Journal of Advanced Zoology

ISSN: 0253-7214

Volume ***45*** *Issue* ***S-2 Year 2024*** *Pages* ***297- 304***

Study the genetic diversity of honey bee species in Nashik district of Maharashtra, India

A. Sonawane[1*] and B. Waykar[2]

[1, 2] Department of Zoology, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, Maharashtra, India*

**Corresponding author's E-mail: sonawaneaparna85@gmail.com*

Abstract

This study was planned to estimate the genetic diversity of honeybee species in Nashik district. Because of climatic variation different crops are taken all over the district. In this study, genetic diversity between Apis dorsata, Apis cerana, Apis florea and Apis mellifera is estimated from the study area. For this, DNA barcoding of samples was performed. Sequences were uploaded to NCBI and accession id numbers were generated. Data was subjected to MEGA X software which was used for computing the evolutionary divergence, transition\transversion bias. Kimura 2 model is used for data analysis. Divergence between honeybee species in Nashik district is ranges from 0.00479 to 0.148774. Maximum divergence is observed in Apis dorsata and Apis florea i.e. 0.148774. Neighbor- joining phylogenetic tree revealed close relatedness between Apis dorsata - Apis mellifera and Apis cerana – Apis florea. Estimated evolutionary divergence in Trimbakeshwar tehsil ranges from 0.102 to 0.130. Apis dorsata and Apis mellifera showing minimum divergence i.e. 0.102 and Apis dorsata and Apis florea show maximum divergence i.e. 0.130. From this we can conclude that Apis dorsata and Apis mellifera are closely related while Apis dorsata and Apis florea are distantly. From Malegaon tehsil, the maximum divergence is 0.149 and minimum is 0.106.

Keywords: *Apis dorsata, cerana, florea, mellifera, genetic, DNA barcoding*

1. Introduction

Honey bee is an important insect for mankind as well as the environment because of beneficial products like honey, wax, propolis and its prime role in pollination. In crop pollination and wild pollination, honeybees have a crucial role. Honeybee influences floral biodiversity and conservation is estimated to affect 80% of wild flora through pollination. India is good in agriculture and takes different crops on a large scale because of suitable climatic conditions. The significance of honeybees in behavioural research lies in their colonial nature, intricate social behaviour, and vast genetic diversity, which remains largely untapped. Molecular marker techniques have the potential to explore this genetic diversity effectively (Dempster and McLean, 1999; Alattal et al., 2019). India, with its diverse crops influenced by climatic variations, is home to four main honeybee species: *Apis dorsata*, *Apis cerana*, *Apis florea*, and the introduced European species, *Apis mellifera*, for commercial apiculture since 1962.

Identifying honeybee subspecies is crucial for both conservation and breeding purposes. DNA barcoding emerges as a valuable tool for swift species identification, complementing traditional taxonomy in biodiversity studies. Despite India's rich insect diversity, only a small percentage (3.73%) has been barcoded among the described 62,429 insect species from 595 families (Shankarganesh E, 2017). India, covering 2% of the Earth's landmass, stands among the top ten mega-diverse nations, hosting 7.10% of the world's insect fauna. However, India contributes only 1.53% to global DNA barcoding efforts, with Costa Rica leading at 77%, followed by South Africa, China, and Mexico.

In India, DNA barcoding efforts have primarily focused on Lepidoptera (26.08%), with hymenoptera representing only 6.40% of sequences (Sreedevi K, et al., 2015). Notably, experts predominantly concentrate on orders like Lepidoptera, Hemiptera, Diptera, and Coleoptera, which collectively account for 80% of the barcoded sequences (Sreedevi K, et al., 2015). While genetic studies extensively cover the western honeybee *A. mellifera* globally due to its importance in apiculture, genetic investigations on *A. dorsata* and *A. cerana* have also been conducted to some extent.

The westward expansion of the Dwarf honeybee, *A. florea*, originally endemic to South East Asia, has been largely unnoticed. Therefore, this study aims to comprehensively analyze all four honeybee species using DNA barcoding, shedding light on their genetic makeup and contributing to the broader understanding of their ecological and evolutionary dynamics.

