Original paper

PHYLOGENETIC UNIQUENESS OF HONEYBEE *APIS CERANA* FROM THE KOREAN PENINSULA INFERRED FROM THE MITOCHONDRIAL, NUCLEAR, AND MORPHOLOGICAL DATA

Rustem A. Ilyasov[1,4]*
Junhyung Park[2]
Junichi Takahashi[3]
Hyung Wook Kwon[4]*

[1] Institute of Biochemistry and Genetics, Ufa Federal Research Centre, Russian Academy of Sciences, Ufa, Russia
[2] Bioinformatics company 3BIGS CO. LTD, Hwaseong, Gyeonggi, Korea
[3] Faculty of Life Sciences, Kyoto Sangyo University, Kyoto, Japan
[4] Department of Life Sciences and Convergence Research Center for Insect Vectors, Incheon National University, Incheon, Korea

*co-corresponding authors: RAI (apismell@hotmail.com), HWK (hwkwon@inu.ac.kr)
Received: 8 February 2018; accepted: 11 September 2018

Abstract

***Apis cerana* is an Eastern honeybee species distributed throughout Asia and closely related to the Western honeybee species *Apis mellifera* distributed across all of Africa, Europe and Western Asia, and subdivided into thirty confirmed subspecies. Currently, *A. cerana* is an endangered bee species in contrast to *A. mellifera.* We sequenced and annotated the complete mitochondrial genome of *A. cerana* from the Jeollanam-do province of South Korea and uploaded to the DDBJ/Genbank database(AP018431). MtDNA sequence is 15.925 bp long, has 84% AT-content and 16% GC-content and contains 22 tRNA genes, 13 protein-coding genes, two ribosomal RNA genes, one AT-rich region and four non-coding intergenic regions (*NC1-4*). All protein-coding genes are started by ATT and ATG codons, except the genes *ATP8* and *ND4*, which started by ATC and ATA, respectively, and are stopped by the common codons TAA and TAG. A comparative analysis of the whole mtDNA sequences of *A. cerana* from Korea and Taiwan, *A. c. cerana* from China and *A. c. japonica* from Japan showed that the genetic divergence of the Korean *A. cerana* sample from subspecies *A. c. cerana* (2.57%) and *A. c. japonica* (2.58%) matched to the level of genetic divergence of mtDNA between animal subspecies (0.8-8%). Based on the comparative analysis of complete mtDNA (~16,000 bp), two nuclear gene *VG* and *EF1-*α sequences (~8,000 bp) and morphological measurements (six parameters), we assumed that Korean *A. cerana*, Chinese *A. c. cerana* and Japanese *A. c. japonica* are different subspecies at an early stage of sub-speciation and could be called further as subspecies of *Apis cerana koreana*.**

Keywords: *Apis cerana, A. c. cerana, A. c. japonica, A. c. koreana*, mitochondrial genome, mtDNA

INTRODUCTION

The Hymenoptera is one of the largest orders of class Insecta comprising over 100,000 species. Genus *Apis* is most familiar for us and contains about ten honeybee species (LaSalle & Gauld, 1993; Tan et al., 2011). *Apis cerana* F. is a bee species distributed mainly in Asia and closely related to species of *Apis mellifera* L. *A. cerana* was earlier considered as a sub-species of honeybees *A. mellifera* (Buttel-Reepen, 1906; Rinderer, 1986; Tan et al., 2011) but later, based on behavioral, morphological and genetic characteristics, *A. cerana* and *A. mellifera* were suggested to be sister species (Cameron, 1993; Engel & Schultz, 1997; Arias & Sheppard, 2005; Raffiudin & Crozier, 2007). *A. cerana* occurs sympatrically with *Apis florea* F. within the same area in southern Asia and inhabits allopatrically with *A. mellifera*, whose native area is Africa, Europe

and central and western Asia (Ruttner, 1988; Crane, 1999; Hepburn & Radloff, 2011; Koetz, 2013). The native area of *A. cerana* encompasses a wide range of climatic zones from southern tropical to northern coniferous taiga forests (Hepburn & Radloff, 2011; Koetz, 2013). The current area of *A. cerana* has expanded across the world due to human interference (Koetz, 2013). *A. cerana* is considered to be subdivided into twenty-two subspecies: *A. c. cerana, A. c. skorikovi, A. c. abaensis, A. c. hainanensis, A. c. indica, A. c. japonica, A. c. bijjieca, A. c. cathayca, A. c. fantsun, A. c. heimifeng, A. c. himalaya, A. c. javana, A. c. johni, A. c. kweiyanga, A. c. maerkang, A. c. nuluensis, A. c. pekinga, A. c. peroni, A. c. philippina, A. c. samarensis, A. c. shankianga, A. c. twolareca* (Pesenko et al., 1989; Zhuang, 1989; Zhen-ming et al., 1992; Diniz-Filho, Malapsina & Pignata, 1993; Engel, 1999; Sugawara, 2000; Hepburn et al., 2001; Takahashi & Yoshida, 2003; Radloff et al., 2010; Takahashi et al., 2016). The taxonomy of *A. cerana* is controversial because some subspecies are not widely accepted. Three *A. cerana* subspecies, *A. c. cerana*, *A. c. japonica* and *A. c. indica,* are predominant and used for apiculture in Asia (Radloff et al., 2010). These subspecies could have originated as a result of migration and following geographic isolation in different parts of Asia. Previously, it was assumed on the basis of mtDNA analysis that subspecies *A. c. japonica* originally had come to Japan from the Korean peninsula (Sugawara, 2000; Takahashi & Yoshida, 2003).

Honeybee *A. cerana* is an important and famous pollinator of crops in Asia and a producer of honey, wax, royal jelly and bee pollen (Behura, 2007). *A. cerana* matches *A. mellifera* in commercial use value and has high potentials for further genetic improvement by selective breeding based on molecular markers. Now *A. cerana* is threatened across its native area due to the spreading of Korean Sacbrood Virus (kSBV) and importing of *A. mellifera* (Choi et al., 2010; Koetz, 2013; Vung et al., 2017). Molecular genetic studies allow the development of basic strategies for *A. cerana* conservation. Mitochondrial DNA (mtDNA) markers are useful instruments in phylogenetic reconstruction and evolutionary research of Hymenoptera (Cornuet, Garnery, & Solignac, 1991). The mtDNA with direct maternal inheritance and without recombination allows research on a phylogeny for the animal species. MtDNA has demonstrated to be a useful tool in inter- and intra-specific phylogenetic studies of honeybees (Garnery, Cornuet & Solignac, 1992; Garnery et al., 1995; Arias & Sheppard, 1996; Songrarn, Sittipraneed, & Klinbunga, 2006; Tan et al., 2011).

The exact number of *A. cerana* subspecies is unclear now, so it is especially exciting to research island and peninsula *A. cerana* populations. *A. cerana* populations from the Korean peninsula and Taiwan Island have not an exact taxonomical status yet. These populations possibly belong to separate *A. cerana* subspecies and have genetic differences from the mainland *A. c. cerana* population and the island *A. c. japonica* population. The previous studies on mtDNA loci have not provided information concerning phylogenetic differentiation of Korean *A. cerana* population (Tanaka et al., 2001; Zhao et al., 2013; Lee et al., 2016). We assumed that differences in Korean *A. cerana* populations could be found through comparative analysis of complete mtDNA. In this work, we determined the complete mtDNA sequence of *A. cerana* from the Jeollanam-do Province of South Korea. Phylogenetic analysis based on comparative analysis of the complete mtDNA sequence of Korean, Japanese, Chinese and Taiwan *A. cerana* population was performed. Our findings can be useful to explain the origin and migration of *A. cerana* populations.

MATERIAL AND METHODS

Adult honeybee workers of *A. cerana* were collected from a hive in an apiary at the Hakjung-ri District, Gokseong-eup Town, Gokseong-gun County, Jeollanam-do Province, South Korea. The genomic DNA was extracted from the thoracic muscle tissue with the Wizard Genomic DNA Purification Kit (Promega, Madison, WI, USA) according to the manufacturer's recommendations. The DNA samples were stored at -20°C until further use.

Genomic libraries were prepared with the

Nextera DNA Library Preparation Kit (ILLUMINA, United States) according to the instructions of the manufacturer. Sequencing was prepared with the NextSeq 500/550 High Output Kit v2 (75 cycles) (ILLUMINA, United States) following the instructions of the kit.

Complete mitochondrial DNA was sequenced with paired-end read runs (2 x 150 bp) using the Illumina's Next Seq 500 (ILLUMINA, United States) following the instructions of the reagents at the Department of Life Sciences of Kyoto Sangyo University. The 1,662,000 reads were obtained with seventy-five medium coverage level and were assembled by the Geneious R9 (BIOMATTERS, New Zealand) and annotated using the MITOS web server (Bernt et al., 2013, Germany), and resultant mtDNA consisting of 15,925 bp was uploaded into the DDBJ/Genbank database (AP018431). The phylogenetic analysis was performed using the MEGA7 software (Kumar et al., 2016) based on the nucleotide sequences of complete mtDNA.

To explore the phylogeny of Korean *A. cerana,* additional *A. cerana* samples were taken from the DDBJ/Genbank database: GQ162109, KM244704 (*A. c. cerana*, China, Yunnan) (Tan et al., 2011), AP017983 (*A. cerana*, China, Jiangsu) (Okuyama et al., 2017), AP017314 (*A. c. japonica*, Japan, Kyoto) (Takahashi et al., 2016), AP017941 (*A. c. japonica,* Japan, Amami), AP017985 (*A. c. japonica,* Japan, Tsushima), AP017984 (*A. cerana*, Taiwan, Taipei) (Okuyama et al., 2017) (Fig. 1). The GQ162109 (*A. c. cerana*, China, Yunnan) (Tan et al., 2011) was used as the reference sequence, and JX982136, KC170303 (*A. florea*, ancestral species) (Lindaur, 1956) - as outgroup sequences.

Pairwise nucleotide sequence divergences were estimated using the Unipro UGENE 1.28 (UNIPRO, Russia) and the CLC Genomics Workbench 11 (CLCbio, Denmark) on the basis of complete mtDNA sequences using the Jukes-Cantor (Jukes & Cantor, 1969), Tamura-Nei (Tamura & Nei, 1993) and the p-distance models (Nei & Kumar, 2000).

Fig. 1. Geographical distribution of *A. cerana* samples used in comparative analysis of the complete mtDNA sequences and two possible ways of their migration to North-East Asia.

Phylogenetic tree constructed using the Reltime method (Tamura et al., 2012) and branch lengths estimated inferred using the Neighbor-Joining method (Saitou & Nei, 1987) based on the Jukes-Cantor model with 1000 bootstrap replications (Saitou & Nei, 1987).

The morphological tests of *A. cerana* samples were performed using the Statistica 8.0 statistical package (StatSoft Power Solutions, Inc., USA) based on the well-known methods of multivariate analysis - factor analysis (principal component analysis, PCA) and cluster analysis (Neighbor-Joining (NJ) clustering method based on the Euclidean distances) of the previously published morphological data (forewing's length and width, cubital index, hindleg length, metatarsal index, and the length of tergite 3+4th) of the Korean (N=40), Chinese (N=124) and Japanese (N=8) *A. cerana* samples (Lee & Choi, 1986; Ruttner, 1988; Ken et al., 2003) which were measured according to the methods described by Ruttner (1988).

RESULTS

The complete mtDNA of Korean *A. cerana* have not been studied and published. The last comparative research of *A. cerana* populations from China, Indonesia, Korea, Malaysia, Russia, Taiwan, Thailand, Vietnam and Japan based on the *NC1* non-coding intergenic region of mtDNA found nine haplotypes of *NC1* region, ten haplotypes of *NC2* region and seventy-eight haplotypes of *COX1* gene which could not be differentiated Korean *A. cerana* (Tanaka et al., 2001; Zhao et al., 2013; Lee et al., 2016).

The complete mtDNA of Korean *A. cerana* is 15,925 bp and contains local DNA regions of high and low compositional complexity. The complexity of DNA depends on features of the sequence A, T, G, C nucleotides. Local DNA regions include different types of nucleotide clusters, some of which contain AT-rich and GC-rich regions, short tandem repeats or aperiodic mosaics of all nucleotides. The low local complexity regions alternate by high local complexity. An equable distribution of high

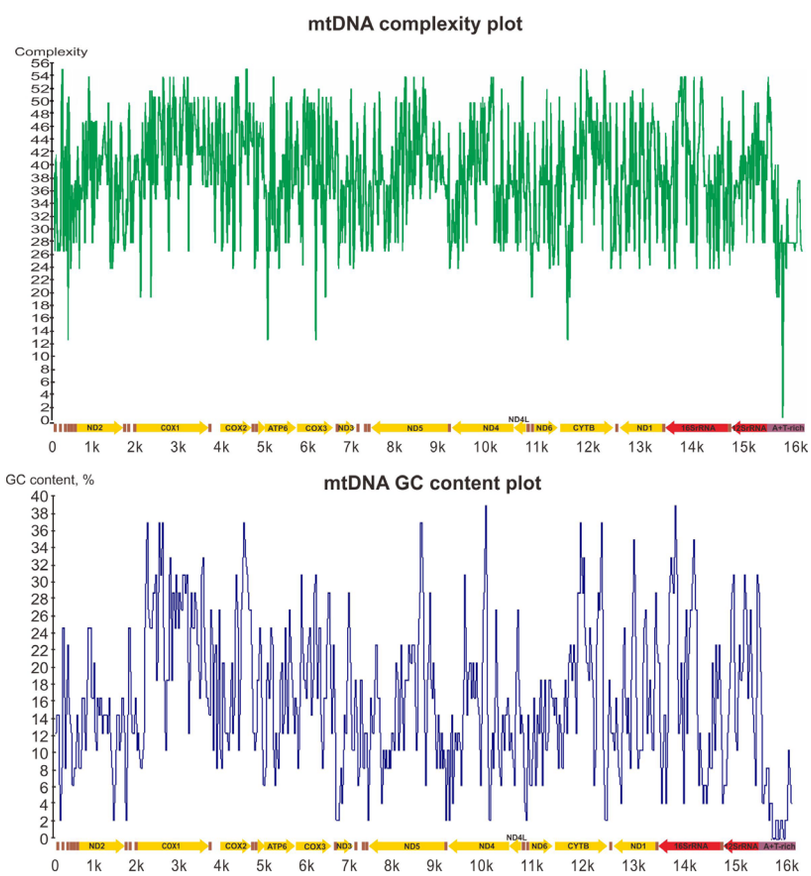

Fig. 2. Complexity and GC-content plots of Korean *A. cerana* complete mtDNA.

and low local complexity loci throughout the complete mtDNA sequence can be observed on the complexity and GC-content plots (Fig. 2).

