



# Genetic Divergence and Phylogeny of North-Western Indian Himalayan Population of Honey Bee (Hymenoptera: Apidae) Inferred from Mitochondrial DNA Sequences

Sunaullah Bhat, Amit Umesh Paschapur\*, A.R.N.S. Subbanna, Johnson Stanley, Jaiprakash Gupta, K.K. Mishra and Sandeep Kumar<sup>1</sup>

Crop Protection Division, ICAR- Vivekananda Parvatiya Krishi Anusandhan Sansthan, Almora, Uttarakhand-263 601, India

<sup>1</sup>Department of Zoology, Kumaun University, SSJ Campus, Almora, Uttarakhand-263 601, India

\*E-mail: [amitp3929@gmail.com](mailto:amitp3929@gmail.com)

**Abstract:** The utilization of molecular techniques to predict the phylogenetic similarity within the honey bee topologically divided these into three groups viz., giant bees, dwarf bees and cavity-nesting bees. Four species of honey bee, i.e., Asian honey bee (*Apis cerana*), dwarf bee (*Apis florea*), the rock or giant honey bee (*Apis dorsata*) and European honey bee (*Apis mellifera*) collected from North Western Himalaya region were molecularly characterized using partial mitochondrial Cytochrome oxidase I (COI). Nucleotide frequency analysis revealed that partial COI gene sequences were A+T biased (>73%), while the amino acid frequency analysis showed higher frequency of leucine (13.71%) and serine (12.27%). Among the tested species of honey bees *A. florea* showed exceptionally higher concentration of Cystine Amino acid. The overall variation in 72 SNPs was revealed by multiple sequence alignment of COI sequences from four species. The results reveal more transversion (56.5%) compared to transition (43.5%). Additionally, the pairwise genetic distance analysis described that *A. cerana* and *A. dorsata* had least genetic distance (0.102). The Ts/Tv value of 0.61 suggests that there is insignificant neutral selection in honeybees of north western Himalayan region of Uttarakhand. Such researches are significantly important to analyze the biodiversity of area, distinguish cryptic species and develop distinct taxonomic tools for integrative taxonomy of Honey bees.

**Keywords:** Honey bees, Molecular characterization, Cytochrome oxidase I, Phylogenetic analysis, Genetic divergence, Indian Himalayas

Pollinators are fundamental vehicles for plant reproduction, ecosystem health, and agricultural production (Ollerton 2017 and Klein et al 2007). The most essential and diverse group of pollinators are the insects, with honey bees specifically regarded the most essential group due to their significant numbers and expertise on floral resources (Ollerton 2017). They contribute up to 73% pollination services provided by animals in total (Michener 2007, Abrol 2009 and Ollerton et al 2011). Worldwide the services of bee pollinators to the agriculture are predicted around \$235- \$577 billion per year (FAO 2018). Honey bees of genus *Apis* (Hymenoptera: Apidae: Apini), comprise around seven species and are ranked as best pollinators due to their eusocial nature (Alqarni et al 2011, Oleksa and Tofilski 2015, Haddad et al 2016, Cridland et al 2017, Eimanifar et al 2018). Generally, all the honey bee species have same biological characters viz., i) the occurrence of one queen in the colony (for egg laying), ii) thousands of workers are responsible for all activities such as foraging, cleaning of the hive, defending the enemy (Abou-Shaara et al 2017), and iii) drones with main role to mate virgin queen (Heidinger et al 2014 and Abou-Shaara et al 2021), iv) the similar thermoregulation capacity (Abou-Shaara et al 2017). These resemblances

direct the occurrence of the common genetic characters between them. These characteristics are based mainly on the coding regions in the nuclear and the mitochondrial DNA (mtDNA) (Abou-Shaara 2019a).

During the last few decades, numerous identification approaches like developing the database of prototypes and comparing the unknown specimen with the well identified specimen within the database are practiced but, with only probable accuracy has been achieved (Ratnasingham and Hebert 2007). Distinguishing among the honey bee subspecies is intricate and necessitates specific knowledge. It is normally based on morphometric characteristics of bees, like head width in frontal view (including eyes), antennal scape length, and hind tibia length (Gruber et al 2013). Angles in the forewing have been used to distinguish within bee groups (Nedic et al 2011 and Kulici et al 2014). Besides these wing morphometrics are also utilized (Silva et al 2015). Automated Bee Identification (Schröder et al 2002) and Digital Automated Identification System (DAISY) (Weeks et al 1999) are two such systems, although they are rarely effective because a new specimen whose prototype does not exist in the library cannot be identified with 100% accuracy. Using molecular techniques, researchers were able to better

understand the genetic diversity pattern, phylogenetics, biogeography, evolution, and population genetics of insects (Susanta 2006 and Andrade- Souza et al 2017). Molecular characterization methods such as DNA barcoding using mitochondrial cytochrome oxidase I genes have shown promise in identifying the bee biota. (Magnacca and Brown 2010, Magnacca and Brown 2012 and Wang et al 2014). The molecular techniques provide numerous other information's like identification of the sympatric and cryptic species within a habitat (Eriksson et al 2017 and Nneji et al 2020), classification of dimorphic sex species (Gibbs 2009 and Sheffield et al 2009), characterizing of species which are otherwise difficult to identify on morphological basic (Gibbs 2009, Rehan and Sheffield 2011 and Williams et al 2012) and determination of the evolutionary relationship and genetic distances within the targeted insect species (Will and Rubinoff 2004 and Kekkonen and Hebert 2014). Use of COI gene, has been utilized successfully to understand the population genetic structure and variations among the insects (Pramual et al 2005 and Patel and Jadhav 2019). Therefore, utilizing this method in combination with other mitochondrial and nuclear markers has unlocked new perspectives for the identification of species and studies related to genetic diversification of insects (Low et al 2016 and Pramual and Nanork 2012). In this study, we investigated the population genetic diversity of *Apis* species of Northwestern Himalayas. We sequenced the cytochrome c oxidase subunit 1 gene (COI) of individuals of *Apis* species collected from different eco-geographic regions in Northwestern Himalayas. The objectives present study was to use the DNA (COI) barcode data to reliably identify and characterize *Apis* species and evaluate the mitochondrial genetic diversity pattern and population structure.

## MATERIAL AND METHODS

**Test insects:** The specimens of four *Apis* species (*A. cerana*, *A. mellifera*, *A. florea* and *A. dorsata*) were collected from crops grown in the Experimental Farm, Hawalbagh of ICAR-Vivekananda Parvatiya Krishi Anusandhan Sansthan (VPKAS), Almora situated at 29°37'N and 79°40'E with an altitude of 1250 m in Uttarakhand state of North Western (NW) Himalayan Region, India. The *Apis* species are dominant pollinators of the crops in the Northwestern Himalayas.