2. Materials and Methods

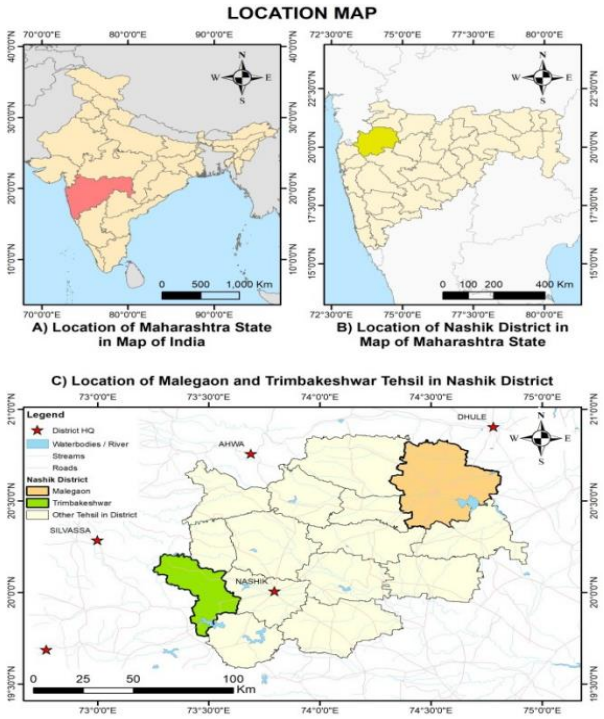
2.1 Study Area

Nashik district was selected as study area because of climatic variation; it is like a mini Maharashtra. From this region, Trimbakeshwar and Malegaon tehsils with different geography were selected. Trimbakeshwar and Malegaon are located oppositely in Nashik district with extreme climatic difference. Distance between both tehsils is 137.8 km. Trimbakeshwar receives approximately 1564 mm of precipitation while Malegaon gets only 82.92mm of precipitation every year. Satmala mountain range which is an integral part of Sahyadri is present in Nashik district. Mountain range occupies a large area in Trimbakeshwar tehsil. Comparatively Malegaon tehsil is plain. Maps of both tehsils show difference in availability of water bodies, poor, moderate and dense vegetation. Due to different climatic conditions, flora and fauna are varied in both regions. Vegetation index of both the tehsils also varies distinctly. That is why these regions are considered to study the genetic variation of honey species; *A. dorsata, A. cerana, A. florea* and *A. mellifera* by DNA barcoding method. Location of sample collection sites and vegetation indices are shown in the tehsil maps. In map, location of water bodies and different vegetation (poor, moderate and dense) are clearly visible which is varied in both tehsils on large scale.

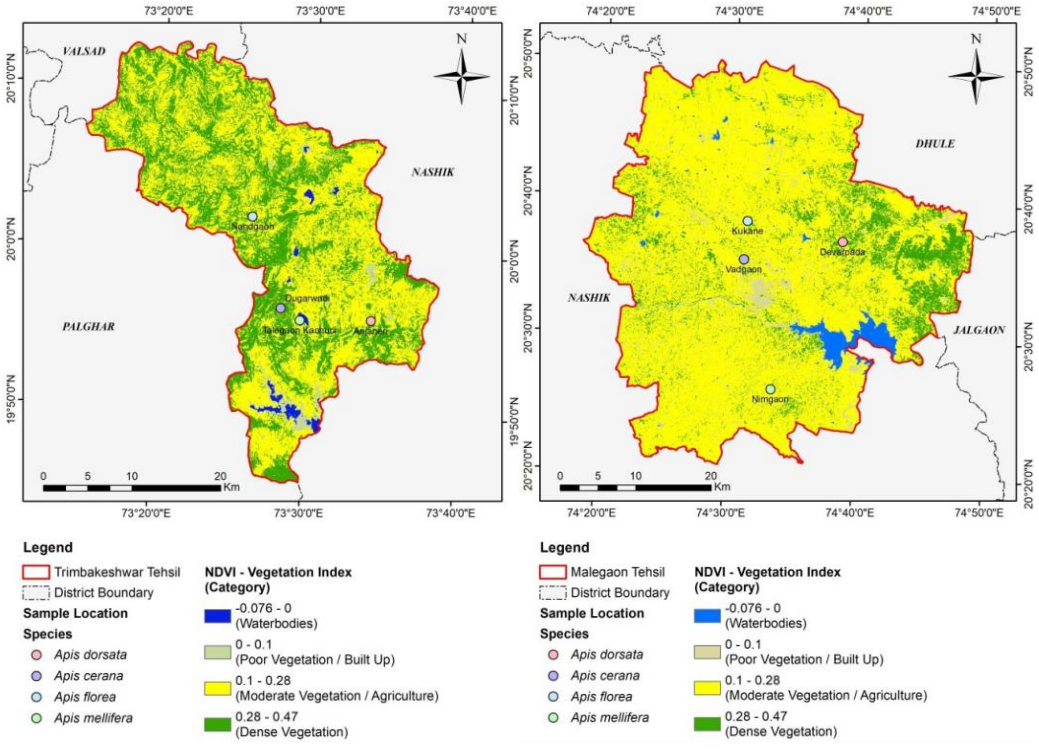
2.2 Sample collection- Honey bee samples were collected from different localities of both tehsils. Table no. 1 shows details of sample collection sites. By using insect collecting aerial net and transferred to insect killing bottle containing cotton wad soaked with chloroform and preserved in 100% alcohol at -40^{0}C in laboratory till *DNA* barcoding.

