

Ken Tan<sup>1</sup>,  
Marina D. Meixner<sup>3\*</sup>,  
Stefan Fuchs<sup>2</sup>,  
Xuan Zhang<sup>1</sup>, Shaoyu He<sup>1</sup>,  
Irfan Kandemir<sup>3†</sup>,  
Walter S. Sheppard<sup>3</sup> &  
Nikolaus Koeniger<sup>2</sup>

<sup>1</sup>Eastern Bee Research Institute  
of Yunnan Agricultural  
University, Kunming, China

<sup>2</sup>Institut für Bienenkunde  
(Polytechnische Gesellschaft),  
Fachbereich Biowissenschaften  
der J.W. Goethe-Universität,  
Frankfurt am Main,  
Karl-von-Frisch-Weg 2, 61440  
Oberursel, Germany

<sup>3</sup>Department of Entomology,  
Washington State University,  
Pullman WA 99164-6382, USA

submitted January 2005

accepted June 2005

# Geographic distribution of the eastern honeybee, *Apis cerana* (Hymenoptera: Apidae), across ecological zones in China: morphological and molecular analyses

**Abstract** The biogeography and intraspecific variability of the eastern cavity-nesting honeybee, *Apis cerana*, are not very well known. We studied the variability of this species in China using morphometrical methods together with restriction and sequence analysis of two different regions of mitochondrial DNA. Samples of *A. cerana* were collected from feral or traditionally managed colonies in 19 locations of the Chinese mainland, covering the main ecological regions. Worker bees from each sample were dissected and morphometric characters were measured. The data were analysed with multivariate statistical procedures. Data were supplemented by previously published Chinese samples from the Oberursel data bank and reference samples of *A. cerana* from adjacent countries. A mitochondrial DNA fragment containing a non-coding region was amplified and analysed with the restriction enzyme *DraI*. This fragment was sequenced for two samples. For a subset of samples, the subunit 2 of the mitochondrial NADH gene was amplified and sequenced. Morphometric analysis revealed a high degree of variation, strongly associated with ecological zones and correlated with geographical and climatic parameters. Two main clusters were apparent, one comprised the bees from the southern tropical seasonal rain forest region, showing strong associations to the bees of Vietnam, Thailand and Myanmar. The second main cluster included the bees from the temperate deciduous broad-leaved forest region, the subtropical evergreen broad-leaved forest zone, the high, cold meadow and steppe region and the North, and showed increasing similarity to the bees of Korea and Japan. In particular, the bees from Qingzang plateau, on the fringe of the Gansu province, were set apart by their exceptional body size, darkness and pilosity. There was no variation in *DraI* restriction patterns within China. Sequence variation of the mitochondrial ND2 region was consistent with geographic patterns of morphological variation. Bees of the South Yunnan region were set apart by characteristically broad abdominal sterna and wax mirrors, with this locally restricted trait transcending the main transition line at the northern limit of the tropical seasonal rain forest region. This northern limit appears to correspond to the separation line between *A. cerana indica* and *A. cerana cerana*.

**Key words** *Apis cerana*, biogeography, China, morphometry, mtDNA

## Introduction

Compared with the western honey bee, *Apis mellifera* L., the biogeography and intraspecific variability of the eastern cavity-

nesting honey bee, *Apis cerana* Fabr., is not well known. The range of *A. cerana* extends from Pakistan to Japan and south to India and the Sunda and Philippine archipelagos. Comparatively few samples and populations have been studied so far and most studies have been restricted to morphometric methods. Based on multivariate statistical analysis of a suite of morphological characters, Ruttner (1988, 1992) suggested the

\*Corresponding author. Email: mmeixner@wsu.edu

†Present address: Department of Biology, Zonguldak Karaelmas University, Zonguldak 67100, Turkey.

recognition of four different groups, *A. c. cerana* from Afghanistan through China, *A. c. indica* from India through Indonesia, *A. c. japonica*, and a group from the eastern Himalaya mountains, referred to as *A. c. skorikovi* by Engel (1999). More recently, molecular markers have been used to study variation within *A. cerana* (Arias *et al.*, 1996; Smith & Hagen, 1996, 1999; de la Rúa *et al.*, 2000). Based on sequence analysis of the mitochondrial intergenic region between the tRNA for Leucine and the COII gene, Smith & Hagen (1996, 1999) proposed subdivision of *A. cerana* into three major lineages, a mainland Asia group that also includes the bees from Japan, a Sundaland group, and a Philippine group. However, results from both morphometric and molecular analyses remain insufficient to yield a consistent explanation of *A. cerana* phylogeography (reviewed in Hepburn *et al.*, 2001). It is particularly troublesome that very little is known about *A. cerana* in China, which includes the species' northern limit of distribution in mainland Asia.

China occupies an extremely large territory, encompassing various geographic regions and climate zones with a high degree of botanical and zoological diversity. Within China there are an estimated 2 million colonies of the eastern honeybee, *A. cerana* (Yaochun, 1993). Of the eight discernible botanical regions, *A. cerana* is primarily restricted to six: cold-temperate deciduous needle-leaved forest, temperate deciduous broad-leaved forest, subtropical evergreen broad-leaved forest, tropical seasonal rain forest, temperate steppe and temperate desert. Few *A. cerana* colonies occur in the small area of high, cold meadow steppe and none in the high, cold semi-desert and desert regions (Peng *et al.*, 1989).

Within this variety of environments, *A. cerana* exhibits extensive variation. Earlier studies separated *A. cerana* in China into five different subspecies, corresponding to the regions of Hailan, Eastern Yunnan, Southern Yunnan, Aba and Xizhang (Tibet) (Yang & Xue, 1982). Later, seven biotypes were described, which included the Palm forest and mountain biotype of Hailan, and the biotypes of Guangdong-Guangxi, Hunan, Yunnan Plateau, Northern and Changbeishan (Yang & Xue, 1986). Peng *et al.* (1989) reviewed the taxonomic studies of China and concluded that the patterns of variation from the north-eastern parts to the southern parts of China could not be reliably reconstructed from existing studies due to methodological and character ambiguities. Thus, the variation of *A. cerana* in China and relationships to the known subspecies of the East Asian mainland remain largely obscure, except for a recent study of bees from Yunnan and Beijing (Tan Ken *et al.*, 2003).