The GC-content (guanine-cytosine content) plot have been constructed for the Korean *A. cerana* complete mtDNA (Fig. 2). GC-content is usually expressed as a percentage value, but sometimes as GC-ratio. The GC-content percentage is calculated as the AT/GC-ratio.

The major properties of the mtDNA of Korean *A. cerana* were calculated (Tab. 1). Overall, it comprises 42.3% A, 41.8% T, 6.2% G, and 9.7% C, contains at most AT (84%) and at least CG (16%), which very similar to *A. c. japonica* (AP017941, AP017985, AP017314), *A. c. cerana* (AP017984, AP017983, KM244704, GQ162109) (Tan et al., 2011; Takahashi et al., 2016; Okuyama et al., 2017), *A. m. ligustica* (NC_001566) (Crozier & Crozier, 1993), *A. m. caucasica* (AP018404), *A. m. carpathica* (AP018403), *A. m. meda* (KY464957) and *A. m. syriaca* (GenBank -KY926882).

Korean *A. cerana* mtDNA is characterized by the highest frequencies of AA, AT, TT and TA dinucleotides (>16% each) and lowest frequencies of GG, GC, CG dinucleotides (<1% each). This may be a consequence of AT-richness of most honeybee mtDNA.

The genes of Korean *A. cerana* were identified on the basis of similarity with the structural organization of *A. mellifera* mtDNA. We constructed the circular physical map of the complete mtDNA of Korean *A. cerana* (Fig. 3). The thirteen protein-coding genes, two ribosomal RNA genes, twenty-two tRNA genes, and AT-rich region

Table 1.

Properties of complete mtDNA of Korean *A. cerana*

Parameter	Count (freguency)
Adenine (A)	6741 (0.423)
Cytosine (C)	1537 (0.097)
Guanine (G)	994 (0.062)
Thymine (T)	6651 (0.418)
GC-content	2531 (0.159)
AT-content	13392 (0.841)
Dinucleotide AA	2974 (0.187)
Dinucleotide AC	477 (0.030)
Dinucleotide AG	358 (0.022)
Dinucleotide AT	2931 (0.184)
Dinucleotide CA	674 (0.042)
Dinucleotide CC	236 (0.015)
Dinucleotide CG	67 (0.004)
Dinucleotide CT	560 (0.035)
Dinucleotide GA	472 (0.03)
Dinucleotide GC	118 (0.007)
Dinucleotide GG	135 (0.008)
Dinucleotide GT	269 (0.017)
Dinucleotide TA	2621 (0.165)
Dinucleotide TC	706 (0.044)
Dinucleotide TG	433 (0.027)
Dinucleotide TT	2890 (0.182)
Length	15925 bp
Weight	4906.855 kDa
Protein-coding genes	13
tRNA	22
rRNA	2

are all at the similar positions relative to the reference *A. cerana* sequence (GQ162109) (Tab. 2). All protein-coding genes are begun by ATT and ATG start codons, except the genes *ATP8* and *ND4*, which started by ATC and ATA, respectively, and are finished by the common TAA and TAG stop codons. We found four non-coding intergenic loci (*NC1-4*) on the circular map of Korean *A. cerana* complete mtDNA as well (Fig. 3).

Korean *A. cerana* complete mtDNA sequence is slightly longer (15,925 bp) than all published *A. cerana* mtDNA sequences AP017985 (15,339 bp), AP017984 (15,376 bp), AP017983 (15,460 bp), KM244704 (15,712 bp), AP017941 (15,788 bp), GQ162109 (15,895 bp), AP017314 (15,917 bp), AP018431 (15,925 bp).

Similar to other *A. cerana* samples (Lee et al., 2016), Korean *A. cerana* mtDNA sequence has four relatively big intergenic regions: *NC1* (228 bp) located between genes *tRNA-Met* and *tRNA-Gln*; *NC2* (89 bp) located between genes *tRNA-Leu*$^{(TAA)}$ and *COX2*; *NC3* (68 bp) located between genes *COX3* and *tRNA-Gly*; *NC4* (51 bp) located between genes *tRNA-Pro* and *ND6*. These regions constitute about 1.3 - 0.3% of the entire mtDNA size.

The Korean *A. cerana* mtDNA (15,925 bp) is slightly shorter than that of *A. mellifera* mtDNA (16,343 bp) and *Drosophila yakuba* (16,019 bp) has a shorter AT-rich control region (565 bp) than seen in *A. mellifera* (827 bp) and *D. yakuba* mtDNA (1,077 bp). All non-coding intergenic sequences in Korean *A. cerana* mtDNA (1298 bp) are slightly larger than that for *D. yakuba* (1,262 bp) and shorter than that for *A. mellifera* (1,639 bp) (Clary & Wolstenholme, 1985; Crozier & Crozier, 1993) (Tab. 2).

Nine protein-coding genes *CYTB, COX1, COX2, COX3, ND2, ND3, ND6, ATP6,* and *ATP8* are transcribed from the light strand, while four protein-coding genes *ND1, ND4*, *ND4L*, and *ND5* are transcribed from the heavy strand of mtDNA. Three overlap points are found in the light strand between genes *ATP8* and *ATP6*, *COX1* and *tRNA-Leu*$^{(TAA)}$, *COX2* and *tRNA-Asp* and one overlap between gene *NAD2* of light strand and gene *tRNA-Cys* of the heavy strand. This over-

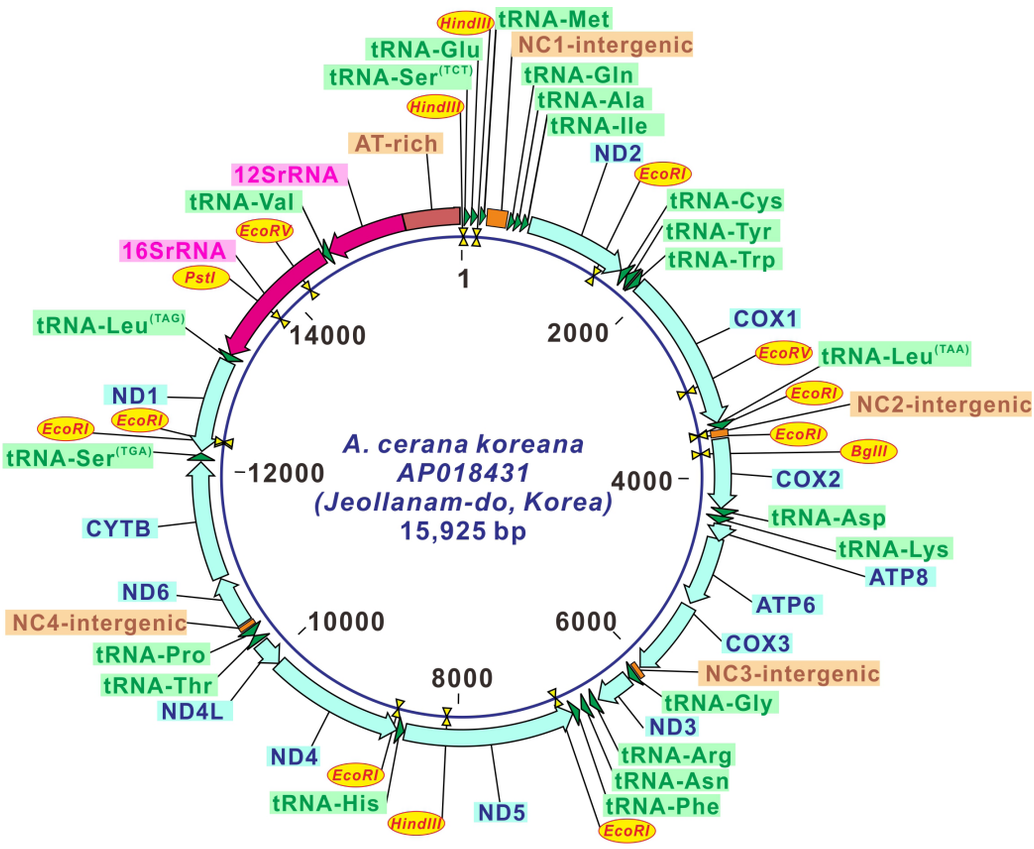

Fig. 3. Circular map of Korean *A. cerana* complete mtDNA.

lapping of mtDNA genes proves its origin from the prokaryotic genome with the polycistronic type of transcription (Tab. 2).

A comparative analysis of the Korean *A. cerana* mtDNA with all other *A. cerana* samples found the greatest divergence from Taiwanese sample, middle divergence from Chinese samples, and least divergence from Japanese samples (Tab. 3). The genes *ATP6, COX1, COX2, COX3, CYTB, ND1, ND2, ND3, ND4, ND5, ND6* are most polymorphic and more informative for phylogenetic studies of *A. cerana*. The genes *ATP8, ND4L,*

Table 2.

Short annotation of complete mtDNA of Korean *A. cerana* (AP018431)

Gene / Region	Strand	Start position	End position	Size, bp
tRNA-Ser(TCT)	plus	1	60	60
tRNA-Glu	plus	64	129	66
tRNA-Met	plus	164	229	66
tRNA-Gln	plus	458	519	62
tRNA-Ala	plus	520	585	66
tRNA-Ile	plus	604	669	66
ND2	plus	670	1665	996
tRNA-Cys	minus	1665	1730	66
tRNA-Tyr	minus	1736	1804	69
tRNA-Trp	plus	1821	1889	69
COX1	plus	1890	3455	1566
tRNA-Leu(TAA)	plus	3451	3520	70
COX2	plus	3610	4290	681
tRNA-Asp	plus	4289	4356	68
tRNA-Lys	plus	4363	4434	72
ATP8	plus	4441	4602	162
ATP6	plus	4584	5261	678
COX3	plus	5300	6079	780
tRNA-Gly	plus	6148	6214	67
ND3	plus	6215	6568	354
tRNA-Arg	plus	6589	6654	66
tRNA-Asn	plus	6690	6757	68
tRNA-Phe	minus	6779	6849	71
ND5	minus	6850	8520	1671
tRNA-His	minus	8528	8593	66
ND4	minus	8601	9929	1329
ND4L	minus	9930	10193	264
tRNA-Thr	plus	10213	10279	67
tRNA-Pro	minus	10295	10372	78
ND6	plus	10424	10930	507
CYTB	plus	10943	12091	1149
tRNA-Ser(TGA)	plus	12115	12181	67
ND1	minus	12194	13108	915
tRNA-Leu(TAG)	minus	13109	13177	69
16S rRNA	minus	13178	14507	1330
tRNA-Val	minus	14508	14574	67
12S rRNA	minus	14575	15360	786
AT-rich region	plus	15361	15925	565
Total				15925

Table 3.

Transitions and transversions (underlined) and amino acid replacements in mtDNA between Korean and other samples of *A. cerana* with positions relative to the reference sequence.