Sr no.	Sample code of honeybees	Location	Latitude	Longitude
Trimbakeshwar Tehsil				
1	A1	Anjaneri	19.93170477	73.57006608
2	A2	Dugarwadi	19.94095556	73.47036989
3	A3	Nandgaon	20.03446674	73.43507143
4	A4	Talegaon Kachurli	19.92905585	73.49194077
Malegaon Tehsil				
1	B1	Devarpada	20.61908799	74.64443723
2	B2	Vadgaon	20.59345052	74.51765938
3	B3	Kukane	20.63996948	74.52043361
4	B4	Nimgaon	20.43727687	74.55846867

Table 1 - Sample collection site details of Trimbakeshwar and Malegaon Tehsils from Nashik districts



Map 1 Location Map of Trimbakeshwar and Malegaon tehsils in Nashik district of Maharashtra, India



Map 2 and Map 3 shows Vegetation indexed Map of Trimbakeshwar and Malegaon tehsils in Nashik district

2.3 Methodology

DNA extraction and PCR amplification of given tissue samples:

DNA extraction- To create a uniform suspension, 1 mL of CTAB lysis buffer was combined with 10 mg of finely ground tissue, followed by the addition of 2 µL of Proteinase K. The samples were then subjected to an incubation period at 60 °C for 30 minutes. After this step, an equal volume of chloroform:isoamyl alcohol (24:1) was introduced, and the mixture was centrifuged. The subsequent step involved the gradual addition of chilled absolute alcohol (2.5 volumes) until DNA spooling was observed. To the spooled DNA, 300 µL of 70% chilled alcohol was added and centrifuged. The resulting pellet was treated with an additional 300 µL of absolute alcohol and subjected to further centrifugation. Following air drying, the DNA sample was dissolved in 150 µL of TE buffer and stored at 4°C. For the assessment of genomic DNA integrity, samples were loaded onto a 0.7% agarose gel and stained with ethidium bromide at a concentration of 5 µg µL-1. Once the integrity of genomic DNA was confirmed through gel electrophoresis, quantification of the samples was performed using Nanodrop.

Gel Electrophoresis- The extracted DNA (10µl each) was then subjected to 60 minutes gel electrophoresis to visualized bands under UV in 1% agarose gel stained with EtBr against Ladder Generuler 1kb.

DNA quantification- The DNA was quantified using nano-drop instrument. The quantity measured in ng/µL is shown in Table 3 below.

Polymerase Chain Reaction- PCR was conducted utilizing COI primers listed in Table 1. A total DNA amount of 100 ng was used for the PCR reaction, and the fixed reaction volume was set at 50 µL. The final concentrations of components in the reaction volume were as follows: 2.5 mM MgCl2, 1X Taq Buffer, 0.2 mMdNTPs, 0.5 µM primers, and 5 U Taq polymerase. The template DNA was maintained at standard concentrations, diluted to a final volume of 50 µL using nuclease-free distilled water.

The PCR procedure was carried out in an Applied Biosystems PCR thermocycler GeneAmp 9700, with the following parameters:

Sequence	Initial denaturation	Denaturation	Annealing	Extension	Final extension
FP:TCGTATAGAGCTCCGTCGACCTG	96°C, 5mins	95°C,1min	54°C,1min	72°C,1min	72°C, 10min
RP:TGGAACAAAACTGGATCGCC		35Cycles			

Table 2 Amplification of DNA using COI Primer

The PCR product was run in 1.5% agarose gel stained with EtBr against Ladder 100bp ruler and visualized under UV light.