We investigated the variation of *A. cerana* in mainland China through morphometric and mtDNA analyses. A suite of 38 morphological characters was combined with analyses of two different regions of mitochondrial DNA: the region between the COI and COII gene that includes an intergenic non-coding sequence, and subunit 2 of the NADH gene. The goal of the study was to recognize geographic and genetic variability of *A. cerana* within China, analyse the relationship of morphometric characters to ecological parameters and to explore the relationship of the honey bees of China to *A. cerana* samples from adjacent regions.

## Materials and methods

### Collection of bee samples

Honey bee samples were collected between 1992 and 2003 from nine different regions throughout China (Fig. 1). The 19 localities were selected to represent different climatic or vegetation zones, ranging from 50 m to 2700 m in elevation (Table I). Honeybees were collected from natural nests or semi-managed hives, such as log-hives or natural cavities. Movable beehives have been introduced into some areas recently, but migratory beekeeping is uncommon.

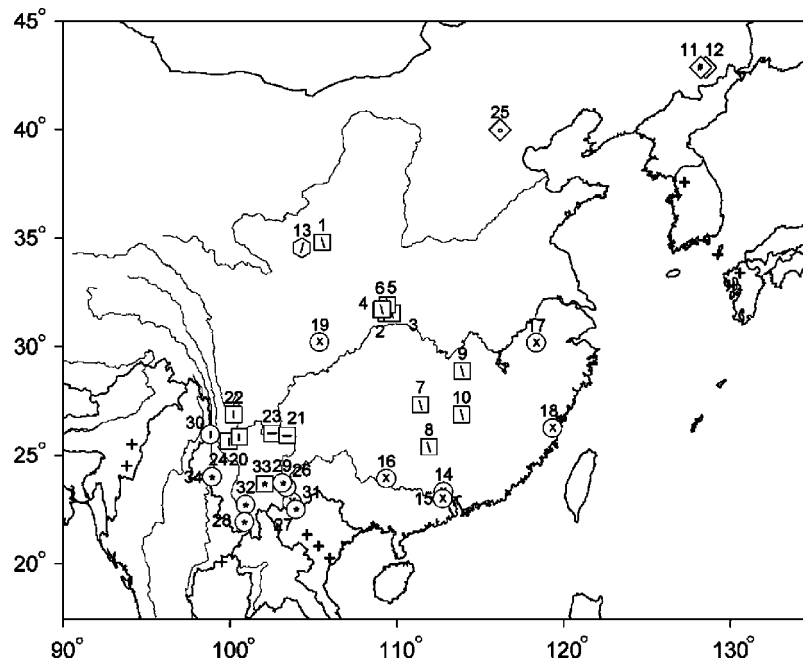
A total of 26 samples were collected. In each locality, between one and six colonies were sampled. Each sample contained 30 worker bees, which were preserved in 75% ethanol. Samples were split and one half was deposited in the Eastern Bee Research Institute of the Yunnan Agricultural University of China, Kunming, and the other half in the bee collection of the Institut für Bienenkunde, Oberursel, Germany.

### Preparations and measurements

All 26 colony samples were analysed at the Institut für Bienenkunde. From each sample, 15 worker bees were dissected for morphometric analysis and measured according to the methods described by Ruttner *et al.* (1978) and Ruttner (1988). Of the 41 morphometric characters listed in Ruttner *et al.* (1978), 38 were measured, excluding length of proboscis (No. 5) and cubital veins of the left wing (No. 29 and 30), resulting in 16 size characters, 11 wing angles, 7 colour characters, 3 hair characters and number of hamuli. Measurements and colour scaling were performed using a stereomicroscope and a computer-aided measuring system based on a video system and measuring program ((Bee2, © Meixner, 2004).

### Statistical analysis of the data

Colony sample means, standard deviation and standard error were computed for each character from the samples by the morphometric measuring program, thus representing estimates for the colony. Colony means for China were combined with data of 30 previously analysed samples from Yunnan and Beijing (Tan Ken *et al.*, 2003). To explore relations to bees outside of China, data of samples from Myanmar (2), Japan (8), Korea (5), Nepal (4), Thailand (8), Vietnam (17), Malaysia (2) and India (5) were included. Locations for all samples are given in Table I. Reference samples were collected from managed, semi-natural or natural nests. These data were taken from the morphometric data bank of the Institut für Bienenkunde, Oberursel, Germany. Four different methods were used to analyse the data, each focusing on specific aspects. To display the general morphological relations between samples, data were submitted to factor analysis and sample scores were plotted on principal component (PC) co-ordinates for visualization. Subsequently, the structuring of morphometric similarities was investigated by hierarchical cluster analysis, including sampling location means for the China mainland samples and group means for the bee populations from adjacent countries. Further, a k-means clustering procedure was performed with



**Figure 1** Sample locations in China used in the analysis. 1–19 sample locations from this study. 20–34 sample locations from Tan Ken *et al.* (2003). Location names and geographical coordinates are given in Table I. Circles: tropical seasonal rain forest ecological zone (TSRF); squares: subtropical evergreen broad-leaved forest ecological zone (SEBLF); Rhombus: temperate deciduous broad-leaved forest zone (TDBLF); hexagon: high, cold meadow and steppe zone (HCMS). Membership of each sample in K-means subcluster 1 to 8 is marked by x, +, –, l, \, /, #, ° within the symbols. For key to Chinese location numbers, see Table 1a, b. Sample locations in adjacent countries within map range are indicated by +.

increasing numbers of clusters until maximal coherence of geographic and ecological zones was obtained. To test the quality of the grouping obtained by factor and cluster analysis, samples were reallocated to their respective groups by discriminant analysis. Relations of morphometric traits to environmental variables including altitude, latitude and longitude, were investigated by correlation and regression analysis. Calculations were performed using the SPSS for Windows 10.00 and Systat 9.00 statistical packages.

### Molecular analyses

Subsets of each sample were transferred to 90% ethanol and shipped to Washington State University where the molecular analyses were performed. Total nucleic acids of one individual worker per sample were extracted according to methods of Arias & Sheppard (1996).

### *Dra*I restriction analysis

A mitochondrial fragment containing the intergenic region between the tRNA<sup>Leu</sup> gene and the second subunit of the cytochrome oxidase gene was amplified using the primer pair E2-H2 (Gamery *et al.*, 1993): E2: 5'-GGC AGA ATA AGT GCA TTG-3', H2: 5'-CAA TAT CAT TGA TGA CC-3'. The 50  $\mu$ L reaction mix consisted of 0.8  $\mu$ M of each primer, 0.2 mM of PCR Nucleotide mix (Boehringer Mannheim), 1.5 mM MgCl<sub>2</sub> (Invitrogen), 1X Reaction Buffer (Invitrogen), 2 U *Taq* Polymerase (Invitrogen) and 2  $\mu$ L of template. The amplification cycle consisted of an initial denaturation step of 5 minutes at 94 °C, followed by 35 cycles of 30 s at 92 °C, 90 s at 47 °C

and 2 min at 63 °C, followed by a final extension step of 10 minutes at 63 °C.

