CHAPTER I

Introductions

Honey bee, the honey maker, is a stringing insect that lives in large social groups. Its symbolic dance language and polyandrous mating system are unique among different bee species (Cameron, 1991). In agricultural cultivation, honey bee helps increasing crop plant production by pollination. Moreover, the products directly derived from honey bee such as honey, pollen, wax and royal jelly are essentially important in food, drug and cosmetic industry. Honey is a famous natural product that contains available nutrients and is one of the most important exports of Thailand. Therefore, honey bee is economically important insect species. The value of the natural honey exportation has been estimated to be between 9.29-44.01 million baht since 1984. This makes an annually additional income to the country (Table 1).

The social structure of honey bee colony is composed of one queen, large number of workers and a few hundred of drones. Both the queen and workers are heterozygous diploid reproduced through sexual fertilization ($2n = 32$) whereas drones, reproducible males, are hemizygotes (haploid individuals) arising from unfertilized eggs. The queen, a long-lived female, evidences a mechanism of controlling offspring by the "Queen pheromone" and consecutively mates with 7 to 17 drones and stores about 5 million spermatozoa in her spermatheca during the mating period. Unlike the queen,

Table 1 The imports and exports of natural honey during 1980 - 1997.

Year	Imports		Exports	
	Quantity (Ton)	C.I.F. Value (Million Baht)	Quantity (Ton)	C.I.F. Value (Million Baht)
1980	46	1.30	0	0
1981	71	2.16	0	0
1982	109	2.95	12	0.23
1983	172	4.41	70	1.10
1984	114	3.26	160	2.30
1985	No Data			
1986	132	4.03	1222	21.02
1987	130	3.84	745	11.11
1988	125	4.42	1750	24.53
1989	146	5.59	704	9.29
1990	166	6.19	2432	31.11
1991	232	8.79	1206	16.96
1992	172	7.30	2407	32.39
1993	230	10.60	2108	28.30
1994	264	12.24	1894	26.94
1995	238	11.10	1908	29.37
1996	326	16.29	2656	44.01
Oct.1997	229	13.29	1944	30.06

Source : Thai Customs Department, Finance Ministry, Thailand

workers are usually sterile, and have many different tasks such as nursing and foraging on age-dependent changes. Normal aged nurses are seven to ten days old, and aged foragers are 21-24 days old (Page and Erickson, 1986 ; Robinson, 1991).

Classification of honey bee can be morphologically defined as follows, (Borror, et al., 1976 ; Gojmerac, 1980).

Kingdom Metazoa

Phylum Arthropoda

Class Insecta

Order Hymenoptera

Super family Apoidea

Family Apidae

Genus *Apis*

Apis species can be grouped on the basis of morphology and behaviour into three lineages: firstly, the cavity-nesting bees, *A. mellifera*, *A. cerana* and *A. koschevnikovi* ; secondly, the dwarf bees, *A. florea* and *A. andreniformis* ; and finally, the giant bees, *A. dorsata* (Smith, 1991). All except *A. koschevnikovi* could be found in Thailand. Four species, including *A. dorsata*, *A. andreniformis*, *A. florea* and *A. cerana*, are native and widely distributed over vast geographic areas whereas *A. mellifera* is the introduced honey bee originating from Europe and Africa. Only *A. mellifera* and *A. cerana* can be brought for commercial beekeeping (Figure 1.1). Their nests are in sheltered cavities with multiple parallel combs. Beekeeping of *A. mellifera* is more sucessful than does *A. cerana* because the former is not aggressive, well

Figure 1.1 Beekeeping of *A. cerana* in Thailand.

studied and simple to manage whereas beekeeping of *A. cerana* is mainly a common house - hold business. Nevertheless, *A. cerana* is still of interest for beekeeping in Thailand because *A. cerana* has several better biological features than does *A. mellifera*. These are composed of disease resistance, less sugar feeding requirement and better climatic adaptability. Moreover, *A. cerana* has a cleaning behaviour to detect, bite, kill and remove mites from their colonies (Wongsiri, Rinderer and Sylvester, 1990). Unfortunately, there have been a limited number of publications concerning the behavioural biology of *A. cerana* which is necessary and useful discipline for understanding its basic behaviour. This needs to be carried out in the near future.