Genes / Regions	China, Yunnan, GQ162109, 15895 bp	China, Yunnan, KM244704, 15712 bp	China, Jiangsu, AP017983, 15460 bp
tRNA-Ser(TCT)	-	-	-
tRNA-Ala	-	573A>G	-
tRNA-Glu	-	-	-
tRNA-Ile	-	-	-
ND2	678A>G ; 885C>T; 1276A>G (Met >Val); 1362C>T; 1579G>A (Asp>Asn); 1614A>G; 1652C>T (Thr >Met); 1642T>A (Leu >Met)	678A>G ; 834C>T; 885C>T; 1276A>G (Met >Val); 1362C>T; 1482C>T; 1614A>G; 1621C>T; 1642T>A (Leu >Met); 1652C>T (Thr >Met);	678A>G ; 831G>A; 834C>T; 885C>T; 1276A>G (Met >Val); 1362C>T; 1482C>T; 1614A>G; 1642T>A (Leu >Met); 1652C>T (Thr >Met);
tRNA-Cys	-	-	-
COX1	2018T>C; 2192T>C; 2198G>A; 2309C>T; 2330T>C; 2423T>C; 2465T>C; 2603T>C; 2643T>C; 2873A>T	2018T>C; 2192T>C; 2309C>T; 2330T>C; 2423T>C; 2465T>C; 2873A>T; 3885G>A	1991G>A; 2018T>C; 2309C>T; 2330T>C; 2423T>C; 2465T>C; 2873A>T; 3885G>A
tRNA-Leu(TAA)	-	-	-
COX2	3651C>T; 3885G>A; 3989A>G (Lys>Ser); 3994T>C; 4060C>T	3989A>G (Lys>Ser); 3994T>C; 4060C>T; 4113A>G	3921A>G; 3989A>G (Lys>Ser); 3994T>C; 4060C>T; 4113A>G
tRNA-Asp	-	-	-
tRNA-Lys		-	-
ATP8	4507C>T	4507C>T	4507C>T
ATP6	4640C>T; 4644T>C; 4670T>C; 4932C>T; 5049T>A (Trp> Arg); 5115T>C; 5142G>T (Ala>Ser)	4640C>T; 4670T>C; 4932C>T; 5049T>A (Trp> Arg); 5142G>T (Ala>Ser)	4640C>T; 4670T>C; 4932C>T; 5049T>A (Trp> Arg); 5142G>T (Ala>Ser)5115T>C;
COX3	5299G>A; 5320G>A; 5332T>C; 5358T>C (Leu >Ser); 5366A>G (Ser>Gly); 5842A>T (Leu >Phe)	5299G>A; 5320G>A; 5332T>C; 5358T>C (Leu >Ser); 5366A>G (Ser>Gly); 5842A>T (Leu >Phe)	5299G>A; 5320G>A; 5332T>C; 5358T>C (Leu >Ser); 5366A>G (Ser>Gly); 5842A>T (Leu >Phe)
ND3	6264A>G (Thr>Ala); 6416C>T; 6418T>C; 6458T>C	6264A>G (Thr>Ala); 6416C>T	6231T>C; 6264A>G (Thr>Ala); 6416C>T
tRNA-Arg	6585G>A; 6601T>C	6585G>A; 6601T>C	6585G>A; 6601T>C
tRNA-Asn	-	-	-
ND5	6822C>T; 6963T>C; 8268A>G; 8309T>C(Ile >Val); 7961T>A (Met>Leu)	6822C>T; 6963T>C; 6981A>G; 7961T>A (Met>Leu); 8268A>G;	6822C>T; 6963T>C; 7842A>G; 7896T>C; 7961T>A (Met>Leu); 8067A>T; 8268A>G; 8309T>C(Ile >Val); 8343A>G; 8392A>T (Met>Lys)
ND4	8735A>G; 8816G>A; 8906T>C; 8964A>G (Phe >Ser); 9053A>G; 9500T>C; 8942T>A; 8969A>T (Phe >Leu); 8972A>T; 9482T>A; 9650 ATT>TAA (Asn>Leu); 9657T>A (Lys>Met); 9704A>T; 9719A>T	8906T>C; 8942T>A; 8964A>G (Phe >Ser); 8969A>T (Phe >Leu); 8972A>T; 9482T>A; 9500T>C; 9626T>C; 9650 ATT>TAA (Asn>Leu); 9657T>A (Lys>Met); 9704A>T; 9719A>T;	8906T>C; 8942T>A; 8964A>G (Phe >Ser); 8969A>T (Phe >Leu); 9053A>G; 8972A>T; 9482T>A; 9626T>C; 9650 ATT>TAA (Asn>Leu); 9657T>A (Lys>Met); 9704A>T; 9719A>T;
ND4L	-	-	-
ND6	10538T>C (Val> Ala); 10630A>G (Ile >Val); 10781G>A (Gly >Glu); 10872C>T	10538T>C (Val> Ala); 10630A>G (Ile >Val); 10872C>T	10538T>C (Val> Ala); 10630A>G (Ile >Val); 10872C>T
CYTB	10935T>C; 11592T>C; 11808T>C; 11871T>C; 11376T>A; 11394A>T	10935T>C; 11376T>A; 11394A>T; 11709A>G; 11871T>C; 11973T>C	10935T>C; 11376T>A; 11394A>T; 11592T>C; 11709A>G; 11871T>C; 11934G>A; 11955T>C; 12010A>T (Met>Leu)
ND1	12346G>A; 12922T>C; 12592A>T (Phe >Leu)	12346G>A; 12592A>T (Phe >Leu); 12922T>C;	12346G>A; 12592A>T (Phe >Leu); 12922T>C; 12934G>A;
tRNA-Leu(TAG)	-	14442A>C	14442A>C
16S rRNA	14452T>C; 14442A>C	14452T>C;	13968T>A; 14121G>A; 14188G>A; 14452T>C

12S rRNA	14866C>T; 15026G>A; 15065T>C	14866C>T; 15026G>A; 15065T>C;	14866C>T; 14967C>T; 15026G>A; 15065T>C;15095C>T
AT-rich region	15432C>T; 15474A>G; 15517T>C; 15548A>G; 15554A>G; 15557A>G; 15601T>C; 15709G>A; 15463T>A; 15587T>A; 15590A>T; 15611T>A;	15432C>T; 15453T>C; 15463T>A; 15474A>G; 15517T>C; 15548A>G; 15554A>G; 15557A>G; 15587T>A; 15590A>T; 15601T>C; 15709G>A; 15451T>A	15432C>T; 15474A>G; 15517T>C; 15463T>A; 15548A>G; 15554A>G; 15557A>G; 15587T>A; 15590A>T; 15601T>C; 15709G>A;
Intergenic regions	278C>T; 307C>T; 323C>A; 339A>G; 5266T>C; 5276G>A; 5273G>T; 5278A>T; 6635T>A; 6550T>A; 6555T>A; 6559A>T; 6566A>T; 6636T>A; 6637T>T; 6651A>T; 6721T>A; 6724T>A; 6725 A>T; 6732T>A; 6733TA>T; 6734A>T; 6735T>A	307C>T; 339A>G; 596T>C; 5266T>C; 5273G>T; 5276G>A; 5278A>T; 6550T>A; 6555T>A; 6559A>T; 6566A>T; 6636T>A; 6651A>T; 6721T>A; 6724T>A; 6725 A>T; 6732T>A; 6733TA>T; 6734A>T; 6735T>A;10379G>A	307C>T; 339A>G; 596T>C; 5266T>C; 5273G>T; 5276G>A; 5278A>T; 6550T>A; 6555T>A; 6559A>T; 6566A>T; 6636T>A; 6651A>T; 6721T>A; 6724T>A; 6725 A>T; 6732T>A; 6733TA>T; 6734A>T; 6735T>A;

Table 3.

Transitions and transversions (underlined) and amino acid replacements in mtDNA between Korean and other samples of *A. cerana* with positions relative to the reference sequence. (continued)

Genes / Regions	Japan, Tsushima, AP017985, 15339 bp	Japan, Kyoto, AP017314, 15917 bp	Japan, Amami, AP017941, 15788 bp	Taiwan, Taipei, AP017984, 15376 bp
tRNA-Ser(TCT)	-	-	-	8T>C
tRNA-Ala	-	-	-	-
tRNA-Glu	-	-	-	116T>C
tRNA-Ile	-	-	-	658C>T; 656T>A
ND2	678A>G ; 885C>T; 1276A>G (Met >Val); 1384A>G (Ile>Val); 1482C>T; 1614A>G	678A>G ; 885C>T; 1276A>G (Met >Val); 1482C>T; 1614A>G	678A>G ; 885C>T; 1276A>G (Met >Val); 1482C>T	738G>A; 750C>T; 831G>A; 858T>C; 885C>T; 906A>G; 984C>T; 1098T>C; 1276A>G (Met >Val); 1281T>C; 1317T>C; 1362C>T; 1452C>T; 1482C>T; 1494T>C; 741A>T; 930A>T; 1173A>C; 1347A>T; 1620A>T
tRNA-Cys	-	-	-	1681C>T
COX1	2018T>C; 2048T>C; 2192T>C	2018T>C; 2048T>C; 2192T>C; 3885G>A	2018T>C; 2048T>C; 2192T>C; 2700G>A; 3885G>A	2015C>T; 2018T>C; 2030T>C; 2060A>G; 2075C>T; 2087T>C; 2105C>T; 2192T>C; 2198G>A; 2201T>C; 2330T>C; 2334T>C; 2378A>G; 2480T>C; 2490C>T; 2573T>C; 2675A>G; 2697C>T; 2700G>A; 2795T>C; 2867T>C; 2873A>T; 2999C>T; 3027T>C; 3081T>C; 3114C>T; 3206C>T; 3215T>C; 3425C>T; 3431C>T; 3433T>C (Met>Thr); 2276T>G; 2462A>T; 2660T>A; 3107T>A; 3269A>C
tRNA-Leu(TAA)	-	-	-	3504T>A
COX2	3885G>A; 3933C>T	-	-	3648A>G; 3778C>T; 3885G>A; 3915C>T; 3927T>C; 3942C>T; 3954C>T; 3989A>G (Lys>Ser); 4012C>T; 4060C>T; 4179C>T; 4278C>T; 3745C>A; 4029T>A; 4209A>T
tRNA-Asp	-	-	-	4346T>A

tRNA-Lys	-	-	-	4381T>C
ATP8	-	-	-	4576G>A (Glu>Lys); 4593C>T; 4597C>T (His>Tyr)
ATP6	5049T>A (Trp> Arg)	5049T>A (Trp> Arg)	5049T>A (Trp> Arg)	4640C>T; 4683GTC>ATT (Val>Ile); 4688C>T; 4742T>C; 4829C>T; 4905T>C; 4937C>T; 4946T>C; 4985T>C; 4998G>A (Val>Ile); 5049T>A (Trp> Arg); 5099C>T; 5138T>C; 5147C>T; 4663T>A (Phe>Tyr); 4682A>T; 4778T>A; 5051A>T; 5258T>A
COX3	5358T>C (Leu >Ser); 5427C>T (Thr>Met); 5476A>T; 5823C>A (Ala>Glu); 5842A>T (Leu >Phe)	5358T>C (Leu >Ser); 5427C>T (Thr>Met); 5476A>T; 5842A>T (Leu >Phe)	5299G>A; 5358T>C (Leu >Ser); 5427C>T (Thr>Met); 5476A>T; 5842A>T (Leu >Phe)	5299G>A; 5311A>G; 5320G>A; 5336T>C; 5358T>C (Leu >Ser); 5377C>T; 5421C>T (Thr>Met); 5429GTA>ACA (Val>Thr); 5432GTA>ACA (Val>Thr); 5452C>T; 5503T>C; 5513T>C; 5531C>T; 5540T>C; 5549C>T; 5578C>T; 5617T>C; 5653T>C; 5842A>T (Leu >Phe); 5917T>C; 6016T>C; 6028A>G; 5417A>T (Met>Leu); 5422T>A (Thr>Met); 5524T>A; 5530T>A; 5632A>C; 5698C>A; 5962A>T
ND3	6399G>A (Val>Ile)	-	-	6209C>T; 6236T>C; 6320T>C; 6410T>C; 6416C>T; 6476C>T; 6494A>G; 6530T>C; 6377A>T
tRNA-Arg	6585G>A; 6601T>C	6585G>A; 6601T>C	6585G>A; 6601T>C	6585G>A; 6619A>T
tRNA-Asn	-	-	-	6667A>T; 6670A>T; 6697A>T
ND5	6922G>A (Thr>Met); 7295A>G; 7455A>G; 7458T>C; 7560A>G; 8076T>C; 8373G>A	7455A>G; 7458T>C; 7560A>G; 8076T>C; 8268A>G; 6886C>A (Ser>Met)	7295A>G; 7455A>G; 7458T>C; 7560A>G; 8076T>C; 8268A>G; 6886C>A (Ser>Met)	6854T>C (Thr>Ala); 6861T>C; 6956C>T (Val>Ile); 6963T>C; 7004C>T (Val>Ile); 7068C>T; 7134A>G; 7176C>T; 7212A>G; 7370C>T (Val>Met); 7557C>T; 7629C>T; 7767A>G; 7785A>G; 7881C>T; 7944T>C; 7946G>A; 7980T>C; 8073T>C; 8166A>G; 8249A>G; 8313C>T; 8322T>C; 7353T>A; 7569T>A; 7725T>A; 7779A>T (Phe>Leu); 7920T>A; 7998T>A; 8004A>T (Phe>Leu); 8067A>T; 8193T>A
ND4	8816G>A; 8906T>C; 8942T>A; 8964A>G (Phe >Ser); 8969A>T (Phe >Leu); 8972A>T; 9443T>C; 9449A>G; 9482T>A; 9614T>C; 9650 ATT>TAA (Asn>Leu); 9657T>A (Lys>Met); 9704A>T; 9719A>T; 9896T>A (Met>Phe); 9898T>A (Met>Phe)	8816G>A; 8906T>C; 8942T>A; 8964A>G (Phe >Ser); 8969A>T (Phe >Leu); 8972A>T; 9443T>C; 9449A>G; 9482T>A; 9614T>C; 9650 ATT>TAA (Asn>Leu); 9657T>A (Lys>Met); 9704A>T; 9719A>T; 9896T>A (Met>Phe); 9898T>A (Met>Phe)	8816G>A; 8906T>C; 8942T>A; 8964A>G (Phe >Ser); 8969A>T (Phe >Leu); 8972A>T; 9443T>C; 9449A>G; 9482T>A; 9614T>C; 9639G>A (Thr>Met); 9650 ATT>TAA (Asn>Leu); 9657T>A (Lys>Met); 9704A>T; 9719A>T; 9896T>A (Met>Phe); 9898T>A (Met>Phe)	8792T>C; 8813T>C; 8816G>A; 8821C>T; 8906T>C; 8909T>C; 8927C>T; 8942T>A; 8964A>G (Phe >Ser); 8969A>T (Phe >Leu); 8972A>T; 9148T>C (Ile>Val); 9155C>T; 9206C>T; 9281T>C; 9286G>A; 9461A>G; 9482T>A; 9506C>T; 9593T>C; 9641T>C; 9650 ATT>TAA (Asn>Leu); 9657T>A (Lys>Met); 9668T>C; 9698T>C; 9704A>T; 9719A>T; 9737A>G; 9770G>A; 9844C>T; 9880A>G; 8801A>C; 9098T>A; 9154A>T (Phe>Ile); 9248A>T; 9269T>A; 9308G>T; 9418A>T (Leu>Met); 9694T>A (Met>Leu); 9889T>A (Ile>Phe)
ND4L	10046A>G	10046A>G	10046A>G	9910A>G; 9956T>C; 9964G>A; 10010T>C; 10115T>C