Ten  $\mu$ L of PCR products were run on a 1.5% agarose gel, stained with ethidium bromide and photographed under UV illumination. 20  $\mu$ L of each positive reaction were digested with the restriction enzyme *Dra*I at 37 °C overnight. Restriction fragments were separated on 10% polyacrylamide gels, stained with ethidium bromide and photographed under UV illumination. This analysis was performed with all samples of the current collection, three additional samples from the previous collection in Yunnan (Tan Ken *et al.*, 2003) and samples from adjacent populations of *A. cerana* from Thailand (20), Nepal (2), Vietnam (2), Beijing (2), and Korea (1). Subsequently, the amplification products of two samples from China were directly sequenced from both directions using a cycle sequencing protocol (Craxton, 1991) and an ABI 310 automated sequencer.

### Sequence analysis of the NADH subunit 2

For a subset of samples (Ganshu1, Ganshu2, Hubei, Sayinpan, Lijian, Hekou, Guangdong, Sichuan), together with reference samples of *A. cerana* from Nepal, Vietnam, Thailand and Sri Lanka, part of the mitochondrial NADH 2 gene was sequenced. We used the polymerase chain reaction to amplify a fragment of about 700 bp. The primers used were developed from published sequences of *Apis cerana* (Arias *et al.*, 1996; Koulianos & Crozier, 1999):

AcerND2F: 5'-TTT ATT CAT AAA TTT TAA AC-3'

AcerND2R2: 5'-AAA TCT AAT TAA TAT ATA A-3'

Location (voucher number)	N	Ecological region	Altitude (m above sea level)	Rainfall (mm/year)	Mean annual temperature (°C)	Latitude	Longitude	K-means cluster
<b>(a)</b>								
1 Ganshu 1 (3255–3226)	2	SEBLF	1200	300	8.00	33°13N	105°52E	5/6
2 Hubei 1 (3219)	1	SEBLF	500	1100	12.00	31°32N	109°23E	5
3 Hubei 2 (3220–3221)	2	SEBLF	790	1100	12.00	31°32N	109°43E	5
4 Hubei 3 (3222)	1	SEBLF	1100	1050	11.50	31°41N	109°05E	5
5 Hubei 4 (3223)	1	SEBLF	1650	1020	10.50	31°55N	109°26E	5
6 Hubei 5 (3224)	1	SEBLF	1800	1020	10.80	31°43N	109°08E	5
7 Hunan 1 (3215)	1	SEBLF	420	1200	13.50	26°80N	110°85E	5
8 Hunan 2 (3216)	1	SEBLF	480	1220	13.50	25°23N	111°56E	5
9 Hunan 3 (3217)	1	SEBLF	260	1300	14.60	28°53N	113°56E	5
10 Hunan 4 (3218)	1	SEBLF	280	1350	14.50	26°52N	113°52E	5
11 Jilin 1 (3233)	1	TDBLF	1000	800	6.20	42°52N	128°31E	7
12 Jilin 2 (3234–3235)	2	TDBLF	1100	800	6.20	42°52N	128°13E	7
13 Ganshu 2 (3227–3228)	2	HCMS	2700	160	2.00	34°33N	104°23E	6
14 Guangdong 1 (3230)	1	TSRF	100	1700	21.50	23°19N	112°48E	1
15 Guangdong 2 (3231–3232)	2	TSRF	50	1700	21.80	23°00N	112°45E	1
16 Guangxi (3212)	1	TSRF	200	1600	21.70	23°55N	109°22E	1
17 Huangshan (3210–3211)	2	TSRF	300	1500	15.60	30°10N	118°21E	1
18 Hujian (3229)	1	TSRF	400	1500	18.50	26°15N	119°20E	1
19 Sichuan (3213–3214)	2	TSRF	450	1300	16.00	31°32N	105°22E	1
<b>(b)</b>								
20 Binchuan (2996–2997)	2	SEBLF	1690	587	18.80	25°51N	100°33E	4
21 Huize (3016)	1	SEBLF	1500	822	12.70	25°55N	103°26E	3
22 Lijian (2998–2999)	2	SEBLF	2680	938	12.60	26°52N	100°13E	4
23 Sayinpan (3008–3009)	2	SEBLF	2250	968	14.00	26°00N	102°31E	3
24 Yuangbi (2994–2995)	2	SEBLF	1810	1077	16.10	25°39N	99°56E	4
25 Beijing (1370–1373)	4	TDBLF	250	644	11.40	39°59N	116°11E	6
26 Caoba (3010–3011)	2	TSRF	1250	827	18.60	23°32N	103°23E	3/2
27 Hekou (3006–3007)	2	TSRF	100	1777	22.60	22°30N	103°57E	2
28 Jianghong (3004–3005)	2	TSRF	600	1207	21.70	21°55N	100°52E	2
29 Kaiyuan (3014–3015)	2	TSRF	1600	813	19.70	23°43N	103°08E	2
30 Lushui (3000–3001)	2	TSRF	1950	1159	15.00	25°57N	98°49E	4
31 Pingbian (2992–2993)	2	TSRF	1400	1648	16.40	22°50N	103°43E	1/4
32 Simao (3002–3003)	2	TSRF	1100	1547	17.70	22°43N	100°56E	2
33 Yuangjian (3017)	1	TSRF	1500	801	19.80	23°41N	102°05E	2
34 ZhenKang (3012–3013)	2	TSRF	1000	1624	18.90	24°00N	98°55E	2
<b>(c)</b>								
Country	N	Location (voucher number)	Latitude	Longitude				
Myanmar	2	Taingyr Shan (1301–1302)	17°18N	95°57E				
India	1	Calcutta (2943)	22°34N	88°21E				
India	2	Manipur (1232–1233)	24°30N	88°21E				
India	1	Darjeeling/Lebong (1300)	26°58N	88°17E				
India	1	Gubudia (1329)	30°21N	80°32E				
Japan	1	Kyushn/Fukuoka (946)	33°23N	130°34E				
Japan	1	Izuhara-Chou (966)	34°14N	129°15E				

**Table 1.** List of locations and numbers of Chinese *A. cerana* samples per location (N), ecological regions, altitude, yearly rainfall, mean annual temperature and geographical positions of sample data used in this study. For abbreviations of ecological regions, see Fig. 1 caption. Cluster membership of K-means cluster analysis is indicated. Split numbers show different clustering within locations. **(a)** Samples from China collected and analysed for this study. **(b)** Sample data from China (Yunnan and Beijing) included from Tan Ken *et al.* (2003). **(c)** Sample data for reference samples from the Oberursel bee collection.