A. cerana can be classified into four subspecies based on morphometric characters, distribution and behaviour including 1) a north subspecies, *A. cerana cerana* found in Afghanistan, China, North India, North Viet Nam and Pakistan 2) a south subspecies, *A. cerana indica* found in south India, Sri Lanka, Bangladesh, Burma, Malaysia, Thailand, Indonesia and Philippines 3) a Japanese subspecies, *A. cerana japonica* found in Japan 4) a Himalayan subspecies, *A. cerana himalaya* found in the South-East Asian mountain and probably in the South-West of China (Ruttner, 1988). The geographic distribution of these species is shown in Figure 2.1. *A. cerana* in Thailand was further grouped using two multivariate statistical analyses of morphometric charateristics. These were carried out by Canonical and Cluster analyses. The former is able to divide *A. cerana* into three distinctive groups consisting

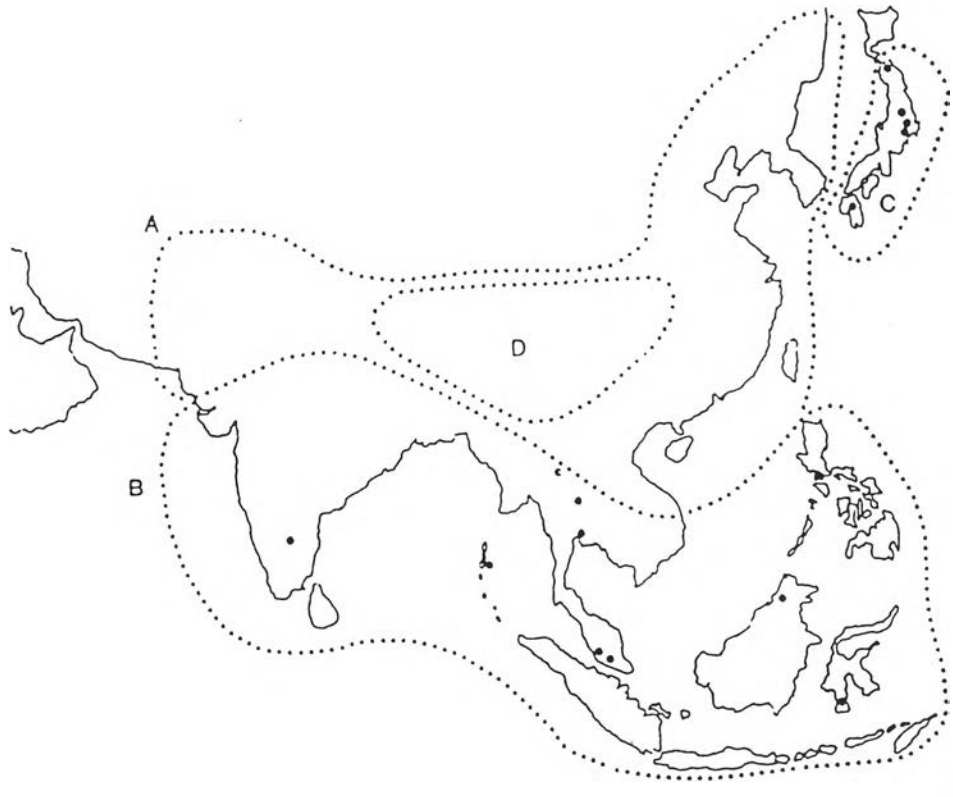

Figure 1.2 Dotted lines indicate the approximate geographic ranges of *A. cerana* subspecies as recognized by Ruttner (1988). A = *A. cerana cerana*; B =*A. cerana indica*;C =*A. cerana japonica*;D = *A. cerana himalaya*

of Northern and Southern latitude bees and the Samui Island bees whereas the latter can allocate *A. cerana* into only Northern and Southern latitude bees (Limbipichai, 1990).

