Chapter II

Materials and Methods

Materials and Chemicals

-Absolute ethanol (Merck, Germany)

-Agarose gel type 1-A low electroendosmosis (Sigma Chemical Co., USA)

-Ammonuim acetate (Merck, Germany)

-Acrylamide (Sigma Chemical Co., USA)

-Bovine Serum Albumin (New England Biolabs, USA)

-Boric acid (Merck, Germany)

-Chloroform (Merck, Germany)

-Citric acid and tri-sodium citrate (Fluka Chemika-Bio Chemika, Switzerland)

-Deoxynucleoside triphosphate (dATP, dCTP, dGTP and dTTP, Perkin Elmer, USA)

-Ethylenediaminetetraacitic acid (EDTA), disodium salt dihydrate (Fluka Chemika-Bio Chemika, Switzerland)

-Dextrose (Sigma Chemical Co., USA)

-Formaldehyde (BDH, England)

-Glacial acetic acid (Merck, Germany)

-Glucose (Sigma Chemical Co., USA)

-$MgCl_2$ (Perkin elemer Cetus, USA)

-Proteinase-K (Gibco BRL life technologies, Inc., USA)

-RNase A (Sigma Chemical Co., USA)

-Silver nitrate (Sigma Chemical Co., USA)

-Sodium carbonate (Fluka Chemika-Bio Chemika, Switzerland)

-Sodium chloride (Merck, Germany)

-Sodium dodecyl sulfate (Sigma Chemical Co., USA)

-Sodium thiosulfate (BDH, England)

-Sucrose (Sigma Chemical Co., USA)

-Tris (hydroxymethyl) aminomethane (Fluka Chemika-Bio Chemika, Switzerland)

Enzyme

-Ampli Taq DNA polymerase (Perkin-Elmer Cetus, USA)

Experimental Procedures

Sampling

Ninety individuals of juvenile *Chelonia mydas* bodies originating from the Andaman Sea and the Gulf of Thailand were used in genetic variation study. Turtles from the former sampling area were forty-one individuals maintained at Sea Turtle Conservation Center (the Third Squadron Fleet, Phangnga). The Gulf of Thailand *C. mydas* specimens were forty-nine individuals maintained at Sea Turtle Conservation Center at Sattahip Naval Base (Chonburi).

The samples were kept at -20° C were transported to the laboratory on ice or, in a few cases, on dry ice (-70° C) and then freezed at -80° C for a long storage when required, *C. mydas* individuals were thawed. Internal organs consisting of heart, liver and kidney from each individual were dissected out and served as tissue sources for genomic DNA isolation.

For multiple paternity study, 6 individuals of juvenile *C. mydas* from the same clutch, originated from the Andaman Sea and hatched in our laboratory, were used. Approximately 20 µl of blood were individually collected from the dorso-cervical sinus region and carefully added into a glass tube containing 2 ml of an acid citrate dextrose solution (ACD, 0.48 % citric acid, 1.32 % sodium citrate and 1.47 % dextrose by weight) to prevent blood coagulation. The mixture was gently but thoroughly mixed by inverting of the tube for

several times and immediately placed on ice (Maniatis, 1989). Blood samples were stored in a 4° C refrigerator until further used.

DNA extraction

Both internal tissues and blood samples were extracted DNA by proteinase K/SDS digestion following by phenol extraction modified from that described in Maniatis (1989). This extraction procedure was surprisingly worked well without the use of chloroform/isoamyl alcohol.

Approximately one gramme of soft tissues (heart, liver and kidney) dissected out individually was pulverised with a mortar and pestle to a fine powder. The tissue was frozen throughout this process by occasionally adding liquid N_2. The powdered tissue was transferred into a tube containing 15 ml of extraction buffer (10 mM Tris-HCl, pH 8.0, 100 mM EDTA, pH 8.0, 20 μg/ml pancreatic RNase and 0.5% SDS) and homogenised briefly to disperse and disrupt any remaining tissue lumps. And then the mixture was mixed thoroughly and incubated at 37° C for 1 hour. A proteinase K solution was added to a final concentration of 300 μg/ml and further incubated at 37° C for 3 hours. An equal volume of buffer equilibrated phenol (saturated in TE buffer; 10 mM Tris-HCl pH 8.0, 1 mM EDTA) was added and gently mixed by inverting of the tube for about 5 minutes. The mixture were centrifuged at 5000xg for 20 minutes at 27° C. The upper (aqueous) phase was carefully decanted into a new sterile tube. Phenol extraction was further repeated for 1-2 times or until the protein interface was not observed. The final extracted

aqueous solution was then collected. DNA precipitation was carried out by an addition of one-third volume of 10 M ammonium acetate followed by two volumes of absolute ethanol. The DNA pellet was hooked out using a glass rod and washed twice for approximately 5 minutes each in 70%, 80% and 90% ethanol, respectively. DNA pellet was air-dried at room temperature and dissolved in 1 ml of TE buffer and incubated at 37° C until completely redissolved. The resulting DNA solution was stored at 4° C until used.

