Review article

Intraspecific categories of *Apis cerana*: morphometric, allozymal and mtDNA diversity

H. Randall Hepburn\textsuperscript{a*}, Deborah R. Smith\textsuperscript{b}, Sarah E. Radloff\textsuperscript{c}, Gard W. Otis\textsuperscript{d}

\textsuperscript{a} Department of Zoology and Entomology, Rhodes University, Grahamstown 6140, South Africa
\textsuperscript{b} Department of Entomology, University of Kansas, Lawrence, Kansas 66045, USA
\textsuperscript{c} Department of Statistics, Rhodes University, Grahamstown 6140, South Africa
\textsuperscript{d} Department of Environmental Biology, University of Guelph, Guelph, Ontario, N1G 2W1, Canada

(Received 4 July 2000; accepted 29 September 2000)

Abstract – An analysis of the infraspecific categories of *Apis cerana* was prepared from the relevant literature on taxonomy, morphometrics, allozyme polymorphism and mtDNA diversity. About 31 putative biometric groups have been proposed and assigned to about eight equivocal “subspecies” and various ecotypes. However nearly half of the area of distribution of *A. cerana* remains unexamined. Allozyme polymorphism is greatest in southeast and lowest in northern and western Asia. About four major mtDNA groups are discernable. There is a very low overall geographic congruity amongst the morphoclasters, allozyme polymorphs and mtDNA clusters. The greatest problems in resolving infraspecific categories in *A. cerana* are inadequate sampling, incompatible differences in sample sizes, character suites, sampling distance, confidence limits and range of geographical scales employed in different studies.

*Apis cerana* / taxonomy / biogeography / Asia / honeybees

1. INTRODUCTION

The history of honeybee classification reflects a slow movement away from the fixed abstractions of the linnaean system to the analysis of population dynamics in multivariate probability terms (Hepburn and Radloff, 1998; Ruttner, 1988). Contemporary classification of honeybees stems from multivariate methods of analysis originally advanced by DuPraw (1964, 1965) and substantially developed by Ruttner (1988),...
Ruttner et al. (1978) and Daly (1991, 1992). The last decade has been particularly fruitful in this regard. Ruttner (1988) completed the daunting task of providing the first multivariate analytical attempt at a comprehensive macroscale synthesis of honeybee classification for the genus *Apis*. This was a major impetus for subsequent mesoscale studies of honeybee morphometrics in *A. mellifera* L. (cf. Hepburn and Radloff, 1998) as well as for the analysis of allozymic and DNA diversity of honeybee populations (Arias and Sheppard, 1996; Cornuet and Garnery, 1991; Smith, 1991; Smith et al., 1991).

In parallel with studies of *A. mellifera*, the Ruttner monograph (1988) also opened a new chapter in the study of the classification and biogeography of the honeybees of Asia. This is particularly evident in a recent spate of Asian regional studies in several journals and monographs (Verma, 1990, 1992). The purpose of this communication is to report the results of a survey on the published literature to date on the infraspecific classification of *A. cerana* Fabr. throughout its entire natural range of some 30 million km². The sympatric occurrence of other *Apis* species with *A. cerana* in southeastern Asia raises a very undesirable spectre: some previous "*A. cerana*" literature could well be contaminated through the inadvertent inclusion of data derived from species other than *A. cerana*. The likelihood of detecting such errors seems remote at best. We have included the results of morphometric studies as well as those on allozymic and DNA diversity. In the end we present a current portrait of putative infraspecific categories of *A. cerana*. There were no formal “Materials and Methods” for this study, solely analyses of the published literature cited in the references. It should be noted that the approach taken was one of trying to resolve distinct groups of *A. cerana* independently of “correcting” their complex taxonomic history in terms of the International Code of Zoological Nomenclature (Engel, 1999).

2. RESULTS AND DISCUSSION

The natural distribution of *A. cerana* based on published citations with reference to specific localities is shown in Figure 1. Literature consulted was predominantly that recoverable from *Apicultural Abstracts* (1950–1999). Positions of localities (closed circles) are only approximate because of map scale; areas in which *A. cerana* has been specifically sought but found absent are indicated with stars (Fig. 1). There are several regions where *A. cerana* undoubtedly occurs but for which there is extremely sparse or no data at all (Afghanistan, much of India, Laos, Cambodia, Myanmar and Sumatra) and other areas where it has recently been introduced (Papua New Guinea) but these are not considered.

2.1. Morphometrics

2.1.1. Western Asia

This region extends from the western borders of Afghanistan to the north and Pakistan to the south at about longitude 60°, thence eastwards below the Himalayan mountain range and across the Indian subcontinent to Myanmar at about longitude 94° (Fig. 2). The extent and quality of information on the honeybee populations of this area (about 4 million km²) is extremely variable and ranges from anecdotal descriptions to full multivariate statistical analyses of morphometric characters.

The only information on the classification of the honeybees of Afghanistan and Pakistan (Fig. 2, area 1) are those of Maa (1953) and Ruttner (1988, 1992) who concluded, respectively, that these bees could not be morphologically or morphometrically discriminated from those of neighbouring China (classically regarded as *A. cerana cerana* (Ruttner et al., 1989). However, extremely few samples from this area were available to Ruttner. Further possible discrimination of these honeybee
Infraspecific categories of *Apis cerana* well as population structure for the honeybees of this region is blurred because of fundamental incompatibilities in the methodologies and forms of analysis employed in studies published to date (Fig. 2, areas 2–3, 7–11, 14–16). This becomes particularly evident from comparisons of three important publications based on the Indian subcontinent. Firstly, Kshirsagar (1983) proposed some seven ecotypes for the bees of India vis-a-vis the more usually accepted “hills” and “plains” varieties or ecotypes (Kapil, 1956; Narayanan et al., 1960, 1961a, b; Ruttner, 1988). Secondly, there is a series of papers by Verma and colleagues (Mattu and Verma, 1983a, b; 1984a, b; Sihanuntavong et al., 1999; Singh et al., 5