**DNA extraction:** The standard methodology of CTAB as described by Subbanna et al (2016) was followed for extraction of genomic DNA from the honey bee specimens. The three pairs of legs of honey bee species were crushed in liquid nitrogen utilizing autoclaved pestle and mortar, the crushed material was transferred to a micro-centrifuge tube.

600 µl of CTAB solution (20 parts of 1M Tris-HCl, 8 parts of 0.5M EDTA, 56 parts of 5M NaCl and 4 parts of CTAB and the final volume makeup 1000ml and the pH adjusted to 8. After this 0.4 parts of β-mercaptaethanol were added along with 1 mg/ml of proteinase K was transferred to the tube and incubated in a water bath at 57°C for 3-5 hours. The contents were vortexed manually after every 20 minutes for thorough degradation of the tissues. The degraded material was treated twice with phenol-chloroform-isoamyl solution (25:24:1) to extract the genomic DNA and ice-cold isopropyl alcohol was used to precipitate the DNA at -20°C for 30 minutes. The DNA pellet obtained after centrifugation at 10000 rpm for 7 minutes was washed with 70% chilled ethyl alcohol to remove the excess salts and was suspended in 40 µl of TE buffer. DNase-free RNase A treatment was followed for 1 h at 37°C to remove the RNA residues. Electrophoresis with 0.8% agarose gel was carried out to visualize intact genomic DNA and the DNA samples were diluted if required to obtain a working solution of 20-25 ng/µl.

**PCR protocol:** The insect specific universal mitochondrial cytochrome oxidase I primers (JM 76 (5-GAGCT GAATTAGG(G/A)ACTCCAGG-3) and JM 77 (5-ATCACCTCC(A/T) CCTGCAGG ATC-3)) were used for amplification of the target region from four *Apis* species. The PCR reaction mix utilized for the study was as follows, 50 ng of DNA template, 200nM of dNTPs, 1mM of each primer, 2.5 units of Taq DNA polymerase and 5µl of PCR reaction buffer was added to make the make a final volume of 50 µl. The PCR reactions were performed in a thermal cycler (Biorad) with an initial 3 min denaturation step at 95°C, followed by 35 amplification cycles consisting of 1 min denaturation at 95°C, 45 seconds annealing at 52°C and 1 min extension at 72°C with an additional final step of extension for 10 min at 72°C. The presence of amplified PCR product was visualized and confirmed in the gel documentation system (Alpha Image Analyzer, Alpha Innotech Corporation) by 1.2% agarose-EtBr 10 mg/ml gel electrophoresis with 2.5 µl PCR product.

**Sequencing and data analysis:** Gel elution columns (Sigma) were used for purification of the amplified products of the target gene. The purified products were sequenced directly by an automated DNA sequencer (ABI 377) following manufacturers guidelines for the Big Dye terminator kit (Applied Biosystems). The sequence thus obtained were aligned with Clustal Omega (1.2.2) multiple sequence alignment (Sievers and Higgins 2018). Further analysis on phylogenetic and molecular evolutionary analyses, pairwise genetic distance among *Apis* species, variation in nucleotide sequences as well as transition/transversion rate ratios were calculated by comparing with the CO1 sequences of other closely related *Apis* spp. in the NCBI GenBank database by

BLASTN. The MEGA X 10.0.5 software (Molecular Evolutionary Genetic Analysis version X) (Kumar et al 2018) was used for the construction of Maximum likelihood evolution tree (Saitou and Nei 1987) utilizing the distance matrix from the alignment. The confidence level of each branch was tested by bootstrapping 1000 replicates generated with random seed. The nucleotide sequences were translated into amino acid sequences with the help of invertebrate mitochondrial genetic code through the ExPASy translate: SIB bioinformatics resource portal (Artimo et al 2012) and were aligned using Clustal omega software (Sievers and Higgins 2018). The variation in the amino acid concentration within and among the species was estimated by computing the amino acid composition by MEGAX 10.0.5 software. The generated sequences of partial mitochondrial COI region were further submitted to NCBI GenBank database through Bankit submission tool (<https://www.ncbi.nlm.nih.gov/WebSub/?tool=genbank>) to acquire the individual accession number (*Apis dorsata* COI-Almora 1 - ON506691, *Apis florea* COI-Almora 1 - ON506690, *Apis mellifera* COI-Almora 1 - ON506267, *Apis cerana indica* COI-Almora 1 - ON506013).

## RESULTS AND DISCUSSION

Four species of honey bees were collected from the premises of experimental form hawalbagh of VPKAS-ICAR, Almora while pollinating various crops. During the present investigation, the partial mitochondrial COI regions of (683-696) bp of *A. cerana*, *A. mellifera*, *A. florea* and *A. dorsata* were amplified. These species were morphologically characterized but to clear the ambiguity at molecular level, the molecular characterization with partial mitochondrial Cox 1 gene were carried out. The sequences of honey bee obtained were BLASTn to determine the resemblance index and recognition of every isolate from North-western Himalayas of India with closely related sequences are displayed in the phylogenetic tree. For mitochondrial COI gene fragment, there was 97-100, 98.53-99.25, 85.28-95.23 and 91.29-99.84 percent nucleotide identity for *A. cerana*, *A. mellifera*, *A. florea* and *A. dorsata* sequences respectively,

when these sequences were analyzed by BLASTn separately in NCBI database. The evolutionary relationship of honey bees based on mitochondrial COI gene categorized the whole assemblage into four main groups (Fig. 1). *Apis cerana indica* voucher specimen Almora form a sub-cluster with 100% similarity with the *Apis cerana* isolate CL24 mitochondrion species form Bangalore, India, collected by Sudhagar et al (2014), *Apis cerana* isolate KYY-41, and *Apis cerana* isolate KYY-42 from China isolated by Wang (2021), *Apis cerana* mitochondrion, a species from USA (having NCBI accession number KU963188.1, MZ191825.1, MZ191824 and NC\_014295.1 respectively). All these species form separate sub-clade in the maximum likelihood tree. *Apis florea* voucher specimen Almora represents 99% similarity in its sequence with *Apis florea* cytochrome oxidase subunit I specimen collected from Bangalore, Karnataka (NCBI accession number KU666428.1), both forming a separate cluster in maximum likelihood tree (Sangeetha et al 2016). *A. dorsata* voucher specimen Almora shows 96% sequence identity with *Apis dorsata* specimen cytochrome oxidase subunit I (NCBI accession number KJ755328) recorded from Indian Institute of Horticultural Research, Hessaraghatta Lake Post, Bangalore, Karnataka (Sudhagar et al 2014). Both these form a separate sub-cluster in maximum likelihood tree.