Country	N	Location (voucher number)	Latitude	Longitude
Japan	3	Shimoagata/Gun (967–969)	34°14N	129°15E
Japan	1	Yamanash (945)	35°33N	138°40E
Japan	2	Tokyo (947–948)	35°41N	138°40E
Korea	5	Seoul/Kyungki (1330–1334)	37°34N	127°15E
Malaysia	2	Malaysia (3236–3237)	4°19N	101°81E
Nepal	1	Nagarkot (1076)	27°41N	85°43E
Nepal	1	Kathmandu (2526)	27°45N	85°16E
Nepal	2	Jumla (1902–1903)	28°24N	83°33E
Thailand	1	Chantaburi (1803)	12°36N	102°07E
Thailand	2	Bangkok (1322, 1563)	13°39N	100°31E
Thailand	3	Khon Kaen (963–965)	16°26N	102°49E
Thailand	1	Lampuu Phitsanuloke (1562)	16°47N	100°16E
Thailand	1	Fang (1321)	20°05N	99°30E
Vietnam	2	Can Tho (2241–2242)	10°02N	105°47E
Vietnam	1	Long Khanh (2240)	10°38N	106°21E
Vietnam	5	Cuc Phuong (2209–2213)	20°15N	105°58E
Vietnam	5	Hoa Binh (1335–1338)	20°49N	105°19E
Vietnam	5	Moc Chau (2214–2218)	21°20N	104°37E

**Table 1** Continued.

Reactions were performed in a total volume of 50  $\mu$ L with a final concentration of 1X reaction buffer, 1.5 mM MgCl<sub>2</sub>, 0.2 mM of each dNTP, 0.8  $\mu$ M of each primer and 2 units of taq polymerase. The amplification cycle consisted of 94 °C (60 s), 44 °C (80 s), 68 °C (120 s) and was repeated 35 times, followed by a final extension step of 5 min at 72 °C. Products were electrophoresed on a 1.5% agarose gel, stained with ethidium bromide and photographed under UV light. Amplification products were directly sequenced from both directions using the cycle sequencing protocol and an ABI 310 automated sequencer. Sequences were deposited in GenBank under Accession numbers: AY849558 – AY849569.

### Sequence alignment and phylogenetic analysis

The sequences were aligned using ClustalX (Thompson *et al.*, 1997) and adjusted manually where necessary. Phylogenetic analyses were performed using MEGA version 2.1 (Kumar *et al.*, 2001). Confidence probabilities for the minimum evolution tree (Rzhetsky & Nei, 1993) were calculated using the interior branch test algorithm of Rzhetsky & Nei (1992).

## Results

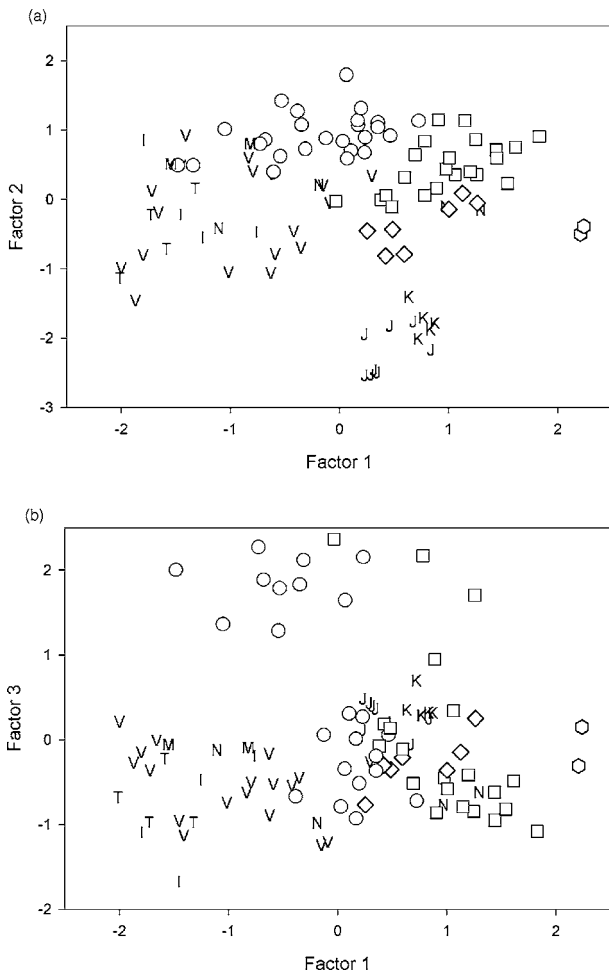
### Morphometric analysis

Factor analysis (FA) of the 38 morphometric characters performed on the 107 sample means yielded three factors with high eigenvalues (> 3.5) that accounted for 64.4% of the total variation in the data. The first factor accounted for 37.6% of the total variation and was positively associated ( $r > 0.6$ ) with the size measures length of abdominal terga 3 and 4, length of sternum 3 and 6, length of wax mirror, length and width of forewing, length of femur and tibia, length and width of metabasitarsus, length of cubital vein 1, and with the hairlength. The second factor accounted for 16.9% of the total variation

in the data, and was positively associated with the distance of the wax mirrors, the width of the dark stripe of the tomentum, the length of the cubital vein 2, the pigment on abdominal terga 3, 4 and scutellum and the wing venation angle A4. It was negatively associated with the width of the tomentum and wing venation angle B4. The third factor accounted for 9.8% of the total variation in the data and was positively associated with the width of abdominal sternum 6, the width of the wax mirror and wing venation angle O26, but negatively associated with wing venation angles N23 and J16.