ND6	10490C>T (Thr>Met); 10630A>G (Ile >Val)	10630A>G (Ile >Val)	10630A>G (Ile >Val); 10747T>C	10413A>G; 10464C>T; 10540C>T; 10554T>C; 10563T>C; 10623C>T; 10630A>G (Ile >Val); 10671C>T; 10761T>C; 10810T>C; 10816C>T; 10824A>G; 10872C>T; 10878A>G; 7961A>T (Leu>Met); 10713A>T (Leu>Phe)
CYTB	11059G>A; 11505T>C; 11583T>C	11059G>A; 11505T>C; 11583T>C	11505T>C; 11583T>C	10959A>G; 11184C>T; 11319A>G; 11371C>T; 11376T>A; 11379T>C; 11394A>T; 11469C>T; 11481T>C; 11505T>C; 11559C>T; 11634T>C; 11682T>C; 11691T>C; 11757C>T; 11775C>T; 11757A>G (Ile>Val); 11844T>C; 11916T>C; 11934G>A; 11944T>C; 11967T>C; 11985T>C; 11061T>A (Val>Ile); 11079T>A; 11181T>A; 11292T>A; 11406C>A; 11790T>A; 11796A>C; 12010A>T (Met>Leu)
ND1	12346G>A	12346G>A; 12511C>T	12346G>A; 12511C>T	12346G>A; 12350C>T (Ser>Asn); 12487C>T; 12574C>T; 12592A>T (Phe >Leu); 12631T>C; 12688T>C; 12765T>C (Ile>Val); 12766T>C; 12775A>G; 12802T>C; 12934G>A; 13038T>C (Ile>Val); 12691T>A; 12691T>A
tRNA-Leu(TAG)	13101A>T	13101A>T	13101A>T	-
16S rRNA	14188G>A; 13869T>A	14174C>T; 14188G>A	14174C>T; 14188G>A	13779C>T; 14037T>C; 14130A>G; 14188G>A; 14243C>T; 14247T>C; 14336C>T; 14441G>A; 14446T>C; 14841T>A; 13179A>T; 13450A>T; 14344A>T; 14345A>T; 14375T>A; 14426C>A; 14447T>A; 14461A>C
12S rRNA	-	15026G>A	15026G>A	14644C>T; 14792C>T; 14793T>C; 14961A>G; 15026G>A; 15048T>C; 15065T>C;15095C>T; 15220T>C
AT-rich region	15432C>T; 15463T>A; 15474A>G; 15517T>C; 15548A>G; 15554A>G; 15557A>G; 15587T>A; 15590A>T; 15601T>C; 15611T>A; 15709G>A;	15474A>G; 15517T>C; 15548A>G; 15554A>G; 15557A>G; 15601T>C; 15709G>A	15432C>T; 15474A>G; 15517T>C; 15548A>G; 15554A>G; 15557A>G; 15601T>C; 15587T>A; 15590A>T; 15709G>A;	15397A>G; 15432C>T; 15474A>G; 15517T>C; 15548A>G; 15463T>A; 15554A>G; 15557A>G; 15587T>A; 15590A>T; 15601T>C; 15611T>A; 15709G>A;
Intergenic regions	281T>C; 307C>T; 339A>G; 3532G>A; 5266T>C; 5273G>T; 5276G>A; 5278A>T; 5268A>G; 6550T>A; 6555T>A; 6559A>T; 6566A>T; 6635T>A; 6636T>A; 6651A>T; 6721T>A; 6724T>A; 6725 A>T; 6732T>A; 6733TA>T; 6734A>T; 6735T>A	307C>T; 339A>G; 5266T>C; 5268A>G; 5273G>T; 3576C>T;;5276G>A; 5278A>T; 6550T>A; 6555T>A; 6559A>T; 6566A>T; 6636T>A; 6651A>T; 6721T>A; 6724T>A; 6725 A>T; 6732T>A; 6733TA>T; 6734A>T; 6735T>A; 6637T>T	307C>T; 339A>G; 5266T>C; 5268A>G; 5273G>T; 5276G>A; 5278A>T; 6110A>T; 6111T>A; 6550T>A; 6555T>A; 6559A>T; 6566A>T; 6636T>A; 6637T>T; 6651A>T; 6721T>A; 6724T>A; 6725 A>T; 6732T>A; 6733TA>T; 6734A>T; 6735T>A	278C>T; 307C>T; 323C>A; 324C>T; 311C>T; 339A>G; 335AG; 355C>T; 406T>C; 594G>A; 595C>T; 5273G>T; 5276G>A; 5278A>T; 6550T>A; 6641A>G; 10356C>T; 10909T>C; 138A>T; 244T>A; 353A>T; 3598A>T; 6555T>A; 6559A>T; 6566A>T; 6635T>A; 6636T>A; 6651A>T; 6721T>A; 6724T>A; 6725 A>T; 6732T>A; 6733TA>T; 6734A>T; 6735T>A;10365A>T; 10911R>G; 12074T>A

12S rRNA, 16S rRNA, and 11 tRNA (*tRNA-Ala, tRNA-Arg, tRNA-Asn, tRNA-Asp, tRNA-Cys, tRNA-Glu, tRNA-Ile, tRNA-Leu$^{(TAA)}$, tRNA-Leu$^{(TAG)}$, tRNA-Lys, tRNA-Ser$^{(TCT)}$*) are the least polymorphic and less informative for phylogenetic studies of *A. cerana*. The remaining 11 *tRNA* genes (*tRNA-Gln, tRNA-Gly, tRNA-His, tRNA-Met, tRNA-Phe, tRNA-Pro, tRNA-Ser$^{(TGA)}$, tRNA-Thr, tRNA-Trp, tRNA-Tyr, tRNA-Val*) are conservative between *A. cerana* subspecies (Tab. 3).

The ratio of transitions/transversions and the number of amino acid substitutions were counted between Korean and all other *A. cerana* samples (Tab. 4). Korean *A. cerana* differs the most from Taiwanese *A. cerana* (AP017984) (258 transitions, 58 transversions and 86 amino acid replacements) and the least from Japanese *A. cerana* (AP017314) (48 transitions, 13 transversions, and 30 amino acid replacements) (Tab. 4).

Four non-coding regions (*NC1-NC4*) were described in the mtDNA of *A. cerana* (Tan, Warrit, & Smith, 2007; Lee et al., 2016). Non-coding region *NC2* in *A. cerana* (86-93 bp) was found for the first time in the *A. mellifera* mtDNA (Crozier & Crozier, 1993; Cornuet et al., 1991)

Table 4.

A ratio of the transitions/transversions between mtDNA of Korean and other *A. cerana* samples. Parentheses contain numbers of amino acid replacements

Gene / Region	China, Yunnan, GQ162109, 15895 bp	China, Yunnan, KM244704, 15712 bp	China, Jiangsu, AP017983, 15460 bp	Japan, Tsushima, AP017985, 15339 bp	Japan, Kyoto, AP017314, 15917 bp	Japan, Amami, AP017941, 15788 bp	Taiwan, Taipei, AP017984, 15376 bp
tRNA-Ser$^{(TCT}$	-	-	-	-	-	-	1/0
tRNA-Ala	-	0/1	-	-	-	-	-
tRNA-Glu	-	-	-	-	-	-	1/0
tRNA-Ile	-	-	-	-	-	-	1/1
ND2	7(3)/1(1)	9(2)/1(1)	9(2)/1(1)	6(2)/0	5(1)/0	4(1)/0	15(1)/5
tRNA-Cys	-	-	-	-	-	-	1/0
COX1	9/1	7/1	7/1	3/0	4/0	5/0	30(1)/6
tRNA-Leu$^{(TAA)}$	-	-	-	-	-	-	0/1
COX2	5(1)/0	4(1)/0	5(1)/0	2/0	-	-	12(1)/3
tRNA-Asp	-	-	-	-	-	-	0/1
tRNA-Lys	-	-	-	-	-	-	1/0
ATP8	1/0	1/0	1/0	-	-	-	3(2)/0
ATP6	7(2)/0	5(2)/0	5(2)/0	1(1)/0	1(1)/0	1(1)/0	14(3)/5(1)
COX3	5(2)/1(1)	5(2)/1(1)	5(2)/1(1)	2(2)/3(2)	2(2)/2(1)	2(2)/2(1)	21(4)/8(3)
ND3	4(1)/0	2(1)/0	3(1)/0	1(1)/0	-	-	8/1
tRNA-Arg	2/0	2/0	2/0	2/0	2/0	2/0	1/1
tRNA-Asn	-	-	-	-	-	-	0/3
ND5	4(1)/1(1)	4/1(1)	7(1)/3(2)	7(1)/0	5/1(1)	6/1(1)	23(4)/9(2)
ND4	6(1)/8(3)	4(1)/8(3)	4(1)/8(3)	6(1)/11(5)	6(1)/10(5)	7(2)/10(5)	23(2)/17(7)
ND4L	-	-	-	1/0	1/0	1/0	5/0
ND6	4(3)/0	3(2)/0	3(2)/0	2(2)/0	1(1)/0	2(1)/0	14(1)/2(2)
CYTB	4/2	4/2	6/3(1)	3/0	3/0	2/0	29(3)/2
ND1	2/1(1)	2/1(1)	3/1(1)	1/0	2/0	2/0	14(3)/1(1)
tRNA-Leu$^{(TAG)}$	-	0/1	0/1	0/1	0/1	0/1	-
16S rRNA	1/1	1/0	3/1	1/1	2/0	2/0	9/9
12S rRNA	3/0	1/0	5/0	-	1/0	1/0	9/0
AT-rich	8/4	9/4	8/3	8/4	7/0	8/2	9/4
Intergenic	5/18	6/15	5/15	7/16	6/16	5/18	14/24
Total	77(41)/ 12(6)	69(37)/ 12(5)	81(37)/ 15(7)	53(32)/ 14(7)	48(24)/ 13(6)	50(29)/ 13(6)	258(70)/ 58(16)

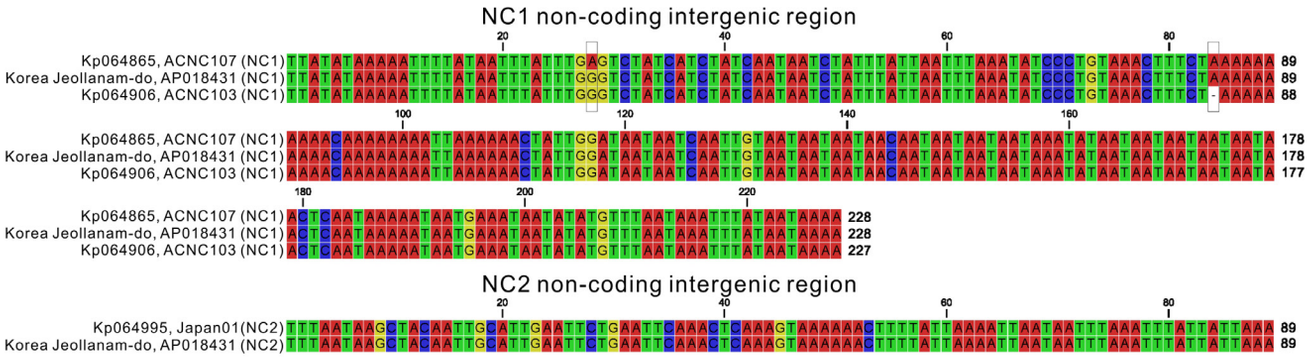

Fig. 4. Alignment of non-coding intergenic regions *NC1* and *NC2* of Korean *A. cerana* (AP018431) with haplotype sequences *ACNC107* (KP064865), *ACNC103* (KP064906) of *NC1* and Japan01 (KP064995) of *NC2*.

and was the longest intergenic spacer observed in *Apis* species. Later three additional unique non-coding regions: *NC1* located between *tRNA-Met* and *tRNA-Gln* (228-231 bp), *NC3* located between *COX3* and *tRNA-Gly* (66 bp), and *NC4* located between *tRNA-Pro* and *ND6* (50 bp) were found in *A. cerana* mtDNA (Lee et al., 2016).

A comparative analysis of *NC1* and *NC2* non-coding intergenic region sequences of Korean and GenBank *A. cerana* found similarity in *NC1* with sequences KP064865 (haplotype *ACNC107*) and KP064906 (haplotype *ACNC103*) and in *NC2* with sequence KP064995 (haplotype Japan01) (Fig. 4).

A pairwise comparison of mtDNA of all *A. cerana* and *A. florea* (out-group) was done. The percent of nucleotide differences and Genetic distances (Jukes-Cantor / Tamura-Nei / p-distance) were counted. Genetic distances between *A. cerana* samples varied from 0.001 / 0.002 / 0.002 to 0.020 / 0.022 / 0.021 (0.90 - 6.21% differences), between *A. cerana* and *A. florea* varied from 0.170 / 0.173 / 0.160 to 0.180 / 0.174 / 0.161 (21.23 - 28.57% differences). These three genetic distances practically do not differ between themselves (Tab. 5).

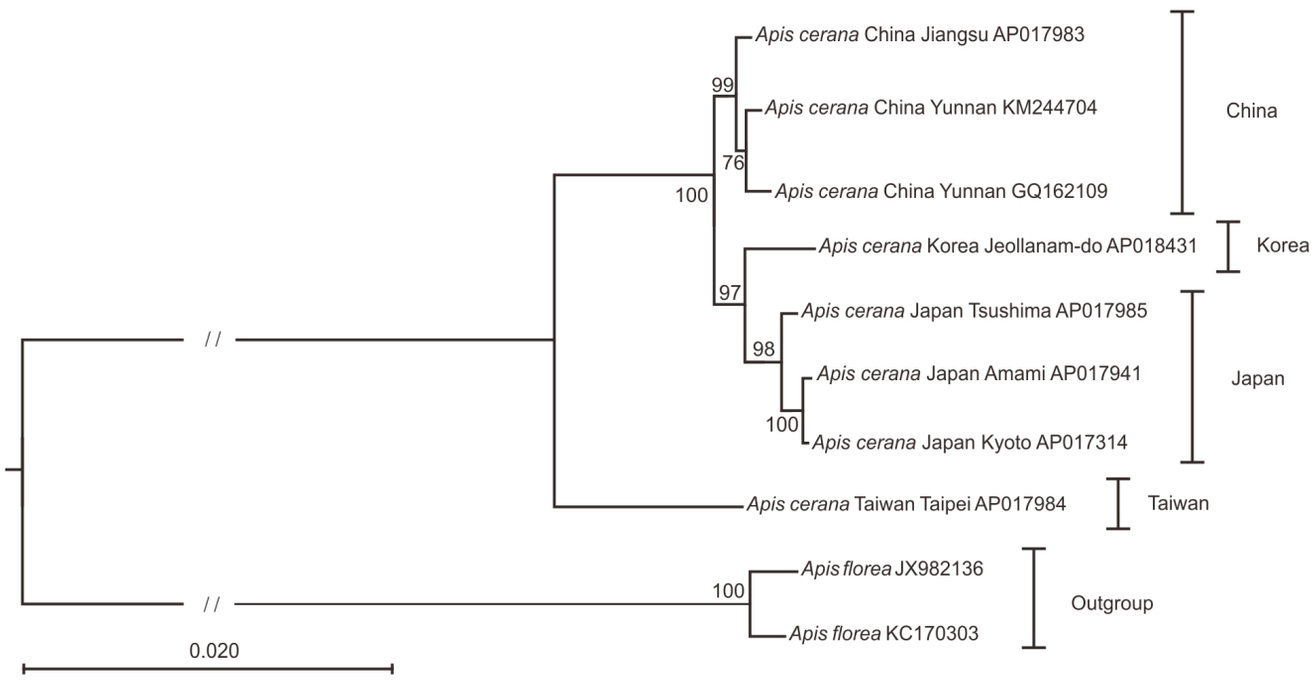

Fig. 5. Phylogenetic relationships of *A. cerana* samples visualized by the Neighbor-Joining phylogenetic tree based on the Jukes-Cantor genetic distances between complete mtDNA sequences.