*A. mellifera* form two sub-cluster, in which our specimen i.e., *Apis mellifera*-COI-Almora 1 represent a separate sub-cluster. Although *A. mellifera* show cent percent identity with *Apis mellifera* voucher PHEL: 4132 from USA (KU601503.1), *Apis mellifera* voucher HY37 (NC001566.1) from USA, *Apis mellifera* anatoliaca mitochondrion (MT188686.1) from USA, *Apis mellifera caucasica* mitochondrial DNA (AP018404.1) from Japan, *Apis mellifera* voucher HY37 from Bangladesh and *Apis mellifera carpatica* isolate 6-27 (MF100917.1) from Russia (Dhami et al 2016, Aslam et al 2016, Aslam et al 2018, Boardman et al 2020 and Ilyasov et al 2019) it did not show close evolutionary relationship with any of these sequences. As the entire mitochondrial DNA sequence was used for the construction of phylogeny, the Cox I region proves to be highly conserved region, through which

**Table 1.** Pair-wise genetic distance analysis between four species of *Apis* spp

| Cox1 nucleotide sequences                  | <i>Apis_cerana_Indica_C</i><br>OI-Almora_1 | <i>Apis_mellifera</i> -COI-<br>Almora_1 | <i>Apis_florea</i> -COI-<br>Almora_1 | <i>Apis_dorsata</i> -COI-<br>Almora_1* |
|--|--|---|--------------------------------------|--|
| <i>Apis_cerana_Indica_C</i> -COI-Almora_1* |  |   |                                      |  |
| <i>Apis_mellifera</i> -COI-Almora_1        | 0.127                                      |   |                                      |  |
| <i>Apis_florea</i> -COI-Almora_1           | 0.255                                      | 0.263                                   |                                      |  |
| <i>Apis_dorsata</i> -COI-Almora_1          | 0.102                                      | 0.114                                   | 0.277                                |  |

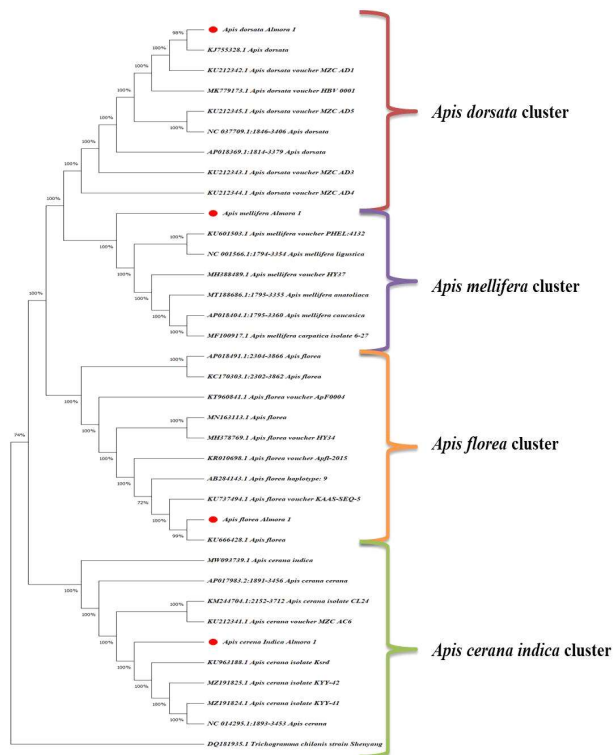
\*The nucleotide sequences of *Apis dorsata*-COI-Almora\_1 and *Apis cerana\_Indica\_C*-COI-Almora\_1 showed no pairwise genetic relationship, thus yielding the Pair-wise genetic distance analysis value of 0.000

molecular characterization of insects can be taken up with higher degrees of specificity.

When the phylogenetic tree was constructed with the MEGA X 10.0.5 software (Fig. 1), it was observed that the four species of *Apis* native to Indian Himalayas formed four separate groups in the maximum likelihood evolution tree. The node support estimated using 1000 bootstrap pseudoreplicates, showed that *A. dorsata*, *A. florea* and *A. mellifera* species evolved together as they showed 100% node value. The results showed that the *A. cerana* species showed 98% node value using 1000 bootstrap pseudoreplicates in the evolutionary analysis.

The pairwise genetic distance analysis carried out between the Four tested *Apis* species, recorded that *A. cerana indica* and *A. dorsata* had least genetic distance of 0.102 followed by *A. dorsata* and *A. mellifera* with genetics distance of 0.114 while the genetic distance between *A. florea* and *A. mellifera* was the highest with the value 0.263. Further to understand the inter-specific diversity among the four experimented species on the basis of 684 bp CO1 sequence, the variation in single nucleotide polymorphism (SNPs) was estimated by aligning the four sequences in CLUSTAL Omega (1.2.4) multiple sequence alignment

software (Fig. 2). It was observed that a total variation in 72 SNP's was recorded, which was large enough to differentiate the species at inter-specific level. The mean number nucleotide frequency amongst the four species of *Apis* was also examined (Fig. 3) and COI sequences were usually A+T biased with the concentration of A+T exceeding 73.49%, while the concentration of G+C was well below 26.51%. The highest concentration was found to be of T/U i.e., 40.57 %. Along with nucleotide frequencies, the transition/transversion rate ratios were also calculated (Table 2) for the 36 species of four *Apis* species and it was reported that the ratios were  $k1 = 1.294$  (purines) and  $k2 = 1.738$  (pyrimidines) and the overall transition/transversion bias was  $R=0.627$ , where  $R = [A^*G*k1 + T^*C*k2]/[(A+G)*(T+C)]$ . The tested honey bee specimen show more transversion (56.5%) compared to transition (43.5). Maximum Composite Likelihood Estimate of the Pattern of Nucleotide Substitution showed maximum base substitution between C to T and vice versa with a maximum of 19.91. To further analyze the variation in amino acid composition among the four sequences of *Apis* species, the sequences were translated by ExPASy translate software and the nucleotide compositions were estimated by MEGA X software (Fig. 4). Majority of deviations in polypeptides were associated with substitution of small amino acids with other. The per cent amino acid composition among the four *Apis* species showed that Leucine and Serine were found at the highest frequencies with an average of 13.71% and 12.27% respectively. Among all the tested specimen Cysteine amino acid was found with higher frequency in *A. florea* only. Moreover, it was found that the variation in amino acid frequency indicates huge variation among the four *Apis* species native to Indian Himalayas. Throughout the world, 20,355 species of bee pollinators has been documented (Ascher and Pickering 2019), their taxonomic identification is still in dearth stages due to lack of distinguishing morphological characters in bee taxonomy (Packer et al 2009), identification of bees become more herculean task for taxonomists due to huge biological diversity and presence of cryptic species (Hines and Williams 2012 and Vamosi et al