Agarose gel electrophoresis

Agarose gel electrophoresis was carried out for two purposes. Firstly, to determine the quality of the isolated DNA samples and secondly, to estimate their DNA concentration.

An appropriate amount of agarose was weighed out and heated to dissolve in an appropriate volume of 1x Tris-Borate-EDTA buffer, TBE (89 mM Tris-HCl, 8.9 mM Boric acid, 2.5 mM EDTA, pH 8.0) to make desired gel concentration. While the agarose was cooling, the ends of the gel bed were sealed and placed on a level bench. When the agarose cooled to about 50° C, the melted agarose was poured into a gel mould (which the comb was already inserted to). When the gel had completely cooled and solidified (usually left at room temperature for 1 hour), the comb was gently removed. The gel was

submerged in a gel chamber containing an enough of TBE that covered the gel to a depth about 1-2 mm.

Extracted DNA was prepared for loading by mixing loading dye buffer (0.1% bromophenol blue, 40% ficoll 400 and 5 mM EDTA, pH 8.0) to each sample and mixed well. The samples were applied into the wells slowly. Ten microlitres of λ/DNA digested with *Hin*d III (λ/*Hin*dIII) were also loaded to the gel and served as a DNA standard (250 ng). A gel bath was connected to a power supply and turned on. DNA migrates into the gel toward the anode or positive lead. Generally, the gel was run at 100 volts until bromophenol blue had migrated about three-fourths of a gel distance (approximately 1-2 hours). After electrophoresis, the gel was stained with 0.25 μg/ml ethidium bromide for 10 minutes. The gel was destained in deionised H_2O for 10-15 minutes to leach out unbound ethidium bromide. The gel was placed on a long wavelength UV transluminator (UVP) and photographed using Kodak TMX-400 film.

Since the intensity of fluorescence resulted from ethidium bromide is proportional to DNA length therefore the quantity of DNA samples could be estimated by comparing their fluorescence to that of a λ-*Hind* III standard DNA. This was rapid and convenient method to estimate the amount of DNA in such samples. DNA concentration of each individual was then diluted to 25 ng/μl.

Amplification of microsatellite DNA using the Polymerase Chain Reaction (PCR)

Five pairs of nucleotide primers originally developed in *C. mydas* from Australia as described in FitzSimmons *et al.* (1995) were used as homologous and heterozygous primers in the present study to amplified microsatellite loci of *C. mydas* from Thailand. The Cm 72, Cm84 and Cc117 primers were purchased from Bioservice Unit, National Center for Genetic Engineering and Biotechnology whereas Cm 3 and Cm58 were from Thaican Co., Ltd. Their sequences were shown in Table 2.1.

Approximately 25 ng of genomic DNA from an individual of *C. mydas* was used as DNA template for a 25 µl PCR reaction mixture containing 0.2, 0.4 or 0.4 µM of DNA primers loci Cm3, Cm72 and Cc117, respectively, 1x *Taq* polymerase buffer (10 mM Tris-HCl, pH 8.4, 50 mM KCl), appropriate concentration of $MgCl_2$ at 4 mM, 3 mM, 3 mM, 4.5 mM and 2.5 mM for Cm3, Cm58, Cm72, Cm84 and Cc117, respectively, 0.16 µg/ml of acetylated bovine serum albumin (BSA) and 100 µM each of dATP, dCTP, dGTP and dTTP. Finally, 0.625 unit of *Taq* polymerase was added into the resulting mixture. The PCR mixture was spun briefly at 2000xg before subjected to the amplification process in a Perkin-Elmer Cetus DNA thermal cycler (model 2400).