Although morphometric analysis is one of reliable identification approaches available, it also has several disadvantages, for example ; some important morphometric characters particularly sizes are environmentally influenced (Smith and Brown, 1988) and because of undefined relationships between phenotypes and genotypes, morphometric characters may not be good indicators for phylogenetic reconstruction between these *Apis* (Cornuet and Garnery, 1991). More importantly, this technique requires experienced scientists to decide whether interested characters are informative (Rinderer, 1986).

Allozyme or isozyme analysis was the first approach used to determine genetic polymorphism in honey bees. The different alleles of variant proteins could be found on their charges under electric field. Among *Apis* species, two enzymes, malate dehydrogenase (MDH) and esterase (EST), were extensively studied. MDH showed that *A. florea* was different from *A. dorsata* and *A. cerana* whereas EST is able to separate *A. cerana* from either *A. dorsata* or *A. florea* (Nunamaker, Wilson and Ahmad, 1984). Likewise, EST can distinguish the genetic difference between *A. cerana* and *A. mellifera* (Tanabe and Tamaki, 1985). Within *A. mellifera*, the frequency of MDH alleles among conspecific populations was also significantly different (Smith and Glenn, 1995 ; Schiff and Shappard, 1996).

Theoretically, genetic polymorphism detected by allozymes is underestimated. This is mainly due to its inability to detect synonymous mutation and some kinds of non-synonymous mutations. The reason for this is that the genetic code is degenerate, therefore a polymorphism at the protein level could be only detected when a nucleotide substitution has resulted in an amino acid replacement affecting the electrophoretic mobility of the investigated protein molecules (Weising *et al.*, 1995). On the other hand, Hymenoptera (haplodiploid organism) such as bee has a lower frequency of enzyme polymorphism than does diploid organism. Therefore, this approach seems to be inappropriate to determine level of genetic polymorphisms in *Apis* species (Metcalf, Marlin and Whitt, 1975 ; Pamilo and Crozier, 1981).

Analysis of DNA polymorphism is the direct approach to study genetic variation in various taxa. DNA can be utilised to examine genetic variation at intraspecific level. Theoretically, various methods based on DNA analysis are consistently increased especially when polymorphism of investigated species are rather limited.

Nuclear DNA polymorphisms.

Nuclear DNA polymorphisms can be resulted from several mechanisms. Single base substitutions, loss (deletion), gain (insertion, duplication) and rearrangment of variable lengths of sequences can be occured within a DNA fragment. The structure of protein coding gene in the nuclear gene consists of coding regions (exons) and non coding region (introns). Exons are highly

conserved. In contrast, much higher polymorphisms can be found in introns (intervening sequences) (Krawczak and Schmidtke, 1994).

Restriction fragment length polymorphisms (RFLP) is one of several techniques used to determine DNA variation based on differences in fragment lengths. When genomic DNA is digested with a restriction enzyme, millions of different DNA fragments are obtained. On the agarose gel, the DNA fragments appear as a continuous smear. The specific fragment can be indentified by hybridization with the specific radiolabeled probe (Davis, Battey and Kuehl, 1994). Initially, the genetic differences between Africanized and European *A. mellifera* were found using RFLPs through Southern analysis (Hall, 1986). Likewise, six difference patterns were observed from *A. cerana* in Thailand when probing total DNA-*Hea* III digestion with repetitive sequence DNA (Uthaisang, 1994). Recently, microsatellite loci have been applied for analysis of DNA polymorphism in *Apis* species. The microsatellite loci characterise a high variance in repeat units of short-tandemly repetitive sequences. Because of their higher mutation rate, microsatellite analysis is a powerful appoach to be used for kinship analysis and intraspecific phylogeny (Estoup, Garnery, Solignac and Cornuet, 1995 ; Estoup, Tailliez, Cornuet and Solignac, 1995).