The honeybees of the sub-Himalayan and Indian regions have been intensively investigated in recent years. However, the current picture of biometric groups and ecotypes as well as population structure for the honeybees of this region is blurred because of fundamental incompatibilities in the methodologies and forms of analysis employed in studies published to date (Fig. 2, areas 2–3, 7–11, 14–16). This becomes particularly evident from comparisons of three important publications based on the Indian subcontinent. Firstly, Kshirsagar (1983) proposed some seven ecotypes for the bees of India vis-a-vis the more usually accepted “hills” and “plains” varieties or ecotypes (Kapil, 1956; Narayanan et al., 1960, 1961a, b; Ruttner, 1988). Secondly, there is a series of papers by Verma and colleagues (Mattu and Verma, 1983a, b; 1984a, b; Sihanuntavong et al., 1999; Singh et al.,

---

Figure 1. Known distribution of *A. cerana* = (circles); reported absence = (stars).
Infraspecific categories of *Apis cerana*

As the bees of southern Pakistan would group together as populations of *A. cerana indica* in the Ruttner system. To complete the tally for western Asia, the honeybees of Afghanistan and northern Pakistan would presumably show close affinities to *A. cerana cerana*. The three ecotypes proposed for Sri Lanka (Fig. 2, area 17) by Fernando (1979) apparently coincide with mainland *A. cerana indica* (Damus and Otis, 1997; Fuchs et al., 1996). However, it must be noted that if the bees of Sri Lanka are indeed to be named *A. cerana indica* these "indica" are not the same biometric "indica" found further east in Malaysia (Damus and Otis, 1997). The current picture of honeybee subspecies, biometric groups, morphoclusters, ecotypes and/or biotypes in western Asia (Fig. 2) can only be regarded as highly suggestive and tentative because they emanate from separate studies which are not cross-compatible.

2.1.2. Northeast Asia

Northeast Asia as here defined includes China, the Manchurian plains of the former USSR, Korea and Japan (Fig. 2). The honeybees of the vast territory of China have been systematically investigated in a lengthy series of publications, principally by Yang and colleagues (Peng et al., 1989; Yang, 1989) who reached a number of conclusions based on analyses of honeybees from more than 1000 localities. They proposed a series of five major biometric groups or morphoclusters: 1. Manipuri bees from Nagaland, Manipur and Mizoram (Fig. 2, area 11); 2. Brahmputra bees from that valley and also from southern Assam and Megahalaya (Fig. 2, area 10); and 3. Himalayan bees from Sikkim, West Bengal, northern and western Assam and Arunachal Pradesh (Fig. 2, area 9). In a complementary study Verma et al. (1994) analysed bees from Nepal and the western Himalayas and could discriminate four biometric morphoclusters: 1. Terai plains bees in Nepal (Fig. 2, area 7); 2. Midland bees in the Nepali midlands (Fig. 2, area 8); 3. Himachali bees from Himachal Pradesh (Fig. 2, area 3); and 4. Kashmiri bees from Kashmir in northern India (Fig. 2, area 2). How the honeybees of Kashmir may relate to those of the northwestern frontier of Pakistan (Muzaffar and Ahmad, 1989) cannot yet be determined.

The mesoscale analyses of Singh et al. (1990) and Verma et al. (1994) were thorough multivariate analyses but performed on entirely unrelated, non-contiguous databases. Therefore, these rigorous mesoscale results are localised and cannot be extrapolated or statistically amalgamated to assess the whole region. In the circumstances this leaves a single macroscale analysis for the region in the study of Ruttner (1988) whose work provides conclusions on far less data and which is presented in a way that precludes any further numerical analysis. In any event, in the Ruttner perspective, the Kashmiri and Himachali bees (Fig. 2, areas 2–3) would be classified as *A. cerana cerana* while those extending across Nepal to the border of Myanmar as populations of *A. cerana himalayana* (Fig. 2, areas 14, 7–11). The majority of the ecotypes proposed by Kshirsagar (1983) as well as the bees of southern Pakistan would group together as populations of *A. cerana indica* in the Ruttner system.

To complete the tally for western Asia, the honeybees of Afghanistan and northern Pakistan would presumably show close affinities to *A. cerana indica*. The three ecotypes proposed for Sri Lanka (Fig. 2, area 17) by Fernando (1979) apparently coincide with mainland *A. cerana indica* (Damus and Otis, 1997; Fuchs et al., 1996). However, it must be noted that if the bees of Sri Lanka are indeed to be named *A. cerana indica* these "indica" are not the same biometric "indica" found further east in Malaysia (Damus and Otis, 1997). The current picture of honeybee subspecies, biometric groups, morphoclusters, ecotypes and/or biotypes in western Asia (Fig. 2) can only be regarded as highly suggestive and tentative because they emanate from separate studies which are not cross-compatible.

2.1.2. Northeast Asia

Northeast Asia as here defined includes China, the Manchurian plains of the former USSR, Korea and Japan (Fig. 2). The honeybees of the vast territory of China have been systematically investigated in a lengthy series of publications, principally by Yang and colleagues (Peng et al., 1989; Yang, 1989) who reached a number of conclusions based on analyses of honeybees from more than 1000 localities. They proposed a series of five major biometric groups or morphoclusters as well as several ecotypes (cf. Peng et al., 1989). The same groups or races have been subsequently supported (Zhen-Ming et al., 1992). There are two aspects to this work. Firstly, there are the original results of the Yang group and secondly, some new analyses and comments on parts of the same database by Peng et al. (1989).

Figure 2 (areas 4, 12, 13, 18, 27) illustrates the distributions of the five honeybee races that emanated from Yang’s group as well as the several ecotypes within them.
Peng et al. (1989) stated that the original Chinese studies did not include sufficient raw data nor details of descriptive statistics with which to re-evaluate the findings. Nonetheless, Peng and colleagues used multivariate methods to re-analyse some of the original Yang data but using a small character suite of only three morphometric features. So, there is an intrinsic difficulty both in the interpretation of the original Yang data as well as in the limited database processed by Peng et al. (1989).