The plots of the sample scores on the three principal component axes are presented in Fig. 2 (a and b). The sample data from China show considerable variability in these plots and samples from the four ecological zones clearly occupy different plot regions. Bees from the tropical forest zone (TSRF) were positioned towards smaller sizes (low factor 1 values) and lighter colour (high factor 2 values) than samples from the subtropical evergreen forest zone (SEBLF), with only little overlap. Within the SEBLF group some substructure can be observed, with larger bees mostly originating from the eastern regions of Hunan, Hubei and Ganshu, while smaller bees were found in north Yunnan (e.g. Lijian, Yuangbi and Sayinpan). Samples from the temperate deciduous broad-leaved forest zone (TDBLF, Beijing and Jilin) were in general darker, but similar in size to those of the SEBLF. Two samples from the fringe of high, cold meadow and steppe zone (HCMS) in the Ganshu province were set apart, due to their large size, dark coloration of the terga and high pilosity scores, which exceeded that of any other sample in this analysis.

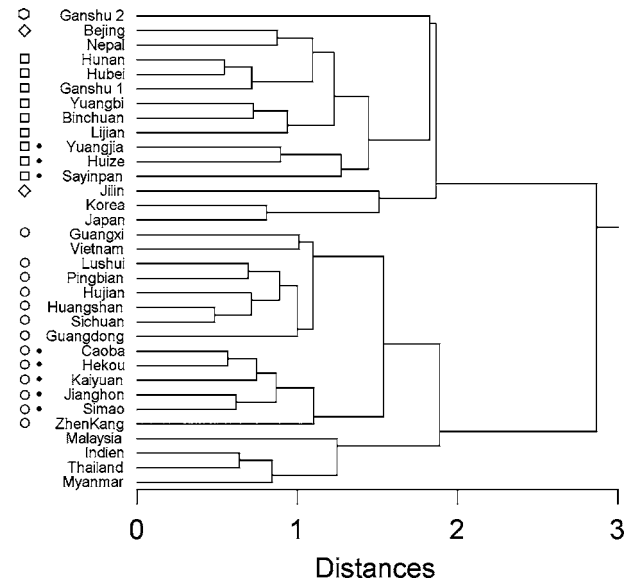
Samples from the TSRF zone showed overlap with the bees of the adjacent southern countries (Vietnam, Thailand, Malaysia, Myanmar, North India), but were lighter coloured and larger, reaching the size of samples from Korea or Japan. Samples from the SEBLF group even exceeded in size those of Korea and Japan, but were still lighter coloured.



**Figure 2** Factor scores (principal component) of colony samples from China and adjacent countries. Different symbols mark ecological zones, as in Fig. 1. K = Korea, I = India, J = Japan, M = Myanmar, N = Nepal, T = Thailand, V = Vietnam. Abscissa: Factor 1. Ordinate: Factor 2 (Fig. 2a) and factor 3 (Fig. 2b), respectively.

In regard to factor 3, mostly characterized by broad sterna and wax mirrors, some samples formed a distinct group with high factor 3 scores, consisting mostly of TSRF zone bees but also including 3 SEBLF samples (Fig. 2b). Regardless of their ecological zone of origin, all these samples came from the region of southern or northern Yunnan.

A cluster analysis was performed on mean values combined for locations or regions, using complete linkage between groups based on z-normalized group means of the characters. The resulting dendrogram (Fig. 3) showed two main clusters. The first cluster combined all sample locations from the tropical seasonal rain forest ecological zone (TSRF) of China and did not contain any samples from the other zones. Samples from Thailand, Myanmar and North India were included into this cluster, but formed a separate sub-branch. Another separate sub-branch was formed by samples from southern Yunnan with high factor 3 scores in the factor analysis, indicating exceptionally wide sterna and wax mirror. Samples from Vietnam were included in the third sub-branch with the rest of the samples from China.



**Figure 3** Hierarchical clustering dendrogram, derived from complete clustering on standardized values averaged for regions. The ecological zones of origin are indicated by symbols as in Fig. 1. Sample locations characterized by high factor 3 scores (see Fig. 2b) are marked by points.

The second main cluster included the sample locations from the three other ecological zones in China together with samples from Japan, Korea and Nepal. Within this cluster, the one sample location from the high, cold meadow and steppe zone (Ganshu2) was set apart, while a sub-cluster incorporated the north-eastern location of Jilin (TDBLF zone) together with Korea and Japan. In this analysis, all of the SEBLF locations were combined into a cluster that included the locations of Beijing (TDBLF) and Nepal. Further differentiation within the SEBLF locations separated the north-eastern group (Hunan, Hubei and Ganshu1) and the north Yunnan group (Yuangbi, Binchuan and Lijian) each into their own sub-cluster. Finally, a third sub-branch was formed, containing locations whose bees were characterized by exceptionally wide sterna and wax mirrors (indicated by high factor 3 scores in the factor analysis).

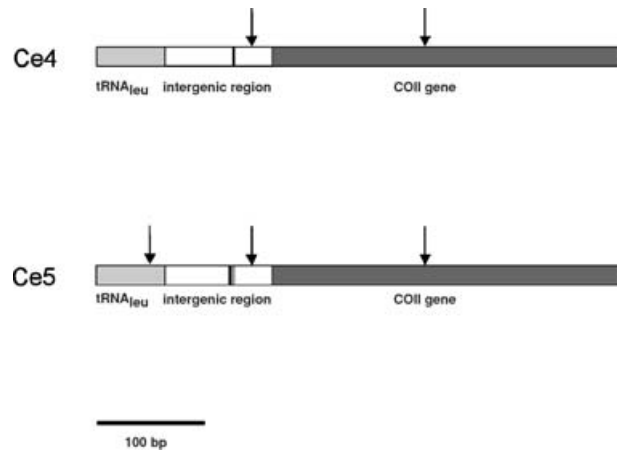
The clusters obtained were re-examined by discriminant analyses (DA). In a first analysis, the main branching between the tropical seasonal rain forest region and the other three regions combined was confirmed for all TSRF samples and all 30 samples from the combined SEBLF, TDBLF and HCMS groups ( $P > 0.99$ , except two samples from TSRF with  $P > 0.95$ , and two samples from the other regions with  $P = 0.94$  and  $P = 0.65$ ). When grouped by the four ecological zones in a second analysis, all 56 samples were reallocated into their respective group with high post-hoc probabilities ( $P > 0.99$ ), indicating a close connection between morphology and ecological zone of origin. In a third analysis, samples from Yunnan that were characterized by distinctly high factor 3 scores in factor analysis were defined as a distinct group (highf3), and were analysed together with the remaining TSRF and SEBLF samples each as a separate group. The 14 highf3, 15 TSRF and 18 SEBLF samples were reallocated into their respective groups with high post-hoc probabilities ( $P > 0.99$ ).