Mitochondrial DNA polymorphisms

Animal mitochondrial DNA (mtDNA) is an extrachromosomal DNA occuring as single circular molecules, generally about 16,000-20,000 bp long, located in mitochondria. Generally, the mitochondrial genome is composed of

13 proteins-coding genes, 22 transfer RNA (tRNA) genes, two ribosomal RNA subunits and noncoding region containing the origin of replication . The genes that code for protein are three subunits of cytochrome c oxidase (CO I, II and III), seven subunits of NADH dehydrogenase (ND1-6 and ND4L), cytochrome b, and two subunits of ATP synthetase (ATPase 6 and 8) (Crozier and Crozier, 1993) (Figure 1.3). Unlike the nuclear DNA, animal mtDNA is maternally inherited. Moreover, recombination is rarely found in this molecules. (Smith and Brown, 1988). The evolutionary change of mtDNA occurs mainly due to nucleotide substitutions rather than deletions and insertions (Nei and Tajima, 1981). The rate of nucleotide substitutions in the mitochondrial genome has been estimated to be about 5 - 10 times more rapidly than that of nuclear DNA (Brown, George and Wilson, 1979). For these reasons, mtDNA is potentially useful for population genetic and molecular systematic studies of animal taxa. An important application on the basis of its high mutation rate is particularly useful in determination of population structure at intraspecific level.

MtDNA-RFLP is a popular technique used to evaluate levels of genetic variation in various species. Both restriction mapping and restriction fragment length polymorphisms of *A. mellifera* mtDNA were reported. All restriction endonuclease used (*Bgl* II, *Bcl* I, *Eco*R I, *Nde* I, *Xba* I and *Hinf* I) revealed polymorphism (Moritz et al., 1986 ; Smith and Brown, 1988 ; Meixner, Sheppard and Poklikar, 1993 ; Sheppard et al., 1996). Heteroplasmy was found in the Brazilian samples, which was 270 bp larger than that in the American samples (Smith and Brown, 1988). Each of restriction sites was

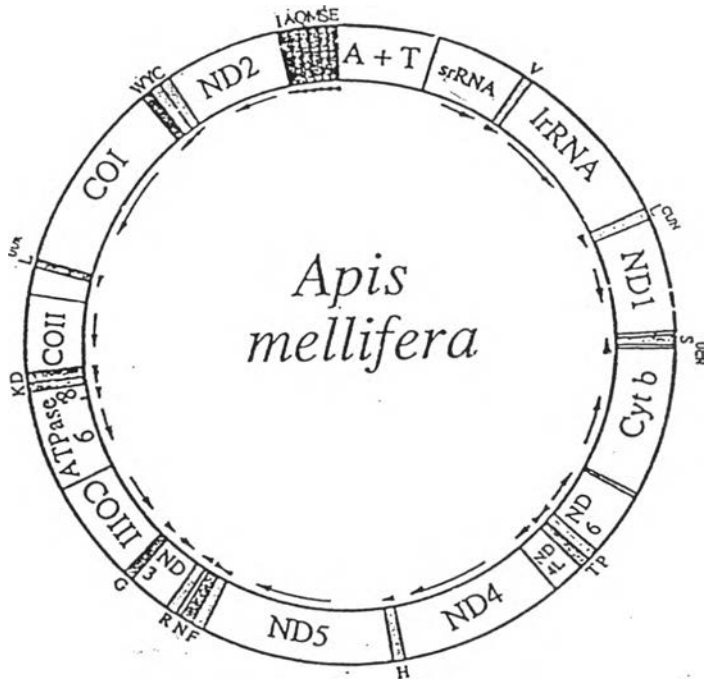

Figure 1.3 Map of mitochondrial genome of the honey bee, *A. mellifera*.

One letter code denoted for tRNA genes corresponding amino acid and arrows indicated the direction of transcription for each coding region.

revealed polymorphism (Moritz et al., 1986 ; Smith and Brown, 1988 ; Meixner, Sheppard and Poklikar, 1993 ; Sheppard et al., 1996). Heteroplasmy was found in the Brazilian samples, which was 270 bp larger than that in the American samples (Smith and Brown, 1988). Each of restriction sites was then mapped onto the mitochondrial genome (Smith and Brown, 1990) (Figure 1.4). Moreover, several genes in the *A. mellifera* mitochondrial genome, e.g. sRNA and lrRNA, CO I, II and III and cytochrome b genes were unambiguously located (Arias and Norbrega, 1991). Likewise, genetic variation of *A. cerana* mtDNA was determined with *Bcl* I, *Bgl* II, *Eco*R I, *Hin*d III and *Spe* I. The results indicated three main matriarchal lineages : one including *A. cerana* from south India, Thailand, Malasia, Borneo and Japan; another consisting of that from the Andaman Islands; and the rest consisting of Luzon *A. cerana* (Smith, 1991).