Although Peng et al. (1989) were unable to support the biotypes of *A. cerana hainanensis*, they did however demonstrate a significant discrimination function (but of low probability) for the five biotypes of *A. cerana cerana* proposed by Yang (Fig. 2, area 4a–e). Peng et al. (1989) concluded that methodological differences between the Yang group and others preclude comparisons of these putative groups (Fig. 2) with those emanating from other honeybee studies in eastern Asia. There are only two possible secure links for this data. There appears to be a safe link between the *A. cerana indica* of southern Yunnan with the honeybees in the subtending Indochina peninsula (this same point emerged from the macroscale studies of Ruttner (1988, 1992). Finally, while Peng et al. (1989) were able to confirm the separateness of *A. cerana cerana* and *A. cerana skorikovi* but not the other three Yang races, they were inclined to accept these other races on the basis of behavioural and other biological characteristics. A very limited amount of data from the honeybees of China were available to Ruttner (1988) and he was only able to state that the bees of northern China was *A. cerana cerana* and those of the southwest a different, unspecified subspecies.

Although *A. cerana* is apparently non-native to the great expanses of the former Soviet Union, even in areas very near northern Afghanistan such as Tajikistan and Khirgizia, it resurfaces in the far eastern region of the Ussuri (or Primorsky) district (Fig. 2, area 6) just eastwards of Manchurian China (Avetisyan, 1960; Lawrjochin, 1960; Pesenko et al., 1989). Although Lawrjochin (1960) suggested that the Ussurian bees were close to *A. cerana japonica*, the Russian literature does not appear to provide any morphometric data on the bees of this region. Ruttner (1988, 1992) apparently did not have access to Ussurian honeybee samples and did not comment either way. Although never published, *A. cerana* apparently also occurs in eastern Mongolia (Choon Thin Yat, personal communication).

The honeybees of peninsular Korea (Fig. 2, area 5) have been analysed in a series of papers by Kwon and colleagues (cf. references in Kwon and Huh, 1992). Basically, they studied samples from fifteen localities in southern Korea and placed them all within the same biometric group. In the absence of re-analysable data, these results cannot be compared with any other Asian work. Ruttner (1988) seemed to regard these bees as morphometrically intermediate between *A. cerana cerana* of the mainland and *A. cerana japonica* in Japan.

The honeybees of the islands of Japan have been extensively analysed over the last century. It is general consensus that these bees are morphometrically completely isolated from others of the *A. cerana* complex (Damus and Otis, 1997; Ruttner, 1988; Sasaki, 1994) as well as in terms of mtDNA haplotypes (Deowanish et al., 1996). This isolation provides a convenient basis for studies of natural population structure in this branch of *A. cerana* as does the fragmentation of Japan itself into a series of islands. Two distinct morphoclusters are currently recognised, one on the islands of Kyushu, Shikoku and Honshu (bees are not native to northern Hokkaido) and another morphocluster occuring only on the small island of Tsushima in the Straits of Korea (Fig. 2, area 31a,b). The honeybees of each of these islands has some unique properties. For example, for southernmost Kyushu, Akahira and Sakagami (1959b) demonstrated a size cline in which the more southerly bees were larger than their northern counterparts.
In infraspecific categories of *Apis cerana* and Otis, 1997; Fuchs et al., 1996; Ruttner, 1988).

In another recent study of this region Damus and Otis (1997) performed multivariate analyses of insular Malaysia and Indonesia and obtained four distinct morphoclasters (Fig. 2, areas 23–26); one isolated island cluster on Timor (area 25) (one of these in extreme southern Sulawesi in area 24 was first noted in Hadisoesilo et al., 1995). The greater part of Indonesia formed one morphocluster (Fig. 2, area 23) with the exception of one small cluster in southern Sulawesi (area 24) and the bees of Sabah, NE Borneo (Fig. 2, area 26) yet another. In the classical literature all of these bees belong to the *A. cerana indica* complex (Ruttner, 1988, 1992), but are sometimes referred to as *A. c. javana* (Damus and Otis, 1997; Engel, 1999).

Damus and Otis (1997) also included bees of the Philippines (Fig. 2, areas 28–30) in their study and concluded that they are morphometrically distinct from the *A. cerana indica* of Indonesia. Moreover, they found that the bees of Luzon were morphometrically distinct from those of Mindanao. Coupling their morphometric data with the mtDNA results obtained by Smith and Hagen (1997) they questioned whether the bees of Luzon actually belong to any of the *A. cerana* groups. We return to this problem in considering morphometrics, mtDNA and allozymes conjointly.

Finally, the most recently analysed island group of honeybees is that of Tilde et al. (2000) who extensively covered the Philippines using standard multivariate methods. They found three distinct morphoclasters (Fig. 2, areas 28–30) corresponding to Luzon island (with highland and lowland ecotypes), another morphocluster on the islands in the Visayas and Mindanao groups and a quite separate cluster on Palawan. The bees of Palawan were quite distinct from the others. All of these bees were tentatively regarded as *A. cerana philippina* by Ruttner (1988).

Whereas, southern bees are lighter in colour than northern ones (Tokuda, 1924). Moreover, at an interlocality sampling distance of less than 100 km, intercolonial variance was low, intracolonial variance high. However, the intercolonial morphometric homogeneity in the variances of honeybees on Kyushu argues for a fairly uniform single population with continuous genetic flow among them.

2.1.3 Southeast Asia

This region extends east of longitude 98° and southwards from about latitude 20° N to Timor below the equator at 10° S. The mainland is about 1.5 million km² (Fig. 2). With the notable exceptions of Laos, Cambodia and Vietnam, it is also that area of Asia for which most of the recent analyses of honeybees have included thorough multivariate statistical analyses as well as analyses of mitochondrial DNA and various allozymes (see below).