A k-means clustering procedure was performed and 8 clusters were obtained by maximizing geographic and ecological zone coherence. The cluster membership of each sample is listed in Table I and incorporated in Fig. 1. In this analysis the tropical seasonal rain forest samples were further subdivided into two clusters (cluster 1 and 2), separating the samples originating from the South Yunnan tropical seasonal rain forest region. The subtropical evergreen broad-leaved forest region split into three clusters, one containing samples from Sayinpan and Huize (but including one sample from Caoba, located in the TSRF zone), a second one including samples from the western SEBLF zone (North Yunnan, cluster 4) and a third one containing those from the eastern SEBLF zone (Hubei, Hunan, Ganshu1; cluster 5). The two temperate deciduous broad-leaved forest zone locations clustered separately (cluster 6 and 7) as did the samples from high, cold meadow and steppe region in Ganshu province (cluster 8). Examining distance measures between group centroids of factor analysis scores revealed proximities between the two subtropical evergreen broad-leaved forest region clusters (cluster 4 and 5), the two Yunnan clusters (2 and 3) and the two temperate deciduous broad-leaved forest zone locations (cluster 6 and 7). Distances further indicated proximities between cluster 4 (North Yunnan TSRF) and cluster 6 (Beijing TDBLF).

The morphological traits of the China samples showed clear relations to geographical and environmental variables. We observed significant correlation ( $P < 0.05$ ) of 26 traits with latitude, of 17 with longitude, of 10 with altitude and of 23 with rainfall. A multiple regression of factor scores on the environmental variables indicated that the strongest influence on factor 1, positively associated with measures of size, was exerted by temperature (standardized coefficient  $-1.13$ ,  $P < 0.0005$ ), followed by latitude and longitude (standardized coefficients  $-0.88$ ,  $P < 0.001$  and  $-0.449$ ,  $P < 0.009$ ), while rainfall or altitude had no separate effect. Relations with factor 2 scores, positively associated with light coloration, were less pronounced, and only latitude, rainfall, and altitude showed significant influences (standardized coefficients  $-0.59$ ,  $P < 0.02$ ,  $0.41$ ,  $P < 0.01$  and  $-0.36$ ,  $P < 0.03$ ).

### DraI restriction analysis

No length polymorphisms were detected in the fragment amplified with the E2 and H2 primer pair in any of the samples from China or from adjacent regions of Asia. The length of the amplified fragment was approximately 480 bp. Digestion with *DraI* resulted in two different restriction patterns that differed in the presence of a single restriction site. One of the restriction patterns consisted of three fragments that were 185 bp, 155 bp and about 140 bp long, similar to the Ce4 haplotype described by De la Rúa *et al.* (2000) in *A. cerana* from Palawan. This pattern was observed exclusively in samples from Thailand where it was predominant (16 of 20 samples, mainly in the southern part of the country). All samples collected from China, regardless of the ecological zone of origin, together with all other samples from adjacent countries, exhibited another restriction pattern. This pattern was named Ce5 and consisted of four fragments of 185 bp, 155 bp, about 93 bp, and 45 bp in length.



**Figure 4** Restriction map of the mitochondrial COI–COII region of *Apis cerana* containing a non-coding intergenic sequence. The Ce4 haplotype was described by De la Rúa *et al.* (2000) for samples from Palawan. Restriction sites are marked by arrows. Open and filled bars indicate insertions and deletions of single base pairs, respectively.

The only variation observed consisted of some minor length polymorphisms within the  $\sim 93$  bp fragment (and the  $\sim 140$  bp fragment, respectively), probably caused by insertions or deletions of single base pairs. Fragment lengths were confirmed with sequence analysis. Restriction maps are shown in Fig. 4.

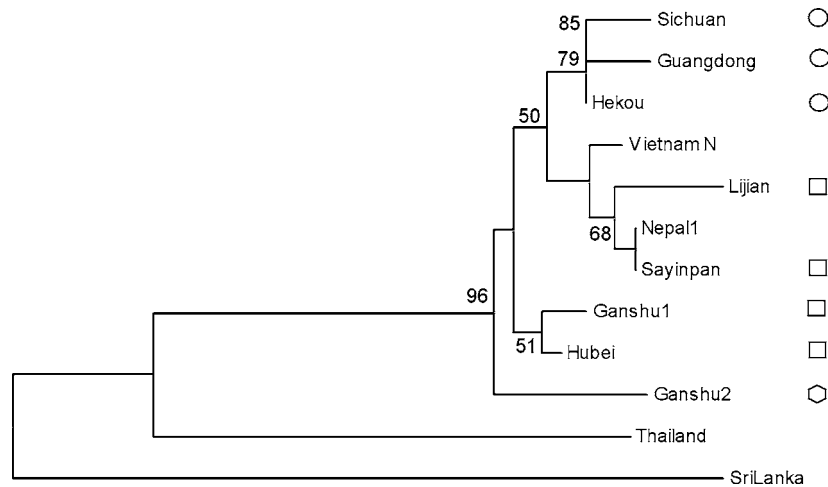
### Sequence analysis of the COI–COII intergenic region

For two samples from China (Ce5), the entire fragment, consisting of part of the tRNA<sup>leu</sup> gene, the noncoding intergenic region and the 5' end of the COII gene was sequenced in both directions. Comparison of our data of the intergenic region with published sequences of the noncoding region in *A. cerana* (Smith & Hagen, 1996) demonstrated identity of one sequence (containing a 93 bp restriction fragment) with the Japan1 haplotype, while the other sequence (containing a 92 bp fragment) was identical with the Thai1 haplotype reported by Smith & Hagen (1996).

### ND2 sequence analysis

The amplification of the mitochondrial ND2 gene produced a fragment of approximately 660 bp. Among all sequenced samples, a fragment of 416 bp was successfully aligned. Within this fragment, 32 sites were variable, 9 of which were informative under parsimony. The sequence divergence was highest between the samples from China and the reference samples from Sri Lanka (4.6%–5%) and southern Thailand (3.1%–3.8%). Within China, the sequence divergence was low, ranging between 0 and 1.2% (including the reference samples from Vietnam and Nepal).

Both minimum-evolution and parsimony trees yielded similar overall topologies (parsimony tree not shown). The minimum-evolution tree (Fig. 5) showed the honey bee populations from China to differ most strongly from those of Sri Lanka and southern Thailand, but combined them into the same main cluster as the reference samples from Nepal and northern



**Figure 5** Minimum-evolution tree. Confidence probabilities above 50% (interior branch test, 2000 replications) are indicated on branches. Symbols for ecological zones of origin as in Fig. 1.