Table 5.

Pairwise comparison complete mtDNA sequences of all *A. cerana* and *A. florea* samples. Upper triangle is Genetic distances (Jukes-Cantor / Tamura-Nei / *p*-distance), lower is percent of differences.

Samples		China, Yunnan, GQ162109	China, Yunnan, KM244704	China, Jiangsu, AP017983	Japan, Kyoto, AP017314	Japan, Amami, AP017941	Japan, Tsushima, AP017985	Taiwan, Taipei, AP017984	Korea, Jeollanam-do, AP018431	*Apis florea* JX982136	*Apis florea* KC170303
		Genetic distances Jukes-Cantor / Tamura-Nei / *p*-distance									
China, Yunnan, GQ162109	Genetic divergence Jukes-Cantor, %		0.001/ 0.002/ 0.002	0.001/ 0.002/ 0.002	0.010/ 0.005/ 0.005	0.010/ 0.005/ 0.005	0.010/ 0.006/ 0.006	0.020/ 0.022/ 0.021	0.010/ 0.007/ 0.007	0.180/ 0.174/ 0.161	0.170/ 0.173/ 0.160
China, Yunnan, KM244704		1.47		0.001/ 0.002/ 0.002	0.001/ 0.004/ 0.004	0.001/ 0.005/ 0.004	0.010/ 0.005/ 0.005	0.020/ 0.022/ 0.021	0.010/ 0.006/ 0.006	0.170/ 0.173/ 0.160	0.170/ 0.172/ 0.160
China, Jiangsu, AP017983		3.06	1.82		0.010/ 0.005/ 0.005	0.010/ 0.005/ 0.005	0.010/ 0.006/ 0.006	0.020/ 0.022/ 0.021	0.010/ 0.007/ 0.007	0.170/ 0.173/ 0.160	0.170/ 0.172/ 0.160
Japan, Kyoto, AP017314		1.78	2.8	3.45		0.001/ 0.001/ 0.001	0.001/ 0.001/ 0.001	0.020/ 0.022/ 0.021	0.010/ 0.005/ 0.005	0.180/ 0.174/ 0.161	0.170/ 0.173/ 0.160
Japan, Amami, AP017941		1.48	2.2	2.86	0.90		0.001/ 0.002/ 0.002	0.020/ 0.022/ 0.021	0.010/ 0.005/ 0.005	0.170/ 0.174/ 0.161	0.170/ 0.173/ 0.160
Japan, Tsushima, AP017985		4.22	3.01	1.53	3.81	3.03		0.020/ 0.022/ 0.021	0.010/ 0.005/ 0.005	0.170/ 0.174/ 0.161	0.170/ 0.173/ 0.160
Taiwan, Taipei, AP017984		5.67	4.53	3.01	5.70	5.12	3.84		0.020/ 0.023/ 0.022	0.170/ 0.175/ 0.162	0.170/ 0.175/ 0.161
Korea, Jeollanam-do, AP018431		1.35	2.38	4.02	1.15	1.88	4.64	6.21		0.180/ 0.174/ 0.161	0.170/ 0.173/ 0.160
Apis florea JX982136		26.97	27.69	28.48	26.92	27.42	29.02	28.57	27.15		0.010/ 0.005/ 0.005
Apis florea KC170303		21.84	20.98	21.24	21.78	21.36	21.79	21.23	22.04	10.13	

Based on the comparative analysis of complete mtDNA sequences of *A. cerana* and *A. florea* samples, the phylogenetic tree has been constructed (Fig. 5). All *A. cerana* samples were subdivided into four separate groups according to their geographical habitat: China, Taiwan, Korea, and Japan. This divergence of mtDNA could be a result of their long-term geographical isolation.

The magnitude of genetic divergences of *A. cerana* samples can be visualized by the median network. The median network was constructed on the basis of transversions in all protein-coding genes of mtDNA in comparison with reference to the *A. cerana* sample (Fig. 6). All transversions on the median network were placed on lines connecting samples. The length of the lines on the median network depends on the number of transversions. Amino acid replacements are presented on the lines of the median network also (Fig. 6).

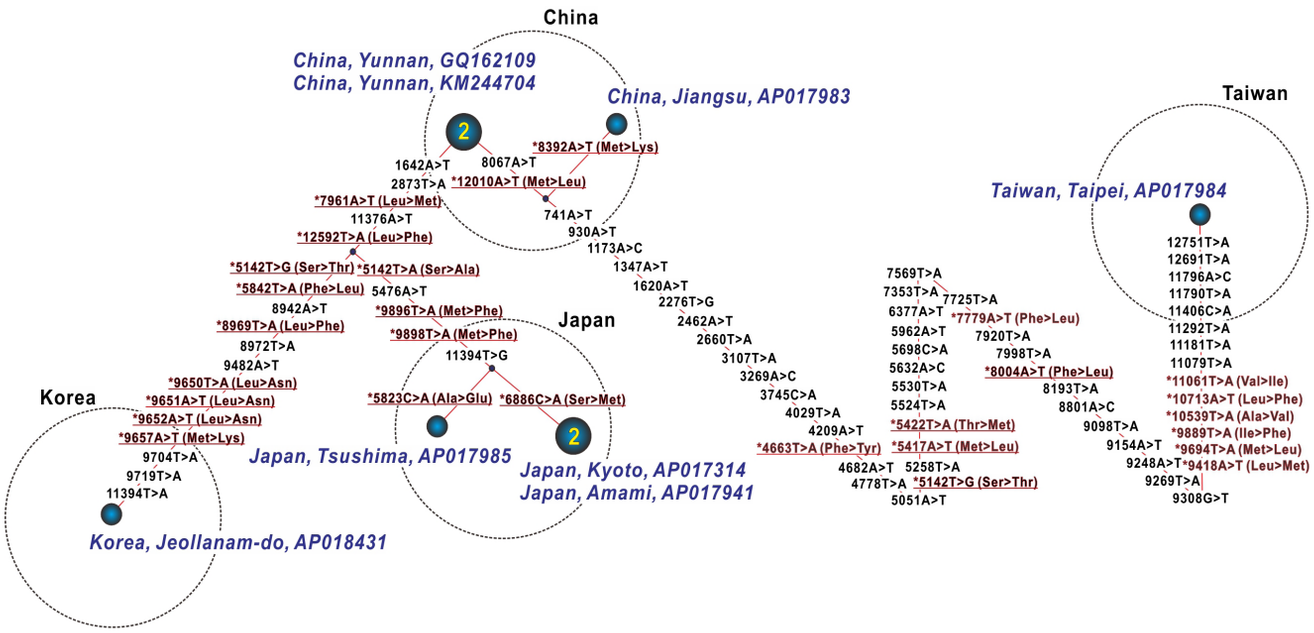

Fig. 6. The genetic divergence of *A. cerana* samples on the median network, constructed on the basis of transversions in all protein-coding genes of mtDNA between the reference and all other *A. cerana* samples.

DISCUSSION

Most DNA sequences can be evaluated using comparative analysis between species, subspecies, ecotypes, and populations. Comparative analysis of mtDNA allows differences to found among all compared samples and is useful for taxonomic, population genetic, phylogenetic and phylogeographic studies. The complexity, GC/AT-content, dinucleotides, repeats, structure and synteny of genes are usually analyzed for each sequenced genomes (Fig. 2).

For the most part, GC-content characterizes the stability of the DNA sequence. The G≡C pair is bound by three hydrogen bonds, while A=T pairs are bound by two hydrogen bonds. Thus, the DNA with lower GC-content is less stable than DNA with higher GC-content. However, the hydrogen bonds themselves do not have a particularly significant impact on stabilization, and the stabilization is due mainly to interactions of base stacking. GC-content in genetics is the percentage of guanine or cytosine nitrogenous bases on a DNA molecule. We have found that GC-content is variable over the entire length of mtDNA of *A. cerana* (Fig. 2). Probably, GC-content variation is the result of a mosaic-like distribution of GC-rich regions. The complexity of *A. cerana* mtDNA is directly proportional to the GC-content and varied over the entire length of mtDNA. GC-content of protein-coding genes in *A. cerana* mtDNA is directly proportional to their length. Thus, long genes have more complexity than short genes. The complexity of DNA can be represented as a heterogeneity, which is characterized by the domains within a domain. The spatial heterogeneity of the nucleotide composition and the long-range correlation of the DNA sequence can increase the complexity of the mtDNA (Fig. 2).

The AT-content in Korean *A. cerana* at 84.1% is similar to AT-content in other insect species: *Abispa ephippium* - 80.6%, *A. c. cerana* – 84%, *A. mellifera* - 84.9%, *Bombus hypocrita* - 85.3%, *B. ignitus* - 86.8%, *Cephus cinctus* – 82%, *Cotesia vestalis* - 87.2%, *Diadegma semiclausum*, - 87.4%, *Enicospilus* sp - 85.2%, *Evania appendigaster* - 77.8%, *Melipona bicolor* - 86.7%, *Orussus occidentalis* - 76.2%, *Philaenus spumarius* – 77%, *Polistes humilis* - 84.7%, *Spathius agrili* - 84, *Vanhornia eucnemidarum* - 80.1% (Tan et al., 2011). Martin (1995) and Belle et al. (2005) showed that high AT-content of animal mtDNA is correlated to metabolic rate. The metabolic rate could affect AT-content through the mutagenic

effects of oxygen free radicals, which are released in the mitochondria during aerobic respiration. Oxygen free radicals have mutagenic effects and lead to the replacement of GC with AT (Lindahl, 1993; Belle et al., 2005). Additionally, high AT-content can be due to frequent replacements of 5-methylcytosine and 6-methylguanine in CG pairs by AT pairs in mtDNA (Clary & Wolstenholme, 1985; Crozier & Crozier, 1993). Dinucleotides are involved in the formation complexity of mtDNA as well. Our calculated frequencies of each nucleotide, dinucleotides and AT/GC-content over the entire length of the mtDNA of Korean *A. cerana* makes it possible to explain DNA complexity. Some dinucleotides can increase the complexity of mtDNA. We found that AT-content and dinucleotides AA, AT, TA, TT are represented in highest frequency in Korean *A. cerana* mtDNA, due to which the complexity of mtDNA decreases to the average level 35% (Tab. 1).

The visualization of complete mtDNA plays an important role in the characterization of the genome because the circular structure allows the quick finding of genes, their flanking regions, and physical co-localization (synteny). The circular map of complete mtDNA can help to compare complete genomes. Thus, the comparative analysis of *A. cerana* and *A. mellifera* (NC_001566) circular structures of mtDNA allow us to find some differences in gene arrangement and direction. The mtDNA map is able to facilitate the comparison of mtDNA of different species, for example between Korean *A. cerana*, *A. mellifera* (NC_001566) (Crozier & Crozier, 1993) and *Drosophila yakuba* (NC_001322) (Clary & Wolstenholme, 1985). The protein-coding genes, the rRNA genes, and AT-rich control region of *A. cerana* are all at the same positions relative to the *D. yakuba* and *A. mellifera* mtDNA maps. However, fourteen tRNA genes of *A. cerana* are in altered positions relative to the *D. yakuba*, two tRNA genes (*tRNA-Glu* and *tRNA-Ser*$^{(TCT)}$) in altered positions relative to the *A. mellifera* and one tRNA gene (*tRNA-Arg*) have a reverse direction. Probably these differences in the synteny of tRNA genes are the result of translocational mutation between these species.

Some overlapping mtDNA genes have incomplete termination codons. Incomplete termination codons are known to be a general feature of animal mtDNAs, which can be completed by polyadenylation mechanism in their transcripts. Only Cnidaria species do not contain overlapping mtDNA genes due to their early divergence before the origin of the cleavage-polyadenylation mechanism. Cnidaria species have an alternative mechanism for the synthesis of mitochondrial mRNAs (Wolstenholme, 1992). The Korean *A. cerana* mtDNA has guanine as the rarest nucleotide (6.2%) and more similar with mtDNA of Echinoderm species than with mtDNA of *D. yakuba* with the rarest nucleotide cytosine (Wolstenholme, 1992).

The transition/transversion ratio (ts/tv) in the mtDNA of Korean *A. cerana* is relative to the reference sequence (GQ162109) is 2. It is known that ts/tv is different in the various species. For example, the average ts/tv in *Mus musculus* is 2 (Lindblad-Toh et al., 2000), in *Pan troglodytes* is 11, in *Gorilla gorilla* 6, in *Homo sapiens* 14 (Belle et al., 2005), in *D. melanogaster*, *D. simuluns*, *D. pseudoobscura* 2 (Moriyama & Powell, 1996), in *D. silvestris* 16 (DeSalle et al., 1987), in *D. yakuba* 5 (Clary & Wolstenholme, 1985), in *D. melanogaster* 2 (Seplyarskiy et al., 2012) and in *A. mellifera* 2 (Crozier & Crozier, 1993). The ratio of the transitions to transversions is higher mainly in mammals than in insects with the exception of *M. musculus* and *D. silvestris*. Thus, ts/tv of Korean *A. cerana* mtDNA is similar with those of four insects species *D. melanogaster*, *D. simuluns*, *D. pseudoobscura, A. mellifera* and one mammal species *M. musculus* and less in comparison with two insects species *D. silvestris, D. yakuba,* and three mammal primate species. Such a high level of ts/tv in some insect species can be explained by a high mutation rate in their genome due to the action of viruses or bacterias like *Wolbachia*. An analysis of intergenic sequences can be very informative in phylogenetic studies since their exclusion from natural selection. A comparative analysis of *NC1* and *NC2* non-coding intergenic region sequences of Korean *A. cerana* (AP018431) with GenBank data found similarity

with haplotypes *ACNC107*, *ACNC103* of *NC1* region and Japan01 of *NC2* region (Fig. 4). The sequence of the non-coding intergenic region *NC1* of Korean *A. cerana* (AP018431) differ from haplotype *ACNC107* by one transition 28A>G, relative to the sequence KP064865, and differ from haplotype *ACNC103* by one deletion 84delA, relative to the sequence KP064865. This new *NC1* haplotype of Korean *A. cerana* (AP018431) had not been known before and is named *ACNC110* (Fig. 4).