3. Data analysis-

Sequencing, alignment- Bees are a part of biodiversity. Study of social insect honeybee has great value. Genetic study of *A. dorsata* was performed by using random amplified polymorphic DNA (RAPD) markers by Qamer et al., 2021. Genomic and transcriptomic analysis of the Asian honeybee *A. cerana* was studied in detail by Diao, Q. et al., 2018. AL attal et al., 2019, Yogesh and Khan, 2014, Issa et al., 2013 worked out genetic variation in *A. mellifera* (Qamer et al., 2021). Although some molecular studies of *A. mellifera* were conducted based on mitochondrial gene segment i.e., cytochrome c oxidase I (Rizwan M. et al., 2018). The exploration of species diversity at the molecular level serves as a blueprint, offering insights into gene expressions and intra-colonial genomic variations influenced by different ecological habitats (Shashank, P. R. et al., 2022). In this study, honey bee species samples underwent barcoding, and Sanger's sequencing was employed for the sequencing process. The resultant amplified sequences were submitted to the NCBI database, and published sequences relevant to the study were retrieved for further analysis. Subsequent to sequence alignment, phylogenetic trees illustrating the evolutionary divergence between sequences from both the tehsil area and each tehsil were constructed using Maximum Likelihood analysis. Genetic divergence was quantified through Maximum Likelihood analysis utilizing the Kimura-2 parameter model. Additionally, Disparity Index was computed to examine variations in base composition biases (Kumar S., and Gadagkar S.R., 2001). Various parameters such as GC content, pairwise distance, and Maximum Likelihood transition/transversion bias were evaluated using MEGA X software to gain a comprehensive understanding of the genetic relationships and evolutionary patterns within the studied honey bee populations.

4. Result and discussion- In this investigation, molecular studies were conducted utilizing the mitochondrial cytochrome oxidase subunit I (COI) gene. The Sanger's dideoxy sequencing methodology was

employed for the sequencing process. The quantified amounts of DNA utilized in the study are detailed in Table No. 3.

Sr. no.	Sample	DNA conc.(ng/µL)	Sequence number	Accession numbers	Base pairs
1	A1	154	SUB12870334Seq1	OQ439640	603bp
2	B1	134	SUB12870334Seq2	OQ439641	536bp
3	A2	187	SUB12870334Seq3	OQ439642	619bp
4	B2	156	SUB12870334Seq4	OQ439643	620bp
5	A3	124	SUB12870334Seq5	OQ439644	560bp
6	B3	95	SUB12870334Seq6	OQ439645	470bp
7	A4	131	SUB12870334Seq7	OQ439646	621bp
8	B4	124	SUB12870334Seq8	OQ439647	613bp

Table 3 DNA Quantification, sequence numbers and accession numbers generated (NCBI)

Accurate identification of insect species holds significant importance in addressing fundamental questions within the realms of ecology, evolution, agro biodiversity, and the conservation of biology (Sreedevi K, et al., 2015). Despite a steady increase in the total number of species documented in India, the growth rate remains notably sluggish (Dempster and McLean, 1999). Honeybees, being colonial insects with intricate social behavior, serve as essential model organisms for behavioral research. Their significance is further underscored by the vast untapped genetic diversity and gene pool, which can be more comprehensively explored through the application of molecular marker techniques (AL Attal et al., 2019). Fig. 1 shows graphical representation of base count of sequences from study area.

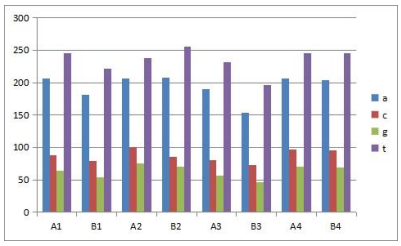


Fig. 1 Graphical representation of base count of sample sequences

Following the extraction of DNA, the samples underwent sequencing, and the resultant data were submitted to the NCBI Gene Bank, generating unique accession ID numbers as detailed in Table No. 3. The amplification of the Cytochrome Oxidase Subunit 1 region was successfully attempted, yielding PCR products that were subsequently sequenced using the Sanger's sequencing method. Figures 2 and 3 depict the genomic DNA of the samples, analyzed on a 0.8% agarose gel against a 1Kb ladder, and the PCR amplified products on a 1.5% agarose gel against a 100 bp ladder, respectively.

