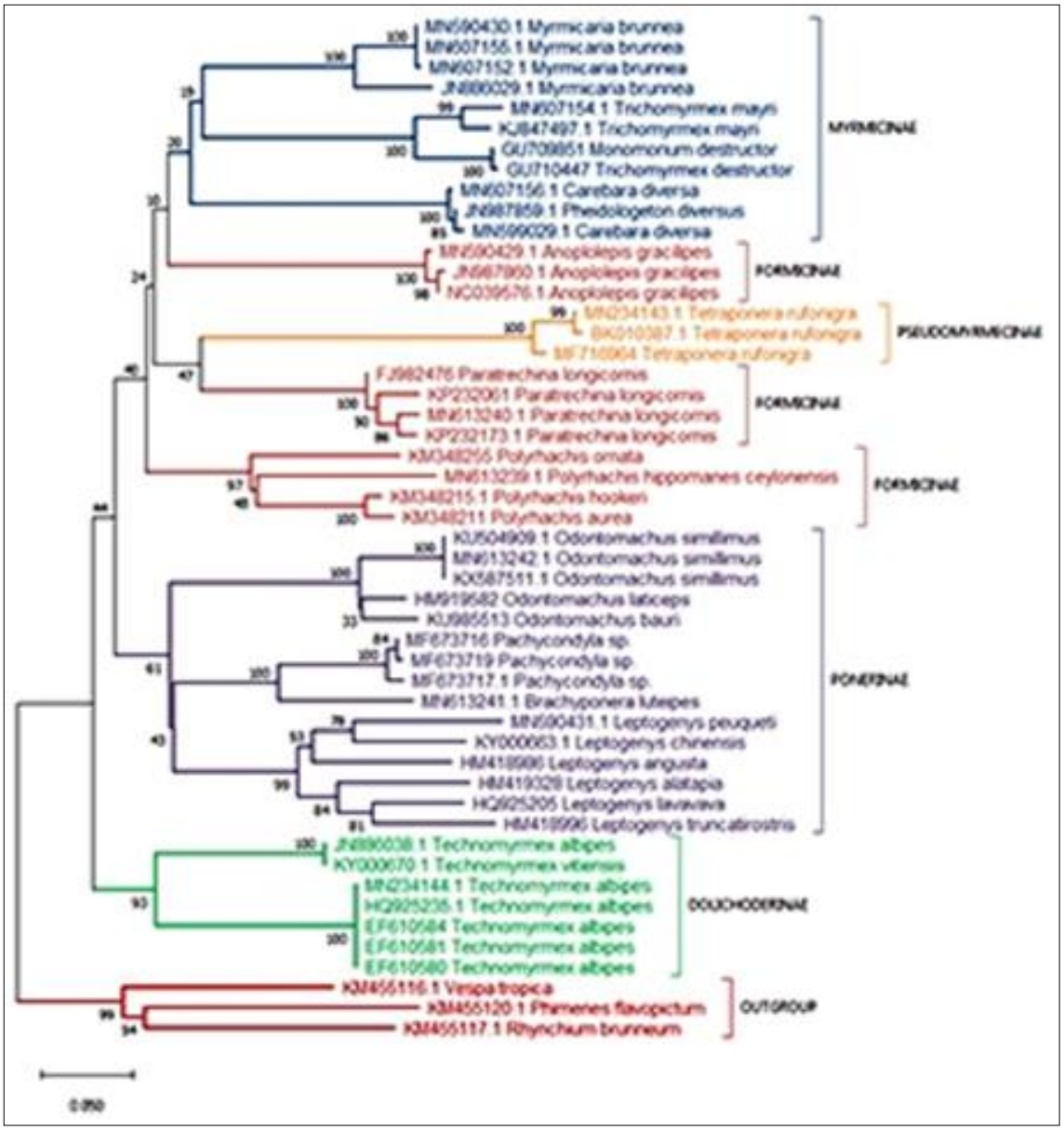

Fig 2: Neighbour Joining representation of the 50 sequences under analysis including outgroup. Different subfamilies are coloured differently

Figure 1.1 *Brachyponera luteipes*,
(A) Body in lateral view, (B) Head in full face view (C) Body in dorsal view

Figure 1.2 *Leptogenys peuqueti*
A) Body in lateral view, (B) Head in full face view (C) Body in dorsal view

Figure 1.3 *Polyrhachis hippomanes ceylonensis*
(A) Body in lateral view, (B) Head in full face view (C) Body in dorsal view

Fig 3: Maximum likelihood representation of the 50 sequences under analysis including outgroup. Different subfamilies are coloured differently.