Moving southwards down peninsular Indochina, the first study of interest concerns Thailand and Malaysia. Sylvester et al. (1998) published a comprehensive morphometric study of the honeybees of this region and unequivocally established four distinct morphoclasters (Fig. 2, areas 19–22) one covering most of Thailand, a second southern Thailand and continental Malaysia, a third at Phuket island and a fourth at Samui island. All four of these morphoclasters could be considered as subsets of what has previously been recognised as *A. cerana indica* (Ruttner, 1988, 1992). Explanations for the distinctness of the Samui and Phuket morphoclasters have been reasonably attributed to founder effects (Sylvester et al., 1998). The meaningfulness of the designation “*A. cerana indica*” (sometimes *A. c. javana*) for these bees is put under considerable pressure when it is remembered that the “*cerana indica*” of Thailand, Borneo and Malaysia are certainly not the same bees called “*cerana indica*” which occur in India and Sri Lanka (Damus and Otis, 1997; Fuchs et al., 1996; Ruttner, 1988).

In another recent study of this region Damus and Otis (1997) performed multivariate analyses of insular Malaysia and Indonesia and obtained four distinct morphoclasters (Fig. 2, areas 23–26): one isolated island cluster on Timor (area 25) (one of these in extreme southern Sulawesi in area 24 was first noted in Hadisoesilo et al., 1995). The greater part of Indonesia formed one morphocluster (Fig. 2, area 23) with the exception of one small cluster in southern Sulawesi (area 24) and the bees of Sabah, NE Borneo (Fig. 2, area 26) yet another. In the classical literature all of these bees belong to the *A. cerana indica* complex (Ruttner, 1988, 1992), but are sometimes referred to as *A. c. javana* (Damus and Otis, 1997; Engel, 1999).

Damus and Otis (1997) also included bees of the Philippines (Fig. 2, areas 28–30) in their study and concluded that they are morphometrically distinct from the *A. cerana indica* of Indonesia. Moreover, they found that the bees of Luzon were morphometrically distinct from those of Mindanao. Coupling their morphometric data with the mtDNA results obtained by Smith and Hagen (1997) they questioned whether the bees of Luzon actually belong to any of the *A. cerana* groups. We return to this problem in considering morphometrics, mtDNA and allozymes conjointly.

Finally, the most recently analysed island group of honeybees is that of Tilde et al. (2000) who extensively covered the Philippines using standard multivariate methods. They found three distinct morphoclasters (Fig. 2, areas 28–30) corresponding to Luzon island (with highland and lowland ecotypes), another morphocluster on the islands in the Visayas and Mindanao groups and a quite separate cluster on Palawan. The bees of Palawan were quite distinct from the others. All of these bees were tentatively regarded as *A. cerana philippina* by Ruttner (1988).
2.2. Allozyme diversity

Numerous studies of allozymes in *Apis mellifera* have shown relatively little allozyme diversity in this species (Cornuet and Garnery, 1991). However such variation as does occur, particularly in cytoplasmic malate dehydrogenase (MDH1, Enzyme Commission number 1.1.1.37), may have important metabolic consequences for honeybee flight (Harrison et al., 1996; Hepburn et al., 1999). MDH1 has proven to be a powerful tool for the investigation of population structure and gene flow in *A. mellifera* (Meixner et al., 1994; Smith and Glenn, 1994) especially when used in conjunction with other polymorphic enzymes, such as non-specific esterases (EST, E. C. number 3.1.1.1) or hexokinase (HK, E. C. number 2.7.1.1). Although allozymes have proved very useful in studies of *A. mellifera*, the study of allozymes in *A. cerana* is at a very early stage, and is beset with problems.

Only small portions of the range of *A. cerana* have been sampled: Pakistan (Nunamaker et al., 1984), Sri Lanka (Sheppard and Berlocher, 1989), Thailand, peninsular Malaysia, southern Sulawesi and the Philippines (Gan et al., 1991), Yunnan, China (Li et al., 1986), Japan (Rozalski et al., 1996; Tanabe and Tamaki, 1985) and Korea (Lee and Woo, 1991; Lee et al., 1989). In addition sample sizes have been small, on the order of 13 or fewer colonies. The majority of studies only investigated variation in MDH1 and/or non-specific esterases (EST).

Three studies (Gan et al., 1991; Lee and Woo, 1991; Sheppard and Berlocher, 1989) carried out a more thorough survey of 10–15 enzyme systems. Although a different suite of enzymes was surveyed in each study there is some overlap, particularly between the Korean (Lee and Woo, 1991) and Sri Lankan (Sheppard and Berlocher, 1989) studies. Not surprisingly, studies that surveyed more enzyme systems detected more variation. All polymorphisms discovered in these studies consisted of one common allele present at a frequency of 86% or higher, and one or more rare alleles. These studies are summarised in Table I.

Unfortunately, it is not possible to combine data from these studies to examine geographic patterns of allozyme variation or draw broader biogeographic conclusions. In this connection differing buffer systems are critical for the detection of allozyme variation and often confound the comparisons of results of different studies. Only the studies by Rozalski et al. (1996) and Sheppard and Berlocher (1989) provided a standardised nomenclature of alleles. In their studies, putative alleles of *A. cerana* were compared to the alleles found in *A. mellifera*, and named according to their relative electrophoretic mobility. Another useful practice followed by these authors was to include *A. mellifera* “standards” on gels, that is, samples with known genotypes.