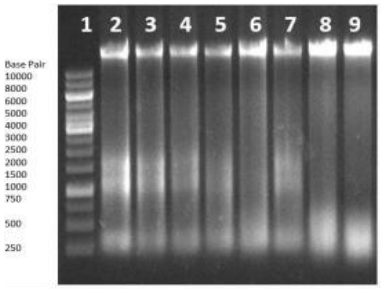


Figure 2

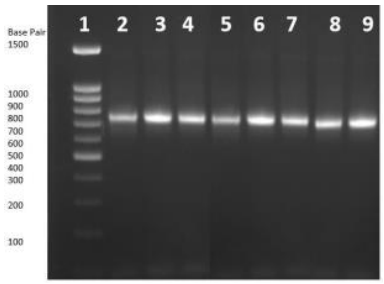


Figure 3

Figure 2 Shows Genomic DNA of the samples were analyzed on 0.8 % agarose gel stained with 0.1 µg/mL of Ethidium Bromide.Lane 1 shows standard 1 Kb ladder, Lane 2 shows gDNA extracted from sample A1, Lane 3 shows gDNA extracted from sample B1, Lane 4 shows gDNA extracted from sample A2, Lane 5 shows gDNA extracted from sample B2, Lane 6 shows gDNA extracted from sample A3, Lane 7 shows gDNA extracted from sample B3, Lane 8 shows gDNA extracted from sample A4, Lane 9 shows gDNA extracted from sample B4

*Figure **3** Shows PCR amplified products were analyzed on 1.5 % agarose gel stained with 0.1 µg/mL of Ethidium Bromide. Lane 1shows standard 100 bp ladder, Lane 2 shows PCR product of sample A1, Lane 3 shows PCR product of sample B1, Lane 4shows PCR product of sample A2, Lane 5 shows PCR product of sample B2, Lane 6 shows PCR product of sample A3, Lane7showsPCRproductofsample B3, Lane 8 shows PCR product of sample A4, Lane 9 shows PCR product of sample B4*

A Neighbor-Joining phylogenetic tree (Fig. 4) was crafted to illustrate the phylogenetic distances among all eight nucleotide sequences originating from both tehsils. Notably, *A. dorsata* and *A. mellifera* exhibited a close relationship, while a similar closeness was observed between *A. cerana* and *A. florea*. In contrast, *A. dorsata* and *A. mellifera* appeared distantly related. To delve deeper into the phylogenetic distinctions, separate NJ-phylogenetic trees (Fig. 5 and Fig. 6) were constructed for each tehsil, elucidating the phylogenetic differences among all four honeybee species within a tehsil. Once again, a pattern of close relatedness between *A. dorsata* and *A. mellifera* was evident.

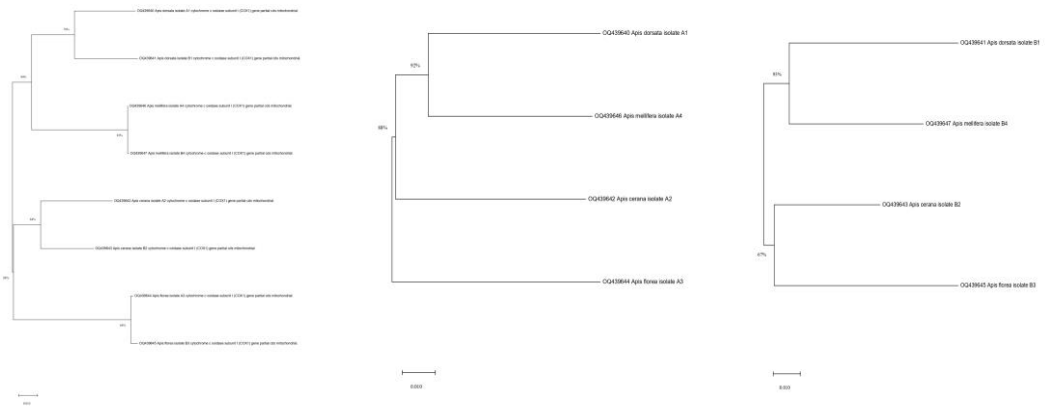


Fig. 4, Fig. 5 and Fig. 6 shows Phylogenetic evolutionary relationship between species of Trimbakeshwar and Malegaon tehsils, phylogenetic evolutionary relationship between species in Trimbakeshwar, phylogenetic evolutionary relationship between species of Malegaon tehsil respectively