The previous investigation of *NC2* region in the Korean *A. cerana* population found four Korean haplotypes (Korea11, Korea12, Korea13, Korea14) and one common Asian mainland haplotype (Japan1) (Lee et al., 2015). The Japan1 haplotype was found in *A. cerana* populations of Japan, China, Vietnam, and Thailand. Korean *A. cerana* (AP018431) along belongs to the common haplotype Japan1. Such a wide distribution of Japan1 haplotype could be explained as the result of the gene flow and the consequence of a common genetic origin of most *A. cerana* populations in mainland Asia. Thus, the *NC2* region could not be used successfully in population genetic studies of *A. cerana.*

A previous comparative analysis of the *COX1* gene of *A. cerana* in Asia found two groups of haplotypes. The first group of *A. cerana* haplotypes was lineage A (Mainland group) with three sublineages - Taiwan, CH, HN, and JKR. The second group of *A. cerana* haplotypes was lineage B (Sundaland group) with two sublineages - Indonesia and IM (Smith, Warrit, & Hepburn, 2004; Zhao et al., 2013). The *COX1* gene of the Korean *A. cerana* (AP018431) sample was identical to the Korean (AF153108) and Japanese (AF153109) *A. cerana* samples belonging to the JKR sublineage of lineage A (Tanaka et al., 2001; Zhao et al., 2013). The joint grouping of Korean and Japanese *A. cerana* can be explained by the common historical origin of all *A. cerana* populations and high conservativeness of the *COX1* gene among subspecies. The *COX1* gene is known to be more useful for interspecific than intraspecific studies.

Studies have shown that mtDNA evolves ten times faster than nuclear DNA, and changes mainly occur in the intergenic control regions (Wilson, 1976; DeSalle et al., 1987; Arias & Sheppard, 1996; Tanaka et al., 2001). The speciation process cannot always be accompanied by changes in coding genes. In most cases, a significant divergence can be found in intergenic control regions (Wilson, 1976), which can be detected by differences in gene expression (Kartavtsev, 2009). For example, some fish species have a low divergence in enzyme sequences (Nei's genetic distance 0.01 - 0.08) (Nei, 1987) and a high difference in enzyme expression (differences 27 - 32%), which can be explained by a divergence in gene control regions (Kartavtsev, 2009).

The speciation process is always accompanied by increased genetic distances and genetic divergence between populations. Closely related taxa are characterized by a similarity of genetic distances and genetic divergence on each level of speciation. Vavilov (1951) stated in his law of homologous series that a particular variation observed in one species is also expected to be available in its related species. From this, related species should have a similar level of genetic variability, which can be observed in different insect species (Tab. 6).

Based on the data from the table, we have counted ranges of inter- and intraspecific genetic distance and genetic divergence in different insect taxa (Tab. 6). We assumed that the level of genetic divergence between individuals within subspecies of insects varied from 0.00% to 0.80%, and the genetic distance from 0.000 to 0.005. The level of genetic divergence between subspecies within insect species varied from 0.80% to 8.00%, and the genetic distance from 0.005 to 0.100. The level of genetic divergence between species within genera of insects varied from 8.00% to 17.00%, and the genetic distance from 0.100 to 0.200 (Tab. 6).

Korean *A. cerana* (AP018431) genetically differed less from Chinese Yunnan (GQ162109) (1.35%), Japanese Kyoto (AP017314) (1.15%), and Japanese Amami (AP017941) (1.88%) samples, and differed more from the Chinese Jiangsu sample (AP017983) (4.02%), Japanese Tsushima sample (AP017985) (4.64%) and Taiwanese

Table 6.

Inter- and intraspecific genetic distance and genetic divergence in different insect taxons

Subspecies and species	Genes / regions	Counting method	Value	Reference
Intraspecific genetic divergence (between subspecies)				
between subspecies and ecotypes of 84 insect species of subfamily Denticollinae	*COX1*	Kimura's 2-parameter	0.78%	Han et al., 2016
between subspecies and ecotypes of 3 insect species (*Toxoptera aurantii, T. citricidus, T. victoriae*)	*COX1*	Kimura's 2-parameter	1.30%	Wang & Qiao, 2009
between subspecies and ecotypes of bee species *A. mellifera* (*A. m. carnica, A. m. ligustica, A. m. sicula, A. m. iberica, A. m. adami, A. m. macedonica, A. m. anatolica, A. m. syriaca, A. m. intermissa*)	*COX1*	Jukes-Cantor	1.80%	Ashokan, 2011
between subspecies of 9 wasp species (*Eupristina koningsbergeri, Walkerella benjamini, Sycoscapter sp., Philotrypesis sp., Sycobia sp., Acophila sp., Sycophila sp., Ormyrus sp., Dibrachys sp.*)	*COX1, CYTB*	Tamura-Nei	2.00%	Xiao et al., 2010
between subspecies and ecotypes of 4 bee species of genus *Apis* (*Apis nuluensis, A. cerana, A. koschevnikovi, A. dorsata*)	*16SRNA, COX1, COX2*	Tamura-Nei	2.40%	Tanaka et al., 2001
between subspecies and ecotypes of bee species *A. cerana* (populations of China, Hainan, Japan, Korea, Russia, Taiwan)	*COX1*	Dest, Jost, 2008	3.00%	Zhao et al., 2013
Average			1.88% (0.78%-3.00%)	
Intraspecific genetic distance (between subspecies)				
between subspecies and ecotypes of bee species *A. cerana* (populations of different regions of Asia)	complete mtDNA	Tamura-Nei	0.005	Takahashi et al., 2016
between subspecies and ecotypes of bee species *A. cerana* (*A. c. cerana, A. c. japonica*)	complete mtDNA	Tamura-Nei	0.006	Takahashi et al., 2016
between subspecies and ecotypes of bee species *A. mellifera* (*A. m. lamarckii, A. m. mellifera, A. m. intermissa*)	complete mtDNA	p-distance	0.012	Eimanifar et al., 2017
between subspecies and ecotypes of bee species *A. mellifera* (*A. m. syriaca, A. m. scutellata, A. m. ligustica, A. m. intermissa*)	complete mtDNA	Tamura-Nei	0.015	Takahashi et al., 2016
between subspecies and ecotypes of 9 insect species (*Athous pucticollis, Hemicrepidius oblongus, H. coreanus, H. seccessus, Ctenicerini sp., Poemnites katalin-bodorae, Actenicerus infirmus, A. alternatus, Selatosomus coreanus*)	*COX1*	Kimura's 2-parameter	0.030	Han et al., 2016
Average			0.013 (0.005-0.030)	

Interspecific genetic divergence (between species)				
between insect species (*Toxoptera aurantii, T. citricidus, T. victoriae*)	*COX1*	Kimura's 2-parameter	8.30%	Wang & Qiao, 2009
between 7 bee species of genus *Apis* (*Apis nuluensis, A. mellifera, A. cerana, A. koschevnikovi, A. dorsata, A. florea, A. andreniofmis*)	*16SRNA, COX1, COX2*	Tamura-Nei	9.00%	Tanaka et al., 2001
between 2 insect species (*Schizaphis rotundiventris, S. graminum*)	*COX1*	Tamura-Nei	10.00%	Shufran et al., 2000
between 84 insect species of subfamily Denticollinae	*COX1*	Kimura's 2-parameter	11.74%	Han et al., 2016
between 2 honeybee species (*A. mellifera, A. cerana*)	complete mtDNA	Jukes-Cantor	17.00%	Tan et al., 2011
Average			11.21% (8.30%-17.00%)	
Interspecific genetic distance (between species)				
between 2 bumblebee species (*Bombus ignitus, B. hypocrita*)	complete mtDNA	Tamura-Nei	0.100	Takahashi et al., 2016
between 2 insect species (*Schizaphis rotundiventris, S. graminum*)	*COX1*	Tamura-Nei	0.160	Shufran et al., 2000
between 6 bee species of genus *Apis* (*A. florea, A. andreniofmis, A. cerana, A. mellifera, A. florea, A. dorsata*)	complete mtDNA	Tamura-Nei	0.200	Takahashi et al., 2016
Average			0.153 (0.100-0.200)	

Taipei sample (AP017984) (6.21%) (Tab. 5). The genetic divergence (Jukes-Cantor genetic distances) between Korean and Japanese *A. cerana* samples averages 2.57% (0.005), between Korean and Chinese *A. cerana* samples 2.58% (0.007), between Korean and Taiwanese *A. cerana* samples 6.21 (0.020), between Taiwanese and Japanese *A. cerana* samples 4.89 (0.022), between Taiwanese and Chinese *A. cerana* samples 4.40% (0.022) and between Chinese and Japanese *A. cerana* samples 2.59% (0.005).

The levels of genetic divergence and distance between Korean, Japanese, Chinese and Taiwanese *A. cerana* samples are matched to the range of the intraspecific level of differences in other insects - 0.80% - 8.00% (0.005 - 0.100). Thus, based on the differences in the complete mtDNA analysis we suggest that Korean, Japanese, Chinese and Taiwanese honeybee samples can be representatives of distinct subspecies of *A. cerana.*

Similar differentiation of Korean, Chinese and Japanese *A. cerana* populations can be obtained through morphology studies. Based on the factor and cluster analysis in Statistica 8.0 (StatSoft Power Solutions, Inc., USA) of morphological data (forewing's length and width, cubital index, hindleg length, metatarsal index, and the length of tergite 3+4th) (Ruttner, 1988), the Korean (N=40), Chinese (N=124) and Japanese (N=8) *A. cerana* samples (Ken et al., 2003; Lee & Choi, 1986) are observed to be grouped separately from each other on the tree and plot diagrams. The Japanese samples most differed from both Chinese and Korean samples. The similar grouping patterns of *A. cerana* on both tree and plot diagrams support morphological data (Fig. 7 A, B).

Analogous differentiation of Korean, Chinese and Japanese populations of *A. cerana* can were obtained through nuclear gene studies. We analysed *VG* (Vitellogenin) (~6,000 bp) and *EF1-*α (Elongation factor 1-alpha) (~2,000 bp) genes of nuclear DNA which had been previously used in the phylogenetic studies of honeybees (Arias, & Sheppard, 2005; Ramírez et al., 2010; Kent et al., 2011; Ilyasov, Poskryakov, & Nikolenko, 2015). On both tree diagrams (Jukes-Cantor distances, Neighbor-Joining method) constructed by CLC Genomics Workbench 11, the Korean, Japanese and Chinese *A. cerana* samples were grouped

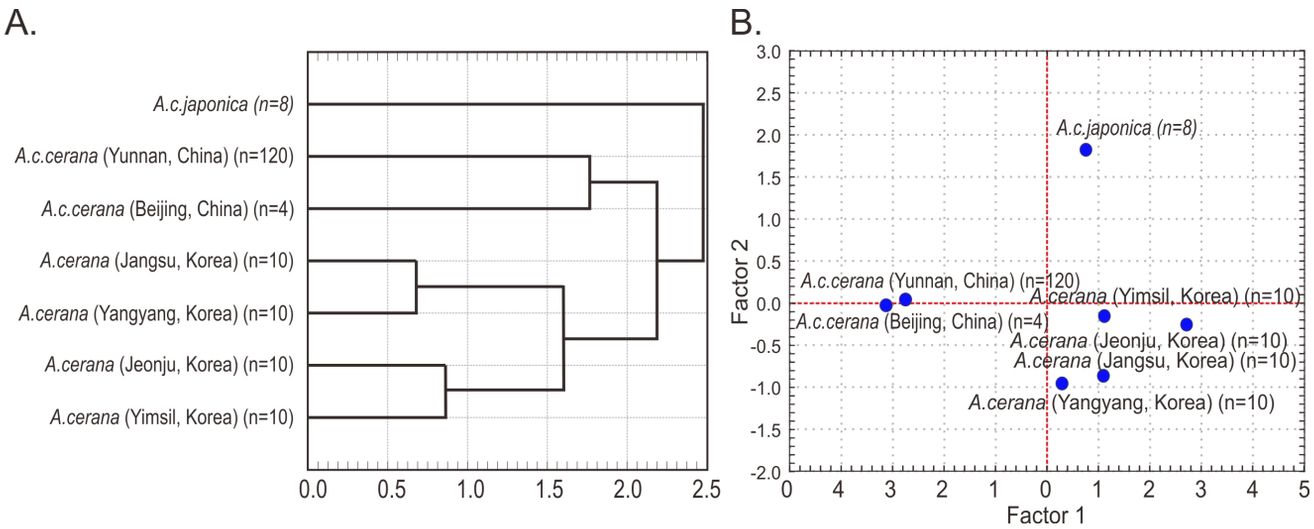

Fig. 7. Differentiation of *A. cerana* samples from China, Japan and Korea by cluster (A) and factor (B) analysis of morphology data. Numbers of analyzed honeybee colonies shown in parentheses.

separately from one another. The species of honeybees *A. mellifera* (Amel_4.5), *A. dorsata* (Apis dorsata 1.3) and *A. florea* (Aflo_1.0) were used as out-group samples (Fig. 8 A, B).