**Fig. 1.** Phylogenetic analysis through maximum likelihood method of four species of *Apis* along with 32 closely associated *Apis* species sequences (*Trichogramma chilonis* strain shenyang was taken as out group)

**Table 2.** Maximum composite likelihood estimate of the pattern of nucleotide substitution of 34 species of *Apis*

| Nucleotides | A     | T     | C    | G    |
|-------------|-------|-------|------|------|
| A           | -     | 11.46 | 4.23 | 4.22 |
| T           | 9.3   | -     | 7.35 | 3.26 |
| C           | 9.3   | 19.91 | -    | 3.26 |
| G           | 12.03 | 11.46 | 4.23 | -    |

2017). Morphological identification is no longer enough to obtain accurate data on the position of a species. Molecular identification using CO1 gene as a genetic barcode has been accepted universally (Herbert et al 2003). Saini and Chandra (2019) observed that , the Indian subcontinent is estimated to have a diverse bee fauna, including over 766 species divided into 71 genera and six families, although the actual figure

may be much higher. Pakrashi et al (2020) conducted a survey of Himalayan bee fauna and molecular characterization showed huge diversity of bees in the Indian Himalayas. Out of the 156 bee species collected through extensive surveys, the bees belonging to the four major families were reported viz., Apidae (40 species), Halictidae (five species), Megachilidae (nine species), and Melittidae

|                         |                         |     |
|-------------------------|-------------------------|-----|
| Florea-COI-Almora       | --- <td>56</td>         | 56  |
| Mellifera-COI-Almora    | CCCATGGTATTAATTCTAGCTTT | 59  |
| Cerena                  | ---AAATATATTTATTAGCTTT  | 52  |
| Dorsata-COI-Almora      | ---AAATTATCTTTTATGCAAT  | 53  |
| * * * * * * * * * * * * |                         |     |
| Florea-COI-Almora       | GTTTATTTAAATATGAATTAT   | 116 |
| Mellifera-COI-Almora    | TATTATTCGA---ATAGAAT    | 116 |
| Cerena                  | GATTGTTGCG---ATAGAAT    | 109 |
| Dorsata-COI-Almora      | TATTATTCGA---ATAGAAT    | 110 |
| * * * * * * * * * * * * |                         |     |
| Florea-COI-Almora       | TATTACATAATGTACAGGT     | 176 |
| Mellifera-COI-Almora    | TAATACAATTGTTACTAGT     | 176 |
| Cerena                  | TAATACAGTTGTAACAAG      | 169 |
| Dorsata-COI-Almora      | TAATACAATTGTCACTAG      | 170 |
| * * * * * * * * * * * * |                         |     |
| Florea-COI-Almora       | GATTGTAGGATGCGGAAAT     | 236 |
| Mellifera-COI-Almora    | AATTGGAGGATTTGGAAAT     | 236 |
| Cerena                  | AATTGGAGGTTTGGAAAT      | 229 |
| Dorsata-COI-Almora      | AATTGGAGGATTTGGAAAT     | 230 |
| * * * * * * * * * * * * |                         |     |
| Florea-COI-Almora       | TCCCCTAAATAAATAATAT     | 296 |
| Mellifera-COI-Almora    | CCCCGAATAAATAATATT      | 296 |
| Cerena                  | TCCTCGAATAAATAATAT      | 289 |
| Dorsata-COI-Almora      | TCCTCGAATAAATAATAT      | 290 |
| * * * * * * * * * * * * |                         |     |
| Florea-COI-Almora       | AAGAAATTTATTTTATCC      | 356 |
| Mellifera-COI-Almora    | AAGAAATTTATTTTATCC      | 356 |
| Cerena                  | AAGAAATTTATTTTATCC      | 349 |
| Dorsata-COI-Almora      | AAGAAATATATATTATCC      | 350 |
| * * * * * * * * * * * * |                         |     |
| Florea-COI-Almora       | TTATTTATATCATTTCAT      | 416 |
| Mellifera-COI-Almora    | ATATTTATATCATTCTTC      | 416 |
| Cerena                  | ATATTTATATCATTTCAT      | 409 |
| Dorsata-COI-Almora      | ATATATATACACTTTCAC      | 410 |
| * * * * * * * * * * * * |                         |     |
| Florea-COI-Almora       | AATTCGTCAATTATAGG       | 476 |
| Mellifera-COI-Almora    | AATTCCTCAATTATAGG       | 476 |
| Cerena                  | AATCTCATCAATTATAGG      | 469 |
| Dorsata-COI-Almora      | TATTTCTTCAATTATAGG      | 470 |
| * * * * * * * * * * * * |                         |     |
| Florea-COI-Almora       | TTCATTAATTATGATCAA      | 536 |
| Mellifera-COI-Almora    | TTCTATAAATTATGACCA      | 536 |
| Cerena                  | TTCATTAAATTATGATCAA     | 529 |
| Dorsata-COI-Almora      | TTCATTAAATTATGATCAA     | 530 |
| * * * * * * * * * * * * |                         |     |
| Florea-COI-Almora       | ACTTGTAATATCTTTACC      | 596 |
| Mellifera-COI-Almora    | ATTAATATATCATTACCT      | 596 |
| Cerena                  | ATTAATATATCCTTCAG       | 589 |
| Dorsata-COI-Almora      | ATTAATATATCATTACCA      | 590 |
| * * * * * * * * * * * * |                         |     |
| Florea-COI-Almora       | TTTTAATACATCATT         | 656 |
| Mellifera-COI-Almora    | TTTTAATACATCATT         | 656 |
| Cerena                  | TTTTAATACATCATT         | 649 |
| Dorsata-COI-Almora      | TTTTAATACATCATT         | 650 |
| * * * * * * * * * * * * |                         |     |

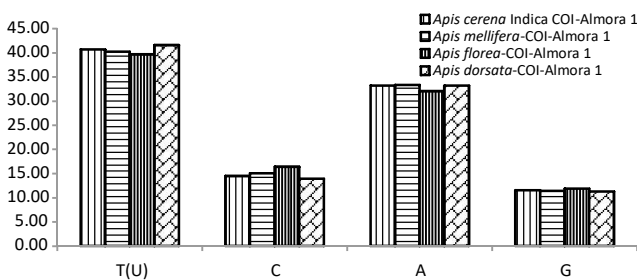
Fig. 2. Multiple sequence alignment of four species of Apis for analyzing the variation in single nucleotide polymorphism (SNPs)

(one species) which are widely distributed all over the Himalayas.