Vietnam. Within this cluster, however, a close correlation between the branching pattern and the geographic regions of sample origin was not strongly supported. Nonetheless, some differentiation was observed. Apart from the sample from the high, cold meadow environment that splits off by itself, a deep branch separates the Yunnan SEBLF and Nepal samples from the eastern SEBLF group which groups closer to the samples from the tropical forest environment and Vietnam.

## Discussion

Our studies confirmed that the morphometric features of the bees of China show a wide range of variation in measurements of size, pigmentation and pilosity. This high variation reflects the ecological diversity of China where the samples were taken. One of the most prominent results of this study was the close relationship found between morphology and ecological zone of origin, as demonstrated by the factor score plots and the cluster analyses. The close connection between the ecological zone and bee morphology was also apparent in the strong correlation of many of the traits with the environmental and geographical variables. The results confirmed the trends that bees get bigger and darker at higher elevation, also found in former studies on *A. mellifera* (Ruttner, 1988; Hepburn *et al.*, 2000), *A. cerana* (Verma *et al.*, 1994; Tan Ken *et al.*, 2003) and *A. florea* (Ruttner *et al.*, 1995).

In the main pattern, honey bee populations of China could be separated into a southern and a northern group, with the southern group containing all samples except one (Yuangjian) of the tropical seasonal rain forest, and a northern group containing the samples of the temperate deciduous broad-leaved forest, the subtropical evergreen broad-leaved forest and the temperate steppe. A significant degree of differentiation could be observed among the bees of the three northern ecological zones. In particular, the bees of the west China Qingzang plateau, on the fringe of high, cold meadow and steppe region in Ganshu province at an altitude of 2700 m (Ganshu 2) took extreme positions in all analyses as they were particularly big, had the longest hairs and the darkest coloration. At a still finer

level of resolution eight subgroups emerged from K-means cluster analysis, which were, with few exceptions, consistent with the former analyses. In particular, the bees characterized by distinctly broad sterna and wax mirrors (South Yunnan) (Tan Ken *et al.*, 2003) transgress the boundaries of two ecological zones, the tropical seasonal rainforest zone (subgroup 2) and the subtropical evergreen broad-leaved forest (subgroup 3).

The grouping pattern discussed above does in part corroborate major groupings in earlier Chinese literature, as reviewed in Peng *et al.* (1989), who discussed 5 races and 7 biotypes. Our samples from Ganshu2 by description and distribution evidently represent members of the 'Aba race', which has been recognized as a separate subspecies, *Apis cerana heimifeng*, by Engel (1999). All other samples, except those from South Yunnan ('Southern Yunnan race'), fall into the 'Eastern race' of Peng *et al.* (1989). While Peng *et al.* (1989) stated that the data available to them did not permit systematic conclusions about the populations investigated, they referred to the 'South Yunnan race' as most likely being *A. c. indica* and quoted the 'Eastern race' as belonging to *A. c. cerana*. The five biotypes described within the 'Eastern race' apparently correspond to subclusters recognized in our study (Guangdong-Guanxi: cluster 1; Hunan: subgroup 5, Yunnan plateau: subgroup 4; Northern: subgroup 6; Chanbeishan: subgroup 7). However, the strong contrasts within this 'Eastern race' do not support the concept of a homogeneous subspecies, but rather suggest further subdivisions, apparently related to the three ecological zones covered. From current knowledge, however, it appears premature to propose a systematic revision.

The restriction and sequence analysis of the mitochondrial COI-COII region did not reveal sufficient variation within China to support or reject the hypotheses based on the morphological data. All samples from China exhibited haplotypes that identified them as belonging to the 'Mainland Asia' group of *Apis cerana* as described by Smith & Hagen (1996, 1999; Smith *et al.*, 2000). This group of closely related haplotypes has been reported from India and Japan with apparently very

little variation, a fact that has been interpreted as a consequence of rapid post-glacial colonization (Smith *et al.*, 2000). This hypothesis is strengthened by the presence of these haplotypes throughout China, but it appears that the variation within this mitochondrial region is insufficient to fully resolve phylogeographic relationships within mainland Asia. The biogeographic transition between the mainland Asia group and the Sundaland group of haplotypes in the non-coding intergenic region (Smith *et al.*, 2000) is reflected in our study by the dominance of the Ce4 restriction pattern in samples from the southern part of Thailand.

Variation of the mitochondrial subunit 2 of the NADH gene has been used to investigate phylogenetic relationships within *Apis mellifera* (Arias & Sheppard, 1996) and some island populations of *A. cerana* (Arias *et al.*, 1996). Prior to this study, no data from samples of the mainland range of *A. cerana* have been available. The sequence divergence of samples within China and between China and the other mainland sources (Vietnam and Nepal) of up to 1.2% is similar to the sequence divergence observed within *A. mellifera* (up to 2%, Arias & Sheppard, 1996). However, the divergence between mainland and insular (Sri Lanka) or peninsular (Thailand) *A. cerana* populations is much higher, as also shown by Arias *et al.* (1996). Data from our study support separation between an 'eastern group' and the populations from Yunnan, and probably the existence of a distinct population in the high, cold meadow environment. However, our data do not yet appear to be comprehensive enough to allow statistically definite conclusions, or to propose or confirm subspecies or lineages of *A. cerana* in China.

### Relations to bees of adjacent countries

Relations of the Chinese bees to *A. cerana* from countries outside of China supported a north-south transition. In the factor analysis plots and in cluster analysis the tropical seasonal rain forest region bees fell close to those of Thailand, Myanmar, Malaysia and Vietnam, while the three northern ecological zones were associated with Nepal, Korea and Japan. It is tempting to associate the northern limit of the tropical seasonal rain forest with the northern limit of *A. c. indica*, which appears to fairly accurately trace the distribution limit of the cluster analysis of the southern branch. The northern branch would then be related to *A. c. cerana*, which is apparent from the close association of the northeastern temperate deciduous broad-leaved forest region samples of Jilin with the bees of Korea, Japan and Nepal. Peng *et al.* (1989) reported that the South Yunnan bees belonged to *A. c. indica*. The current study suggests, based on results of morphometry and ND2 analysis, a clear coherence between these and the more eastern tropical seasonal rain forest samples (Guangdong-Guanxi, cluster 1), a possibility that is also supported by behavioural descriptions (Peng *et al.*, 1989). Thus, a transition between *A. c. indica* and *A. c. cerana* along the northern limit of the tropical seasonal rain forest appears arguable, although too few samples were analysed to be conclusive. Interestingly, it appears that the main transition zone for mitochondrial DNA might be even further to the south. That is, COI-COII restriction analysis

showed two different patterns in Thailand, with the second pattern predominant in the south of Thailand. ND2 data also demonstrated high sequence divergence between samples from China and Thailand (3.5% compared to around 1.0% within China).