Table 2.1 Five primer sequences used in amplified microsatellite DNA

Locus	Array	primer sequence	T_m
Cm3I	$(CA)_{13}$	CTA CCA TGA GAT GGG ATG TG	60° C
Cm3II	$(CA)_{13}$	TTT TCT CCA TAA ACA AGG CC	56° C
Cm58I	$(CA)_{13}$	GCA GTA CAC TCG GTA TTT AT	56° C
Cm58II	$(CA)_{13}$	TGA AAG TGA CAG GAT GTA CC	58° C
Cm72I	$(CA)_{33}$	CTA TAA GGA GAA AGC GTT AAG ACA	66° C
Cm72II	$(CA)_{33}$	CCA AAT TAG GAT TAC ACA GCC AAC	68° C
Cm84I	$(CA)_{15}$	TGT TTT GAC ATT AGT CCA GGA TTG	66° C
Cm84II	$(CA)_{15}$	ATT GTT ATA GCC TAT TGT TCA GGA	64° C
Cc117I	$(CA)_{17}$	TGT TTA ACG TAT CTC CTG TAG CTC	68° C
Cc117II	$(CA)_{17}$	CAG TAG TGT CAG TTC ATT GTT TCA	66° C

The reaction was initially denatured for 2.5 minutes at 95° C followed by further 35 cycles with a 95° C denaturing step for 45 seconds, a 60 second annealing step at 55° C, 55° C, 66° C, 64° C or 61° C for loci Cm3, Cm58, Cm72, Cm84 and Cc117, respectively and a 72° C extension step for 90 seconds. The final extension of the PCR reaction was carried out for 5 minutes at 72° C. The amplification products were stored at 4° C and electrophoretically analysed for their molecular sizes as soon as possible.

A DNA size-standard used was "pGEM-3Zf (+) sequencing marker" which was provided as a control DNA in cycle-sequencing kit (Promega). A 16 µl reaction mixture (prepared as a set of 4 tubes labelled as A, C, G, T) containing 4 µl of pGEM-3Zf(+) (1 µg/µl), 3.2 µl of 5x DNA sequencing buffer, 3.6 µl of pUC/M13 forward primer (4.5 pmol), 2 µl of d/ddNTP mix (for each of A, C, G, T reaction), 3.4 µl of steriled H_2O and 5 units of sequencing grade *Taq* DNA polymerase was prepared as recommended by the manufacturer. PCR was carried out in a Perkin-Elmer Cetus DNA thermal cycler (model 2400) for 60 cycles consisting of 30 seconds at 95°C and 30 seconds at 70° C.

At the end of the PCR reaction, three microlitres of loading dye buffer (10 mM NaOH, 95% formamide, 0.05% bromophenol blue, 0.05% xylene cyanol) and one microlitre of 10x TE (100 mM Tris-HCl, pH 8.0, 10 mM EDTA) was added to each of the PCR-amplified marker tubes.

Size estimation of amplified microsatellite DNA using electrophoresis of sequencing polyacrlyamide gels

Glass plates were thoroughly cleaned with detergent, rinsed with deionised H_2O and air-dried. The longer glass plate was siliconised with glass coating solution, Rain-X (Unelko Co., USA) and further cleaned with detergent. The plate was then thoroughly rinsed with deionised H_2O and air-dried. The shorter glass plate was wiped with binding solution (0.0118 mM γ-methacryloxypropyltri-methoxysilane in 95% ethanol and 0.5% glacial acetic acid) and left at room temperature for 4-5 minutes. The short glass plate was carefully wiped with 95% ethanol to remove the remaining binding solution. Plates were placed together with spacers and their edges were sealed with sticky tape.

Standard sequencing gel (6% acrylamide monomer) was prepared and used to determined the sizes of *C. mydas* microsatellite DNA. An enough amount of ingredients for a 50 ml gel composed of 0.084 M acrylamide monomer, 2 mM bis-acrylamide, 0.3 M urea and 1x TBE was prepared. The dissolved ingredients were suction-filtered. To initial the gel polymerisation, 235 µl of 10% ammonium persulphate and 46.8 µl of TEMED (*N, N, N', N'* tetramethylethylenediamine) was added to acrylamide mixture, gently swirled and quickly poured into the gel apparatus. The comb was inserted. The polymerisation process were allowed to be complete for 2-3 hours. The comb was then removed. Urea was flushed out of the wells prior to loading samples.