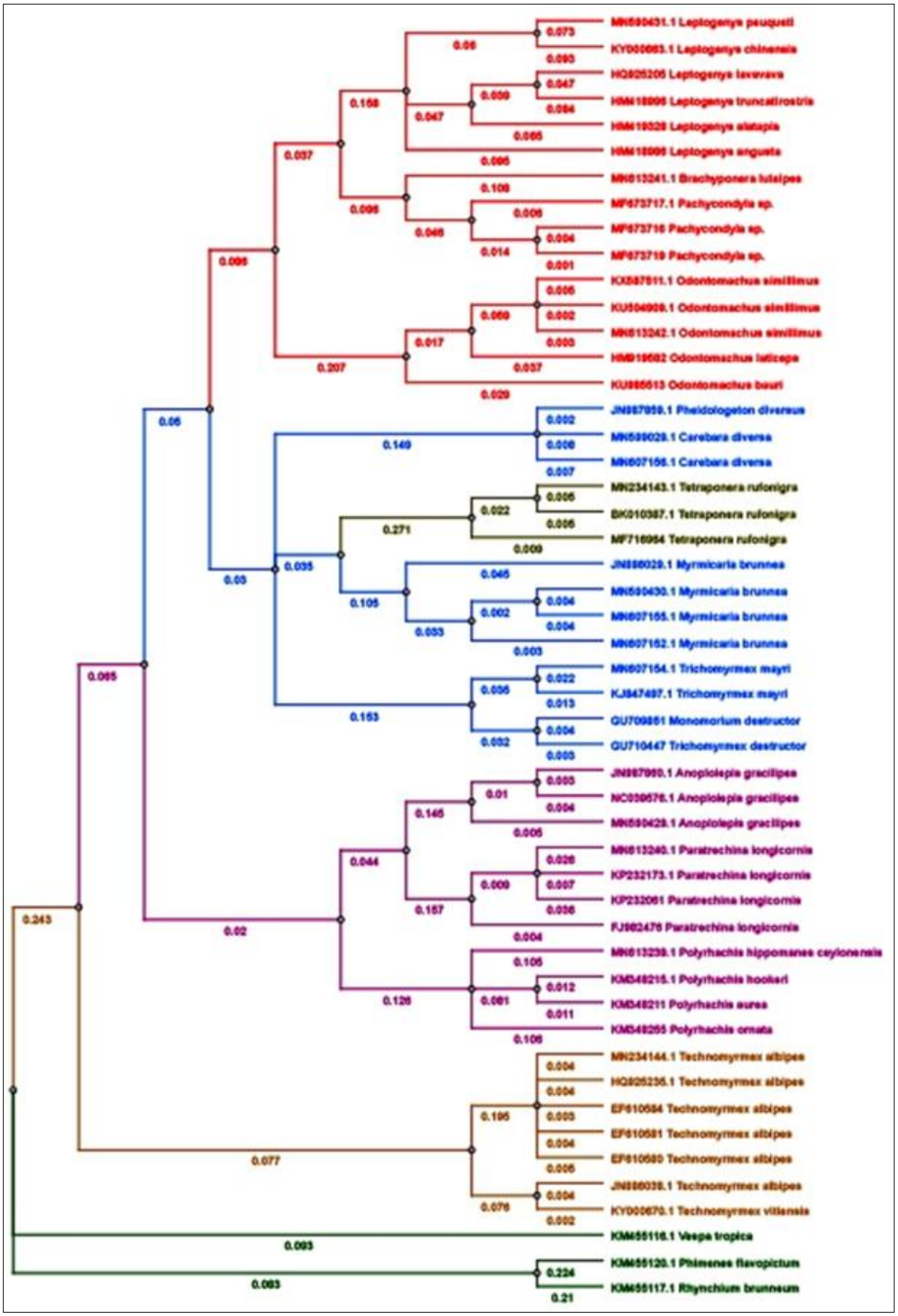

Fig 4: Phylogenetic relationship representation of 50 sequences under analysis including outgroup using Bayesian Inference. Different subfamilies are coloured differently.