To evaluate evolutionary divergence between the all 8 sequences of both tehsils; Trimbakeshwar and Malegaon, 8 nucleotide sequences were considered. There were a total of 656 positions in the final dataset. Analyses were conducted using the Kimura 2-parameter model (Kimura M. 1980). Codon positions included were 1st+2nd+3rd+Noncoding. All ambiguous positions were removed for each sequence pair (pair wise deletion option). Evolutionary analyses were conducted in MEGA X (Kumar S. et al 2018). Divergence between *A. dorsata* of both the tehsils is 0.0658, *A. cerana* is 0.0660, *A. florea* is 0.0047and no divergence is observed in *A. mellifera*. The estimated Transition/Transversion bias(R) is 0.76. Substitution pattern and rates were estimated under the Kimura (1980) 2-parameter model. The nucleotide frequencies are A = 25.00%, T/U = 25.00%, C =25.00%, and G = 25.00%. For estimating ML values, a tree topology was automatically computed. The maximum Log likelihood for this computation was -1889. 095.In the Trimbakeshwar tehsil, the pair wise distance analysis between *A. dorsata, A. cerana, A. florea,* and *A. mellifera* revealed divergence ranging from 0.102 to 0.130. Specifically, *A. dorsata* and *A. mellifera* exhibited the minimum divergence (0.102), indicating a close relationship, while *A. dorsata* and *A. florea* displayed the maximum divergence (0.130), suggesting a more distant association. This analysis suggests that *A. dorsata* and *A. mellifera* are closely related, whereas *A. dorsata* and *A. florea* are more distantly related. The estimated Transition/Transversion bias (R) for species in the Trimbakeshwar tehsil was determined to be 0.71. Substitution patterns and rates were assessed using the Kimura (1980) 2-parameter model, with nucleotide frequencies reported as A = 25.00%, T/U = 0.00%, C = 25.00%, and G = 25.00%. The tree topology for Maximum Likelihood (ML) values was automatically computed, yielding a maximum Log likelihood of -1568.828 for the dataset consisting of 636 positions. Furthermore, the Maximum Composite Likelihood transition/transversion bias was calculated, with nucleotide frequencies reported as 33.73% (A), 40.32% (T/U), 15.27% (C), and 10.69% (G). The transition/transversion rate ratios were determined as k1 = 1.218 (purines) and k2 = 2.312 (pyrimidines). The overall transition/transversion bias (R) was found to be 0.754, calculated as R = [AGk1 + TCk2]/[(A+G)*(T+C)]. In the Malegaon tehsil, the genetic divergence analysis revealed a maximum divergence of 0.149 and a minimum divergence of 0.106 among the studied honeybee species. For the estimation of substitution patterns in species from Malegaon tehsil, the nucleotide frequencies were reported as A = 25.00%, T/U = 25.00%, C = 25.00%, and G = 25.00%. The Maximum Likelihood (ML) values were computed with an automatically generated tree topology, resulting in a maximum Log likelihood of -1519.115 for the dataset. Furthermore, for the Maximum Composite Likelihood transition/transversion bias, the nucleotide frequencies were determined as 33.55% (A), 41.26% (T/U), 15.23% (C), and 9.95% (G). The transition/transversion rate ratios were calculated as k1 = 0.53 (purines) and k2 = 1.601 (pyrimidines). The overall transition/transversion bias (R) was found to be 0.481, calculated as R = [A*G*k1 + T*C*k2]/[(A+G)*(T+C)]. Evolutionary divergence between sequences was estimated using the Kimura 2 parameter model (Kimura M. 1980), including codon positions 1st+2nd+3rd+Noncoding. Ambiguous positions were removed for each sequence pair using the pairwise deletion

option. The genetic variation among all four species from each tehsil was compared, and phylogenetic trees were constructed using the neighbor-joining technique.

5. Conclusion- The molecular estimation of genetic barcoding in honey bees focused on the mitochondrial COX I gene. *A. mellifera*, due to widespread beekeeping practices, has been extensively studied on a global scale. Subsequent studies have been conducted on *A. dorsata* and *A. cerana*, with comparatively less research dedicated to *A. florea*. This research encompasses all four species, conducting barcoding and analyzing the genetic variance. The obtained data has been submitted to the NCBI, and genetic divergence is computed through Maximum Likelihood analysis. Evolutionary divergence is observed and compared among honeybee species from both tehsil areas. The estimation of genetic variation reveals close relatedness between *A. dorsata* and *A. mellifera*, a noteworthy finding from molecular studies. Maximum divergence is noted in *A. dorsata* and *A. florea*. The generated data holds promise for laying the foundation for fundamental insights into the molecular study of the genus *Apis*, offering valuable information for future research endeavors.

