CHAPTER 4

DISCUSSION

The ATPase6-ATPase8 gene was chosen for studying of genetic variation and population structure of the honeybees, *A. cerana* in Thailand because it was used in the analysis of mtDNA variation in several animals and plant including human. Theoretically, ATPase genes have relatively high evolutionary rate and provide useful genetic variation data for phylogenetic studies as discuss below.

Study of sequence variation among different phenotypic forms of brown trout (*Salmo trutta*) from the North of Italy using the control region and ATPase6 indicated that level of genetic variation obtained from the protein coding genes like ATPase6 was slightly lower than that observed from the control region. Nevertheless, phylogenetic reconstruction based on both gene portion provided topologically comparable trees (Giuffra *et al.*, 1994). Chow and Ushiama (1995) studied global population structure of albacore (*Thunnus alalunga*) by using RFLP analysis of the mitochondrial ATPase gene. The results suggested greater gene flow between albacore of the northern and southern hemispheres than does between the Atlantic and Pacific Oceans. In the last year, Vitic and Strobeck (1996) studied genetic population structure and stock identification of lake trout (*Salvelinus namaycush*) in West-Central Canada using mitochondrial DNA polymorphisms. Several different genotypes were found from amplified fragment length polymorphisms (PCR-RFLP) coupling with DNA sequencing of the

ATPase6-COIII region. A total of 17 mitochondrial haplotypes were identified. Common haplotypes were overlappingly distributed across geographic regions whereas unique haplotype were observed in each population with relatively low frequencies indicating high gene flow in this species. Additionally, intraspecific phylogeny of sugarcane (*Saccharum officinarum* L.) were examined using RFLP approach by probing digested total DNA with Cytochrome Oxidase (CO) subunits 11, 12, 13, Cytochrome B and ATPase6 in mitochondrial genome as well as ATPase9, ATPase alpha and 18S+5S in the nuclear genome, respectively (Aljonabi *et al.*, 1994).

From all described above, the ATPase6-ATPase8 mtDNA gene was then chosen for this study.

4.1 Methodology in this study

Genetic variation and population structure of the honeybees, *A. cerana* in Thailand was carried out using the restriction fragment length polymorphism (RFLP) of PCR amplified ATPase6-ATPase8 in the mitochondrial genome. Three restriction endonuclease including two having six-base recognition sequences; *Ssp*I (AAT↓ATT) and *Vsp*I (AT↓TAAT), and one having four-base recognition sequence, *Taq*I (TC↓GA), were used to determine levels of genetic diversity of conspecific *A. cerana* population in Thailand. The three individuals of *A. mellifera* were also included in the experiment for interspecific comparison.

To obtain the mitochondrial haplotypes, total genomic DNA was extracted from thoracic tissue using a modification of a standard method described in 2.7. ATPase6-ATPase8 mtDNA gene of *A. cerana* was then

amplified via PCR, with a set of primers specifically designed from *A. mellifera* sequences. Since *A. mellifera* and *A. cerana* were closely related (Garnery *et al.*,1991), these primers were able to used for amplification of ATPase6-ATPase8 from both species. The resulting PCR product was 825 bp long in both *A. cerana* and *A. mellifera*.

Total DNA of honeybees from individual thoraces was used as template in PCR for ATPase6-ATPase8 mtDNA amplification. Based on the fact that primers were specifically designed from ATPase6-ATPase8 gene in mitochondrial genome of closely related taxon such as *A. mellifera* and high temperature of an annealing step (50°C) was used in PCR to amplify only of ATPase6-ATPase8 gene without nonspecific fragment as shown in Figure 3.5. Amplification of ATPase6-ATPase8 gene from both total DNA and mitochondrial DNA yielded a product having identical molecular length but no amplified product from nuclear genome used as template. As a result, total DNA could be used as a template in PCR for amplified this gene because it is easy to be isolated with high yield from individual thorax and required of a short time period.

The amplified product of *A. mellifera* was confirmed to be ATPase6-ATPase8 gene. From RFLP analysis, resulting fragments produced with several restriction enzymes could be correctly inferred from the restriction map of ATPase6-ATPase8 gene of this species. Additionally, the amplified product of *A. cerana* was sequenced from 5'end and then aligned with ATPase6-ATPase8 gene of *A. mellifera* mtDNA, resulting in approximately 83% base pair matching in this

region. This results supported that the amplified product of both species, *A. mellifera* and *A. cerana*, were homologous.

In *A. cerana* samples, length heteroplasmy was observed intraindividually. Nonetheless, a 825 bp fragment presented in all *A. cerana* and *A. mellifera* investigated samples even in heteroplasmic individual. Accordingly, an amplified 825 bp fragment of ATPase6-ATPase8 gene was the most common product, so that this fragment was further employed for RFLP analysis. *A. cerana* individuals showing heteroplasmy in ATPase6-ATPase8 were only found in the South and Samui Island. In heteroplasmic samples, a band at 825 bp was eluted from agarose gel electrophoresis using QIAQUICK column.