The current description of both the morphological patterning and molecular variation remains incomplete due to the limited number of samples available but serves as a starting point for further investigation. In particular, the bees from the Qingzang plateau in western China, on the fringe of the high, cold meadow and steppe region (Ganshu 2), which stood out as exceptionally large, dark and hairy, suggest that surprising variation remains to be found in the northern regions of central China.

### Acknowledgements

We wish to thank the Yunnan Agricultural University and Yunnan University, Yunnan Animal Husbandry Administrative Bureau, the colleagues Liu Yiqiu, Ling Zhuncheng, You Fangdong and Zhang Xuepeng for invaluable help with the collection of bee samples. We further thank B. Springer for assistance with the bee preparations. The Project was supported by a visitor's grant of the J.-W. Goethe Universität, Frankfurt am Main, to Tan Ken.

### 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**, 557–566.
- ARIAS, M.C., TINGEK, S., KELITU, A. & SHEPPARD, W.S. 1996. *Apis nuluensis* Tingek, Koeniger and Koeniger, 1996 and its genetic relationship with sympatric species inferred from DNA sequences. *Apidologie* **27**, 415–422.
- BEE2. MORPHOMETRIC SOFTWARE. © 1996–2004, A. Meixner and M.D. Meixner, Frankfurt, Germany.
- CRAXTON, M. 1991. Linear amplification sequencing: a powerful method for sequencing DNA. *Methods* **3**, 20–24.
- DE LA RÚA, P., SIMON, U.E., TILDE, A.C., MORITZ, R.F.A. & FUCHS, S. 2000. MtDNA variation in *Apis cerana* populations from the Philippines. *Heredity* **84**, 124–130.
- ENGEL, M. 1999. The taxonomy of recent and fossil honey bees (Hymenoptera: Apidae: *Apis*). *Journal of Hymenopteran Research* **8** (2), 165–196.
- GARNERY, L., SOLIGNAC, M., CELEBRANO, G. & CORNUET, J.M. 1993. A simple test using restricted PCR-amplified mitochondrial DNA to study the genetic-structure of *Apis mellifera*. *Experientia* **49** (11), 1016–1021.
- HEPBURN, H.R., RADLOFF, S.E. & OGHIAKE, S. 2000. Mountain bees of Africa. *Apidologie* **31**, 205–221.
- HEPBURN, H.R., SMITH, D.R., RADLOFF, S.E. & OTIS, G.W. 2001. Intraspecific categories of *Apis cerana*: morphometric, allozymal and mtDNA diversity. *Apidologie* **32**, 3–24.
- KEN, T., FUCHS, S., KOENIGER, N. & RUIGUANG, Z. 2003. Morphological characterization of *Apis cerana* in the Yunnan Province of China. *Apidologie* **34**, 553–562.
- KOULIANOS, S. & CROZIER, R.H. 1999. Current intraspecific dynamics of sequence evolution differs from long-term trends and can account for the AT-richness of honeybee mitochondrial DNA. *Journal of Molecular Evolution* **49**, 44–48.
- KUMAR, S., TAMURA, K., JAKOBSEN, I.B. & NEI, M. 2001. MEGA2: Molecular Evolutionary Genetics Analysis software. *Bioinformatics* **17**(12), 1244–1245.

- PENG, Y.S., NASR, M.E. & LOCKE, S.J. 1989. Geographical races of *Apis cerana* Fabricius in China and their distribution. Review of recent Chinese publications and a preliminary statistical analysis. *Apidologie* **20**, 9–20.
- RUTTNER, F. 1988. *Biogeography and Taxonomy of Honeybees*. Springer-Verlag, Berlin.
- RUTTNER, F. 1992. *Die Naturgeschichte der Honigbienen*. Ehrenwirth, Munich.
- RUTTNER, F., MOSSADEGH, M.S. & KAUHAUSEN, D. 1995. Distribution and variation in size of *Apis florea* F. in Iran. *Apidologie* **26**, 477–486.
- RUTTNER, F., TASSENCOURT, L. & LOUVEAUX, J. 1978. Biometrical-statistical analysis of the geographic variability of *Apis mellifera*. *Apidologie* **9**, 363–381.
- RZHETSKY, A. & NEI, M. 1992. A simple method for estimating and testing minimum-evolution trees. *Molecular Biology and Evolution* **9**, 945–967.
- RZHETSKY, A. & NEI, M. 1993. Theoretical foundation of the minimum-evolution method of phylogenetic inference. *Molecular Biology and Evolution* **10**, 1073–1095.
- SMITH, D.R. & HAGEN, R.H. 1996. The biogeography of *Apis cerana* as revealed by mitochondrial DNA sequence data. *Journal of the Kansas Entomological Society* **69** (4), suppl. 294–310.
- SMITH, D.R. & HAGEN, R.H. 1999. Phylogeny and biogeography of *Apis cerana* subspecies: testing alternative hypotheses. In: HOOPINGARNER, R. and CONNOR, L., Eds., *Apiculture for the 21st century*. Wicwas Press, Cheshire, pp. CT, 60–68.
- SMITH, D.R., VILLAFUERTE, L., OTIS, G. & PALMER, M. 2000. Biogeography of *Apis cerana* F. and *A nigrocincta* Smith: insights from mtDNA studies. *Apidologie* **31**, 265–279.
- THOMPSON, J.D., GIBSON, T.J., PLEWNIAK, F., JEANMOUGIN, F. & HIGGINS, D.G. 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* **24**, 4876–4882.
- VERMA, L.R., MATTU, V.K. & DALY, H.V. 1994. Multivariate morphometrics of the Indian honeybee in the northwest Himalayan region. *Apidologie* **15**, 203–223.
- YANG, G. & XUE, S. 1982. Geographic distribution and variation of China *Apis cerana* (in Chinese). *China Beekeeping* **3**, 17–19.
- YANG, G. & XUE, S. 1986. Geographic distribution of *Apis cerana* Fabr. and their subspecies diversity (in Chinese). *Journal of Yunnan Agriculture University* **1**, 86–92.
- YAOCHUN, C. 1993. *Apiculture in China*. China Agricultural Publication, Beijing, 21. (In Chinese)