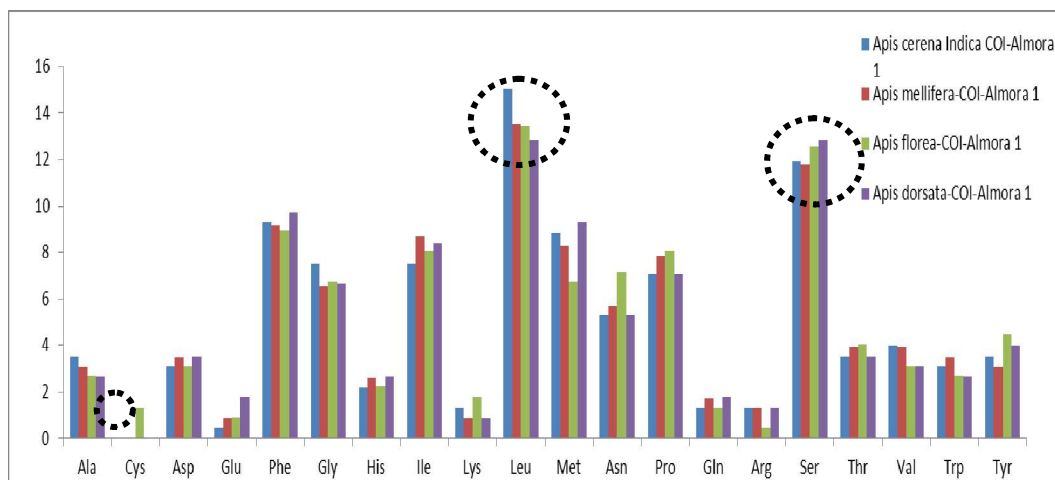
Saini et al (2019) discovered a new species of wild bee (*Melitta indica*) (Hymenoptera: Melittidae) from North Western Himalayan region of Uttarakhand through molecular characterization and DNA barcoding methods that has sparked the importance of molecular techniques in understanding the bees species delimitation in several areas of the world (Smith et al 2008; Butcher et al 2012). Through morphological Characterization it is difficult to identify a specimen (Gibbs 2009, Rehan and Sheffield 2011 and Williams et al 2012), appropriate identification of the specimens up to species level (Schmidt et al 2015), clarifying level of speciation of various species associated with various taxonomic levels (Trunz et al 2016 and Oh et al. 2013), association of sexes in dimorphic species classification (Gibbs 2009 and Sheffield et al 2009), cryptic species detection (Packer and Taylor 1997) and determining the evolutionary relationship and genetic distances among the targeted insect species (Kekkonen and Hebert 2014). Besides the above-highlighted applications molecular techniques, these support taxonomists in integrative

taxonomy to decode the biological diversity of the given region (Dayrat 2005). Globally, nine species have been commonly recognized under the genus *Apis*, including *A. andreniformis*, *A. cerana*, *A. dorsata*, *A. florea*, *A. koschevnikovi*, *A. laboriosa*, *A. mellifera*, *A. nigrocincta* and *A. nuluensis* (Koeniger et al 2011). Based on recent studies, the origin of *Apis* spp. has been predicted from tropical regions, and of all the honey bee species including *A. mellifera* and *A. cerana* have spread northward into the temperate zone, explaining their large scale role in agriculture (Ji 2021). Using molecular techniques to predict the phylogenetic similarity with the help of mitochondrial genes, strongly supported the basic topology that these groups of honeybee species belong to three main clusters: giant bees (*A. dorsata* and *A. laboriosa*), dwarf bees (*A. andreniformis* and *A. florea*) and cavity-nesting bees (*A. mellifera*, *A. cerana*, *A. koschevnikovi*, *A. nuluensis* and *A. nigrocincta*) (Koeniger et al., 2011 and Ji 2021).

In present study investigated the genetic diversity and population structure among the honeybee populations in North western Himalayan region. In our study, characterization of four *Apis* sp., three native to India viz., *A. cerana*, *A. florea* and *A. dorsata* and one introduced i.e., *A. mellifera* utilizing molecular techniques like DNA barcoding and phylogenetic analysis with the closely related *Apis* species obtained from NCBI data base were carried out. The phylogenetic hierarchy of the *Apis* species from Northwestern Himalayas and closely related sequences obtained from NCBI form four separate cluster in the maximum likelihood tree. All the sequences fall into respective groups, which are in agreement with Arias and Sheppard (2005). The result highlighted that *A. cerana* voucher Almora show close resemblance with *Apis cerana* species from India, China and USA. Similarly *A. mellifera*



**Fig. 3.** Average nucleotide frequency among the four species of *Apis*



**Fig. 4.** Variation in amino acid composition of four *Apis* species native to North Western Himalayas

show resemblance with species from USA, Bangladesh, China, Japan and Russia. This genetic variation in mitochondrial COI represent that this gene has high average genetic diversity. Similar results were obtained by Ballard and Whitlock (2004). Oldroyd et al (2006) suggested that mitochondrial COI genes are well suited for the characterization of phylogenetic divergence within species. Meixner et al (2013) revealed that this method is effective in inferring bee identification and evolution of subspecies. Packer and Ruz (2017) and Kevan and Packer (2009) suggested that COX 1 genes are the commonly used universal genes for the identification of honey bee species worldwide. Ozdil and Ilhan (2012) also followed DNA barcoding technique for characterizing honey bee subspecies in Turkey. The transition/transversion (Ts/Tv) ratio is instrumental in inferring the direction and magnitude of natural selection. The Ts/Tv value of 0.61 suggests that there is insignificant neutral selection in honeybees of north western Himalayan region of Uttarakhand. The ratio of more than 1, implies a positive or Darwinian selection; whereas the ratio less than 1, implies purifying selection and a ratio of one indicates neutral (i.e., no) selection (Chalapathy et al 2014). However the positive and purifying selection at different points within the gene or at different times along its evolution may cancel each other out giving an average value that may be lower, equal or higher. Thus, it could be inferred that the honeybees of north western Himalayan region of Uttarakhand might be in the verge of diversification. Moreover, the exact transition rate (43.5%) and transversion rate (56.5%) suggests no significant alteration representing frequent migration, mutations, gene flow amongst these populations. Transitions do not contribute heavily to genetic divergence, whereas transversion generate significant impact on the evolution of species. The comparable values of transitions and transversions in the current study suggest the possible occurrence of genetic divergence over evolutionary time scales. Chalapathy et al (2014) showed similar result in the honey bees from Karnataka, India. Like Willis et al (1992), suggested honeybees are AT biased and the mean number of A+ T and C+G satisfied of the findings is in the ratio of approximately (73.49:26.51) 3:1 respectively. Willis et al (1992) additionally discussed that AT rich sequence may be the result of numerous factors including small effective population size, selection drift, and mitochondrial polymerase inefficiency. Historical sign suggest rapid changes in the morphology of *Apis* and these rapid changes are predicted between upper Eocene and Oligocene approximately 10 million years ago (Culliney 1983). The A+T bias pointed out for various codons has been elucidated either as a mutational practice favouring the buildup of those

nucleotides or the result of an ineffective repair system (Chalapathy et al 2014).