Tentative among-region comparisons can be made for MDH and EST, the two enzymes most commonly surveyed. Japan and Pakistan samples showed only a single MDH1 allele, while samples from all other locations showed two alleles. In Korea and Sri Lanka, the “fast” allele was reported to be more common than the “slow” allele (this information was not provided for the Thai, Malay, Indonesian and Philippine samples). It is possible that MHD1107 from Japan, MHD1109 from Sri Lanka, and MDH1 fast from Korea and Thai, Malay, Indonesian and Philippine samples all correspond to the same common electromorph, while MHD1175 from Sri Lanka, and MDH1 slow from Korea and Thai, Malay, Indonesian and Philippine samples all correspond to the same rarer allele, but this cannot be confirmed without more direct comparisons. Some evidence of geographic variation in allele frequency is apparent in EST. This enzyme was reported to be monomorphic in Pakistan, China and Korea. Japan, Sri Lanka, and the Thai, Malay, Indonesian and Philippine samples each showed two
Table I. Summary of allozyme studies of *Apis cerana* (where alleles are named according to their relative mobility, authors used *A. mellifera* standards).*

<table>
<thead>
<tr>
<th>Locality</th>
<th>Sample size</th>
<th>Enzymes</th>
<th>Polymorphism</th>
<th>Alleles</th>
<th>Frequency</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rawalpindi, Pakistan</td>
<td>12 colonies, 100 bees total</td>
<td>EST</td>
<td>No</td>
<td>MDH1</td>
<td>No</td>
<td>Nunamaker et al., 1984</td>
</tr>
<tr>
<td>Meng La, southwest Yunnan, China</td>
<td>100 bees</td>
<td>EST</td>
<td>No</td>
<td></td>
<td></td>
<td>Li et al., 1986</td>
</tr>
<tr>
<td>Japan, 9 locations</td>
<td>13 colonies, 12–48 bees/colony, 405 workers, 48 drones</td>
<td>EST</td>
<td>Yes</td>
<td>EST73</td>
<td>0.96a</td>
<td>Rozalski et al., 1996</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>EST93</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>MDH1</td>
<td>0.86</td>
<td>Lee et al., 1989</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>MDH1slow</td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td>Korea</td>
<td>5 colonies, 27–30 bees/colony</td>
<td>EST</td>
<td>No</td>
<td></td>
<td></td>
<td>Lee and Woo, 1991</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>MDH1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>MDH1fast</td>
<td>0.86</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>MDH1slow</td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACPH, APH, EST, α-GPDH, HK, IDH, ME, ODH, PGM &amp; XDH all monomorphic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sri Lanka, 10 locations</td>
<td>10 colonies, ≥ 15 workers/colony</td>
<td>ACON2</td>
<td>Yes</td>
<td></td>
<td>ACON2100</td>
<td>Sheppard and Berlocher, 1989</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACON2114</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>EST86</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>EST95</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ME110</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ME91</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>MDH1109</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.86–1.00b</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.03–0.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.97–1.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.03–1.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.97–1.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.97–1.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.95–1.00</td>
<td></td>
</tr>
</tbody>
</table>
Table I. Continued.

<table>
<thead>
<tr>
<th>Locality</th>
<th>Sample size</th>
<th>Enzymes</th>
<th>Polymorphism</th>
<th>Alleles</th>
<th>Frequency</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bangkok, Thailand</td>
<td>unspecified</td>
<td>EST</td>
<td>No</td>
<td>EST&lt;sup&gt;70&lt;/sup&gt;</td>
<td>most common</td>
<td>Gan et al., 1991</td>
</tr>
<tr>
<td>Peninsular Malaysia</td>
<td></td>
<td>FUM</td>
<td>Yes</td>
<td>EST&lt;sup&gt;100&lt;/sup&gt;</td>
<td>least common</td>
<td></td>
</tr>
<tr>
<td>Sabah, Borneo</td>
<td></td>
<td>α-GPDH</td>
<td>Yes</td>
<td>1 common</td>
<td>4 rare alleles</td>
<td></td>
</tr>
<tr>
<td>south Sulawesi</td>
<td></td>
<td>GLDH</td>
<td>Yes</td>
<td>1 common</td>
<td>2 rare alleles</td>
<td></td>
</tr>
<tr>
<td>Indonesia</td>
<td></td>
<td>MDH</td>
<td>Yes</td>
<td>2 common</td>
<td>1 rare allele</td>
<td></td>
</tr>
<tr>
<td>Luzon, Philippines</td>
<td></td>
<td>SUDH</td>
<td>Yes</td>
<td>1 common</td>
<td>1 or 2 rare alleles</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>APH, ACPH, 6-PGD, SHDH</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Enzymes mentioned in text, with abbreviation and enzyme commission numbers in parentheses: ACON = aconitase (4.2.1.2); ACPH = acid phosphatase (3.1.3.2); ALDO = aldolase (4.1.2.13); APH = alkaline phosphatase (3.1.3.1); ARGK = arginine kinase (3.3.8.9); EST = non-specific esterase (3.1.1.1); FUM = fumarase (4.2.1.2); G-3-PDH = glyceraldehyde-3-phosphate dehydrogenase (1.2.1.12); GLDH = glucose dehydrogenase (1.1.1.47); α-GPDH = α-glycerophosphate dehydrogenase (1.1.1.3); β-HBDH = β-hydroxybutyric acid dehydrogenase (1.1.1.30); HK = hexokinase (2.7.1.1); IDH = isocitrate dehydrogenase (1.1.1.42); LAP = leucine amino peptidase (3.4.1.1); MDH1 = cytoplasmic malate dehydrogenase (1.1.1.37); ME = malate dehydrogenase (1.1.1.40); ODH = octanol dehydrogenase (1.1.1.73); 6-PGD = 6-phosphogluconate dehydrogenase (1.1.1.44); PGI = phosphoglucone isomerase (5.3.1.9); PGM = phosphoglucomutase (2.7.5.1); SUDH = succinate dehydrogenase (1.3.99.1); TPI = triose phosphate isomerase (5.3.1.1); XDH = xanthine dehydrogenase (1.2.1.37).

<sup>a</sup> Genotypes of queens were inferred from worker and drone genotypes, allele frequencies estimated from 13 queen genotypes.

<sup>b</sup> Range of allele frequencies found in colonies.
Infraspecific categories of *Apis cerana*

Philippines: de la Rúa et al., 2000; Smith et al., 2000). These studies have used several techniques for detection of variation. The earliest study surveyed restriction enzyme cleavage sites over the entire mitochondrial genome of *A. cerana* samples (Smith, 1991), while more recent studies PCR-amplify fragments of the mitochondrial genome, and screen for variation in restriction enzyme cleavage sites or sequence (Arias et al., 1996).