Although RFLP of the whole mitochondrial genome is useful for studies of genetic variation in *A. cerana* population, this technique requires large amount of pure mtDNA. Furthermore, the procedures were more complicated with using radioisotope for visualisation of the results. As a result, RFLP analysis of whole mitochondrial genome is tedious and time consuming.

Presently, polymerase chain reaction (PCR) is an ingenious new approach for molecular biology. With this technique, millions of DNA can be synthesised from DNA target by the activity of thermostable DNA polymerase (usually *Taq* polymerase) within a few hours. PCR reaction

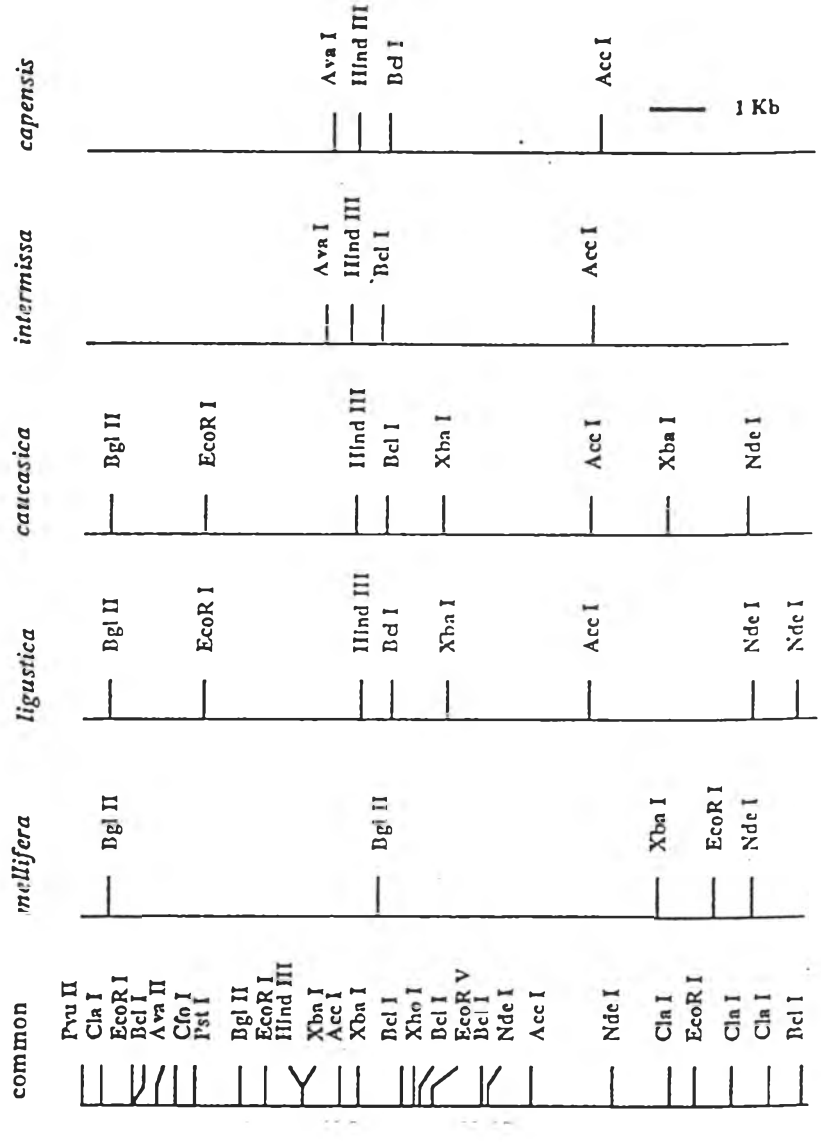

Figure 1.4 A restriction enzyme cleavage map for mitochondrial DNA of *A. mellifera* subspecies. Restriction sites common to all subspecies were shown at the bottom.

contains three important steps including ; denaturation of double stranded DNA template at high temperature, annealing to allow primers to form hybrid molecules at optimal temperature, and extension at working temperature of DNA polymerase. PCR can also be utilised for direct sequencing of amplified fragments (Innis *et al.*, 1990). The PCR-amplified fragment can be further employed for analysis of restriction site and fragment length polymorphisms.