## CONCLUSION

The comparative analysis of honey bee populations was carried out, taking into account the residual phylogenetic relationships. As the Indian Himalayas are undisturbed as well as unexplored ecosystems on the earth where biodiversity has not been studied comprehensively due the hilly location where taking up survey is very much difficult. Our investigation not only forms the point of commencement for study of bee fauna in Indian Himalayas, but also opens the doors of opportunities to discover the faunal diversity of the Himalayas.

## ACKNOWLEDGEMENT

The authors are thankful to Indian Council of Agricultural Research (ICAR) and the Director, ICAR-VPKAS, Almora for providing financial assistance and laboratory facilities for conducting experiments successfully.

## REFERENCES

- Abou-Shaara HF 2019. Utilizing bioinformatics to detect genetic similarities between African honey bee subspecies. *Journal of Genetics* **98**(4): 1-7.
- Abou-Shaara HF, Abbas AS, Al-Kahtani SN, Taha EKA, Khan KA, Jamal ZA and El Den Nasser MG 2021. Exploring the non-coding regions in the mtDNA of some honey bee species and subspecies. *Saudi Journal of Biological Sciences* **28**(1): 204-209.
- Abou-Shaara HF, Owayss AA, Ibrahim, YY and Basuny NK 2017. A review of impacts of temperature and relative humidity on various activities of honey bees. *Insectes Sociaux* **64**(4): 455-463.
- Abrol DP 2009. Plant-pollinator interactions in the context of climate change: An endangered mutualism. *Journal of Palynology* **45**: 1-25.
- Ajao AM, Nneji LM, Adeola AC, Oladipo SO, Ayoola AO, Wang YY and Olademeji YU 2021. Genetic diversity and population structure of the native Western African honeybee (*Apis mellifera adansonii* Latreille, 1804) in Nigeria based on mitochondrial COI sequences. *Zoologischer Anzeiger* **293**: 17-25.
- Alqarni AS, Hannan MA, Owayss AA and Engel MS 2011. The indigenous honey bees of Saudi Arabia (Hymenoptera, Apidae, *Apis mellifera jemenitica* Ruttner): Their natural history and role in beekeeping. *ZooKeys* **134**: 83-98.
- Arias MC and Sheppard WS 2005. Phylogenetic relationships of honey bees (Hymenoptera: Apinae: Apini) inferred from nuclear and mitochondrial DNA sequence data. *Molecular Phylogenetics and Evolution* **37**(1): 25-35.
- Artimo P, Jonnalagedda M, Arnold K, Baratin D, Csardi G, De Castro, E and Stockinger H 2012. ExPASy: SIB bioinformatics resource portal. *Nucleic Acids Research* **40**(W1): W597-W603.
- Ascher JS and Pickering J 2019. Discover life bee species guide and world checklist (Hymenoptera: Apoidea: Anthophila). Available from: <https://www.discoverlife.org> (Accessed on June 2020).
- Aslam AFM, Sarker S and Howlader AJ 2018. Submitted (24-MAY-2018) Zoology, *DNA Barcoding* Laboratory, Department of Zoology, Jahangirnagar University, Savar, Dhaka 1342, Bangladesh