All of the more recent studies have focused on one region of the honey bee mitochondrial genome, from the cytochrome oxidase I gene (COI) to the cytochrome oxidase II gene (COII). Between COI and COII, lie the leucine tRNAUUR gene and a non-coding sequence that is apparently unique to *Apis* (Cornuet et al., 1991). The non-coding region is small in *A. florea*, *A. andreniformis* and *A. dorsata* (on the order of 24–32 bases), but larger in the cavity-nesting bees (89 to 97 in *A. cerana*, 94 in *A. koschevnikovi*, ~200–900 in *A. mellifera*). Because it is non-coding, this sequence is free to evolve rapidly, and provides information analysable at the intraspecific level.

A simple comparison of *P*, the proportion of polymorphic loci, among sites is difficult. Most studies only examined Mdh and Est, and the three studies that examined more loci did not examine the same set of enzymes. Empirically, some enzymes (e.g. Mdh1, and phosphoglucomutase, Pgm) frequently show variation, while other enzymes are less likely to show variation, so that the proportion of polymorphic loci will be biased by the set of enzymes surveyed. At present insufficient information on enzyme polymorphism (or lack thereof) is available from which to make any major inferences about geographic variation in allele frequencies in *A. cerana*.

2.3. Mitochondrial DNA diversity

MtDNA haplotypes have proven to be an incisive tool for unravelling the population structure of *A. mellifera* (Hall and Smith, 1991; Smith et al., 1991). By comparison, studies of the mtDNA of *A. cerana* and other Asian honeybees are in their infancy, even though *A. cerana* occupies an area comparable in size to that of *A. mellifera*. To date, nine studies have been published on the mtDNA of *A. cerana*. Most of these are comparative, surveying samples from numerous geographic locations but with relatively few samples per location (de la Rúa et al., 2000; Deowanish et al., 1996; Smith, 1991; Smith and Hagen, 1997, 1999; Smith et al., 2000). Some provide a more intensive survey of variation in a particular geographic location (Thailand: Deowanish et al., 1996; Sihanuntavong et al., 1999; Philippines: de la Rúa et al., 2000; Smith et al., 2000). These studies have used several techniques for detection of variation. The earliest study surveyed restriction enzyme cleavage sites over the entire mitochondrial genome of *A. cerana* samples (Smith, 1991), while more recent studies PCR-amplify fragments of the mitochondrial genome, and screen for variation in restriction enzyme cleavage sites or sequence (Arias et al., 1996).

All of the more recent studies have focused on one region of the honey bee mitochondrial genome, from the cytochrome oxidase I gene (COI) to the cytochrome oxidase II gene (COII). Between COI and COII, lie the leucine tRNAUUR gene and a non-coding sequence that is apparently unique to *Apis* (Cornuet et al., 1991). The non-coding region is small in *A. florea*, *A. andreniformis* and *A. dorsata* (on the order of 24–32 bases), but larger in the cavity-nesting bees (89 to 97 in *A. cerana*, 94 in *A. koschevnikovi*, ~200–900 in *A. mellifera*). Because it is non-coding, this sequence is free to evolve rapidly, and provides information analysable at the intraspecific level.

In addition to this region, Sihanuntavong et al. (1999) also examined PCR-amplified fragments containing portions of the genes for the small and large subunit ribosomal RNA genes (ssRNA and lsRNA).

Comparative, macroscale studies have employed both restriction fragment length polymorphisms (de la Rúa et al., 2000; Deowanish et al., 1996; Smith, 1991) and DNA sequence of the non-coding region (de la Rúa et al., 2000; Smith and Hagen, 1997, 1999; Smith et al., 2000) and there is only partial overlap among these studies in the geographic regions sampled. Nonetheless, results of these studies are largely congruent. Groups detected by all comparative studies are: (1) mainland Asia including Japan; (2) Sundaland (including southern or peninsular Thailand and the island of Samui); (3) Palawan (Philippines); and (4) the oceanic islands of the Philippines (Fig. 3).
The Asian mainland group contains a large number of related haplotypes. It includes samples from Japan, Korea, China (Hong Kong, Yunnan), Nepal, Vietnam (northern and southern), northern Thailand, and some of the bees from India. Deowanish et al. (1996) were also able to discriminate samples from Honshu Island, Japan, from their other mainland Asian samples by a HaeIII restriction site polymorphism in a fragment of mtDNA including part of the leucine tRNA gene, the non-coding region, and the 5' end of the COII gene. Studies using the sequence of the non-coding region alone (Smith and Hagen, 1997, 1999; Smith et al., 2000) were unable to discriminate Japanese bees from the bees of Korea and other parts of the mainland.

The Sundaland group of haplotypes includes samples from peninsular Thailand.
Infraspecific categories of *Apis cerana* (Deowanish et al., 1996; Smith and Hagen, 1997, 1999; Smith et al., 2000), and with morphometric data (Limbipichai, 1990; Sylvester et al., 1998). Thus the northern and central Thai samples belong to the Asian mainland group, while the southern Thai, Samui and probably the Phuket samples belong to the Sundaland group.

The shift from Asian mainland to Sundaland haplotypes occurs in the Isthmus of Kra at the so-called Kra ecotone, where there is a transition from evergreen rainforest (south of the imaginary line joining the cities of Kangar and Pattani) to more seasonal, semi-evergreen forest (Whitmore, 1984). The Bilauktaung mountain range, which forms part of the boundary between Myanmar (Burma) and Thailand in the Malay peninsula, may also provide some impediment to gene flow between Mainland and Sundaland populations.