Crozier, Koulianos and Crozier (1991) amplified a 485 bp fragment of cytochrome b gene from *A. mellifera* mitochondrial genome. The amplification product was then restricted with *Bgl* II resulting in an ability to discriminate the African from non-African lineages. The results were in agreement with those from PCR-RFLP of lrRNA (738 bp) and CO I (1044 bp) genes, respectively. More specifically described, an *Eco*R I site in the lrRNA gene and an *Xba* I site in the CO I gene were found only in *A. mellifera* from the east European, but a *Hinc* I site in the CO I gene was only observed in the west European (Hall and Smith, 1991). A region between CO I-CO II gene (inter CO I-CO II region) was probably the most popular mtDNA portion to be studied in *A. mellifera*. Length polymorphisms of amplified DNA were found both between and within *A. mellifera* subspecies. The region was sequenced and analysed. It comprised two units : P (54 bp, 100 % A + T) or P_0 (67 bp) and different number of Q (196 bp, 93.4 % A + T). This region exihibited at least 7 length variants and can be explained by the combinations of P_0Q, P_0QQ, P_0QQQ, PQ, PQQ, PQQQ and Q (Crozier, Crozier and Mackilay, 1989 ; Cornuet, Garnery and Solignac, 1991; Garnery *et*

al., 1993). The inter CO I-CO II region was proved to be highly polymorphic portion of *A. mellifera* mitochondrial genome when digested with other restriction enzymes (*Xba* I and *Dra* I). Hall and Smith (1991) reported *Xba* I restriction site found only in the east European race of *A. mellifera.* More recently, over twenty-one different haplotypes were observed when the PCR product of inter CO I-CO II region amplified from the Moroccan and Spanish *A. mellifera* was digested with *Dra* I. The results were then subjected to phylogenetic reconstruction among them to indicate genetic differences its subspecies (Garnery *et al.*, 1993;Garnery *et al.*,1995).

Thus, PCR-RFLP is powerful technique to study genetic variation in *Apis* species. The advantages of this technique are mainly due to its rapid and simple experimental procedures obviating the requirement to use radioactive material such as ^{32}P. More importantly, the technique does not require large amount of the DNA which seems to be one of the most limiting factors on population genetic studies of *Apis* species.

Theoretically, DNA sequencing can be carried out by either chain termination or chemical degradation method. The chain termination method is much more convenient and produced by newly synthesising DNA chains with DNA polymerase. The reaction is terminated by an incorporation of a modified nucleotide (dideoxynucleotide triphosphate) added into the reaction. The complete sequences of honey bee, *A. mellifera* mitochondrial genome was reported to be 16,343 bp long and extremely A+T rich (84.9 % AT content). Transversions were about twice as abundant as transitions when

comparing to *Drosophila* and *Apis* protein-encoding genes (Crozier and Crozier, 1993). Then, intraspecific phylogenetic tree constructed from the sequences of ATPase 6, CO III, cytochrome b, and ND 2 genes showed a significant association between *A. mellifera* subspecies (Koulianos and Crozier, 1997).

The aim of this thesis is to determine levels of polymorphisms in various mtDNA regions such as lrRNA gene, sRNA gene and inter CO I-CO II region using PCR-RFLP approach. The results obtained will be used to construct intraspecific phylogenetic tree of *A. cerana* originating from geographically different conspecific locations of such a species in Thialand. A representative of a gene region (lrRNA gene) will then be selected for further analysis on DNA sequencing. The information obtained from this and PCR-RFLP will be compared.