- Aslam AFM, Sarker S and Howlader AJ 2018. Direct Submission, Submitted (24-MAY-2018) Zoology, DNA Barcoding Laboratory, Department of Zoology, Jahangirnagar University, Savar, Dhaka 1342, Bangladesh (NCBI accession number MH388489.1 of *A. mellifera* cited on 5-08-2021) <https://www.ncbi.nlm.nih.gov/nuccore/MH388489.1>
- Ballard JW and Whitlock MC 2004. The incomplete natural history of mitochondria. *Molecular Ecology* **13**(4): 729-744.
- Behura SK 2006. Molecular marker systems in insects: current trends and future avenues. *Molecular Ecology* **15**(11): 3087-3113.
- Boardman L, Kimball RT, Braun EL, Fuchs S, Grünewald B and Ellis JD 2020. Mitochondrial genome of *Apis mellifera anatoliaca* (Hymenoptera: Apidae)—the Anatolian honey bee. *Mitochondrial DNA Part B* **5**(2): 1876-1877.
- Breeze TD, Bailey AP, Balcombe KG and Potts SG 2011. Pollination services in the UK: How important are honeybees? *Agriculture, Ecosystems and Environment* **142**(3-4): 137-143.
- Butcher BA, Smith MA, Sharkey MJ and Quicke DL 2012. A turbo-taxonomic study of Thai *Aleiodes* (Aleiodes) and *Aleiodes* (Arcaleiodes) (Hymenoptera: Braconidae: Rogadinae) based largely on COI barcoded specimens, with rapid descriptions of 179 new species. *Zootaxa* **3457**(1): 1-232.
- Chalopathy CV, Puttaraju HP and Sivaram V 2014. A pilot study on genetic diversity in Indian honeybees-*Apis cerana* of Karnataka populations. *Journal of Entomology and Zoology Studies* **2**(3): 7-13.
- Cridland JM, Tsutsui ND and Ramírez SR 2017. The complex demographic history and evolutionary origin of the western honey bee, *Apis mellifera*. *Genome Biology and Evolution* **9**(2): 457-472.
- Culliney TW 1983. Origin and evolutionary history of the honeybees. *Apis Bee World* **64**: 29-38.
- Dayrat B 2005. Towards integrative taxonomy. *Biological journal of the Linnean society* **85**(3): 407-417.
- Dhami MK, Dsouza M, Waite DW, Anderson D and Li D 2016. Real-Time PCR assay for the identification of the Brown Marmorated Stink Bug (*Halyomorpha halys*). *Frontiers in molecular biosciences* **3**: 5.
- Eriksson JS, Blanco-Pastor JL, Sousa F, Bertrand YJK and Pfeil BE 2017. A cryptic species produced by autopolyploidy and subsequent introgression involving *Medicago prostrata* (Fabaceae). *Molecular Phylogenetics and Evolution* **107**: 367-381.
- Food and Agriculture Organization (FAO) 2018. Food and Agriculture Organization of the United Nations. Why bees matter: the importance of bees and other pollinators for food and agriculture. Available at: <http://www.fao.org/3/i9527en/i9527en.pdf>
- Gibbs J. 2009. Integrative taxonomy identifies new (and old) species in the *Lasioglossum* (Dialictus) tegulare (Robertson) species group (Hymenoptera, Halictidae). *Zootaxa* **2032**(1): 1-38.
- Gruber K, Schöning C, Otte M, Kinuthia W and Hasselmann M 2013. Distinct subspecies or phenotypic plasticity? Genetic and morphological differentiation of mountain honey bees in East Africa. *Ecology and Evolution* **3**(10): 3204-3218.
- Heidinger IM, Meixner MD, Berg S and Büchler R 2014. Observation of the mating behavior of honey bee (*Apis mellifera* L.) queens using radio-frequency identification (RFID): Factors influencing the duration and frequency of nuptial flights. *Insects* **5**(3): 513-527.
- Hines HM and Williams PH 2012. Mimetic colour pattern evolution in the highly polymorphic *Bombus trifasciatus* (Hymenoptera: Apidae) species complex and its comimics. *Zoological Journal of the Linnean Society* **166**(4): 805-826.
- Hu P, Lu ZX, Haddad N, Noureddine A, Loucif-Ayad W, Wang Y Z, Zhao RB, Zhang AL, Guan X, Zhang HX and Niu H 2016. Complete mitochondrial genome of the Algerian honeybee, *Apis mellifera intermissa* (Hymenoptera: Apidae). *Mitochondrial DNA A* **27**: 1791-1792.
- Ilyasov R, Nikolenko A, Tuktarov V, Goto K, Takahashi JI and Kwon HW 2019. Comparative analysis of mitochondrial genomes of the honey bee subspecies *A. m. caucasica* and *A. m. carpathica* and refinement of their evolutionary lineages. *Journal of Apicultural Research* **58**(4): 567-579.
- Ji Y 2021. The geographical origin, refugia, and diversification of honey bees (*Apis* spp.) based on biogeography and niche modeling. *Apidologie* **52**(2): 367-377.
- Kek SP, Chin NL, Tan SW, Yusof YA and Chua LS 2017. Molecular identification of honey entomological origin based on bee mitochondrial 16S rRNA and COI gene sequences. *Food Control* **78**: 150-159.
- Kekkonen M and Hebert PD 2014. DNA barcode-based delineation of putative species: Efficient start for taxonomic workflows. *Molecular Ecology Resources* **14**(4): 706-715.
- Klein AM, Vaissiere BE, Cane JH, Steffan-Dewenter I, Cunningham SA, Kremen C and Tscharntke T 2007. Importance of pollinators in changing landscapes for world crops. *Proceedings of the royal society B: biological sciences* **274**(1608): 303-313.
- Koeniger N, Koeniger G and Smith D (2011). Phylogeny of the genus *Apis*. In *Honeybees of Asia* (pp. 23-50). Springer, Berlin, Heidelberg.
- Kulici M, Bajrami Z and Kume K 2014. Genetic local differentiation of *Am carnica* population as well as subspecies *Am macedonica*, *Am ligustica*, *A m mellifera*, *A m caucasica*, in Germany, Alpine Region, Austria, Croatia, Serbia, Northern Kosovo, Albania and Macedonia. *Journal of Natural Sciences Research* **4**: 53-60.
- Kumar S, Stecher G, Li M, Knyaz C and Tamura K 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution* **35**(6): 1547.
- Low VL, Takaoka H, Pramual P, Adler PH, Yacob Z, Huang Y T, Da Pham X, Ramli R, Chen CD, Wannaket A and Sofian-Azirun M 2016. Delineating taxonomic boundaries in the largest species complex of black flies (Simuliidae) in the Oriental Region. *Scientific Reports* **6**(1): 1-8.
- Magnacca KN and Brown MJ 2010. Mitochondrial heteroplasmy and DNA barcoding in Hawaiian Hylaeus (Nesoprotopis) bees (Hymenoptera: Colletidae). *BMC Evolutionary Biology* **10**(1): 1-16.
- Magnacca KN and Brown MJ 2012. DNA barcoding a regional fauna: Irish solitary bees. *Molecular Ecology Resources* **12**(6): 990-998.
- Meixner MD, Pinto MA, Bouga M, Kryger P, Ivanova E and Fuchs S 2013. Standard methods for characterizing subspecies and ecotypes of *Apis mellifera*. *Journal of Apicultural Research* **52**(4): 1-28.
- Nedic N, Jevtic G, Jez G, Anđelkovic B, Milosavljevic S and Kostic M 2011. Forewing differentiation of the honey bees from Serbia. *Biotechnology in Animal Husbandry* **27**(3): 1387-1394.
- Nneji LM, Adeola AC, Ayoola AO, Oladipo SO, Wang YY, Malann YD, Anyaele O, Nneji IC, Rahman MM and Olori CS 2020. DNA barcoding and species delimitation of butterflies (Lepidoptera) from Nigeria. *Molecular Biology Reports* **47**(12): 9441-9457.
- Oh HK, Yoon HJ, Lee JY, Park JS and Kim I 2013. Population genetic structure of the bumblebee, *Bombus ignitus* (Hymenoptera: Apidae), based on mitochondrial COI gene and nuclear ribosomal ITS2 sequences. *International Journal of Industrial Entomology* **27**(1): 142-158.
- Okuyama H, Takahashi J, Ilyasov R and Goto K 2017. Direct Submission, Submitted (16-OCT-2017) Contact: Hisashi Okuyama Kyoto Sangyo University, Faculty of Life Sciences; Kamigamo-Motoyama, Kita-ku, Kyoto, Kyoto 608-8555, Japan (NCBI accession number AP018404.1 of *A. mellifera* cited on 5-08-2021) <https://www.ncbi.nlm.nih.gov/nuccore/1767334393>
- Oldroyd BP, Reddy MS, Chapman NC, Thompson GJ and Beekman M 2006. Evidence for reproductive isolation between two colour morphs of cavity nesting honey bees (*Apis*) in south India. *Insectes Sociaux* **53**(4): 428-434.