The islands of the Philippines are home to a diverse collection of *A. cerana* populations, belonging to at least two mitochondrial lineages: the Palawan group and the oceanic Philippine islands group (Fig. 3). Mesoscale studies of Philippine bees from the islands of Palawan, Luzon, Mindanao, Panay, Negros, Cebu and Leyte were carried out by de la Rúa et al. (2000) and Smith et al. (2000). Both employed a set of colony samples (47 and 29 colonies, respectively) provided by the Bee Program of the University of the Philippines, Los Banos; Smith et al. (2000) also employed an additional 10 samples provided by other collectors. Palawan, Luzon and Mindanao had the best coverage, while Negros, Panay, Leyte and Cebu were each represented by 1–3 colonies. Both groups PCR-amplified a region of the mitochondrial genome that included the leucine tRNA gene, the non-coding region and the 5' end of COII. The approach used by de la Rúa et al. (2000) was to digest the amplification products with the restriction enzyme *DraI* to detect variation, and then sequence exemplars of each restriction pattern. This approach enables rapid screening...
of large numbers of colonies. Using this method they detected four haplotypes: two on Palawan, one on Luzon, and one on the remaining islands. Smith et al. (2000) sequenced the non-coding region of each sample; this is slower, but maximises the variation detected. Using this approach, 11 haplotypes were detected. The broad results of the two studies are congruent, showing haplotype frequency differences between three regions: Palawan, Luzon, and Mindanao plus the other islands. The studies differ in the amount of variation detected, particularly in Panay, Negros, Cebu and Leyte.

The distributions and relationships of these haplotypes can also be interpreted in light of changes in sea level during the Pleistocene (Heaney, 1985, 1986, 1991; Heaney and Rickart, 1990). During low sea level periods, islands of the Philippines were joined into larger units, or “mega islands”: Greater Palawan, Greater Luzon, Greater Mindanao (which includes modern Mindanao, Samar, Leyte, and Bohol), and Greater Negros-Panay (which includes modern Negros, Panay, Cebu and Masbate). Greater Luzon and Greater Mindanao may have been joined in the mid-Pleistocene, though this is not certain. (Today the trench separating Luzon and Samar is only 140 m deep, but this is currently a region of geological uplift, so the trench may have been deeper during the Pleistocene.)

Today Palawan is separated from nearby Borneo by a 145 m trench; thus Greater Palawan would have been united with the Asian mainland through Borneo in the mid Pleistocene, when sea levels were 160 m lower than present, but not in the late Pleistocene, when sea levels were only 120 m lower than present. The Palawan haplotype group includes two haplotypes found only on Palawan, and two found on Panay and northwestern Mindanao. Neighbour-joining and parsimony analyses (Smith et al., 2000) show the Palawan sequences more closely related to Sundaland and mainland Asian sequences than to other sequences from the Philippines.

The other Philippine islands have had no above-water connection to the Asian mainland, though two island chains, the Palawan chain between Borneo and Mindoro, and the Sulu Archipelago between Borneo and Mindanao, may have formed “stepping stones” between Borneo and the oceanic islands of the Philippines. The haplotypes on these islands appear to be the most divergent from other A. cerana haplotypes.

A current working hypothesis on population structure based on variability of mtDNA haplotypes is shown in Figure 3. Combining all of the currently available data it would appear there are four major groups of mtDNA lineages. These are (1) an Asian mainland mtDNA lineage, and three additional mtDNA lineages occurring on lands connected to the mainland for successively shorter periods of time: (2) Sundaland (connected to the mainland during the mid and late Pleistocene), (3) Palawan (connected only during the mid-Pleistocene), and (4) the oceanic islands of the Philippines (never connected to the mainland). The Indonesian island of Sulawesi too, was never connected to the mainland. On Sulawesi we find two exceedingly similar cavity-nesting bee species, A. nigrocincta and A. cerana with Sundaland mtDNA. The affinities of the “yellow” or “plains” bees of India and Sri Lanka are still uncertain, as are those of haplotypes lacking most of the non-coding sequence (all samples from Taiwan, and some from Sulawesi and the Philippines).

Discrete groups can be recognised within some of the major groups of mtDNA lineages. Within the large Asian mainland group, some haplotypes from “black” or “hill” bees form a distinct group; some authors have also been able to discriminate the bees of Japan from the other mainland samples. Within the Sundaland group, samples from Samui, peninsular Malaysia, and Borneo form distinct clusters (Fig. 3, areas a–c).
3. CONCLUDING REMARKS

Although a relatively large amount of information on the infraspecific classification of *A. cerana* is available, it varies quite considerably in quality and kind. Understandably, it ranges from the anecdotally descriptive accounts of earlier decades through univariate methods of analysis and finally to the more recent application of complete multivariate analyses of morphometric characters for specific regions. The greatest obstacle to a reasoned synthesis of infraspecific categories in *A. cerana* at present is that the results of most of such studies cannot be collated and unified because of fundamental and incompatible differences in statistical analyses, sample sizes, character suites, morphocluster confidence limits, the critical elements of sampling distance and extent of geographical coverage (Daly, 1991, 1992; Ruttner, 1988; Ruttner et al., 1978). Recent studies of *A. mellifera* show that morphocluster formation and inclusivity (correct classification) are highly sensitive to sampling distance intervals. Indeed, the length of a transect may obscure small biometric groups when the between-group variation is considerably larger than the within-group variation. Likewise, varying the limits of confidence applied to the ellipses and the discriminant *a posteriori* probabilities from low to high also decreases the numbers of colonies correctly assignable to morphoclusters (Diniz-Filho et al., 2000; Hepburn et al., 2000; Radloff and Hepburn, 1998). In the circumstances, specific recommendations for an approach to future morphometric analyses of *A. cerana* populations are currently being prepared (Hepburn et al., in preparation). Resolution of this problem is further exacerbated by the fact that the relevant literature largely ignores the recommendations of the International Code of Zoological Nomenclature (Engel, 1999). Nonetheless it is too premature to apply such rules until such time as the groups of *cerana* bees can be treated in a single, unified analysis.