4. Discussion

The tropics support the enormous biodiversity in India and hence the generation of barcodes of Indian fauna is important. There is only a single published study on DNA barcoding of Indian ants using mitochondrial COI sequence till date. Though this study elucidated the mitochondrial COI sequence for 16 ant species collected from Karnataka, India (Ojha *et al*., 2014) [15], these sequences were reviewed and rechecked in the present study. Out of the 15 sequences, representing 11 species, elucidated by the present study; 5 species were already barcoded from India. These 5 barcode sequences from India (with accession numbers JN886029,

JN987859, JN987860, JN886034, JN886038) were not shown as the nearest neighbour when the alignment search for the respective sequences from present study representing same species was done using BLAST. On further check these 5 barcodes and the set of other 11 barcodes produced from the same study were commented unverified by the GenBank. Three of these sequences, JN886029, JN987860 and JN886038 representing *Myrmicaria brunnea*, *Anoplolepis gracilipes* and *Technomyrmex albipes* respectively were found as a different lineage though clustered with conspecific sequences of the present study. The JN886038 suggested as barcode of *Technomyrmex albipes* was found to be 99.68% similar to *Technomyrmex vitiensis* species barcode from India with accession number KY000670 and these sequences delineated separately from the conspecific clade of *Technomyrmex albipes* in phylogenetic trees. *Technomyrmex albipes* are distinguished from *Technomyrmex vitiensis* in that they possess comparatively short scape, short and broad mesonotum with convex mesonotal dorsal outline, blunt or non-angular propodeal dorsum and declivity. All these character difference represented by less than 1% sequence divergence indicates the barcode sharing among these species. This can happen as both of these species belong to a species complex that is hard to identify. Hence the *Technomyrmex albipes* barcode generated by the present study shall be proposed as the valid barcode of this species from India. Another unambiguous barcode deposit from India is that of *Paratrechina longicornis* with accession number JN886034, for which a second barcode is deposited in the present study. The JN886034 on search with BLAST tool gave *Lepisiota frauenfeldi* as the nearest neighbor with 98.69% similarity. All the first seven nearest neighbours of this sequence were the same species that is, *Lepisiota frauenfeldi*. The *Paratrechina longicornis* barcode generated in the present study gave 97.96% similarity to another barcode representing same species and all the first seven nearest neighbours were the same species confirming the correct identification of the specimen. Accordingly the other barcode deposit from India could be resulted from an incorrect identification or contamination. The coherent clustering of conspecific and congeneric sequences in the phylogenetic tree further confirms the efficiency of COI gene in molecular identification of Formicidae. All the terminal nodes of phylogenetic trees shows high nodal value and this decrease as the nodes get deeper. For sister species the nodes exhibited extremely high nodal support values.

5. Conclusion

It is already reported that COI is best at explicating species level connections in arthropods, but resolving relationships beyond family or further cannot be done perfectly with COI (Waugh, 2007) [20]. Though the NJ, ML and Bayesian tree generated in this study shows paraphyly within and between few subfamilies, it cannot be accounted as such since COI gene fragment comparison alone cannot resolve the phylogenetics and evolutionary history. A discussion on phylogeny of ants is not aimed here whereas the molecular identification of ants with COI as a tool is emphasized by the present study. This study could not generate multiple conspecific sequences for all the species collected. The authors plan to generate multiple sequences for each and every species that have barcoded so that this would be helpful for further researches focusing on phylogeography and population genetics. Barcode generation of many other ant species from the study area will follow as a continuation of the present study. As far as every branch of biology is concerned, a comprehensive DNA database is very important.