One sample of Chinese *A. c. cerana* (ApisCC1.0) have a similarity with Japanese *A. c. japonica* (Apiscer_1.0) but they still differ at the subspecies level. Chinese samples of *A. cerana* vary from one another. The territory of China is assumed to be possibly inhabited by different

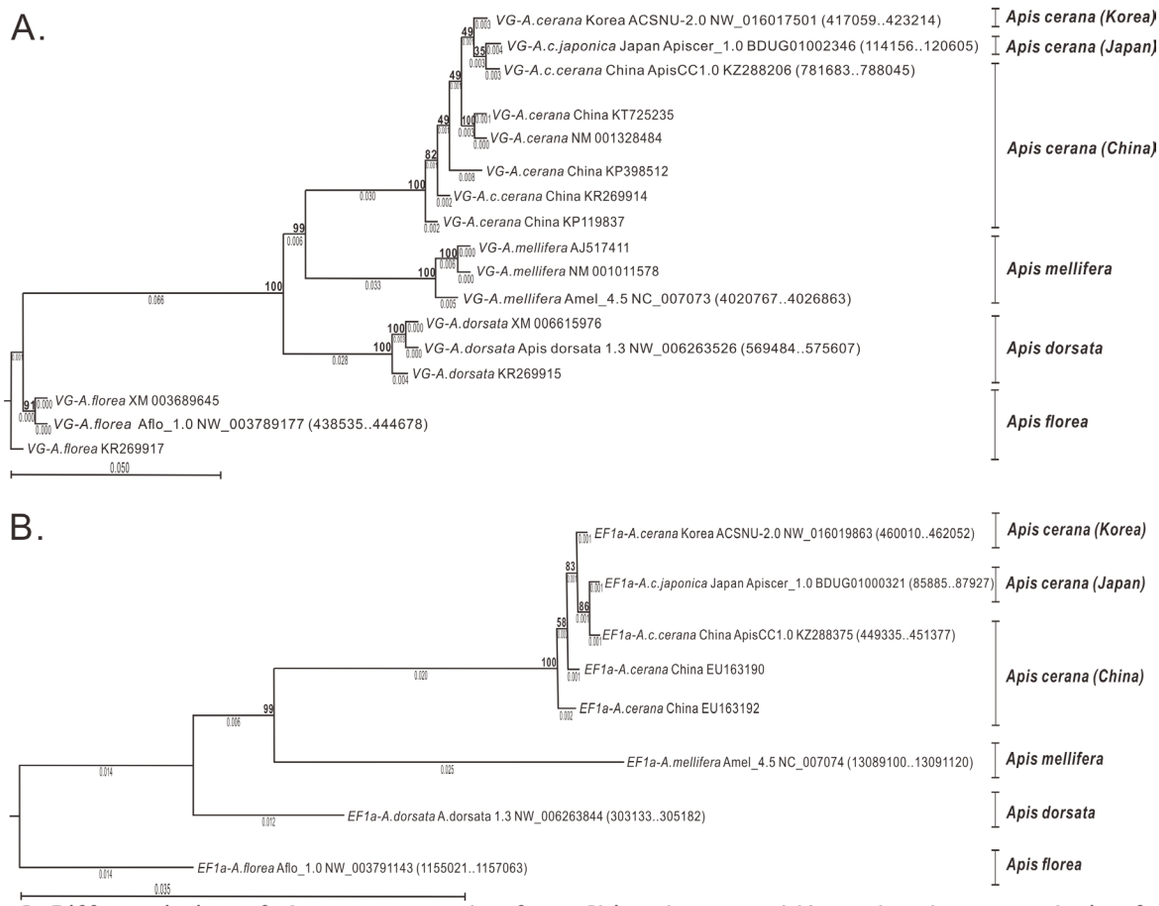

Figure 8. Differentiation of *A. cerana* samples from China, Japan and Korea by cluster analysis of *VG* (A) and EF1-α (B) gene nucleotide sequences. Positions of the genes on the scaffolds in honeybee genome assemblies shown in parentheses.

subspecies of *A. cerana* due natural selection of bees in different climatic zones (Tanaka et al., 2001; Ken et al., 2003; Smith, Warrit, & Hepburn, 2004; Tan, Warrit, & Smith, 2007; Zhao et al., 2013; Lee et al., 2016).

The previous statement that subspecies *A. c. japonica* originally came to Japan from the Korean peninsula (Sugawara, 2000; Takahashi & Yoshida, 2003) does not contradict this assumption that *A. cerana* from Korea and Japan are distinct subspecies. The common origin of different subspecies is possible and it is a law of evolutionary transformation. The process of speciation always begins from the division of a common population and further isolation, and the isolation time and distance between populations play a decisive role. The level of divergence between populations depends on the time of isolation because as it increases so does genetic divergence due to the accumulation of mutations in each population. Genetic divergence between populations comprises both neutral and non-neutral variations. Genetic variation between spatially separated populations is generated by the stochastic and non-stochastic accumulation of mutations and the hitchhiking of neighboring neutral variation at loci under selection in consequence of genetic drift and natural selection.

It is evaluated that the divergence in time between *A. cerana* populations of the mainland and Pacific Ocean islands is about 17,000 years (Voris, 2000; Zhao et al., 2013). There is the geographical isolation by sea and mountains and isolation by distance between Korean, Japanese and Chinese *A. cerana* populations. The distance between the Korean and Japanese populations of *A. cerana* is about 180-388 km. During this period their isolation could be sometimes partially disrupted by human activity and climatic changes, which could not totally destroy their aboriginal gene pool. Two possible ways are assumed for the historical migration of *A. cerana* into north-east Asia (Fig. 1). The first way was the predominantly natural migration of *A. cerana* in the period when the sea level fell below 100 meters about 0.02 and 0.25 million years ago due to glaciation (Voris, 2000; Zhao et al., 2013) (Fig. 1). The second way was a predominantly artificial migration of *A. cerana* in the period, when there was a mass migration of humans to north-eastern Asia across the sea about 0.02 and 0.03 million years ago (Jinam et al., 2012) (Fig. 1).

There are different hypotheses about the divergence time of *A. cerana* and *A. mellifera.* Morphological studies assumed that the divergence of *A. cerana* and *A. mellifera* occurred two million years ago (Ruttner & Maul, 1983), mtDNA studies four million years ago (Smith, 1990) and nuclear DNA studies six million years ago (Willis, Winston, & Honda, 1992; Wallberg et al., 2014). The divergence time of *A. mellifera* subspecies was assumed at 0.30 million years ago on the basis of 8.3 million SNPs (Wallberg et al., 2014).

It is assumed that the 2.3% difference in mtDNA corresponds to the divergence one million years ago (HsuChen, Kotin, & Dubin, 1984; DeSalle et al., 1987; Arias & Sheppard, 1996; Tanaka et al., 2001). According to this assumption, Korean, Japanese, and Chinese *A. cerana* diverged about one million years ago, and Taiwanese *A. cerana* about two million years ago. We assumed that *A. cerana* subspecies divergence occurred about 0.3 - 1 million years ago, similarly to *A. mellifera.* On the phylogenetic tree, Korean and Japanese *A. cerana* samples were grouped together in one common cluster (Fig. 5). However, within this cluster Korean and Japanese *A. cerana* subdivided separately. Chinese *A. cerana* samples were combined into the second distinct cluster. Korean *A. cerana* was located in the middle position between Japanese and Chinese *A. cerana* groups. The middle position of Korean *A. cerana* can be explained by the historical process of migration from mainland Asia to north-east Asia into the Japanese archipelago through Korean peninsula according to the first way of migration (Fig. 1) (Takahashi & Yoshida, 2003). Taiwanese *A. cerana* were located separately from all Japanese and Chinese *A. cerana* groups on the phylogenetic tree (Fig. 5).

On the median network graph, all *A. cerana* samples were subdivided into four groups. Taiwanese *A. cerana* sample most differs

from the reference sequence (GQ162109) and contains fifty-seven transversions in protein-coding genes, thirteen of which led to the amino acid replacements. The genetic remoteness of Taiwanese *A. cerana* from other mainland and island *A. cerana* populations has been observed by other researchers on the basis of *COX1* mtDNA gene polymorphism (Zhao et al., 2013). All Chinese *A. cerana* samples on the median network combined into one common group. Chinese *A. cerana* sample from Jiangsu are located separately from all the other Chinese *A. cerana* samples on both the phylogenetic tree and the median network and differs from them by three transversions in positions 8067A>T and 8392A>T (Met>Lys) in *ND5* gene and by 12010A>T (Met>Leu) in *CYTB* gene, relative to the reference sequence (Fig. 6). The Japanese *A. cerana* sample from Tsushima is located separately on both the phylogenetic tree and the median network and differs from other Japanese *A. cerana* samples by transversions 5823C>A (Ala>Glu) in the *COX3* gene and by 6886C>A (Ser>Met) in the *ND5* gene. This difference can be explained by a different time of isolation from one another. Korean and Japanese *A. cerana* samples combined into one common branch, but they are located separately within it. The split point of Korean and Japanese *A. cerana* is assumed to be a common ancestor, which differs from the reference sequence by five transversions. Korean and Japanese *A. cerana* samples differ from each other by twenty transversions in protein-coding genes, twelve of which led to the amino acid replacements (Fig. 6).

Thus, based on comparative analysis of complete mtDNA (~16,000 bp), two nuclear gene *VG* and *EF1-*α sequences (~8,000 bp) and morphological measurements (six parameters) we assume that Korean *A. cerana*, Chinese *A. c. cerana* and Japanese *A. c. japonica* are different subspecies at an early stage of sub-speciation. The representatives of the Korean population of *A. cerana* could be called further as a subspecies of *Apis cerana koreana*. Additional population genetic analysis must be provided to prove the gene pool homogeneity of *A. c. koreana* population.

ACKNOWLEDGMENTS

This work was carried out with the support of the Postdoctoral Fellowships in the Incheon National University (2017-2019), of the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (No. 2016R1A2B3011742), of the Cooperative Research Program for Agriculture Science and Technology Development (RDA) (Project Number: PJ012526). We are grateful to Dr. Hisashi Okuyama for kindly providing data for our analysis.

REFERENCES

Arias, M.C., & Sheppard, W.S. (1996). Molecular phylogenetics of honey bee subspecies (*Apis mellifera* L.) Inferred from Mitochondrial DNA Sequence. *Molecular Phylogenetics and Evolution, 5*(3), 557-566. DOI: 10.1006/mpev.1996.0050

Arias, M.C., & Sheppard, W.S. (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. DOI: 10.1016/j.ympev.2005.02.017

Ashokan, K.V. (2011). Molecular phylogenetic study on *Apis mellifera* subspecies inferred from Cytochrome Oxidase I sequence. *Indian Journal of Fundamental and Applied Life Sciences, 1*(4), 193-202.

Behura, S.K. (2007). Analysis of nuclear copies of mitochondrial sequences in honeybee (*Apis mellifera*) genome. *Molecular Biology and Evolution, 24*(7), 1492-1505. DOI: 10.1093/molbev/msm068

Belle, E.M., Piganeau, G., Gardner, M., & Eyre-Walker, A. (2005). An investigation of the variation in the transition bias among various animal mitochondrial DNA. *Gene, 355*, 58-66.

Bernt, M., Donath, A., Jühling, F., Externbrink, F., Florentz, C., Fritzsch, G., ... Stadler, P.F. (2013). MITOS: improved de novo Metazoan mitochondrial genome annotation. *Molecular Phylogenetics and Evolution, 69,* 313-319. DOI: 10.1016/j.ympev.2012.08.023

Buttel-Reepen, H. (1906). *Mitteilungen aus dem zoologischen museum in Berlin.* Berlin: Springer Netherlands.

Cameron, S.A. (1993). Multiple origins of advanced eusociality in bees inferred from mitochondrial DNA sequences. *Proceedings of the National Academy of Sciences, 90*, 8687–8691. DOI: 10.1073/pnas.90.18.8687

Choi, Y.S., Lee, M.Y., Hong, I.P., Kim, N.S., Kim, H.K., Lee, K.G., Lee, M L. (2010). Occurrence of sacbrood virus in Korean apiaries from *Apis cerana* (Hymenoptera: Apidae). *Journal of Apiculture, 25*(3), 187-191. DOI: 10.17519/apiculture.2017.11.32.4.345

Clary, D.O., & Wolstenholme, D.R. (1985). The mitochondrial DNA molecule of *Drosophila yakuba*: nucleotide sequence, gene organization, and genetic code. *Journal of Molecular Evolution, 22,* 252-271. DOI: 10.1007/BF02099755

Cornuet, J.M., Garnery, L., & Solignac, M. (1991). Putative origin and function of the intergenic region between COI and COII of *Apis mellifera* L. mitochondrial DNA. *Genetics, 128*, 393-403. DOI: 10.1051/apido:2007007

Crane, E. (1999). *The world history of beekeeping and honey hunting*. New York: Routledge.

Crozier, R.H., & Crozier, Y.C. (1993). The mitochondrial genome of the honeybee *Apis mellifera*: complete sequence and genome organization. *Genetics, 133*(1), 97-117. DOI: 10.1111/j.1365-2583.1993.tb00131.x

De Salle, R., Freedman, T., Prager, E.M., & Wilson, A.C. (1987). Tempo and mode of sequence evolution in mitochondrial DNA of Hawaiian Drosophila. *Journal of Molecular Evolution, 26*, 157-164. DOI: 10.1007/BF02111289

DeSalle, R., Freedman, T., Prager, E.M., & Wilson, A.C. (1987). Tempo and mode of sequence evolution in mitochondrial DNA of Hawaiian *Drosophila*. *Journal of Molecular Evolution, 26*(1–2), 157–164. DOI: 10.1007/BF02111289

Diniz-Filho, J.A.F., Malapsina, O., & Pignata, M.I.B. (1993). Geographic variation in *Apis cerana indica* F.: a spatial autocorrelation analysis of morphometric patterns. *Journal of Apicultural Research, 32,* 65-72. DOI: 10.1080/00218839.1993.11101289

Eimanifar, A., Kimball, R.T., Braun, E.L., Moustafa, B.M., Haddad, N., Fuchs, S., Grünewald, B., Ellis, J.D. (2017). The complete mitochondrial genome of the Egyptian honey bee, Apis mellifera lamarckii (Insecta: Hymenoptera: Apidae). *Mitochondrial DNA Part B, 2*(1), 270-272. DOI: 10.1080/23802359.2017.1325343

Engel, M.S. (1999). The taxonomy of recent and fossil honey bees (Hymenoptera Apidae Apis). *Journal of Hymenoptera Research, 8*(2), 165-196. DOI: 10.1007/978-1-4614-4960-7_18

Engel, M.S., & Schultz, T.R. (1997). Phylogeny and behavior in honey bees (Hymenoptera: Apidae). *Annals of the Entomological Society of America, 90*(1), 43-53. DOI: 10.1093/aesa/90.1.43

Garnery, L., Cornuet, J.M., & Solignac, M. (1992). Evolutionary history of the honey bee *Apis mellifera* inferred from mitochondrial DNA analysis. *Molecular Ecology, 1*(3), 145-154. DOI: 10.1111/j.1365-294X.1992.tb00170.x

Garnery, L., Mosshine, E.H., Oldroyd, B.P., & Cornuet, J.M. (1995). Mitochondrial DNA variation in Moroccan and Spanish honey bee populations. *Molecular Ecology, 4*(4), 465-472. DOI: 10.1111/j.1365-294X.1995.tb00240.x

Han, T., Lee, W., Lee, S., Park, I.G., Park, H. (2016). Reassessment of species diversity of the subfamily Denticollinae (Coleoptera: Elateridae) through DNA barcoding. *PLoS ONE, 11*(2), 1-28. DOI: 10.1371/journal.pone.0148602

Hepburn, H.R., & Radloff, S.E. (2011). *Honeybees of Asia*. Berlin, Heidelberg, Germany: Springer-Verlag. DOI: 10.1007/978-3-642-16422-4

Hepburn, H.R., Radloff, S.E., Verma, S., & Verma, L.R.