There were twenty-three restriciton endonuclease; *Bam*HI, *Bfr*I, *Bgl*I, *Bgl*II, *Bst*EII, *Bst*NI, *Cla*I, *Hae*III, *Hin*cII, *Hin*dIII, *Kpn*I, *Mlu*NI, *Nde*I, *Pst*I, *Pvu*II, *Rsa*I, *Sal*I, *Sca*I, *Sma*I, *Swa*I, *Tha*I, *Xba*I and *Xho*I which showed no restriction sites in *A. cerana* ATPase6-ATPase8 mtDNA while *Hin*fI could digest only two specimens (S02 and S60, of the South) out of all 181 specimens. *Tru*9I gave DNA fragments which were too small to determine consistently by agarose gel electrophoresis. Five restriction enzymes, *Acs*I, *Alu*I, *Dra*I, *Eco*RI and *Sau*3AI generated the same fragment pattern in all except 2 specimens , S02 and S60, of the South. Based on produced variable patterns and consistent results, three enzymes (*Ssp*I, *Taq*I and *Vsp*I) were chose to be used for population studies. *Vsp*I, *Ssp*I and *Taq*I produced six, five and two different restriction patterns, respectively.

In this study, there were 13 single haplotypes observed from 181 individuals analyzed with 3 restriction enzymes (*Taq*I, *Ssp*I and *Vsp*I).

Six of these were common haplotypes in which 4 haplotypes were only represented by a single colony. Moreover, one of these (Haplotype C of *Vsp*I digest) was a population specific haplotype found only in the Samui Island. One possible explanation of this was that it is possible to employ this population specific haplotype as a molecular marker for conservative program of *A. cerana* in Thailand.

A total number of 10 composite haplotype of mtDNA were then generated and were subjected to phylogenetic tree construction based on the genetic distance among composite haplotypes using UPGMA. The mtDNA composite haplotypes could then be clustered into 3 groups, the Northern (North, North-East and Central), the Southern (South and the Samui Island), and a group containing two colonies (S02 and S60) of South *A. cerana* samples, respectively. On the basis of this result, it was clear that the Northern has a different gene pool from the Southern *A. cerana*. Surprisingly, the estimate of genetic distance between two haplotypes (AEE, AEF) of S02 and S60 from the South and other haplotypes in *A. cerana* was greater than that of between AFG of *A. mellifera* and all haplotypes except AEE and AEF, this strongly suggesting that S02 and S60 individuals may not be in *A. cerana*.

4.2 Heteroplasmy

Typically, an amplified ATPase6-ATPase8 gene in *A. cerana* and *A. mellifera* were 825 bp in which was equally to that obtained from a complete sequences of *A. mellifera* mitochondrial genome (Crozier and Crozier, 1993) which was evaluated to be 825 bp.

In this study, the amplified product of 60 *A. cerana* individuals from the South and 18 individuals from the Samui Island had 2 to 4 bands (825, 900, 925 and 950 bp). These were sequenced for approximately 58-65 bp. Multiple alignment of each these bands sequences with the ATPase6-ATPase8 sequences of *A. mellifera*, show 82% sequences similarity suggesting that these four bands were the same amplified products of ATPase6-ATPase8 gene in *A. cerana* mitochondrial genome but size different resulted from insertions or deletions of mtDNA. This evidence is called heteroplasmy.

Contaminated DNA from parasites (e.g. *Calex* species) in *A. cerana* may be the case for heterogeneous in amplified mtDNA. However, this hypothesis were not true for this study as proved by amplification of ATPase6-ATPase8 from eight *A. cerana* individual from the same colony gave the same amplification pattern..

More recently, within-individual differences in mtDNA (heteroplasmy) have been documented in domestic cattles (Hauswirth and Laipis, 1982; Olivo *et al.*, 1983; Hauswirth *et al.*, 1984) and in a laboratory strain of *Drosophila mauritiana* (Solignae, Monnerot and Mounolou, 1983). In cattles, the heteroplasmic differences are due to both base substitutions and small insertions and deletions (i.e., from 1 to 10 bp). In *D. mauritiana* , the mtDNA patterns of its two subpopulations differ by 500 bp. This heteroplasmic difference was due to the presence of a tandemly duplicated 500 bp unit whlch is present in either two or three copies.

Heteroplasmy, when it occurs, is usually due to differences in the copy number of tandemly repeated sequences among molecules.

Heteroplasmy due to sequence differences is more rare (<1%) and can usually be attributed to proximal mutational events (Olivo *et al.*, 1983).

4.3 DNA divergence within and among *A. cerana* populations

The haplotypic diversity of 5 geographic areas of *A. cerana* mtDNA ranged from 0.073-0.508 (Table 3.7). The lowest value of haplotype diversity was observed in South (0.0726) because they have one common haplotype and two private haplotypes which were found in one individual each. The highest value of haplotype diversity was observed in the Samui Island (0.5085) because of high frequencies of their two composite haplotypes. Although, North-East have four composite haplotype but the low value of haplotype diversity was observed because of low frequencies of their composite haplotypes. The average haplotype diversity for *A. cerana* was 0.2004±0.0063 suggesting low level of genetic polymorphism in *A. cerana* mtDNA.