6. Acknowledgements

This work was supported by Kerala State Council for Science, Technology and Environment under the financial assistance in the form of research fellowship. Authors thank Dr. Paul D. N. Hebert and Dr. Mikko Pentinsaari for clarification on the validity of unverified barcodes in GenBank. The supports received from Dr. Aniruddha Marathe (Ant taxonomist), Dr. Priyadarshanan Dharmarajan (Senior Fellow) and Mr. Aswaj Punnath (Researcher) of Ashoka Trust for Research in Ecology and Environment, Bangalore, in specimen identification and Mr. Ranjith AP (Researcher) in collection and identification of specimens are gratefully acknowledged.

7. Disclosure statement

The authors declare that they have no competing interests.

8. References

1. Bharti H, Guenard B, Bharti M, Economo EP. An updated checklist of the ants of India with their specific distributions in Indian states (Hymenoptera, Formicidae). *ZooKeys*, 2016; 551:1-83.
2. Bingham CT. The fauna of British India, including Ceylon and Burma. Hymenoptera. Ants and Cuckoo-Wasps. London: Taylor and Francis. 1903; 2:506.
3. Bolton B. Synopsis and classification of Formicidae. Memoirs of the American Entomological Institute, 2003; 71:1-370.
4. Bolton B, Alpert G, Ward PS, Nasrecki P. Bolton's Catalogue of ants of the world. Harvard University Press, Cambridge, Massachusetts, 2006, 1758-2005.
5. Brandon TB, Donat A, Leeanne EA, Roberto C, William FBLB Jr., Jacques HCD *et al*. Field techniques for the study of ground-dwelling ants: An overview, description, and evaluation, Book: ANTS, Standard methods for measuring and monitoring Biodiversity, Edited by Donat Agosti, Jonathan D. Majer, Leeanne E. Alonso and Ted R Schultz, 2000.
6. Brown WL Jr. Ant taxonomy. *In:* Kessel, E.W. (Ed.) A century of progress in the natural sciences, 1853–1953. California Academy of Sciences, San Francisco, 1955, 569-572.
7. Felsenstein J. Confidence limits on phylogenies: An approach using the bootstrap. Evolution. 1985; 39:783-791.
8. Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R. DNA primers for amplification of mitochondrial Cytochrome C oxidase subunit I from diverse metazoan invertebrates. Molecular marine biology and biotechnology. 1994; 3:294-9.
9. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic acids symposium series. 1999; 41(41):1979-2000.
10. Hebert PDN, Cywinska A, Ball SL, de Waard JR. Biological identifications through DNA barcodes, Proceedings of the Royal Soceity B: Bioogical Sciences. 2003; 270:313-321.

11. Holldobler B, Wilson EO. The ants. Harvard University Press, Cambridge, Massachusetts. 1990; xii+732pp.
12. Huelsenbeck J, Ronquist F, Nielsen R, Bollback J. Bayesian Inference of Phylogeny and Its Impact on Evolutionary Biology. Science. 2001; 294:2310-2314.
13. Kimura M. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. Journal of Molecular Evolution. 1980; 16:111-120.
14. Kumar S, Glen S, Michael L, Christina K, Tamura K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Molecular Biology and Evolution. 2018; 35:(6):1547-1549.
15. Ojha R, Jalali SK, Ali TM, Venkatesan T, Prosser SW, Kumar NK. DNA barcoding of Indian ant species based on COX1 gene. IJBT. 2014; 13(2):165-171.
16. Okonechnikov K, Golosova O, Fursov M. Unipro UGENE: a unified bioinformatics toolkit. Bioinformatics. 2012; 28:1166-1167.
17. Saitou N, Nei M. The neighbor-joining method: A new method for reconstructing phylogenetic trees. Molecular Biology and Evolution, 1987; 4:406-425.
18. Sanger F, Coulson AR. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. Journal of Molecular Biology. 1975; 94(3):441-448.
19. Ward PS. Phylogeny, classification, and species-level taxonomy of ants (Hymenoptera: Formicidae). Zootaxa. 2007; 1668(1):549-563.
20. Waugh J. DNA barcoding in animal species: progress, potential and pitfalls. BioEssays. 2007; 29(2):188-197.