- Ollerton J 2017. Pollinator diversity: distribution, ecological function, and conservation. *Annual Review of Ecology, Evolution, and Systematics* **48**: 353-376.
- Ollerton J, Winfree R and Tarrant S 2011. How many flowering plants are pollinated by animals?. *Oikos* **120**(3): 321-326.
- Özdil F and İlhan F 2012. Phylogenetic relationship of Turkish *Apis mellifera* subspecies based on sequencing of mitochondrial cytochrome C oxidase I region. *Genetics and Molecular Research* **11**(2): 1130-1141.
- Packer L and Ruz L 2017. DNA barcoding the bees (Hymenoptera: Apoidea) of Chile: species discovery in a reasonably well known bee fauna with the description of a new species of *Lonchopria* (Colletidae). *Genome* **60**(5): 414-430.
- Packer L, Gibbs J, Sheffield C and Hanner R 2009. DNA barcoding and the mediocrity of morphology. *Molecular Ecology Resources* **9**: 42-50.
- Pakrashi A, Kundu S, Saini J, Tyagi K, Chandra K, and Kumar V 2020. Molecular identification of selected bees from the Indian Himalaya: A preliminary effort. *Journal of Asia-Pacific Entomology* **23**(4): 883-889.
- Pramual P and Nanork P 2012. Phylogenetic analysis based on multiple gene sequences revealing cryptic biodiversity in *Simulium* multistriatum group (Diptera: Simuliidae) in Thailand. *Entomological Science* **15**(2): 202-213.
- Pramual P, Kuvangkadilok C, Baimai V and Walton C 2005. Phylogeography of the black fly *Simulium tani* (Diptera: Simuliidae) from Thailand as inferred from mtDNA sequences. *Molecular Ecology* **14**(13): 3989-4001.
- Punchihewa RWK, Koeniger N, Kevan PG and Gadawski RM 1985. Observations on the dance communication and natural foraging ranges of *Apis cerana*, *Apis dorsata* and *Apis florea* in Sri Lanka. *Journal of Apicultural Research* **24**(3): 168-175
- Ratnasingham S and Hebert PD 2007. BOLD: The barcode of life data system (<http://www.barcodinglife.org>). *Molecular Ecology Notes* **7**(3): 355-364.
- Rehan SM and Sheffield CS 2011. Morphological and molecular delineation of a new species in the *Ceratina dupla* species-group (Hymenoptera: Apidae: Xylocopinae) of eastern North America. *Zootaxa* **2873**(1): 35-50.
- Safonkin AF, Triseleva TA and Bykova TO 2017. Direct Submission Submitted (12-MAY-2017) Molecular Diagnostic Center, Severtsov Institute of Ecology and Evolution, Leninsky 33, Moscow 119071, Russia (NCBI accession number MF100917.1 of *A. mellifera* cited on 5-08-2021) <https://www.ncbi.nlm.nih.gov/nuccore/MF100917.1>
- Saini J and Chandra K 2019. Current status of diversity and distribution of bee pollinators (Apoidea: Hymenoptera) of India. *ZSI ENVIS Newsletter* **25**(1-4): 22-25.
- Saitou N and Nei M 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* **4**(4): 406-425.
- Sangeetha BG, Asokan R, Jayaprakas CA, Arunkarthick S, Aravintharaj R, Chaitanya BN, Rebijith KB and Niveditha M 2016. Submitted (26-JAN-2016) Biotechnology, Indian Institute of Horticultural Research, Hesaraghatta, Bengaluru, Karnataka 560088, India
- Schroder S, Wittmann D, Drescher W, Roth V, Steinhage V and Cremers AB 2002. The new key to bees: automated identification by image analysis of wings. Pollinating bees—the Conservation Link Between Agriculture and Nature, Ministry of Environment, Brasilia, 209-218.
- Sheffield CS, Hebert PD, Kevan PG and Packer L 2009. DNA barcoding a regional bee (Hymenoptera: Apoidea) fauna and its potential for ecological studies. *Molecular Ecology Resources* **9**: 196-207.
- Sievers F and Higgins DG 2018. Clustal Omega for making accurate alignments of many protein sequences. *Protein Science* **27**(1): 135-145.
- Subbanna ARNS, Stanley J, Jain SK, Bhatt JC and Bisht JK 2016. Phylogeny and genetic divergence of Indian Himalayan population of *Anomala dimidiata* (Coleoptera: Scarabaeidae) inferred from mitochondrial DNA sequences. *Agricultural research* **5**(1): 64-71.
- Sudhagar S, Rami Reddy PV and Sridhar V 2014. Direct Submission, Submitted on (27-APR-2014) Entomology, Indian Institute of Horticultural Research, Hesaraghatta Lake Post, Bangalore, Karnataka 560089, India (NCBI accession number KJ755328 of *A. dorsata* cited on 5-08-2021) <https://www.ncbi.nlm.nih.gov/nuccore/672241652>
- Sudhagar S, Venkata Rami Reddy P and Sridhar V 2016. Submitted (22-MAR-2016) Entomology, Indian Institute of Horticultural Research, Hesaraghatta Lake Post, Bangalore, Karnataka 560089, India
- Tan H W, Liu GH, Dong X, Lin R Q, Song HQ, Huang SY, Yuan ZG, Zhao GH and Zhu XQ 2011. The complete mitochondrial genome of the Asiatic cavity-nesting honeybee *Apis cerana* (Hymenoptera: Apidae). *Plos one* **6**(8): e23008.
- Tang M, Tan M, Meng G, Yang S, Su XU, Liu S, Song W, Li Y, Wu Q, Zhang A and Zhou X 2014. Multiplex sequencing of pooled mitochondrial genomes- a crucial step toward biodiversity analysis using mito-metagenomics. *Nucleic acids research* **42**(22): e166-e166.
- Trunz V, Packer L, Vieu J, Arrigo N and Praz CJ 2016. Comprehensive phylogeny, biogeography and new classification of the diverse bee tribe Megachilini: Can we use DNA barcodes in phylogenies of large genera?. *Molecular Phylogenetics and Evolution* **103**: 245-259.
- Vamosi JC, Gong YB, Adamowicz SJ and Packer L 2017. Forecasting pollination declines through DNA barcoding: The potential contributions of macroecological and macroevolutionary scales of inquiry. *New Phytologist* **214**(1): 11-18.
- Wang Y, Jodoin PM, Porikli F, Konrad J, Benezeth Y and Ishwar P 2014. An expanded change detection benchmark dataset. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops* (pp. 387-394).
- Wang M 2021. Submitted (13-MAY-2021) Key Laboratory of Entomological Biopharmaceutical R&D, Dali University, wanhua road, Dali, Yunnan 671000, China
- Weeks PJD, O'Neill MA, Gaston KJ and Gauld ID 1999. Automating insect identification: Exploring the limitations of a prototype system. *Journal of Applied Entomology* **123**(1): 1-8.
- Will KW and Rubinoff D 2004. Myth of the molecule: DNA barcodes for species cannot replace morphology for identification and classification. *Cladistics* **20**(1): 47-55.
- Williams NM, Regetz J and Kremen C 2012. Landscape-scale resources promote colony growth but not reproductive performance of bumble bees. *Ecology* **93**(5): 1049-1058.
- Willis LG, Winston ML and Honda BM 1992. Phylogenetic relationships in the honeybee (genus *Apis*) as determined by the sequence of the cytochrome oxidase II region of mitochondrial DNA. *Molecular Phylogenetics and Evolution* **1**(3): 169-178.
- Young AM, Kohl PL, Rutschmann B, Steffan-Dewenter I, Brockmann A and Dyer FC 2021. Temporal and spatial foraging patterns of three Asian honey bee species in Bangalore, India. *Apidologie* **52**(2): 503-523.