(2001). Morphometric analysis of *Apis cerana* populations in the southern Himalayan region. *Apidologie, 32*(5), 435-447. DOI: 10.1051/apido:2001142

HsuChen, C.C., Kotin, R.M., & Dubin, D.T. (1984). Sequences of the coding and flanking regions of the large ribosomal subunit RNA gene of mosquito mitochondria. Nucleic Acids Research, 12(20), 7771-7785. DOI: 10.1093/nar/12.20.7771

Ilyasov, R.A., Poskryakov, A.V., & Nikolenko, A.G. (2015). New SNP markers of the honeybee vitellogenin gene (Vg) used for identification of subspecies *Apis mellifera mellifera* L. *Russian Journal of Genetics*, *51*(2), 163-168. DOI: 10.1134/S1022795415020088

Jinam, T., Nishida, N., Hirai, M., Kawamura, S., Oota, H., Umetsu, K., Saitou, N. (2012). The history of human populations in the Japanese Archipelago inferred from genome-wide SNP data with a special reference to the Ainu and the Ryukyuan populations. *Journal of Human Genetics, 57*(12), 787-795. DOI: 10.1038/jhg.2012.114

Jukes, T.H., & Cantor, C.R. (1969) Evolution of protein molecules. In Munro, H. N., editor, Mammalian protein metabolism. (21-132 pp.). New York: Academic Press.

Kartavtsev, Y.F. (2009). Molecular evolution and population genetics. (280 pp.). Vladivostok: Far Eastern University Publishing.

Ken, T., Fuchs, S., Koeniger, N., & Ruiguang, Z. (2003). Morphological characterization of *Apis cerana* in the Yunnan Province of China. *Apidologie*, *34*(6), 553-561. DOI: 10.1051/apido:2003049

Kent, C.F., Issa, A., Bunting, A.C., & Zayed, A. (2011). Adaptive evolution of a key gene affecting queen and worker traits in the honey bee, *Apis mellifera*. *Molecular Ecology*, *20*(24), 5226-5235. DOI: 10.1111/j.1365-294X.2011.05299.x

Koetz, A.H. (2013). Ecology, behaviour and control of *Apis cerana* with a focus on relevance to the Australian incursion. *Insects, 4*, 558-592. DOI: 10.3390/insects4040558

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution, 33*, 1870-1874. DOI: 10.1093/molbev/msw054

LaSalle, J., & Gauld, I.D. (1993). Hymenoptera: their diversity and their impact on the diversity of other organisms. In *Hymenoptera and biodiversity*. Wallingford, United Kingdom: CAB International.

Lee, J.Y., Wang, A.R., Choi, Y.S., Thapa, R., Kwon, H.W., Kim, I. (2016). Mitochondrial DNA variations in Korean Apis cerana (Hymenoptera: Apidae) and development of another potential marker. *Apidologie*, *47*(1), 123-134. DOI: 10.1007/s13592-015-0381-y.

Lee, M.L., & Choi, S.Y. (1986). Biometrical studies on the variation of some morphological characters in Korean Honeybees, *Apis cerana* F. and *A. mellifera* L. *Korean Journal of Apiculture*, *1*(1), 5-23.

Lindahl, T. (1993). Instability and decay of the primary structure of DNA. *Nature, 362*, 709-715. DOI: 10.1038/362709a0

Lindauer, M. (1956). Uber die Verstandigung bei indischen Bienen. *Zeitschrift für vergleichende Physiologie, 38,* 521-557.

Lindblad-Toh, K., Winchester, E., Daly, M.J., Wang, D.G., Hirschhorn, J.N. (2000). Large-scale discovery and genotyping of single nucleotide polymorphisms in the mouse. *Nature Genetics, 24,* 381-386. DOI: 10.1038/74215

Martin, A., (1995). Metabolic rate and directional substitution in animal mitochondrial DNA. *Molecular Biology and Evolution, 12,* 1124-1131. DOI: 10.1093/oxfordjournals.molbev.a040286

Moriyama, E.N., & Powell, J.R. (1996). Intraspecific nuclear DNA variation in *Drosophila*. *Molecular Biology and Evolution, 13,* 261-277. DOI: 10.1093/oxfordjournals.molbev.a025563

Nei, M. (1987). Molecular evolutionary genetics. (512 pp.). New York, USA: Columbia University Press.

Nei, M., & Kumar, S. (2000). Molecular evolution and phylogenetics. New York: Oxford University Press. DOI: 10.1046/j.1365-2540.2001.0923a.x

Okuyama, H., Wakamiya, T., Fujiwara, A., Washitani, I., Takahashi, J. (2017). Complete mitochondrial genome of the honeybee *Apis cerana* native to two remote islands in Japan. *Conservation Genetics Resources, 9*(4), 557-560. DOI: 10.1007/s12686-017-0721-5

Pesenko, Y.A., Lelei, A.S., Radchenko, V.G., Filatkin, G.N. (1989). Chinese wax-bee, *Apis cerana cerana* F. (Hymenoptera, Apidae) in the Far East of the USSR (in Russian). *Entomologicheskoe Obozrenie, 68*, 527-548.

Radloff, S.E., Hepburn, C., Hepburn, H.R., Fuchs, S., Hadisoesilo, S., Tan, K., Engel, M.S., Kuznetsov, V. (2010). Population structure and classification of *Apis cerana*. *Apidologie, 41*, 589-601. DOI: 10.1051/apido/2010008

Raffiudin, R., & Crozier, R.H. (2007). Phylogenetic analysis of honey bee behavioral evolution. *Molecular Phylogenetics and Evolution, 43*(2), 543–552. DOI: 10.1016/j.ympev.2006.10.013

Ramírez, S.R., Roubik, D.W., Skov, C.E., & Pierce, N.E. (2010). Phylogeny, diversification patterns and historical biogeography of euglossine orchid bees (Hymenoptera: Apidae). *Biological Journal of the Linnean Society, 100*(3), 552-572. DOI: 10.1111/j.1095-8312.2010.01440.x

Rinderer, T.E. (1986). *Bee genetics and breeding*. San Diego: Academic Press. DOI: 10.1016/B978-0-12-588920-9.50011-X

Ruttner, F. (1988). *Biogeography and taxonomy of honeybees*. Berlin, Heidelberg, Germany: Springer-Verlag. DOI: 10.1016/0169-5347(89)90176-6

Ruttner, F., & Maul, V. (1983). Experimental analysis of the reproductive interspecific isolation of *Apis mellifera* L. and *Apis cerana* Fabr. *Apidologie, 14*(4), 309-327. DOI: 10.1051/apido:19830405

Saitou, N., & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution, 4,* 406-425.

Seplyarskiy, V.B., Kharchenko, P., Kondrashov, A.S., & Bazykin, G.A. (2012). Heterogeneity of the transition/transversion ratio in *Drosophila* and *Hominidae* genomes. *Molecular Biology and Evolution, 29*(8), 1943-1955. DOI: 10.1093/molbev/mss071

Shufran, K.A., Burd, J.D., Anstead, J.A., & Lushai, G. (2000). Mitochondrial DNA sequence divergence among greenbug (Homoptera: Aphididae) biotypes: Evidence for host-adapted races. *Insect Molecular Biology, 9*(2), 179–184. DOI: 10.1046/j.1365-2583.2000.00177.x

Smith, D.R. (1990). Mitochondrial DNA and biogeography of *Apis* species. Proceedings, XI Congress, IUSSI, India, 113-114.

Smith, D.R., Warrit, N., & Hepburn, H.R. (2004). *Apis cerana* from Myanmar (Burma): unusual distribution of mitochondrial lineages. *Apidologie*, *35*, 637–644. DOI: 10.1051/apido:2004058

Songrarn, O., Sittipraneed, S., & Klinbunga, S. (2006). Mitochondrial DNA diversity and genetic differentiation of the honeybee (*Apis cerana*) in Thailand. *Biochemical Genetics, 44* (5-6), 256-269. DOI: 10.1007/s10528-006-9030-5

Sugawara, M. (2000). Feral colonies of Japanese honey bees, *Apis cerana japonica* and their life history. 2. Natural nests and swarming. *Mitsubachi Kagaku (Honeybee Science), 21*(1), 35-39.

Takahashi, J., Wakamiya, T., Kiyoshi, T., Uchiyama, H., Yajima, S., Kimura, K., Nomura, T. (2016). The complete mitochondrial genome of the Japanese honeybee, *Apis cerana japonica* (Insecta: Hymenoptera: Apidae). *Mitochondrial DNA Part B, 1*(1), 156-157. DOI: 10.1080/23802359.2016.1144108

Takahashi, J., & Yoshida, T. (2003). The origin of Japanese honey bee *Apis cerana japonica* inferred from mitochondrial DNA (in Japanese). *Mitsubachi Kagaku (Honeybee Science), 24*(2), 71-76.

Tamura, K, & Nei, M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution, 10*(3), 512-526. DOI: 10.1093/molbev/msl149

Tamura, K., Battistuzzi, F.U., Billing-Ross, P., Murillo, O., Filipski, A., Kumar, S. (2012). Estimating divergence times in large molecular phylogenies. *Proceedings of the National Academy of Sciences of the United States of America, 109*(47), 19333-19338. DOI: 10.1073/pnas.1213199109

Tan, H.W., Liu, G.H., Dong, X., Lin, R.Q., Song, H.Q., Huang, S.Y., Yuan, Z.G., Zhao, X.Q. (2011). The Complete Mitochondrial Genome of the Asiatic Cavity-Nesting Honeybee *Apis cerana* (Hymenoptera: Apidae). *PLoS ONE, 6*(8), e23008. DOI: 10.1371/journal.pone.0023008

Tan, K., Warrit, N., & Smith, D.R. (2007). Mitochondrial DNA diversity of Chineses *Apis cerana*. *Apidologie, 38*, 238-246. DOI: 10.1051/apido:2007008

Tanaka, H., Roubik, D.W., Kato, M., Liew, F., Gunsalam, G. (2001). Phylogenetic position of *Apis nuluensis* of northern Borneo and phylogeography of *A. cerana* as inferred from mitochondrial DNA sequences. *Insectes Sociaux, 48*(1), 44-51. DOI: 10.1007/PL00001744

Vavilov, N.I. (1951). The origin, variation, immunity and breeding of cultivated plants. (364 pp.). Waltham, Massachusetts: Chronica Botanica; New York: Stechert-Hafner.

Voris, H.K. (2000). Maps of Pleistocene sea levels in Southeast Asia: shorelines, river systems and time durations. *Journal of Biogeography, 27*, 1153-1167. DOI: 10.1046/j.1365-2699.2000.00489.x

Vung, N.N., Lee, M.L., Lee, M.Y., Kim, H.K., Kang, E.J., Kim, J.E., Choi, Y.S. (2017). Breeding and Selection for Resistance to Sacbrood Virus for *Apis cerana*. *Journal of Apiculture, 32*(4), 345-352. DOI: 10.17519/apiculture.2017.11.32.4.345

Wallberg, A., Han, F., Wellhagen, G., Dahle, B., Kawata, M., Haddad, N., ... Webster, M.T. (2014). A worldwide survey of genome sequence variation provides insight into the evolutionary history of the honeybee *Apis mellifera*. *Nature Genetics, 46*(10), 1081-1088. DOI: 10.1038/ng.3077

Wang, J.F., & Qiao, G.X. (2009). DNA barcoding of genus *Toxoptera* Koch (Hemiptera: Aphididae): Identification and molecular phylogeny inferred from mitochondrial COI sequences. *Insect Science, 16*(6), 475-484. DOI: 10.1111/j.1744-7917.2009.01270.x

Willis, L.G., Winston, M.L., & Honda, B.M. (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. DOI: 10.1016/1055-7903(92)90013-7

Wilson, A.C. (1976). Gene regulation in evolution. In Ayala, F. J. editor, Molecular evolution. (225-234 pp.). Sunderland, Massachusetts: Sinauer Associates Inc.

Xiao, J.H., Wang, N.X., Li, Y.W., Murphy, R.W., Wan, D.G., Niu, L.M., ... Huang, D.W. (2010). Molecular approaches to identify cryptic species and polymorphic species within a complex community of fig wasps. *PLoS ONE, 5*(11). DOI: 10.1371/journal.pone.0015067

Zhao, W., Tan, K., Zhou, D., Wang, M., Cheng, C., Yu, Z., Miao, Y., He, S. (2013). Phylogeographic analysis of *Apis cerana* populations on Hainan Island and southern mainland China, based on mitochondrial DNA sequences. *Apidologie, 45*(1), 21-33. DOI: 10.1007/s13592-013-0223-8

Zhen-ming, J., Yang, G., Huang, S., Liang, S., Ren, Z. (1992). The advancement of beekeeping science and technology in China. In Verma, L. R., editor, Honeybees in mountain agriculture. (274 pp). New Delhi: Oxford and IBH Publishing.