When ready, the samples mixed with loading dye buffer (10 mM NaOH, 95% formamide, 0.05% bromophenol blue, 0.05% xylene cyanol) and 10x TE buffer at the ratio of 30:10: 3. The mixing solution was heated at 95° C for 10 minutes and immediately snap-cooled on ice before loading. Four and eight microlitres of pGEM sequencing marker and amplified DNA samples were then loaded into the gel. Electrophoresis was carried out with constant voltage 1,150 volts for 3, 4.5 and 5.5 hours for Cm3, Cm72 and Cm84, respectively.

Silver Staining

At the end of electrophoresis, the plates were dismantled. The short plate containing a sequencing gel was placed in a fixing/stopping solution (10% acetic acid) to fix the DNA bands and to leach out urea which can block signal for silver staining if still remaining in the gel. This was carried out on an orbital shaker for approximately 20 minutes at room temperature.

The gel was then agitation-washed with double-distilled H_2O 3 times for 10 minutes each. The gel was soaked in a staining solution (0.6 mM of silver nitrate and 0.056% formaldehyde) and well agitated for 30 minutes followed by briefly rinsed with ultrapure H_2O for 5-10 seconds before developing of the gel. The gel was then soaked in chilled developing solution (0.03 M sodium carbonate, 0.056% formaldehyde and 0.14 mM sodium thiosulfate) until amplification bands were clearly visualised. Developing of the gel was terminated by adding an enough amount of chilled fixing/stopping solution and left for 20 minutes. The stained gel was briefly rinsed with double-distilled

H_2O to remove the remaining fixing/stopping solution. The gel was then air-dried and can be kept permanently at room temperature until results were confirmed. Finally, the gel was photographed using Kodak TMX 100.

Data statistical analysis

Assumption

An electrophoretic pattern of each *C. mydas* individuals were scored for each locus and referred to a genotype. Therefore, the genetic status could be divided into homo and heterozygotic states (theoretically, one or two bands, respectively). Nevertheless, stuttered bands were commonly observed in dinucleotide microsatellite. Accordingly, scoring of the gel can be carried out unambiguously by making an assumption that a real band reflecting an actual allele was the most intense band among its neighbour group of stuttered bands.

Each allele was designed using its fragment size which compared to the sequencing marker (in base pair, bp) As a result, each single individual was recorded to be either homo- and heterozygote. The allelic stages were recorded from each individual for each locus.

Allele frequencies

The frequency of a given allele in a population for diploid organisms can be estimated as

$$p = \frac{2N_{AA} + N_{Aa}}{2N}$$

where N_{AA} and N_{Aa} are number of homo- and heterozygotes for such an allele and N is number of investigated individuals.

Hardy-Weinberg equilibrium

Theoretically, each investigated population of diploid organisms are determined for Hardy-Weinberg (H-W) equilibrium for each locus to illustrate that mating is actually occurred at random and genotypes are produced by random union of male and female gametes.

Basically, H-W distribution was tested for each locus in each population using the chi-square for goodness of fit as follow

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

where O and E represent observed and expected genotype frequencies, respectively. Practically, Hardy-Weinberg conformation in this study was calculated using an exact test implemented in Genepop. The theory behind this was described by Gua and Thompson (1992).

Gametic disequilibrium (Weir, 1996)

Gametic disequilibrium (the so called, linkage analysis) is a study of associations between alleles at different loci when two or more loci are considered together. Any associations found is referred to linkage

disequilibrium therefore the null hypothesis of this test is that gametic association are in equilibrium.

The probability, Pr (n_{AB}) of the gametic array under the hypothesis of no linkage disequilibrium (= linkage equilibrium) is

$$\Pr(n_{AB}) = \Pr(n_{AB}, n_{A\bar{B}}, n_{\bar{A}B}, n_{\bar{A}\bar{B}})$$

$$= \frac{(2n)!(p_A p_B)^{n_{AB}}(p_A p_{\bar{B}})^{n_{A\bar{B}}}(p_{\bar{A}} p_B)^{n_{\bar{A}B}} p_{\bar{A}} p_{\bar{B}})^{n_{\bar{A}\bar{B}}}}{n_{AB}! n_{AB}! n_{\bar{A}B}! n_{\bar{A}\bar{B}}}$$

and the probability of the two allele arrays are

$$\Pr(n_A n_{\bar{A}}) = \frac{(2n)!}{n_A! n_{\bar{A}}!}(p_A)^{n_A}(p_{\bar{A}})^{n_{\bar{A}}}$$

$$\Pr(n_B n_{\bar{B}}) = \frac{(2n)!}{n_B! n_B!}(p_B)^{n_B}(p_{\bar{B}})^{n_{\bar{B}}}$$