The average nucleotide diversity among populations in *A. cerana* was approximately 2.15874% (Table 3.8) which was much higher than that within population (0.2852%) (Table 3.7) implying strong genetic differentiation within this species.

Pairwise comparisons in nucleotide divergence (Table 3.8) within the Northern and the Southern *A. cerana* were quite low (0.0780-0.112 for the former and 0.8537 for the latter) whereas large nucleotide divergence were observed between each of the Northern and each of the Southern. The average nucleotide divergence *A. cerana* in the present study was 0.018735±0.0000253. Phylogenetic reconstruction based on the genetic distance and nucleotide divergence was carried out using

UPGMA. The phenogram obtained from this method indicated that 5 geographic regions of *A. cerana* could be clearly divided into 2 distinct evolutionary lineages, the Northern and the Southern. The former was composed of North, North-East and Central while the latter was composed of South and the Samui Island.

From the value of genetic distance, haplotype diversity and nucleotide divergence, it was clearly that the Northern was reproductively isolated from the Southern *A. cerana*.

Remarkably, the nucleotide divergence between *A. mellifera* and *A. cerana* from the Southern locations were higher than that between *A. mellifera* and *A. cerana* from the Northern locations which implied that Southern *A. cerana* accumulated of mutations greater than Northern *A. cerana*. This resulted suggested that Southern *A. cerana* are the ancestor.

4.4 Geographic heterogeneity frequencies distribution

A chi-square (χ^2) analysis, a Monte Carlo simulation was used to analyze heterogeneity in distribution frequencies of 10 composite haplotypes among five difference geographic areas of *A. cerana* in Thailand (North, North-East, Central, South and the Samui Island).

Geographic heterogeneity between each pairwise comparison of the Northern *A. cerana* were not significant different (P=0.1463-0.4788) (Table 3.5), indicating that *A. cerana* from North, North-East and Central were the same population (= the same gene pool). Composite haplotypes from these three locations were pooled. Geographic heterogeneity was then reanalyzed using a Monte Carlo simulation. Highly significant differences in haplotype frequencies between the Northern and each of

the Southern *A. cerana* (South and Samui Island) were observed ($P<0.0001$) (Table 3.5). This mean that the Northern *A. cerana* is not the same population as the Southern *A. cerana* indicating strong population differentiation in this species. With the Southren *A. cerana*, geographic heterogeneity between South and the Samui Island was highly significant different ($P<0.0001$) (Table 3.5) implied that South or Samui Island can be further allocate to be different population.

Based on geographic heterogeneity analysis, it was clear that *A. cerana* from five different geographic areas were allocated into three genetic population, the Northern (North, Northern and Central), South and Samui Island, respectively.

4.5 Do the two specimens from the South are different species?

Private haplotypes of all restriction enzymes digestion except with *Taq*I and *Sau*3AI were observed from SSouth , S02 and S60, in the South. Additionally, genetic distances of two composite haploytpes (AEE and AEF) and single composite haplotype of *A. mellifera* (AFG) in comparison to those of *A. cerana* were highly different (Table 3.6). The estimated genetic distance of composite haplotypes between SSouth and each of *A. cerana* was approximately 3.89%-5.93%. Genetic differences between *A. mellifera* (AFG) and each of the Northern *A. cerana* was 3.40%-6.19% which was much lower than pariwise comparison on genetic distances between AFG and BBC or BBB (7.33% and 10.8%, respectively). It was then suspected that the specimens S02 and S60 may not be *A. cerana* as their genetic distances with other *A. cerana* individuals were larger than that between AFG (of *A. mellifera*) and all

of the remaining haplotypes. Large differences in nucleotide divergence of SSouth and the Northern and the Southern *A. cerana* were also observed (Table 3.10). This also supported the suspicion of sympatnetic (overlapping) of two species which are morphologically similar and their taxonomic status goes undetected by classical taxonomic methods. Nevertheless, this conclusion are required to be confirmed by a larger sample sizes because nucleotide divergence estimated from SSouth and *A. mellifera* was from 2 and 3 individuals, respectively.

Two individuals of SSouth might be *A. koschevnikovi* previously known to occur only on Borneo and Sumatra because its morphology is the closest to *A. cerana* resulting in some problems to clearly dissociate this two species morphologically. Therefor, further investigation using know *Apis* species specimens as references would be an alternative way to clarified the taxonomic status of SSouth samples.

ATPase6-ATPase8 gene of *A. cerana* has proved to be powerful to in analysis of genetic variation and population structure of *A. cerana* in Thailand. Results from this study provided a fundamental information on genetic obviously of this species. This is essential for further genetic studies. However, the study about genetic variation and population structure should be analyze with more than one approach along with other basic biological disciplines for instance morphology, life history, mating behavior. The resulting data will be much more reliable than from only one approach.