This state of affairs is worrying for a number of reasons. As examples, note that the ecotypical studies of the honeybees of the Indian subcontinent reported by Kshirsagar (1983) cannot be rationalised with the multivariate studies of Singh et al. (1990) or of Verma et al. (1994) because of different methods of analysis, completely different databases and the absence of the raw databases. Similarly, the thorough and precise studies of Singh et al. (1990) on the honeybees of the western Himalayas cannot be added to the similarly precise work of Verma et al. (1994) on eastern Himalayan bees because of fundamentally different databases and the absence of sufficient raw data to provide for new and combined analyses. Similarly, equally thorough studies touching on Malaysia, Indonesia and the Philippines (Damus and Otis, 1997; Sylvester et al., 1998; Tilde et al., 2000) cannot be amalgamated into a single synthesis for this southern Asian region. Moreover, there remain as yet unsampled regions of *A. cerana* distribution which are quite considerably greater in extent than those that have been sampled.

Numerous anomalies become particularly apparent if a composite of the geographical distribution of all putative subspecies of *A. cerana* is attempted as in Figure 4. For example, the apparent disjunct distribution for *A. c. cerana* (separated by *A. c. skorikovi*) is unlikely to prove real given detailed analyses. Similarly the reality of *A. c. skorikovi* and *A. c. abaensis* derives from one set of authors (Peng et al., 1989) but is not considered by others (Ruttner, 1988, 1992). Moreover, it is readily apparent that most recent authors clearly distinguish minimally two different sets of what is currently termed *A. c. indica*, one to the west and the other eastern in distribution, the latter sometimes being called “*javana*” (Fig. 4).

Clarification of this anomaly can only be achieved when there is a sufficient database for Myanmar, Laos, Cambodia and Vietnam. Many other more localised anomalies are apparent.
The putative morphometric portrait of *A. cerana* shown in Figure 2 with a classification structure as in Figure 4 cannot be considered a reality as of yet. Extraordinary gaps must be filled for Afghanistan, Pakistan, west and central India, Myanmar, Laos, Cambodia, Vietnam, southern and northeastern China and Indonesia, the Andaman and Nicobar islands and perhaps Mongolia. Given these uncertainties, a comparison of morphometric results with those of mtDNA are equally disconcerting. There is a reasonable geographic congruity only between the two character sets for honeybees of the island systems Japan, Philippines, Taiwan, Phuket and Samui in Thailand and for peninsular Malaysia and Korea. All other morphoclusters and mtDNA clusters show little or no geographical correspondence. Turning to enzyme polymorphism (Fig. 3), the available data, as for mtDNA, is extremely sparse. In any event, it appears that low enzyme polymorphism occurs towards the two end points of *A. cerana* distribution (e.g., northeastern China and northern India). Conversely, high enzyme polymorphism occurs among the island systems (Sri Lanka, Sulawesi and the Philippines) and the Indochina peninsula. The massive continental mainland remains completely unexplored in this regard.
Infraspecific categories of *Apis cerana*

ACKNOWLEDGEMENTS

We thank P. de la Rúa, P. Neumann, Choon Thin Yat and S. Wongsiri for their help and advice with this manuscript. We especially thank P. Munn of IBRA for providing relevant literature.

Résumé – Les catégories infraspécifiques d’*Apis cerana*: diversité morphologique, allozymique et de l’ADNmt. La répartition actuelle d’*A. cerana* est continentale (Fig. 1). L’analyse des travaux publiés sur la taxonomie et la morphologie de cette espèce fournit environ 31 groupes biométriques possibles (Fig. 2) ou peut-être huit sous-espèces avec plusieurs écotypes (Fig. 4). Ces résultats proviennent d’un amalgame d’études allant de l’analyse anecdotique à l’analyse multivariée. Aussi une synthèse complète des catégories infraspécifiques définies par la morphométrie est elle actuellement exclue en raison de différences fondamentales et incompatibles entre les études quant à la taille de l’échantillon, les choix des caractères, la couverture géographique, la distance d’échantillonnage, les limites de confiance, la méthodologie statistique et l’absence courante d’une base de données continentale unifiée.

Les études de polymorphisme allozymique portant sur *A. cerana* sont très récentes mais elles rencontrent des problèmes. Peu de régions ont été étudiées, les échantillons sont de petite taille et le nombre de systèmes enzymatiques analysés est limité ; il s’agit principalement de la malate déshydrogénase et d’estérases non spécifiques. Mais l’étendue de la variation est liée au nombre de système enzymatiques testés. Les résultats des diverses études ne peuvent pas non plus être combinés car il n’existe aucune nomenclature normalisée des allèles. Il est donc difficile de déterminer les proportions de locus polymorphiques, ce qui exclut toutes déductions importantes concernant les catégories infraspécifiques.

Les analyses des haplotypes d’ADNmt sont peu nombreuses, limitées en échantillonnage et portent uniquement sur le gène de la cytochrome oxydase. Néanmoins plusieurs groupes distincts d’haplotypes émergent des études comparatives : (i) l’Asie continentale et le Japon, (ii) le Sundaland (péninsule malaysienne + îles de la Sonde, Bornéo) y compris la Thaïlande, (iii) Palawan (îles Philippines) et (iv) les îles océaniques des Philippines (Fig. 3). La répartition de ces groupes peut s’interpréter en relation avec les variations du niveau des mers au Pléistocène.

À cause d’incompatibilités méthodologiques entre les études morphométriques, allozymiques et d’ADNmt et de lacunes importantes dans l’échantillonnage géographique, il n’est pas encore possible d’établir des catégories infraspécifiques significatives pour *A. cerana*. De même, il n’est pas encore possible de faire des déductions importantes concernant les catégories infraspécifiques en combinant les données morphométriques, allozymiques et d’ADNmt, ni de fournir des dénominations acceptables du point de vue de la nomenclature pour les sous espèces d’*A. cerana*.


**REFERENCES** 


Infraspecific categories of *Apis cerana*


To access this journal online:
www.edpsciences.org