Taking the ratio of these quantities gives the probability of the gametic numbers conditional on the allele numbers

$$\Pr n_{AB}, n_{A\bar{B}}, n_{\bar{A}B}, n_{\bar{A}\bar{B}} | n_A, n_B = \frac{n_A! n_{\bar{A}}! n_B! n_{\bar{B}}}{n_{AB}! n_{AB}! n_{\bar{A}B}! n_{\bar{A}\bar{B}}}$$

Exact test for multiple alleles are expressed as the probability of a set of count $\{n_{uv}\}$ for A_uB_v gametes, conditional on the allele counts $\{n_u\}$ and $\{n_v\}$ is

$$\Pr(\{n_{uv}\}|\{n_u\},\{n_v\}) = \frac{\prod_u n_u! \prod_v n_v!}{(2n)! \prod_{u,v} n_{uv}!}$$

and the significant level is obtained by permuting the allele at one of the loci.

Genetic variation

The number of alleles per locus and their allele frequencies was evaluated across loci. Genetic variation within populations (in this case within the Andaman Sea and the Gulf of Thailand) was measured in terms of heterozygosity (H) which is the proportion of heterozygous individuals in all investigated specimens, therefore observed heterozygosity can be measured as

$$H = \sum(N_{Ab} / N) / n$$

where n is the number of investigated loci.

When the populations investigated are in Hardy-Weinberg equilibrium, heterozygosity (surprisingly called as *expected heterozygosity*) can be calculated as

$$h = 1 - \sum x_i^2$$

where x_i is the frequency of i^{th} allele at a given locus. Accordingly, expected heterozygosity across all loci is the mean of h from each locus.

Genetic distance

Genetic distance is a measure of gene diversity between populations which is expressed as a function of genotype frequency. Several numbers of

formulations which have been proposed but one of the most popular estimation is Nei's coefficient of genetic identity (I) between two taxa (Nei, 1972)

For population X and Y, the probability of identity of two randomly chosen genes at a single locus (Jk) is

$$j_x = \sum x_i^2 \text{ and } j_y = \sum y_i^2$$

where x_i and y_i are the frequencies of the i^{th} alleles at a given locus in population X and Y, respectively.

The probability of identity of a gene at the same locus in population X and Y is

$$j_{xy} = \sum x_i y_i$$

The normalised identity between population X and Y with respect to all loci is

$$I = \frac{J_{XY}}{(J_X J_Y)^{1/2}}$$

where J_X, J_Y and J_{XY} represent arithmetic means of j_x, j_y and j_{xy}, respectively which are taken overall loci.

Genetic identity between X and Y populations is then converted to Nei's standard genetic distance using the formulation

$$D = -ln\ (I)$$

Geographic heterogeneity test

To determine whether there was significance heterogeneity in distribution of allele frequencies between geographically separated populations, a χ^2 test based on Monte Carlo simulations was carried out for ten thousand times using Monte 4.1 on REAP.

In this pseudo-chi square analysis, the χ^2 value was calculated from the original data following by the calculation of χ^2 from simulated data. The probability of an analysis accounts the number of times that the simulated data set are exceeded or equal to the original χ^2 value by chance divided to the total number of simulation.

Estimation of population structure and gene flow level (Wright et al., 1978)

F-estimate, F_{ST} is a standard parameter for measuration of population structure (or interpopulation diversity). It can be calculated from the formulation;

$$F_{ST} = \text{(Actual variance)/(Limitting variance)}$$

where,

Actual variance = *Total variance - Sampling variance*

$$\textit{Total variance} = r^{-1}\sum(p-\bar{p})^2$$

Sampling variance = $r^{-1}\sum(2n)^{-1}\,\bar{p}(1-\bar{p})$

Limitting variance = $\bar{p}(1-\bar{p})$

and p represents the frequency of an allele at a locus in a population whereas $\bar{p}$ is the average allelic frequency over all populations at that allele and that locus, n is the number of individuals sampled from that locus in that population and r represents the number of investigated populations.

Under neutral theory, gene flow (the genetic exchange between populations resulting from migration of individuals or their gametes) is of interest because the level of interpopulation diversity is a function of the number of reproductively successful migrants exchanged among populations. On the basis of F-statistics, the absolute number of individuals transferred between populations per generation (N_m) can be estimated as

$$F_{ST} \cong 1/(1+4N_m) \text{ or } N_m \cong (1-F_{ST})/4F_{ST}$$