Declaration- The author declares that have they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References-

1. *Alattal et al., 2019 Y. Alattal, M. Alsharhi, A. Alghamdi, S. Fuchs Characterization of Socotra (Yemen) honey bees (Apis mellifera) using morphometric and genetic markers Bull. Insectol., 72 (2) (2019), pp. 281-285*
2. Alattal, Y., Alsharhi, M., Alghamdi, A., Fuchs, S., 2019. Characterization of Socotra (Yemen) honey bees (*Apis mellifera*) using morphometric and genetic markers.Bull. Insectol. 72 (2), 281–285.
3. Dempster and McLean,1999, Alattal (Journal of King Sauduni)
4. Diao, Q., Sun, L., Zheng, H., Zeng, Z., Wang, S., Xu, S., ... & Wu, J. (2018). Genomic and transcriptomic analysis of the Asian honeybee *Apis cerana* provides novel insights into honeybee biology. *Scientific Reports*, *8*(1), 822.
5. Khan KA. Genetic diversity and phylogenetic relationship among the western and the Asian honey bees based on two mitochondrial gene segments (COI and ND5). Saudi J Biol Sci. 2021 Dec;28(12):6853-6860. doi: 10.1016/j.sjbs.2021.07.062. Epub 2021 Jul 29. PMID: 34866985; PMCID: PMC8626216.
6. Kimura M. (1980).A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. Journal of Molecular Evolution16:111-120.
7. Kojima, Y., Toki, T., & Moritz, R. F. (2011). Molecular phylogeny of the honeybee *Apis mellifera* subspecies from Japan based on the mitochondrial tRNAleu-cox2 intergenic region. Apidologie, 42(6), 696-705.
8. Kumar S. and Gadagkar S.R. (2001). Disparity Index: A simple statistic to measure and test the homogeneity of substitution patterns between molecular sequences. Genetics 158:1321-1327.
9. Kumar S., Stecher G., Li M., Knyaz C., and Tamura K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across
10. Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. Molecular Biology and Evolution, 35(6), 1547-1549.
11. Morse, R. A., & Calderone, N. W. (2000). The value of honey bees as pollinators of US crops in 2000. Bee Culture, 128(3), 1-15.
12. Qamer, S., Al-Abbadi, A. A., Sajid, M., Asad, F., Khan, M. F., Khan, N. A., ... & Mohammed, O. B. (2021). Genetic analysis of honey bee, *Apis dorsata* populations using random amplified polymorphic DNA (RAPD) markers. *Journal of King Saud University-Science*, *33*(1), 101218.
13. Rizwan M., Li Z., Nie H., Qasim M., Raza M., Hassanyar A. High mitochondrial diversity of *Apis mellifera* under COI gene from China and Pakistan. Appl. Ecol. Environ. Res. 2018;16:2933–2945.
14. Ruttner, F. (1988). Biogeography and Taxonomy of Honeybees. Springer Science & Business Media.
15. Shankarganesh E (2017) Insect biodiversity: The teeming millions-Are view. Bull EnvPharmacol Life Sci6(3):101–105
16. Shashank,P.R., Naveena, N.L., Rajgopal, N.N. et al.DNA barcoding of insects from India: Current status and future perspectives. MolBiolRep49, 10617–10626 (2022).https://doi.org/10.1007/s11033-022-07628-2
17. Sreedevi K, Meshram N, Shashank PR (2015) Insect taxonomy- basics to barcoding. In: Chakravarthy AK (ed) New Horizons in Insect Science. Towards Sustainable Pest Management. Springer India
18. Winston, M. L. (1991). The Biology of the Honey Bee. Harvard University Press. Oldroyd, B. P. (2007). What's Killing American Honey Bees? PLoS Biology, 5(6), e168.
19. ZSI (2012) COPXI publications. www.zsi.gov.in/Cop-11/cop-11.html/[Accessedon Nov 2021]
20. ZSI (2020) COPXI publications. www.zsi.gov.in/Cop-11/cop-11.html/[AccessedonNov 2021]